Archives of Microbiology (2024) 206:76
https://doi.org/10.1007/s00203-023-03799-z

ORIGINAL PAPER

Exploring *Nocardia's* ecological spectrum and novel therapeutic frontiers through whole-genome sequencing: unraveling drug resistance and virulence factors

Shaslinah Nathar[1] · Raji Rajmichael[1] · Chitra Jeyaraj Pandian[2] · Hemavathy Nagarajan[1] · Ahila Mathimaran[1] · Jemima D. Kingsley[3] · Jeyakanthan Jeyaraman[1]

Received: 25 August 2023 / Revised: 14 December 2023 / Accepted: 15 December 2023 / Published online: 24 January 2024

Abstract

Nocardia farcinica is the leading pathogen responsible for nocardiosis, a life-threatening infection primarily affecting immunocompromised patients. In this study, the genomic sequence of a clinically isolated *N. farcinica* sample was sequenced. Subsequently, the assembled genome was annotated to identify antimicrobial resistance and virulence genes, as well as plasmid and prophages. The analysis of the entire genome size was 6,021,225 bp, with a GC content of 70.78% and consists of 103 contigs and N50 values of 292,531 bp. The genome analysis revealed the presence of several antimicrobial resistance genes, including RbpA, mtrA, FAR-1, blaFAR-1, blaFAR-1_1, and rox. In addition, virulence genes such as relA, icl, and mbtH were also detected. The present study signifies that *N. farcinica* genome is pivotal for the understanding of antimicrobial resistance and virulence genes is crucial for comprehending resistance mechanism, and developing effective strategies to combat bacterial infections effectively, especially adhesins and toxins. This study aids in identifying crucial drug targets for combating multidrug-resistant *N. farcinica* in the future.

Keywords *Nocardia farcinica* · DNA sequencing · Antimicrobial resistance gene · Drug resistance and virulence gene

Introduction

Antibiotic resistance is a global healthcare concern, and *Nocardia farcinica* poses a significant threat due to its inherent resistance to multiple antibiotics. The genus of *Nocardia* belongs to the actinomycetes, a group of aerobic bacilli that are found commonly in soil and water (Mehta and Shamoo 2020; Conville et al. 2017). Although, there are more than 80 species in *Nocardia*, approximately 54 species notably *N. nova* complex, *N. abscessus* complex, *N. transvalensis* complex, *N. farcinica*, *N. asteroides* type VI (*N. cyriacigeorgica*), *N. brevicatena*/*N. paucivorans* complex, and *N. brasiliensis* are pathogenic to humans (Duggal and Chugh 2020). In *Nocardia*, the pathogenesis mechanism is not completely understood (Ji et al. 2020). *Nocardia* species are regarded in the aerobic actinomycete group and virulence in *Nocardia* has been ascribed to its ability to survive and grow in various human cells and evade the immune response by producing antioxidant enzymes (catalase/superoxide dismutase (SOD)), inhibiting formation phagolysosome complex, reducing levels of phosphomonoesterases II in tissue macrophages, secreting toxins and hemolysin (in few cases) (Mehta and Shamoo 2020; Conville et al. 2017). Since the disease is difficult to diagnose and can be left untreated, it can spread to other organs of the body, including the spine and brain (Kövér et al. 2023). Nocardiosis of the brain or spinal cord leads to mortality for more than 85% of them (https://www.cdc.gov/nocardiosis/infection/index.html). Nocardiosis infections can spread through injuries to the subcutaneous tissue. It may result in closely related cellulitis, pyoderma, abscess formation, and *Staphylococcal* or *Streptococcal* infections. However, disseminating the

Communicated by Yusuf Akhter.

✉ Jeyakanthan Jeyaraman
jjbioinformatics@gmail.com

[1] Structural Biology and Bio-Computing Lab, Department of Bioinformatics, Alagappa University, Science Block, Karaikudi 630 003, Tamil Nadu, India

[2] Department of Biotechnology, Dr. Umayal Ramanathan College for Women, Karaikudi 630 003, Tamil Nadu, India

[3] Orbito Asia Diagnostics Private Limited Coimbatore, Coimbatore 641 045, Tamil Nadu, India

infection via lymphatics to the regional lymph nodes may provide insights into lymphocutaneous Nocardiosis (Duggal and Chugh 2020). *Nocardia* infections constitute an important threat to human health, and the diagnosis and etiology of this disease are very important (Zhang et al. 2023). Regretfully, the research on the role of the majority of the genes in *Nocardia* due to has not been yet studied (Ji et al. 2022). Currently, there are few studies on the *N. farcinica* virulence factors, and many potential virulence factors are yet to be discovered. Virulence factors are as follows, Superoxide Dismutase (SOD) and Catalase: *N. farcinica* prevents phagosome-lysosome binding, reduces lysosomal enzyme action in macrophages, neutralizes phagosomal acidification, and inhibits the oxidative killing mechanisms of phagocytes. This resistance is due to the production of surface-associated superoxide dismutase and high levels of catalase. Cell wall glycolipids: Complex cell wall glycolipids act as a virulence factor in *N. farcinica.* In chronic granulomatous disease, neutrophils and macrophages inhibit the burst of oxidative metabolism during phagocytosis, thus reducing the intracellular killing of catalase-positive bacteria such as *N. farcinica*. Although *N. farcinica* may exhibit resistance to a typical oxidative burst. Mammalian cell entry protein family: Mce1C and Mce1D protein (Mammalian cell entry protein family) are situated in the cell wall of *N. farcinica*. Both Mce1C and Mce1D genes are expressed at the level of protein and mRNA and evoke antibody activity during the period of infection. Mce1C and Mce1D proteins inhibit the expression of proinflammatory cytokines and prevent the NF-κB and MAPK signaling pathways, thereby inhibiting the innate immune response. Mycolic acid: *N. farcinica* are actinomycetes, commonly referred to as mycolata, that contain mycolic acid, which helps to provide a defense mechanism for pathogens to resist the host immune systems. rox gene: In *N. farcinica*, the rox gene plays a secondary function in rifampicin resistance by initiating the process of rifampicin breakdown and producing a novel metabolite in the first phase. The rox gene of *N. farcinica* expresses a rifampicin monooxygenase that can change rifampicin into the molecule 2-N-hydroxy-4-oxo-rifampicin, which has resulted in reduced antibiotic activity. phospholipase C: Phospholipase C has the role of destroying tissue, Hemolysin (toxic proteins): destroying red blood cells, lipases, and proteases (Cai et al. 2022; Ji et al. 2022).

Antibiotic resistance is an alarming issue and a worldwide threat faced by the healthcare community (CDC 2019). *N. farcinica* remains a virulent species and exhibits intrinsic resistance to many antibiotics, including third-generation cephalosporins (Mehta and Shamoo 2020). *N. farcinica* is a gram-positive organism that can cause life-threatening infections because of the organism's trend to disseminate rapidly and resist antibiotics (Wu et al. 2021). There is currently no defined procedure for antimicrobial susceptibility testing for guiding clinical therapy, and antimicrobial susceptibility varies depending on the species of *Nocardia* isolates (Lee et al. 2021). It has a unique antibiotic susceptibility pattern and resistance to antimicrobial agents (Gao et al. 2021). Globally, *N. farcinica* is the most common *Nocardia* species that produce pulmonary infections in humans (Yetmar et al. 2023) with a mortality rate of 10–31% in Asia, Europe, and North America (Jiao et al 2021). The majority of sulfonamides, that is trimethoprim–sulfamethoxazole (TMP–SMX), are the standard treatment for nocardiosis based partially on the results of a few retroactive reviews. These indicated found that patients receiving regimens containing sulfonamides had a trend toward longer survival. In patients with sulfonamide allergy or resistance, as well as in cases where clinical treatment failures have been documented, alternative or combination therapy is chosen. The objective of that study was to assess the factors linked to in-hospital mortality in patients with pulmonary nocardiosis, including risk factors, clinical, radiographic, and microbiologic features, as well as outcomes (Rahim et al. 2023). It is extremely virulent and is known to be naturally resistant to numerous antibiotics (Adapa et al. 2020). Antimicrobial susceptibility testing is presently suggested for all *Nocardia* spp. isolates before starting antimicrobials, because sensitivities are commonly difficult to predict, and first-line antibiotic therapy may not tolerate for patients. Cotrimoxazole, imipenem, linezolid, and amikacin recommended as initial therapeutic for nocardiosis, were the antibiotics most commonly utilized and exhibited the most favorable sensitivity patterns. Cotrimoxazole has been thought to be the initial therapeutic for nocardiosis because most studies reveal great sensitivity rates to this drug; however, the number of cotrimoxazole-resistant *Nocardia* is reported to be increasing. Compared with some of the previous studies, we discovered higher rates of resistance to imipenem, ciprofloxacin, and amoxicillin–clavulanate, although resistance to ceftriaxone was lower than other reports (Besteiro et al. 2023). Difficulties in timely detection and challenging diagnoses often lead to treatment delays, resulting in poorer patient outcomes (Wu et al. 2021). *N. farcinica* can lead to various clinical manifestations, including brain abscess, keratitis, bacteremia, and infections in the lungs, kidneys, and skin (Bell et al. 2019). In certain instances, *N. farcinica* has been identified as the causative agent of peritonitis in immunocompromised individuals. The clinical manifestation depends on the patient's medical history, as nosocomial infections caused by *N. farcinica* are frequently encountered in patients

undergoing chemotherapy, dialysis, or receiving treatments for conditions like HIV infection and autoimmune diseases. The most preferred antibiotic for treating *N. farcinica* infections is trimethoprim–sulfamethoxazole (SXT). Yet, because of its rapidly evolving antibiotic resistance, a combination of SXT with other antibiotic medications, including amikacin, imipenem–cilastatin, and moxifloxacin, is currently utilized. Despite all these therapeutic efforts, treated patients often experience a relapse in the disease prognosis, and the mortality rate remains around 39% (Adapa et al. 2020). Sulfonamides, aminoglycosides, β-lactams (penicillins, carbapenems, and cephalosporins) and β-lactam/β-lactamase inhibitors, quinolones, macrolides, and tetracyclines are the six main categories of antimicrobial substances presently in clinical use. Aminoglycoside antibiotics like Amikacin and Tobramycin act as bactericidal by disrupting the 30S ribosomal subunit, causing faulty protein synthesis. β-Lactam antibiotics (Penicillins, Carbapenems, and cephalosporins like Cefotaxime, Ceftriaxone, Cefixime, Cefuroxime) and the β-Lactam-β-lactamase inhibitor amoxicillin–clavulanic acid attain bactericidal activity by impeding bacterial cell wall construction by preventing transpeptidase responsible for catalysis of peptidoglycan cross-linking. Macrolide antibiotics such as Erythromycin and Clarithromycin employ diverse bactericidal mechanisms, with clavulanic acid safeguarding amoxicillin from degradation by deactivating a broad spectrum of β-lactamases. In contrast, the macrolides erythromycin and clarithromycin inhibit protein synthesis by attaching to the 50S ribosomal subunit. Oxazolidinones, exemplified by Linezolid, function as antibacterial agents by preventing the synthesis of bacterial proteins. These antibiotics, such as quinolones (e.g., ciprofloxacin, moxifloxacin), sulfonamides (sulfamethoxazole, trimethoprim/TMP–SMX), and tetracyclines (doxycycline, minocycline), disrupt bacterial processes. Quinolones inhibit gyrase, essential for DNA folding, sulfonamides block the folate pathway, and tetracyclines interfere with anticoagulant and antacid treatments (Nouioui et al. 2020).

In the present study, whole-genome sequencing was performed using the Illumina sequencing platform to decipher the genomic basis of virulence factors and antimicrobial resistance. The attained sequence data were further assembled and annotated utilizing various bioinformatics tools. This research demonstrates the importance of genome-based analysis in assessing the potential health risks associated with emerging *Nocardia* pathogens. The abundance of bacterial genome sequence information has facilitated various distinct methods for identifying therapeutic targets. The objective of the study is to perform the whole-genome sequencing of *N. farcinica* to detect resistance to first- and second-line anti-nocardiosis drugs.

Materials and methods

Isolation and culture conditions of *Nocardia farcinica*

Nocardia farcinica, isolated from sputum samples, was cultured at Frontier Lifeline Hospitals, Chennai, Tamil Nadu, India. Conventional culture methods were employed using Luria-Bertani (LB) medium for bacterial cultivation at 37 °C for 42 h. Gram stain, colony morphology, and biochemical analysis were employed to confirm the identity of the acquired Gram-positive *N. farcinica* (Fig. 1a). The strain was cultured on blood agar, and after 3 days of incubation at 37 °C, sufficient growth was observed. Characteristic dry, chalky, gray, and wrinkled colonies appeared on the blood agar (Fig. 1b).

Antibiotic susceptibility test

The antibiotic sensitivity test was performed in Orbito Asia Diagnostics, Coimbatore, Tamil Nadu, India using the Kirby–Bauer disc diffusion method on cation-adjusted LB agar plates. The inoculums *of N. farcinica* were prepared in accordance with Clinical and Laboratory Standards Institute (CLSI) standards. The antibiotic disc used, include vancomycin (VA), tobramycin (TOB), tetracyclines (TE), teicoplanin (TEI), spectinomycin (SPT), amikacin (AK), ofloxacin (OF), fusidic acid (FC), carbenicillin (CB), imipenem (IPM), ciprofloxacin (CIP), levofloxacin (LE), netillin (NET), polymyxin-B (PB), rifampicin (RIF), streptomycin (S), gentamicin (GEN), kanamycin (K), penicillin-G (P), norfloxacin (NX), nalidixic Acid (NA), erythromycin (E), clindamycin (CD), and chloramphenicol (C) was procured from Himedia. The results were observed after 72 h of incubation. The diameter of the inhibition zone was measured for each antibiotic disc and compared with thresholds. A 'non-susceptible' isolate was considered resistant (R) and a 'susceptible' isolate was defined as Sensitive (S). All the tests were performed in duplicates (Mozrall et al. 2022).

