Genetic and culture diversity of bacteria of Loktak lake and descriptions of a few new taxa

Thesis submitted for the degree of
Doctor of Philosophy

By

L Lhingjakim Khongsai
(Reg. No. 16LPPH07)

Research Supervisor
Prof. Ch. Venkata Ramana

Department of Plant Sciences
School of Life Sciences
University of Hyderabad
Hyderabad – 500046
INDIA

July, 2023

UNIVERSITY OF HYDERABAD

(A Central University established in 1974 by an Act of the Parliament) HYDERABAD 500046

प्रतिष्ठित संस्थान
INSTITUTION OF EMINENCE

CERTIFICATE

This is to certify that **Ms. L Lhingjakim Khongsai** has carried out the research work embodied in the present thesis under the supervision and guidance of **Prof. Ch. Venkata Ramana** for the full period prescribed under the Ph.D. ordinances of this University. We recommend her thesis entitled **"Genetic and culture diversity of bacteria of Loktak lake and descriptions of a few new taxa"** for submission for the degree of Doctor of Philosophy to the University.

Prof. Ch. Venkata Ramana Ph.D
Department of Plant Sciences School of Life Sciences
University of Hyderabad, P.O. Central University,
Hyderabad - 500 046, India

24/7/2023

HEAD
Department of Plant Sciences

HEAD
Dept. of Plant Sciences
School of Life Sciences
University of Hyderabad
Hyderabad-500 046 INDIA

24/7/23

DEAN
School of Life Science

संकाय अध्यक्ष / Dean
जीव विज्ञान संकाय / School of Life Sciences
हैदराबाद विश्वविद्यालय / University of Hyderabad
हैदराबाद / Hyderabad-500 046.

24/7/2023

SUPERVISOR
Prof. Ch. Venkata Ramana

UNIVERSITY OF HYDERABAD
(A Central University established in 1974 by an Act of the Parliament) HYDERABAD 500046

प्रतिष्ठित संस्थान
INSTITUTION OF EMINENCE

DECLARATION

I, **L Lhingjakim Khongsai**, hereby declare that this thesis entitled **"Genetic and culture diversity of bacteria of Loktak lake and descriptions of a few new taxa"** submitted by me under the guidance and supervision of **Prof. Ch. Venkata Ramana** is an original and independent research work. I, hereby declare that this work is original and has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

L. Lhingjakim

L Lhingjakim Khongsai
(Ph.D. Student)

Prof. Ch. Venkata Ramana Ph.D
Department of Plant Sciences School of Life Sciences
University of Hyderabad, P.O. Central University,
Hyderabad-500 046, India
21/7/2023

Supervisor
Prof. Ch. Venkata Ramana

UNIVERSITY OF HYDERABAD

(A Central University established in 1974 by an Act of the Parliament) HYDERABAD 500046

प्रतिष्ठित संस्थान
INSTITUTION OF EMINENCE

CERTIFICATE

This is to certify that the thesis entitled **"Genetic and culture diversity of bacteria of Loktak lake and descriptions of a few new taxa"** submitted by **Ms. L Lhingjakim Khongsai**, bearing registration number **16LPPH07**, in partial fulfilment of the requirements for the award of Doctoral of Philosophy in Department of Plant Sciences, School of Life Sciences, University of Hyderabad, is a bonafide work carried out by her under my guidance and supervision. This thesis is free from plagiarism and has not been submitted in any part or in full to this or any other University or Institute for the award of any degree or diploma.

Parts of the thesis have been:

A. Authored in the following publications:

1. Kumar, G. *, **Lhingjakim, K. L**. *, Uppada, J. Ahamad, S., Kumar, D., Kashif, G.M., Sasikala, C., and Ramana, C.V. (2021). *Aquisphaera insulae* sp. nov., a new member in the family *Isosphaeraceae*, isolated from the floating island of Loktak lake and emended description of the genus *Aquisphaera*. *Antonie. Leeuwenhoek. J. Microbiol.* 114: 1465–147. https://doi.org/10.1007/s10482-021-01615-6 (*, the authors contributed equally)

2. **Lhingjakim, K. L**. *, Smita, N. *, Kumar, G.*, Jagadeeshwari, U., Ahamad. S., Sasikala, C., and Ramana, C.V. (2022). *Paludisphaera rhizosphaereae* sp. nov., a new member of the family *Isosphaeraceae*, isolated from the rhizosphere soil of *Erianthus ravennae*. *Antonie. Leeuwenhoek. J. Microbiol.* 115: 1073–1084. https://doi.org/10.1007/s10482-022-01758-0 (*, the authors contributed equally)

3. **Lhingjakim, K. L.**, Sasikala, C. and Ramana, C.V. (2022). Draft Genome Sequence of *Glutamicibacter* sp. Strain JC586, Isolated from Soil Sediment of the Floating Islands of Loktak Lake, India. *Microbiol. Resour. Announc.* Sep 20: e00746-22.

B. Presented at the following conferences:

1. Presented poster at the "59th Annual Conference of Association of Microbiologists of India" (Dec. 2018), University of Hyderabad, Hyderabad.
2. Presented an oral presentation at "National Conference on Frontiers in Plant Biology" (Jan. 2019), University of Hyderabad, Hyderabad.

Further, the student has passed the following courses towards the fulfilment of the coursework requirement for Ph.D. degree.

S. No.	Course	Name	Credits	Results
1	PL801	Analytical techniques	4	Pass
2	PL802	Research ethics, biosafety, data analysis and biostatistics	4	Pass
3	PL803	Lab work and seminar	4	Pass

Prof. Ch. Venkata Ramana Ph.D
Department of Plant Sciences, School of Life Sciences
University of Hyderabad (UGC Central University)
Hyderabad - 500 046, India

HEAD
Department of Plant Sciences

HEAD
Dept. of Plant Sciences
School of Life Sciences
University of Hyderabad
Hyderabad-500 046 INDIA

DEAN
School of Life Science

संकाय अध्यक्ष / Dean
जीव विज्ञान संकाय / School of Life Sciences
हैदराबाद विश्वविद्यालय / University of Hyderabad
हैदराबाद / Hyderabad-500 046.

SUPERVISOR
Prof. Ch. Venkata Ramana

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervisor, ***Prof. Ch. Venkata Ramana*** *for his valuable guidance, and unwavering support throughout the completion of my Ph.D. His expertise and profound knowledge of the subject matter, and innovative approach to research have pushed me to continually strive for excellence and restored my confidence in pursuit of my research goals. I am also grateful for his exceptional support and guidance during my difficult period in my academic journey. With the unexpected and untimely passing of my former guide, Prof. Ch. Venkata Ramana graciously took me in and become my guiding light. I am immensely grateful for his compassion, constant encouragement and suggestions without which this thesis would not have been possible. Above all, I am truly fortunate to have him as my supervisor and deeply grateful for shaping me into who I am today.*

I would also like to honour the memory of late professor, ***Prof. K. Seshagiri Rao****. Although no longer with us, His impact on my academic journey deserves utmost recognition and gratitude. I am forever grateful for the wisdom and guidance that Prof. K. Seshagiri Rao imparted during our time together.*

I would like to acknowledge and thank my doctoral committee members ***Prof. S. Rajagopal*** *and* ***Prof. K. Gopinath*** *for their valuable and constructive suggestions during my work.*

I am also grateful to former Deans, ***Prof. P. Reddanna****,* ***Prof. Ramaiah****,* ***Prof. S. Dayananda****, the current Dean,* ***Prof. N. Shiva Kumar*** *School of Life Sciences; and the former Head,* ***Prof. Ch. Venkata Ramana****,* ***Prof. Padmaja****, and present Head,* ***Prof. S. Rajagopal****, Dept. of Plant Sciences, for extending their support in all possible ways.*

I am particularly grateful to ***Prof. Ch. Sasikala****, CEN, IST, JNTU for allowing me to work in her lab. Her generosity in sharing her laboratory facilities have been invaluable in shaping the outcome of my PhD programme. Her insightful comment and constant encouragement have been truly inspiring.*

Infrastructural support provided by UGC-SAP, DBT-CREBB and DST-FIST to the Dept. of Plant Sciences are highly acknowledged. Financial support from CSIR, DBT-TATA and IOE projects are acknowledged.

I also thank my former and senior lab members (Dr. Laksmi Prasuna, Dr. Mujahid M, Dr. Ganesh M, Dr. Suresh G, Dr. Ashif Ali, Dr. Dhanesh K, Dr. Shabbir A, Dr. Indu, Dr. Deepshikha G, Dr. Gaurav K, Dr. Anusha, Amulya Sai) for their support and valuable suggestions.

I am also grateful like to my present labmates (Krishna, Smita, Sreya, Ipshita, Ria, Vignesh, Sushmita, Mahima, Ramudu, Hari, Sohini, Bhavana, Chandramani, Pragya). Their positive attitude, motivation, and willingness to lend a helping hand had been a source of inspiration throughout my PhD journey. I am also grateful to lab mates from JNTUH (Jagadeeshawri, Richa and Sangeetha ma'am) for their help and cooperation during my research.

I would like to extend my heartful appreciation to my friends (HU family, Dr. Sandeep Das, Dr. Amita Pal, Dr. Jampa, Bempi, Seiboi, Lalboi, Mahati, Joycy). Their motivation and uplifting words during challenging times have been a source of strength and their friendship have been invaluable to me, both personally and academically.

I am also grateful for the diverse wildlife that we call our campus home. Special mention to the beloved dogs and cats (Mimi, Browny, Chotu, Bibi, Nungahnu, Abby, Blacky, Happy, Kuku) who have brought joy, comfort and companionship during my stay in the university. I would like to thank the university administration for fostering a campus environment that embraces these furry companions.

The help and cooperation of the non-teaching staff in SLS, hostels and administrative building is highly acknowledged.

I am also thankful for all the help I received from dad, mom, uncle Brojen, Maj. Hegin, Kui, SI Goulen, for their valuable assistance during the sampling phase of my research.

To my family, I am thankful and blessed for the love, patience, and the unwavering support throughout my academic journey. I am also grateful to my siblings for their belief in me and for being my rock in fulfilling my dreams. I would also like to express my gratitude to my uncles, aunts, and community elders who always fervently prayed for me.

I would like to dedicate this thesis in loving memory of my late brother (L Thangkhai Shoney) who set the foundation of my education. His love, guidance and belief in my capabilities played an instrumental role in shaping my academic pursuits.

Lastly, I am profoundly grateful to the Almighty God for His blessings and guidance during this academic endeavour.

L Lhingjakim Khongsai

“Nature is not a place to visit, it is home”

- Garry Snider

Dedicated to my beloved brother (L) L Thangkhai Shoney and my supervisor

Contents

Abbreviations	Expansion
AAI	Average amino acid identity
AL	Aminolipid
ANI	Average nucleotide identity
antiSMASH	Antibiotics & Secondary Metabolite Analysis Shell
BP	Beibl and Pfennig media
BGI	Beijing Genomics Institute
BLASTp	Protein Basic Local Alignment Search Tool
BPGA	Bacterial Pan Genome Analysis tool
cas	CRISPER-associated protein CBM Carbohydrate-binding module
CDS	Coding DNA sequences
CE	Carbohydrate esterase
C/N	Carbon/Nitrogen
COG	Clusters of Orthologous Groups of Proteins
CPR	Candidate Phyla Radiation
CRISPR	Clustered regularly interspaced short palindromic repeats
CSP	Conserved signature protein
˚C	Degree Celsius
*d*DDH	Digital DNA-DNA Hybridisation
DNA	Deoxyribonucleic acid
DPG	Diphosphatidylglycerol
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
EDTA	Ethylenediaminetetraacetic acid
ENA	European Nucleotide Archive
FAME	Fatty acid methyl ester
FESEM	Field scanning electron microscope
gL-1	Gram per litre
G+C	Guanine + cytosine
GGDH	Genome to genome distance hybridization
GH	Glycosyl hydrolase GL Glycolipid
GPS	Global positioning system
GT	Glycosyl transferase
h	Hour
HPLC	High-pressure liquid chromatography
IAA	Indole-3-acetic acid
ICNP	International Code of Nomenclature of Prokaryotes
ICM	Inner cell mass
IJSEM	International Journal of Systematic and Evolutionary Microbiology
IMG	Integrated Microbial Genomes
IUCN	International Union for Conservation of Nature
KACC	Korean Agricultural Culture Collection
KEGG	Kyoto Encyclopedia of Gene and Genomes
KCTC	Korean Collection for Type Cultures
KLNJ	Keibul Lamjao National Park
KO	KEGG Orthology
L	Litre
LCB	Local collinear blocks
LDA	Loktak Development Authority
LOK	Loktak
LTP	The Living Tree Project
Mbp	Mega base pair
MEGA	Molecular evolutionary genetics analysis
MLSA	Multilocus Sequence Analysis
MLST	Multilocus Sequence Typing

MK	Menaquinone
Δ*T*m	Melting temperature
µg	Microgram
µg-l	Microgram per litre
mg	Milligram
ML	Maximum likelihood
µl	Microlitre
ml	Milliliter
µm	Micrometer
mM	Millimolar
MoEFCC	Ministry of Environment, Forest and Climate Change
MLSA	Multilocus sequence analysis
MUSCLE	Multiple Sequence Comparison by Log-Expectation
NCBI	National center for biotechnology information
NDMS	Non-metric multi-dimensional scaling
Nm	Nanometer
NWA	National Water Act
nt	Nucleotide
OD	Optical density
OGRI	Overall genomic relatedness indices
OPA	o-Phthaldialdehyde
Ortho ANI	Orthologous average nucleotide identity
OTU	Operational taxonomic unit
PATRIC	PathoSystems Resource Integration Center
PCR	Polymerase chain reaction
PDA	Photodiode array detector
PFGE	Pulse field gel electrophoresis
PGAP-X	Pan-genome analysis pipeline
PGAweb	Pan genomes analyse web server
PICRUSt	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
PL	Phospholipid
POCP	Percentage of conserved protein
Q	Ubiquinone
RAST	Rapid Annotations using Sub-Systems Technology
RDP	Ribosomal Database Project
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
SEM	Scanning electron microscopy
SRA	Sequence read archive
TAE	Tris-base, acetic acid and EDTA
TE	Tris-EDTA
TLC	Thin layer chromatography
TPKS	Type 1 polyketide synthases
TYGS	Type Strain Genome server
UBCG	Up-to-date bacterial core gene
v/v	Volume per volume
w/v	Weight per volume
WGS	Whole genome sequencing
WISA	Wetlands International - South Asia

