

## Insight into *Saccharothrix espanaensis* Grown in Different Media: Transcriptome and Natural Product Analysis

Suzan Samra<sup>1</sup>, Bogdan Tokovenko<sup>2</sup>, Monika Kaszubowska<sup>1</sup>, Roman Makitrynskyi<sup>1</sup>, Ahmad Alali<sup>1</sup>, Yvonne Schmidt<sup>1</sup>, Elisabeth Welle<sup>1</sup>, Sandra Gross<sup>1</sup>, Thomas Paululat<sup>3</sup>, Andriy Luzhetskyy<sup>2,4</sup> and Andreas Bechthold<sup>1\*</sup>

<sup>1</sup>Department of Pharmaceutical Biology and Biotechnology, Institute of Pharmaceutical Sciences, Albert Ludwigs University, 79104 Freiburg, Germany

<sup>2</sup>Helmholtz Institute for Pharmaceutical Research Saarland, Saarland University Campus, Building C2.3, 66123 Saarbrücken, Germany

<sup>3</sup>Department of Chemistry and Biology, Organic Chemistry II, Faculty IV, University of Siegen, Adolf-Reichwein-Straße 2, 57068 Siegen, Germany

<sup>4</sup>Department of Pharmaceutical Biology, Saarland University, Building C2.3, 66123 Saarbrücken, Germany

\*Corresponding author: Bechthold A, Department of Pharmaceutical Biology and Biotechnology, Institute of Pharmaceutical Sciences, Albert-Ludwigs-University, 79104 Freiburg, Germany, Tel: +49-761-2038371; Fax: +49-761-2038383; E-mail: andreas.bechthold@pharmazie.uni-freiburg.de

Received: July 03, 2018; Accepted: July 20, 2018; Published: August 01, 2018

Copyright: © 2018 Samra S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

In this study, we show that cultivating of *Saccharothrix espanaensis* in the amino acid-rich SPY medium led to the detection of antibacterial and antifungal products. The cultivation in GYM-, HA- or NL111-medium did not lead to the production of these compounds. Four products were isolated from *Saccharothrix espanaensis*-SPY culture and determined as the known compounds anthranilic acid, indol-3-carboxylic acid, benzoic acid, and phenylacetic acid. In order to investigate the influence of SPY medium on gene transcription and to identify candidate genes involved in the biosynthesis of the purified compounds a comparative RNA sequencing analysis of *Saccharothrix espanaensis* cultivated in different media was performed. Total transcriptome analysis demonstrated that media composition has an influence on the expression of several genes. Among them are genes encoding enzymes belong to amino acid metabolism enzymes. Furthermore, based on the collected data genes involved in the biosynthesis of anthranilic acid, indol-3-carboxylic acid, benzoic acid, and phenylacetic acid could be predicted. Deletion of the genes *BN6\_01860* and *BN6\_67950* revealed their importance in the biosynthesis of indole-3-carboxylic acid and benzoic acid, respectively.

**Keywords:** *Saccharothrix espanaensis*; Transcriptome; Tryptophan; Phenylalanine

### Introduction

Variation of cultivation parameters can have a great impact on the production of microbial natural products. The deliberate elaboration of cultural parameters, to augment the metabolite diversity of a strain, has been called OSMAC (one strain many compounds) [1]. Following the OSMAC approach varying media composition led to the production of novel and potentially bioactive natural products from actinomycetes [1]. Different media have been described for the cultivation of actinomycetes and in many of them a successful production of natural products has been reported [2]. Furthermore, it was found that the addition of the rare earth, scandium, causes antibiotic overproduction in some actinomycetes strains [3].

*Saccharothrix espanaensis* (*S. espanaensis*) is a soil-dwelling Gram-positive bacterium belonging to the family *Pseudonocardiaceae* from the order *Actinomycetales* [4], known to include producer strains of an impressive number of biologically active compounds, many of which are clinically-used as antibiotics, anticancer and immunosuppressive agents [5]. *S. espanaensis* has been described as producer of saccharomicin A and saccharomicin B [6], which are two heptadecaglycoside antibiotics exhibiting potent antibacterial activity both *in vitro* and *in vivo* against multiply resistant strains of *Staphylococcus aureus* as well as vancomycin-resistant *enterococci* [7]. Unfortunately, attempts to produce both compounds under laboratory conditions failed recently for so far unknown reasons.

Recently the complete genome sequence of *S. espanaensis* was published. The circular genome comprises 9,360,653 bp, encoding 8,427 genes. The genomic sequencing revealed the presence of 26 biosynthetic gene clusters predicted to produce secondary metabolites, but it also revealed 1292 genes having no clear orthologs in the database [8].

Under laboratory conditions *S. espanaensis* was not able to produce any biologically active compounds. In order to optimize the production condition of this strain it was grown in different media such as SPY, GYM, HA and NL111. Cultivating the strain in SPY medium led to the formation of several compounds which were either not produced or produced in a low amount when other media were used. Four of these compounds were isolated and the structure of each was elucidated. The rare element scandium was added to the media in order to enhance the production of the different metabolites [3]. RNA sequencing (RNA-seq) analysis was performed after structure elucidation of all four compounds, not only to find genes involved in the biosynthesis of these compounds but also to detect other genes that are expressed significantly differently under different nutrient conditions. The GusA reporter gene system was also performed to verify the difference in the expression of some genes. Moreover, deletion of the genes *BN6\_01860* and *BN6\_67950* could prove their importance in the biosynthesis of indole-3-carboxylic acid and benzoic acid, respectively.

## Materials and Methods

### Strains, plasmids and culture conditions

*S. espanaensis* DSM44229 strain was purchased from DSMZ. *E. coli* XL1-blue (Stratagene) was used for cloning and *E. coli* ET12567 x pUZ8002 for intergeneric conjugation [9]. The strains used for antimicrobial activity testing were *Bacillus subtilis* COHN ATCC6051 (*B. subtilis*) [10], *Streptomyces albus* J1074 (*S. albus*) [11], *E. coli* Turbo (New England Biolabs), *Candida parapsilosis* (*C. parapsilosis*) DSM 5784, and *Fusarium verticillioides* DSM 62264 (*F. verticillioides*). The composition of the media used in this study was per 1 liter medium as follows: LB (growth medium of *E. coli* and *B. subtilis*): 21 g dry powder LB medium. Trypticase soy broth (TSB) medium (growth medium of *S. espanaensis* and *S. albus*): 30 g Caso Bouillon. SPY medium: 30 g soluble starch, 26 g trypton/pepton, 4 g polypepton, 2 g yeast extract, 15 mg MgSO<sub>4</sub>·7H<sub>2</sub>O, 50 mg CaCl<sub>2</sub>·2H<sub>2</sub>O, 35 mg ZnSO<sub>4</sub>·7H<sub>2</sub>O, 12 mg FeSO<sub>4</sub>·5H<sub>2</sub>O, 16 mg CuSO<sub>4</sub>·5H<sub>2</sub>O and 12 mg MnSO<sub>4</sub>·4H<sub>2</sub>O. GYM medium: 4 g glucose, 4 g yeast extract, 10 g malt extract, 1 g polypeptone, 2 g NaCl and 11.2 g MOPS. HA medium: 4 g glucose, 4 g yeast extract, and 10 g malt extract. NL111 medium: 20 g meat extract, 100 g malt extract and 10 g CaCO<sub>3</sub>. The media SPY, GYM, HA and NL111 were supplemented with 70 µM ScCl<sub>3</sub>·6H<sub>2</sub>O (Sigma-Aldrich, USA) and were used for cultivating *S. espanaensis*. YM medium (growth medium of *C. parapsilosis*): 3 g yeast extract, 3 g malt extract, 5 g peptone from soybean, 10 g glucose; PGA medium (growth medium of *F. verticillioides*): 26 g potato glucose. MS medium was used for the intergeneric conjugation: 20 g D-Mannitol, 20 g soya flour, and after autoclaving 20 µM CaCl<sub>2</sub> were added. The pH value of all media was adjusted with 1 N HCl or 1 N NaOH to 7.2 prior to autoclaving. 21 g (per 1 liter medium) agar were added before autoclaving to obtain solid media. Apramycin, spectinomycin and phosphomycin were used when appropriate with final concentrations of 50 µg/ml and 200 µg/ml. All media components were purchased from Roth, Germany; the trace elements used in SPY medium were purchased from Merck, Germany. The shuttle plasmid pKC1132 [12] was used to construct gene disruption mutants via homologous recombination. The plasmid pGUS was used for the GusA reporter gene system assay [13].