Genomic DNA extraction and quantification

The genomic DNA extraction and quantification was carried out at Genotypic Technology Pvt. Ltd., Bengaluru-560094, Karnataka, India. *N. farcinica* was grown overnight (37 °C; 42 h) and the DNA from the *Nocardia* cell pellet was extracted using a DNA extraction kit Qiagen DNeasy Blood and Tissue Kit (Cat No. 69506). The cell pellet was re-suspended in Lysozyme procured from Sigma (Cat. No. L7651)

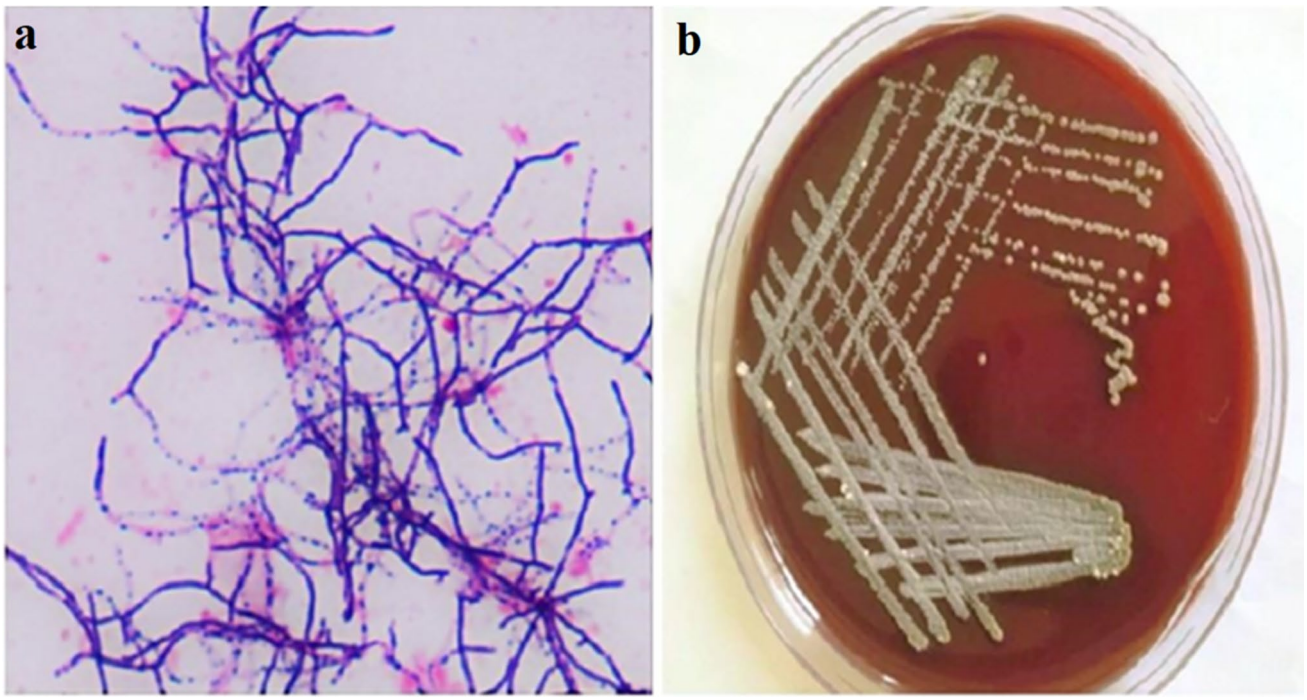


Fig. 1 **a** Gram staining of *N. farcinica* revealed Gram-positive bacteria characterized by rod-shaped morphology and visible branches. **b** Colonial characteristics of *N. farcinica* on blood agar plates include pale *yellow pigmentation*, elevated growth, a rough surface, and opacity (color figure online)

recommends a concentration of 10 mg/ml. The enzyme activity is optimal at 37 °C and for hydrolysis of the peptidoglycan layer, the sample is incubated for 30 mins. The protocol as recommended by Sigma is incorporated. The bacterial suspension treated with AL buffer and Proteinase K was incubated at 56 °C for 2 h followed by RNaseA (Cat. No. 2101076; MP Biomedicals) treatment for 20 min at 65 °C. The lysate was mixed with half the volume of absolute ethanol loaded into the spin column, and placed in a 2 ml collection tube. The tube was centrifuged at 8000 rpm for 1 min, and the flow through was discarded. The remaining wash step was followed according to the manufacturer's protocol. DNA was eluted in 10 mM Tris HCl, pH 8.0. Genomic DNA concentration and purity were measured (Thermo Scientific; 2000) using the Nanodrop Spectrophotometer, and DNA integrity and amount of DNA were analyzed using Agarose gel electrophoresis and Qubit dsDNA HS assay kit (Cat No: Q32854), respectively. Agarose Gel Electrophoresis parameters for DNA integrity analysis using DNA were loaded on 1% gel and electrophoresis was performed at 100 volts. The pure samples with optimal yield and concentration were considered suitable for Illumina and Nanopore library preparation.

Strain purity check

The purity of the bacterial strain was assessed using the 16S rRNA gene with PCR amplification conducted using 30–50 ng of genomic DNA as a template and 16S rDNA primers, 27 Forward (AGAGTTTGATCCTGGCTCAG) and 1492 Reverse (TACGGCTACCTTGTTACGACTT), and Takara ExTaq in a 25 µl reaction mix were used 1.5 kb PCR product was generated, purified, and used for Sanger sequencing. Column-based PCR clean-up kit (Genetix) purification method used for the 1.5 kb PCR product before Sanger sequencing (Naveed et al. 2023).

The PCR Conditions are as follows:

Step 1: Initial Denaturation 98 °C, 2 min
Step 2:Denaturation 98 °C, 20 s
Step 3: Annealing 60 °C, 30 s
Step 4: Extension 72 °C, 30 s
Step 5: Go to step 1, 30 times
Step 6: Final extension 72 °C, 1 min.

Library construction and genome sequencing

Library construction was carried out at Genotypic Technology company using QIASeq FX DNA Library Preparation protocol (Cat#180475) by following the manufacturer's instructions. The libraries were sequenced on Illumina Nova Seq 6000 (Illumina, San Diego, USA) using 150 bp paired-end chemistry following the manufacturer's instructions.

50 ng of Qubit quantified DNA was enzymatically fragmented, end-repaired, and A-tailed in a one-tube reaction using the FX Enzyme Mix provided in the QIASeq FX DNA kit. The end-repaired and adenylated fragments were subjected to adapter ligation, whereby an index-incorporated Illumina adapter was ligated, to generate a sequencing library. The library was subjected to 6 cycles of Index-PCR (Initial Denaturation at 98 °C for 2 min, cycling (at 98 °C for 20 s, 60 °C for 30 s, 72 °C for 30 s) and final extension at (72 °C for 1 min) to enrich the adapter-tagged fragments. Finally, the amplified library was purified using JetSeq Beads (Bio, # 68031) followed by a library quality control check. Illumina-compatible sequencing library was quantified by Qubit fluorometer (Thermo Fisher Scientific, MA, USA) and its fragment size distribution was analyzed on Agilent 2200 TapeStation.

Genome sequencing, assembly, and annotation

The obtained genomic DNA of *N. farcinica* was sequenced on Illumina Nova Seq 6000 using 150 bp paired-end chemistry following the manufacturer's instructions (Cortese et al. 2021). A total of ~7.1 million Illumina sequencing data was

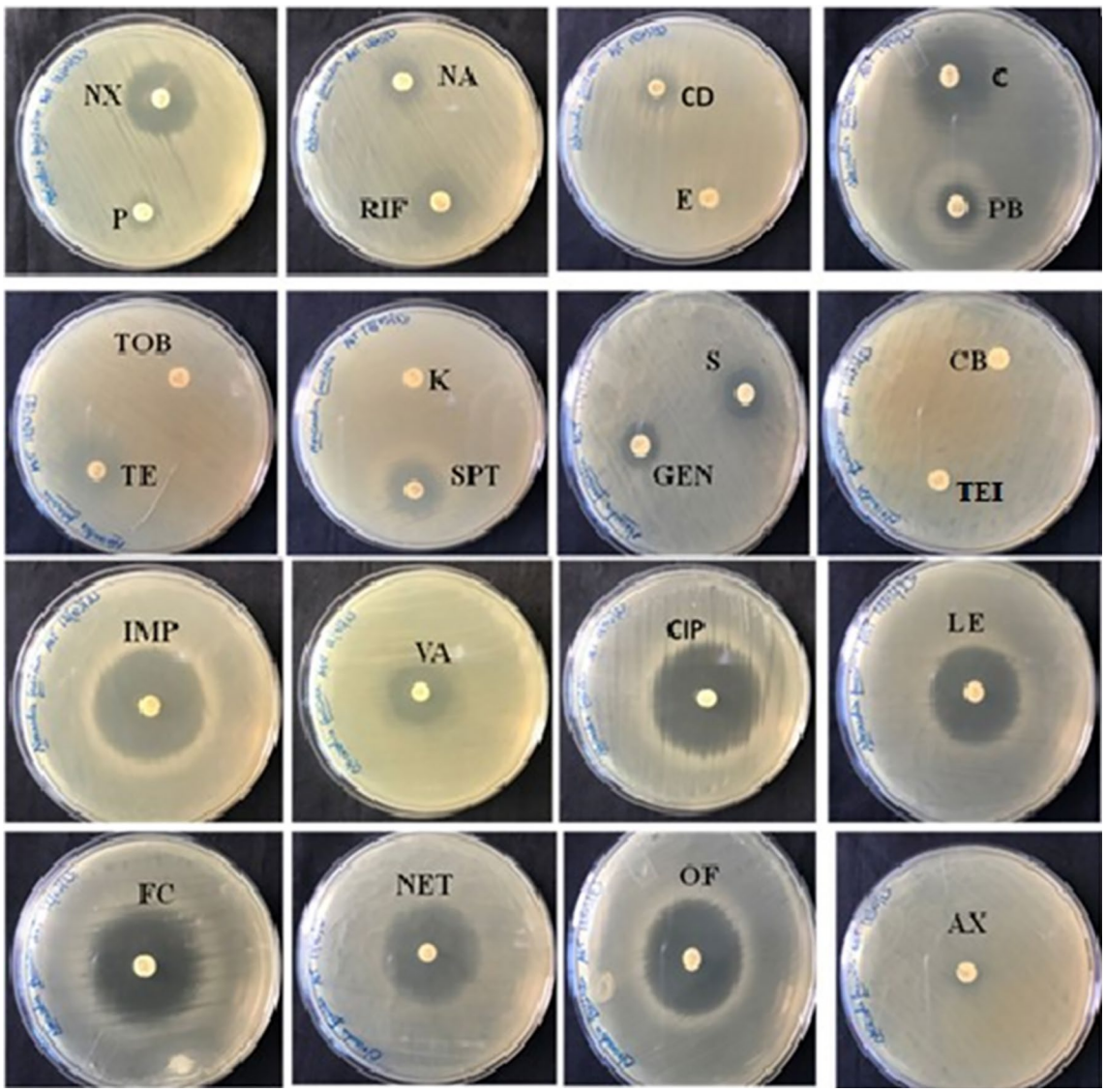


Fig. 2 Antibiotic susceptibility testing of *N. farcinica* with disc diffusion methods. *N. farcinica* antibiotic susceptibility: disc diffusion analysis reveals distinctive zones of inhibition. The study unveils distinctive zones of inhibition, highlighting the bacterium's response to various antibiotics

generated for the *Nocardia* sample. The obtained raw fastq files were subjected to adapters removal using fastp (version 0.23.0). Further, the quality of raw reads was examined through the FastQC tool. Further, Unicycler (version 0.5.0) was considered to identify the good quality reads with the *de novo* assembled short-read data, based on the default settings (Mozrall et al. 2022; Juraschek et al. 2021; Irfan et al. 2023). The purpose of utilizing the unicycler (0.5.0) is to filter out contigs of lesser deep, yielding clean assemblies even if the read set has a low level of contamination (https://github.com/rrwick/Unicycler). The assembled genome was evaluated using the Quality Assessment Tool for Genome Assemblies (QUAST) software employing a default parameter setting of minimum contig length which was set at 500 (Mozrall et al. 2022). QUAST generates a complete set of metrics and statistics to evaluate the accuracy of a genome assembly. This provides different statistics such as N50, L50, contigs, Scaffolds, and total length of the assembly that help the researchers to decipher the complete assembly (Sharma et al. 2023). The Rapid Annotation using Subsystem Technology (RAST) server (version 2.0) and Prokka (Prokaryotic Genome Annotation) software were utilized to annotate the

Table 1 Antibiotic susceptibility test and interpretation thresholds

S. no	Antibiotics	Dosage	Zone of Inhibition (mm)	Susceptible (S)	Resistance (R)	Interpretation threshold (mm)*	
						S	R
1	Vancomycin VA	30 mcg	19.5 ± 0.7	S	–	≥17	–
2	Norflaxacin NX	10 mcg	22.5 ± 0.7	S	–	≥17	≤12
3	Chloramphenicol C	30 mcg	20 ± 1.4	S	–	≥18	≤12
4	Spectinomycin SPT	100 mcg	13.5 ± 2.1	S	–	>14	<10
5	Ofloxacin OF	5 mcg	31 ± 1.4	S	–	≥16	≤12
6	Netillin (Netilmicin sulfate) NET	30 mcg	32 ± 1.4	S	–	≥15	>12
7	Levofloxacin LE	5 mcg	30 ± 1.4	S	–	≥20	<17
8	Fusidic acid FC	10 mcg	34 ± 1.4	S	–	<22	>17
9	Imipenem IPM	10 mcg	33 ± 1.4	S	–	≥24	<17
10	Ciprofloxacin CIP	5 mcg	37.5 ± 0.7	S	–	≥25	<22
11	Polymyxin-B PB	300 units	12 ± 1.4	S	–	≥11	>9
12	Rifampicin RIF	5 mcg	No zone	–	R	>20	<16
13	Streptomycin S	10 mcg	No zone	–	R	≥22	≤14
14	Gentamicin GEN	10 mcg	No zone	–	R	≥15	≤10
15	Kanamycin K	30 mcg	No zone	–	R	≥24	≤20
16	Penicillin-G P	10 units	No zone	–	R	>29	<28
17	Tobramycin TOB	10 mcg	No zone	–	R	≥18	<16
18	Tetracyclines TE	30 mcg	No zone	–	R	≥19	≤14
19	Teicoplanin TEI	30 mcg	No zone	–	R	≥17	≤14
20	Nalidixic Acid NA	30 mcg	No zone	–	R	≥19	≤13
21	Erythromycin E	15 mcg	No zone	–	R	≥22	<17
22	Clindamycin CD	2 mcg	No zone	–	R	>21	<14
23	Carbenicillin CB	100 mcg	No zone	–	R	>15	<11
24	Amikacin AK	30 mcg	No zone	–	R	≥17	<15

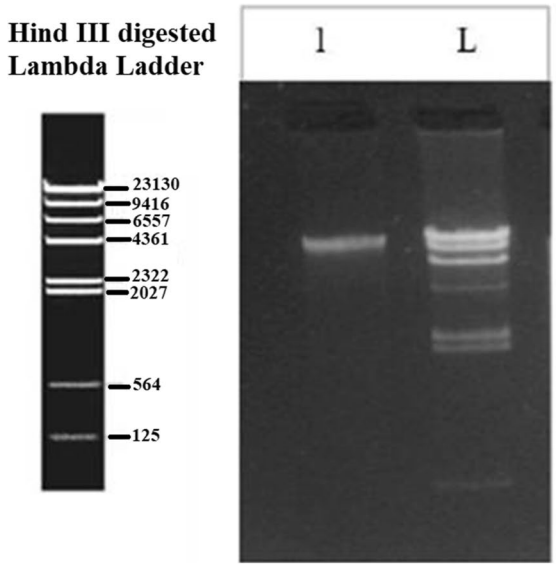
*Deciphering antibiotic susceptibility: unveiling interpretation criteria and thresholds through disc diffusion method

S susceptible, *R* resistance

Table 2 Estimated DNA concentration and purity

S. no	Sample ID	Nanodrop QC				Qubit QC		
		ng/µl	260/280	260/230	Yield (ng)	Qubit Conc. (ng/µl)	Volume (µl)	Yield (ng)
1	SO_10894_N1	71.9	1.72	2.21	1438	22.8	20	456

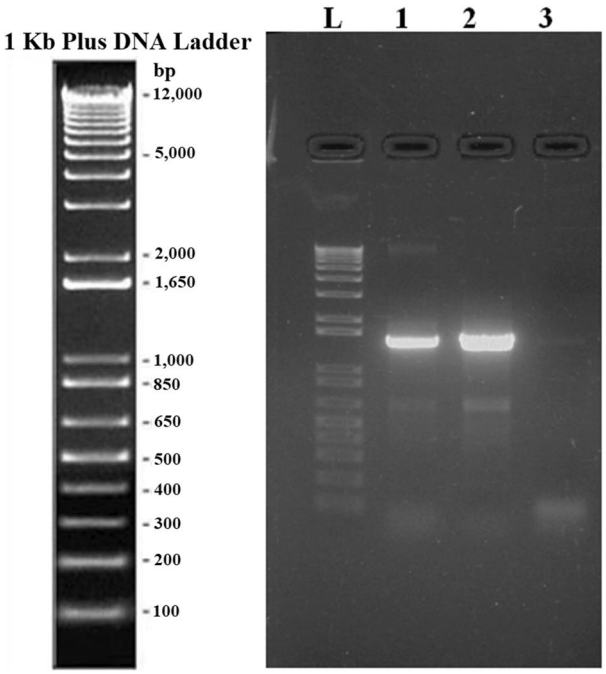
Estimating concentration and assessing purity with nanodrop spectrophotometer analysis



LANES	SAMPLES
L	Hind III digested Lambda Ladder
1	SO_10894_N1

Fig. 3 Agarose gel electrophoresis (AGE) of the *N. farcinica* DNA sample. Illustration of the *N. farcinica* DNA using agarose gel electrophoresis (AGE), with 'I' denoting the sample lane and 'L' representing the molecular weight ladder

assembled *N. farcinica* genome (Madaha et al. 2020; Zhao et al. 2020). The RAST (https://rast.nmpdr.org/rast.cgi) are fully automated with default parameters and Prokka (https://github.com/tseemann/prokka) parameters with default settings such as similarity *e* value cut-off ('1$e-09$'), minimum coverage on query protein (80), minimum contig size (1) (Aziz et al. 2008; Overbeek et al 2014; Brettin et al. 2015; Pei et al 2021; Seemann 2014). The antimicrobial resistance genes and virulent factors were predicted using the ABRicate (version 1.0.1) program with default parameters settings (Seemann 2022) (https://github.com/tseemann/abricate). ABRicate is a software widely used for mass screening of contigs to identify antimicrobial resistance and virulence genes. This software encompasses NCBI, Resfinder, CARD (comprehensive antibiotic resistance database), and Virulence Factor DataBase (VFDB) (Zakaria et al. 2021). The presence of plasmids was computationally predicted through the PLSDB server based on the default parameters of the search strategy of mash screen, maximum *p* value 0.1 and minimum identity 0.99 (https://ccb-microbe.cs.



