Draft genome sequence of *Actinocorallia populi* A251^T, an actinomycetes producing polyketides and nonribosomal polypeptides

Xiaoguang Li¹ · Zhikuan Wang¹ · Longzhan Gan¹ · Yongqiang Tian¹

Received: 16 December 2019 / Accepted: 14 January 2020 / Published online: 30 January 2020
© King Abdulaziz City for Science and Technology 2020

Abstract

In this work, we present the draft genome sequence of *Actinocorallia populi* A251^T consisting of 8,253,402 bp with a G + C content of 71.5 mol%. The genome sequence includes 3 5S rRNA genes, 54 tRNA genes, 28 snRNA genes and 19 CRISPRs in 26 contigs. Using the Rapid Annotation using Subsystem Technology, a total of 7766 coding DNA sequences, which were assigned to 436 subsystems, were found in the genome. And 6743 protein-coding sequences with designated functions were assigned to 19 categories in the Cluster of Orthologous Groups database. The genome contains 20 gene clusters responsible for the synthesis of secondary metabolites, including two type I polyketide synthase (PKS) gene clusters, two type III PKS gene clusters, three nonribosomal peptide synthetase (NRPS) gene clusters and two hybrid PKS/NRPS gene clusters. These results indicate that strain A251 has the ability to produce several nonribosomal polypeptides and diverse polyketides, which was verified by liquid chromatography-mass spectrometry analysis of the extraction from fermentation broth. In addition, lactones, indole and many unknown natural products are present in the metabolites. Bioactivity testing revealed anticancer and antidiabetic activities of the crude extract towards the cancer cell line HepG2. In summary, the genomic sequences and bioactive fermentation products demonstrate the potential of strain A251 in biotechnology.

Keywords Draft genome · *Actinocorallia populi* · Genomic annotation · Crude extract

Abbreviations

CDSs	Coding DNA sequences
PKS	Polyketide synthase
PK	Polyketide
NRPS	Nonribosomal polypeptide synthetase
NRP	Nonribosomal polypeptide
BGC	Biosynthetic gene cluster

Introduction

The genus *Actinocorallia*, belongs to the family *Thermomonosporaceae* (Zhang et al. 2001), is described as rare environmental bacteria with nine valid species. *Actinocorallia* strains are Gram-stain-positive, aerobic and can form spore chains without mobility (Iinuma et al. 1994), and widely distributed, including in natural caves (Lee 2006), the root of *Populus adenopoda* (Li et al. 2018), the head of *Lasius fuliginosus* L (Liu et al. 2016) and soil from Bangkok (Iinuma et al. 1994). *Actinocorallia*, a member of the actinomycetes, has great potential as an important source of new bioactive substances, such as enzymes, antibiotics, polyketides and nonribosomal peptides. These metabolites have anticancer, antibacterial, immunosuppression and other pharmacological activities (Demain and Sanchez 2009). Several compounds were isolated from the liquid fermentation of *Actinocorallia aurantiaca*, including two new furan-containing polyketides, aurantiadioic acids A and B, and aurantoic acid A. All of the last three compounds showed weak anti-NO activity and aurantoic acid A also displayed antiviral activity (Han et al. 2019).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-020-2070-4>) contains supplementary material, which is available to authorized users.

✉ Yongqiang Tian
yqtian@scu.edu.cn

¹ Key Laboratory of Leather Chemistry and Engineering, Ministry of Education, College of Biomass Science and Engineering, Sichuan University, No. 24 South Section 1, Yihuan Road, Chengdu 610065, People's Republic of China

Up to now, numerous actinomycetes have been isolated and classified using polyphasic taxonomy, but useful strains are often shelved or miss-assigned owing to the limitations of conventional and commercial phenotyping test methods. With the rapid development and increasing availability of next-generation sequencing technologies, an increasing number of whole genomes from actinomycetes have been sequenced, assembled and annotated, resulting in genetic information for many novel compounds. On the other hand, genome mining technique is currently being used to explore uncharacterized genes. For example, in silico analysis of *Streptomyces* sp. strain 130 revealed genes associated to the synthesis of extracellular enzymes, including chitinase, xylanase and cellulase (Kumar et al. 2019). In another recent work, Shao et al (2019) reported the genome characterization of *Streptomyces albogriseolus* LBX-2 to explore mechanisms underlying the ability of the strain to degrade polyethylene.