### Antimicrobial activity of *Saccharothrix espanaensis* culture extract and purified compounds

*In vitro* antimicrobial activity of *S. espanaensis* extract and single purified compounds was measured by disk diffusion method. All tested samples were dissolved in methanol and applied to sterile paper disks (6 mm in diameter). The paper disks treated with each sample were placed on top of the agar plates each of them seeded with one test organism. The inhibition zones of mycelia growth were measured after 18 h incubation at 37°C (*B. subtilis*, *E. coli* Turbo and *C. parapsilosis* DSM 5784), at 28°C (*S. albus* J1074) and after 48 h incubation at 28°C (*F. verticillioides*).

### Analytical HPLC/ESI-MS of *Saccharothrix espanaensis* extract

The HPLC-MS analysis was performed on an Agilent 1100 series LC/MS system with electrospray ionization (ESI) and mass spectrometry (HPLC/ESI-MS). The LC system was equipped with a waters X-bridge reverse phase column (4.6 × 100 mm, 3.5 µm). Compounds were resolved using a linear gradient of 5/95 (acetonitrile/

water, 0.5% acetic acid) to 95/5 (acetonitrile/water, 0.5% acetic acid) for 32 min (0.5 ml/min). The final solvent composition was held for 4 additional minutes.

### Isolation of natural compounds produced by *Saccharothrix espanaensis*

*S. espanaensis* was incubated in 50 × 500 ml flasks, each one contains 200 ml SPY medium, at 28°C on a rotary shaker at 180 rpm for 4 days. The culture broth was centrifuged, and the supernatant was acidified to pH 3.5 by addition of HCl. Culture fluid was extracted with 25 l ethyl acetate and filtrated using filter paper (MN 615. Ø 185 mm Machery-Nagel). The ethyl acetate phase was concentrated to dryness by using a rotary evaporator. After resuspension in 50 ml methanol/water (40:60) the crude extract was transferred to five Oasis HLB 20 35ss columns (6 g) (Waters). Compounds were eluted with increasing gradient of methanol in water (40:60, 50:50, 60:40, 65:35, 70:30, 80:20, 90:10, 100:0). The 50% and 60% methanol fractions, which exhibited antifungal and antimicrobial activity against both Gram-negative and Gram-positive bacteria, were concentrated to dryness by using a rotary evaporator. After resuspension of each fraction in 5 ml methanol further purified using preparative thin layer chromatography (TLC) (silica gel 60 F254, Macherey-Nagel) with a solvent system of dichloromethane/methanol/acetic acid (90/10/0.5 [v/v/v]). In total 5 plates were used.

The separated bands on the TLC plates were visualized using a UV detector at 254 nm and 360 nm. Silica gel bands of the 50% methanol fraction at the position of R<sub>f</sub> 0.65 and 0.7, which showed antimicrobial activity, were scrapped off and extracted four times with 100% methanol, then dried using a rotary evaporator. The same purification step was done for the silica gel bands of the 60% methanol fraction at the positions of R<sub>f</sub> 0.41 and 0.56. Further separation of the single compounds from the silica gel bands was done using preparative HPLC. A concentrated methanol solution of crude compound was applied to the column and eluted with acetonitrile/0.5% acetic acid as buffer A and water/0.5% acetic acid as buffer B at a flow rate of 2 mL min<sup>-1</sup>. Fractions containing the compounds were further purified by gel filtration on a Sephadex LH20 column (GE Healthcare) with 100% methanol. After evaporating the solvent purified compounds were dissolved in CD<sub>3</sub>OD and analyzed by <sup>1</sup>H-NMR (400 MHz) and <sup>13</sup>C-NMR (100 MHz) on a Bruker DRX-500 spectrometer (Bruker, Karlsruhe, Germany). 2D-NMR (<sup>1</sup>H/<sup>1</sup>H-COSY and HMBC) and ESI-MS were also used to confirm the structure of the compound.

### Structure elucidation of purified compounds from *Saccharothrix espanaensis* extract

Nuclear Magnetic Resonance spectra (NMR) of the purified compounds were measured on a Bruker AV 400 MHz or Varian VNMR-S 600 MHz spectrometer. NMR spectra were recorded in CD<sub>3</sub>OD at 22°C (Bruker) or 25°C (Varian), acetone-d<sub>6</sub> at 25°C or DMSO-d<sub>6</sub> at 35°C (supplementary material S3). The spectra were referenced based on solvent signals using the following values: CD<sub>3</sub>OD: δ<sub>H</sub>: 3.30 ppm, δ<sub>C</sub>: 49.0 ppm; Acetone-d<sub>6</sub>: δ<sub>H</sub>: 2.05 ppm, δ<sub>C</sub>: 29.8 ppm and DMSO-d<sub>6</sub>: δ<sub>H</sub>: 2.50 ppm, δ<sub>C</sub>: 39.5 ppm. NMR spectroscopic data for anthranilic acid, indol-3-carboxylic acid, benzoic acid, and phenylacetic acid are consistent with previously reported spectra. They are given in Supporting Information (Table S3).

## RNA sequencing

The pre-culture of *S. espanaensis* was prepared by inoculating 300 µl of 25% saccharose culture of *S. espanaensis* in 50 ml TSB medium. The pre-culture was grown for 48 h on a rotary shaker at 180 rpm at 28°C. Subsequently 3 ml of the pre-culture were transferred into two flasks, one of them containing 100 ml GYM medium and the other 100 ml SPY medium, both media were supplemented with 100 µM ScCl<sub>3</sub>·6H<sub>2</sub>O. The flasks were placed on the shaker for 4 days (96 h) at 180 rpm at 28°C. The entire liquid content of each flask was poured in 50 ml Falcon-Tube and spun at 5000 rpm for 10 min. Supernatants were extracted with an equal volume of ethyl acetate, then dried by vacuum and analyzed using HPLC-MS. Wet pellets were frozen at -80°C and stored on dry ice for library construction and sequencing in the following days. RNA sequencing was performed by Vertis Biotechnologie AG (Freising, Germany). Ribo-Zero Bacteria kit from Epicentre (Madison, USA) was used to remove the ribosomal RNA molecules from the isolated total RNA. The rRNA depleted RNA was fragmented with RNase III and samples were poly(A)-tailed using poly(A) polymerase. Then, the 5'PPP structures were removed using tobacco acid pyrophosphatase (TAP). Afterwards, an RNA adaptor was ligated to the 5'-monophosphate of the RNA. First-strand cDNA synthesis was performed using an oligo(dT)-adapter primer and M-MLV reverse transcriptase. The resulting cDNAs were amplified using a high-fidelity DNA polymerase, purified, and fragments in the size range of 250-500 bp were sequenced on an Illumina HiSeq 2000 system (San Diego, CA, USA) using 100 bp read length.

A total of 26378549 and 25087449 short reads (for GYM and SPY media, respectively) were received for analysis. Of these, 94% (GYM) and 90% (SPY) passed initial quality and length filtering with prinseq [14]. Reads were then mapped using Novoalign V3.00.04 (see supplementary material S1 for mapping statistics), and quantified using in-house software. For gene GC content and length bias control we have used the EDASeq R package [15]. Neither bias plots nor log-fold-change plots revealed significant dependencies between GC/length and read counts, so no EDASeq normalization was applied. Array Quality Metrics [16] quality report has not revealed any problems with the samples. Differential expression analysis was performed with DESeq package for the R statistical environment [17]. A comprehensive list of the 4827 loci from *S. espanaensis* and their respective transcription levels can be found in the supplementary material S4.