LANES	SAMPLE ID	Amplicon Observed (bp)
L	1 Kb + Ladder	-
1	SO_10894_NI	~ 1500 bp
2	Positive Control	~ 1500 bp
3	No Template Control	-

Fig. 4 Amplification of ~1500 bp observed for sample. Gel visualization of ~1500 bp 16S rRNA amplicon bands in *N. farcinica.* This image captures the distinct bands corresponding to approximately 1500 base pairs, representing the amplified 16S rRNA in *N. farcinica*

CACGTGGGTGATCTGCCCTGTACTTCGGGATAAGCCTGGGAAACTGGGTCTAATACCGGATATGACCTTACATCGCAT
GGTGTTTGGTGGAAAGATTTATCGGTACAGGATGGGCCCGCGGCCTATCAGCTTGTTGGTGGGGTAATGGCCTACCAA
GGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGC
AGCAGTGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATGACGGCCTTCGGGTTGTA
AACCTCTTTCGACAGGGACGAAGCGCAAGTGACGGTACCTGTAGAAGAAGCACCGGCCAACTACGTGCCAGCAGCCGC
GGTAATACGTAGGGTGCGAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTTGTAGGCGGTTTGTCGCGTCGTCCGTG
AAAACTTGGGGCTCAACCCCAAGCTTGCGGGCGATACGGGCAGACTTGAGTACTGCAGGGGAGACTGGAATTCCTGGT
GTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGGTCTCTGGGCAGTAACTGACGCTGAGAA
GCGAAAGCGTGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGCGCTAGGTGTGGGTTTC
CTTCCACGGGATCCGTGCCGTAGCTAACGCATTAAGCGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAG
GAATTGACGGGGGCCCGCACAAGCGGCGGAGCATGTGGATTAATTCGATGCAACGCGAAGAACCTTACCTGGGTTTGA
CATACACCGGAAACCTGCAGAGATGTAGGCCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTG
TCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCTGTGTTGCCAGCGCGTTATGGCGGGGACTCGC
AGGAGACTGCCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTCCAGGGCTTCACA
CATGCTACAATGGCCGGTACAGAGGGCTGCGATACCGTGAGGTGGAGCGAATCCCTTAAAGCCGGTCTCAGTTCGGAT
CGGGGTCTGCAACTCGACCCCGTGAAGTTGGAGTCGCTAGTAATCGCAGATCAGCAACGCTGCGGTGAATACGTTCCC
GGGCCTTGTACACACCGCCCGTCACGTCATGAAAGTCGGTAACACCCGAAGCCGGTGGCCTAACCCCTTGTGGGAGGG
AGCCGTCGAA

Fig. 5 Identification of 16 s rRNA gene sequencing. Visualization of the 16S rRNA sequence of *Nocardia farcinica*

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Nocardia farcinica strain ARSS8 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1516	MT317187.1
☑	Nocardia farcinica strain M-5 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1431	MT256155.1
☑	Nocardia farcinica strain ARS8 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1516	MN100049.1
☑	Nocardia farcinica strain W6977 chromosome, complete genome	Nocardia farcinica	2468	7404	100%	0.0	100.00%	6291633	CP031418.1
☑	Nocardia farcinica genome assembly NCTC11134, chromosome : 1	Nocardia farcinica	2468	7404	100%	0.0	100.00%	3648762	LN868938.1
☑	Nocardia farcinica strain DSM 43298 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1411	MW563849.1
☑	Nocardia farcinica strain DSM 46005 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1411	MW563848.1
☑	Nocardia farcinica strain DSM 43244 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1412	MW563844.1
☑	Nocardia farcinica strain DSM 43003 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1416	MW563843.1
☑	Nocardia farcinica strain 169 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1394	MW563805.1
☑	Nocardia farcinica strain SZ 1509 chromosome, complete genome	Nocardia farcinica	2468	7404	100%	0.0	100.00%	6229422	CP086599.1
☑	Nocardia farcinica strain DSM 43665 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1419	OK103762.1
☑	Nocardia farcinica strain PCM 2712 16S ribosomal RNA gene, partial sequence	Nocardia farcinica	2468	2468	100%	0.0	100.00%	1457	KC478309.1

Fig. 6 Sequence identity analysis using 16 s rRNA sequences of *N. farcinica*. The BLAST result of the 16S ribosomal RNA of *Nocardia farcinica* sequence shows 100% identity with *N. farcinica*. This significant result underscores the high level of sequence similarity, suggesting a robust association and potential taxonomic association. The bacterial strain that appeared at the BLAST analysis results with maximum query coverage and percent identity is identified as the strain

unisaarland.de/plsdb/plasmids/search_form/seq/) and the sequences were visualized using the proksee tool (https://proksee.ca) (Tian et al. 2022). PLSDB is a database that contains 13,789 plasmid records that have been collected from the NCBI nucleotide database. The PSLDB is widely utilized for the prediction of plasmids within the bacterial genome assemblies (Galata et al. 2019). In Proksee bacterial assembled and annotated genomes from different third-party tools are visualized in one graphical map. Here, plasmids can also be visualized utilizing the proksee server (Grant et al. 2023). Moreover, DNA–DNA hybridization analysis was performed using type strain genome server (TYGS) based on the default settings (https://tygs.dsmz.de/user_requests/new). The rationale behind performing DNA–DNA hybridization analysis using the type TYGS is that allowed for the genome-based replacement of taxonomic graded techniques including DNA: DNA hybridization (DDH), 16S rRNA gene sequencing, G + C-content measurement, and multi-locus sequence analysis. This server creates digital DNA: DNA hybridization values, which are utilized to identify the degree of genetic resemblance between two bacterial strains. This approach is utilized for identifying new

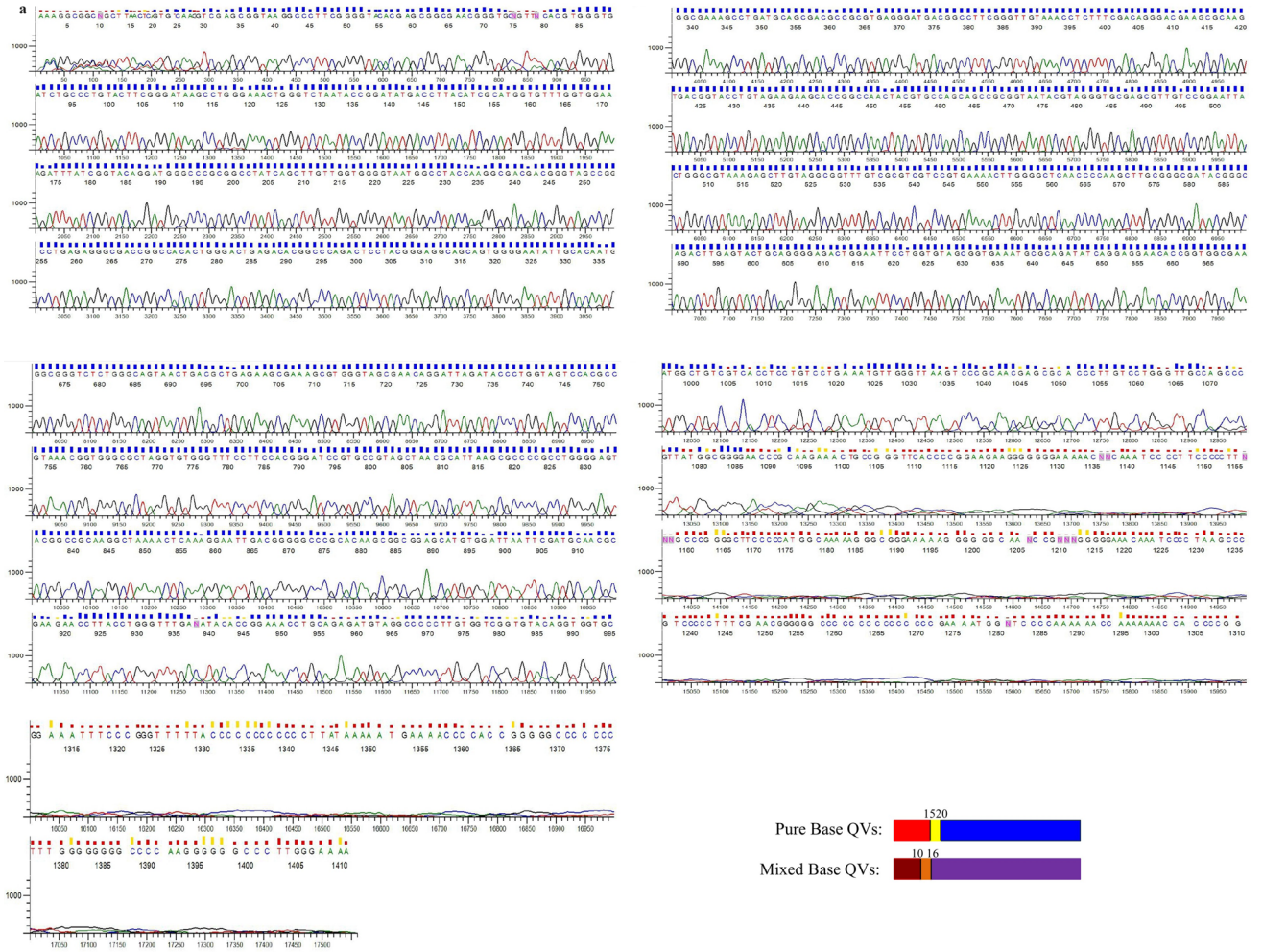


Fig. 7 **a**, **b** Representation of a chromatogram depicting the outcomes of 16S rRNA forward and reverse sequencing in the *N. farcinica* strain, offering insights into the genetic composition. Peaks and patterns in the chromatogram illustrate the sequencing results, contributing to the molecular characterization of the bacterial strain. *Red*, *yellow*, and *blue* indicated the pure base; whereas, brown, orange, and violet indicated the mixed base (color figure online)

bacterial species and for identifying the degree of genetic resemblance between two bacterial strains (Meier-Kolthoff and Göker 2019). The average nucleotide identity score was calculated using the FastANI tool in the proksee server. The Analysis of Average Nucleotide Identity (ANI) is a whole-genome similarity metric that facilitates taxonomic analysis at high resolution across thousands of genomes from varied phylogenetic groups. ANI is a reliable and useful metric for determining the association between two genomes. FastANI is more accurate for both complete and draft genomes, and three orders of magnitude faster in comparison to alignment-based methods (Jain et al. 2018). Identification of the prophage sequences was performed in the PHASTER server (https://phaster.ca/). PHASTER (PHAge Search Tool Enhanced Release) is an important improvement to the popular PHAST web server with default parameters for rapidly finding and annotating the prophage sequences in the genomes of bacteria and plasmids. Three types of prophages are identified by PHASTER: intact, incomplete, and questionable. Prophage prediction confidence levels are reflected in these categories. Here, questionable prophages have less confidence, therefore intact prophages are more confidently predicted as complete. These are the criteria used for prophage sequence identification in the PHASTER server (Arndt et al.2016; Zhou et al. 2011). The presence of secondary metabolite gene clusters was detected using the antiSMASH server with default settings (version 7.0.0) (https://antismash.secondarymetabolites.org/#!/start). Gene profiles for particular kinds of gene clusters are stored in hidden markov models (HMMs). Gene clusters are determined based on the co-occurrence of genes involved in secondary metabolism utilizing the ClusterFinder algorithm. The manually curated clusters of biosynthetic genes and related metadata are available in the MIBiG database.

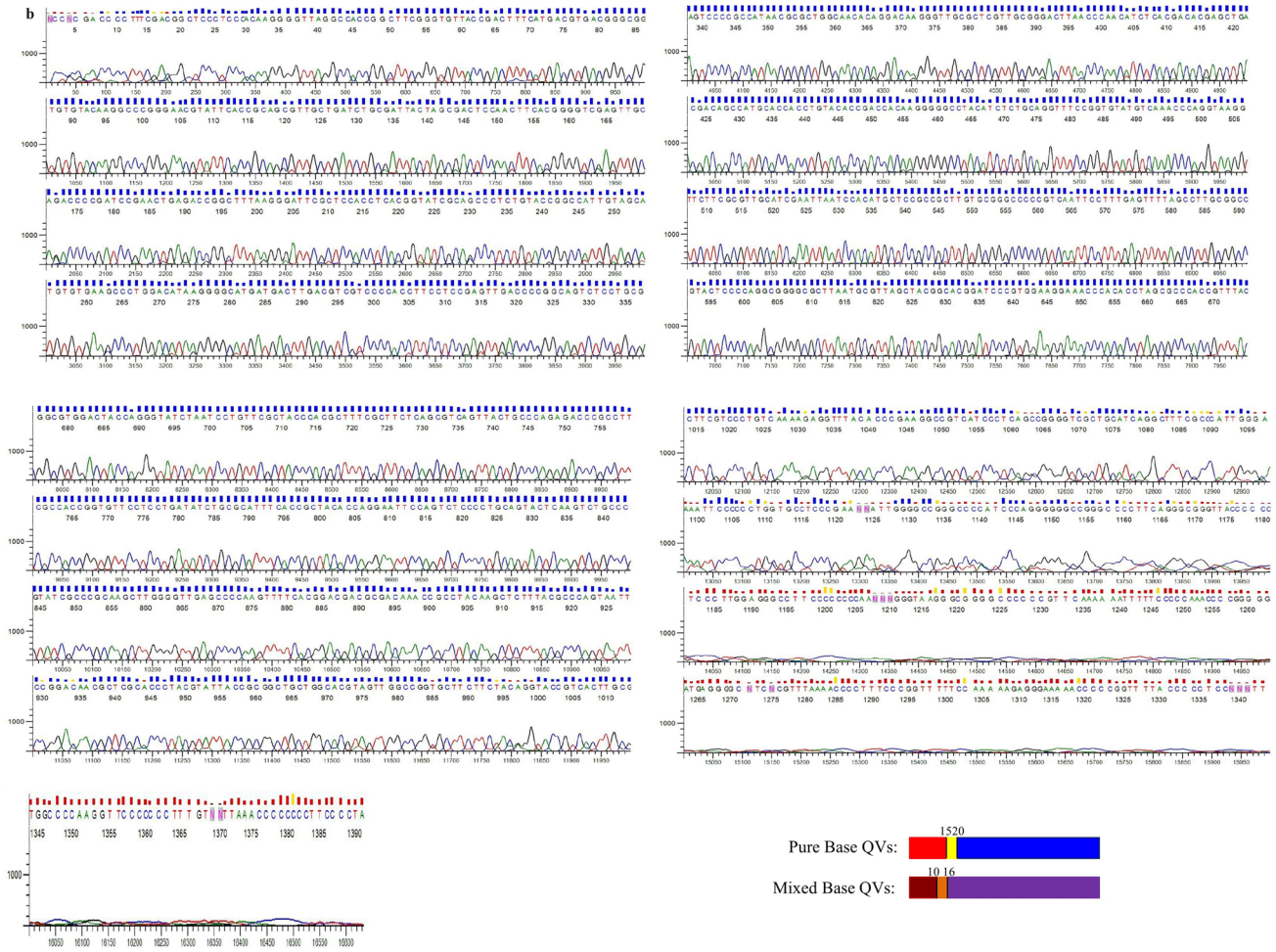


Fig. 7 (continued)

Table 3 Illumina read statistics

Raw reads	Sequencing coverage	Processed reads	Read retained (%)
7,181,520	308X	7,142,784	99.4

Illumina read statistics for a detailed analysis unveiling insights into sequencing quality and depth

The antiSMASH database consists of gene clusters found with antiSMASH version 4 in over 6000 completed bacterial genomes. All publicly available genomes are analyzed using the ClusterBlast algorithm, which identifies similar gene clusters. These are the databases used in antiSMASH for detecting secondary metabolite gene clusters in bacterial genomes (Blin et al. 2023).