Previously, an endophytic actinomycete, designated A251^T, was isolated from the root of *Populus adenopoda* Maxim, a medicinal plant in China. The isolate was characterized using a polyphasic approach and we proposed it as a novel species within the genus *Actinocorallia* with the valid name *Actinocorallia populi* (Li et al. 2018). The novel grew optimally at 28 °C and pH 7.0, showed nitrate reduction activity, was able to liquefy gelatin and use a variety of carbon and nitrogen sources. As currently there are no reports on the genomic characteristics and biosynthetic potential of *Actinocorallia* species, we sequenced the genome of strain A251 and explored its genetic characteristics. The fully sequenced genome of this rare strain will be a valuable tool for obtaining specific information about the production of bioactive compounds, and will allow identification of novel natural products while avoiding the time- and resource-consuming repeated “discovery” of known compounds.

Genome sequencing, annotation and bioinformatic analyses

A. populi A251^T was cultured in 500 mL of ISP2 broth medium (Shirling and Gottlieb 1966) in shake flasks at 28 °C for 3 days. Cells were harvested by centrifugation at 8000×g, and then washed three times with ultra-pure water. Genomic DNA was isolated from fresh cells and purified as described previously (Pitcher et al. 1989). DNA quality was assessed by running the sample on a 1% (w/v) agarose gel, and DNA quantification was determined using a Nanodrop Life Spectrophotometer (Thermo Fisher Scientific, USA).

Genomic samples were sequenced using Solexa paired-end sequencing technology (Illumina HiSeq system) at Shanghai Personalbio Bioinformatics Technology Co., Ltd. A total of 12,659,012 paired-end reads (2 × 150 bp)

were obtained, with an average sequencing depth of 221-fold. The reads data were filtered by quality-control steps into 12,398,438 (accounted for 97.9% of the total reads) high-quality reads, which were assembled using the SOAPdenovo version 2.04 (Luo et al. 2012) into 25 scaffolds or 26 contigs with sizes varying from 1508 bp to 1,405,409 bp. Finally, the genomic data have been deposited at DDBJ/ENA/GenBank under the accession number PVZV00000000 and automatically annotated by the NCBI Prokaryotic Genome Annotation Pipeline.

Strain A251 was compared to other type species at the genomic level by employing average nucleotide identity (ANI) and digital DNA–DNA hybridization (dDDH), using the Orthologous Average Nucleotide Identity Tool (Lee et al. 2016) and the Genome-to-Genome Distance Calculator 2.1 web service (Meier-Kolthoff et al. 2013), respectively. Additionally, clustered regularly interspaced short palindromic repeats (CRISPRs) were identified by CRISPR finder (Grissa et al. 2007). Noncoding RNA, such as transfer RNA (tRNA), ribosome RNA (rRNA) and small nuclear RNA (snRNA) were predicted using tRNAscan-SE version 1.3.1 (Lowe and Eddy 1997), rRNAmmer (Lagesen et al. 2007) and Rfam database (Burge et al. 2013), respectively.

The Rapid Annotation using Subsystem Technology (RAST) web server (<https://rast.nmpdr.org/rast.cgi>; Aziz et al. 2008) was applied to annotate the coding DNA sequences (CDSs) in the draft genome. The functional annotation of each CDS was made by homology search against Non-Redundant Protein Database (NR; Li et al. 2002), Cluster of Orthologous Groups (COG; Tatusov et al. 2003), Swiss-prot (Amos and Rolf 2000), and Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa et al. 2004) databases. Besides, all carbohydrate-active enzymes (CAZymes) were predicted by dbCAN meta-server (Zhang et al. 2018), and antiSMASH (antibiotics and Secondary Metabolite Analysis Shell; Medema et al. 2011) was used to predict the biosynthetic gene clusters and secondary metabolites.

