University of South Bohemia in České Budějovice

Faculty of Science

Genome analysis of *Kutzneria* sp. strain BCCO 10_1627 and detection of antifungal secondary metabolites

Bachelor Thesis

Linda Jernej

Supervisor: Dr.nat.techn. Erika Corretto

Co-supervisor: RNDr. Alica Chroňáková, Ph.D.

Institute of Soil Biology, Biology Centre CAS

České Budějovice 2019

Jernej L., 2019: Genome analysis of *Kutzneria* sp. strain BCCO 10_1627 and detection of antifungal secondary metabolites. Bc. Thesis, in English 41 pages, Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation

Secondary metabolites, produced by various microbes, are used since decades for treatment of diseases. The bacterial genus *Kutzneria* is known for synthesizing compounds with antifungal activity, which might be used against pathogenic fungi, causing agricultural relevant diseases. This thesis deals with the genome analysis of the strain *Kutzneria* sp. BCCO 10_1627 for prediction of secondary metabolite gene cluster. Moreover, the antifungal activity of the strain was determined experimentally.

Declaration

I hereby declare that I have worked on my bachelor's thesis independently and used only the sources listed in the bibliography.

I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full form to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

České Budějovice, 15.8.2019 .....

Linda Jernej

Acknowledgements

I would like to express my sincere thanks to my thesis supervisor Dr.nat.techn. Erika Corretto for her guidance during all experiments and her answers to all my questions. She was an excellent teacher in all bioinformatic topics, and she thought me to work with bacteria and fungi.

Moreover, I want to thank RNDr. Alica Chroňáková, Ph.D. and the whole team of the Institute of Soil Biology for creating a pleasant work environment and for their help in any situation. Especially, I want to thank Engy Ahmed, Ph.D. for providing all fungi strains and keeping them alive.

List of abbreviations

amglyccycl	Aminoglycoside/aminocyclitol
bp	base pairs
EtOAc	ethyl acetate
GC %	content of guanine and cytosine in percentage
kb	kilo base pairs
MetOH	methanol
N50	median statistic such that 50% of the entire assembly is contained
NGS	next-generation sequencing
Nrps	Non-ribosomal peptide synthetase
ORF	open reading frame
PCR	polymerase chain reaction
PKS	Polyketide synthase
PPi	inorganic pyrophosphate

Abstract

Gram-positive and aerobic *Kutzneria* species belong to the family *Pseudonocardiaceae*. Through their secondary metabolism, they produce compounds with antifungal and antibacterial activity, which are useful not only in medicine, but also for the treatment of plant diseases.

In this thesis, genome sequencing and analysis of *Kutzneria* sp. BCCO 10_1627 was performed. As a result, we gained insight into the phylogenetic relations of the strain and its potential secondary metabolite production. The genome of *Kutzneria* sp. BCCO 10_1627 has a size of 11,664,045 bp divided in 125 contigs, which are coding for 10,338 genes. 4.5 % of genes are coding for secondary metabolites biosynthesis, transport and catabolism. 57 gene secondary metabolite gene clusters were predicted with antiSMASH. Phylogenetic analysis of the strain revealed the close relation of the *Kutzneria* genus to *Actinokutzneria* and *Lechevaleria* rather than to *Streptomyces.* The antifungal activity of BCCO 10_1627 against different plant pathogenic fungi was investigated via two antifungal screenings: using living cultures and ethyl acetate or methanol metabolite extracts of cultures grown in different media. Living culture of *Kutzneria* sp. BCCO 10_1627 showed antifungal activity against all 8 tested pathogenic fungi. In the second screening, antifungal activity was found only in ethyl acetate and methanol extracts, but not in the filtered supernatant.

Concluding, the tested strain *Kutzneria* sp. BCCO 10_1627 has a significant biosynthetic potential, producing natural antimicrobial compounds. The tested strain might be used in the future as provider for antifungal and antibiotic compounds in medicine and agriculture.

Table of contents

1 Introduction

1.1 Next generation sequencing platforms

Sequencing of genomes provides valuable information about the genetic potential of organisms. Thus, the improvement of sequencing methods is ongoing since decades. Sanger sequencing was the first method developed at the end of the 70' (Sanger et al., 1977). The automated Sanger sequencing method involves the use of a DNA polymerases, nucleotides and labelled dideoxynucleotides, a sequencing primer. The purified DNA template is sequenced via cycles of annealing, labeling, and termination steps. The sequenced DNA fragments are visualized using, capillary electrophoresis and fluorescence detection. This cost-extensive and time-consuming sequencing method is considered as a first-generation sequencing method (Metzker, 2009). Nowadays, next-generation sequencing (NGS) methods replaced the Sanger sequencing technique, which was used for the previous three decades (Lal and Seshasayee, 2014). With NGS methods whole genomes can be sequenced faster and simpler. Therefore, sequencing and investigation of microbial genomes is of growing interest, providing increasing information about the Bacterial domain (Behjati and Tarpey, 2013). Besides the sequencing of genomes, NGS is used for disease gene identifications, ChIP-seq, metagenomics and gene expression analysis (Buermans and den Dunnen, 2014).,

Various platforms for genome sequencing were developed over the years, with different working principles. The main advantages of NGS are that the DNA fragments no longer have to be cloned into a vector and multiple samples can be sequenced at the same time, thus making the process significantly faster. The greatest limitation of many NGS platforms is given by the cutting of the DNA into fragments, which leads to NGS read lengths shorter than with Sanger sequencing (Lal and Seshasayee, 2014).

Second-generation sequencing methods include different approaches to detect the incorporation of the correct nucleotide during the sequencing process. 454, IonTorrent and Illumina sequencing belong to this generation. In comparison to the first-generation, 454 uses no labelled radio- or fluorescence-marked nucleotides, instead luminescence occurs from pyrophosphate PPi release when a nucleotide is incorporated. Illumina sequencing still uses labelled nucleotides and IonTorrent is detecting the different nucleotides by a change in voltage, caused by a change in pH value. Second-generation sequencing is still a sequence by synthesis technique and requires the action of DNA polymerase. PCR needs to be done as a preliminary

step to multiply DNA fragments. While the first-generation methods can read up to 1000 kb, second-generation methods produce reads up to maximal 700 bp. PacBio sequencing machines belong to the third generation of sequencing methods. This platform is able to sequence single DNA molecules in real time. Thus, no DNA amplification via PCR is needed. Reads up to 10 kb pairs are produced using this technique (Heather and Chain, 2016). A fourth generation of genome sequencing called Nanopore was recently developed. While not only PCR is not needed, also no DNA polymerase is required. A single DNA molecule can be sequenced with reads up to 200 kb (Mignardi and Mats, 2014).

Among these NGS methods, one of the most frequently used is the Illumina platform. DNA is fragmented and short adapters are ligated at the ends of the DNA fragments, so that they can attach to the sequencing flow-cell. Each fragment is amplified by bridge PCR to get spatially separated clusters of identical fragments. Fluorescent labelled nucleotides are then added to the sequencing reaction. Every nucleotide carries a different fluorophore, that is detected by a camera (Lal and Seshasayee, 2014). Among the different Illumina sequencing machines, MiSeq is often used for sequencing small size genomes, like the ones of Bacteria. One of the most significant benefit of MiSeq is the enhanced read length form up to 2 x 150-300 paired end reads (Liu et al., 2012). Sequencing results in a large volume of sequencing data, that requires adequate bioinformatics tools to gain the desired information (Lal and Seshasayee, 2014).

1.2 Genome assembly and annotation

Although a bacterial genome is rather small, compared to a eukaryotic one, next generation sequencing results in a large output of data. Before starting the assembly, the quality of the sequencing data has to be determined. As a tool for quality assessment FastQC can be used. The program shows the quality scores of all reads and provides other statistics like the sequence length distribution and sequence GC % content (Wignett and Andrews, 2018). Bases and reads with low quality scores ($\leq 20 - 28$) are usually trimmed or excluded using for instance the program Trimmomatic (Ekblom and Wolf, 2014) (Bolger et al., 2014).

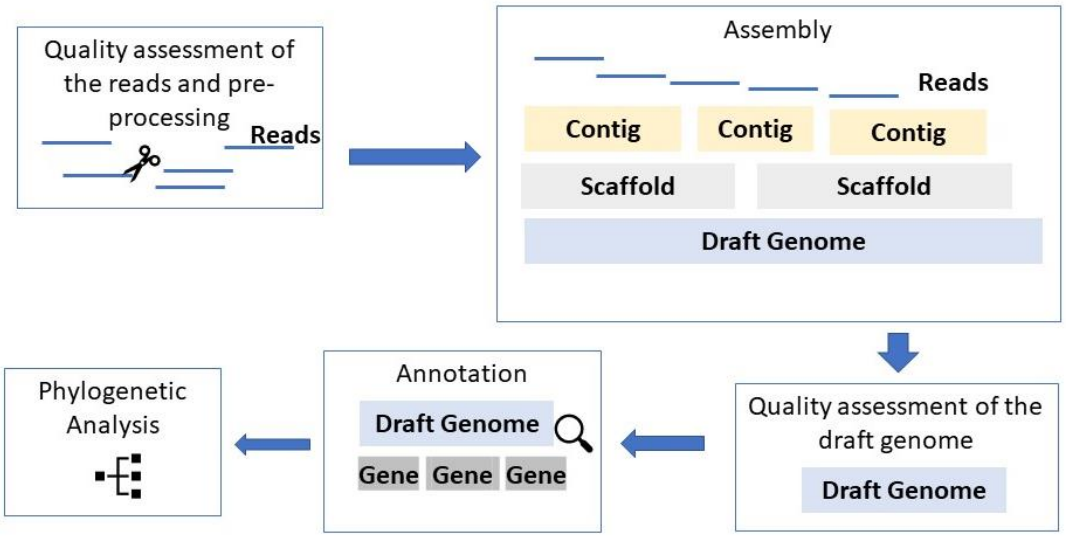


Figure 1: Workflow of *de novo* genome assembly and annotation, followed by taxon assignment via phylogenetic analysis.

After read pre-processing, the assembly can be started. If no genome of a closely related species is available as reference, a *de novo* assembly approach has to be applied. In case of Illumina sequencing data, short reads are obtained. First, they are assembled into contigs and eventually into a complete genome. A common tool for *de novo* assembly of bacterial genomes is SPAdes (Bankevich et al., 2012). It uses an algorithm based on a multisized de Bruijn graph. The de Bruijn method uses k-mers (fragments of DNA with *k* bases) to assemble the whole genome. The k-mers are split into *k-1* fragments, overlaps between these fragments are obtained. It is essential to choose a proper size of k-mer, as a too large *k-mers* may fail in detecting overlaps and too small will make the graph too twisted to analyse (Bankevich et al., 2012). When the overlapping step is completed, contigs are formed. Contigs are long sequences, containing no gaps, which can be assembled into even larger structures called scaffolds. Joining scaffolds will result in a complete genome (Del Angel et al., 2018).

Afterwards, the quality of the genome should be assessed with evaluation tools like QUAST (Gurevich et al., 2013). The quality of a genome assembly can be determined validating different parameters. The total number of contig and the total number of bases (total length) indicate whether the sequencing was successful. Another important parameter is the N50 value, which specifies the length of the largest contig accounting for at least 50 % of all contigs. A high N50 value indicates a good assembly, whereas a low one is characteristic of low-quality genomes fragmented in several contigs. The GC % content of a genome gives the percentage of guanine and cytosine bases divided by all bases and is specific for certain taxa of

related microorganisms. Moreover, the number of genes can be used to validate the completeness: roughly 1000 genes for 1 Mb of genome. Furthermore, finding all rRNA and tRNA essential for cell functions and protein building, as well as important single copy marker genes in the genome are indicators of its completeness and can show the presence of eventual contaminations (Gurevich et al., 2013). Figure 1 illustrates the genome assembly and annotation process.