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

1. INTRODUCTION

Microorganisms are ubiquitous in nature, underpinning various ecological functions with enormous diversity. They can inhabit almost every habitat including soil, air, water and plant-associated and can be found in extreme environmental conditions like cold, saline, drought, hot springs, and acidic or alkaline soil. The microbiomes present in lake ecosystems serve as an ideal bioresource for agriculture, industry, and other related fields (Yadav et al. 2018). Freshwater habitats such as lakes, streams, rivers, and wetlands offer valuable ecosystem services to humans like drinking water, fisheries, storage, and release of terrestrial carbon and also act as treasure inventory for the extraction of biotechnological and pharmaceutically important bioactive compounds. Lakes are also the connection point of interaction between humans and numerous host organisms (Heino et al. 2021). In-depth genomic studies of numerous aquatic habitats like Lake Soyang (South Korea), Lough Neagh Lake (Ireland), Soda Lake (Russia), Lonar Lake, Pangong Lake and Pangong Lake (India), Salt- lake (China), and Antarctic Lake, have revealed the microbial diversity and functional profiling of these lakes (Kaushal et al. 2022). Freshwater lakes are commonly established in low-lying areas that receive their water from streams, rivers, and nearby sites. They provide a different environment for microbes in comparison to other aquatic environments (moving water and oceans) and the microbes in such environments carry out important processes like organic matter decomposition, sequestration of inorganic compounds, and stabilizing the biogeochemical cycle. Research on the microbiome of freshwater lake environments has also been conducted globally. e.g. Vembanad Lake, India (Chandran et al. 2011), Lake Baikal, Russia (Kurilkina et al. 2016), Toolik Lake, Alaska, USA (Bahr et al. 1996), Lake Fuchskuhle, Germany (Glockner et al. 2006), Adirondack lake, New York (Hiorns et al. 1997), Lake Soyang, South Korea (Zwart et al. 2002), Lake

Loosdrecht, The Netherlands (Zwart et al. 1998), Lake Kasumigaura, Japan (Tamaki et al. 2005), Lake Taihu, China (Wu et al. 2007), Lake Tanganyika, Africa (De Wever et al. 2005), Lake Cadagno, Switzerland (Bosshard et al. 2003).

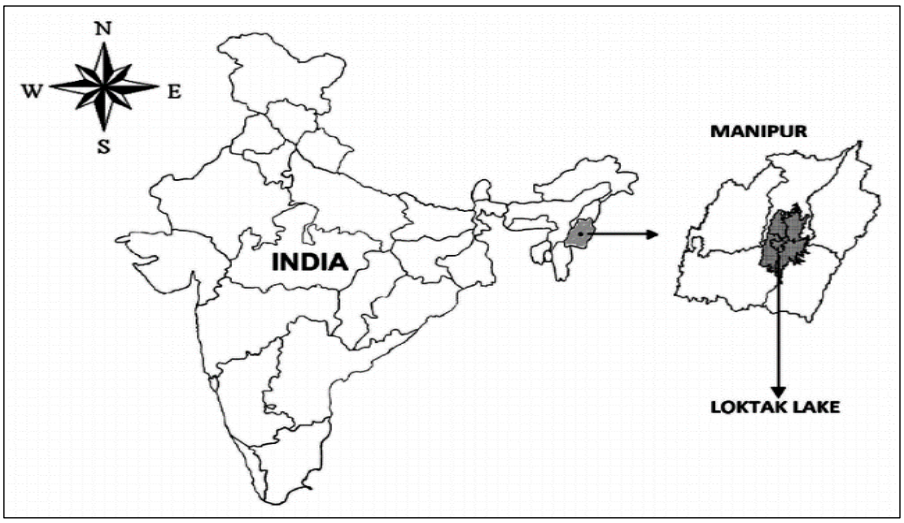
In freshwater ecosystems, bio-resources are largely unexplored, especially in the eastern part of India. However, studies subjected to lakes and other water bodies such as mangroves, riverine, estuarine, hot-springs and coastal and marine ecosystems of different regions of India are available in the literature. For instance, the evaluation of the microbiome of the hyper-saline Sambhar salt lake (Mehta et al. 2021), exploring the taxonomic and functional diversity of bacterial communities of Himalayan freshwater lake i.e. Dal lake (Ahmad et al. 2021), bacterial diversity of Roopkund glacier and glacial lake (Pradhan et al. 2010) profiling of bacterial community of Chilika lake (Pramanik et al. 2015), understanding the aquatic microbiome of the Indian Sundarban mangrove areas (Dhal et al. 2020), bacterial diversity of Yumthang hot spring, North Sikkim (Panda et al. 2017), prospecting bacterial diversity of cold deserts north western Himalayas (Chandratal lake, Dashair lake, Rohtang pass, and Beas-river) (Yadav et al. 2015) and prokaryotic diversity of tropical coastal sand dunes ecosystem (Shet and Garg 2021). Lakes also provide essential and valuable ecosystem services (raw water supplies, food, hydropower, tourism, climate regulation, etc.) to human existence and economies besides being a biodiversity hotspot (Sterner et al. 2020).

1.1. Loktak lake

Loktak lake is the largest freshwater wetland in the Bishnupur district of Manipur, India (National Wetland Atlas 2013). It is located in the north-eastern region of India (a biodiversity hotspot in the Indo-Burma region) (Meyers et al. 2000) and is rich in vast biological resources, yet the people of this zone are devoid of modern industrial

infrastructures. It has an area of 24,672 ha with a 12 km buffer area mainly consisting of agricultural land and some forest area (NWA 2013) and is located between 93°46' - 93°55' E and 24°25' - 24°42' N and occupies about 84 % of the total state area. The lake is oval-shaped with a length of 13 km and a width of 26 km, having a depth between 0.5-4.58 m (average depth of 2.7 m) (Laishram and Dey 2014) surrounded by mountainous ranges with steep slopes (Singh et al. 2010). The southern part of Loktak Lake wetland forms the Keibul Lamjao National Park, the world's only floating island. It is composed of a continuous mass of floating phumdis with an area of approx.40 sq. km (Prasad and Chhabra 2001) (Fig. 1a and 1b).

a)



b)

Fig. 1. **a)** Map showing Loktak lake, Manipur, India. (Source information adopted from Rai and Singh 2014).
b) Aerial view of Loktak lake. Source: Kim)

1.1.1. Conservation status

Wetlands are defined as "areas of marsh, bog, fen or water, whether natural or manmade, temporary or permanent, with still or moving, fresh, brackish or salt, including areas of seawater, whose depth at low tide does not exceed six meters (Ramsar Convention 2013). Wetlands are crucial and serve numerous vital functions like regulating the biogeochemical cycle, hydrological regimes, and sustaining human habitats to control groundwater recharging and alleviate flooding problems (Ramchander and Rahul 2021). According to the Ramsar Convention (1971), India is home to 75 Ramsar sites or wetlands designated to be of international importance covering an area of 13,26,677 ha (MoEFCC). Indian wetlands cover an entire range of ecosystems which includes high-altitude Himalayan lakes, flood plains of rivers, saline, arid and semi-arid regions, lagoons, backwaters and estuaries, mangrove swamps, coral reefs, marine wetlands, etc. (Prasad et al. 2002). North-east India rich in biodiversity accounts for four Ramsar sites i.e., the Deepor Beel in Assam, Rudra Sagar in Tripura, Pala in Mizoram, and the Loktak Lake in Manipur (Rai and Singh 2014). Loktak lake was recognised as a Wetland of International Importance (Ramsar site no.463) on 16th June, 1990.

1.1.2. Importance of Loktak lake

1.1.2.1. Ecological importance

1.1.2.1.1. Phumdi

Loktak lake is famous for its unique feature called floating islands which is of different sizes and thickness locally, known as Phumdis. Phumdis are the heterogenous mass of soil, vegetation, and organic matter at various stages of decomposition ranging in thickness from a few centimetres to 2-3 m (Kangabam et. al 2018). Phumdis float on lake

water with about $^{1}/_{5}$th of thickness above and $^{4}/_{5}$th under the water's surface. Phumdis appear in three distinct vertical zones stacked one above the other. The topmost root zone is 0-15 cm thick, followed by the mat zone of 25-65 cm and the lowermost peat zone of 0-25 cm (Takhelmayum and Gupta 2011). Phumdis are classified into three categories: i) Phumdi Live (PL) mostly, consisting of para-grass and plants like *Brachairia mutica, Alternathara* both live and dry para grass and plants and iii) Phumdi Dry (PD), a combination of plants, dead and dry para grass (Singh et al. 2003).

Phumdis play a vital role in the ecological process and functions of the lake (Fig. 2a). They act as a biological sink for the essential nutrients and regulate the lake's water quality and nutrient dynamics by absorbing the nutrients and accumulating them in their tissues (WISA and LDA 2002). Artificial floating islands or circular fish culture ponds commonly known as "Athaphums" not only create a habitat for fish and a platform for bird nesting but also enhance the water quality of the lake (Singh et al. 2011) (Fig. 2b) Phumdi is also home to various aquatic, semi-aquatic, and terrestrial plants, including one hundred twenty-eight species from forty-six families that shelter fifty-four fish species belonging to seventeen families (Singh and Singh 1994), aiding in pollution control (Meitei and Prasad 2015). A total of twenty-seven wild edible and twenty-five medicinal Phumdi species are associated with the locals living around the lake (Meitei and Prasad 2015). Phumdi plants are categorized according to how the local communities used the plant parts: plants for food, medicinal value, cultural significances, fodder and fuel, raw materials for handicrafts, athaphoom and house building material (Meitei and Prasad 2015). Phumdis also help in maintaining the lake's surface water temperature (Kangabam et al. 2018).

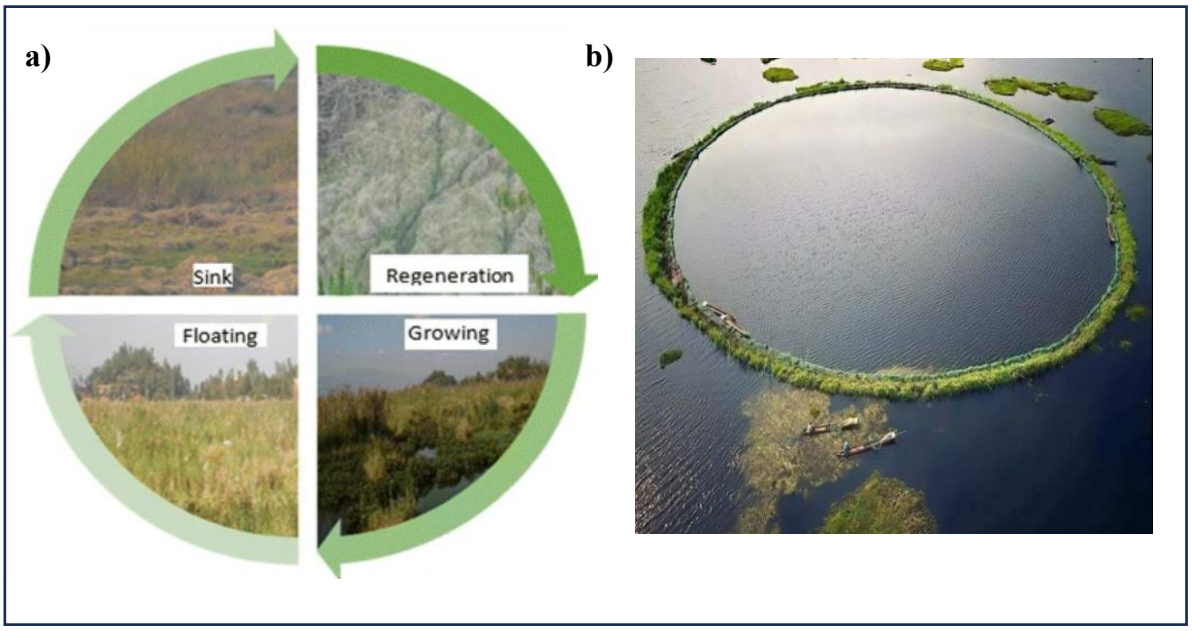


Fig. 2. a) Seasonal stages of Phumdis **b)** Circular fish ponds (Athaphums) (Source information adapted from Kangabam et al. 2017)

1.1.2.2.2. Socio-economic importance

Loktak lake is vital in maintaining the region's ecological, social, and economic security. The lake serves as a significant food source for the local people and urban hamlets that primarily depend on fishing and farming (Kosygin and Dharmendra 2009). It is also the only water source for the neighbourhood settlement (Laishram and Dey 2014). The Loktak Hydropower Project dam, constructed in 1983 as a part of the National Loktak Multipurpose Hydro-Electric project, facilitates the generation of power and irrigation (Singh and Khundrakpam 2011). This project is an essential source of energy for the entire North Eastern India. Loktak lake has been given the status "Lifeline of Manipur" due to its significance in socio-economic and cultural life (LDA 2002).