Results

Antibiotic susceptibility test

The results of the antibiotic susceptibility test with 24 antibiotics for *N. farcinica* are represented in (Fig. 2). *N. farcinica* was found susceptible to 13 antibiotics, namely vancomycin (VA), tobramycin (TOB), tetracyclines (TE),

Table 4 Description of libraries

S. no	Sample ID	Qubit conc. (ng/µl)	Vol (µl)	Yield (ng)	Index 1	Index 1 sequence	Index 2	Index 2 sequence
1	SO_10894_N1	21.4	10	214	D710	TCCGCGAA	D501	TATAGCCT

Illumina sequencing exposing statistical insights into *Nocardia farcinica* libraries

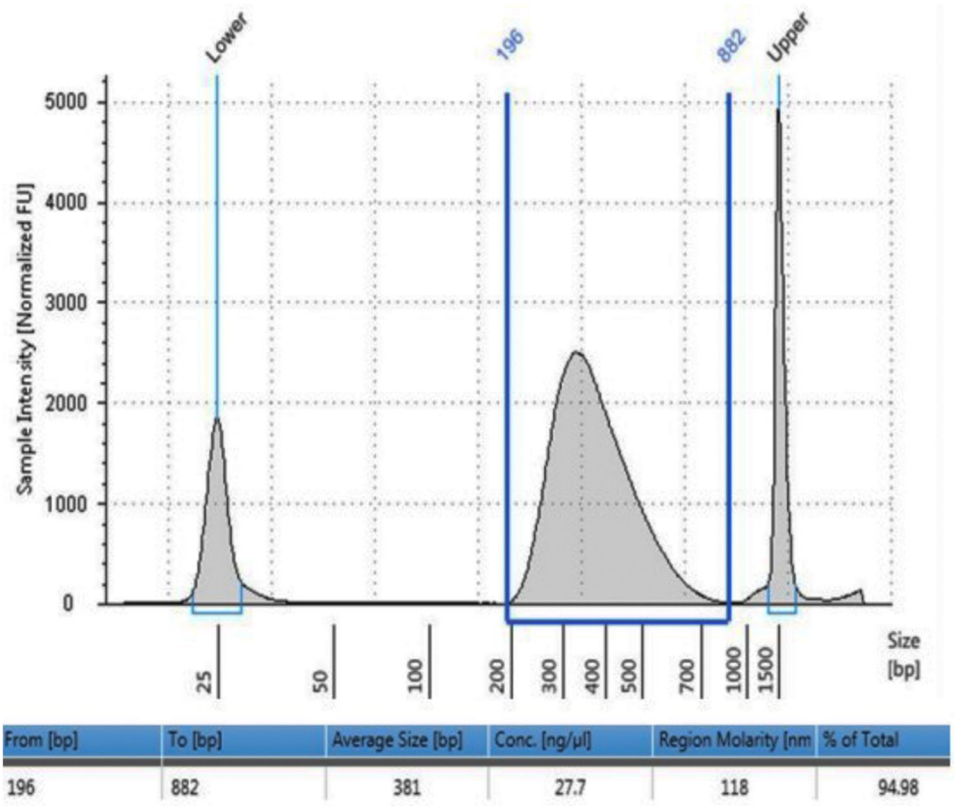


Fig. 8 TapeStation profile of a sequencing library. Exploring genomic libraries is achieved through TapeStation profiling of a sequencing library. The investigation of genomic libraries using TapeStation profiling provides valuable insights into library composition and quality

teicoplanin (TEI), spectinomycin (SPT), amikacin (AK), ofloxacin (OF), fusidic acid (FC), carbenicillin (CB), imipenem (IPM), ciprofloxacin (CIP), levofloxacin (LE), netillin (NET) and found resistant to 11 antibiotics, namely polymyxin-B (PB), rifampicin (RIF), streptomycin (S), gentamicin (GEN), kanamycin (K), penicillin-G (P), norfloxacin (NX), nalidixic acid (NA), erythromycin (E), clindamycin (CD), and chloramphenicol (C). The resistance and susceptibility patterns of *N. farcinica* are represented in (Table 1).

Genomic DNA extraction and quantification

The samples that passed the quality assessment with optimal yield and concentration were deemed suitable for Illumina and Nanopore library preparation (Table 2 and Fig. 3)

Strain purity check

The 16s rRNA sequence was subjected to a BLAST search against the nucleotide collection (nr/nt) database for analyzing the purity of the strain. The BLAST results show a percentage identity value and higher total score of 100% and 7404, respectively, which indicates that the sample was a pure resemblance match to a known bacterial species *N. farcinica* with high similarity. Amplification of 16S rRNA for strain purity test by Sanger sequencing (Fig. 4).

Sequence obtained from forward and reverse primer is used to generate contig for which BLAST analysis (NCBI) is performed against nr/nt database. The bacterial strain that appeared at the top of the BLAST analysis results with maximum query coverage and percent identity is identified as the strain (Figs. 5 and 6). Chromatogram for 16 s rRNA forward and reverse sequencing (Fig. 7a, b).

Library construction and genome sequencing

Thus, 7181520 raw readings were found, and its GC content of 70.7%. The Illumina data generated for the sample had ~7.1 million reads with a sequencing coverage of 308X (Table 3).

The Illumina-compatible sequencing library for the samples showed an average fragment size of 381 bp as well as sufficient concentration for obtaining desired sequencing data. Lists the concentration of libraries obtained and indices used. TapeStation profile of a sequencing library (Table 4 and Fig. 8).

Adapter was removed by fastp and evaluated by QUAST (Quality Assessment Tool for Genome Assemblies)

Automatic adapter fastp trims both single-end and paired-end Illumina data in which each pair of sequences is

Table 5 Overview of genomic assembly statistics

	Assembly
# contigs (≥0 bp)	103
# contigs (≥1000 bp)	63
# contigs (≥5000 bp)	46
# contigs (≥10,000 bp)	36
# contigs (≥25,000 bp)	28
# contigs (≥50,000 bp)	20
Total length (≥0 bp)	6,131,470
Total length (≥1000 bp)	6,118,221
Total length (≥5000 bp)	6,082,203
Total length (≥10,000 bp)	6,011,357
Total length (≥25,000 bp)	5,861,903
Total length (≥50,000 bp)	5,579,449
# contigs	72
Largest contig	997,868
Total length	6,123,581
GC (%)	70.78
N50	292,531
N90	58,441
auN	503,033.6
L50	5
L90	19
# N's per 100 kbp	0.00

All statistics are based on contigs of size ≥500 bp unless otherwise noted (e.g., "# contigs (≥0 bp)" and "Total length (≥0 bp)" include all contigs)

Provides a detailed analysis of genomic assembly statistics, offering insights into genome structure. The assembly includes 103 contigs (≥0 bp) with varying lengths. Key metrics at different length thresholds (≥1000, ≥5000, ≥10,000, ≥25,000, and ≥ 50,000 bp) elucidate assembly continuity, while N50, GC content (%), and the largest contig size contribute to understanding quality and completeness. L50 and N90 parameters further characterize assembly continuity at different scales

identified based on their overlap (Supplementary 1 and 2). Subsequently, a draft genome assembly metrics analyzed using QUAST showed a total length of 6,123,581 bp, contigs of 103, N50 of 29,253 bp, GC content of 70.78%, and 63 contigs greater than 1000 bp (Table 5 and Fig. 9).

Annotation results by rapid annotation subsystems technology RAST and Prokka

The genome annotations were predicted by Prokka and RAST. According to prokka, was found 6,131,470 bp, 5683 protein coding sequences, 4 rRNA, 7 repeat regions, 60 tRNA, and 1 tmRNA. Rapid Annotations Subsystems Technology found that 5945 protein-coding sequences belonged to 302 subsystems and 54 rRNA (Table 6). The functional analysis obtained from RAST (Fig. 10) revealed that the genome had

a) cofactors, vitamins, prosthetic groups, and pigments: 185 genes
b) cell wall and capsule: 24 genes
c) Virulence, Disease, and Defense: 58 genes
d) potassium metabolism: 8 genes
e) Miscellaneous: 34 genes
f) membrane transport:36 genes
g) acquisition and metabolism: 7 genes
h) RNA metabolism: 55 genes
i) Nucleosides and Nucleotides:79 genes
j) Protein metabolism: 186 genes
k) Regulation and cell signaling: 16 genes
l) Secondary metabolism: 5 genes
m) DNA metabolism: 96 genes
n) Fatty acids, lipids, and isoprenoids: 203 genes
o) Nitrogen metabolism: 23 genes
p) Dormancy and sporulation: 1 gene
q) Respiration: 114 genes
r) Stress response: 49 genes
s) Metabolism of aromatic compounds: 62 genes
t) Amino acid and derivatives: 364 genes
u) Sulfur metabolism: 13 genes
v) Phosphorus metabolism: 24 genes
w) Carbohydrates248 genes

Identification of antimicrobial resistance and virulence gene

CARD (Comprehensive Antibiotic Resistance Database) analysis for antimicrobial resistance genes in *N. farcinica* of the genome, results in the presence of rifamycin resistance (RbpA gene), macrolide; penam resistance (mtrA gene), and penam resistance (FAR-1 gene). Resfinder-predicted resistance genes encoding blaFAR-1 resistance (blaFAR-1_1). NCBI-predicted resistance genes responsible for rifampin resistance (rox) and beta-lactam resistance (blaFAR-1). Virulent factors such as relA, icl, and mbtH genes were found in VFDB. The presence of both virulent and resistant genes in *N. farcinica's* genome (Table 7 and Fig. 11).

Plasmid identification and genetic analysis of plasmids

The PLSDB server v2.1.1 was used to detect the contigs that could belong to a plasmid. The results showed that *N. farcinica* had a chromosome of 6.1mb and one predicted plasmid (Accession: NZ_LN868939.1) with a sequence length

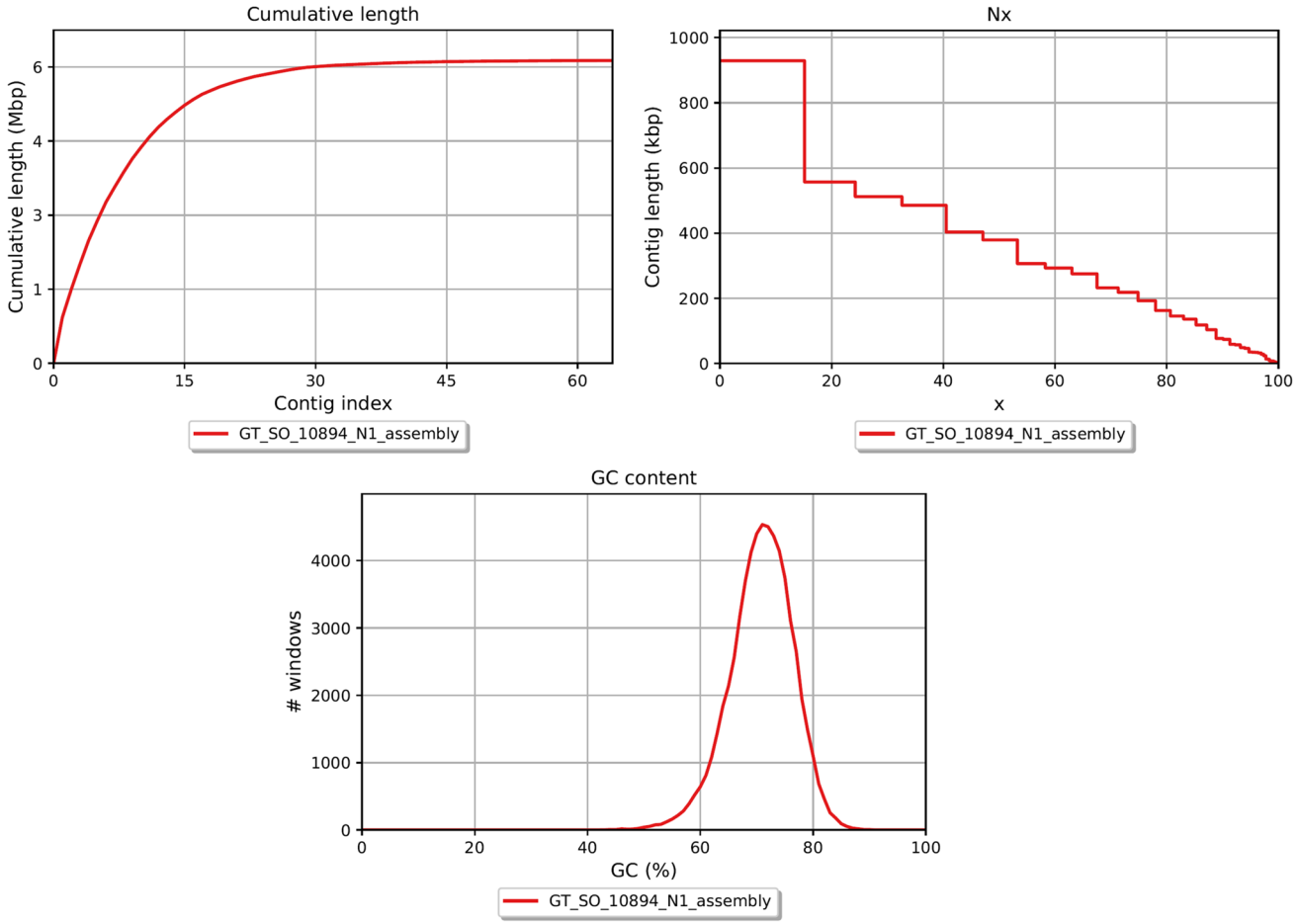


Fig. 9 A Comprehensive genomic evaluation through QUAST analysis of *N. farcinica*. Represents a detailed genomic assessment using QUAST analysis, providing accurate and detailed insights into the structure and quality of the *N. farcinica* genome

is 2,657,929 bp, respectively (Table 8). Proksee server was utilized to visualize the plasmid maps of the *N. farcinica* (Fig. 12).

Prophage regions in the genome

Three prophages were identified in the chromosome through the PHASTER tool and the obtained results are shown as follows.

