

# High-Quality Draft Genome Sequence of the Actinobacterium *Nocardia terpenica* IFM 0406, Producer of the Immunosuppressant Brasilicardins, Using Illumina and PacBio Technologies

Anina Buchmann,<sup>a,b</sup> Michael Eitel,<sup>c</sup> Pierre Koch,<sup>c</sup> Paul N. Schwarz,<sup>d</sup> Evi Stegmann,<sup>b,d</sup> Wolfgang Wohleben,<sup>b,d</sup> Marcin Wolański,<sup>e</sup> Michał Krawiec,<sup>e</sup> Jolanta Zakrzewska-Czerwińska,<sup>e,f</sup> Carmen Méndez,<sup>g</sup> Alma Botas,<sup>g</sup> Luz Elena Núñez,<sup>h</sup> Francisco Morís,<sup>h</sup> Jesus Cortés,<sup>h</sup> Harald Gross<sup>a,b</sup>

Department of Pharmaceutical Biology, Institute of Pharmaceutical Sciences, University of Tübingen, Tübingen, Germany<sup>a</sup>; German Centre for Infection Research (DZIF), Partner Site Tübingen, Tübingen, Germany<sup>b</sup>; Department of Medicinal and Pharmaceutical Chemistry, Institute of Pharmaceutical Sciences, University of Tübingen, Tübingen, Germany<sup>c</sup>; Department of Microbiology and Biotechnology, Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, Tübingen, Germany<sup>d</sup>; Faculty of Biotechnology, University of Wrocław, Wrocław, Poland<sup>e</sup>; Ludwik Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wrocław, Poland<sup>f</sup>; Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias, Universidad de Oviedo, Oviedo, Spain<sup>g</sup>; EntreChem S.L., Oviedo, Spain<sup>h</sup>

**The bacterium *Nocardia terpenica* IFM 0406 is known as the producer of the immunosuppressant brasilicardin A. Here, we report the completely sequenced genome of strain IFM 0406, which facilitates the heterologous expression of the brasilicardin biosynthetic gene cluster but also unveils the intriguing biosynthetic capacity of the strain to produce secondary metabolites.**

Received 17 October 2016 Accepted 21 October 2016 Published 15 December 2016

**Citation** Buchmann A, Eitel M, Koch P, Schwarz PN, Stegmann E, Wohleben W, Wolański M, Krawiec M, Zakrzewska-Czerwińska J, Méndez C, Botas A, Núñez LE, Morís F, Cortés J, Gross H. 2016. High-quality draft genome sequence of the actinobacterium *Nocardia terpenica* IFM 0406, producer of the immunosuppressant brasilicardins, using Illumina and PacBio technologies. *Genome Announc* 4(6):e01391-16 doi:10.1128/genomeA.01391-16.

**Copyright** © 2016 Buchmann et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Harald Gross, [harald.gross@uni-tuebingen.de](mailto:harald.gross@uni-tuebingen.de).

Actinobacteria of the genus *Nocardia* are often found to be human and animal pathogens that cause pulmonary, cutaneous and subcutaneous diseases (1–3). However, species in this genus also produce a plethora of potent bioactive natural compounds with therapeutic potential. For example, *Nocardia terpenica* IFM 0406 (formerly referred to as *Nocardia brasiliensis* IFM 0406) is known to produce the antifungal and immunosuppressant brasilinolides A–C (4, 5) and brasilicardins A–D (6–9). The immunosuppressant brasilicardin A is of particular interest because its potency is superior to today's standard immunosuppressive drugs for organ transplantation, but it is less toxic, since it mediates its activity by a different mode of action (10). The development of this promising lead structure is currently hampered due to supply issues. The producer strain only shows a low production titer and is furthermore categorized as a biosafety-level 2 organism. Both of these facts make the fermentative production of brasilicardin A expensive and elaborate. Due to its complex stereochemistry, a total synthesis has not been achieved to date (11–13). Therefore, we want to overcome this hurdle by heterologous expression of the corresponding biosynthetic gene cluster (14) in a nonpathogenic producer strain. In order to obtain the full biosynthetic pathway, including the related precursor metabolic pathways, we sequenced the genome of *N. terpenica* IFM 0406.