1.2.3. Biodiversity of the lake

Loktak lake is rich in biodiversity and supports numerous floras and faunas which include some of the globally threatened species. The lake was designated a Ramsar Site (a wetland site recognized to be of international importance) in 1990 due to its status as an ecological hotspot with a remarkable diversity of flora and fauna (NWA 2013). Keibul Lamjao National Park (KLNP), the world's only national park and last natural habitat of the Manipur brow- antlered deer *Rucervus eldii eldii* (Gray et al. 2015), locally known as Sangai is present in the southern part of the lake (Kangabam et al. 2017) (Fig. 3a). It is also the only habitat in India for a rare wild rice species Manchurian Wild Rice (*Zizania latifolia*) (NWA 2013). Loktak lake has a rich biodiversity which includes two hundred thirty-three macrophytes and four hundred twenty-five species of animals, including one hundred seventy-six invertebrates and two hundred forty-nine vertebrates. The invertebrates include one hundred fifty species of arthropods, sixteen species of annelids, and ten species of molluscs. The vertebrate consists of one hundred six bird species and thirty-two mammal species, six amphibian species, forty-three aquatic and semi-aquatic herbal medicinal plants, twenty-five species of amphibian flora, diversity of Cladocera, and identified fifty-one species belonging to twenty-one genera and seven families, one hundred eighty-six species of zooplankton. It also includes one hundred twenty rotifers belonging to thirty-six genera and nineteen families, the highest Rotifera ever found in an Indian subcontinent aquatic habitat, one hundred seventeen species of butterflies from seventy-nine genera and eighteen subfamilies, one hundred eighty-six species of fish. Lake. According to the IUCN Red List, *Osteobrama belangeri*, previously believed to be extinct in the wild, was discovered in the Loktak region. Loktak lake is also home to one hundred sixteen bird species, including twenty-one waterfowl migrating from the northern

hemisphere (Trishal and Manihar 2004). It also serves as an important breeding site for migratory and waterfowl birds. (Fig. 3b)

a) b)

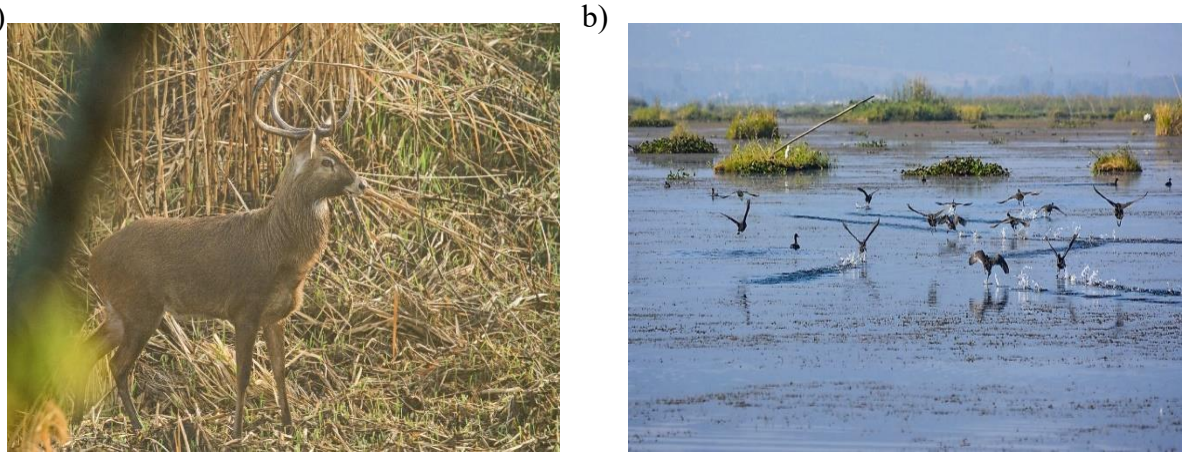


Fig. 3.a) Sangai deer (*Rucervus eldii eldii*) and b) Birds landing in Loktak lake
Source: http:// Kangla online, and
https://www.worldatlas.com/articles/what-is-a-phumdi.html

1.2.4. Current status of the lake

Loktak lake has undergone a substantial ecological transformation resulting from extensive anthropogenic activities in and around the lake. One of the main issues is the proliferation of Phumdis due to the construction of the Ithai barrage in 1983, which hinders the flushing of Phumdis, thereby, deteriorating the lake and increasing eutrophication in the open water area. It also leads to the thinning of Phumdis in the Keibul Lamjao National Park, which indirectly threatens the habitat of Sangai, where a population of only two hundred six had been reported according to the 2013 census (Kangabam et. al 2015). During the monsoon season (July-August), the Phumdis float on the lake's surface when the water level is high, and during the dry season (February-March), the Phumdis sink and

become attached to the lake bed, where they absorb nutrients essential for the growth of vegetation. However, the construction of the Ithai barrage also changed the life cycle of Phumdis or floating islands. Currently, due to the persistent storage of water in the park, Phumdis float throughout the year, including the winter season, which prevents the nutrient uptake by the Phumdis thereby reducing their thickness. This is a major concern as the thinning of Phumdis will submerge the highly endangered deer species Sangai (*Rucervus eldii eldii*) by increasing the chances of habitat destruction from its only habitat in the world (Kangabam et al. 2018). Urbanization, increased settlement, municipal waters deposited through the Imphal river, bathing, washing of clothes and utensils, fertilizers and pesticide run-off from the surrounding agricultural lands also contributed to the pollution of Loktak lake (Laishram and Dey 2014). Thus, the ecologically and economically important lake has been degrading in recent years, leading to its inclusion in the Montreux Record (a register of wetland sites on the Ramsar list that is threatened by technological developments, pollution, or human interference) in 1993, considering its deteriorating ecosystem (Roy and Majumder 2019).

1.2. Bacterial taxonomy

Bacterial taxonomy may be defined as the study of the diversity of microorganisms with the aim of organizing and prioritizing in an orderly manner (Truper 1991). It consists of three independent, but interrelated disciplines, namely classification, nomenclature and identification (Cowan 1965) which is done by following the enactment given in the Prokaryotic Code 1990 (revised in 2008) by the International Code of Nomenclature of Prokaryotes (Parker et al. 2019). Bacterial taxonomy provides details on species' identity and individuality in relation to its clinical, industrial and ecological niche. Taxonomy is one aspect of the broader field of systematics. Bacterial systematics and taxonomy are

frequently used synonymously with each-other, however, systematics is a much broader scope, which includes data on bacterial morphology, molecular biology and biochemistry, physiology, pathogenic potential, metabolic products, epidemiology, and ecological niches to characterize, arrange and classify bacteria (Moor et al. 2010).

1.2.1. Importance of bacterial taxonomy

Taxonomic tools are used to study microbial diversity, which is essential in the food industry, agriculture, bioremediation, bioprospection of enzymes, and antibiotics production (Helene et al. 2022). According to the present scenario, the phylogenetic analysis of the 16S rRNA gene resolves the novelty of the bacterial taxa, which correlates with the phenotypic, chemotaxonomic, and genomic studies. Bacterial taxonomy helps in the differentiation at the strain level, for example, *Shigella* strains have maintained a separate genus name for historical and medical reasons, although they belong to the subspecies of *Escherichia coli*. However, to avoid confusion in the medical community, it remains unchanged (Abram et al. 2021). To prevent any kind of taxonomic error, an organism should be carefully identified before it is used in any type of investigation. It also helps in the reclassification of existing bacterial genera to other taxonomic units (e.g., *Stenotrophomonas maltophilia* was first described as *Bacterium bookeri* (1943), later redesignated as *Pseudomonas maltophilia* (1958) and *Xanthomonas maltophilia* (1981); and finally, genus *Stenotrophomonas* was proposed in 1993). Additionally, many of the key databases that the scientific community and the general public today rely on the names ascribed to organisms as their primary entry point (Margos et al. 2020).

However, due to many limitations in the use of % similarity in 16S rRNA gene sequence analysis, DNA–DNA hybridization (DDH), and mol % of G + C content methods, polyphasic taxonomy is facing severe competition from the data produced by whole-

genome sequences. Microbial genomics taxonomy powered by next-generation sequencing provides quick and affordable approaches to access genomic information along with computational tools. It helps in species identification and classification. For example, Haley et al. 2010 used a genomic taxonomy approach for the description of the new species, *Vibrio meniscus* and *Vibrio paris* which are phylogenetically closely related to *Vibrio cholerae* and *Vibrio mimicus*, respectively. Microbial taxonomy also has an impact on medical microbiology as it can lose the visibility of certain organisms with small changes. For example, renaming *Streptococcus bovis*, a species linked to colon carcinoma, to *Streptococcus gallolyticus* (Wout and Bijlmar 2005). Such issues can be avoided by reporting the new name, followed by the old name in brackets, e.g., *Streptococcus gallolyticus* [*Streptococcus bovis*]. By developing standardized tools for taxonomic identification and classification, it has built a global platform for studying and disseminating reliable data about thousands of species across national boundaries, thereby reducing duplication. Thus, in light of advancements in molecular biology and data handling techniques that may aid in identifying microbes of ecological, medical, and industrial value, microbial taxonomy has thus seen a spectacular revolution in recent years (Priest and Goodfellow 2000).

1.2.2. Advancement in bacterial taxonomy

The history of bacterial systematics can be categorized into four distinct phases:

a) **1872 - 1900**

This is the phase of early descriptions based on chemical reactions, pathogenic potential, and morphology. Anton *van* Leeunwenhoek (1675) first

Cohn demonstrated the hierarchical designation of bacteria into genera and species using the paradigm proposed by Carl Linnaeus. Koch (1881) described

the plating technique which initiated the era of pure culture (Blevins and Bronze 2010), which forms the basis of bacterial classification (Logan 1994). The invention of petri dish improved the plating method by R.J. Petri.

b) **1900 - 1955**

In this phase, bacterial physiology and ecology were first explored and described. In 1923, the Society of American Bacteriologists (which later changed its name to the American Society for Microbiology) presented a report on the characterization and classification of bacterial types. This report served as the foundation for Bergey's Manual, the definitive work on bacterial taxonomy even today. In the late 1950s, numerical taxonomy emerged, aiming to create a standard technique for the classification of organisms (Sneath 1989).

c) **1955 - 1980**

In this era, many new approaches were developed. For instance, chemotaxonomy was integrated into species description. DNA-DNA hybridization was proposed by McCarthy and Bolton (1961) as a method of genetic material comparison (Mc Carthy et al. 1963). In the 1970s, Carl Woese discovered the existence of a third kingdom i.e., Archaea, besides Bacteria and Eukaryotes by cataloguing and comparing the partial 16S rRNA gene sequences (Woese 2004).

d) **1980 - till date**

This phase is known as the modern era, where advanced DNA techniques are incorporated into the species description. By the 1980s, the list of validly published species had reached 2500 (Skerman et al. 1980). Nucleic acid analyses, such as 16S rRNA sequence analysis, protein-encoding gene sequence

analysis, and gene profiling techniques, had impacted bacterial taxonomy highlighting the enormous amount of microbial diversity (Yarza P et al. 2014). The advent of Next Generation Sequencing (NGS) in recent years has allowed the rapid sequencing of genomes at affordable rates by directly generating *in-silico* values using the overall genome relatedness index (OGRI) (Carro et al. 2021). 66,303 16S rRNA gene sequences and 191,909 genomes have been sequenced as per the EzBioCloud database (Yoon et al. 2017), and a total of 2286 species have been validly published from 2021 to 2022 (https://lpsn.dsmz.de/statistics).

1.2.3. Approaches of Bacterial taxonomy

Methods for investigating microbiota from a given ecosystem can be either culture-dependent or culture-independent. It is impossible to understand the diversity of microorganisms without using various perspectives and analytical tools. Only with the aid of techniques and potent tools is it is able to distinguish between the morphology, physiology, and phylogenetics of microorganisms. These methods can be divided into two parts: i) Culture-dependent study: This study is based on distinct cellular, metabolic, morphological and physiological traits of microorganisms (Varjani et al. 2018). It also evaluates microbial metabolism and contributions to ecosystem function in ways that cannot be achieved using culture-independent approaches. For instance, the first bacterium known to perform complete nitrifications was discovered in *Nitrospira* cultures (Daims et al. 2015); a reversible TCA cycle was discovered in *Thermosulfidibacter takaii* which was not detected by metagenomics (Nunoura et al. 2018). Culture-dependent methods provide a clearer path to taxonomic resolution and functional assessments (Demko et al. 2021); ii) Culture-independent study: This study allowed the identification of hitherto unknown groups of species, such as the Candidate Phyla Radiation (CPR) bacteria, which have

different metabolisms and ecological roles (Brown et al. 2015), and the Asgard archaea (Zaremba-Niedzwiedzka et al. 2017). It provides more extensive diversity estimates and can also be used to determine metabolic requirements (Tripp et al. 2008). Culture-independent approaches helped clarify potential microbial roles; however, culture-based studies are still required for understanding microbial characteristics and phenotypes (Naha et al. 2018). The two types of culture-based studies are discussed below:

1.2.3.1. Culture-independent study

1.2.3.1.1. Metagenome analysis

Bacteria carry out numerous important biogeochemical activities that give them the potential to alter and regulate water quality in aquatic habitats. However, our understanding of their community interactions, functional potential and genetic variability is still restricted, the reason being due to the difficulty in culturing the abundant lake bacteria (Newton et al. 2011). Culture-based methods have been shown to be insufficient for estimating the variety of microbial variety since majority of the microorganisms are still difficult to cultivate (Moreira et al. 2011). To overcome the traditional methods, metagenomics now offers a feasible tool for high-throughput environment community monitoring (Sogin et al. 2006) and has made it possible to explore the diversity of microorganisms in unexplored areas and enabled a better understanding of the extreme and uncharted realms. (Kirk et al. 2004; Cowan et al. 2015). Metagenomics helps in processing genetic information of organisms that are challenging to isolate into pure cultures or that have never been cultivated previously (Simon and Daniel 2011). Numerous metagenomics studies have been conducted from different environments and some of the recently reported studies include the psychrophilic Pangong Lake (Rathour et al. 2017), Upper Mississippi River (Minnesota) (Staley et al. 2014), Amazon Basin freshwater lakes (Toyama et al.