The draft genome of strain A251 comprises 8,253,402 bases, with a G + C content of 71.5 mol%. Five rRNA (5S, 16S and 23S) genes, 54 tRNA genes, 28 snRNA genes and 19 CRISPRs were located in the genome (Table 1). Comparison of the genome of strain A251 with the publicly released genome of *Actinocorallia herbida* DSM 44254^T (accession number RJKE00000000) showed that the length of the latter was 9.5 Mb and composed of 1 contig, whereas the draft genome of strain A251 was higher in genome coverage. Furthermore, the maximum values of the orthoANI, original ANI and dDDH between strain A251 and the closely related species were found to be 78.5%, 77.4%, and 22.6% (Table 2), respectively, which were all below the accepted thresholds (Richter and Rosselló-Móra 2009; Wayne et al. 1987), indicating that strain A251 had unique genomic properties.

Table 1 Summary of the genome's assembly and annotation information

Property	Term
Finishing quality	High-quality draft
Assembly size (bp)	8,253,402
Number of contigs	26
N50 (bp)	749,123
Fold coverage (×)	221
GC content (mol%)	71.5
Protein coding genes	7766
5S rRNA	3
16S rRNA	1
23S rRNA	1
tRNA	54
snRNA	28
CRISPR	19
Genes annotated with NR	7180
Genes annotated with COG	6743
Genes annotated with KEGG	4975
Genes annotated with Swiss-prot	4863
Bioproject	PRJNA431515
Genbank ID	PVZV01000001- PVZV01000026

Using RAST, a total of 7766 coding DNA sequences were predicted, with the gene length of 7,290,987 bp, accounting for 88.3% of the draft genome. And 436 subsystems were determined by RAST server (Fig. 1), the most abundant being carbohydrates (626 genes), followed by amino acids and derivatives (531 genes), cofactors, vitamins, prosthetic groups, pigments (445 genes), protein metabolism (313 genes) and membrane transport (157 genes). Of them, *A. populi* A251^T contains 60 specific functional genes that encode some substances to against several antibiotics and toxic compounds including 16, 4, 8, 1 genes resistant to cobalt–zinc–cadmium, fluoroquinolones, arsenic, chromium compounds, respectively, and 10 genes for copper homeostasis. Strain A251 possesses 13 iron acquisition and metabolism genes, 1 gene encoding siderophores. In the previous study, strain A251 showed significant protease activity when using gelatin as a substrate, which may be attributed to the presence of 37 genes involved in protein degradation. The majority of them belong to metalloproteases (EC 3.4.17; 8 genes), aminopeptidases (EC 3.4.11; 6 genes), dipeptidases (EC 3.4.13; 1 gene) and serine endopeptidases (EC3.4.21; 2 genes). Four denitrifying reductase gene clusters, nitrate and nitrite ammonification (14 genes) and ammonia assimilation (10 genes) from strain A251 have

Table 2 The ANI and dDDH values between strain A251^T and the closely related species

<i>Actinocorallia populi</i> A251 ^T vs	ANI (%)		dDDH (%)		
	OrthoANI	Original ANI	Formula 1	Formula 2	Formula 3
<i>Actinocorallia herbida</i> DSM 44254 ^T	78.5	77.4	22.6	22.4	21.6
<i>Actinomadura hibisica</i> NBRC 15177 ^T	75.2	74.2	16.2	21.1	16.1
<i>Actinomadura kijaniata</i> NBRC 14229 ^T	75.3	73.9	16.0	21.2	16.0
<i>Actinomadura oligospora</i> ATCC 43269 ^T	74.4	73.0	14.8	21.1	15.0
<i>Actinomadura chibensis</i> NBRC 106107 ^T	74.4	73.0	14.9	20.6	15.1