Genome sequencing was carried out using the Illumina HiSeq 2500 and Pacific Biosciences RS II sequencing platforms. Genomic DNA was isolated (ZR fungal/bacterial DNA mini prep kit, Zymo Research) and a paired-end library with bar coding was constructed (Nextera DNA sample preparation kit). The HiSeq paired-end data were processed and filtered, and a quality analysis

was performed. This resulted in 6,564,821 reads with an average length of 509 bp (approx. 360× coverage). The quality of the Illumina FASTQ sequences was enhanced by trimming off low-quality bases, which was followed by *de novo* assembly using CLC Genomics Workbench version 7.0.4 to yield contigs. The optimal *k*-mer size was determined using KmerGenie (15). In addition, a 10-kb genomic sublibrary of strain IFM 0406 was produced and sequenced with PacBioSMRT technology. The resultant continuous long-read (CLR) data were filtered using the SMRT Analysis software suite. In summary, one run produced 113,652 filtered reads with average read lengths of 3,262 bp (approx. 40× coverage). The contigs were linked and placed into super-scaffolds based on the alignment of the PacBio CLR reads using BLASR (16) and SSPACE-LongRead scaffolder version 7.0.4 (17). The gapped regions within the super-scaffolds were closed using GapFiller version 1.10 (18).

The final gap-closed genome comprises three scaffolds with a total size of 9,276,856 bp and a G+C content of 68.5%. The genome contains 8,755 open reading frames, including 76 tRNA genes and nine rRNA operons. Bioinformatic analyses using antiSMASH version 3.0 (19) revealed that IFM 0406 possessed, besides the putative brasilicardin (14) and brasilinolide (20) gene cluster, 47 further orphan natural product biosynthesis gene clusters. Insight into the genome of strain IFM 0406 not only reveals a high potential to produce secondary metabolites, but will also enable and facilitate the heterologous expression and biotechnological production of brasilicardins in a nonpathogenic host system.

**Accession number(s).** This genome sequence has been deposited in EMBL/GenBank under the accession number

LWGR0000000. The version described in this paper is the first version.

## ACKNOWLEDGMENT

We thank BaseClear BV (Leiden, Netherlands) for sequencing the *Nocardia terpenica* genome, especially Walter Pirovano and his team for helpful advice.

## FUNDING INFORMATION

This work, including the efforts of Carmen Mendez, was funded by Ministerio de Economía y Competitividad (MINECO) (PCIN-2014-066). This work, including the efforts of Francisco Morris, was funded by Ministerio de Economía y Competitividad (MINECO) (PCIN-2014-097). This work, including the efforts of Harald Gross and Pierre Koch, was funded by Bundesministerium für Bildung und Forschung (BMBF) (FKZ 031A568A). This work, including the efforts of Wolfgang Wohlleben, was funded by Bundesministerium für Bildung und Forschung (BMBF) (FKZ 031A568B). This work, including the efforts of Jolanta Zakrzewska-Czerwińska, was funded by MNiSW | Narodowe Centrum Badań i Rozwoju (NCBR) (ERA-NET-IB/NeBrasCa/10/2015).