1. "> 2 ", "length = 928,982" bps, "depth = 1.08×" 0 intact, o questionable, and 1 incomplete region is presented in the chromosome (Fig. 13). Prophage regions 1 with the start: 596,655, End:604,096, CDs:8, predicted type: incomplete, and GC % contents of 68.23%, respectively. PHASTER tool revealed that the aforesaid regions were incomplete as they scored <70. The gene function of region 1 (incomplete prophage regions), was found to be essential for phage activity while phage-like protein, hypothetical protein, and fiber protein were identified to play a crucial role in the same.
2. "> 3 ", "length = 775,693" bps, "depth = 0.94×" 0 intact, o questionable, and 1 incomplete region is presented in the chromosome (Fig. 14). Prophage regions 1 with the start: 76,665, End:85,470, CDs: 8, predicted type: incomplete, and GC% contents of 66.90%, respectively. PHASTER tool revealed that the aforesaid regions were incomplete as they scored <70. The gene function of region 1 (incomplete prophage regions), was found to be essential for phage activity, while 3 phage-like proteins and 5 hypothetical proteins were identified to play a crucial role in the same.
3. "> 6 ", "length = 274,970" bps, "depth = 1.00×" 0 intact, o questionable, and 1 incomplete region is presented in the chromosome (Fig. 15). Prophage regions 1 with the start:202,280, End:210,331, CDs:10, pre-

Table 6 Decoding *N. farcinica* for a comprehensive overview of genomic insights

Genome	*Nocardia farcinica* (Taxonomy ID: 37,329)
Domain	Bacteria
Taxonomy	Bacteria; Terrabacteria group; Actinobacteria; Actinomycetia; Corynebacteriales; Nocardiaceae; *Nocardia*; *Nocardia farcinica*
Size	6,131,470
GC Content	70.8
N50	292,531
L50	5
Number of contigs (with PEGs)	103
Number of Subsystems	302
Number of coding sequences	5945
Number of RNAs	54

This table provides a detailed snapshot of the genomic profile of *N. farcinica* (Taxonomy ID: 37329). The bacterium belongs to the domain Bacteria and falls within the Terrabacteria group, Actinobacteria phylum, Actinomycetia class, Corynebacteriales order, and Nocardiaceae family. With a genome size of 6,131,470 base pairs and a GC content of 70.8%, *Nocardia farcinica* exhibits distinctive genomic characteristics. The N50 value, indicating the length of the contig at which 50% of the total genome is covered, is 292,531 base pairs, with an associated L50 of 5 contigs. The genome is assembled into 103 contigs, featuring 5945 coding sequences and 54 RNAs. Furthermore, an exploration of 302 subsystems offers insights into the functional diversity of the genome. This comprehensive overview enhances our understanding of the genetic makeup of *Nocardia farcinica,* laying the groundwork for further investigations into its biology, evolution, and potential applications in various fields

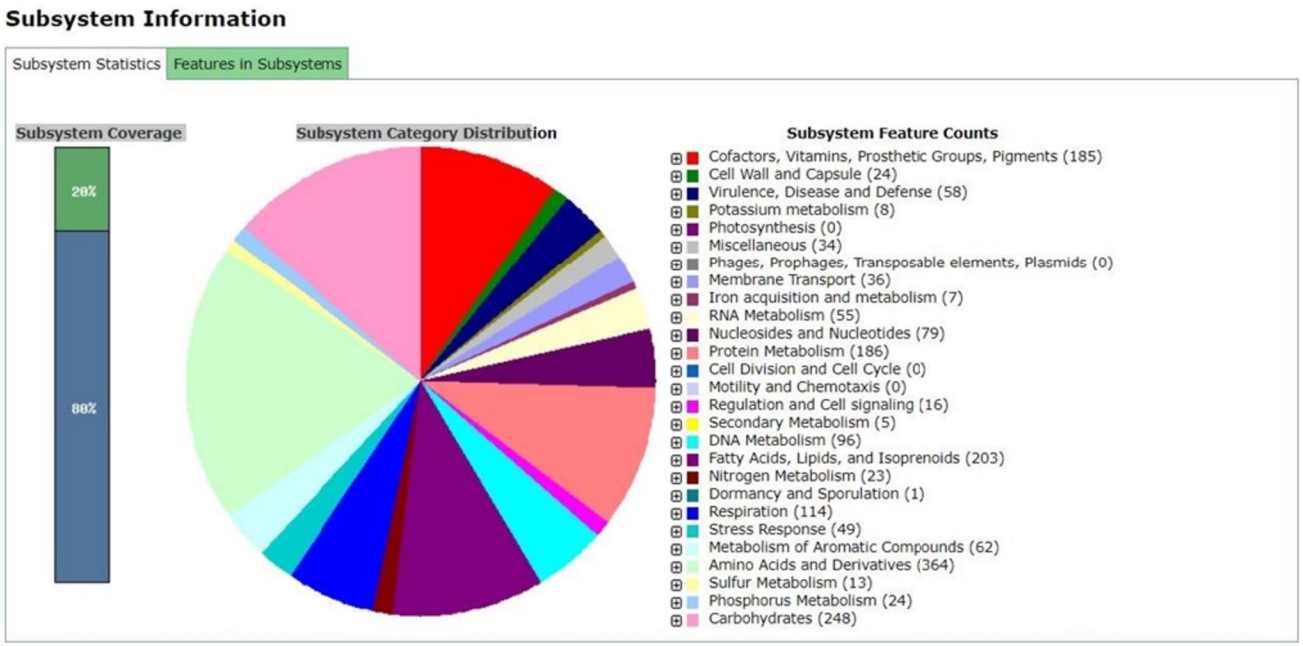


Fig. 10 Subsystem distribution in different categories of *Nocardia farcinica.* Subsystem coverage indicates the total genes in the subsystems. Each part of the pie graph represents different functions and proportions of genes

dicted type: incomplete, and GC% contents of 72.62%, respectively. PHASTER tool revealed that the aforesaid regions were incomplete as they scored <70. The gene function of region 1 (incomplete prophage regions) was found necessary for phage activity and 2 phage-like proteins and 8 Hypothetical proteins were identified to play a crucial role in the same.

Table 7 Exploration of antimicrobial resistance and virulence genes for a comprehensive understanding of microbial defense mechanisms

Seq	Start	End	Strand	Gene	Coverage	Gaps	coverage%	identity%	Database	Accession	Product	Resistance
1	800,102	800,433	–	RbpA	1–332/345	0/0	96.23	82.83	CARD	HQ203032:0–345	RNA-polymerase binding protein which confers resistance to rifampin	Rifamycin
10	157,458	158,138	–	mtrA	7–687/687	2/2	98.98	81.08	CARD	AL123456.3:3,627,349–3,626,662	MtrA is a transcriptional activator of the MtrCDE multidrug efflux pump of Neisseria gonorrhoeae	Macrolide; penam
3	514,063	514,956	+	FAR-1	1–894/894	0/0	100	99.22	CARD	AF024601.1:302–1196	FAR-1 is a class A beta-lactamase gene found in *Nocardiafarcinica*	Penam
1	383,961	385,382	–	Rox	1–1422/1422	0/0	100.00	99.30	NCBI	NG_052045.1	Rifampin monooxygenase Rox	Rifampin
3	514,024	514,956	+	blaFAR-1	1–933/933	0/0	100.00	99.25	NCBI	NG_049089.2	Class A extended-spectrum beta-lactamase FAR-1	Beta-Lactam
3	514,063	514,956	+	BlaFAR-1_1	1–894/894	0/0	100.00	99.22	Resfinder	AF024601	blaFAR-1	-
1	237,530	239,772	-	relA	121–2369/2373	5/26	94.1	80.03	VFDB	NP_217099	(relA) Probable GTP pyrophosphokinase RelA (ATP:GTP 3'-pyrophosphotransferase) (PPGPP synthetase I) ((P)PPGPP synthetase) (GTP diphosphokinase) [RelA (VF0287)] [Mycobacterium tuberculosis H37Rv]	-
2	113,417	114,693	–	icl	12–1285/1287	3/5	98.91	83.8	VFDB	YP_177728	(icl) Isocitrate lyase Icl (isocitrase) (isocitratase) [Isocitrate lyase (VF0253)] [Mycobacterium tuberculosis H37Rv]	-
9	25,075	25,281	+	mbtH	2–208/216	0/0	95.83	80.68	VFDB	NP_216893	(mbtH) putative protein MbtH [Mycobactin (VF0299)] [Mycobacterium tuberculosis H37Rv]	

Detailed overview of antimicrobial resistance and virulence genes in *N. farcinica*, offering key details such as gene location, orientation, coverage, gaps, and associated resistance

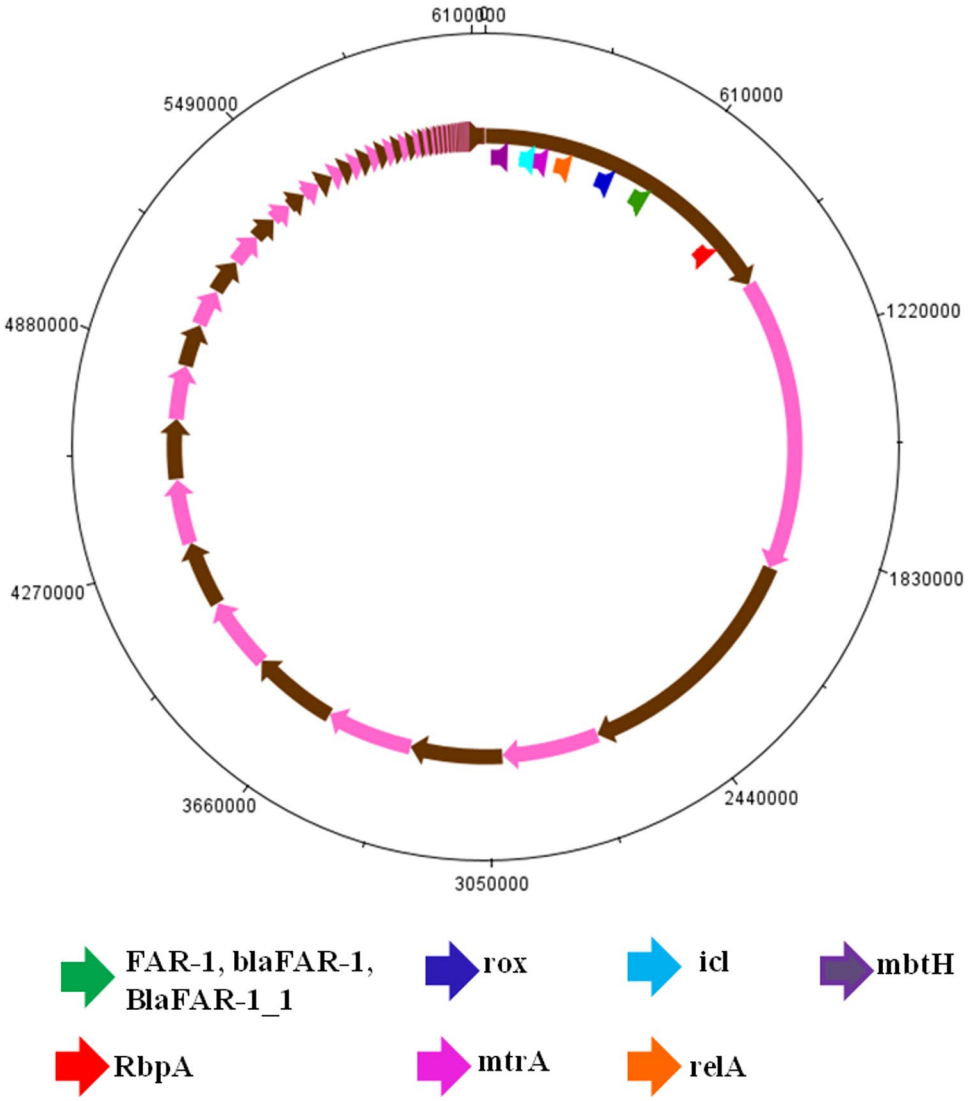


Fig. 11 Distribution of antibiotics resistant and virulent genes in the genome of *N. farcinica*. The distribution pattern of antibiotic-resistant and virulent genes within the genome of *Nocardia farcinica.* The analysis provides valuable insights into the genetic elements contributing to antibiotic resistance and virulence in this bacterium

Annotation of DNA–DNA Hybridization

TYGS was used to infer the taxonomic status within a particular group of the bacterial dataset. Formulae d0 and d6 were used to measure the similarity in gene content and formula d4 reports a similarity based on sequence identity. Based on the whole-genome phylogenetic result it was observed that the draft bacterial genome (assembly. fasta) was closely related to the *N. farcinica NBRC 15532* strain (Fig. 16).

Average nucleotide identity (ANI)

The similarity analysis was carried out between the assembled genome and reference genome (*N. farcinica* NCTC1134, *N. farcinica* DSM43257, *N. farcinica* NBRC 15532) using FastANI which showed the identity of 99.0739, 99.0679, and 99.0454%, respectively, against the reference genomes. The red line segment denotes the reciprocal mapping between the query and reference genome, indicating their evolutionary conserved regions (Fig. 17a–c).

Identification of gene clusters involved in secondary metabolite

Gene clusters involved in the bioactive compound synthesis, type such as 4 Terpene, ranthipeptide, NRP-metallophore, 12 NRPS, 2 NRPS-like, 4 T1PKS, hgIE-KS, ectoine, redox-cofactor, amino polycarboxylic acid, 2 T3PKS, NAPAA, aryl polyene, Ripp-like, furan (Table 9).

Terpene: It has been found that the terpene compounds in *Nocardia* have antibiotic and cytostatic properties. Terpenoids and meroterpenoids produced by actinomycete that have significant antibacterial activity are of interest as

Table 8 Identification of plasmids in *N. farcinica* utilizing PSLDB Server

Plasmid accession	Identity	Shared hashes	Median multiplicity	*p* value	Topology	Length	GC%	Taxon	Kingdom	Phylum	Class	Order	Family	Genus	Species
NZ_LN868939.1	0.990132	812	1	0	circular	2,657,929 bp	70.21%	*Nocardia farcinica*	Bacteria (2)	Actinobacteria (201,174)	Actinomycetia (1760)	Corynebacteriales (85,007)	Nocardiaceae (85,025)	*Nocardia* (1817)	*Nocardia farcinica*(37,329)

The PSLDB Server implies the analysis of plasmids in *N. farcinica* assembly emphasizing efficiency and clarity in the identification process. The overall analysis provides an overview of plasmid identification, circular topology, unique identity, and taxonomic classification within actinobacteria in the specified bacterium

a source of new antibiotics efficient against drug-resistant pathogenic bacteria. It is an important pathway in actinomycetes' secondary metabolism to synthesize terpenes and terpenoids.

NRPS and RIPP-like:

Nocardia genomes can generate a variety of secondary metabolites, siderophores, antibiotics, and other small bioactive molecules due to the high number of non-ribosomal peptide synthases (NRPS) in each strain.

RiPPs (ribosomal synthesized and post-translationally modified peptides) and NRPS (ribosomal synthesized and post-translationally modified peptides) are produced by *Nocardia* strains.

Aryl polyene:

Any potential damage that might result as a consequence of exposure to reactive oxygen species is protected by the aryl propene.