## Generation of *Saccharothrix espanaensis* reporter gene expression mutants

The Plasmids pGUS × p01860, pGUS × p68440 and pGUS × p78320 were constructed as follows. 300 to 400 bps fragments, which are located before the genes *BN6\_78320*, *BN6\_01860* and *BN6\_68440* and supposed to include the promoters of these genes, were PCR-amplified with the corresponding primers (Table 1) introducing 5' XbaI and 3' KpnI sites and cloned into XbaI- and KpnI- digested pGUS plasmid to create the above mentioned expression vectors. The resulting vectors, as well as the empty vector pGUS, were conjugated into *S. espanaensis* as previously described [18]. The apramycin- and spectinomycin-resistant exconjugants were grown in TSB medium at 28°C and prepared for Gus assay.

Primer	Sequence
Primer for single crossover mutants	
KO-01860-f	TATATCTAGACGCGGCTGGGCGACAAG
KO-01860-r	TATAGGATCCAGGCAGTGCTCGTTGAAC
KO-67950-f	TAATTCTAGATCGCGCTGGCCTGCCAC
KO-67950-r	TATAGGATCCAGGCACCGCACGGTCTC
Primer for promoters for GUS assay	
P-01860-f	TATATCTAGACGAAGGCGACGTCCG
P-01860-r	TCTAGGTACCGAGCTGGACGTTATC
P-78320-f	TACATCTAGAACCGCGTGTCTCGAC
P-78320-r	TATCGGTACCTCGACTCTGGCTGGTG
P-68440-f	TATATCTAGAGCAGCGCGCCTTGGTG
P-68440-r	TAGAGGTACCGCTTCGACACACCAC

**Table 1:** Primer used in this study.

## Gus assay

Each of the *S. espanaensis* expression mutants containing the plasmids pGUS × p01860, pGUS, pGUS × p68440 and pGUS × p78320 were inoculated into apramycin- and spectinomycin-containing TSB medium. After 48 h 1 ml of the culture was inoculated to 60 ml of SPY and 60 ml of GYM media and for 4 days cultivated. After cultivation, 30 ml of cultures were pelleted, washed once with distilled water, resuspended in 10 mL of Gus buffer 1 (50 mM phosphate buffer [pH 7.0], 5 mM dithiothreitol [DTT], 0.1% Triton X-100) then the French pressure cell press was used to break the cell walls. Lysates were centrifuged at 5,000 rpm for 10 min at 4°C. A 500 µl sample of the lysate was mixed with 500 µl of Gus buffer 3 (50 mM phosphate buffer [pH 7.0], 5 mM DTT, 0.1% Triton X-100 supplemented with 2 mM p-nitrophenyl-beta-D-glucuronide). The absorption was spectrophotometrically measured at 415 nm for 30 min (Ultraspec 2100pro UV/Visible Spectrophotometer, Amersham Biosciences). As a reference, a 1:1 mixture of lysate and Gus buffer1 was used. The enzymatic activity was calculated as previously described [19].

## Generation of *Saccharothrix espanaensis* strain deletion mutants

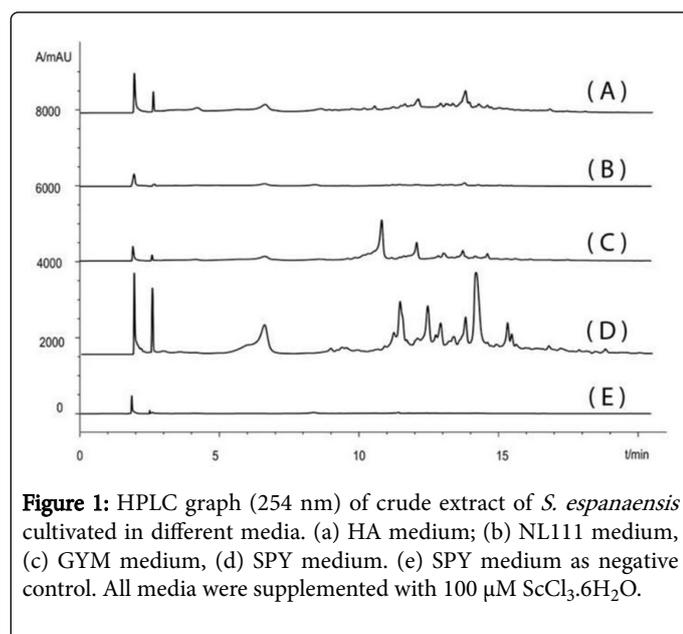
The *BN6\_01860* and *BN6\_67950* disruption vectors pKC1132 × KO01860 and pKC1132 × KO67950 were constructed as follows. Internal fragment from each gene was PCR-amplified with the corresponding primers introducing 5' XbaI and 3' BamHI sites. The primers used in this study are listed in Table 1. The resulting fragments were cloned into XbaI- and BamHI-digested pKC1132 plasmid to create the above mentioned disruption vectors. The constructed vectors were introduced into *S. espanaensis* by intergeneric conjugation using *E. coli* ET12567/pUZ8002 as previously described [18]. The single crossover mutants were grown in apramycin-containing TSB medium at 28°C. For production assay the resulting

mutants, together with the wild-type strain, were grown in SPY medium at 28°C for 4 days. After adjusting the pH to 3.5 the supernatants were extracted with an equal volume of ethyl acetate, and the dried crude extracts were analyzed by HPLC/ESI-MS.

## Results

### Purification, characterization and structure elucidation of four compounds produced by *Saccharothrix espanaensis*

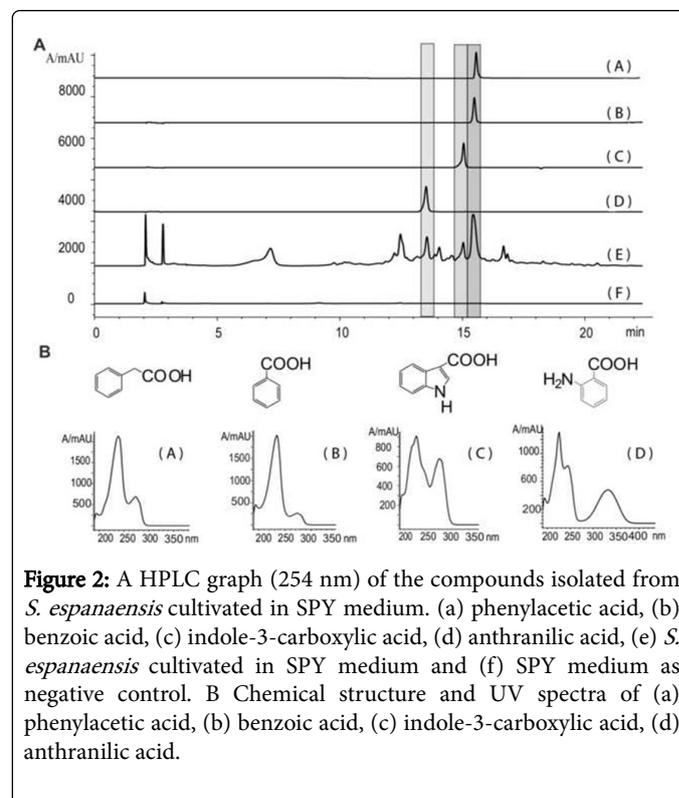
*S. espanaensis* DSM44229 was grown in different media as described in Material and Methods. Crude extracts were analyzed by HPLC/MS. As shown in Figure 1, several compounds were produced by *S. espanaensis* cultivated in SPY medium which were either not produced in any of the other used media or produced only in a low amount. Antibacterial and antifungal activity was observed in the crude extract of *S. espanaensis* cultivated in SPY medium but not in GYM medium. The diameter of the observed inhibition zones were ca. 10 mm for *E. coli*, *B. subtilis* and *C. parapsilosis* and 23 mm for *F. verticillioides*. The addition of scandium chloride to the media could enhance the production.



**Figure 1:** HPLC graph (254 nm) of crude extract of *S. espanaensis* cultivated in different media. (a) HA medium; (b) NL111 medium, (c) GYM medium, (d) SPY medium. (e) SPY medium as negative control. All media were supplemented with 100  $\mu$ M  $\text{ScCl}_3 \cdot 6\text{H}_2\text{O}$ .