Fig. 1 An overview of the subsystems of annotated genome of strain *A. populi* 251^T based on RAST annotation. Pie chart showing functional annotation of various subsystem genes

been noted, and the 28 genes may be the main factors for its reduction of nitrate to nitrite. Within the metabolism of aromatic compounds, strain A251 harbors the genes for degrading salicylate ester (7 genes), quinate (1 gene), biphenyl (7 genes), benzoate (2 genes) and p-hydroxybenzoate (1 gene), and can be used for biodegradation of toxic compounds. Finally, 157 genes predicted to be involved in stress response were also identified, suggesting that *A. populi* A251^T can adapt to a variety of environments.

A general overview of biological features encoded in the genome was obtained using COG, and 6743 CDSs with designated functions were assigned to 19 categories (Table S1). We can see that in *A. populi* A251^T genome, the number of genes related to transcription (K), signal transduction mechanisms (T), energy production and conversion (C), and amino acid transport and metabolism (E) were more than other function-related genes. Furthermore, the CDSs were further annotated and divided into 43 functional classes according to KEGG pathway analysis (Fig. S1). The metabolic networks of strain A251 was obtained, including for purine, carbon, pyrimidine, nitrogen, amino acid and fatty acid metabolism, which could ensure the diversity of metabolites. Among all the KEGG pathways, carbon metabolism accounted for the largest proportion, and *A. populi* A251^T has the complete oxidative phosphorylation, glycolysis,

gluconeogenesis, tricarboxylic acid cycle and pentose phosphate pathway.

Genes related to carbohydrate-active enzymes were predicted by CAZy database, which returned 239 results including glycoside hydrolases (GHs, 82 genes), glycosyl transferases (GTs, 82 genes), polysaccharide lyases (PLs, 1 gene), carbohydrate esterases (CEs, 48 genes) and auxiliary activities (AAs, 26 genes). The presence of GHs and PLs in the genome indicated the potential of strain A251 to assimilate a wide range of polysaccharides such as agar, arabinoxylan and carrageenan. In the GH analysis, several hydrolases, such as glucosaccharase, lactase, agarase, galactase and α -L-rhamnosidase, were found to have signal peptide sequences, thus they can be secreted outside the cell membrane.

A. populi A251^T can produce many important secondary metabolites, and a total of 20 gene clusters were predicted to be involved in secondary metabolism by antiSMASH (Fig. S2). They are mainly focused on nonribosomal peptides, ectoine, and diverse polyketides (type I, II and III). By the comparison of antiSMASH, although several biosynthetic gene clusters (BGCs) are highly homologous or completely matched with known secondary metabolite synthesis genes, such as geosmin, ectoine and coelichelin, most have very low similarity with known clusters (Table 3). A cluster of biosynthetic genes from strain A251 share the highest gene

Table 3 Putative gene clusters coding for secondary metabolites in *A. populi* A251^T

Type	Length (bp)	Most similar known cluster	% of genes show similarity	MiBIG BGC-ID
T2pks-T1pks-Otherks	55,714	Hedamycin	50	BGC0000233-c1
T1pks	46,243	—	—	—
Terpene	22,196	Geosmin	100	BGC0000661-c1
Terpene	20,937	Frankiamicin	14	BGC0001197-c1
Terpene	21,029	Chlortetracycline	5	BGC0000209-c1
Bacteriocin	10,819	—	—	—
Indole-Lantipeptide	33,971	—	—	—
Butyrolactone	11,812	Paromomycin	5	BGC0000712-c1
Nrps	62,854	Laspertomycin	9	BGC0000379-c1
T3pks	41,113	Kendomycin	25	BGC0001066-c1
T3pks	41,050	Alkylresorcinol	66	BGC0000282-c1
Ectoine	10,396	Ectoine	100	BGC0000853-c1
Siderophore	12,924	—	—	—
Nrps	49,336	Enduracidin	4	BGC0000341-c1
Lasso peptide	22,546	A54145	3	BGC0000291-c1
Nrps	43,180	Kirromycin	3	BGC0001070-c1
T1pks-Nrps	38,088	Collismycin	7	BGC0000973-c1
T1pks-Nrps	107,233	Coelichelin	90	BGC0000325-c1
Terpene	20,900	Isorenieratene	18	BGC0001227-c1
T1pks	45,829	Maduropeptin	3	BGC0001008-c1