Discussion

Nocardia, as a facultative intracellular pathogen infecting the brains and lungs of immunosuppressed patients with consequences that could be at risk of mortality. *Nocardia* is difficult to identify because of its prolonged incubation time and it is undiagnosed, unrecognized, and neglected because of its non-specific symptoms of infection. *Nocardia* can enter and survive within host cells, including macrophages and epithelial cells, and resist the host immune system by generating several virulence factors, including hemolysin and superoxide dismutase (Ji et al. 2022). The use of potent antibiotics to treat *N. farcinica* leads to adverse side effects and paves the way for drug resistance. There is an urgent need to identify alternative drug targets capable of combating the multi-drug resistance through antibiotics. This research aims to design potent inhibitors of *N. farcinica* infections, minimizing or eliminating side effects. As per the previous report of Ji et al. (2022), the *Nocardia* virulence factors such as mce, hbha, and nfa34810 play a significant role in adhesion and invasion. Complete genome sequence data indicated that the genome of *N. farcinica* consists of many inducible virulence genes, including catalases and nbt, which may play important roles during the infection process (Ji et al. 2022). This study embarked on whole-genome sequencing of *N. farcinica* strains resistant to commonly employed antibiotics. This comprehensive exploration unveiled six pivotal antibiotic resistance genes—RbpA, mtrA, FAR-1, rox, blaFAR-1, and BlaFAR-1_1. The RbpA gene resistance phenotype is Rifampicin or Rifampin. The Rifampicin or Rifampin antibiotic plays a major role in inhibiting DNA-dependent RNA polymerase, leading to the suppression of RNA synthesis and cell death (Newell et al. 2006; Hu et al. 2012; Wang et al. 2020). The FAR-1 is a class A β-lactamase gene and its resistance phenotype is penicillin

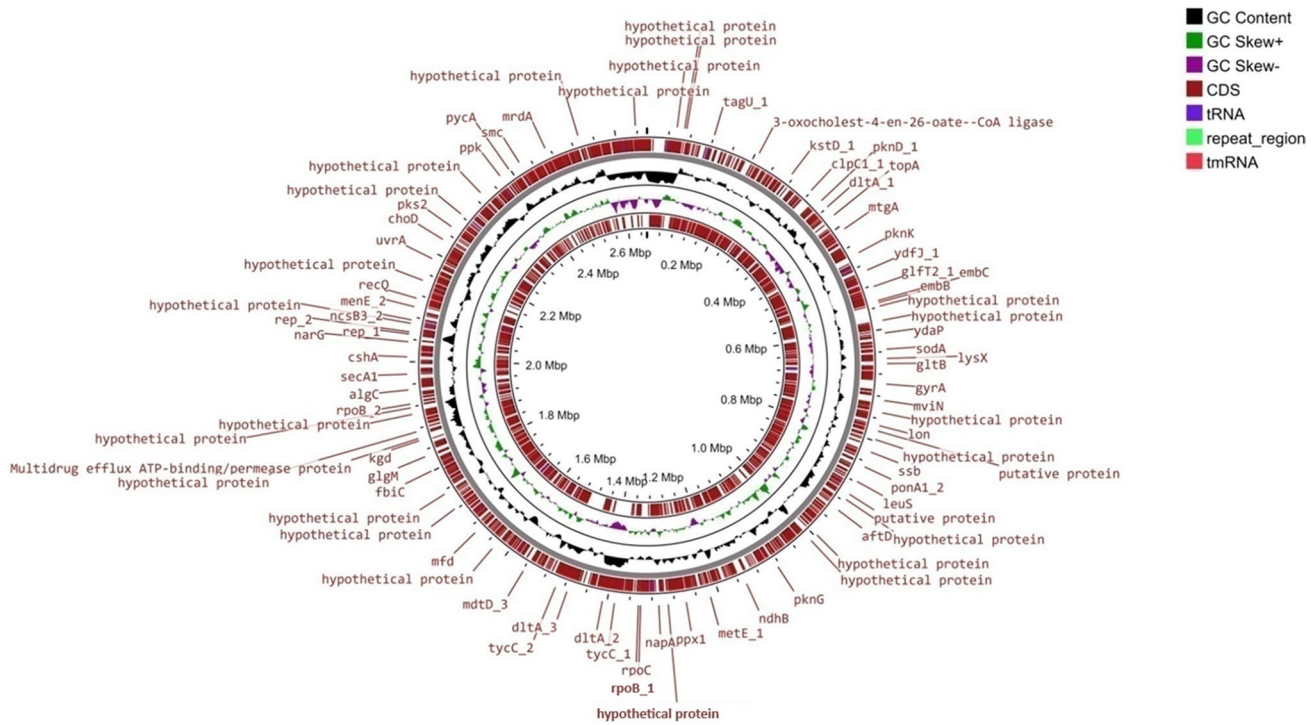


Fig. 12 The plasmid map of *Nocardia farcinica* for a comprehensive exploration using proksee tool. This illustrative figure showcases the plasmid map of *Nocardia farcinica*, generated through the utilization of the Proksee tool. The intricately annotated map provides a detailed depiction of key genetic elements, including coding sequences, structural features, and functional annotations associated with the identified plasmid

antibiotic. All penicillins are beta-lactam antibiotics in the penam sub-group and are used in the treatment of bacterial infections caused by Gram-positive organisms. The resistance mechanism of FAR-1 is antibiotic inactivation (De Pascale and Wright 2010; Wang et al. 2020). The resistance phenotype associated with the mtrA gene is characterized by its ability to resist macrolide antibiotics, particularly erythromycin, through a mechanism involving drug efflux (Wang et al. 2020; Sun et al. 2014). In *N. farcinica*, Rox breaks down Rifampicin, creating a new metabolite and confers resistance to Rifampin by inactivating the antibiotic through a monooxygenase mechanism (Hoshino et al. 2010). The blaFAR-1 gene resistance phenotype is Beta-Lactam. Notably, penicillin resistance was associated with the presence of blaFAR-1 genes, respectively (Wang et al. 2020). Moreover, the whole-genome sequencing of multidrug-resistant *N. farcinica* revealed the existence of three vital virulence factor genes—relA, Icl, and mbtH. In this era of antibiotic resistance, our findings underscored the need to explore novel approaches and strategies to combat AMR. This study bridges the gap between genetic intricacies and therapeutic challenges by elucidating the genetic basis of multi-drug resistance in *N. farcinica*. Additionally, by uncovering significant virulence factors, we gain insights into the pathogen's strategies for thriving. In summary, this research constitutes a pivotal stride toward addressing the pressing need for innovative interventions against *N. farcinica* infections. It explores both drug resistance and virulence factors revealing insights that could lead to potential therapeutic breakthroughs. Empowered with this profound understanding, the goal is to identify potent lead compounds, paving the way to outsmart *N. farcinica's* multi-drug resistance. The significance of the *Nocardia* Species lies in their emergence as human pathogens which closely resemble other mycolic acid-containing genera of the order Actinomycetales, particularly *Mycobacterium tuberculosis*. As a result, misdiagnosis and it is possible for therapies to fail in clinical settings. This research aimed to address these gaps by using a collection of clinical isolates and investigating potential identification approaches, antimicrobial susceptibility patterns, and resistance mechanisms in *Nocardia* species. The outcomes of this research will play an important role in improving the diagnosis and effective treatment of infections caused by *Nocardia* species. Additionally, this research aims to establish clear and standardized criteria for assessing antimicrobial susceptibility, aiding in more precise interpretations.

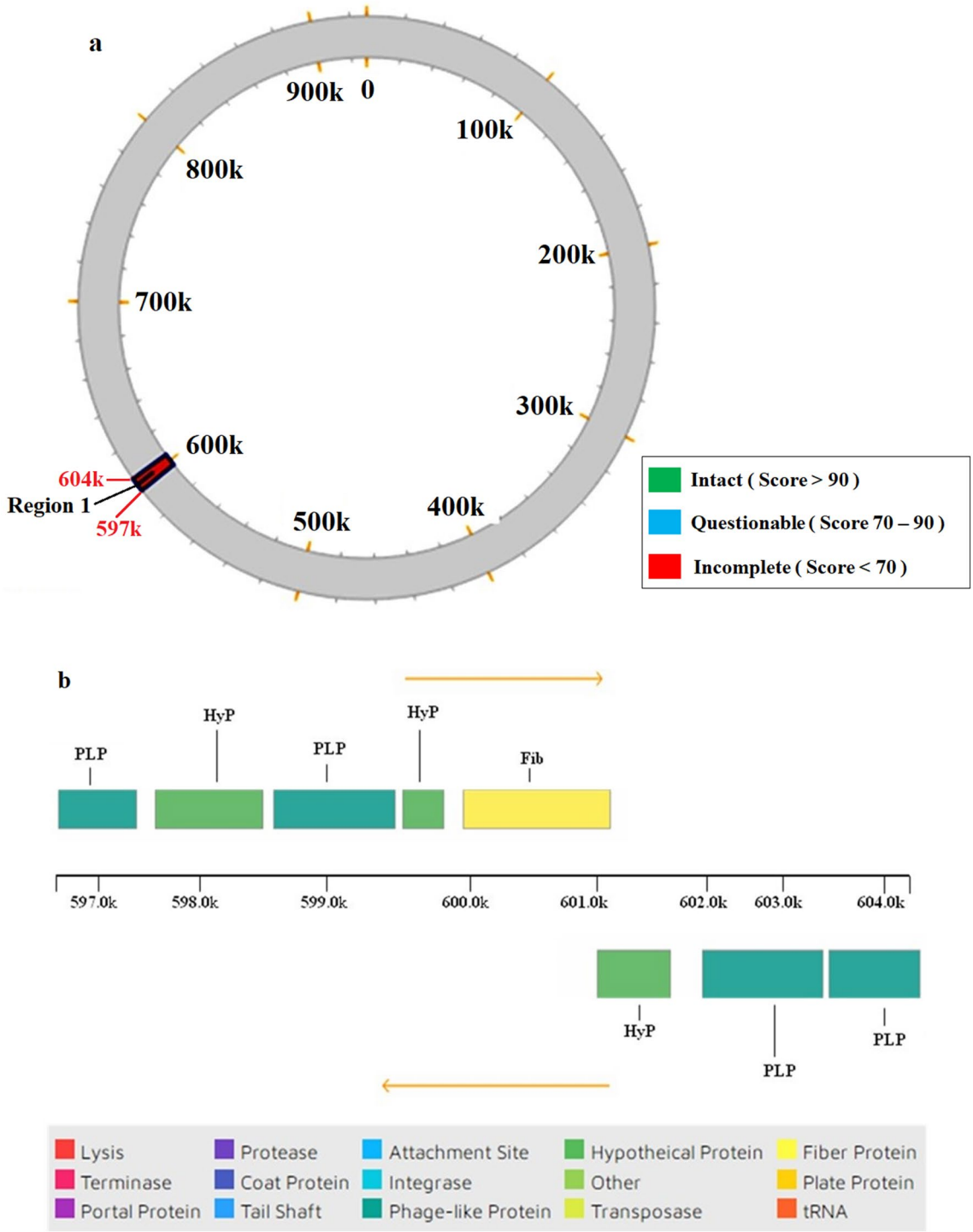


Fig. 13 **a** A total of 1 prophage region is positioned in the chromosome. *Green* indicates the intact prophage regions (score > 90), *blue* represents the questionable prophage regions (score 70–90), and *red* specifies the incomplete prophage regions (score < 70). **b** Structure of one intact prophage region. Genes are colored based on the predicted functions (color figure online)

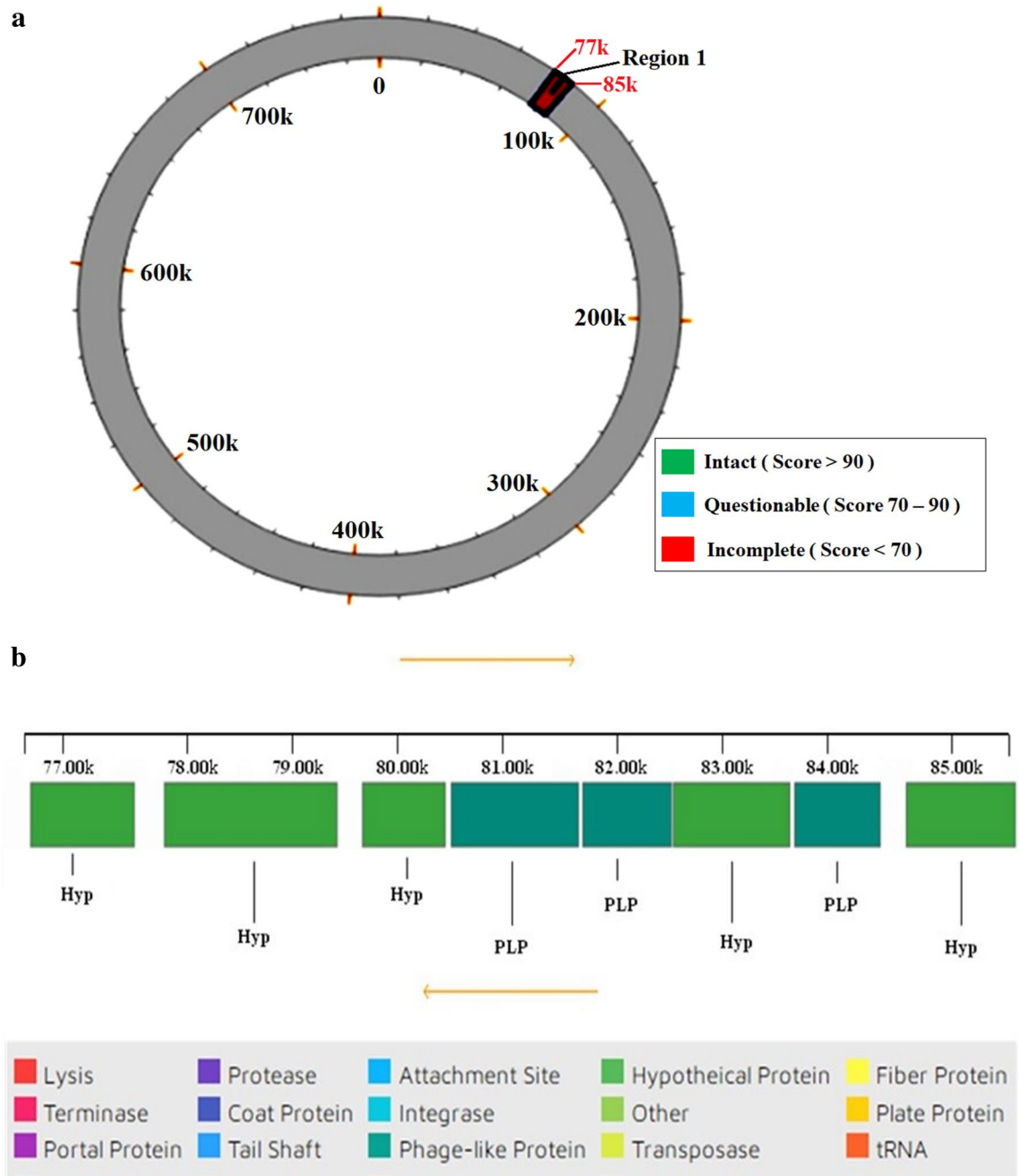


Fig. 14 **a** A total of 1 prophage region is positioned in the chromosome. *Green* indicates intact prophage regions (score >90), *blue* represents questionable prophage regions (score 70–90), and *red* specifies incomplete prophage regions (score <70). (**b**) Structure of one intact prophage region. Genes are colored based on the predicted functions (color figure online)

Conclusion

In the present study, the research has uncovered the complex genetic makeup of *N. farcinica* through genome analysis. This research has uncovered important genetic components, such as the putative virulence factors relA, icl, and mbtH besides shedding light on pathogen's versatility by identifying resistance genes to beta-lactams, macrolides, penams, and rifampin. This study has predicted the genetically closest match as the *N. farcinica* NBRC 15532 strain. The complete genome characterization also offers prospective directions for the development of novel drug targets by