Before genes can be assigned to a function, open reading frames (ORFs) need to be identified in the genome sequence. Prokaryotic genomes are small and do not include introns in their genome; hence gene identification is simpler. Nevertheless, some ORFs are very small and can overlap. This might lead to misidentified genes. Genes ORFs can be found either searching for structural characteristics (for example, start and stop codons, specific nucleotide patterns) or by comparing the identified sequences to already known genes. Several annotation pipelines combine a structural and functional strategy to predict and annotate genes (Ekblom and Wolf, 2014). A tool for bacterial and archaeal genome annotation is RAST (Rapid Annotation using Subsystem Technology), which uses subsystems to easily browse genes and their annotations. Additionally, genes can be compared to others within the SEED environment (Aziz et al., 2008). Predicting orthologues is possible with the eggNOG tool (Huerta-Cepas et al., 2016). Orthologues are groups of genes having similar functions. Comparing orthologous sequences gives insight about the functional potential of a bacteria. Moreover, prediction of the evolution of a gene can be made.

1.3 Actinobacteria

Actinomycetes, a genus group of bacteria affiliated to the phylum of *Actinobacteria*, means “ray fungi” in old Greek (Sowani et al., 2017). Due to their live cycle and morphology some Actinomycetes such as *Streptomyces* were mistaken for fungi, although being bacteria (Doroghazi and Metcalf, 2013). They are gram-positive, filamentous and occur in soil, aquatic environments, composts and food. More than 150 genera of Actinomycetes exist. They are morphologically similar, but phylogenetically different (Sowani et al., 2017). The GC % content of Actinomycetes is usually over 55%. *Actinobacteria* are producers of antibiotics and antifungal compounds (for example fungicidin) but also the deadliest pathogens known, such as *Myobacterium tuberculosis*. Besides their antimicrobial activity, Actinomycetes’ secondary metabolites are used as plant growth promoters, herbicidal and antineoplastic (Doroghazi and Metcalf, 2013). The many secondary metabolites also allow the bacteria survive in ecological

conditions, that are normally not occupied by bacteria, in both terrestrial and aquatic ecosystems, including animals' gastrointestinal tracts (Prakash et al., 2013).

1.4 Bacterial secondary metabolites

Secondary metabolites are organic compounds, produced in nature and acting for instance as antibiotics or antifungals. The organisms produce these chemicals as defence against other organisms, as agents of symbiosis, as metal transporters, as signalling effector (Croteau et al., 2000). The production of secondary metabolites is stimulated in natural conditions by a change in environmental factors. When cultivating the organisms in laboratories, secondary metabolism usually starts when the stationary growth phase is reached (Demain and Fang, 2000). Especially microbial secondary metabolites are often used as medicines as they have among other functions anti-tumour, cholesterol-lowering or antibiotic activity (Medema et al., 2011).

Fungal infections are often the cause of fruit rot and plant death, such as the Southern corn leaf blight or the grey mould. Of available drug types active against mycotic diseases only a small portion are antifungal antibiotics (Bharti et al., 2010). *Actinobacteria*, especially *Streptomyces* and related strains, are considered beneficial microbes in agriculture, as they boost the plant growth and protect plants from fungal diseases. These helper bacteria are convenient when used in plant-microbe symbiosis to achieve more sustainable agricultural practises. Research on new biofertilizer and biopesticides is still ongoing (Vurukonda et al., 2018).

With the progress in sequencing methods and annotation, microbial genomes are more accessible. Thus, secondary metabolites with antifungal activity produced by bacteria can be found also with the help informatics tools (Rebets et al. 2014). The characterization of secondary metabolites is challenging, due to their biochemical heterogeneity. AntiSMASH (antibiotics & Secondary Metabolite Analysis Shell) predicts secondary metabolites gene clusters and their function, using not only a database with known genes, but it can also identify new clusters using the ClusterFinder algorithm (Blin et al., 2017).

1.5 The genus *Kutzneria*

Kutzneria species are aerobic and gram-positive bacteria belonging to the family *Pseudonocardiaceae.* Their sporangia are globose and large with a diameter of 10 to 48 µm. *Kutzneria* have spherical, non-motile spores, which are rod-shaped or ovoid. Their 50 µm long sporangiophores are formed by the septation of unbranched hyphae. Cell walls of *Kutzneria* contain N-acetylated muramic acid, galactose as sugar and meso-diaminopimelic acid. Naturally, *Kutzneria* occur mostly in soils (Labeda and Whitman, 2015).

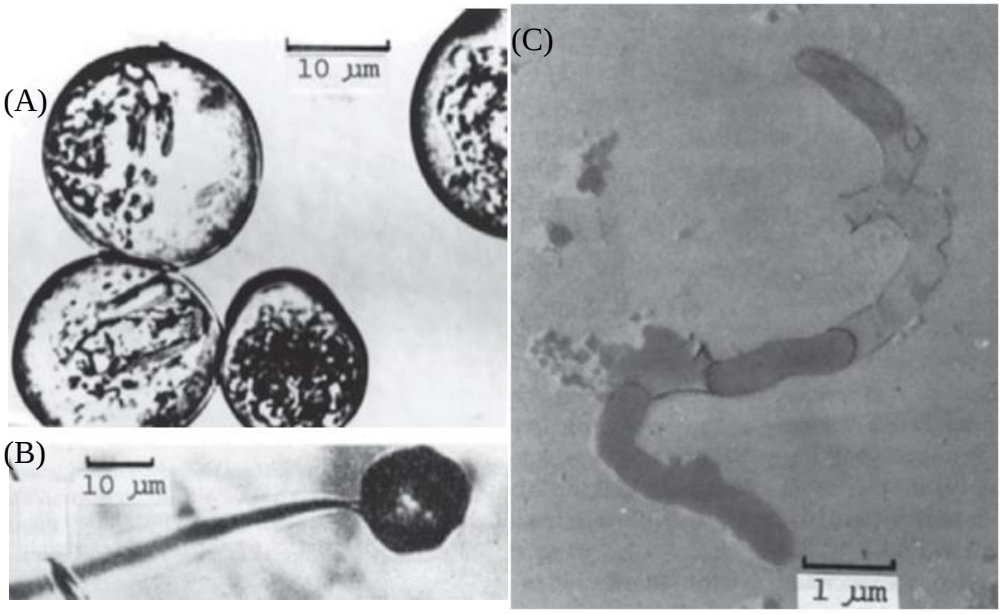
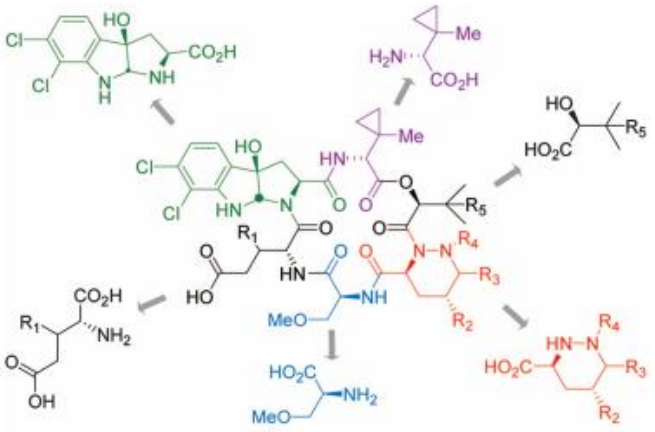


Figure 2: (A) sporangia, (B) sporangiophores and (C) spores *of Kutzneria kofuensis* (Labeda, 2015).

Until now only one genome of *Kutzneria* was completely sequenced. *Kutzneria albida,* isolated from Japanese soil, produces aculeximycin, a compound active against bacteria, fungi and mosquito larvae. The strain‘s genome is circular, with no extrachromosomal replication. The genome of *K. albida* revealed the most secondary metabolite gene clusters among all actinobacterial genomes. Besides the cluster for aculeximycin, 45 secondary metabolite gene clusters were found, making *K. albida* a rich source of natural antimicrobial compounds (Rebets et al., 2014).

Kuztneria are known to produce secondary metabolites called kutznerides, which were found to be antifungal and antimicrobial compounds. The cyclic hexadepsipeptides were first investigated in *Kutzneria* sp. 774 (Fujimori et al., 2007). Nine structurally different, but related forms of kutznerides were isolated (Figure 3). Cyclic depsipeptides commonly occur as secondary metabolites in various kinds of organisms, such as fungi, bacteria and sponges, and

have shown to be anticancer, antiviral, antibacterial, anti-inflammatory and antifungal (Fujimori et al., 2007). Previous studies by Broberg et al. have shown that especially kutznerides 1-4 have inhibited the growth of several root rotting fungi (Broberg et al., 2006).



	R_1	R_2	R_3	R_4	R_5
Kutzneride 1	(S)-OH	H	H	H	Me
Kutzneride 2	(S)-OH	Cl	H	H	Me
Kutzneride 3	(R)-OH	H	H	H	Me
Kutzneride 4	(R)-OH	H	π-bond		Me
Kutzneride 5	(R)-OH	H	H	H	H
Kutzneride 6	(R)-OH	OH	π-bond		Me
Kutzneride 7	(S)-OH	H	H	H	H
Kutzneride 8	(R)-OH	Cl	H	H	Me
Kutzneride 9	(S)-OH	H	π-bond		Me

Figure 3: All nine different structures of kutznerides found in Kutzneria sp. 774 (Fujimori et al., 2007).

1.5.1 *Kutzneria* sp. BCCO 10_1627

The strain used in this work is affiliated to the genus *Kutzneria*. It is deposited in the Culture Collection of Soil Actinomycetes České Budějovice of the Institute of Soil Biology, Biology Centre CAS, v. v. i (www.actinomycetes.cz). The strain showed strong antifungal activity against various strains. Moreover, it inhibits *Saccharomyces cerevisiae, Bacillus subtilis* and *Staphylococcus aureus.* Resistance was shown to ampicillin, penicillin, cephalosporin, gentamycin, tetracycline (10) and vancomycin. On the other hand, BCCO 10_1627 is sensitive to ciprofloxacin, erythromycin, rifampicin, streptomycin and chloramphenicol. The strain was transient to amoxicillin and tetracycline (30).

A preliminary PCR screening using specific primers showed that *Kutzneria* sp. BCCO 10_1627 has a short chain length factor gene, which might be involved in the synthesis of polyene compounds (data not published). These molecules have a wide range of bioactivities: antifungal, antibacterial, cytotoxic and anti-inflammatory (Odds et al., 2003)

Due to these interesting characteristics the strain was chosen for further investigations: (i) whole genome sequencing, (ii) test its antifungal activity in more detail.

2 Aims of the study

The aim of this thesis was to test the antifungal activity of the strain *Kutzneria* sp. BCCO 10_1627 using experimental, as well as bioinformatic tools. The specific goals of this work were:

in silico analyses

1) genome assembly and annotation

2) phylogenetic analysis using the 16S rRNA gene as molecular marker

3) identification of potential secondary metabolite gene clusters using bioinformatic tools

experimental work

4) investigation of its antifungal activity against different plant pathogenic fungi

a) using living cultures

b) using metabolite extracts from culture grown in different media

3 Materials and Methods

3.1 Isolation and Storage of *Kutzneria* sp. BCCO 10_1627

The strain *Kutzneria* sp. BCCO 10_1627 was isolated in 2014 from forest soil of Mount Cameroon (Cameroon) at an average altitude of 350 m. The isolation medium was YEME and the growth medium M2 agar. It is part of the Culture Collection of Soil Actinomycetes České Budějovice (CCSACB, www.actinomycetes.cz) of the Institute of Soil Biology, Biology Centre CAS. The strain was persevered and stored at -80°C as well as at -150 °C.