2016), marine waters (Brown et al. 2009), wastewater (Sanapareddy et al. 2009), soils (Jones et al. 2009), contaminated water (Das et al. 2017). Metagenomics is defined as the genetic analysis of the genome from microbial populations in environmental samples without the necessity for prior cultivation (Oulas et al. 2015). It is based on the principle of directly isolating DNA from a complex environmental sample containing a variety of microbiota in order to determine the true microbial composition of that environment (Ahmad et al. 2019). The term "metagenomics" was used first by Handelsman in 1998. The Next Generation Sequencing (NGS) made this metagenomic research more accessible via targeted metagenomics i.e.,16S amplicon sequencing (Amrane and Lagier 2018). NGS has provided new insights into the structure and functions of the microbiome (Jiang et al. 2019; Ning and Tong 2019). This technique has been used to identify extremely rare populations that are present in very low numbers and may be crucial for the functional diversity and stability of ecosystems (Sogin et al. 2006; Staley et al. 2013). Functional metagenomics studies have potential applications in biotechnological applications, food, and pharmaceutical industries for discovering novel bioactive compounds, including antimicrobials and enzymes capable of catalyzing novel reactions (Coughlan et al. 2015). The two main approaches for microbiome analysis are 16S rRNA gene amplicon-based and shotgun metagenomics.

1.2.3.1.1.1. Metagenomics approach

1.2.3.1.1.1.1 Amplicon based

Amplicon based approach is straightforward and cost-effective for profiling the taxonomic composition of microbial communities. It is also applicable to detect rare taxa (Peterson et al. 2021, Stefanini and Cavalieri 2018). In this approach, DNA is extracted, and a specific region (V3-V4) of the highly conserved bacterial gene, like the rRNA gene

is amplified and sequenced (Oulas et al. 2015). 16S rRNA gene (Liu et al. 2007), three ITS regions in fungi: ITS1-5.8S rRNA-ITS2, 26S rRNA gene, and 18S rRNA gene (Xu 2016) are currently used for amplicon-based metagenomics analysis as these regions are suitable for species differentiation. These genes are identified based on several curated databases: Greengenes, SILVA, RDP, and LTP (Santamaria et al. 2012). A series of steps needs to be checked once the amplicon-based 16S rRNA are sequenced, which includes trimming of low-quality bases, matching and stitching paired reads, removal of chimeras, filtering of contaminated sequences, identification and assigning taxonomic OTUs (Callahan et al. 2016; Caporaso et al. 2010; Schloss et al. 2009). Mothur is one of the frequently used software for microbial population analysis (Schloss et al. 2009). The prediction of the functional profile of a given population is also possible with tools like PICRUST (Langille et al. 2013) based on the use of the Greengenes reference database (DeSantis et al. 2006) and the functional composition of the reference genomes described in IMG (Markowitz et al. 2012). The annotated gene counts per sample are linked to Kyoto Encyclopedia of genes and genomes (KEGG) orthology (KO) accession numbers (Kanehisa et al. 2004), and the functions are inferred.

1.2.3.1.1.1.2. Shotgun-based

This approach analyses untargeted sequences of all the microbial genomes, including the unculturable microorganisms present in a sample. Shotgun metagenomics sequencing can be used for profiling the taxonomic composition and functional annotations to characterize the biological functions associated with microbial communities (Quince et al. 2017). Shotgun metagenomics studies involve the following steps: i) collection, processing, and sequencing of samples; ii) pre-processing of the sequencing read; iii) sequence analysis for profiling taxonomic, functional, and genomic attributes iv) post-

processing analysis and validation (Quince et al. 2017). Contigs can be assembled in two routes, either by the reference-based assembly (using one or more genomes as a map) or *de novo* – based assembly (based on computational power). Binning is an integral part of the assembly where the contigs are grouped into individual genomes and assigned the groups to species pr genus either by the distribution of k-mer sequences or by homology-based binning (Oulas et al. 2015). Binning plays a crucial role in metagenomics analysis, where it may provide insight into the presence of novel genomes and different numbers and types of taxa in the community (Sharpton 2014).

1.2.3.1.2. Sequencing Platforms

The development of Sequencing techniques has provided insight into complex microbial environments, and several genomes have been published. Frederick Sanger followed the first method of DNA sequencing by Maxam and Gilbert in 1977 for the discovery of the chain termination method. Sanger sequencing has been used for over three decades (Meera et al. 2019). The advent of new and efficient detecting tools, the miniaturization of available technologies, and the emergence of various sequencing technologies have made sequencing, resequencing, and data comparison much faster (Nowrousian 2010). Different sequencing platforms are available, which are listed below:

i) First-generation sequencing

Sanger sequencing (1977) based on the chain termination method helped better understand cellular systems and diseases. This method has the advantage of not requiring cloning before sequencing and using unique primers for each reaction (Smith et al.1986). However, some of the drawbacks of this approach include difficulty in addressing various applications of genomics, sequence quality, labour intensiveness, amount, and rate of data creation.

ii) Second-generation sequencing (Next generation techniques)

The various platforms used for sequencing include a) Roche/454 sequencing, b) Ion torrent: proton/PGM sequencing, c) Illumina sequencing, and d) Solid sequencing. NGS technology is capable of generating high-throughput sequence data (Qiang-long et al. 2014), is fast and economical, has a wide range of detection, and provides discreet results (Meera et al. 2019).

iii) Third-generation sequencing

Third-generation sequencing technologies include a) Pacific Biosciences (PacBio) Single Molecule Real Time (SMRT) sequencing, b) the Illumina Tru-seq Synthetic Long-Read technology, and c) the Oxford Nanopore Technologies sequencing platform. New insights into evolution and diversity have been made possible by third-generation technologies as they produce highly precise *de novo* microbial genomic assemblies (Koren et al. 2013, Loman et al. 2015) and the reconstruction of dozens of plants and animal genomes (Berlin et al. 2015).

1.2.3.2. Culture-dependent study

1.2.3.2.1. Polyphasic taxonomy

The first attempt at microbial classification by Schildkraur et al. 1962 based on single-stranded DNA paved the way for the development of polyphasic classification. Colwell 1970, coined the term polyphasic taxonomic which refers to the integration of genotypic, chemotypic, and phenotypic information for the delineation of taxa. Due to the high sequence conservation of the 16S rRNA gene, the delineation of taxa above genus does not have a clear definition. In contrast, species and subspecies level delineation were accommodated by chemotaxonomic techniques and recently with genomic sequence

comparison (Kampfer and Glaeser 2012, Konstantinidis and Tiedje 2005). Advances in polyphasic approach have become an important tool in microbial systematics for classifying bacteria into new genera and species previously placed into invalid taxa. The following criteria for describing valid taxa are discussed below: a) genotypic, b) phenotypic, and c) chemotaxonomic characterization.

1.2.3.2.1.1. Genotypic analysis

1.2.3.2.1.1.1. Marker gene-based

1.2.3.2.1.1.1.1. 16S rRNA

The 16S rRNA gene, which has a length of 1500bp, is composed of nine variable regions (V-regions), V1 to V9 is interspersed between a highly conserved region, which enables the use of universal primers to study genes across the bacterial and archaeal domains (Woese et al. 1990). Stackebrandt and Goebel in 1994 recommended 97 % of 16S rRNA gene sequence identity for species delineation, which led to the abolishment of the DNA-DNA hybridization as the DDH values were never higher than 70 % (Stackebrandt and Goebel 1994). However, this value was re-evaluated in 2005 (Stackebrandt and Ebers 2006), where rather than 97 %, a 16S rRNA gene sequence similarity of 98.7 % should be defined at which DDH experiments are required to determine the genomic uniqueness (Stackebrant and Erko 2011). The standardized threshold for prokaryotic lineages based on 16S rRNA gene sequence identities at the genus, family, order, class, and phylum is 94.5 %, 86.5 %, 82 %, 78.5 %, and 75 % (Yarza et al. 2014). Despite being extremely helpful for classifying bacteria, 16S rRNA gene sequencing has weak phylogenetic power at the species level and weak discriminatory power for some genera (Mignard 2006).

1.2.3.2.1.1.1.2. Multilocus Sequence Analysis (MLSA)

MLSA was introduced by Maiden et al. in 1998 for epidemiology and population genetic research of pathogenic bacteria (Maiden et al. 1998), although it offered a fresh perspective to the understanding of inter- and intra-specific genomic relatedness (Stackebrandt et al. 2002). MLSA is based on multilocus sequence typing (MLST), where phylogenetic trees are constructed using partial sequences of genes that code for proteins with conserved functions (housekeeping genes), and phylogenies are then inferred (Maiden 2006). According to Schleifer 2009 and Gevers et al. 2005, MLSA can replace DDH for species delineation. MLSA was first evaluated in 107 isolates of *Neisseria meningitidis*, where they analysed 11 housekeeping genes with a fragment size of ~470bp (Maiden 1998). Applications of MLSA for species delineation have been applied in *Burkholderia cepacian* where a difference of 3 % for concatenated gene sequences were observed, which is equivalent to 70 % DDH threshold (Vandamme 1997), genus *Streptomyces* based on five housekeeping genes (*atpD, gyrB, recA, rpoB, trpB*) (Rong and Huang 2014), genus *Pseudomonas* based on concatenated gene sequences of 16S rRNA gene with three protein-coding genes, *gyrB*, *rpoB* and *rpoD* and several other taxa (Mulet et al. 2012).

1.2.3.2.1.1.2. Genome-based

Topological variation frequently occurs from single gene-based phylogenetic analyses (Castresana 2007, Delsuc et al. 2005). As a result, the definition of taxa produced significant gaps. To overcome this problem, taxonomists suggested phylogenomics as the best method for examining evolutionary relationship among taxa (Chan et al. 2012; Mateo-estrada et al. 2019). Phylogenomics is the inference of a genome-based phylogenetic tree from a set of core genes (Eisen and Fraser 2003). Based on the genomic data availability, research groups have identified core genes ranging from 64-104 for constructing the phylogenomic tree through software like UBCG (Na et al. 2018), bcgTree (Ankenbrand

and Keller 2016), PhyloSif (Darling et al. 2014). Genome-based classification provides better resolution than 16S rRNA gene because of the more significant proportion of genome being used in comparison, which improves the phylogenetic signal (Johnson 2019). Recently, classification methods that make use of genome sequences include similarity measures between pairs of genomes either at the level of encoded proteins (average amino acid identity) (Konstantinidis 2005) or nucleotides (average nucleotide identity) (Goris et al. 2007) and digital DNA–DNA hybridization (Auch et al. 2010) are used primarily for defining and identifying species

1.2.3.2.1.1.2.1. *In- silico/Digital* DNA-DNA Hybridization

DNA-DNA Hybridization (DDH) is a wet lab method based on the denaturation and renaturation capability of DNA between two genomes. DDH similarity of below 70 % has been recommended for species delineation (Wayne et al. v1987), and it can be omitted if the 16S rRNA gene sequence similarity is more than 98.7 % (Stackebrandt 1994). However, the drawbacks of DDH were tedious, laborious, and error-prone (Klenk and Goker 2010). With the advancement in genome sequencing technology, scientists have replaced the wet lab DDH with computation based with *In- silico/Digital* DNA-DNA Hybridization. This method is resistant to error even in the presence of many paralogous genes, significant repeats, and shortened genomes (Henz 2005).

1.2.3.2.1.1.2.2. G+C content (mol%)

G+C content is the amount of cytosines and guanines in the genome relative to the total number of nucleotides (Rossello-Mora and Amann 2001). It is one of the most used features in the taxonomic delineation of taxa (Meier-Kolthoff et al. 2014) and was first proposed by Sueoka (1962). Different methods, such as traditional (thermal denaturation,

melting profiles, HPLC, RT-PCR) and *In-silico* (genome-based), have been used to analyze the variables that could affect the G+C content (Meier-Kolthoff et al. 2014). The variation in G+C within a species can range from 3-5 %. However, Kolthoff et al. 2014 showed that G+C content, if computed from genome sequences, fluctuates no more than 1 % within species.

1.2.3.2.1.1.2.3 Average Nucleotide identity (ANI)

ANI is a computational tool that measures the genomic relatedness between strains for replacing DNA-DNA Hybridization (DDH) (Konstantinidis and Tiedje 2005). It also serves as a tool for demarcating species of Archaea and bacteria. In this method, the query genome (fragmented into 1020 bp) is searched against the subject genome where at least 70 % of the length is aligned with more than 30 % identity (Goris et al. 2007). Software tools for ANI calculation include BLASTN (ANIb) (OrthoANIb), MUMmer (ANIm), and USEARCH (OrthoANIu) (Lee et al. 2016). While ANIm and OrthoANIu process was faster among the ANI calculating tools, ANIb and orthoANIu were the best in terms of accuracy (Yoon et al. 2017)

1.2.3.2.1.1.2.4. Average amino acid identity (AAI)

Average amino acid identity (AAI) measures pairwise genomic relatedness because resolution gradually decreases at the nucleotide level (Rodriguez and Konstantinidis 2014). AAI values are also effective at predicting the similarity of functional gene content (Konstantiidis and Tiedje 2005). AAI offers a significantly higher resolution than ANI in situations where the species average ANI score is 80 % or less and they share less than or equal to 30 % of the gene content. According to the suggestions of Goris et al. (2007) and Konstantinidis et al. (2017), two strains belong to separate genera if their AAI scores range

from 45 to 65 % and different families if their AAI scores fall below 45 %. However, according to recent findings, AAI values belonging to different genera that are related to one another might range between 60-80 % (Luo et al. 2014; Orata et al. 2018).