Secondary metabolite types detected by antiSMASH. T1pks: Type I PKS; T2pks: Type II PKS; T3pks: Type III PKS; Otherks: other types of PKS cluster; PKS: polyketide synthase; NRPS: Nonribosomal polypeptide synthetase; Bacteriocin: Bacteriocin or other unspecified ribosomally synthesis and post-translationally modified peptide product (RiPP) cluster; Lasso peptide: Lasso peptide cluster

cluster similarity to coelichelin (90%) of *Streptomyces coelicolor* A3 (Challis and Rave 2000). Their gene clusters are consistent in the core biosynthetic domains (ctg133 and SCO0492), and they also have six transport-related genes, two additional biosynthetic genes near the core biosynthetic gene. However, the biosynthetic pathway of strain A251 is more complicated with more additional biosynthetic genes and regulatory genes (Fig. S3). Most notably, some BGCs of NRPS are conserved and lack closely related empirically characterized homologs compared with known biosynthetic gene clusters in the antiSMASH database. For instance, a characterized BGC for the laspartomycin (BGC0000379-c1) has a mere 9% gene match to the homologous BGC in *A. populi*, indicating that BGC may encode a novel NRPS. The presence of RAST-annotated iron acquisition and antibiosis genes, as well as antiSMASH predicted BGCs, confirm the genetic potential of strain A251 to synthesize siderophores and multiple classes of antimicrobial compounds.

Extraction, identification and activity tests of fermentation products

To further elucidate the application potential of strain A251, we performed mass spectrometry identification and activity experiments on its metabolites. Strain A251 was inoculated into 1 L fermentation medium [per litre: 10 g glucose, 2 g yeast extract, 20 g oat meal, 3 g CaCO₃, 25 g starch and 10 g peptone (pH 7.2–7.4)] at 28 °C for 7 days with continuous shaking at 220 rpm. After centrifugation at 7000×g for 20 min, the cells were frozen with liquid nitrogen, ground and mixed with the fermentation broth. The mixed sample was extracted with 4 volumes of methanol, stirred for 30 min and left standing at 4 °C for 1 h. The upper organic solvent extract was retained, while the aqueous layer was discarded after centrifugation. The organic phase was concentrated under reduced pressure using a rotary evaporator, and vacuum-dried to obtain a crude product which was experimentally evaluated to detect direct and indirect activity tests. The crude extract was dissolved in DMSO and diluted to 10 µg mL⁻¹ with sterile ultrapure water. The anticancer activity of the dry extract was tested in vitro by the SRB assay (Skehan et al. 1990; Vichai and Kirtikara 2006) in the human cancer cell HepG2. Meanwhile, the ability of the crude extract to protect against type II diabetes was measured according to the previous publication (Sauma et al. 2006).

The crude extract of strain A251 could inhibit the growth of cancer cell HepG2 throughout the anticancer activity experiment, and promote the growth of cell lines in the antidiabetes activity test. Thus, some substances in the crude extract from strain A251 have the ability to against cancer and diabetic. For understanding the active ingredients of crude extract, dissolved it in 80% (v/v) methanol,