## REFERENCES

1. Wilson JW. 2012. Nocardiosis: updates and clinical overview. *Mayo Clin Proc* 87:403–407. <http://dx.doi.org/10.1016/j.mayocp.2011.11.016>.
2. Xia L, Cai J, Wang B, Huang Y, Jian J, Lu Y. 2015. Draft genome sequence of *Nocardia seriolae* ZJ0503, a fish pathogen isolated from *Trachinotus ovatus* in China. *Genome Announc* 3(1):e01223-14. <http://dx.doi.org/10.1128/genomeA.01223-14>.
3. Tan CK, Lai CC, Lin SH, Liao CH, Chou CH, Hsu HL, Huang YT, Hsueh PR. 2010. Clinical and microbiological characteristics of nocardiosis including those caused by emerging *Nocardia* species in Taiwan, 1998–2008. *Clin Microbiol Infect* 16:966–972. <http://dx.doi.org/10.1111/j.1469-0691.2009.02950.x>.
4. Tanaka Y, Komaki H, Yazawa K, Mikami Y, Nemoto A, Tojyo T, Kadowaki K, Shigemori H, Kobayashi J. 1997. Brasilinolide A, a new macrolide antibiotic produced by *Nocardia brasiliensis*: producing strain, isolation and biological activity. *J Antibiot* 50:1036–1041. <http://dx.doi.org/10.7164/antibiotics.50.1036>.
5. Komatsu K, Tsuda M, Tanaka Y, Mikami Y, Kobayashi J. 2004. Absolute stereochemistry of immunosuppressive macrolide brasilinolide A and its new congener brasilinolide C. *J Org Chem* 69:1535–1541. <http://dx.doi.org/10.1021/jo035773v>.
6. Shigemori H, Komaki H, Yazawa K, Mikami Y, Nemoto A, Tanaka Y, Sasaki T, In Y, Ishida T, Kobayashi J. 1998. Brasilicardin A. A novel tricyclic metabolite with potent immunosuppressive activity from actinomycete *Nocardia brasiliensis*. *J Org Chem* 63:6900–6904. <http://dx.doi.org/10.1021/jo9807114>.
7. Komaki H, Nemoto A, Tanaka Y, Takagi H, Yazawa K, Mikami Y, Shigemori H, Kobayashi J, Ando A, Nagata Y. 1999. Brasilicardin A, a new terpenoid antibiotic from the pathogenic *Nocardia brasiliensis*: fermentation, isolation and biological activity. *J Antibiot* 52:13–19. <http://dx.doi.org/10.7164/antibiotics.52.13>.
8. Komatsu K, Tsuda M, Shiro M, Tanaka Y, Mikami Y, Kobayashi J. 2004. Brasilicardins B–D, new tricyclic terpenoids from actinomycete *Nocardia brasiliensis*. *Bioorg Med Chem* 12:5545–5551. <http://dx.doi.org/10.1016/j.bmc.2004.08.007>.
9. Komatsu K, Tsuda M, Tanaka Y, Mikami Y, Kobayashi J. 2005. SAR studies of brasilicardin A for immunosuppressive and cytotoxic activities. *Bioorg Med Chem* 13:1507–1513. <http://dx.doi.org/10.1016/j.bmc.2004.12.029>.
10. Usui T, Nagumo Y, Watanabe A, Kubota T, Komatsu K, Kobayashi J, Osada H. 2006. Brasilicardin A, a natural immunosuppressant, targets amino acid transport system L. *Chem Biol* 13:1153–1160. <http://dx.doi.org/10.1016/j.chembiol.2006.09.006>.
11. Coltart DM, Danishefsky SJ. 2003. Novel synthetic approach to the 8,10-dimethyl-*anti-syn-anti*-perhydrophenanthrene skeleton. *Org Lett* 5:1289–1292. <http://dx.doi.org/10.1021/ol034213f>.
12. Jung ME, Perez F, Regan CF, Yi SW, Perron Q. 2013. *Se*-phenyl prop-2-ene-selenoate: an ethylene equivalent for Diels-Alder reactions. *Angew Chem Int Ed* 52:2060–2062. <http://dx.doi.org/10.1002/anie.201208294>.
13. Jung ME, Chamberlain BT, Koch P, Niazi KR. 2015. Synthesis and bioactivity of a brasilicardin A analogue featuring a simplified core. *Org Lett* 17:3608–3611. <http://dx.doi.org/10.1021/acs.orglett.5b01712>.
14. Hayashi Y, Matsuura N, Toshima H, Itoh N, Ishikawa J, Mikami Y, Dairi T. 2008. Cloning of the gene cluster responsible for the biosynthesis of brasilicardin A, a unique diterpenoid. *J Antibiot* 61:164–174. <http://dx.doi.org/10.1038/ja.2008.126>.
15. Chikhi R, Medvedev P. 2014. Informed and automated *k*-mer size selection for genome assembly. *Bioinformatics* 30:31–37. <http://dx.doi.org/10.1093/bioinformatics/btt310>.
16. Chaisson MJ, Tesler G. 2012. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* 13:238. <http://dx.doi.org/10.1186/1471-2105-13-238>.
17. Boetzer M, Pirovano W. 2014. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* 15:211. <http://dx.doi.org/10.1186/1471-2105-15-211>.
18. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with Gap-Filler. *Genome Biol* 13:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.
19. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* 43:W237–W243. <http://dx.doi.org/10.1093/nar/gkv437>.
20. Chiu HT, Weng CP, Lin YC, Chen KH. 2016. Target-specific identification and characterization of the putative gene cluster for brasilinolide biosynthesis revealing the mechanistic insights and combinatorial synthetic utility of 2-deoxy-L-fucose biosynthetic enzymes. *Org Biomol Chem* 14:1988–2006. <http://dx.doi.org/10.1039/c5ob02292d>.