1.2.3.2.1.1.2.5. Percentage of conserved proteins (POCP)

The proportion of conserved proteins (POCP) between two strains is used to evaluate their evolutionary and phenotypic distance since the average nucleotide sequence identity between two genomes is insufficient for genus demarcation in prokaryotic species. Two strains of the same genus have POCP values of more than 50 %. The POCP value can range between 0 % to 100 % depending on how closely the two genomes share the protein contents (Qin et al. 2014). However, recent studies showed the 50 % POCP threshold is not suitable for delineating genera (Orata et al. 2018). Another tool that has been suggested for restricting taxonomic at the genus level is AAI (Konstantinidis and Tiedgje 2005). AAI values of 60-80 % and 40-69 % POCP are presently considered the genus boundaries (Aliyu et al. 2016). With the advancement of high-throughput bacterial genome sequencing (Loman et al. 2012), prokaryotic genus demarcation will advance significantly with the incorporation of whole-genome data into the present taxonomy system (Qin et al. 2014).

1.2.3.2.1.12.6. Pan-genome analysis

Pan-genome is the repertoire of orthologous and distinct genes present in certain groups of organisms (Costa el al. 2020). Pan-genome is categorized into core (collection of gene families shared by all genomes) and accessory (set of genes present in one or more genomes but not all of them) (Mclnerney et al. 2017). Core genomes are necessary for clade development, like metabolic pathways and genetic information, whereas accessory genomes are crucial for understanding variances in the clade genomes and evolutionary

paths (Tettelin et al. 2005). Pan-genome analysis plays a significant role in phylogenetic analysis and species evolution (Livingstone et al. 2018; Gonzales-Siles et al. 2020). It also plays important role in pathogenicity, vaccine development, and adaptability of bacteria (Livingstone et al. 2018; Gonzales-Siles et al. 2020). Pan-genome can be open (the number of gene families increases with the addition of new genomes) or closed (the number of gene families remains constant) (Costa et al. 2020). Online tools for pan-genome analysis include BPGA (https://sourceforge.net/projects/bpgatool/) (Chaudhari et al. 2016) PGAPX (http://pgaweb.vlcc.cn/pgapx/) (Zhao et al. 2018) and PGAweb (http://pgaweb.vlcc.cn) (Chen et al. 2018).

1.2.3.2.1.1.2.7. Genome annotation

Genome annotation identifies functional elements along a genome's nucleotide sequence (Joseph and Sergi 2019). Genome annotation consists of two main components, i.e., structural annotation and functional annotation (Eijigu and Jung 2020). Structural annotation consists of introns, exons, promoters, pseudogenes, RNA genes (rRNA, tRNA, miRNA, snRNA and snoRNA, CRISPR/Cas, etc. (Boburque et al. 2018). Gene prediction comprises three approaches: *ab inito*, homology, and combined methods. *Ab inito* approach identifies coding or noncoding regions, promoters, intron-exon regions based on a statistical model (hidden Markov model). Homology-based method uses discovered similarities to align the sequence with expressed sequence tags (EST), complementary DNA (cDNA), or protein evidence and then uses those results to predict genes. The combined method comprises *ab anito* and homology base approaches for gene prediction. Several nucleotide and protein sequence databases for structural annotation include GenBank (Sayers et al. 2019), European Nucleotide Archive (ENA) (Brooksbank et al.

2014), DNA Databank of Japan (DDBJ) (Kodama et al. 2018), UniProt, etc. (Consortium 2019).

Functional annotation is the association of biological data with gene or protein sequences discovered through structural annotation (Eijugu and Jung 2020). It is achieved by using alignment tools like BLAST, which searches a protein database for high-scoring alignments. Many databases, including Kyoto Encyclopedia of Genes and Genomes (KEEG) (Kanehisa and Goto 2000) and Gene Oncology (GO) (Botstein et al. 2000) are used widely for gene function. Structural pipelines such as MAKER2 (Holt et al. 2011), CAT (Fiddes et al. 2018), NCBI Eukaryotic Annotation Pipeline (Thibaud-Nissen et al. 2013), BRAKER1 (Hoff et al. 2016) and functional pipelines like Prokka, Rapid Annotation using Subsystems Technology (RAST) are used for the annotation of prokaryotic genomes.

1.2.3.2.1.2. Phenotypic analysis

The first modern attempt to systematically classify bacteria based on their phenotypic properties began with the first edition of Bergey's Manual of Determinative Bacteriology in 1923, which categorized bacteria into a nested hierarchical classification to indicate different levels of relatedness (Hugenholtz et al. 2021). Phenotypic studies help in the description of taxa, from subspecies to the family level (Garrity 2016). environmental factors have a significant impact on how genes are expressed as phenotypes; therefore, a comparative analysis 1of the novel taxa along with their type strains should be studied under similar growth conditions (Tindall et al. 2010). Phenotypic studies include morphological and metabolic (physiological and biochemical features) characterizations discussed below.

1.2.3.2.1.2.1. Morphological characterization

Morphological traits analyzed and applied for taxonomic studies include colony morphology like shape, size, colour, gram-staining properties, sporulation, flagellation, etc.

1.2.3.2.1.2.2 Metabolic characterization

Metabolic characterization includes both physiological and biochemical features. It includes observing the organism's growth at different substrates like carbon, nitrogen, vitamins, and various growth conditions such as temperature, salinity, pH, oxygen conditions (aerobic or anaerobic), enzyme activities, antibiotic susceptibility, etc. (Tindall et al. 2010).

1.2.3.2.1.3. Chemotaxonomic analysis

1.2.3.2.1.3.1. Polar lipids

Polar lipids are amphiphilic lipids characterized by two hydrophobic tails and a hydrophilic, polar head group (Anto et al. 2020). Some of the primary functions of polar lipids includes maintaining cell homeostasis, signalling, and other developmental process (Bale et al. 2019). Glycerophospholipids are the main components of cell membranes which include phosphatidylcholine, phosphatidylethanolamine, phosphatidylinositol, and phosphatidylserine derived from *de-novo* synthesis (FA glycerol backbone) (Langelier et al. 2010). Non-phosphorous membrane polar lipids include glycolipids and amino lipids (Dugail et al. 2017). Polar lipids analysis has been a powerful tool in chemotaxonomy initiated by Asselineau, O'Leary and Kates (Lechevalier and Moss 1977). Several instances of resolving taxonomic status utilizing polar lipids as chemotaxonomic makers include reassigning some *Rhodobium* members to a new genus (Urdiain et al. 2008) and merging the three *Krokinobacter* members into the genus *Dokdonia* (Yoon et al. 2005).

1.2.3.2.1.3.2. Cellular Fatty acid analysis

Cellular fatty acids are carboxylic acids with saturated/unsaturated aliphatic chains which help maintain membrane stability and functionality (Singer 1972). It also serves as cell storage material and is involved in cell signalling (de Carvalho and Caramujo 2018). Under severe environmental conditions, microorganisms maintain their biological functions by altering their protein, hopanoid, sterol, and carotenoid contents by altering the lipid composition of their cellular membranes (de Carvalho and Caramujo 2018). Fatty acids have been extensively used to characterize bacterial taxa because of their variation in chain length, double-bond location, and substituent group (Suzuki et al. 1993). FAs between 9 to 20 carbon atoms, saturated/unsaturated, hydroxy FA, branched FA (iso, aneiso, methylated), cyclo-FA are used for the characterization of genera and species of bacteria by fatty acid methyl ester (FAME) analysis (Sasser 1990; Kunitsky et al. 2006). FAME analysis was developed by Sasser using a technique called Sherlock microbial identification system (MIDI Inc.), which is a fully automated gas chromatography. MIDI Inc quantifies fatty acid based on retention time which hydrolyses fatty acid and undergoes methylation to produce methyl esters (Kunitsky et al. 2006). It was initially developed to identify fatty acids in bacteria; however, it is currently used to identify fatty acids in fungi and some nematodes along with bacteria (Sekora et al. 2009).

1.2.3.2.1.3.3. Isoprenoid quinones

Isoprenoid quinones occur in the cytoplasmic membranes of proteobacteria and eukaryotes. They are composed of a quinone head group (naptho or benzo ring) to which a polyprenyl tail (isoprene units) is attached. They are an essential part of electron transport chain and oxidative phosphorylation. Based on the nature of head group and their midpoint redox potential, quinones are divided into naphthoquinones (menaquinone and

phylloquinone) and benzoquinones (ubiquinone, plastoquinone and rhodoquinone) (Abby et al. 2020). UQ and PQ are considered high potential quinone, whereas MK and RQ are low-potential quinones (Schoepp-Cothenet et al. 2013). Quinones are used as taxonomic markers to distinguish bacteria at different taxonomic levels due to their large variability side chains such as saturation, different lengths, and hydrogenation (Collins and Jones 1981).

1.2.3.2.1.3.4. Cell wall amino acid

Bacterial cell wall is made up of peptidoglycan, a huge macromolecule containing acylated amino sugars and three to six different amino responsible for the rigidity of the cell wall (Warth and Strominger 1971; Strominger et al. 1967). Peptidoglycan is built out of glycan strands (alternating *N*-acetylglucosamine and *N*-acetylmuramic acid linked by b-1-4 bonds) crossed linked through shot peptides (Ghuysen 1968). Schleifer and Kandler (1972) used a tri-digital approach to classify peptidoglycans and then demonstrated the relationship between taxonomic categories and peptidoglycan types. At the generic level, the composition of the cell wall amino acid aids in the taxonomic classification of gram-positive bacteria (Yokota 1992). According to Takeuchi and Yakota 1989, the distribution of polysaccharides differs among coryneform bacteria, which can serve as a taxonomic marker among these bacteria for classification at the generic level. However, a comparison of the sugar contents of cell-wall polymers among the different genera showed that no specific sugar profiles existed characteristics of the genus.

1.2.3.2.1.3.5. Polyamines

Polyamines are primordial polycations found in eukaryotes, bacteria, and archaea. They are primarily recognised for their essential role in cell growth and proliferation,

biofilm formation, cellular translation, and biosynthesis of natural products. Polyamines are mainly linear, flexible aliphatic chains with two or more amine groups which include diamines 2 1,4-diaminobutane (putrescine, Put), 1,3-diaminopropane (Dap) and 1,5-diaminopentane (cadaverine, Cad), triamines spermidine (Spd), sym-norspermidine (Nspd) and sym-homospermidine (Hspd), the tetraamines norspermine (Nspm), thermospermine (Tspm) and spermine (Spm), the uncommon triamines aminobutylcadaverine and aminopropylcadaverine and the uncommon tetraamine aminopropyl homospermidine (Michael 2016). Polyamine patterning has been used for distinguishing and defining taxonomic relatedness at the species and above the genus level (Vandammme 1996). The α, β, γ, and δ subclasses of *Proteobacteria* are divided based on the absence or presence of norspermidine, homospermidine or spermidine; 2-OH putrescine or diaminopropane. Spd is the most prevalent triamine; however, several bacteria from different phyla only produce Hspd, while considerably fewer bacteria only generate Nspd (Hamana and Matsuzuki 1992). Put is the most common diamine in bacteria; however, cad is widely present in proteobacteria. Gram-positive bacteria like *Bacillus* contains putrescine and spermidine (Michael 2016).

1.3. Definition of the problem

Falling into the Indo-Burma hotspot region, Loktak lake, located in Manipur, Northeast India, is a promising area for microbial diversity studies. This region has a diverse ecological niche which serve as a home for several endemic species and also provide unique and ecological niches for the evolution of microorganisms due to its varying physiography (Myers et al. 2000). However, due to its rigorous and undulating terrain, microbial resources of this region have been poorly explored and are yet to be known to the scientific world (Giri et al. 2022). Only a few studies based on the microbial taxonomy

of Loktak lake have been explored and it is still assumed that extensive and in-depth studies on the microbial taxonomy of the lake is still less known. Numerous limnological studies focusing on biodiversity and management of Loktak are reported (Singh and Singh 1994; Trishal and Manihar 2004; Singh et al. 2013; Meitei and Prasad 2015), but no significant work has been carried out on the bacterial diversity of the lake. Previous studies on Loktak lake includes the culture-independent whole genome shotgun sequencing based metagenomic approach by Puranik et al. (2016) where they explore the favoured microbial community associated with Phumdi in Loktak lake. Another study was taken up by Kangabam et al. (2020) which provides insight into the seasonal profile of bacterial metagenomes to understand variation in microbial communities across the land use in Loktak lake. In our study, a metagenomic approach based on 16S rRNA gene amplicon sequencing was employed to provide insights into the diversity of bacterial communities of Loktak lake. Our study aims at identification and classification of the microbial communities of Loktak lake and the associated functional potential of this habitat, which will serve as base data on the census of the lake microbiome. This study could be of great importance in understanding their role in various industrial, agricultural, and medicinal applications and could serve as a habitat for many economically important bacteria. Cataloguing the microbial diversity of Loktak Lake needs an in-depth study on the microbial taxonomy which was focused in this thesis with the following objectives:

- To stud the bacteria diversity of Lokta lake using 16S rRNA gene amplicon-based metagenomic analysis
- To study the phylogenomic and polyphasic characterization of new taxa, if any

MATERIALS AND METHODS

2. MATERIALS AND METHODS

2.1. Laboratory wares

Erlenmeyer flasks, Petri plates, round bottom flasks, test tubes, beakers, measuring cylinders, screw caps, slides, test tubes, storage bottles, spreaders, pipettes and PCR tubes were used as laboratory glassware and plastic wares manufactured by various companies (Borosil, SCHOTT Duran®, Tarson, Riviera, Lab systems, Eppendorf, Thermo Fisher, and Anumbra).