3.2 DNA Extraction, Genome Assembly and Annotation

DNA was extracted following the instructions of the Wizard Genomic DNA purification kit by Promega. An additional second centrifugation step was performed before resuspending the DNA. The sequencing of the *Kutzneria* sp. BCCO 10_1627 genome was performed by the Laboratory of Environmental Microbiology at the Institute of Microbiology, CAS, Prague. The library was prepared using the TruSeq PCR free LT library preparation kit (Illumina) and quantified with the KAPA library quantification kit (Roche). The library was sequenced using the Illumina MiSeq platform (Reagent kit v2, paired-end, 300 bp).

The obtained reads were analysed, assembled and annotated resulting in a draft genome sequence. FastQC (Wignett and Andrews, 2018) was used to check the read quality. The number of total reads, read length and GC % content were determined. Moreover, the quality score of each nucleotide was obtained. In order to delete traces of the enterobacteria phage phiX174 (CP004084.1), Bowtie2 (Langmead and Salzberg, 2012) was used. Reads were aligned against the genome of the phage; matching sequences were discarded. Using Trimmomatic (Bolger et al., 2014) adapters and low quality or N-bases were removed from the sequences. The reads were scanned with a 4-base-wide sliding window. Bases with an average quality score below 15 were removed. Merging paired-end reads was done with FLASH (Magoc and Salzberg, 2011).

Finally, genome assembly was performed with SPAdes (Safonova et al., 2014) using default settings. The k-mer length was increased in increments of 22 until the k-mer length reached 127. Using QUAST (Mikheenko et al., 2018), the complete genome sequence was evaluated. Number of contigs, total length, N50 value and the GC % content of the genome were obtained. Qualimap was used to determine the genome coverage (Okonechnikov et al.,

2015). Using Aragon (Laslett and Canback, 2004) tRNAs were identified. The presence of each codon was verified to determine the genome's completeness.

The genome was annotated using the online tools RAST (Aziz et al., 2017) and antiSMASH (Blin et al., 2017). The complete sequence was uploaded to the platforms and analysed using implemented browsers. RAST identified protein-encoding, tRNA and rRNA genes and determined the corresponding functions. antiSMASH was used to predict secondary metabolic gene clusters of the genome. Clusters of orthologous genes were identified using eggNOG (Huerta-Cepas et al., 2016). Moreover, the presence of 31 different marker genes was determined using the Amphora webserver (Kerepesi et al., 2014).

3.3 Phylogenetic Analysis

For the phylogenetic analysis, the first 25 related organisms and 5 *Streptomycetes* with 16S rRNA gene matching the one of BCCO 10_1627 were selected using the NCBI Nucleotide BLAST tool (Altschul et al., 1990). In addition, the strains used in the antifungal tests were considered for this analysis with the exception of strain BCCO 10_1008 (Table 1). The identification of this strain is still ongoing. The 16S rRNA gene of BCCO 10_1627 was selected from the GenBank file obtained from the RAST annotation. It was found in contig 66 and has a length of 1576 bp. *E. coli* O157:H7 str. Sakai (BA000007.3) was used as an outgroup.

Table 1: List organisms used for the phylogenetic analysis, including NCBI accession number (Acc. Number), identity percentage and matching base pairs to the total length of BCCO 10_1627 16S rRNA sequence (1576 bp).

Organism	Acc. Number	Identity (%)	Matching base pairs
Blast hits for 16S rRNA			
Kutzneria chonburiensis strain SMC 256	NR_145619.1	99%	1468/1481
Kutzneria buriramensis strain A-T 1846	NR_109430.1	99%	1445/1448
Kutzneria albida strain DSM 43870	NR_122053.1	98%	1482/1506
Kutzneria kofuensis strain NRRL B-24061	NR_041733.1	98%	1482/1507
Kutzneria albida strain DSM 43870	NR_121729.1	98%	1477/1506
Kutzneria albida strain NRRL B-24060	NR_044273.1	98%	1475/1503

Saccharothrix lopnurensis strain YIM LPA2h	NR_145947.2	97%	1471/1510
Saccharothrix yanglingensis strain Hhs.015	NR_117283.1	97%	1467/1512
Saccharothrix hoggarensis strain SA181	NR_109103.1	97%	1464/1508
Actinosynnema mirum strain DSM 43827	NR_074438.1	97%	1463/1509
Saccharothrix ecbatanensis strain UTMC 00537	NR_109447.1	97%	1466/1515
Saccharothrix coeruleofusca strain NRRL B-16115	NR_041736.1	97%	1465/1515
Actinosynnema mirum strain IMSNU 20048	NR_041822.1	97%	1460/1510
Lentzea waywayandensis strain NRRL B-16159	NR_114495.1	97%	1459/1509
Saccharothrix syringae strain NRRL B-16468	NR_041739.1	97%	1461/1513
Lechevalieria flava strain NRRL B-16131	NR_041737.1	97%	1460/1512
Umezawaea endophytica strain YIM 2047X	NR_145951.1	97%	1454/1505
Allokutzneria multivorans strain YIM 120521	NR_109545.1	96%	1461/1515
Allokutzneria albata strain R 761-7	NR_025570.1	97%	1448/1495
Actinosynnema pretiosum strain NBRC 15621	NR_112681.1	97%	1434/1476
Saccharothrix saharensis strain SA152	NR_108320.1	96%	1456/1510
Lechevalieria aerocolonigenes strain NRRL B-3298	NR_028738.1	96%	1452/1508
Saccharothrix espanaensis strain LL-C19004-NS29	NR_102474.1	96%	1453/1508
Actinosynnema pretiosum subsp. auranticum strain NBRC 15620	NR_041600.1	97%	1429/1474
Actinokineospora globicatena strain NRRL B-24048	NR_024963.1	96%	1451/1511

Blast hits for 16S rRNA of *Streptomycetes*			
Streptomyces gardneri strain NBRC 3385	NR_041204.1	94%	1379/1470
Streptomyces bryophytorum strain NEAU-HZ10	NR_146707.1	93%	1410/1517
Streptomyces thermodiastaticus strain JCM 4840	NR_112048.1	93%	1414/1527
Kitasatospora aureofaciens strain KACC 20180	NR_042792.1	93%	1413/1527
Streptomyces calidiresistens strain YIM 78087	NR_134195.1	93%	1404/1516

Strains used in the antifungal assay			
Streptomyces sp. BCCO 10_219	not deposited	91%	1257/1373
Streptomyces avidinii BCCO 10_322	KP718507.1	91%	1263/1385
Streptomyces sp. BCCO 10_521	not deposited	91%	1258/1379
Kribbellas sp. BCCO 10_952	not deposited	90%	1287/1423
Streptomyces sp. BCCO 10_996	not deposited	91%	1160/1274
Streptomyces sp. BCCO 10_1189	not deposited	91%	1264/1379
Streptomyces sp. BCCO 10_1500	not deposited	91%	1268/1387
Streptomyces sp. BCCO 10_1537	not deposited	92%	1285/1388
Streptomyces sp. BCCO 10_1597	not deposited	92%	1269/1373
Streptomyces capoamus BCCO 10_1636	KP718517.1	91%	1268/1388
Streptomyces hundungensis BCCO 10_1666	KP718561.1	91%	1254/1375
Streptomyces sp. TR42	not deposited	92%	1412/1532
Streptomyces lydicus WYEC 108	FJ222807.1	90%	1362/1506

Sequences were aligned with Muscle (Edgar, 2004). Two trees were constructed in MEGAX (Kumar et al., 2018): one using the Maximum Likelihood algorithm based on the Tamura-Nei model and with bootstrap 1000 (Figure 6A); the second tree was constructed using Neighbour Joining with bootstrap 1000 (Figure 6B).

3.4 Growth Media

Potato Dextrose Agar (PDA, HiMedia) plates were used to cultivate the fungi. Bacterial strains were cultivated in M2, GYM or malt extract medium. For the preparation of agar plates, 20 g L^{-1} agar were added to the recipes described in Table 2.

Table 2: Growth media composition.

Ingredient	Amount
M2	
Malt Extract	10 g
Yeast	4 g
GYM	
Malt Extract	10 g
Yeast	4 g
N-Z-Amine	1 g
Sodium Chloride	2 g
OB Salts	3 mL
Malt extract	
Malt Extract	20 g
Peptone	4 g

The pH was adjusted to 7.2-7.3 with 5 M NaOH for every medium. All media were filled up to 1 L with distilled water. When needed, to GYM and M2 medium also 4 g L^{-1} of glucose were added after sterilization.

3.5 Antifungal Assay

The antifungal activity of the *Kutzneria* sp. BCCO 10_1627 was determined with two different screenings. In the first one, the interaction between the actinomycete and several fungi was assessed, using living microorganisms. In the second screening the antifungal efficacy of BCCO 10_1627 was assessed using filtrated supernatants and metabolite extracts. In addition, 14 different bacterial strains were used for comparison (Table 3): one *Kribella* strain and 13 belonging to the genus *Streptomyces*. The 10 different fungal strains used for the assay are listed in Table 4. All actinomycetes and fungal strains are part of the culture collections

(www.micromycetes.cz, www.actinomycetes.cz) of the Institute of Soil Biology, Biology Centre CAS.

Table 3: List of strains used in the antifungal assay, including their isolation source.

Identification	Strain	Isolation Source
Kutzneria sp.	BCCO 10_1627	Soil, Mt. Cameroon
Streptomyces lydicus	WYEC 108	Roots, GB
Streptomyces sp.	BCCO 10_219	Soil, Czech Republic
Streptomyces avidinii	BCCO 10_322	Soil, Czech Republic
Streptomyces sp.	BCCO 10_521	Soil, USA
Kribbella sp.	BCCO 10_952	Soil, USA
Streptomyces sp.	BCCO 10_996	Soil, USA
Streptomyces sp.	BCCO 10_1008	Bentonite clay, Czech Republic
Streptomyces sp.	BCCO 10_1189	Soil, Germany
Streptomyces sp.	BCCO 10_1500	Soil, Czech Republic
Streptomyces sp.	BCCO 10_1537	Soil, Slovenia
Streptomyces sp.	BCCO 10_1597	Soil, Mt. Cameroon
Streptomyces capoamus	BCCO 10_1636	Soil, Mt. Cameroon
Streptomyces hundungensis	BCCO 10_1666	Soil, Mt. Cameroon
Streptomyces sp.	TR42	Sputum, Czech Republic

Table 4: List of plant pathogenic fungi used in the antifungal assay.

Identification	Strain	Plant disease
Fusarium oxysporum	BCCO 20_2866	Numerous diseases for food crops
Bipolaris sorokiniana	BCCO 20_1571	Southern corn leaf blight of maize
Penicillium spinulosum	BCCO 20_389	Rotting leaves
Trichoderma harzianum	BCCO 20_606	Dieback, root rot, post-harvest fruit rot
Fusarium oxysporum	BCCO 20_605	Numerous diseases for food crops
Alternaria alternata	BCCO 20_609	Leaf spot, rots
Botrytis cinerea	BCCO 20_912	Grey mould, stem rot, fruit rot
Geotrichum candidum	BCCO 20_1313	Fruit rot (tomato, peach, citrus)
Chaetomium globosum	BCCO 20_2527	Food spoilage in storage
Monographella cucumerina	BCCO 20_2872	Vascular wilt disease

3.5.1 First Screening

For the first screening living organisms were used. The bacterial strains were grown on M2 agar and the fungi on PDA. After 7 days, circular cut-outs of the fungi and bacteria were prepared and placed onto a malt extract agar plate. The fungi were placed in the middle of the plate with three strains of bacteria in 3 cm distance. Figure 4 depicts the used plate layout.