Four compounds were isolated from the crude extract of *S. espanaensis* cultivated in SPY medium. The structures of these four compounds have been determined to be benzoic acid, phenylacetic acid, anthranilic acid and indole-3-carboxylic acid using  $^1\text{H}$ ,  $^{13}\text{C}$ ,

COSY, HSQC and HMBC NMR spectroscopic data (Figure 2). When trypton/pepton and polypepton were eliminated from the medium, the formation of all four compounds could not be detected.



**Figure 2:** A HPLC graph (254 nm) of the compounds isolated from *S. espanaensis* cultivated in SPY medium. (a) phenylacetic acid, (b) benzoic acid, (c) indole-3-carboxylic acid, (d) anthranilic acid, (e) *S. espanaensis* cultivated in SPY medium and (f) SPY medium as negative control. B Chemical structure and UV spectra of (a) phenylacetic acid, (b) benzoic acid, (c) indole-3-carboxylic acid, (d) anthranilic acid.

### Differential expression analysis of RNA-seq of *Saccharothrix espanaensis* cultivated in SPY medium and in GYM medium

A comparative RNA-seq analysis of *S. espanaensis* cultivated in SPY medium and in GYM medium was performed to detect genes which are expressed significantly differently under various nutrient conditions (Supplementary material S4). In total, a non-zero transcription (at least in one of the conditions) was observed for 8301 genes. Transcription changed less than 2 fold (up or down) for 3596 genes. More genes were significantly higher expressed in GYM medium than in SPY medium. 99 genes showed 100 fold overexpression in GYM medium in comparison to 10 genes in SPY medium. Also the number of genes with an overexpression factor between 5 and 100 fold was significantly higher in GYM medium (873) compared to SPY medium (558) (Table 2). The 10 genes with the highest overexpression factor in each medium are shown in Table 3.

GYM	SPY	Overexpression factor (fold)
Number of genes being overexpressed in one of both media		
99	10	>100
442	182	<100 to >10
431	376	<10 to >5
1324	1499	<5 to >2

3596		2
252	90	Genes transcribed only in one medium

**Table 2:** Comparison of gene transcription of *Saccharothrix espanaensis* in two different media.

Genes overexpressed in GYM medium			Genes overexpressed in SPY medium		
Gene ID	Putative function of the deduced protein	Fold change	Gene ID	Putative function of the deduced protein	Fold change
<i>BN6_60890</i>	Secreted protein	2046	<i>BN6_05370</i>	Superoxide dismutase, Mn-dependent	702
<i>BN6_64430</i>	Chitin binding domain 3 protein	1718	<i>BN6_28080</i>	Microsomal epoxide hydrolase	403
<i>BN6_60880</i>	Secreted protein	1589	<i>BN6_28110</i>	Transposase	338
<i>BN6_30060</i>	Hypothetical protein	1562	<i>BN6_28100</i>	Elongation factor GreAB	331
<i>BN6_46130</i>	Alpha/beta hydrolase fold containing protein	1497	<i>BN6_28120</i>	FAD dependent oxidoreductase	307
<i>BN6_31430</i>	Peptidase S8/S53, subtilisin kexin sedolisin	1328	<i>BN6_49230</i>	Ribonucleoside-Diphosphate reductase	218
<i>BN6_43590</i>	MFS transporter arabinose efflux permease	1306	<i>BN6_26530</i>	Sulfide quinone reductase	191
<i>BN6_46200</i>	Hydroxymethylglutaryl- CoA synthase	1060	<i>BN6_33060</i>	NADPH- dependent FMN reductase	140

**Table 3:** Highly overexpressed genes of *Saccharothrix espanaensis* cultivated in GYM or SPY medium.

Among the 99 genes with a high overexpression rate in GYM medium are 8 genes encoding protease/peptidase proteins, 6 genes encoding proteins with a chitinase/chitin binding domain, 4 genes encoding secreted proteins and 9 transporter genes. In SPY medium among the overexpressed genes are two ribonucleoside diphosphate reductase genes, one superoxide dismutase gene, one epoxide hydrolase gene, one lysine exporter gene and one sulfide quinone reductase gene. Moreover, 7 genes encoding degrading enzymes are more than 20-fold overexpressed in SPY compared to GYM medium.

### RNA sequencing reveals genes putatively involved in the biosynthesis of anthranilic acid, indole-3-carboxylic acid, benzoic acid and phenyl acetic acid

Automatically annotated genes of *S. espanaensis* were searched against the protein sequence databases to identify genes involved in the biosynthesis of anthranilic acid, indole-3-carboxylic acid, benzoic acid and phenylacetic acid. The search was focused on proteins known to be involved in the biosynthesis of these compounds in other organisms. By combining the genome analysis with the transcriptomic results, hypothetical biosynthetic pathway of three of the purified compounds are proposed.

#### Biosynthesis of benzoic acid

In plants and bacteria, it has been reported that benzoic acid is produced from L-phenylalanine through different pathways [20,21].

Genes coding for biochemically characterized enzymes involved in the biosynthesis of benzoic acid from different organisms were used as queries. In *Streptomyces maritimus*, the first enzyme involved in the production of benzoic acid is phenylalanine ammonia lyase (PAL), encoded by encP. It converts L-phenylalanine to trans cinnamic acid [22]. *S. espanaensis* has five genes coding for enzymes, which are homologous to the encP protein, but only one of them, which is *BN6\_75400* is high expressed in SPY medium. This gene is about 16.6 time more expressed in SPY medium compared to GYM medium. The next step during benzoic acid biosynthesis is the activation of cinnamic acid by an acyl-CoA ligase leading to cinnamoyl-CoA. In the case of *Streptomyces coelicolor* the enzyme ScCCL, encoded by SCO4383, catalyzes this step [23]. BLASTp analyses showed that in *S. espanaensis* there are eight genes coding for CoA ligase and are homologous to ScCCL, one of them *BN6\_31010* is overexpressed in SPY medium. During the next biosynthetic step cinnamoyl-CoA is converted into  $\beta$ -ketoacyl propionyl-CoA by the bifunctional enzyme cinnamoyl-CoA hydratase-dehydrogenase, which is encoded by PhCHD in *Petunia hybrid* [24]. *S. espanaensis* has two genes *BN6\_18530* and *BN6\_67950* coding for enzymes that are homologous to the protein PhCHD. Both genes are more expressed in SPY medium. The next biosynthetic step involves the conversion of  $\beta$ -ketoacyl propionyl-CoA to benzaldehyde-CoA by a ketoacyl thiolase that is encoded by encJ from *Streptomyces maritimus* [25] and PhKAT1 from *Petunia hybrida*. BLAST search revealed that the deduced amino acid sequences of five genes from *S. espanaensis* are similar to PhKAT1 and EncJ. Four of these genes *BN6\_07730*, *BN6\_73090*, *BN6\_18540* and *BN6\_67940* are more

expressed in SPY medium (supplementary material S2). In continuation benzaldehyde-CoA is converted to benzaldehyde, which is subsequently oxidized to benzoic acid. The enzyme BALDH from *Antirrhinum majus* has a benzaldehyde dehydrogenase activity [26]. In addition, the enzyme benzaldehyde dehydrogenase II, encoded by *xylC*, from *Acinetobacter calcoaceticus* oxidizes benzaldehyde to benzoic acid and has esterase activity [27]. BLAST search shows that ten genes from *S. espanaensis* code for enzymes that are homologous to both BALDH and XylC (supplementary material S2), seven of them, *BN6\_30230*, *BN6\_77230*, *BN6\_34260*, *BN6\_78480*, *BN6\_13450*, *BN6\_61630* and *BN6\_34200* are more expressed in SPY medium (supplementary material S2).