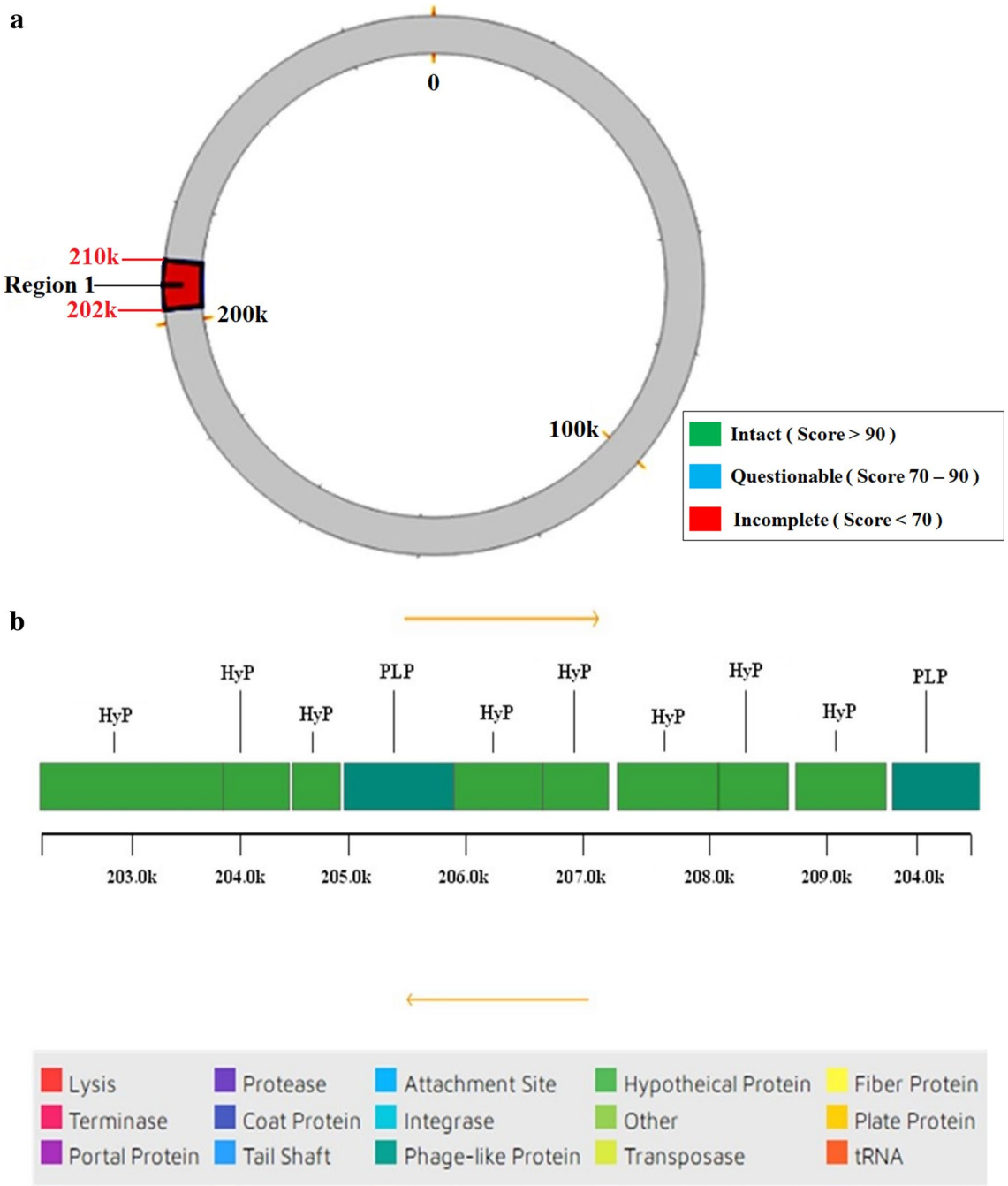


Fig. 15 **a** A total of 1 prophage region is positioned in the chromosome. *Green* indicates the intact prophage regions (score >90), *blue* represents the questionable prophage regions (score 70–90), and *red* specifies the incomplete prophage regions (score <70). **b** Structure of one intact prophage region. Genes are colored based on the predicted functions (color figure online)

providing invaluable insights into the clinically isolated *N. farcinica* sample. The investigation of the molecular characteristics and resistance mechanisms of *Nocardia* species is aided by whole-genome sequencing technology. In the future, subtractive genomics, comparative genomics, and reverse vaccinology approaches will be utilized to predict pathogen-specific potent drug and vaccine targets within significant metabolic pathways. (Khan et al. 2023; Afzal et al. 2023). Furthermore, virtual screening strategies will be employed to identify the therapeutic drug targets in

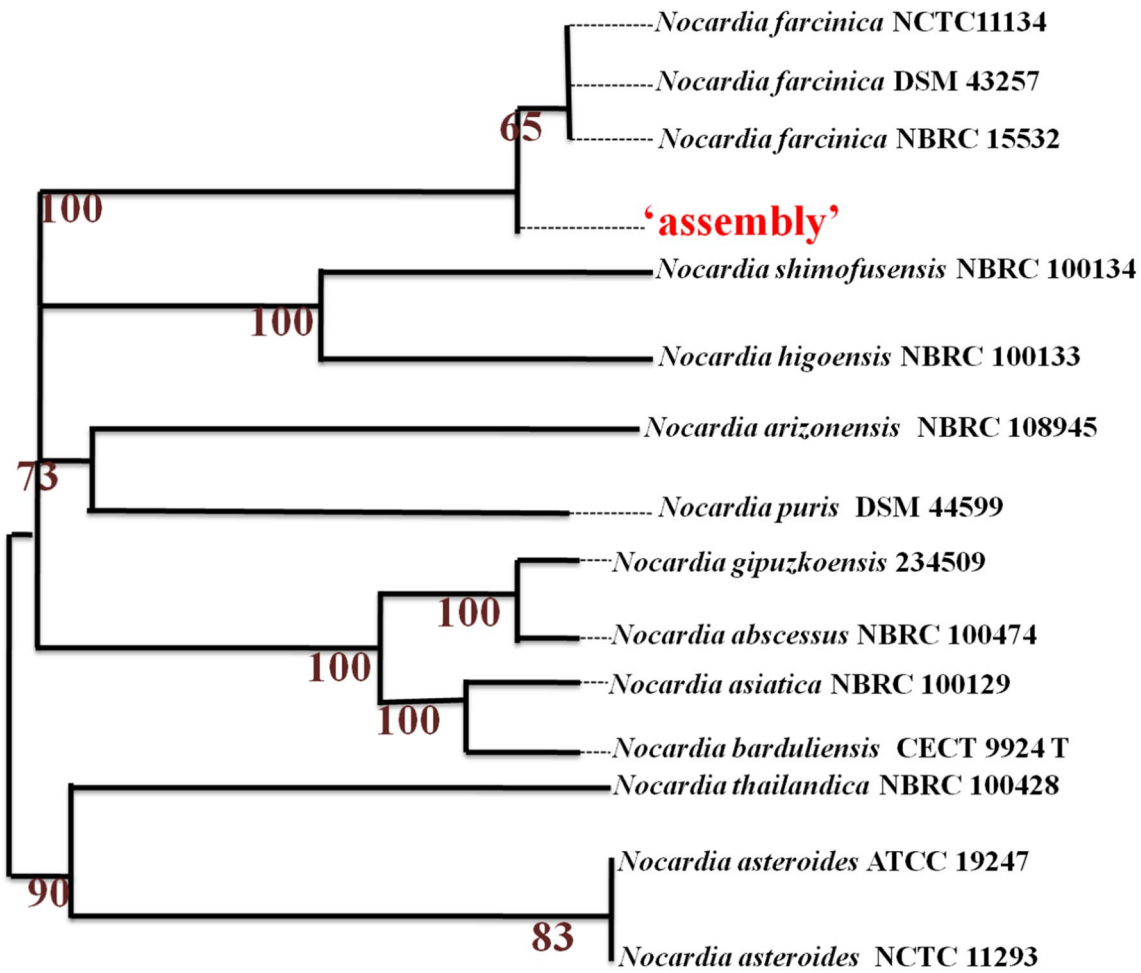


Fig. 16 DNA–DNA hybridization identification using type strain genome server (TYGS). The whole genome-based phylogenetic tree phylogenies built with TYGS are inferred using FastME 2.1.4 with a BioNJ starting tree and subtree pruning and regrafting post processing

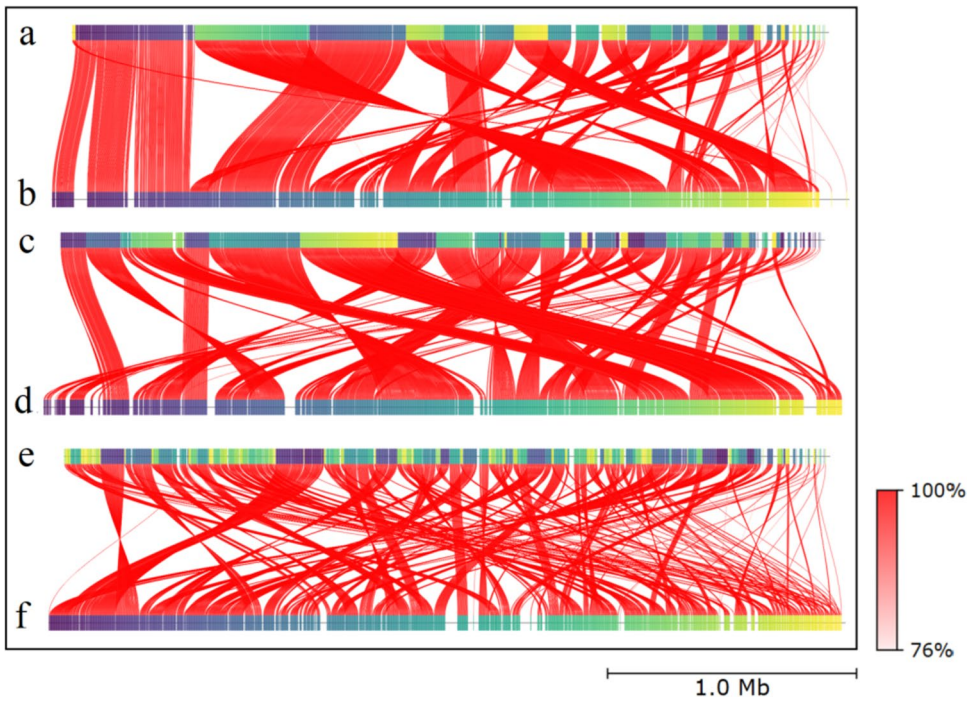


Fig. 17 Similarity analysis between the assembled genome and reference strains. **a**, **c e** Indicates the query, **b** Indicates the *N. farcinica* NCTC1134, **d** Indicates the *N. farcinica* DSM 43257, and f indicates *N. farcinica* NBRC 15532, utilizing Proksee to ascertain the average nucleotide identity (FastANI). Each red line division represents a reciprocal map between the two genes, showing their evolutionarily conserved sequences

Table 9 Identification of secondary metabolite in *N. farcinica* using antiSMASH

Region	Type	From	To	Most similar known cluster		Similarity
Region 1.1	Terpene	420,064	441,173	lymphostin/neolymphostinol B/lymphostinol/neolymphostin	NRP + Polyketide	16%
Region 1.2	Ranthipeptide	475,233	496,728	polyoxypeptin	NRP + Polyketide	8%
Region 1.3	NRP-Metallophore, NRPS	567,623	618,268	coelbactin	NRP	54%
Region 1.4	NRPS	764,532	818,136	hedamycin	Polyketide	9%
Region 1.5	T1PKS, hgIE-KS	867,687	939,724	Nataxazole	Polyketide	7%
Region 2.1	TIPKS	740,121	785,319			
Region 3.1	NRPS	1	54,956	Thiazostatin/watasemycin A/watasemycin B/2-hydroxyphenylthiazoline enantippyochelin/isopyochelin	NRP	13%
Region 3.2	Ectoine	109,067	119,462	ectoine	Other	100%
Region 3.3	Redox cofactor, T3PKS	546,119	602,238	Endophenazine A/endophenazine B	Other Phenazine	33%
Region 4.1	amino polycarboxylic-acid	226,371	240,126	EDHA	Other	55%
Region 4.2	T3PKS	281,413	307,612			
Region 5.1	NRP-Metallophore, NRPS, terpene	202,695	292,531	Nacobactin NA/nacobactin NA 10152B	NRP + Polyketide	100%
Region 6.1	T1PKS	29,367	85,950	lobosamide A/lobosamide B/lobosamide C	Polyketide	13%
Region 7.1	T1PKS	21,346	67,806			
Region 11.1	Terpene	115,713	136,906	isocrenieratene	Terpene	18%
Region 12.1	NAPAA	10,382	44,323	£-Poly-L-Lysine	NRP	100%
Region 13.1	aryl polyene	29,984	71,324	echoside A/echoside B/echoside C/echoside D/echoside E	NRP	11%
Region 14.1	NRPS	1	101,330	corynecin III/corynecin I/corynecin II	Other	40%
Region 17.1	NRPS,RIPP-like	1	24,291	actinoallolide A	Polyketide	20%
Region 17.2	Furan	41,346	62,336	Vazabitide A	NRP	10%
Region 18.1	NRPS	1	50,007	Bottromycin A2	RiPP:Bottromycin	12%
Region 20.1	Terpene	13,741	34,919	carotenoid	Terpene	18%
Region 26.1	NRPS-like	1	28,938			
Region 30.1	NRPS	1	23,425			
Region 33.1	NRPS	1	17,641	cinnapeptin	NRP	10%
Region 37.1	NRPS	1	8907	WS9326	NRP	5%
Region 47.1	NRPS-like, NRPS	1	4492	sarpeptin A/sarpeptin B	NRP	25%
Region 50.1	NRPS	1	3338			
Region 52.1	NRPS	1	2223			

Decoding secondary metabolites in *N. farcinica* through comprehensive identification with the antiSMASH Server. The analysis of biosynthetic gene clusters associated with the production of secondary metabolites

future studies (Irfan et al. 2023; Hassan et al. 2023; Hassan et al. 2023).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00203-023-03799-z.

Author contributions SN: Conceptualizations, data curation, formal analysis, visualization, writing—original draft, writing—review and editing, RR: Conceptualizations and data curation, NH: writing—review and editing, AM: AST analysis, JJ: Conceptualizations, methodology, supervision, funding acquisition, writing—review and editing, CJ and JDK: resources, sample collection, and sequencing.

Funding The author thank MHRD—RUSA 2.0, New Delhi (F.24e51/2014-U), TANSCHE (RGP/2019-20/ALU/HECP-0049 dated: 27/04/2021), DST-INDO-TAIWAN (GITA/DST/TWN/P-86/2019), Department of Biotechnology Bioinformatics Centre (BIC)-No. BT/PR40154/BTIS/137/34/2021, DBT-NNP-N0.BT/PR40156/BTIS/137/54/2023, DST-Fund for Improvement of S&T Infrastructure in Universities & Higher Educational Institutions (FIST) (SR/FST/LSI-667/2016) (C), and DST-Promotion of University Research and Scientific Excellence (PURSE phase II) (No. SR/PURSE Phase 2/38(G), 2017), NH greatly acknowledges the Department of Science and Technology (DST), for financial assistance through DST/INSPIRE Fellowship/2019/IF190083. Policy (TNMulti-Gen). RR

thankfully acknowledges the Indian Council of Medical Research [AMR/Fellowship/24/2022-ECD-II] for a Senior Research Fellowship. MA sincerely acknowledges the Indian Council of Medical Research, New Delhi for providing financial assistance through Research Associate (AMR/Fellowship/39/2022-ECD-II).

Data availability We deposited at the NCBI databases raw reads under the SRA are SRX17912461. The associated bioproject and biosample accession numbers are PRJNA890792 and SAMN31293886.

Declarations

Conflict of interests There is no conflict of interest.

References

Adapa S, Naramala S, Gayam V, Kapoor R, Raju M, Patel P, Konala VM (2020) An unusual presentation from a sporadic partially acid-fast aerobic actinomycete resistant to common antibiotics. J Investig Med High Impact Case Rep 8:2324709619899598. https://doi.org/10.1177/2324709619899598

Afzal M, Hassan SS, Sohail S, Camps I, Khan Y, Basharat Z, Karim A, Aurongzeb M, Irfan M, Salman M, Morel CM (2023) Genomic landscape of the emerging XDR *Salmonella Typhi* for mining druggable targets clpP, hisH, folP and gpmI and screening of novel TCM inhibitors, molecular docking and simulation analyses. BMC Microbiol 23(1):25. https://doi.org/10.1186/s12866-023-02756-6

Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS (2016) PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44(W1):W16–W21. https://doi.org/10.1093/nar/gkw387

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal, M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75

Bell M, McNeil MM, Brown JM (2019) *Nocardia*species (Nocardiosis) [Internet]. Antimicrobe, 11 July. http://www.antimicrobe.org/b117.asp. Accessed 11 July 2019

Besteiro B, Coutinho D, Fragoso J, Figueiredo C, Nunes S, Azevedo C, Teixeira T, Selaru A, Abreu G, Malheiro L (2023) Nocardiosis: a single-center experience and literature review. Br J Infect Dis:102806.

Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, van Wezel GP, Medema MH, Weber T (2023) antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res 51(W1):W46–W50. https://doi.org/10.1093/nar/gkad344

Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA 3rd, Stevens R, Vonstein V, Wattam AR, Xia F (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365

Brown JM, Cowley KD, Manninen KI, McNeil MM (2007) Phenotypic and molecular epidemiologic evaluation of a *Nocardia farcinica* mastitis epizootic. Vet Microbiol 125(1–2):66–72. https://doi.org/10.1016/j.vetmic.2007.04.044

Cai Q, Huang Y, Zhou L, Hu N, Liu Y, Guo F, Liu Q, Huang X, Zhang Y, Zeng L (2022) A complete genome of *Nocardia terpenica* NC_YFY_NT001 and pan-genomic analysis based on different sources of *Nocardia spp.* isolates reveal possibly host-related virulence factors. Infect Drug Resis 15:7259–7270. https://doi.org/10.2147/IDR.S384673

Centers for Disease Control and Prevention (2019) Antibiotic resistance: a global threat [Internet], 5 July. https://www.cdc.gov/features/antibiotic-resistance-global/index.html

Conville PS, Brown-Elliott BA, Smith T, Zelazny AM (2017) The complexities of *Nocardia* taxonomy and identification. J Clin Microbiol 56(1):e01419-e1517. https://doi.org/10.1128/JCM.01419-17

Conville PS, Brown JM, Steigerwalt AG, Lee JW, Anderson VL, Fishbain JT, Holland SM, Witebsky FG (2004) *Nocardia kruczakiae* sp. nov., a pathogen in immunocompromised patients and a member of the "*N. nova complex*". J Clin Microbiol 42(11):5139–5145. https://doi.org/10.1128/JCM.42.11.5139-5145.2004

Cortese IJ, Castrillo ML, Onetto AL, Bich GÁ, Zapata PD, Laczeski ME (2021) Denovo genome assembly of Bacillus altitudinis 19RS3 and *Bacillus altitudinis* T5S-T4, two plant growth-promoting bacteria isolated from *Ilex paraguariensis* St. Hil. (yerba mate). PloS One 16(3):e0248274. https://doi.org/10.1371/journal.pone.0248274

De Pascale G, Wright GD (2010) Antibiotic resistance by enzyme inactivation: from mechanisms to solutions. Chembiochem 11(10):1325–1334. https://doi.org/10.1002/cbic.201000067

Duggal SD, Chugh TD (2020) Nocardiosis: a neglected disease. Med Princ Pract 29(6):514–523. https://doi.org/10.1159/000508717

Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39(Web Server issue):W29–W37. https://doi.org/10.1093/nar/gkr367

Galata V, Fehlmann T, Backes C, Keller A (2019) PLSDB: a resource of complete bacterial plasmids. Nucleic Acids Res 47(D1):D195–D202. https://doi.org/10.1093/nar/gky1050

Gao S, Liu Q, Zhou X, Dai X, He H (2021) Primary cutaneous nocardiosis due to *Nocardia farcinica*: a case report of an often overlooked infection. Infect Drug Resis 14:1435–1440. https://doi.org/10.2147/IDR.S306161

Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen CY, Graham M, Van Domselaar G, Stothard P (2023) Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res 51(W1):W484–W492. https://doi.org/10.1093/nar/gkad326

Hassan SS, Jamal SB, Radusky LG, Tiwari S, Ullah A, Ali J, Behramand de Carvalho PVSD, Shams R, Khan S, Figueiredo HCP, Barh D, Ghosh P, Silva A, Baumbach J, Röttger R, Turjanski AG, Azevedo VAC (2018) The druggable pocketome of *Corynebacterium diphtheriae*: A new approach for in silico putative druggable targets. Front Genet 9:44. https://doi.org/10.3389/fgene.2018.00044

Hassan SS, Shams R, Camps I, Basharat Z, Sohail S, Khan Y, Ullah A, Irfan M, Ali J, Bilal M, Morel CM (2023) Subtractive sequence analysis aided druggable targets mining in *Burkholderiacepacia complex* and finding inhibitors through bioinformatics approach. Mol Diversity 27(6):2823–2847. https://doi.org/10.1007/s11030-022-10584-5

Hoshino Y, Fujii S, Shinonaga H, Arai K, Saito F, Fukai T, Satoh H, Miyazaki Y, Ishikawa J (2010) Monooxygenation of rifampicin catalyzed by the rox gene product of *Nocardia farcinica*: structure elucidation, gene identification and role in drug resistance. J Antibiot 63(1):23–28. https://doi.org/10.1038/ja.2009.116

Hu Y, Morichaud Z, Chen S, Leonetti JP, Brodolin K (2012) *Mycobacterium tuberculosis* RbpA protein is a new type of transcriptional activator that stabilizes the σ A-containing RNA polymerase holoenzyme. Nucleic Acids Res 40(14):6547–6557. https://doi.org/10.1093/nar/gks346

Irfan M, Tariq M, Basharat Z, Abid Khan RM, Jahanzaeb M, Shakeel M, Nisa ZU, Shahzad M, Jahanzaib M, Moin ST, Hassan SS, Khan IA (2023) Genomic analysis of Chryseobacteriumindologenes and conformational dynamics of the selected DD-peptidase. Res Microbiol 174(1–2):103990. https://doi.org/10.1016/j.resmic.2022.103990

Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9(1):5114. https://doi.org/10.1038/s41467-018-07641-9

Ji X, Zhang X, Li H, Sun L, Hou X, Song H, Han L, Xu S, Qiu X, Wang X, Zheng N, Li Z (2020) Nfa34810 facilitates *Nocardia farcinica* invasion of host cells and stimulates tumor necrosis factor alpha secretion through activation of the NF-κB and mitogen-activated protein kinase pathways via toll-like receptor 4. Infect Immun 88(4):e00831-e919. https://doi.org/10.1128/IAI.00831-19

Ji X, Han L, Zhang W, Sun L, Xu S, Qiu X, Fan S, Li Z (2022) Molecular, cellular and neurological consequences of infection by the neglected human pathogen Nocardia. BMC Biol 20(1):251. https://doi.org/10.1186/s12915-022-01452-7

Jiao M, Deng X, Yang H, Dong J, Lv J, Li F (2021) Case report: a severe and multi-site *Nocardia farcinica* infection rapidly and precisely identified by metagenomic next-generation sequencing. Front Med 8:669552. https://doi.org/10.3389/fmed.2021.669552

Juraschek K, Borowiak M, Tausch SH, Malorny B, Käsbohrer A, Otani S, Schwarz S, Meemken D, Deneke C, Hammerl JA (2021) Outcome of different sequencing and assembly approaches on the detection of plasmids and localization of antimicrobial resistance genes in commensal *Escherichia coli*. Microorganisms 9(3):598. https://doi.org/10.3390/microorganisms9030598

Khan A, Sohail S, Yaseen S, Fatima S, Wisal A, Ahmed S, Nasir M, Irfan M, Karim A, Basharat Z, Khan Y, Aurongzeb M, Raza SK, Alshahrani MY, Morel CM, Hassan SS (2023) Exploring and targeting potential druggable antimicrobial resistance targets ArgS, SecY, and MurA in *Staphylococcus sciuri* with TCM inhibitors through a subtractive genomics strategy. Funct Integr Genomics 23(3):254. https://doi.org/10.1007/s10142-023-01179-w

Kövér Z, Nordskag VJ, Bán Á, Gajdács M, Urbán E (2023) The role of *Actinomyces spp.* And related organisms in cervicofacial infections: pathomechanism, diagnosis and therapeutic aspects. Anaerobe:102767

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

Lebeaux D, Bergeron E, Berthet J, Djadi-Prat J, Mouniée D, Boiron P, Lortholary O, Rodriguez-Nava V (2019) Antibiotic susceptibility testing and species identification of Nocardia isolates: a retrospective analysis of data from a French expert laboratory, 2010–2015. Clin Microbiol Infect 25(4):489–495. https://doi.org/10.1016/j.cmi.2018.06.013

Lee EK, Kim J, Park DH, Lee CK, Kim SB, Sohn JW, Yoon YK (2021) Disseminated nocardiosis caused by *Nocardia farcinica* in a patient with colon cancer: a case report and literature review. Medicine 100(29):e26682. https://doi.org/10.1097/MD.0000000000026682

Madaha EL, Mienie C, Gonsu HK, Bughe RN, Fonkoua MC, Mbacham WF, Alayande KA, Bezuidenhout CC, Ateba CN (2020) Whole-genome sequence of multi-drug resistant *Pseudomonas aeruginosa* strains UY1PSABAL and UY1PSABAL2 isolated from human broncho-alveolar lavage, Yaoundé, Cameroon. PloS One 15(9):e0238390. https://doi.org/10.1371/journal.pone.0238390

Mehta HH, Shamoo Y (2020) Pathogenic *Nocardia*: a diverse genus of emerging pathogens or just poorly recognized? PLoS Pathog 16(3):e1008280. https://doi.org/10.1371/journal.ppat.1008280

Meier-Kolthoff JP, Göker M (2019) TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10(1):2182

Mozrall AR, Miranda RR, Kumar G, Wadsworth CB, Hudson AO (2022) Isolation, whole-genome sequencing, and annotation of two antibiotic-producing and resistant bacteria, *Enterobacter roggenkampii* RIT 834 and *Acinetobacter pittii* RIT 835, from disposable masks collected from the environment. Microbiol Resour Announc:e0075722. https://doi.org/10.1128/mra.00757-22

Naveed M, Mahmood S, Aziz T, Azeem A, Hussain I., Waseem M, Ali A, Alharbi M, Alshammari A, Alasmari AF (2023) Designing a novel chimeric multi-epitope vaccine subunit against *Staphylococcus argenteus* through artificial intelligence approach integrating pan-genome analysis, in vitro identification, and immunogenicity profiling. J Biomol Struct Dyn:1–16. https://doi.org/10.1080/07391102.2023.2256881

Newell KV, Thomas DP, Brekasis D, Paget MS (2006) The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on *Streptomyces coelicolor*. Mol Microbiol 60(3):687–696. https://doi.org/10.1111/j.1365-2958.2006.05116.x

Nouioui I, Cortés-Albayay C, Neumann-Schaal M, Vicente D, Cilla G, Klenk HP, Marimón JM, & Ercibengoa M (2020) Genomic virulence features of two novel species *Nocardia barduliensis sp. nov.* and *Nocardia gipuzkoensis sp. nov.*, isolated from patients with chronic pulmonary diseases. Microorganisms 8(10):1517. https://doi.org/10.3390/microorganisms8101517

Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R (2014) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42(Database issue):D206–D214. https://doi.org/10.1093/nar/gkt1226

Pei T, Liu Y, Du J, Huang K, Deng MR, Zhu H (2021) *Croceicoccus gelatinilyticus* sp. nov., isolated from a tidal flat sediment. Arch Microbiol 204(1):93. https://doi.org/10.1007/s00203-021-02703-x

Petar P, Beaman BL, Ravalison M, Boiron P, Rodríguez-Nava V (2015) *Nocardia* and actinomyces. Mol Med Microbiol:731–752. https://doi.org/10.1016/b978-0-12-397169-2.00040-8

Qiu F, Ma Z, Zhong R, Huang H, Wang Y, Liu H (2022) Case report: disseminated nocardiosis caused by *Nocardia vulneris* in a patient with macroglobulinemia. Front Public Health 10:866420. https://doi.org/10.3389/fpubh.2022.866420

Rahim Y, Khan J, Shahid S, Awan S, Irfan M (2023) Clinical characteristics, outcomes, and factors associated with mortality in *Nocardia pneumonia*: 18 years' real-world data from a tertiary care hospital in Karachi. Pak Respir Investig 61(2):254–260

Rathish B, Zito PM (2023) Nocardia. [Updated 2023 Aug 7]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing, Jan. https://www.ncbi.nlm.nih.gov/books/NBK560872/

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics (Oxford, England) 30(14):2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Seemann T (2022) ABRicate, version 1.0.1. https://github.com/tseemann/abricate. Accessed 30 June 2021

Sharma T, Kumar R, Kalra JS, Singh S, Bhalla GS, Bhardwaj A (2023) Galaxy ASIST: a web-based platform for mapping and assessment of global standards of antimicrobial susceptibility: a case study in *Acinetobacter baumannii* genomes. Front Microbiol 13:1041847. https://doi.org/10.3389/fmicb.2022.1041847

Sun J, Deng Z, Yan A (2014) Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. Biochem Biophys Res Commun 453(2):254–267. https://doi.org/10.1016/j.bbrc.2014.05.090

Tian C, Xing M, Fu L, Zhao Y, Fan X, Wang S (2022) Emergence of uncommon KL38-OCL6-ST220 carbapenem-resistant

Acinetobacter pittii strain, co-producing chromosomal NDM-1 and OXA-820 carbapenemases. Front Cell Infect Microbiol 12:943735. https://doi.org/10.3389/fcimb.2022.943735

Tripodi MF, Durante-Mangoni E, Fortunato R, Cuccurullo S, Mikami Y, Farina C, Utili R (2011) In vitro activity of multiple antibiotic combinations against *Nocardia*: relationship with a short-term treatment strategy in heart transplant recipients with pulmonary nocardiosis. Transplant Infectious Disease: an Official Journal of the Transplantation Society 13(4):335–343. https://doi.org/10.1111/j.1399-3062.2010.00588.x

Wang H, Cheng F, Dong XT, Li XF, Wang YJ, Hao MJ (2020) Zhonghuayu fang yixue za zhi [Chin J Prev Med] 54(12):1441–1447. https://doi.org/10.3760/cma.j.cn112150-20200922-01226

Wick RR, Judd LM, Gorrie CL, Holt KE (2017) Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13(6):e1005595. https://doi.org/10.1371/journal.pcbi.1005595

Williams ST, Goodfellow M, Alderson G, Wellington EM, Sneath PH, Sackin MJ (1983) Numerical classification of Streptomyces and related genera. J Gen Microbiol 129(6):1743–1813. https://doi.org/10.1099/00221287-129-6-1743

Wu J, Li X, Zhang T, Lin X, Chen YC (2021) Disseminated *Nocardia farcinica* involves the spinal cord: a case report and review of the literature. BMC Infect Dis 21(1):1224. https://doi.org/10.1186/s12879-021-06905-y

Yetmar ZA, Khodadadi RB, Chesdachai S, McHugh JW, Challener DW, Wengenack NL, Bosch W, Seville MT, Beam E (2023) Mortality after nocardiosis: risk factors and evaluation of disseminated infection. Open Forum Infect Dis 10(8):ofad409. https://doi.org/10.1093/ofid/ofad409

Zakaria Z, Hassan L, Ahmad N, Husin SA, Ali RM, Sharif Z, Sohaimi NM, Garba B (2021) Discerning the antimicrobial resistance, virulence, and phylogenetic relatedness of salmonella isolates across the human, poultry, and food materials sources in Malaysia. Front Microbiol 12:652642. https://doi.org/10.3389/fmicb.2021.652642

Zhang X, Ruan Y, Liu W, Chen Q, Gu L, Guo A (2021) Whole genome sequencing and genome annotation of *Dermacoccusabyssi strain* HZAU 226 isolated from spoiled eggs. Genomics 113(1 Pt 2):1199–1206. https://doi.org/10.1016/j.ygeno.2020.12.005

Zhang J, Zhu Y, Sun Y, Han X, Mao Y (2023) Pathogenic Nocardia amamiensis infection: a rare case report and literature review. Heliyon 9(7):e17183. https://doi.org/10.1016/j.heliyon.2023.e17183

Zhao G, Luo Z, Wang Y, Liu J, Wu D, Zhang L, Yang X (2020) Draft genome sequencing and annotation of a low-virulence Morganella morganii strain CQ-M7, a multidrug-resistant isolate from the giant salamander in China. J Glob Antimicrob Resist 20:248–252. https://doi.org/10.1016/j.jgar.2019.08.012

Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS (2011) PHAST: a fast phage search tool. Nucleic Acids Res 39(Web Server issue):W347–W352. https://doi.org/10.1093/nar/gkr485