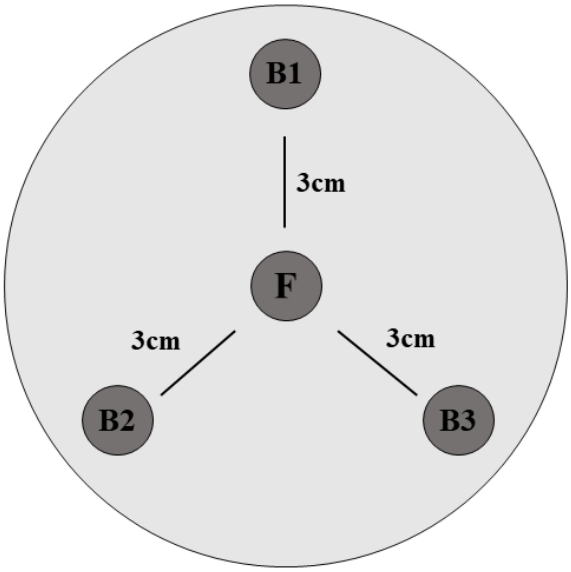


Figure 4: Layout of plates of the first antifungal screening with the fungus (F) in the middle surrounded by different strains of bacteria (B1, B2 and B3).

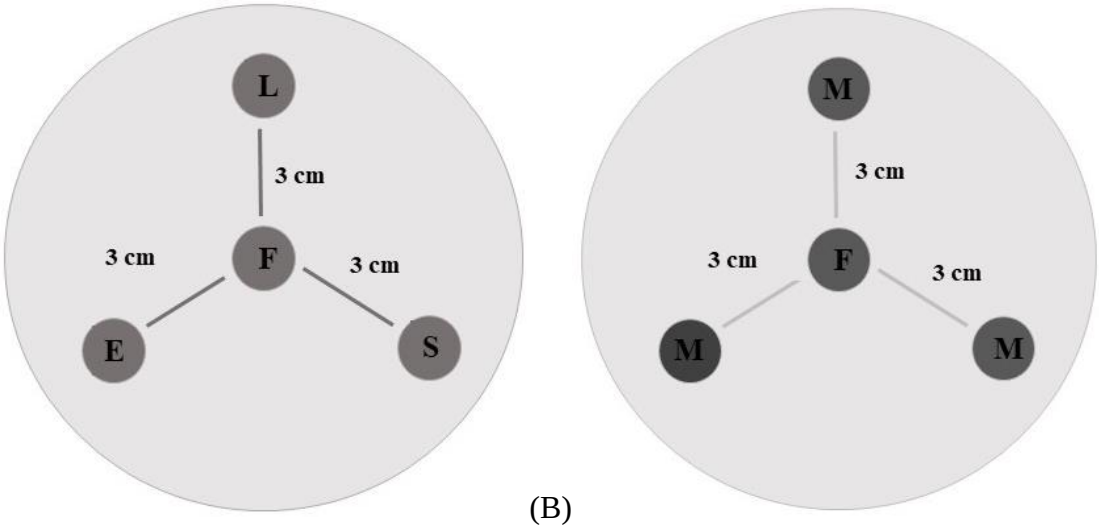
For each combination of bacteria and fungus triplicates were prepared. Plates were incubated at 28 °C for either 4 or 7 days, depending on their growth rate. Afterwards, the antifungal activity was determined by measuring the inhibition zone around the bacteria.

3.5.2 Second Screening

The second screening was performed only with the bacterial strains *Kutzneria* sp. BCCO 10_1627 and *Streptomyces lydicus* WYEC 108. The following fungi were selected for the second screening: *Fusarium oxysporum* BCCO 20_2866, *Fusarium oxysporum* BCCO 20_605, *Bipolaris sorokiniana* BCCO 20_1571, *Trichoderma harzianum* BCCO 20_606, *Alternaria alternate* BCCO 20_609, *Geotrichum candidum* BCCO 20_1313, *Chaetomium globosum* BCCO 20_2527 and *Monographella cucumerina* BCCO 20_2872.

The fungi were grown on PDA plates for 7 days. After a pre-culture step, the bacteria were grown for 72 hours in GYM and malt extract liquid medium as well as in GYM + 0.4% glucose and malt extract + 0.4% glucose. Growth cultures were prepared in duplicates. After the incubation time, the duplicates were pooled. Subsequently, the pooled cultures were divided in four different fractions: liquid culture (L), filtrated supernatant (S), ethyl acetate (E) and

methanol (M) extract. The layout of the plates for the second screening is displayed in Figure 5.



(A) (B)

Figure 5: (A) Layout of the plates of the second screening. The fungus (F) is surrounded by the liquid culture (L), the supernatant (S) and the EtOAc extract (E). Triplicates of each combination were made. (B) Plate layout for the methanol extract (M).

25 µL liquid culture (L) was added directly onto the agar plate. The supernatant (S) was obtained by centrifuging 2 mL culture for 5 min at 10,000 rpm at 4 °C. Cells were discarded and the supernatant was filtered with 0.3 µL syringe filters, then 25 µL were added onto filter discs, which were placed on the agar plate.

The ethyl acetate (EtOAc) extract (E) were prepared by centrifuging the cultures for 10 min at 5000 rpm at 4 °C, biomass and supernatant were extracted separately. The supernatant was filtered into a fresh tube, then the solution was saturated with NaCl to approximately 5 M. One third of the volume EtOAc was added. Subsequently, the mixture was shaken for 30 min at 300 rpm in the dark. The organic phase was collected, and the water phase was extracted once more. The biomass was mixed with 6.7 mL of acetone and shaken for 30 min in the dark. The precipitate was spun down (15 min, 5000 rpm, 4 °C) and the liquid phase was collected and evaporated. 4 mL of EtOAc were added to the dry material and the mixture was shaken again for 30 min in the dark. The organic phase was collected, pooled with the organic phase from the supernatant extraction and evaporated. The remaining crude material was dissolved in 1 mL of chloroform. 25 µL of the chloroform mixture were added on a filter disc, which was placed on the agar plate.

Furthermore, the fourth fraction of the culture was extracted with methanol (MetOH). 6.7 mL of MetOH were added to the freeze-dried culture and shaken (300 rpm) in the dark for 30 min. The precipitate was spun down (15 min, 5000 rpm, 4 °C) and the liquid was collected. Subsequently, the liquid phase was evaporated. The remaining crude material was dissolved in 1 mL of methanol and 25 µL of the mixture were added on a filter disc, which was placed on an agar plate. The used plate layout is illustrated in Figure 5 (B).

4 Results and Discussion

4.1 Phylogenetic Analysis

The phylogenetic trees of Figure 6 were constructed according to section 3.3. Both phylogenetic tree constructing methods gave consistent results, indicating that BCCO 10_1627 is indeed part of the *Kutzneria* genus. The closest described organism is *Kutzneria buriramensis* A-T 1846. Together with *Allokutzneria*, *Kutzneria* species form a distinct group closer to actinomycetes like *Actinosynnema*, *Lentzea*, *Lechevaliera* and *Saccharotrix*.

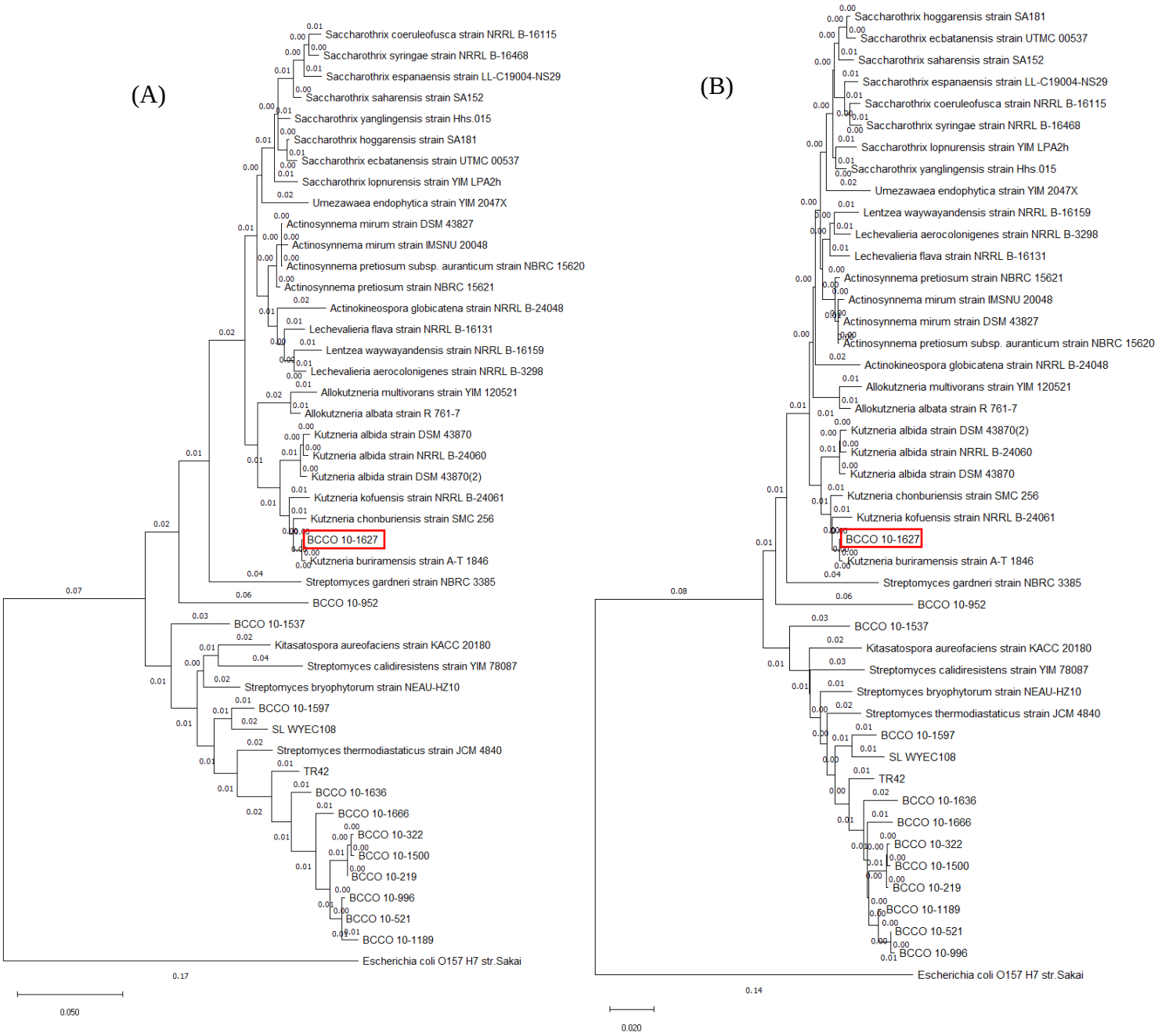


Figure 6: Phylogenetic trees. (A) Maximum likelihood and (B) Neighbor joining trees calculated using bootstrap 1000 in MEGA X.

4.2 Genome Statistics

The genome analysis was done according to section 3.2. The complete genome of *Kutzneria* sp. BCCO 10_1627 has a total length of 11,664,045 bp divided in 125 contigs. 71 contigs were longer than 1000 bp, while 84 were longer than 500 bp. The GC % content of the strain is 70.4 %. The N50 value was found to be 373,347. The genome contains 10,338 coding sequences and 82 tRNA, 1 tmRNA and 6 rRNA genes. Table 5 gives a summary of the main genome statistics values.

Table 5: Genome statistics summary for *Kutzneria* sp. BCCO 10_1627.

Statistics	Values
Total length	11,664,045 bp
Total number of contigs	125
Contigs > 1000 bp	71
Contigs > 500 bp	84
N50 value	373,347
GC %	70.4 %
Coverage	31.44X ± 11.21
CDS	10,338
rRNA	6
tRNA	82
tmRNA	1

The only complete genome of the *Kutzneria* genus found in the literature is from *Kutzneria albida* (Rebets et al., 2014). The genome of *Kutzneria albida* DSM 43870^{T} is 9,874,926 bp long, thus shorter than the genome of BCCO 10_1627. Consequently, less coding genes (8,822) were identified for *K. albida* than for BCCO 10_1627. The GC % content of both genomes is similar with 70.6 % for *K. albida* and 70.4 % for BCCO 10_1627.