### Biosynthesis of phenylacetic acid

L-phenylalanine degradation via “Ehrlich pathway” in microorganisms serves as a biosynthetic pathway for phenylacetic acid [28]. The first step of this pathway involves deamination of L-phenylalanine and formation of phenylpyruvate by a phenylalanine transaminase. The Enzyme ArAT-II encoded by OOC\_04737 from *Thermococcus litoralis* catalyzes this reaction [29]. BLASTp analysis revealed that the sequence of ArAT-II is homologous to the deduced amino acid sequence of seven genes in *S. espanaensis*. Five of them: *BN6\_01860*, *BN6\_68610*, *BN6\_82750*, *BN6\_08560* and *BN6\_79710* are upregulated in SPY medium (supplementary material S2). In the next step of L-phenylalanine degradation phenyl pyruvic acid is decarboxylated to phenylacetaldehyde by phenylpyruvate decarboxylase, which is encoded by *ipdC* in *Azospirillum brasilense* [30]. BLAST analysis reveals that the deduced amino acid sequences of five genes from *S. espanaensis* are similar in their amino acid sequence to PPDC, encoded by *ipdC*. Three of these genes, *BN6\_50800*, *BN6\_70260* and *BN6\_67160*, are more expressed in SPY medium (see supplementary material S2). The final step involves oxidation of phenylacetaldehyde to phenylacetate by a phenylacetaldehyde dehydrogenase. In *E. coli*, it has been shown that this step is catalyzed by *PadA* [31]. There are six genes in *S. espanaensis* coding for enzymes that are similar to *PadA*, four of them: *BN6\_34260*, *BN6\_30230*, *BN6\_34200* and *BN6\_77230* are more expressed in SPY medium (supplementary material S2).

### Biosynthesis of anthranilic acid

In order to find out the genes that code for enzymes putatively involved in the biosynthesis of anthranilic acid, genes coding for the three enzymes of “tryptophan to anthranilate” pathway from *Streptomyces coelicolor* were used as queries [32]. In *Streptomyces coelicolor* the first enzyme of this pathway is a tryptophan 2,3-dioxygenase (TDO), coded by *SCO3646* that cleaves tryptophan to yield N-formyl-kynurenine. The second step involves a removal of the formyl group and forming of L-kynurenine by N-formyl-kynurenine formamidase (KFA), coded by *SCO3644*. In the final step, L-kynurenine is converted into anthranilate by a kynureninase, coded by *SCO3645*. The three genes *BN6\_03100*, *BN6\_14890* and *BN6\_03090* from *S. espanaensis* are coding for enzymes that are homologous to *SCO3646*, *SCO3644* and *SCO3645*, respectively. All three genes are significantly more expressed in SPY medium compared to GYM medium (supplementary material S2).

On the other hand, the two genes *BN6\_68440* and *BN6\_00210*, which potentially encode for anthranilate synthase component 1 and 2, respectively, are responsible for the forming of anthranilate from

chorismic acid in the first step of tryptophan biosynthesis pathway. Both genes are overexpressed in SPY medium.

### Biosynthesis of indole-3-carboxylic acid

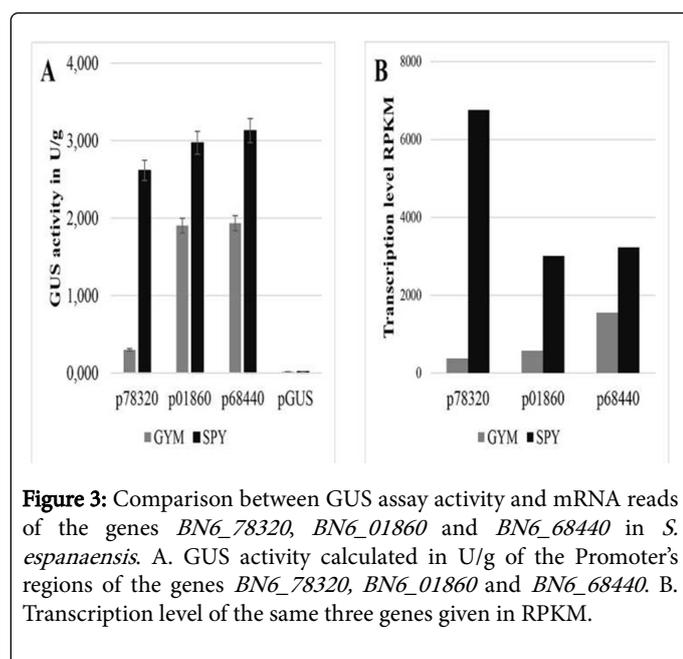
In bacteria, it has been proved that indole-3-carboxylic acid is a degradation product of L-tryptophan [33]. The complete biosynthetic pathway of indole-3-carboxylic acid is not studied so far, but it was reported that L-phenylalanine is converted in three steps to indole-3-acetic acid, which is catalyzed by different enzymes to indole-3-carboxylic acid. The biosynthesis of indole-3-acetic acid has been well studied in various plant-associated bacteria [34]. The first step of this pathway is the transfer of the amino group from L-tryptophan to 2-oxoglutarate by an aminotransferase, forming indole-3-pyruvate and glutamate. In *Azospirillum brasilense* two genes *hisC1* and *hisC2* coding for tryptophan aminotransferase and are homologous to the amino acid sequence of the two genes *BN6\_01860* and *BN6\_68610* from *S. espanaensis* (supplementary material S2). In SPY medium both genes *BN6\_01860* and *BN6\_68610* are more expressed. The *IpdC*, encoded by *ipdC* from *Azospirillum brasilense* is a pyruvate decarboxylase and can use both indole-3-pyruvate and phenylpyruvate as substrate [35,36]. This enzyme catalyzes the decarboxylation of indole-3-pyruvate to yield indole-3-acetaldehyde. Its homologous enzymes in *S. espanaensis* are already mentioned in the phenylacetic acid biosynthesis (supplementary material S2). The next step is the oxidation of indole-3-acetaldehyde to indole-3-acetic acid by an indole-3-acetaldehyde dehydrogenase. It has been demonstrated that this step in the fungus *Ustilago maydis* is carried out by the enzyme *lad1* [37]. *S. espanaensis* has five genes coding for enzymes that are homologous to *lad1* and are overexpressed in SPY medium compared to GYM medium. The genes responsible for the biosynthesis of indole-3-carboxylic acid from indole-3-acetic acid are unclear and there is, to our knowledge, no studies about this biosynthetic pathway.

**Disruption of *BN6\_01860*, *BN6\_67950* genes in *Saccharothrix espanaensis*:** In order to investigate the role of *BN6\_01860* in the biosynthesis of indole-3-carboxylic acid and the role of *BN6\_67950* in the biosynthesis of benzoic acid we disrupted both genes by single cross over homologous recombination and compared the production of the resulted mutants with the production of the wildtype strain using HPLC/ESI-MS analysis. The pKC1132 plasmid was used to construct plasmids carrying internal 0.78 to 1.52-kb fragments of the target *BN6\_01860* and *BN6\_67950* genes. The plasmids pKC1132 × KO01860 and pKC1132 × KO67950 were introduced into *S. espanaensis* by intergeneric conjugation. The gene *BN6\_01860*, which encodes a histidinol-phosphate aminotransferase, is suggested to be responsible for the deamination of L-tryptophan and L-phenylalanine. HPLC analysis of the crude extract from the *S. espanaensis*  $\Delta$ *BN6\_01860* mutant demonstrated that indole-3-carboxylic acid was not produced. Furthermore, a decrease in the production of phenylacetic acid production has been also observed. The gene *BN6\_67950* that encode a 3-hydroxyacyl-CoA dehydrogenase could catalyze the conversion of cinnamoyl-CoA to  $\beta$ -ketoacyl propionyl-CoA in the biosynthesis of benzoic acid. In *S. espanaensis*  $\Delta$ *BN6\_67950* mutant a decrease in the production of benzoic acid has been observed.

### GusA reporter gene assay verify RNA-seq analysis results

Three different genes from *S. espanaensis* were randomly chosen to examine their transcription level using the GUS reporter gene test system [13] after cultivating the strain in GYM and in SPY medium.

0.3-0.4 kb-promotor regions belonging to the genes *BN6\_01860*, *BN6\_68440* and *BN6\_78320* were cloned in the transcription fusion plasmid pGUS. The constructed plasmids, pGUS × 01860, pGUS × p68440 and pGUS × p78320 were introduced into *S. espanaensis* by intergeneric conjugation. After cultivating each of the resulting mutants in both SPY and GYM media, GUS activity was determined in U/g. *S. espanaensis* × pGUS was used as negative control. GUS assay results showed that the genes were higher expressed in SPY medium than in GYM medium (Figure 3).