centrifuged, and then took the supernatant for LC–MS analysis at Beijing Novogene Company. The raw data were properly processed, and used to predict the molecular formulas of the metabolites by comparison with the mzCloud and ChemSpider databases. Based on the comparison results, a variety of known polyketides and nonribosomal polypeptides were detected in crude extract (Table S3), as well as metabolites of the following types: lactone, indole, pyrimidine, nicotianamine, curacin A and unidentified natural products. These metabolites are roughly consistent with the predictions of antiSMASH and demonstrate the diversity of the metabolites produced by strain A251. On the other hand, the polyketides in the metabolites could display anti-cancer and anti-diabetes activities (Packter 1993; Sun and Deng 2006). Also, curacin A, an antitubulin natural product isolated from the cyanobacterium *Lyngbya majuscula* (Chang et al. 2004), occurred in the metabolites of strain A251, which provided a new source for its separation and extraction. In addition, nicotianamine could help plants transport, absorb and balance certain metal ions (Wiren et al. 1999).

Conclusion

In this study, we determined the draft genome sequence of *A. populi* A251^T and used the sequence to perform bioinformatics analyses. A total of 7766 protein coding genes and 87 RNA genes were identified in 26 contigs. Our results revealed the function of predicted genes in the genome of A251^T through different databases, and showed the diversity of bioactive compounds produced by the endophyte. Furthermore, the anticancer and antidiabetes activities were mainly attribute to polyketides and non-ribosomal polypeptides which we detected in the fermentation products of strain A251. Also, many unknown compounds, such as polyketides, nonribosomal peptides and antibiotics, are waiting to be discovered and applied. To our knowledge, this is the first genomic analysis from a species within the genus *Actinocorallia*. Importantly, our work has provided genomic data and experimental evidence to support the view that *A. populi* A251^T is a promising bacterium-producing bioactive compound.

Strain informations

The strain A251 has been deposited at Japan Collection of Microorganisms (JCM, JCM 32178^T) and China General Microbiological Culture Collection Center (CGMCC, CGMCC 4.7421^T).

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession PVZV00000000.

Acknowledgements We thank the Institute of Materia Medica of Chengdu Diao pharmaceutical Corp. for conducting the anticancer and antidiabetic activity screening. This work was supported by the National Key Research and Development Program of China (2017YFB0308401); and the Opening Project of Key Laboratory of Leather Chemistry and Engineering (Sichuan University), Ministry of Education (20826041C4159).

Author contributions All the authors contributed to this work. YT supported and guided the experiments; XL performed the experiments and analyzed the data, and wrote the paper; ZW and LG participated in the data analysis. All authors reviewed and approved the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that there are no conflicts of interest.

Ethical statement This work does not contain any studies with human participants or animals performed by any of the authors.