2.1.1. Distilled, double-distilled, and milli-Q water

For routine media preparations and buffer solutions, distilled and double-distilled were utilized. For DNA/RNA processing and TLC, Milli Q water was used, obtained from the School of Chemistry and School of Life Sciences, University of Hyderabad.

2.1.2. Chemical, solvents and laboratory kits

Analytical grade chemicals and solvents from Sigma-Aldrich, HiMedia, Qualigens, SR LIFESCIENCES, Amresco, Merck, and Thermo Fisher were procured and used. Ready-to-use kits were used for DNA extraction (Nucleopore Fungus Bacteria Kit, Genetix brand) and gel purification (SureExtract PCR kit, Genetix brand).

2.2. Buffers, standard solution and Determination of pH

Buffer solution and standard solutions were prepared using double distilled water and Milli-Q water. Phosphate buffer (KH_2PO_4 / K_2HPO_4) and bicarbonate buffer ($NaHCO_3$ /NaOH) at various pH were prepared and sterilized by autoclaving 1 ml of 1M Tris-HCl (pH 8), 0.2 ml of 0.5M EDTA (pH 8) with a volume made up to 100 ml and 50X Tris Acetate-EDTA (TAE) buffer [24.2 g Tris-base, 5.71 ml glacial acetic acid, 10 ml 0.5 mM EDTA (pH 8), final volume 100 ml]. 20X Sodium saline citrate (SSC) buffer was prepared using 75 g of NaCl, 0.77 g of sodium citrate, and a final volume of 100 ml (pH 7, adjusted

with HCl). The pH of the buffer and growth medium was checked using pH strips (Hi Media make) and a digital pH meter (Digisun Electronics D1-707).

2.3. Sterilization

The necessary glassware was first washed with detergent (Teepol solution) and immersed in the chromic acid solution for an entire night. Sterilization of wares, buffers and growth media was done by autoclaving (SA52, Ketan) at 121° C, 15 psi for 15 min manufactured by LTE scientific. 0.2 µm acrodisc syringe filter manufactured by Pall Life Sciences and Millex was used for filter sterilization of heat-sensitive substances such as antibiotics, glucose, and NAG solution.

2.4. Sample collection

Samples (soil sediments, rhizosphere soils of plants in and around Phumdis) for the study were collected from various locations of Loktak lake: east, west, north and south (Fig. 4). Human activity, directly and indirectly, impacts all sample locations, such as tourists, household waste, aquaculture and hydroelectricity project. The sampling sites' pH, temperature and salinity were measured using a pH meter, thermometer and conductivity meter. The GPS locations of the sample sites were also recorded during sampling. On-site sample inoculation for enrichment was done in a 30 ml serum vial containing *Planctomycetota*-specific media. The samples were also preserved in formalin for a metagenome-based study, and the remaining samples were transported to the lab and refrigerated at 4 ˚C for further analysis.

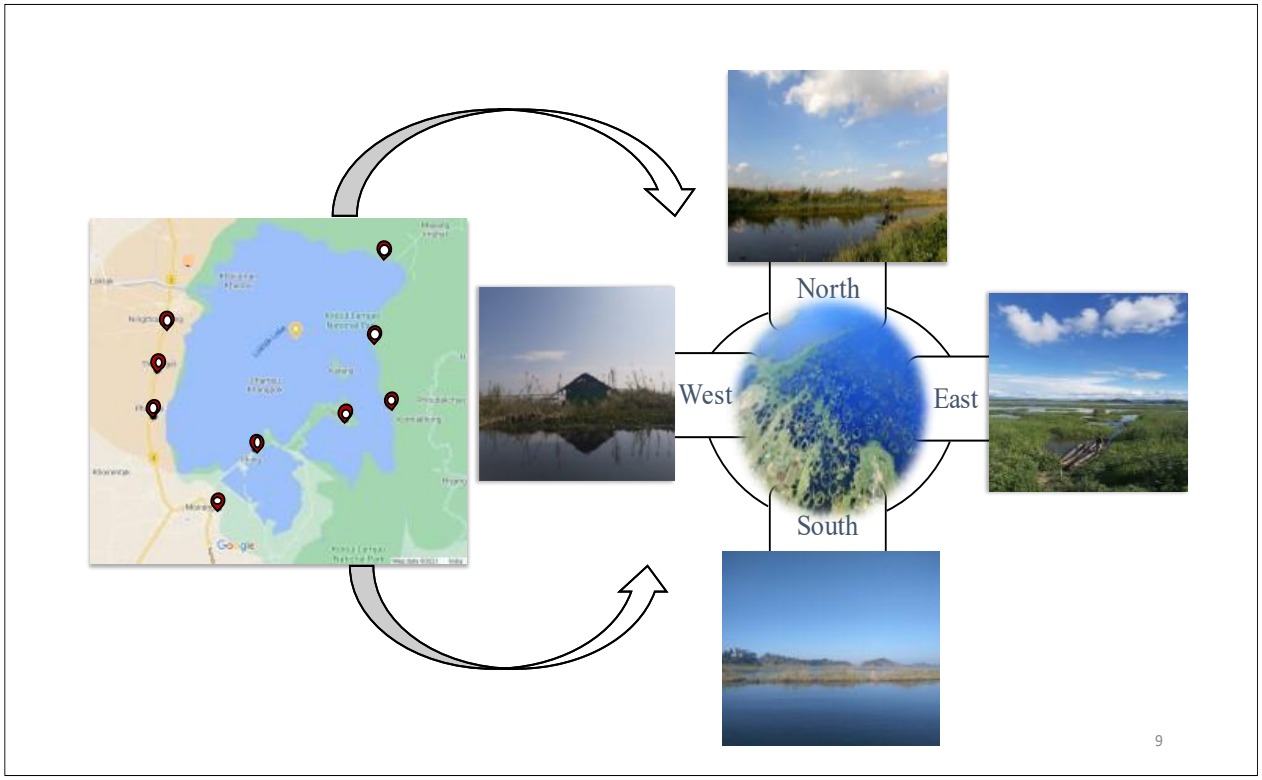


Fig. 4. Sampling sites of Loktak lake

2.5. Amplicon-based metagenomic analysis

2.5.1. DNA extraction from soil samples

1g of soil sample collected from the four sampling sites was used to extract metagenomic DNA using MO BIO's Power Soil DNA Isolation kit (QIAGEN) following the manufacturer's instructions. The concentration of the DNA of the four samples was checked by loading 1µl of extracted DNA in nanodrop to determine the A_{260}/A_{280} ratio. The DNA samples were analysed using 1 % (w/v) agarose gel and visualized under a UV transilluminator (Universal Hood II, BIO-RAD). The extracted DNA was preserved at -20 ˚C until further analysis.

2.5.2. Library preparation and Illumina MiSeq sequencing of the 16S rRNA gene

DNA samples were outsourced to Research and Testing Laboratory (RTL) LLC (Lubbock, TX, USA) and Eurofins, Bangalore, India for Illumina MiSeq (2 x 300 bp) sequencing as per the 16S metagenomic sequencing library preparation code. The V1-V3 regions of DNA were amplified using specific primers, which were F-V1-V3 (F-5'AGAGTTTGATCMTGGCTCAG3') and R-V1-V3 (R-5' ATTACCGCCGGCTGTGG 3').

2.5.3. Downstream analysis and taxonomic assessment

The raw sequence reads generated was uploaded to the NCBI SRA database via the FTP server by creating a bio-project at the NCBI portal. Following the acceptance of the bio-project, a bio-sample was made by filling in the samples' relative information. NCBI accession numbers were assigned to all the sequencing data. The fastq generated from sequencing was further processed using mothur software, an open-source bioinformatics pipeline (Schloss et al. 2009). The sequences were stitched into single-end reads/contigs and then screened. Sequence length less than 300 bp, homopolymers longer than 8 bp, unclear base calls, and incorrect primers were eliminated. All the common sequences were binned to generate a set of distinct sequences to enable counting in each sample, followed by aligning unique sequences (V1-V3 regions) against the SILVA database (Quast et al. 2013). Following alignment filtering, reads were de-noised using a pre-clustered procedure to exclude sequences that were probably the result of sequencing errors. Next, putative chimeric sequences were found and eliminated using the algorithm for chimaera VSEARCH (Rognes et al. 2016). Sequences from each sample obtained after a successful quality check and filtering were grouped with a minimum confidence of 0.8 using RDP

Bayesian classifier, with each cluster corresponding to a species (Wang et al. 2007). The resulting sequences were clustered into OTUs, and finally, a taxonomy of each OTU was generated. The constructed OTU table was modified to correspond with the current official nomenclature (List of Prokaryotic names with Standing in Nomenclature: LPSN-www.bacterio.net). Rarefaction analysis was performed using a single rarefaction function of mothur. The diversity indices such as Inverse Simpson, Shannon Diversity, and observed richness sub-sampling were used to estimate alpha diversity by using PAST v3.18 (Hammer et al. 2001). Non-metric multidimensional scaling (NMDS) using Bay-Curtis distance was used for the examination of community relationships among all four samples. A heat map depicting relative abundance was also constructed using Morpheus, a web-based tool (hhtps://software.broadinstitute.org/Morpheus). The unclassified sequences from each dataset were extracted, followed by alignment using CLUSTAL W and finally, new phylotypes were inferred from the trees using the iTOL software (www.itol.embl.de).

2.6. PICRUSt analysis

Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) gauges the gene families contributed by bacteria or archaea to a metagenome using a 16S rRNA gene and a database of reference genomes (Langile at al. 2013). PICRUSt can form an important platform to look into the intricacies of bacterial community structure and function in an environment (Fan et al. 2017). It consists of two steps: i) gene content inference: It makes use of 16S copy number and gene content annotations from reference bacterial genomes in the IMG database ii) metagenomic inference: it is based on the MOTHUR's OTU table, where OTU IDs are tips in the reference OTU tree, the copy number of the marker gene, and the gene content of each OTU, and produces a metagenome table. The OTU table from Mothur was annotated in the

Kyoto Encyclopedia of Genes and Genomes (KEGG) to predict the functional analysis of the bacterial microbiome derived from amplicon-based metagenome data.

2.7. Characterization of bacterial strains

2.7.1. Isolation and purification of bacterial strains

1g of soil sample was serially diluted to 10^{-4} and plated on Biebl and Pfenning (BP) media with pyruvate (3 gL^{-1}) as the source of carbon (Table.1). After 2-3 days of incubation at 28 ˚C, distinct colonies were selected and purified by repeated streaking on the aforementioned media. The purified colonies were subsequently streaked and maintained at the nutrient agar (NA) plate (Table. 3).

Table. 1. Modified Biebl and Pfennig's (MBP) media composition

Components	**g/L**
$CaCl_2.H_2O$	0.05
$MgSO_4.7H_20$	0.2
Yeast extract	1.0
NaCl	0.4
KH_2PO4	0.5
NH_4Cl	0.6
Sodium pyruvate	3.0
Ferric citrate	5 ml from a 0.2 mg / 10 ml stock
Peptone	1.0
Vitamin B_{12}*	1 ml from a 0.2 mg / 10 ml stock
SL7	1 ml

Table. 2. SL7 components

Components	**mg / L**
$CoCl_2.6H_2O$	200
$MnCl_2.6H_2O$	100
$NaMoO_4.2H_2O$	40
$ZnCl_2$	70
$H3B0_3$	60
$CuCl_2.H_20$	20
$NiCl_2.6H_2O$	20
HCl	25 % v/v; 1ml

Table. 3. Nutrient Broth media composition

Media composition	**g / L**
Beef extract	1
Yeast extract	1.5
Peptone	5
NaCl	5

Adjust the pH to 7 and autoclave

2.7.2. Enrichment and isolation of *Planctomycetota* members

For enrichment and purifications of *Planctomycetota*, MYM media (Kaushik et al. 2020) was used which was optimized in accordance with the environmental conditions (pH, temperature and salinity) of the sample (Table. 4). The serum vials from the on-site sample enrichment were sealed with parafilm and incubated at 25 ˚C for 3-4 months.

Table. 4. MYM media composition (Kaushik et. 2020)

Media composition	**g / L**
Hutner's basal salts	20 ml
Yeast extract	0.1 g
$(NH_4)_2SO_4$	0.5 g
Glucose	2.0 g
N-Acetyl glucosamine (NAG)	1.0 g
Vitamin's solution	10 ml
Distilled water	970 ml
Agar (Sigma made)	20 g (2 %)

Table. 5. Metal 44 components

Components	**mg/ L**
Na-EDTA	250
$ZnSO_4.H_2O$	1095
$FeSO_4.7H_20$	500
$MnSO_4.H_2O$	154
$CuSO_4.5H_2O$	39.2
Co $(NO_3)2.6H_20$	24.8
$Na_2B_40_7.10H_20$	17.7
Autoclaved distilled water	1000 ml

Note: To prevent the metal ions from precipitating, dissolve Na-EDTA and add a few drops of concentrated H_2S0_4 to the solution. METAL 44 can be filtered sterilize and stored at 4 ˚C for up to 3 months.

Table. 6. Hutner's Basal salts compositions:

Components	**g/ L**
Nitrilotriacetic acid (NTA)	10.0
$MgSO_4.7H_20$	29.7
$CaCl_2$. $2H_20$	3.335
$(NH_4)6MoO_7O_{24}.4H_2O$	9.250 mg
$FeSO_4.7H_2O$	99.0 mg
"Metals 44"	50 ml
Autoclaved distilled water	950 ml

Note: Nitrilotriacetic acid (NTA) was dissolved in distilled water and pH was adjusted to 7.0 with KOH (~7.3 g/litre). The remaining salts were independently dissolved. This solution can be stored at 4 ˚C until precipitated.