**Figure 3:** Comparison between GUS assay activity and mRNA reads of the genes *BN6\_78320*, *BN6\_01860* and *BN6\_68440* in *S. espanaensis*. A. GUS activity calculated in U/g of the Promoter's regions of the genes *BN6\_78320*, *BN6\_01860* and *BN6\_68440*. B. Transcription level of the same three genes given in RPKM.

### Transcriptome analysis of putative biosynthetic gene clusters

*S. espanaensis* complete genome sequencing revealed the presence of 26 biosynthetic gene clusters potentially encoding secondary metabolites [8]. Aside from saccharomicin A and B gene cluster there are terpene, PKS, NRPS, PKS/NRPS hybrid and other types of clusters [8]. Transcriptome analysis showed that most of the genes from the saccharomicin gene cluster are not expressed which explains the difficulty of producing the saccharomicins A and B under laboratory conditions. All genes of one PKS/NRPS hybrid cluster (cluster 8) were significantly overexpressed in GYM medium. Furthermore, 7 genes from cluster 8, which have high transcription level in GYM medium, are not transcribed in SPY medium. The deduced amino acid sequences of 16 genes of this cluster are similar to proteins involved in leinamycin biosynthesis [38]. None of the other cluster was differently expressed in either SPY or GYM medium.

### Discussion

It has been shown that media composition can influence the production of natural products in microorganisms. In some cases production of a compound may only be possible when the strain is cultivated in a particular medium. Gene clusters, which were not expressed in one medium, can be highly expressed in another [1]. Today, finding new and novel compounds requires the use of different media during the screening process. This was demonstrated in our study where *S. espanaensis* was cultivated in different media and the generation of novel compounds in SPY medium was observed. SPY

medium contains starch and a higher proportion of amino acids than other media used in our study. The produced compounds determined to be phenylacetic acid, benzoic acid, indole-3-carboxylic acid and anthranilic acid. Anthranilic acid and indole-3-carboxylic acid may originate from the degradation of the aromatic amino acids tryptophan [33,39]. Whereas benzoic acid and phenylacetic acid may originate from phenylalanine [21,40].

In order to investigate the influence of SPY medium on gene expression total transcriptome sequencing was performed. Strand-specific total RNA sequencing was performed 96 hours after inoculation to establish the differences in gene expression between SPY and GYM medium. Interestingly, many genes encoding proteases or transporter proteins were found to be highly expressed in GYM medium, since GYM medium contains a higher portion of undigested proteins.

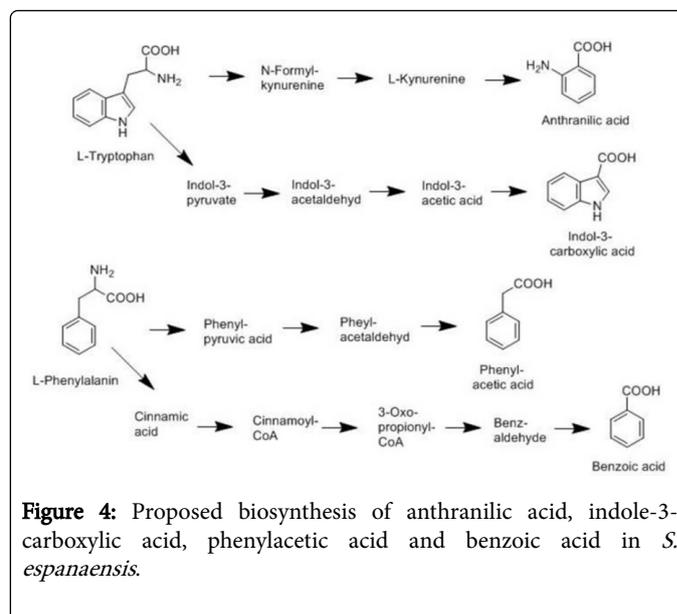
There were also highly overexpressed genes in SPY medium as well. These encoded enzymes involved in biodegradation of different compounds such as amino acids. Among them are seven genes encoding amino acid degrading enzymes such as a D-amino acid hydrogenase, a methionine gammalyase, NAD-specific glutamate dehydrogenase, a proline dehydrogenase, a glycine dehydrogenase, and an arginine deiminase. This can be explained by the high concentration of amino acids in SPY medium. Also a gene encoding an epoxide hydrolase is highly overexpressed. Epoxide hydrolases convert epoxides, which are intermediates in the degradation of aromatic compounds, to transdihydrodiols which are used as substrates by further degrading enzymes. Other overexpressed gene is that of a superoxide dismutase and two genes encoding ribonucleoside diphosphate reductase. Superoxide dismutases are enzymes considered to be important in the protection of aerobes against oxidant damage [41]. As degradation processes lead to oxidative stress in the cell overexpression of superoxide dismutase genes is an important protection mechanism. Ribonucleoside diphosphate reductases catalyze the formation of deoxyribonucleotides from ribonucleotides maintaining controlled dNTP pool sizes to the fidelity of DNA replication. It has been described that organic molecules trigger the expression of a ribonucleoside diphosphate reductase gene in *E. coli* [42].

When we looked for genes encoding for enzymes putatively involved in the biosynthesis of benzoic acid, phenylacetic acid, anthranilic acid and indole-3-carboxylic acid we detected a few genes which were highly overexpressed. Based on the transcriptome analysis results and the comparison with the published biosynthetic pathways we assume that the four purified compounds are degradation products of the aromatic amino acids L-phenylalanine and L-tryptophan.

L-phenylalanine degradation serves as a biosynthetic pathway for both benzoic acid and phenylacetic acid [21,40]. Different catabolism routes of phenylalanine lead to the production of benzoic acid [20], which has antifungal and antibacterial activity against both Gram positive and Gram negative bacteria [43,44]. The gene *encP* from *Streptomyces maritimus* encodes a rare bacterial phenylalanine ammonia-lyase [22] and is homologous to the predicted protein sequence of a hisitidine ammonia-lyase, encoded by *BN6\_75400* in *S. espanaensis*. The *BN6\_75400* gene was strongly overexpressed in SPY medium indicating that in *S. espanaensis* phenylalanine is converted to cinnamic acid by this enzyme. Further genes which might be involved in the conversion of cinnamic acid to benzoic acid are coding enzymes homologous to an acyl-CoA ligase gene, encoded by *SCO4383* from *Streptomyces coelicolor* [23], a 3-cinnamoyl-CoA hydratase-

dehydrogenase gene, encoded by CHD from *petunia hybrid* [24], a 3-ketoacyl-CoA thiolase, encoded by *encJ* from *Streptomyces maritimus* [25] and a putative aldehyde dehydrogenase, encoded by *xylC* from *Acinetobacter calcoaceticus* [27]. Different genes from *S. espanaensis* with predicted similar function to the above mentioned genes are more expressed in SPY medium than in GYM medium (Table 3), indicating that benzoic acid is also a phenylalanine degradation product in *S. espanaensis*. It has been verified that *Streptomyces maritimus* can use starch as a carbon source for benzoate production [45] but the mechanism was not described. SPY medium contains an excess of starch that might indicate that *S. espanaensis* also assimilates starch for production of benzoate in a similar way. The inactivation of *BN6\_67950* gene that encodes for a 3-hydroxyacyl-CoA dehydrogenase led to a decrease in the production of benzoic acid. In *Streptomyces maritimus* it has been shown that genes involved in fatty acid  $\beta$ -oxidation can take over the function of genes responsible for the biosynthesis of the benzoic acid [45]. That could explain the results of *BN6\_67950* inactivation, since the production of benzoic acid did not stop but only decreased.