References

- Amos B, Rolf A (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res* 28:45–48
- Aziz RK, Bartels D, Best AA et al (2008) The RAST server: rapid annotations using subsystems technology. *BMC Genom* 9:75–90
- Burge SW, Daub J, Eberhardt R et al (2013) Rfam 11.0: 10 years of RNA families. *Nucl Acids Res* 41:D226–D232
- Challis GL, Ravel J (2000) Coelichelin, a new peptide siderophore encoded by the *streptomyces coelicolor* genome: structure prediction from the sequence of its non-ribosomal peptide synthetase. *FEMS Microbiol Lett* 187:111–114
- Chang Z, Sitachitta N, Rossi JV et al (2004) Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium. *J Nat Prod* 67:1356–1367
- Demain AL, Sanchez S (2009) Microbial drug discovery: 80 years of progress. *J Antibiot* 62:5–16
- Grissa I, Vergnaud G, Pourcel C (2007) Crisprfinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucl Acids Res* 35:52–57
- Han KY, Wu X, Jiang CL et al (2019) Three new compounds from the actinomycete *Actinocorallia aurantiaca*. *Natural Products and Bioprospecting* 9:351–354
- Iinuma S, Yokota A, Hasegawa T et al (1994) *Actinocorallia* gen. nov., a new genus of the order Actinomycetales. *Int J Syst Bacteriol* 44:230–234
- Kanehisa M, Goto S, Kawashima S et al (2004) The KEGG resource for deciphering the genome. *Nucl Acids Res* 32:D277–D280
- Kumar M, Kumar P, Das P et al (2019) Draft genome of *Streptomyces* sp. strain 130 and functional analysis of extracellular enzyme producing genes. *Mol Biol Rep* 46(5):5063–5071
- Lagesen K, Hallin PF, Rødland EA et al (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucl Acids Res* 35:3100–3108
- Lee SD (2006) *Actinocorallia cavernae* sp. nov. isolated from a natural cave in Jeju, Korea. *Int J Syst Evol Microbiol* 56(5):1085–1088
- Lee I, Kim YO, Park SC et al (2016) OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103
- Li W, Jaroszewski LA, Godzik A (2002) Tolerating some redundancy significantly speeds up clustering of large protein databases. *Bioinformatics* 18:77–82
- Li XG, Wang ZK, Lu FJ et al (2018) *Actinocorallia populi* sp. nov., an endophytic actinomycete isolated from a root of *Populus adenopoda* (maxim.). *Int J Syst Evol Microbiol* 68:2325–2330
- Liu C, Li Y, Ye L et al (2016) *Actinocorallia lasiicapitis* sp. nov. an actinomycete isolated from the head of an ant (*Lasius fuliginosus* L.). *Int J Syst Evol Microbiol* 66:2172–2177
- Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucl Acids Res* 25:955–964
- Luo R, Liu BH, Xie YL et al (2012) Soapdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1:18
- Medema MH, Blin K, Cimermancic P et al (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucl Acids Res* 39:W339–W346
- Meier-kolthoff JP, Auch AF, Klenk HP et al (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinf* 14:60
- Packer NM (1993) The polyketide metabolites, D. O'Hagan Ellis Horwood, New York. *Int J Biol Macromol* 15(2):134
- Pitcher DG, Saunders NA, Owen RJ (1989) Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett Appl Microbiol* 8:151–156
- Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 106:19126–19131
- Sauma L, Sc M, Karin G et al (2006) PPAR- γ response element activity in intact primary human adipocytes: effects of fatty acids. *Nutrition* 22:60–68
- Shao HH, Chen MJ, Fei XT et al (2019) Complete genome sequence and characterization of a polyethylene biodegradation strain, *Streptomyces Albogriseolus* LBX-2. *Microorganisms* 7:379
- Shirling EB, Gottlieb D (1966) Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol* 16:313–340
- Skehan P, Storeng R, Scudiero D et al (1990) New colorimetric cytotoxicity assay for anticancer-drug screening. *J Natl Cancer I* 82:1107–1112
- Sun YH, Deng ZX (2006) Polyketides and combinatorial biosynthetic approaches. *Chin J Antibiot* 1:6–14
- Tatusov RL, Fedorova ND, Jackson JD et al (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinf* 4:41
- Tatusova TA, DiCuccio M, Badretdin A et al (2016) NCBI prokaryotic genome annotation pipeline. *Nucl Acids Res* 44(14):6614–6624
- Vichai V, Kirtikara K (2006) Sulforhodamine b colorimetric assay for cytotoxicity screening. *Nat Protoc* 1:1112–1116
- Wayne LG, Brenner DJ, Colwell RR et al (1987) International committee on systematic bacteriology. Report of the ad hoc committee on the reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* 37:463–464
- Wiren NV, Klair S, Bansal S et al (1999) Nicotianamine chelates both FeIII and FeII. Implications for metal transport in plants. *Plant Physiol* 119:1107–1114
- Zhang Z, Kudo T, Nakajima Y et al (2001) Clarification of the relationship between the members of the family *Thermomonosporaceae* on the basis of the 16S rDNA, 16S–23S rRNA internal transcribed spacer and 23S rDNA sequences and chemotaxonomic analyses. *Int J Syst Evol Microbiol* 51:373–383
- Zhang H, Yohe T, Huang L et al (2018) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucl Acids Res* 6:W95–W101