Table. 7. Vitamin solution compositions

Components	**mg / L**
Vitamin solution	0.2
Thiamine-HCl.2H_20_2	10
Folic acid	1
Ca-pantothenate	10
Biotin	4.0
Riboflavin	10
Nicotinamide	10
p-Aminobenzoic acid	10
Pyridoxine hydrochloride	10
Autoclaved distilled water	1000 ml

Note: Vitamin solution can be stored at 4 ˚C for a year after filter sterilization.

2.7.3. Purification and maintenance of *Planctomycetota*

The enrichment samples were spread out individually on the Petri plate containing the MYM solid media containing 2 % agar. Petri plates were incubated at a temperature of 23 ˚C to 25 ˚C, and every three days, all plates were examined for the development of pinkish colonies. The pink colonies were further purified on Petri plates containing the same MYM solid media. In addition, several glossy and white colonies were also purified in order to look for colonies of *Planctomycetota* without pigment. All the isolated strains were maintained in the MYM solid media by repeatedly sub-culturing at intervals of one to two months.

2.7.4. Culture preservation

The purified cultures were maintained either in glycerol stocks or agar stabs. Glycerol stocks were prepared by adding 1 ml of axenic cultures grown in nutrient broth in 1 ml of 100 % autoclaved glycerol. The stocks were stored at 4 °C overnight before being

moved to -20 °C for extended preservation. For agar stabs, 3 ml of NB containing 2 % agar was added to 5 ml centrifuge tubes and allowed to solidify in a slant. The cultures were stabbed into the media and incubated for 2-3 days at room temperature and were maintained at 4 ˚C to preserve them for a longer period. The purity and viability of the cultures were checked after every three months by sub-culturing the stabs.

2.7.5. Phenotypic studies

2.7.5.1. Morphological characterization

Colony morphologies like colour, size, shape, margin and elevation were observed in NA media under grown optimum conditions.

2.7.5.1.1. Microscopic observation

Morphological characters of the cell (shape, size, cell division) were observed using phase contrast microscope at JNTUH, Hyderabad (Olympus-B2011).

2.7.5.1.2. Field emission scanning electron microscopy (FESEM)

1ml of exponential phase culture was centrifuged at 7000 g for 10mins at 4 ˚C. The cell pellet after further washing with phosphate buffer (0.05 mM, pH 7.2), was centrifuged at 7000 g for 10 mins at 4 ˚C. The washed pellet was fixed in 2.5 % glutaraldehyde solution and incubated for 6 hours at 4 ˚C. The cells were pelleted and sequentially dehydrated with an increased ethanol concentration from 10 to 100 % at 20 % intervals. 10 µl samples (100 %) on mini glass slides were mounted on stab with adhesive tape. SEM stabs were kept for sputtering (4 minutes) and finally visualized under the FESEM (Philips XL30) facility, School of Physics, University of Hyderabad.

2.7.5.1.3. Negative staining for Transmission Electron Microscopy (TEM)

The cells of *Planctomycetota* were negatively stained by 2 % uranyl acetate, and the staining protocol was done according to Kumar et al. 2020 followed by visualization of the grid under TEM microscopy (H-7500 Hitachi) facility of CCMB, Hyderabad.

2.7.5.1.4. Transmission electron micrograph sections for ICM structures

The ultrathin section of *Planctomycetota* cells at the exponential phase was outsourced at RUSKA Diagnostics, Hyderabad and the protocol was followed according to Kumar et al. 2020. Finally, the sections of the Planctomycetal cells were mounted on copper grids and visualized under the TEM (H-7500) facility of CCMB, Hyderabad.

2.7.5.2. Physiological characterization

2.7.5.2.1. Chemotrophic growth mode

Chemotrophic growth mode was checked in Erlenmeyer flask (250 ml) containing 100 ml of BP media supplemented with pyruvate as a carbon source (0.35 % w/v) and was incubated at 30 ˚C in a shaker at 150 rpm.

2.7.5.2.2. Growth at different temperatures

The culture was incubated at various temperature ranges (4, 8, 10, 16, 20, 25, 28, 30, 35, 37, 40 and 45 ˚C) in their respective media and growth was observed.

2.7.5.2.3. Growth at different pH

Growth of culture in different pH (6-11) was checked in NB using a buffer system: for pH 5-8, KH_2PO_4/K_2HPO_4 (phosphate buffer) and for pH 9-11, $NaHCO_3/NaOH$ (bicarbonate buffer) was used. It was then kept at 30 ± 2 °C for incubation.

For *Planctomycetota*, cultures were tested at different pH ranges (4-10) in buffered (Bondoso et al. 2015) liquid growth media at 25 ˚C.

2.7.5.2.4. Growth at different NaCl% concentration

The bacterial cultures were inoculated in their respective media where different concentrations (0-15 %) of NaCl (w/v) were added at an interval of 1 %. Growth turbidity was measured at 540 nm using a colourimeter (Systronics) against the blank (media without culture). An increase in optical density was considered an increase in growth.

2.7.5.2.5. Vitamin B_{12} requirement

The colonies of *Planctomycetota* were inoculated in the MYM media supplemented with sole filter-sterilized vitamin B_{12} solutions (0.2 µm acro disc syringe filter, Pall make). The cultures were repeatedly streaked without vitamin B_{12} (0.02 mM) with glucose (0.1 % w/v) and ammonium sulphate (0.1 % w/v) as carbon and nitrogen sources, respectively.

2.7.5.2.6. Carbon sources utilization

Various carbon sources (pyruvate, D-glucose, D-cellobiose, D-rhamnose, D-mannitol, D-galactose, lactose, ribose, maltose, sucrose, sorbitol, inositol, L-alanine, histidine, gluconate, L-proline, D-xylose, trehalose, acetate, citrate, methanol, propionate, fructose, glycerol, ethanol, malate, D-mannose, arabinose) were used as a source of carbon in their respective media. Pyruvate was replaced in the BP media at (0.3 % w/v). For MYM media, $(NH_4)_2S0_4$ (0.1 % w/v) was used as a sole nitrogen source, and growth was tested with the above carbon substrates at a concentration of (0.1 % w/v).

2.7.5.2.7. Nitrogen sources utilizations

Different nitrogen sources (D-valine, L-leucine, L-lysine, L-tryptophan, L-tyrosine, D-alanine, L-methionine, DL-ornithine, yeast extract, casamino acid, L-isoleucine, NAG, L-aspartic acid, L-serine, cysteine, L-phenylalanine, DL-threonine, L-glutamic acid, L-histidine, L-glycine, D-alanine) were used as a sole nitrogen source to test its utilization in their respective media. In BP media, NH_4Cl was replaced as a nitrogen source at 7 mM. Whereas, in MYM media, D glucose (0.1 % w/v) was used as a sole nitrogen source, and the cell growth was tested on the different nitrogen sources at 0.1 % (w/v).

2.7.5.3. Biochemical characterization

2.7.5.3.1. Oxidase and catalase test

Oxidase discs procured from HiMedia (DD018) were used by following the manufacturer's instructions. The catalase test was evaluated by observing the appearance of bubbles after adding H_2O_2 (3 % v/v).

2.7.5.3.2. Nitrate reduction and H_2S production

Nitrate reduction was evaluated using HiMedia (DD041) discs as per the manufacturer's instructions. Sulphide indole motility agar media was used to measure H_2S generation. The blackening of media along the inoculation's stab line indicates a positive result.

2.7.5.3.3. Hydrolysis of compounds

Starch agar (HiMedia M1075), Gelatin agar (HiMedia M920) and urea broth (HiMedia M112) were used to hydrolyse starch (0.2 % w/v), gelatin (3 % w/v) and urea (2 % w/v), respectively. The hydrolytic capability of various compounds such as Tween 20, Tween

40, Tween 80, cellulose, casein, chitin, and was tested at 0.1 % (w/v) and substituted as substrates in BP medium. Hydrolysis of phytagel was also examined by streaking the cultures in solidified media containing 2 % (w/v) phytagel and incubating for 4 weeks (Kaushik et al. 2020).

2.7.5.3.4. Enzyme activity

Enzyme activity was assayed by using the API ZYM kit (Biomerieux, France) in accordance with the manufacturer's protocol.

2.7.6. Chemotaxonomic characterization

2.7.6.1. Polar lipid profiling

1 g of the lyophilized cells from each strain was used to extract the polar lipids as described by Kates (1986). 5 ml of 0.3 NaCl was added to the 1 g lyophilized cells and kept for 15 mins at room temperature. 8-10 ml of methanol was added to dissolve the pellets and vortexed for 5 minutes. The mixture was then steam heated till the l whole methanol gets evaporated, and then it was allowed to cool. 10 ml of chloroform was added to the dry pellets and then vortexed for 10 minutes. Then, 5 ml of 0.3 % NaCl was added and a well-separated aqueous phase was formed after centrifugation for 10 minutes at 8 ˚C. The lower (chloroform) layer was transferred into a round bottom flask (250 ml) and concentrated using a rotary flash evaporator (LABOROTA 4001, Heildolph). The dried extract, which contained polar lipid, was ultimately dissolved in 0.5-1.0 ml of chloroform and methanol (2:1). Finally, 10-20 µl was spotted on the TLC plate (7 x 7 cm in size), and visualization of lipids was done by dipping the TLC plate in ethanolic molybdophosphoric acid (5 % w/v) for 20 seconds followed by hot air drying.

Solvent system for the first dimension

Chloroform: methanol: water (30:12:1.5)

Solvent system for the second dimension

Chloroform: methanol: glacial acetic acid: water (35:7:6:1)

The polar lipids were further categorized into specific groups:

i) Dragendorff's reagent

This reagent was used to distinguish choline containing lipid molecules which give an orange colour spot. A mixture of solution A (5 ml), solution B (5 ml), acetic acid (20 ml) and H_2O (70 ml) was used for spraying on plates

Solution A: 1.7 g of bismuth nitrate (basic), was added to 100 ml of acetic acid, and H_2O (1:4)

Solution B: 40 g of potassium iodide to be dissolved in H_2O

ii) Ninhydrin reagent

0/1 % (w/v) of ethanolic ninhydrin was sprayed on TLC plate followed by hot air drying. The appearance of pink dots on the plates indicated the presence of amino lipids.

ii) Molybdenum blue reagent

This reagent was used to visualize phosphate-containing lipids which give off blue colour upon spotting. The reagent mixture of solution A (20 ml), solution B (20 ml), and distilled water (40 ml) was prepared and used for spraying.

Solution A: 40.1 g of molybdenum trioxide was added to 1 L of 25 N H_2SO_4

Solution B: 1.78 g of powdered molybdenum was added to 0.5 L of 25 N H_2SO_4 and slowly boiled for 20 min.

iii) Ninhydrin reagent

0.1 % (w/v) of ethanolic ninhydrin was sprayed on TLC and air dried. The visualization of pink spots on the TLC plate indicates the presence of amino lipids.

2.7.6.2. Cellular fatty acid profiling

Cellular fatty acids were methylated, separated, and identified in accordance with the guidelines of the Microbial Identification System (Microbial ID; MIDI 6.0 version; method, RTSBA6) (Sasser 1990). It was outsourced to Royal Research labs, Secunderabad, India.

2.7.6.3. Quinone analysis

1g of the fresh pellet was taken where 15 ml of acetone and 2 ml of distilled were added. The mixture was sonicated (50 % power) for 30 minutes and 20 ml of acetone was added again followed by overnight incubation. The solution was centrifuged at 1000 g for 20 minutes and stored at 4 °C. The mixture was then filtered using Whatman filter paper and concentrated using the rotary flash evaporator. To the dry filtrate, 20 ml of water and 20 ml of hexane in the 1:1 was added and vortex vigorously. The solution was allowed to phase out into 2 layers for 20-30 min. n-Hexane layer was carefully emptied out and Na_2SO_4 was added to eliminate the remnant water. These steps were repeated three times for maximum extraction. The n-hexane layer was then dried, 0.5 ml of ethanol was added to dissolve the extracted quinone and the sample was analysed as per the method outlined by Xie and Yokota (2003).

2.7.6.4. Cell wall amino acid analysis

Cell wall amino acids were analysed as described by (McKerrow et al. 2000; Schleifer 1985). The fresh pellet was washed and suspended in 50 mM phosphate buffer (pH 7.2) and then sonicated (50 % power) for 20 minutes at 4 ˚C. The severed pellets were washed with phosphate buffer and centrifuged (8000 g at 4 ˚C for 10 minutes) at room temperature. This step was repeated 5 times, followed by washing pellets with autoclaved milli-Q water 3 times. Then, the total lysate was centrifuged for 60 minutes at 10000 g to collect the disrupted cell wall, and the gathered cell supernatant was centrifuged (100 g) for 60 minutes. The remaining portion of the cell was solvated in 2-3 ml of 12 N HCl, hydrolysed (12 hours autoclaved), and transferred into a beaker to evaporate HCl by carefully heating it on the hotplate. Finally, the remainder was dissolved in 1 ml of sterile milli-Q water and neutralized (NaOH). Finally, the dried residue was dissolved in 1 ml of sterile milli-Q water and neutralized (NaOH). Derivatization of the cell wall hydrolysed amino acid and standard amino acids was carried out by the o-phthaldialdehyde (OPA) reagent. OPA reagent was prepared by:

i) Dissolving 9 mg sodium tetraborate (HiMedia) in 4.5 ml H_2O with gentle heating

ii) Dissolving 66 mg of N-acetyl-L-cysteine and 27 mg of o-phthaldehyde (27 mg) (Sigma Aldrich), in 0.5 ml methanol, then adding in (i)

Finally, the derivatized samples were filtered (0.2 µl filter, Icon pall supor @200) and analysed with C18 column, phosphate buffer (50 mM pH 7.2): acetonitrile (ratio 85:15) as the mobile phase, with a flow rate of 0.4 ml/min (UV array of detector 334 nm).