RAST (Aziz et al., 2017) was used to annotate the genome of BCCO 10_1627. A total of 10,338 genes were identified and assigned to different subsystem categories as shown in Figure 7.

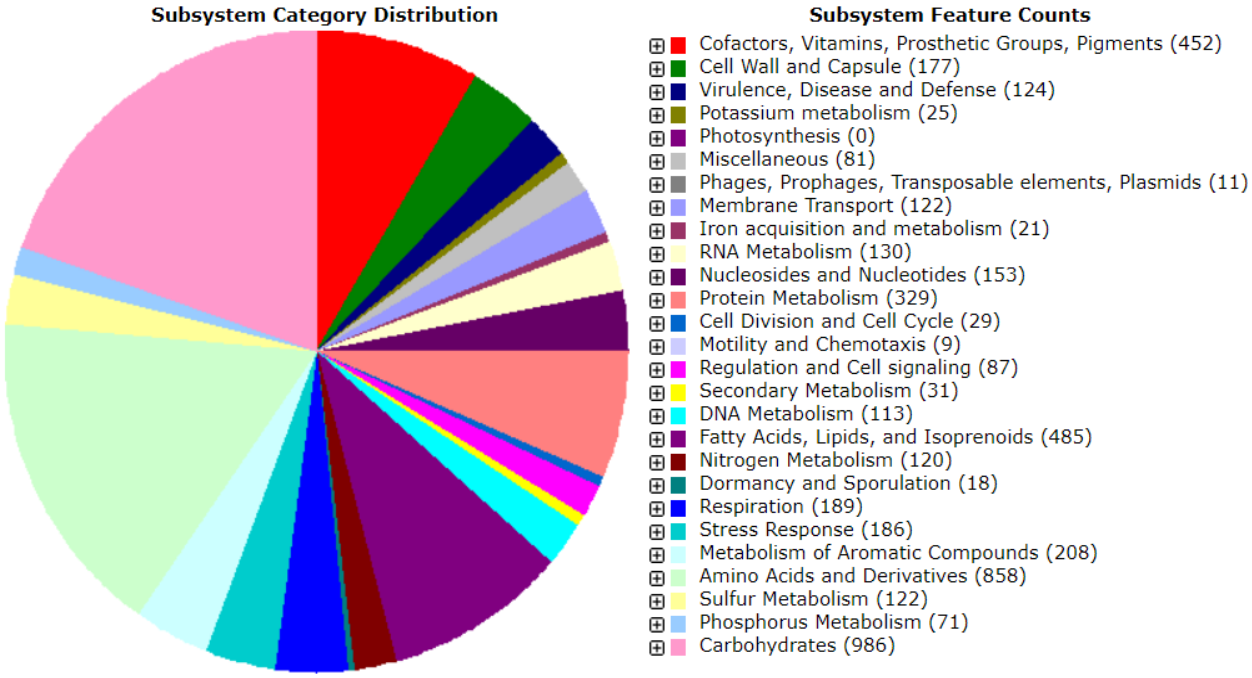


Figure 7: Gene functions as provided by RAST for *Kutzneria* sp. BCCO 10_1627.

Clusters of orthologous genes were identified with eggNOG (Huerta-Cepas et al., 2016). 22 different cluster functions were assigned and are summarized in Table 6. The completeness of the genome was again proven, as all crucial functions were found. Overall, BCCO 10_1627 and *Kutzneria albida* share similar functions. For instance, approximately 4.5 % of BCCO 10_1627 gene functions are dedicated to secondary metabolite biosynthesis, transport and catabolism compared to 4.4% in *Kutzneria albida*.

Table 6: Clusters of orthologous gene families assigned using eggNOG (Huerta-Cepas et al., 2016) for the genome of *Kutzneria* sp. BCCO 10_1627, including the function, total number of orthologous groups and the corresponding percentage (assigned genes/total number of genes).

Function	Total number of orthologous groups	Percentage %
RNA processing & modification	1	0.01
chromatin structure & dynamics	1	0.01
energy production & conversion	460	4.95
cell cycle control, cell division, chromosome portioning	50	0.54
amino acid transport & metabolism	523	5.62
nucleotide transport & metabolism	124	1.33
carbohydrate transport & metabolism	685	7.36
coenzyme transport & metabolism	189	2.03
lipid transport & metabolism	326	3.50
translation, ribosomal structure & biogenesis	278	2.99
transcription	1120	12.04
replication, recombination, repair	219	2.35
cell wall/membrane/envelope biosynthesis	283	3.04
cell motility	2	0.02
post translational modification, protein turnover & chaperons	216	2.32
inorganic ion transport & metabolism	326	3.50
secondary metabolites biosynthesis, transport & catabolism	422	4.54
unknown function	2324	24.99
signal transduction mechanism	366	3.94
intracellular trafficking, secretion & vesicular transport	33	0.35
defence mechanism	181	1.95
cytoskeleton	1	0.01
no function assigned	1171	12.59

4.3 Prediction of secondary metabolite gene clusters

The identification of the secondary metabolite gene clusters was done using antiSMASH (Blin et al., 2017). Table 7 lists all predicted secondary metabolite gene clusters, with at least 50 % of genes showing similarity to known clusters. In addition, other 57 secondary metabolite gene clusters were assigned by antiSMASH, summarized in the appendix (Table S1).

Table 7: Secondary metabolites gene clusters predicted in BCCO 10_1627 genome having at least 50 % of genes showing similarity to known clusters.

Cluster No.	Type	Percentage of genes showing similarity	Most similar known cluster (MIBiG)
70	Bacteriocin-Nrps	100	Kutznerides (BGC0000378)
73	Indole-Thiopeptide-T2pks-Lantipeptide-T1pks	71	Erdasporin (BGC0001336)
86	Cf_fatty_acid-Lantipeptide	100	Ectoine (BGC0000853)
87	Ectoine	100	Ectoine (BGC0000853)
96	Oligosaccharide-T1pks	58	Aculeximycin
100	Terpene	100	Geosmin (BGC0000661)
101	T3pks-Nrps	92	Ristomycin
106	Cf_fatty_acid-T1pks	50	Vicenistatin
111	Nrps	80	Scabichelin (BGC0000423)
113	Indole	92	BE-54017 (BGC0001333)
114	T2pks	92	BE-54017 (BGC0001333)

For cluster 70, the predicted chemical compound is kutznerides, a known antimicrobial against *Drechslera sorokiniana*, a pathogenic fungus. The compound, a cyclic hexadepsipeptides, is described in detail in the Introduction (Rebets et al., 2014). Cluster 73 might encode for genes involved in the synthesis of erdasporine, an antibiotic compound. In cluster 96, 58 % of the genes show similarity with the one for the production of aculeximycin, an antibiotic known from *Streptosporangium albidum* (Ikemoto et al., 1983). Geosmin (cluster 100) is a non-antimicrobial secondary metabolite, responsible for earthy tastes and the special smell of the bacteria (Gerber and Lechevalier, 1965). Scabichelin, (cluster 11) is a well-known antibiotic, found in *Streptomyces halstedii* (Kodani et al., 2013). Some of the compounds produced by these gene clusters might be involved in the previously observed antibacterial activity or in the antifungal activity observed in this study. Further investigation is needed to link these clusters to the corresponding compound and bioactivity.

4.4 Antifungal Assay

The antifungal assays were performed according to section 3.5. In order to compare the antifungal efficiency of BCCO 10_1627, 14 actinomycetes strains were tested to investigate their antifungal effect against 10 pathogenic fungi strains. Among the selected actinomycetes, *S. lydicus* (Yuan and Crawford, 1995) was used as reference, since it is a well-recognized antifungal agent and is used as biocontrol agent, commercially available. All experiments were performed on agar plates. The antifungal activity was determined by measuring the formed inhibition zone around the bacteria.

4.4.1 First Screening

The first screening was done using the living bacterial strains. Table 9 gives the measured inhibition zones in mm. Fungal strains 389 and 912 did not grow during the experiments, hence they were excluded from this study.

Table 9: Mean values of measured inhibition zones (in mm); tests were performed in triplicates for all bacterial and fungal strain combinations. "-":no inhibition zone; "+"distinct inhibition zone, that was not measurable because of a strange colony shape; "w": very small inhibition zones that were not measurable.

		Fungi Strain							
		2866	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
Bacteria Strain	**1627**	5.50 (±1.5)	9.67 (±0.47)	+	6.00 (±0.82)	10.67 (±1.70)	10.00 (±0.00)	7.67 (±0.47)	11.67 (±0.94)
	SL	6.00 (±1.00)	8.67 (±0.47)	4.00 (± 0.00)	5.33 (±0.47)	8.00 (±0.82)	4.33 (±0.47)	3.67 (±0.47)	7.33 (±0.47)
	219	-	-	-	-	3.00 (±0.00)	-	-	-
	322	-	w	-	w	w	w	-	w
	521	-	3.67 (±0.47)	-	3.00 (±0.00)	3.00 (±0.00)	3.00 (±0.00)	-	3.00 (±0.82)
	952	w	+	-	4.00 (±0.00)	w	+	-	2.00 (±0.00)
	996	6.00 (±0.00)	7.67 (±1.25)	2.00 (±0.00)	3.67 (±1.25)	w	2.00 (±0.00)	-	5.00 (±1.41)
	1008	w	4.33 (±0.47)	-	w	2.00 (±0.00)	1.00 (±0.00)	-	1.00 (±0.00)
	1189	-	w	-	w	w	2.00 (±0.00)	-	w
	1500	-	-	-	-	+	w	-	-
	1537	-	-	--	-	w	-	-	w
	1597	7.33 (±0.94)	9.67 (±0.47)	+	5.00 (±0.00)	6.67 (±1.89)	7.00 (±0.00)	8.33 (±0.94)	6.67 (±1.70)
	1636	-	w	-	-	-	-	-	-
	1666	-	w	-	w	0.50 (±0.50)	2.00 (±2.00)	-	-
	TR42	6.33 (±1.25)	8.33 (±1.25)	+	5.67 (±0.47)	9.00 (±1.41)	7.00 (±0.00)	5.33 (±1.25)	8.00 (±0.82)