Phenylacetic acid has been described in literature as an antifungal and antibiotic agent [46] as well as a natural flavor compound [40]. Production of phenylacetic acid from L-phenylalanine has been reported in bacteria [28,40]. It was suggested that in the phototrophic bacterium *Rubrivivax benzoatilyticus* JA2 phenylacetic acid is produced from the catabolism of L-phenylalanine via “Ehrlich pathway” [47]. “Ehrlich pathway” describes the catabolism of certain amino acids by three steps: Transamination, decarboxylation and reduction or oxidation of fusel aldehydes into fusel alcohols or acids (Figure 4). *S. espanaensis* genome contains five gene coding for aromatic aminotransferase enzymes, which are homologous to the phenylalanine transaminase (ArAT-II) from *Thermococcus litoralis*, and are upregulated in SPY medium. One of these five enzymes should be responsible for the conversion of phenylalanine to phenyl pyruvate. For the decarboxylation of phenyl pyruvate in *S. espanaensis* there are three genes coding for enzymes that are homologous to the phenyl pyruvate decarboxylase (PPDC) from *Azospirillum brasilense*. These three genes are overexpressed in SPY medium, indicating that one or more of them is responsible for the catalysis of this step in the biosynthesis. One of four genes from *S. espanaensis*, which encode for enzymes with high homology to the phenylacetaldehyde dehydrogenase (PadA) from *E. coli* should be responsible for the oxidation of phenylacetaldehyde to phenylacetic acid. These genes are upregulated in SPY medium.



In *Streptomyces coelicolor*, it has been demonstrated that tryptophan catabolism leads to the production of anthranilate through a three-step degradation “tryptophan to anthranilate” pathway, which is a branch of kynurenine pathway [32]. In *S. espanaensis* transcriptome analysis revealed that the three genes *BN6\_03100*, *BN6\_14890* and *BN6\_03090* encode for three enzymes, which might be involved in the production of anthranilic acid from L-tryptophan through the same pathway (Figure 4). Like in *Streptomyces coelicolor*, the presence of tryptophan has a positive feedback on the expression of all three genes [32]. Although a gene encoding for an anthranilic acid synthase is slightly upregulated in SPY medium the activity of this enzyme might be allosteric inhibited by the accumulation of L-anthranilic acid. Surprisingly, even genes encoding an anthranilate phosphoribosyltransferase and a phosphoribosylanthranilate isomerase, which are two enzymes participating in the biosynthesis of tryptophan from anthranilic acid, were two times more expressed in SPY medium compared to GYM medium. This indicates that in *S. espanaensis*, like in *Streptomyces coelicolor* [48], tryptophan biosynthesis is not subjected to a negative control by the levels of tryptophan but it is also controlled by the general growth requirement of the cell.

The last product isolated during our studies is the indolic tryptophan metabolite indole-3-carboxylic acid. Indole-3-carboxylic acid is reported as an antifungal and antibiotic agent [49]. Furthermore, indole-3-carboxylic acid has been described as an auxin related molecule without hormonal activity [50]. In *Chromobacterium violaceum* it has been proved that indole-3-carboxylic acid is a degradation product of indole-3-acetic acid, which is produced from L-tryptophan [33] (Figure 4). The biosynthesis of indole-3-acetic acid from L-tryptophan has been demonstrated in different microorganisms [37,51]. In *S. espanaensis* we postulate that L-tryptophan is even the starting molecule for indole-3-carboxylic acid formation. The gene *BN6\_01860* that encodes a histidinolphosphate aminotransferase shows in its amino acid sequence high homology to AAT1 from *Azospirillum brasilense*. The latter is responsible for conversion of L-tryptophan to indole-3-pyruvic acid. Other enzymes that are potentially involved in the catabolism of indole-3-pyruvic acid to indole-3-acetic acid are encoded by genes with homology to *ipdC*

from *Azospirillum brasilense* and *iad1* from *Ustilago maydis*. Further genes, which are required to complete this pathway, are unclear and there are so far no studies about them.

The inactivation of the gene *BN6\_01860*, which encodes a histidinol-phosphate aminotransferase, showed that indole-3-carboxylic acid was not produced anymore and a decrease in the production of phenylacetic acid has been observed. This result indicates that this gene is involved in the biosynthesis of both indole-3-carboxylic acid and phenyl acetic acid.

## Conclusion

Our results show that variation in cultivation conditions of *S. espanaensis* led to production of biologically active compounds. Furthermore, we could show that the compounds purified from SPY medium derive from the amino acids L-tryptophan and L-phenylalanine, which are both medium components. We could also show that the gene *BN6\_01860* is important in the biosynthesis of indole-3-carboxylic acid and phenylacetic acid. Transcriptome results indicate genes either involved in primary metabolism or in the formation of small natural products were significantly overexpressed in SPY medium. They also showed that most of the genes involved in the biosynthesis of saccharomicins A and B are not transcribed either in SPY or in GYM medium, which explains the difficulties in producing these both antibiotics under laboratory conditions.

## Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgement

Financial support for this research was provided by the RTG 1976 (cofactor dependent enzymes) funded by the Deutsche Forschungsgemeinschaft (DFG). We also thank the Albert-Ludwigs University of Freiburg for supporting this study.

## Supporting Information

Spectroscopic data of the compounds and information about biosynthetic genes and RNA-Seq read mapping and quantification statistics are listed in the supporting information for further reference.

## References

1. Bode HB, Bethe B, Höfs R, Zeeck A (2002) Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. *Chem Bio Chem* 3: 619-627.
2. Tormo JR, García JB, DeAntonio M, Feliz J, Mira A (2003) A method for the selection of production media for actinomycete strains based on their metabolite HPLC profiles. *J Ind Microbiol Biotechnol* 30: 582-588.
3. Kawai K, Wang G, Okamoto S, Ochi K (2007) The rare earth, scandium, causes antibiotic over production in *Streptomyces* spp. *FEMS Microbiol Lett* 274: 311-315.
4. Labeda DP, Kroppenstedt RM (2000) Phylogenetic analysis of *Saccharothrix* and related taxa: proposal for Actinosynnemataceae fam. no. *Int J Syst Evol Microbiol* 50: 331-336.
5. Chaudhary HS, Soni B, Shrivastava AR, Shrivastava S (2013) Diversity and versatility of actinomycetes and its role in antibiotic production. *J Appl Pharmaceutical Science* 3: 83-94.
6. Kong F, Zhao N, Siegel MM, Janota K, Ashcroft JS, et al. (1998) Saccharomicins, Novel Heptadecaglycoside Antibiotics Effective against Multidrug-Resistant Bacteria. *J Am Chem Soc* 120: 13301-13311.
7. Singh MP, Petersen PJ, Weiss WJ, Kong F, Greenstein M, et al. (2000) Saccharomicins, Novel Heptadecaglycoside Antibiotics Produced by *Saccharothrix espanaensis*: Antibacterial and Mechanistic Activities. *Antimicrob Agents Chemother* 44: 2154-2159.
8. Strobel T, Al-Dilaimi A, Blom J, Gessner A, Kalinowski J, et al. (2012) Complete genome sequence of *Saccharothrix espanaensis* DSM 44229 (T) and comparison to the other completely sequenced Pseudonocardiaaceae. *BMC Genomics* 13: 465.
9. <https://www.jic.ac.uk/science/molmicro/Strepmanual/Manual.htm>
10. Spizizen J (1958) Transformation of Biochemically Deficient Strains of *Bacillus Subtilis* by Deoxyribonucleate. *Proc Natl Acad Sci USA* 44: 1072-1078.
11. Chater KF, Wilde LC (1976) Restriction of a bacteriophage of *Streptomyces albus* G involving endonuclease *salI*. *J Bacteriol* 128: 644-650.
12. Bierman M, Logan R, O'Brien K, Seno ET, Rao RN, et al. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* 116: 43-49.
13. Myronovskyi M, Welle E, Fedorenko V, Luzhetskyy A (2011) Beta-Glucuronidase as a sensitive and versatile reporter in Actinomycetes. *Appl Environ Microbiol* 77: 5370-5383.
14. Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27: 863-864.
15. Risso D, Schwartz K, Sherlock G, Dudoit S (2011) GC-Content Normalization for RNA-Seq Data. *BMC Bioinformatics* 12: 480.
16. Kauffmann A, Gentleman R, Huber W (2009) Array Quality Metrics-A bioconductor package for quality assessment of microarray data. *Bioinformatics* 25: 415-416.
17. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11: 106.
18. Gessner A, Heitzler T, Zhang S, Klaus C, Murillo R, et al. (2015) Changing biosynthetic profiles by expressing *bldA* in *Streptomyces* strains. *Chem Bio Chem* 16: 2242-2252.
19. Siegl T, Tokovenko B, Myronovskyi M, Luzhetskyy A (2013) Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metab Eng* 19: 98-106.
20. Abd El-Mawla A, Beerhues L (2002) Benzoic acid biosynthesis in cell cultures of *Hypericum androsaemum*. *Planta* 214: 727-733.
21. Hertweck C, Moore BS (2000) A Plant-like Biosynthesis of Benzoyl-CoA in the Marine Bacterium "Streptomyces maritimus". *Tetrahedron* 56: 9115-9120.
22. Xiang L, Moore BS (2002) Inactivation, complementation, and heterologous expression of *encP*, a novel bacterial phenylalanine ammonia-lyase gene. *J Biol Chem* 277: 32505-32509.
23. Kaneko M, Ohnishi Y, Horinouchi S (2003) Cinnamate:coenzyme A ligase from the filamentous bacterium *Streptomyces coelicolor* A3(2). *J Bacteriol* 185: 20-27.
24. Qualley AV, Widhalm JR, Adebesein F, Kish CM, Dudareva N (2012) Completion of the core  $\beta$ -oxidative pathway of benzoic acid biosynthesis in plants. *Proc Natl Acad Sci USA* 109: 16383-16388.
25. Xiang L, Moore BS (2003) Characterization of Benzoyl Coenzyme A Biosynthesis Genes in the Enterocin-Producing Bacterium *Streptomyces maritimus*. *J Bacteriol* 185: 399-404.
26. Long MC, Nagegowda DA, Kaminaga Y, Ho KK, Kish CM, et al. (2009) Involvement of snapdragon benzaldehyde dehydrogenase in benzoic acid biosynthesis. *Plant J* 59: 256-265.
27. MacKintosh RW, Fewson CA (1988) Benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase II from *Acinetobacter calcoaceticus*. Substrate specificities and inhibition studies. *Biochem J* 255: 653-661.
28. Prasuna ML, Mujahid M, Sasikala C, Ramana CV (2012) L-Phenylalanine catabolism and Lphenyllactic acid production by a phototrophic bacterium, *Rubrivivax benzoatilyticus* JA2. *Microbiol Res* 167: 526-531.