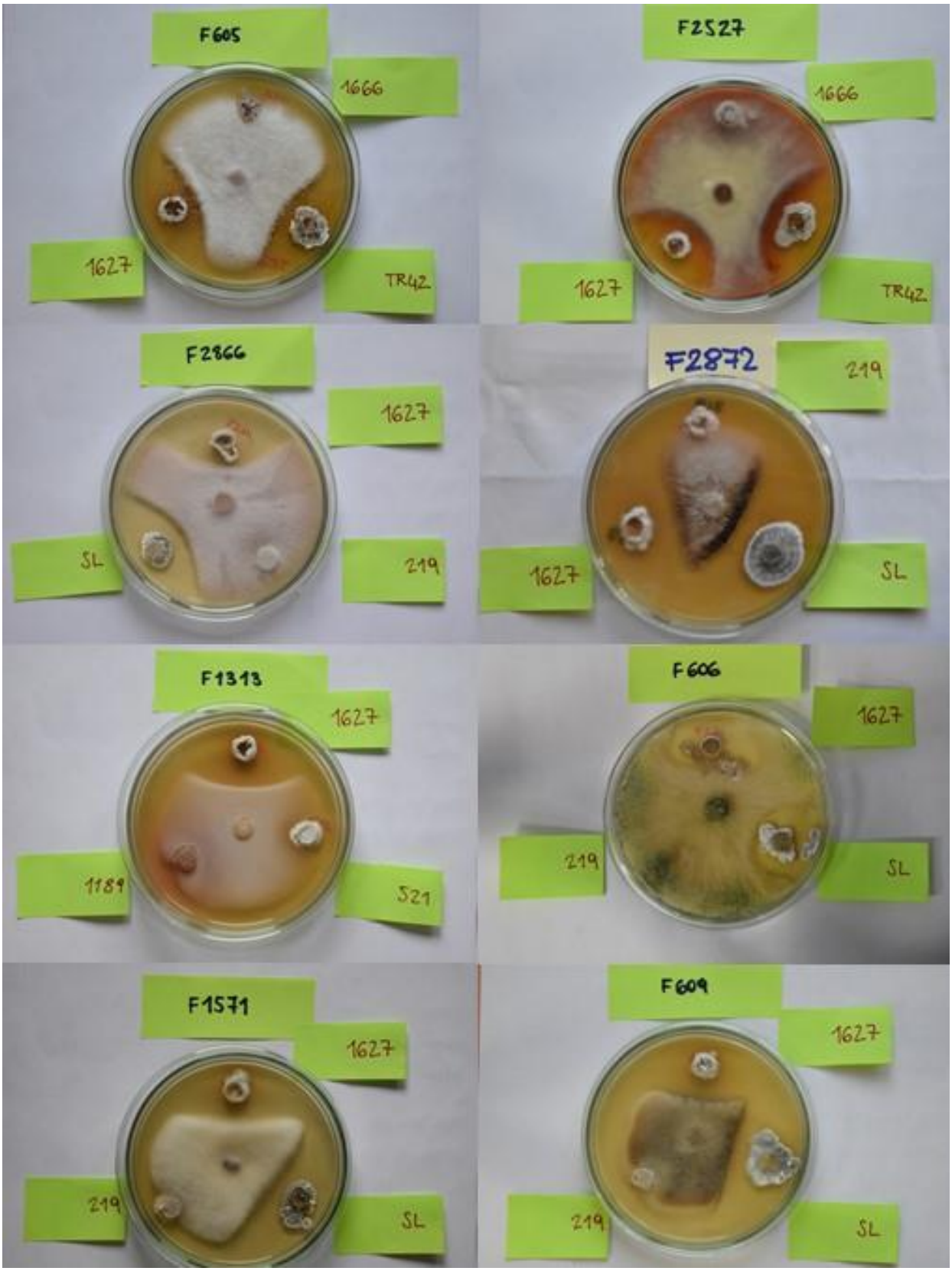


Figure 8: Antifungal activity of BCCO 10_1627 against the selected plant pathogenic fungi.

Figure 8 shows all fungi, which were inhibited by BCCO 10_1627. Compared to other bacterial strains, *Kutzneria* sp. BCCO 10_1627 showed antifungal activity against all fungal strains. Against the fungal strain BCCO 20_606 only weak activity could be observed. In most combinations, the inhibition zones around BCCO 10_1627 were larger than the ones of the reference strain *Streptomyces lydicus* WYEC 108 (SL). Strains *Streptomyces* sp. BCCO 10_1597 and *Streptomyces* sp. BCCO 10_996 showed antifungal activity comparable to BCCO 10_1627 and *Streptomyces lydicus* against most fungal strains. For instance, *Chaetomium globosum* BCCO 20_2527 was inhibited only by the strains SL, BCCO 10_1627, BCCO 10_1597 and TR42, which together with BCCO 10_996 inhibited the growth of *Fusarium oxysporum* BCCO 20_2866 and *Trichoderma harzianum* BCCO 20_606 as well.

These results indicate that BCCO 10_1627 has a strong antifungal activity against a wide range of plant pathogenic fungi. This activity is probably caused by the production of one or more secondary metabolites.

4.4.2 Second screening

In the second screening the antifungal potential of metabolites extracted from BCCO 10_1627 and SL cultures was assessed.

As described in the materials and methods section, the growth cultures were divided into four fractions to assess their antifungal activity: Liquid cultures (L), filtered supernatant (S), methanol extracts (M) and ethyl acetate extracts (E) were cultivated and screened against the fungi strains. To evaluate the effect of different growth media on the production of secondary metabolites, the strains were grown in malt extract, GYM, malt extract +0.4 % glucose and GYM +0.4 % glucose media.

The screening with the strain BCCO 10_1627 showed that the filtered supernatant as well as the EtOAc extracts did not inhibit the growth of the different fungi. For the four different growth media, inhibition zones could only be determined with the liquid cultures. No consistent difference in media could be determined with these experiments. The strain was growing equally well in malt and GYM medium. Moreover, there was no increased growth when glucose was added. Table 10 summarizes the results of the second screening with BCCO 10_1627.

Table 10: Results of the second screening with BCCO 10_1627. Triplicates of each plate were prepared, the mean values of measured inhibition zones (in mm) for the strain BCCO_1627 are shown. "-":no inhibition zone; "+":distinct inhibition zone, that was not measurable because of a strange colony shape.

		Fungi Strain							
	Liquid	**2866**	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
Medium	**Malt extract**	2.33 (±0.47)	5.33 (±1.25)	-	2.00 (±0.86)	3.67 (±0.47)	5.33 (±0.47)	3.33 (±0.47)	8.33 (±0.94)
	GYM	5.50 (±1.50)	5.00 (±0.00)	-	3.00 (±0.52)	5.50 (±0.50)	4.33 (±0.47)	+	6.00 (±1.63)
	Malt extract+ 0.4 %glucose	2.00 (±0.00)	5.33 (±1.25)	-	1.67 (±0.47)	1.67 (±0.47)	3.67 (±0.47)	5.00 (±0.816)	5.00 (±1.41)
	GYM + 0.4 %glucose	+	6.00 (±0.82)	-	2.50 (±0.50)	4.00 (±0.82)	3.67 (±0.94)	2.67 (±0.47)	5.00 (±0.00)
	Supernatant	**2866**	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
	Malt extract	-	-	-	-	-	-	-	-
	GYM	-	-	-	-	-	-	-	-
	Malt extract+ 0.4 %glucose	-	-	-	-	-	-	-	-
	GYM + 0.4 %glucose	-	-	-	-	-	-	-	-
	EtOAc	**2866**	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
	Malt extract	-	-	-	-	-	3.00 (±1.00)	-	-
	GYM	-	-	-	-	-	-	-	-
	Malt extract+ 0.4 %glucose	-	-	-	-	-	1.00 (±2.00)	-	-
	GYM + 0.4 %glucose	-	-	-	-	-	1.00 (±2.00)	-	-
	MetOH	**2866**	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
	Malt extract	1.50 (±0.50)	-	-	4.67 (±2.50)	8.67 (±2.05)	1.50 (±0-82)	1.67 (±0.82)	-
	GYM	-	-	-	-	-	-	-	-
	Malt extract+ 0.4 %glucose	3.50 (±1.67)	12.00 (±1.25)	5.00 (±1.42)	3.33 (±0.94)	6.33 (±0.47)	6.67 (±0-47)	12.33 (±0.47)	-
	GYM + 0.4 %glucose	-	-	-	-	-	-	-	-

MetOH extracts from malt extract cultures showed increased inhibition zones compared to other media. The huge zones around BCCO 20_1571 and BCCO 20_2527 can be explained with additional slow growth of the fungi. The greatest inhibition zones were measured for the fungi strains BCCO 20_1571 and BCCO 20_2872 with the liquid culture. These findings correlate with the first antifungal screening. The supernatant and EtOAc extract did not inhibit the fungal growth. The only exception was found to be with the fungal strain BCCO 20_1313. Rather small and hardly distinguishable inhibition zones were measured for the EtOAc extraction. Figure 9 depicts the inhibition zones of fungi caused by BCCO 10_1627. Fungal strain BCCO 20_606 was not completely inhibited by BCCO 10_1627, as observed in the first screening. Figure 10 shows inhibition using MetOH extracts.

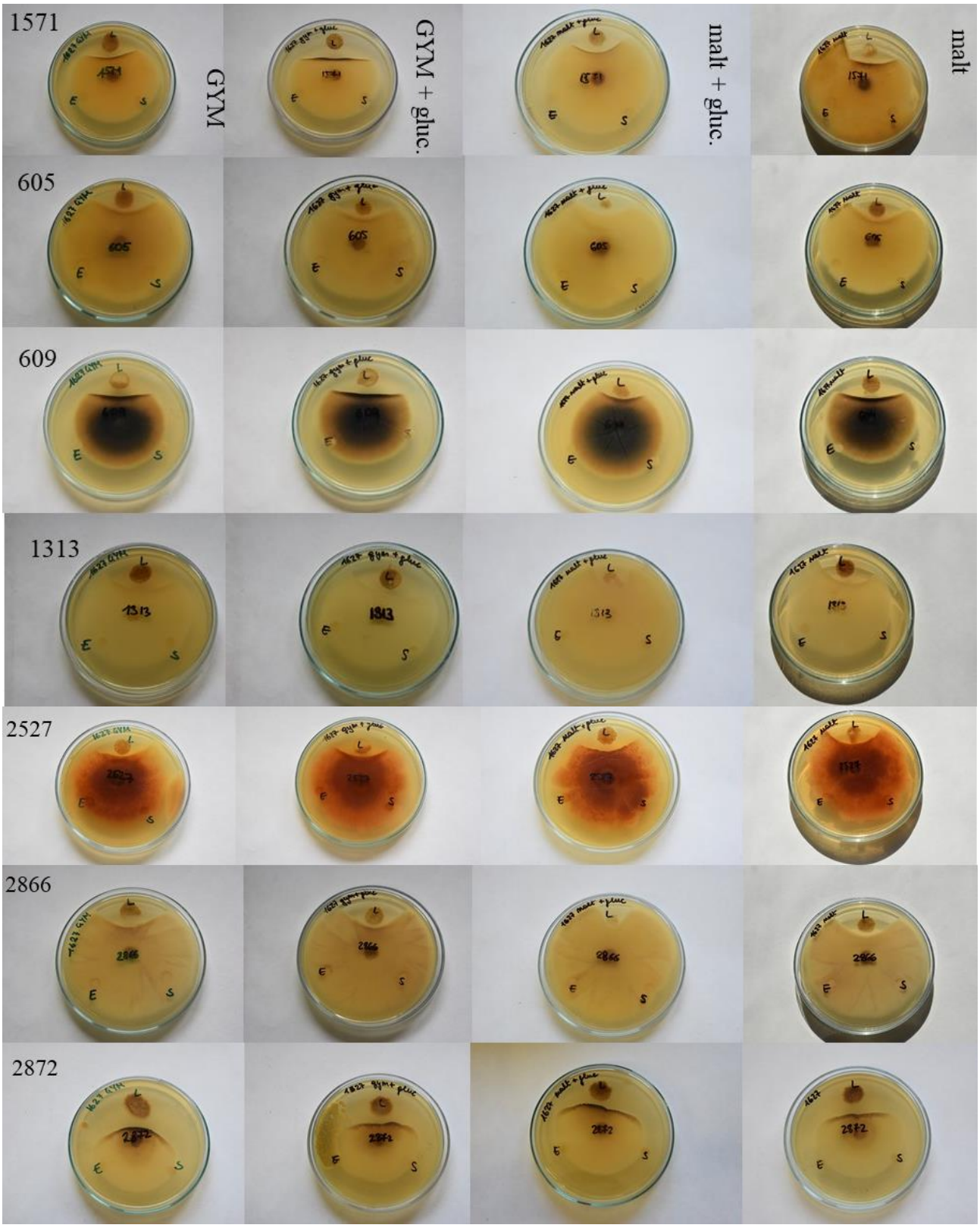


Figure 9: Plate layout of the second screening for BCCO 10_1627 and all selected fungi displaying the liquid fraction (L), supernatant (S) and the EtOAc extracts (E) with the four different used media.

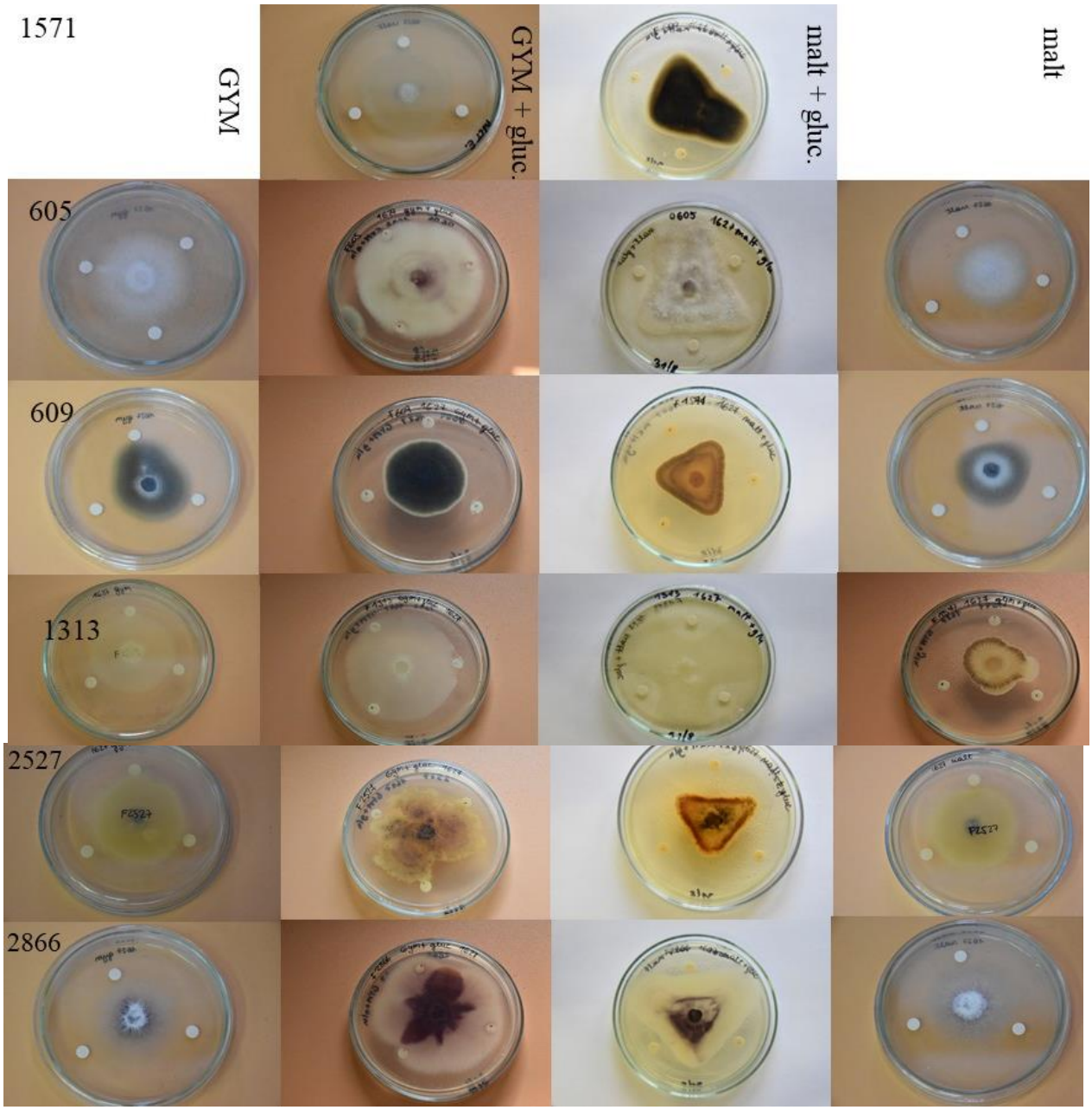


Figure 10: Plate layout of the second screening of BCCO 10_1627 and all selected fungi displaying the methanol extracts and the four used media.

The results for the strain SL are similar to the ones for BCCO 10_1627 (Table 11). There is no observable difference in media preferences of the strain. The liquid cultivation did inhibit the fungi, while neither the filtered supernatant, nor EtOAc extraction did, with the exception for the fungal strain BCCO 20_1313. In contrast to the results of BCCO 10_1627, the methanol extracts did not inhibit the fungal strains.

Table 11: Triplicates of each plate were prepared, the mean values of measured inhibition zones (in mm) for the strain SL and used fungal strains. - means that there was no inhibition zone, + stands for a distinct inhibition zone, that was not measurable because of a strange colony shape.

		Fungi Strain							
	Liquid	**2866**	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
Medium	**Malt extract**	3.00 (±0.00)	8.00 (±1.41)	3.67 (±0.47)	4.50 (±0.50)	6.33 (±1.25)	3.33 (±0.47)	2.50 (±0.50)	6.33 (±0.47)
	GYM	3.33 (±0.47)	8.67 (±0.47)	+	3.33 (±0.47)	6.00 (±.047)	5.33 (±0.82)	5.67 (±0.94)	6.67 (±1.25)
	Malt extract+ 0.4 %glucose	3.33 (±0.94)	8.33 (±0.47)	+	2.67 (±0.47)	6.33 (±0.82)	4.00 (±0.47)	4.00 (±0.00)	4.67 (±0.47)
	GYM + 0.4 %glucose	2.00 (±0.00)	7.67 (±0.94)	+	4.67 (±0.47)	5.67 (±.047)	5.00 (±0.00)	5.00 (±0.82)	5.33 (±0.47)
	Supernatant	**2866**	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
	Malt extract	-	-	-	-	-	-	-	-
	GYM	-	-	-	-	-	-	-	-
	Malt extract+ 0.4 %glucose	-	-	-	-	-	-	-	-
	GYM + 0.4 %glucose	-	-	-	-	-	-	-	-
	EtOAc	**2866**	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
	Malt extract	-	-	-	-	-	2.00 (±0.00)	-	-
	GYM	-	-	-	-	-	-	-	-
	Malt extract+ 0.4 %glucose	-	-	-	-	-	1.67 (±0.47	-	-
	GYM + 0.4 %glucose	-	-	-	-	-	-	-	-
	MetOH	**2866**	**1571**	**606**	**605**	**609**	**1313**	**2527**	**2872**
	Malt extract	-	-	-	-	-	-	-	-
	GYM	-	-	-	-	-	-	-	-
	Malt extract+ 0.4 %glucose	-	-	-	-	-	-	-	-
	GYM + 0.4 %glucose	-	-	-	-	-	-	-	-

In summary, the second antifungal screening of SL and BCCO 10_1627 showed that only MetOH extracts of BCCO 10_1627 were effective. On the other hand, EtOAc extracts did only inhibit the fungal strain BCCO 20_1313. The secondary metabolites produced by BCCO 10_1627 appear to be better soluble in MetOH, than in EtOAc, hence they might be polar. The liquid cultures did inhibit the fungi, although with smaller zones, than for the solid cultures. Possibly less cells in the liquid culture than in the solid were available.

The second antifungal screening needs to be repeated in order to gain significant results. The EtOAc extractions should be improved in order to obtain higher concentrations of secondary metabolites. The MetOH extractions of malt extract cultures must be performed in a bigger scale to ensure its antifungal activity.

5 Conclusion

The genome sequencing of *Kutzneria* sp. BCCO 10_1627 resulted in a draft genome having a total length of 11,664,045 base pairs. The GC % of the strain lies at 70.4 %. The total number of genes is 10,338. Moreover, according to eggNOG 4.5 % of genes are dedicated to secondary metabolism. antiSMASH predicted 11 clusters with more than 50 % of genes showing similarity to known genes. Other 57 secondary metabolic gene clusters were predicted. Among the found clusters is one for kutznerides biosynthesis, a strong antifungal compound produced by *Kutzneria* species. At genome level, similarities with *Kutzneria albida* were observed.

The phylogenetic analysis of the strain showed the close relation of the genus *Kutzneria* to *Actinokutzneria, Lechevaleria, Actionosynnema* and *Saccharothrix* rather than to *Streptomyces.* The phylogenetic trees calculated using the Neighbour Joining and Maximum Likelihood methods are in agreement.

Kutzneria sp. BCCO 10_1627 showed to have outstanding antifungal activity in the first antifungal screening. In comparison to the other tested bacterial strains, BCCO 10_1627 inhibited all tested fungi. Likewise, the strains *Streptomyces lydicus* WYEC 108 and *Streptomyces* sp. BCCO 10_TR42 produced distinct inhibition zones around all selected fungi. Moreover, *Streptomyces* sp. BCCO 10_1597 and *Streptomyces* sp. BCCO 10_996 inhibited the growth of most fungi. Further experiments on antifungal activity need to be done with these strains.

The second antifungal screening showed that the liquid cultures of BCCO 10_1627 had weaker antifungal activity than the solid cultures. Whereas, the supernatant fraction did not show any inhibition zones. Most probably the liquid fraction was too diluted, hence the antifungal compounds could not show any activity. The extracts of EtOAc did not give significant results, while the MetOH extraction confirmed antifungal activity. The secondary metabolites with antifungal activity obviously are better soluble in MetOH than in EtOAc. The second screening might be improved in further experiments using bigger extraction volumes. Generally, the first screening and the liquid fraction did show greater antifungal activity, hence the living cells tend to produce more secondary metabolites when a fungus is present.

To be able to prospectively use the secondary metabolites of *Kutzneria* sp. BCCO 10_1627 as antifungal agent, the compounds need to be extracted and analysed in further experiments. The MetOH extraction can be fractionated by HPLC and analysed further to identify the antifungal agents. Nevertheless, the antifungal activity of strain BCCO 10_1627 was proven through bioinformatic tools as well as experimental work.

6 References

Altschul S.F.., Gish W., Miller W., Myers E.W., Lipman D.J., Basic local alignment search tool. *J Mol Biol., 215(3),* pages 403-410, (1990)

Aziz R. K., Bartels D., Best A. A., DeJongh M., Disz T., Edwards R. A., Zagnitko O., The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics, 9(1),* pages 75, (2008), doi: 10.1186/1471-2164-9-75

Bankevich A., Nurk S., Antipov D., Gurevich A. A., Dvorkin M., Kulikov A. S., Pevzner P. A. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology, 19(5),* pages 455–477, (2012), doi: 10.1089/cmb.2012.002

Behjati S., Tarpey, P. S., What is next generation sequencing? *Archives of Disease in Childhood - Education & Practice Edition, 98(6),* pages 236–238, (2013), doi: 10.1136/archdischild-2013-304340

Bharti A., Kumar V., Gusain O., Bisht G. S., Antifungal Activity of Actinomycetes isolated from Garhwal Region. *Journal of Sci. Eng. & Tech. Mgt., 2 (2),* (2010)

Blin K., Wolf T., Chevrette M. G., Lu X., Schwalen C. J., Kautsar S. A., Suarez Duran H. G., Los Santos E. LC., Uk Kim H., Nave M., Dickschat J. S., Mitchelle D. A., Shelest E., Breitling R., Takano E., Lee S. Y., Weber T., Medema M. Hm., antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification, *Nucleic Acids research, 45(1),* pages 36-41, doi: 10.1093/nar/gkx319.