29. Andreotti G, Cubellis MV, Nitti G, Sannia G, Mai X, et al. (1994) Characterization of aromatic aminotransferases from the hyperthermophilic archaeon *Thermococcus litoralis*. *Eur J Biochem* 220: 543-549.
30. Spaepen S, Versees W, Gocke D, Pohl M, Steyaert J, et al. (2007) Characterization of Phenylpyruvate Decarboxylase, Involved in Auxin Production of *Azospirillum brasilense*. *J Bacteriol* 189: 7626-7633.
31. Ferrández A, Prieto MA, García JL, Díaz E (1997) Molecular characterization of PadA, a phenylacetaldehyde dehydrogenase from *Escherichia coli*. *FEBS Letters* 406: 23-27.
32. Zummo FP, Marineo S, Pace A, Civiletti F, Giardina A, et al. (2012) Tryptophan catabolism via kynurenine production in *Streptomyces coelicolor*: identification of three genes coding for the enzymes of tryptophan to anthranilate pathway. *Appl Environ Microbiol* 94: 719-728.
33. Davis PJ, Gustafson ME, Rosazza JP (1976) Formation of indole-3-carboxylic acid by *Chromobacterium violaceum*. *J Bacteriol* 126: 544-546.
34. Jijon-Moreno S, Marcos-Jimenez C, Pedraza RO, Ramirez-Mata A, de Salamone IG, et al. (2015) The ipdC, hisC1 and hisC2 genes involved in indole-3-acetic production used as alternative phylogenetic markers in *Azospirillum brasilense*. *Antonie van Leeuwenhoek* 107: 1501-1517.
35. Versées W, Spaepen S, Vanderleyden J, Steyaert J (2007) The crystal structure of phenylpyruvate decarboxylase from *Azospirillum brasilense* at 1.5 Å resolution. Implications for its catalytic and regulatory mechanism. *The FEBS Journal* 274: 2363-2375.
36. Zimmer W, Wesche M, Timmermans L (1998) Identification and isolation of the indole-3-pyruvate decarboxylase gene from *Azospirillum brasilense* Sp7: Sequencing and functional analysis of the gene locus. *Curr Microbiol* 36: 327-331.
37. Reineke G, Kahmann R, Basse CW (2008) Indole-3-acetic acid (IAA) biosynthesis in the smut fungus *Ustilago maydis* and its relevance for increased IAA levels in infected tissue. *Mol Plant Pathol* 9: 339-355.
38. Strobel T (2013) Genomische und molekularbiologische Untersuchungen zum Biotransformation swirt *Saccharothrix espanaensis*. PhD Thesis. Abert-Ludwigs University, Freiburg, Germany.
39. Salcher O, Lingens F (1980) Metabolism of tryptophan by *Pseudomonas aureofaciens* and its relationship to pyrrolnitrin biosynthesis. *J Gen Microbiol* 121: 465-471.
40. Helinck S, Le Bars D, Moreau D, Yvon M (2004) Ability of Thermophilic Lactic Acid Bacteria to Produce Aroma Compounds from Amino Acids. *Appl Environ Microbiol* 70: 3855-3861.
41. Cabiscol E, Tamarit J, Ros J (2000) Oxidative stress in bacteria and protein damage by reactive oxygen species. *Int Microbiol* 3: 3-8.
42. Sitjes J, Ysern P, Barbe J, Llagostera M (1992) Induction of ribonucleoside diphosphate reductase gene transcription by chemicals in *Escherichia coli*. *Mutagenesis* 7: 47-49.
43. Cueva C, Moreno-Arribas MV, Martín-Álvarez PJ, Bills G, Vicente MF, et al. (2010) Antimicrobial activity of phenolic acids against commensal, probiotic and pathogenic bacteria. *Res Microbiol* 161: 372-382.
44. Krebs HA, Wiggins D, Stubbs M, Sols A, Bedoya F (1983) Studies on the mechanism of the antifungal action of benzoate. *Biochem J* 214: 657-663.
45. Noda S, Kitazono E, Tanaka T, Ogino C, Kondo A (2012) Benzoic acid fermentation from starch and cellulose via a plant-like  $\beta$ -oxidation pathway in *Streptomyces maritimus*. *Microbial Cell Factories* 11: 49.
46. Hwang BK, Lim SW, Kim BS, Lee JY, Moon SS (2001) Isolation and In Vivo and In Vitro Antifungal Activity of Phenylacetic Acid and Sodium Phenylacetate from *Streptomyces humidus*. *Appl Environ Microbiol* 67: 3739-3745.
47. Hazelwood LA, Daran JM, van Maris AJ, Pronk JT, Dickinson JR (2008) The Ehrlich pathway for fusel alcohol production: a century of research on *Saccharomyces cerevisiae* metabolism. *Appl Environ Microbiol* 74: 2259-2266.
48. Palazzotto E, Renzone G, Fontana P, Botta L, Scaloni A, et al. (2015) Tryptophan promotes morphological and physiological differentiation in *Streptomyces coelicolor*. *Appl Environ Microbiol* 99: 10177-10189.
49. Kavitha A, Prabhakar P, Vijayalakshmi M, Venkateswarlu Y (2010) Purification and biological evaluation of the metabolites produced by *Streptomyces* sp. TK-VL\_333. *Res Microbiol* 161: 335-345.
50. Szponarski W, Guibal O, Espuna M, Doumas P, Rossignol M, et al. (1999) Reconstitution of an electrogenic auxin transport activity mediated by *Arabidopsis thaliana* plasma membrane proteins. *FEBS letters* 446: 153-156.
51. Castro-Guerrero J, Romero A, Aguilar JJ, Xiqui ML, Sandoval JO, et al. (2011) The hisC1 gene, encoding aromatic amino acid aminotransferase-1 in *Azospirillum brasilense* Sp7, expressed in wheat. *Plant Soil* 356: 139-150.