Bolger A. M., Lohse M., Usadel B., Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics, btu170,* (2014)

Broberg A., Menkis A., Vasiliauskas R., Kutznerides 1−4, Depsipeptides from the Actinomycete Kutzneriasp. 744 Inhabiting Mycorrhizal Roots of Piceaabies Seedlings. *Journal of Natural Products, 69(1),* pages 97–102, (2006), doi: 10.1021/np050378g

Buermans H. P. J., den Dunnen J. T., Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 1842(10),* pages 1932–1941, (2014), doi:10.1016/j.bbadis.2014.06.015

Croteau R., Kutchan T. M., Lewis N. G., Natural Products (Secondary Metabolites). *Biochemistry & Molecular Biology of Plants, B. Buchanan, W. Gruissem, R. Jones, Eds Chapter 24*., pages 1250-1316, (2000)

Del Angel V., Hjerde E., Sterck L., Capella-Gutierrez S., Notredame C., Vinnere Pettersson O., Lantz H., Ten steps to get started in Genome Assembly and Annotation. *F1000Research, 7, 148,* (2018), doi: 10.12688/f1000research.13598.1

Demain A. L., Fang A., The Natural Functions of Secondary Metabolites. *Advances in Biochemical Engineering/Biotechnology,* pages 1–39. (2000). doi: 10.1007/3-540-44964-7_1

Doroghazi J. R., Metcalf, W. W., Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. *BMC Genomics, 14(1),* page 611, (2013), doi: 10.1186/1471-2164-14-611

Edgar R. C., MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 32(5),* pages 1792–1797, (2004), doi: 10.1093/nar/gkh340

Ekblom R., Wolf J. B. W., A field guide to whole-genome sequencing, assembly and annotation. *Evolutionary Applications, 7(9),* pages 1026–1042, (2014), doi :10.1111/eva.12178

Fujimori D. G., Hrvatin S., Neumann C. S., Strieker M., Marahiel M. A., & Walsh C. T., Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proceedings of the National Academy of Sciences, 104(42),* pages 16498–16503, (2007), doi: 10.1073/pnas.0708242104

Gerber N. N., Lechevalier H. A., Geosmin, an Earthy-Smelling Substance Isolated from Actinomycetes. *American Society for Microbiology Journals,* (1965)

Gurvevich A., Saveliev V., Vyahhi N., Tesler G., QUAST: quality assessment tool for genome assemblies. *Bioinformatics, 29(8),* pages 1072-1075, (2013), doi: 10.1093/bioinformatics/btt086

Heather J. M., Chain B., The sequence of sequencers: The history of sequencing DNA. *Genomics, 107(1),* pages 1–8, (2016), doi: 10.1016/j.ygeno.2015.11.003

Huerta-Cepas J., Szklarczyk D., Forslund K., Cook H., Heller D., Walter M. C., Rattei T., Mende D. R., Sunagawa S., Kuhn M., Jensen L. J., von Mering C., Brok P., eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucl. Acids Res., 44(1),* pages 286-293, (2016), doi: 10.1093/nar/gkv1248

Ikemoto T., Matsunaga H., Konishi K., Okazaki T., Enokita R., Torikata A., Aculeximycin, a new antibiotic from Streptosporangium albidum. I. taxonomy of producing organism and fermentation. *The Journal of Antibiotics, 36(9),* pages 1093–1096, (1983), doi: 10.7164/antibiotics.36.1093

Kerepesi C., Banky D., Drolmusz V., Methods AmphoraNet: The webserver implementation of the AMPHORA2 metagenomic workflow suite. *Gene, 533(2),* pages 538-540, (2014), doi: 10.1016/j.gene.2013.10.015

Kodani S., Bicz J., Song L., Deeth R. J., Ohnishi-Kameyama M., Yoshida M., Challis G. L. Structure and biosynthesis of scabichelin, a novel tris-hydroxamate siderophore produced by the plant pathogen Streptomyces scabies 87.22. *Organic & Biomolecular Chemistry, 11(28),* page 4686, (2013), doi:10.1039/c3ob40536b

Kumar S, Stecher G, Li M, Knyaz C, and Tamura K., MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution, 35*, pages 1547-1549, (2018)

Labeda D. P., Kutzneria. *Bergey's Manual of Systematics of Archaea and Bacteria*, pages 1–8, (2015), doi:10.1002/9781118960608.gbm00180

Labeda D., Whitman W. B., Statutes of the International Committee on Systematics of Prokaryotes. Subcommittees and the Judicial Commission on Prokaryote Nomenclature of the Bacteriology and Applied Microbiology Division of the IUMS International. *Journal of*

Systematic and Evolutionary Microbiology 65- pages 1093-1100, doi: 10.1099/ijs.0.000019, (2015)

Lal A., Seshasayee A. S. N., The Impact of Next-Generation Sequencing Technology on Bacterial Genomics. *A Systems Theoretic Approach to Systems and Synthetic Biology II: Analysis and Design of Cellular Systems,* pages 31–58, (2014), doi: 10.1007/978-94-017-9047-5_2

Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. *Nature Methods, (9),* pages 357-359, (2012)

Laslett D., Canback B., ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequence. *Nucleic Acids Res., 32(1),* pages 11-16, (2004) doi: 10.1093/nar/gkh152

Lin L., Li Y., Li S., Hu N., He Y., Pong R., Lin D., Lu L., Law M., Comparison of Next-Generation Sequencing Systems. *Journal of Biomedicine and Biotechnology, 2012*, 11 pages (2012), doi: 10.1155/2012/251364 2012

Magoc T., Salzberg S., FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics 27:21,* pages 2957-2963, (2011)

Medema M. H., Blin K., Cimermancic P., de Jager V., Zakrzewski P., Fischbach M. A., Breitling R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research, 39(suppl_2),* pages 339– 346, (2011), doi: 10.1093/nar/gkr466

Metzker M. L., Sequencing technologies – the next generation. *Nature Reviews genetics, 11(1),* pages 31-46, (2009), doi: 10.1038/nrg2626

Mignardi M., Nillson E., Fourth-generation sequencing in the cell and the clinic. *BioMed Central Ltd.*, (2014), doi: 10.1186/gm548

Mikheenko A., Prjibelski A., Saveliev V., Antipov D., Gurevich A., Versatile genome assembly evaluation with QUAST-LG, *Bioinformatics, 34 (13),* pages 142-150, doi: 10.1093/bioinformatics/bty266

Odds F. C., Brown A. J. P., Gow, N. A. RAntifungal agents: mechanisms of action. *Trends in Microbiology, 11(6),* pages 272–279, (2003), doi:10.1016/s0966-842x(03)00117-3

Okonechnikov K., Conesa A., García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics, btv566*. (2015). doi: 10.1093/bioinformatics/btv566

Prakash D., Nawani N., Prakash, M., Bodas M., Mandal A., Khetmalas M., Kapadnis B., Actinomycetes: A Repertory of Green Catalysts with a Potential Revenue Resource. *BioMed Research International, 2013*, pages 1–8, (2013), doi: 10.1155/2013/264020

Rebets Y., Tokovenko B., Lushchyk I., Rückert C., Zaburannyi N., Bechthold A., Luzhetskyy A., Complete genome sequence of producer of the glycopeptide antibiotic Aculeximycin Kutzneria albida DSM 43870T, a representative of minor genus of Pseudonocardiaceae. *BMC Genomics, 15(1),* page 885, doi: 10.1186/1471-2164-15-885

Safonova Y., Bankevich A., Pevzner P.A., dipSPAdes: Assembler for Highly Polymorphic Diploid Genomes. *Research in Computational Molecular Biology. RECOMB, 8394*, (2014)

Sanger F., Nicklen S., Coulson A. R., DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, 74(12)*, pages 5463–5467, (1977), doi: 10.1073/pnas.74.12.5463

Sowani H., Kulkarni M., Zinjarde S., Javdekar V., Gordonia and Related Genera as Opportunistic Human Pathogens Causing Infections of Skin, Soft Tissues, and Bones. *The Microbiology of Skin, Soft Tissue, Bone and Joint Infections*, pages 105–121, (2017), doi: 10.1016/b978-0-12-811079-9.00007-0

Vurukonda S. S. K. P., Giovanardi D., Stefani E., Plant Growth Promoting and Biocontrol Activity of Streptomyces spp. as Endophytes. *International Journal of Molecular Sciences, 19(4)*, page 952, (2018), doi: 10.3390/ijms19040952

Wingett SW, Andrews S., FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res. (7)*, page 1338, (2018), doi: 10.12688/f1000research.15931.2.

Yan W. M., Crawford D. L., Characterization of streptomyces lydicus WYEC108 as a potential biocontrol agent against fungal root and seed rots. *American Society for Microbiology Journals*, (1995)

7 Appendix

Table S1: 57 secondary metabolite gene clusters assigned by antiSMASH

Type	From	To	Most similar known cluster		Similarity
terpene	68,221	88,285			
indole	196,526	217,641	Fortimicin	saccharide	9%
ladderane	518,872	560,101	Ishigamide	nrps-t2pks	22%
terpene	639,625	665,539	Hopene	terpene	46%
NRPS-like	735,431	777,941	Svaricin	nrps-t1pks	6%
bacteriocin	1,063,515	1,074,315			
NRPS-like, butyrolactone	1,113,636	1,156,587	Salinipostin	other	55%
butyrolactone	1,181,943	1,192,914	Colabomycin	t2pks	4%
lanthipeptide	1,360,085	1,383,168			
terpene	1	14,645	Grincamycin	t2pks-saccharide	5%
NRPS-like	21,672	65,643	Azalomycin F	t1pks	8%
oligosaccharide	601,524	625,024			
NRPS-like	55,133	99,209	Elaiophylin	t1pks	12%
bacteriocin	252,476	263,291			
terpene	187,466	208,536	SF2575	t2pks-saccharide-other	6%
T1PKS, NRPS-like	381,943	428,137	Ansamitocin	t1pks	7%
NRPS	69,801	127,939			
terpene	304,195	326,3	Rabelomycin	t2pks-saccharide	4%
bacteriocin	375,353	387,161			
T1PKS,hglE-KS	205,009	255,729			

NRPS-like, lanthipeptide, NRPS	56,889	165,389	Dynemicin	t1pks+t2pks	3%
NRPS, bacteriocin, other	36,278	218,951	Kutznerides	NRPS	100%
NRPS,bacteriocin	223,133	289,263	Ecumicin	NRPS	21%
thiopeptide, LAP, lanthipeptide, T2PKS, T1PKS, NRPS-like, indole	39,794	164,933	Erdasporine	alkaloid	71%
bacteriocin	216,397	228,331			
amglyccycl, PKS-like, T1PKS	83,678	163,644	ML-449	t1pks	25%
T1PKS	257,967	301,008	Clifednamide	nrps-t1pks	20%
T1PKS	205,735	251,911	Macrotermycins	polyketide	7%
NRPS, T1PKS, bacteriocin	51,283	115,924	Tyrobetaine	NRPS	46%
lanthipeptide	179,814	203,641	Chlorizidine A	nrps-t1pks	7%
ectoine	220,667	231,068	Ectoine	other	100%
NRPS-like	99,631	142,88	Herboxidiene	t1pks+t3pks	4%
NRPS-like	177,175	220,999	Feglymycin	NRPS	10%
indole, terpene	108,109	133,657	Xiamycin	terpene-alkaloid	13%
T2PKS, hglE-KS	36,657	130,153	Rubrolone A / rubrolone B	alkaloid	22%
oligosaccharide, T1PKS	152,11	204,047	Aculeximycin	t1pks	50%
NRPS, ladderane	1	65,925	WS9326	NRPS	27%
terpene	617	22,791	Geosmin	terpene	100%
NRPS, T3PKS	75,83	156,833	Avoparcin	NRPS	89%
butyrolactone	87,454	98,14			
T3PKS	94,497	131,082	Alkyl-O-Dihydrogeranyl-Methoxyhydroquinones	terpene-t3pks	28%

T1PKS	1	68,537	Fluvirucin b2	t1pks	44%
butyrolactone	12,855	23,829			
terpene	43,061	63,93	Meridamycin	nrps-t1pks	5%
NRPS	69,52	118,254	Scabichelin	NRPS	70%
T1PKS, NRPS	1	106,776	Aculeximycin	t1pks	35%
indole	1	21,433	BE-54017	alkaloid	85%
T2PKS	23,781	96,239	BE-7585A	t2pks	23%
T3PKS, betalactone, NRPS	8,824	71,027	Furaquinocin A	terpene-t3pks	8%
NRPS	1	42,015	Mildiomycin	other	11%
CDPS	10,747	31,466			
terpene, NRPS	1	52,099	Paulomycin	other	5%
NRPS, ladderane	48	61,205	Ishigamide	nrps-t2pks	55%
terpene	41,92	59,99			
T2PKS	9,539	47,431	Mithramycin	t2pks-saccharide	20%
lassopeptide	9,67	25,986	Saquayamycin	polyketide	5%
NRPS	1	19,919			