2.7.6.5. Polyamine analysis

According to Dion and Herbst (1970), polyamines were extracted from cultures. Cultures at the logarithmic phase were harvested by centrifugation (10000 g, 4 °C, 10 min), and the pellet was then washed three times with 50 mM phosphate buffer (pH 7.5). The pellet was resuspended in 0.2 M perchloric acid and sonicated (50 % power) for 15 minutes. The disrupted cells were neutralized with 1M $NaHCO_3$ and centrifuged for 10 min at 4 °C**.** Polyamines in the supernatant were derivatized with dansyl chloride (3 mg ml-1), whereas the unbound dansyl chloride was removed by adding proline (50 mM). It was extracted three times by toluene, separated and evaporated in a rotary flash evaporator, and dissolved in HPLC-grade methanol. Polyamines were identified by using HPLC with a photodiode array detector (PDA), C18 Luna column (5 µm, 250×4.6 mm). A linear gradient method was applied to separate polyamines. The mobile phase was composed of solvents A (acetic acid in water, 1 % v/v) and B (acetonitrile (ACN): methanol, 4:1). Polyamines were detected by their retention time, peak absorption spectra of standards along with spiking using internal standards.

2.7.6.6. Extraction and characterization of Violacein like pigment

The violacein-like pigment was extracted from 48 hours grown cell suspension in nutrient broth according to Sasidharan et al. (2015) with some minor modifications. 10 ml methanol was added to 40 ml of the broth and incubated for 30 minutes with shaking conditions. 30 ml of methanol lysate was treated with 10 ml of ethyl acetate and mixed well, and violacein pigment present on the top organic layer was extracted using a separating funnel. The extracts were then concentrated at 40 ˚C using a rotary evaporator and air-dried. The dried extract was then characterized using UV–Vis and Fourier transform infrared spectroscopy (FTIR).

2.7.7. Genotypic studies

2.7.7.1. Marker-based phylogenetic studies

2.7.7.1.1. 16S rRNA gene amplification and identification

DNA was extracted from the pure cultures using the commercial DNA isolation kit (Nucleo-pore gDNA Fungal Bacterial Mini Kit, from M/s. Genetix Biotech Asia Pvt. Ltd, India). The isolated DNA was further used for PCR amplification and genome sequencing. The following primers were used for the amplification of 16S rRNA gene: F8/R1492, F8/R1525, F27/R1492, F27/1525, F40/R1388. The PCR amplification was carried out using the master mix from Takara (EmeraldAmp GT PCR). The PCR reaction mixture has the following components (Table. 8). The PCR reaction mixture components:

Component	Volume (µl)
Master mix	25
Forward primer	2
Reverse primer	2
Milli-Q	16
Template DNA*	5

*From the isolated DNA

The 16S rRNA gene was amplified under the following conditions:

Step	Temp (˚C)	Time (min)
Initial denaturation	94	10
Denaturation	94	1
Annealing	54	54
Elongation	72	1.5
Number of cycles:32		
Final elongation	72	15

The amplified product was visualized in the agarose gel prepared in TAE buffer along with a DNA marker (1Kb, Thermo Scientific) at 50 volts for 40 minutes. The bands were visualized by ethidium bromide staining and observed under BIO-RAD gel documentation system. The amplified PCR products were outsourced to Agrigenome Labs

Pvt.Ltd. for sequencing, which was carried out using ABI 3730 XL DNA Analyser automated sequencer. The obtained raw sequences (.ab1 format) were assembled together using the DNA Lasergee SeqMan Pro programme. The BLAST search using the EzBioCloud database was used to search the assembled (1300-1500 bp) sequence identity with the existing bacteria. MUSCLE in MEGA X (Edgar 2004) was used for multiple sequence alignment, whereas the pairwise nucleotide sequence alignment tool (www.ezbiocloud.net/tools/pairAlign) was used for pairwise alignment.

2.7.7.1.2. Phylogenetic studies

The 16S rRNA gene sequences of all strains under consideration obtain from amplicon-based sequencing, and sequence extracted using ContEst16s (https://www.ezbiocloud.net/tools/contest16s) and NCBI databases were aligned using MUSCLE implemented in MEGA 7.0 (Kumar et al. 2016). Kimura two-parameter was used to calculate the distance in a pairwise deletion process (Kimura 1980). The phylogenetic tree was constructed using neighbour-joining (NJ) (Saitou and Nei 1987), maximum likelihood (ML) (Yang 1997) and minimum evolution (ME) (Rzhetsky and Nei 1992) in the MEGA 7 software with a bootstrap value of 1000 replication (Felsenstein 1985).

2.7.7.2. Genome-based phylogenetic studies

2.7.7.2.1. Multilocus sequence analysis

The complete length sequences of nucleotide and amino acids of 8 protein-coding housekeeping genes (*atpD*, *dnaK*, *gryB*, *recA*, *rpoA*, *rpoB*, *rpoD*, and *tryB*) were extracted from the genomes of the taxa under study (NCBI database) which were concatenated and aligned by the MUSCLE program for the MLSA analysis. Phylogenetic trees for the MLSA

were also constructed for the 16S rRNA gene. Additionally, the values of concatenated gene sequence identity between strains were also computed.

2.7.7.2.2. *In-silico*/*digital* DNA-DNA hybridization

In-silico/*digital* DNA-DNA hybridization has replaced the error-prone traditional DNA-DNA hybridization based on the genome sequences of the species. *d*DDH values among the strains were computed using the Genome-to-Genome Distance Calculator (GGDC) 3.0 (http://ggdc.dsmz.de/ggdc.php) (Meier-Kolthoff et al. 2022).

2.7.7.2.3. Orthologous Average nucleotide identity

The OrthoANI was calculated using the online tools in the EzBioCloud (https://www.ezbiocloud.net/tools/ani) (Yoon et al. 2017).

2.7.7.2.4. Average amino acid identity

Average amino acid identity between two species was calculated using the AAI calculator from Kostas lab (Rodriguez and Konstantinidis 2014).

2.7.7.2.4. Percentage of conserved proteins

POCP was calculated using the BLASTP program where the query genome was aligned with the reference genome for the identification of conserved regions (Qin et al. 2014). The query genome proteins were considered to be conserved if the BLAST matches E value<1e-5, greater or 40 % identity, and more than 50 % alignment. The POCP value between the two genomes was calculated using the formula [(C1+C2)/(T1+T2)] *100 %, where C1 and C2 represents conserved proteins, T1 and T2 represents total number of proteins of the two genomes under study.

2.7.7.2.5. UBCG-based phylogenomic tree

Up-to-date bacterial core gene (UBCG) tool developed by Chun's lab was used for constructing the phylogenomic tree using 92 core genes from the publicly available genome sequences of organisms (Na et al. 2018). Three main steps are involved in UBCG based phylogenomic tree: i) Identification of the core genes - Based on the CDS identified from the whole genome sequences, hidden Markov models (HMM) searches the conserved genes in the dataset ii) numerous sequence alignments - The MAFFT algorithm aligns each gene sequence, and then the alignments are concatenated, and gaps containing places are filtered and iii) display of the phylogenetic tree are - it involves the construction of a phylogenomic tree using the concatenated sequence of the 92 core genes called UBCG tree (RAxML / FastTree). The tree was constructed with the alignment files generated by FASTA by using MEGA7 software.

2.7.7.2.6. Pan-genome analysis

Pan-genome studies were conducted using Bacterial Pan-Genome Analysis (BPGA), an ultra-fast pan-genome analysis pipeline (Chaudhari et al. 2016). The first step is the preparation step, where the input files are modified by inserting the genome ID for clustering. The second step involves clustering genes based on sequence identity into orthologous clusters by USEARCH. Thirdly, the binary matrix is prepared based on the presence or absence of clusters. The last step consists of pan-genome profiling, which determines the genes shared following the insertion of each genome one at a time. It is represented as core/pan-genome profile curves which provide genome-wise core, accessory, unique and exclusively absent gene numbers.

2.7.7.2.7. Genome annotation and comparison

Genome annotation was done using several bioinformatic tools like RAST Server, PATRIC, NCBI, KBase and NCBI. The annotation data for each study were downloaded and retrieved as excel files from the relevant servers and examined for the following:

i) Whole genome alignments: In this method, a query genome is used as a reference and is aligned against the given set of genomes. The genomes were aligned using Mauve (genome alignment), a feature of the PATRIC database, to identify several maximum matches and local collinear blocks (LCBs)

ii) Orthologous-based gene content comparison: Here, a set of genomes are compared to determine which genome include certain gene families (Setubal et al. 2018). It is represented as venn diagram with the help of OrthoVenn2 online software (Xu et al. 2019).

iii) Function-based gene content comparison: Protein description database such as cluster of orthologous groups of proteins (COG) in genome sequences are used for genome comparisons between genomes based on functions. Also, dbCAN2 (http://cys.bios.niu.edu/dbCAN2), a meta server for automated Carbohydrate-active enzyme annotation was used the annotation of structural families of associated carbohydrate-binding and catalytic modules.

iv) Secondary metabolite biosynthetic gene analysis: For the analysis of secondary metabolites, antiSMASH 5.0 bacterial version (https://antismash.secondarymetabolites.org), a freely available online tool was used for the *in-silico* identification for the presence of putative genetic clusters that are involved in biosynthesis of secondary metabolites (Blin et al. 2019).

2.7.7.3. Genome sequences used in the study

The type strains of different taxa (genus *Paludisphaera*, *Sinomonas*, *Glutamicibacter*, *Janthinobacteria* and *Comamonas*) were downloaded from the NCBI database. The names of organisms, source of isolation and accession numbers used in the study are listed in Table. 9.

Table. 9. Microorganisms name, strain number, source of isolation and accession number

Organisms	**Isolation source**	**NCBI accession number**
Paludispaera rhizophaerae JC665^T	Rhizosphere soil	JAALCR000000000
Paludisphaera sp. JC747	Algae	JAHPZK000000000
Paludisphaera borealis PX4^T	Peat bog	NZ CP019082
Paludisphaera soli JC670^T	Soil	NZ JAALJI010000001
"Paludisphaera sp. SH-PL62"	Soil	NZ CP011273
Aquisphaera giovannonii OJF2^T	Sediments	NZ CP042997
Aquisphaera insulae JC669^T	Rhizosphere soil	JAALJH010000001
Singulisphaera acidiphila DSM1865^T	Oxic layer of the bog Obukhovskoe	NC 019892
Tautonia plasticadhaerens ElP^T	Algae	NZ CP036426
Tautonia sociabilis GM2012^T	Hot spring microbial mat	NZ RYZH01000100
Tautonia marina JC650^T	Sediment	NZ WEZF01000010
"*Tautonia rosea* JC636"	Sediment	NZJACEHG010000001
"Tautonia rosea JC649"	Sediment	NZ JACEHI010000001
Tautonia rosea JC657^T	Sediment	NZJABBYO010000001
Isosphaera pallida ATCC 43644^T	Algal mat	NC 014962
Alienimonas californiensis CA12^T	Sea water	NZ CP036265
Glutamicibacter rhizosphaerae JC586^T	Rhizosphere	VHIN01000027
Glutamicibacter mishrai SR-52^T	Poultry farm	CP032549
Glutamicibacter halophytocola KLBMPT5180^T	Root of *Limonium sinense*	CP012750
Glutamicibacter creatinolyticus LGCM259^T	Abcess of a mare	CP034412
Paeniglutamicibacter psychrophenolicus DSM 15454^T	Alpine ice cave	JAGIOE000000000
Paeniglutamicibacter kerguelensis DSM 15797 ^{T}	Sea water	JAGIOF000000000
Paeniglutamicibacter antarcticus W2^T	Sediment	CZJT00000000
Rhodococcus ruber DSM43338^T	Sewage sludge	LRRL01000001
Sinomonas cellulolyticus JC656^T	Rhizosphere soil	JAERRC000000000
Sinomonas sp. KCTC 49339	Soil	BNCM01000001
*Sinomonas_notoginsengisoli*_KCTC 29237^T	Rhizosphere soil	JAIZDP000000000
Sinomonas atrocyanea CGMCC1.1891^T	Air	BMKT00000000
Sinomonas_atrocyanea KCTC 3377^T	Mangrove soil	CP014518
Sinomonas_atrocyanea NBRC 12956^T	Soil	BJMO00000000
Sinomonas mesophilia MPKL26^T	Soil	MTIC00000000

Sinomonas albida LC13^{T}	Forest soil	SWDW00000000
Sinomonas sp. Bin6.19	Soil	JABFXL000000000
Sinomonas gamaensis NEAU-HV1^{T}	Soil	SSCL00000000
Sinomonas sp. NEAU-HV1	Soil	SSCL00000000
Sinomonas susongensis A3^{T}	Weathered biotite	NZ_SSNH00000000
Sinomonas humi MUSC117^{T}	Soil	JTDL00000000
Micrococcus terreus CGMCC 1.7054^{T}	Forest soil	FPCG00000000
"*Candidatus* Janthinobacterium sediminis" JC611^{T}	Sediment	JAFLJN010000000
Janthinobacterium rivuli FT68W^{T}	Stream	WFLJ01000100
Janthinobacterium lividum EIF1^{T}	Soil	CP048832
Janthinobacterium violaceinigrum FT13W^{T}	Stream	WFLI01000100
Janthinobacterium tructae SNU WT1^{T}	Kidney of rainbow trout	CP041185
Janthinobacterium svalbardensis PAMC 27463 ^{T}	Glacier ice	CP023422
Janthinobacterium psychrotolerans S3-2 ^{T}	Freshwater pond	LOCQ01000062
Janthinobacterium aquaticum FT58W ^{T}	Stream	WFLK01000010
Janthinobacterium agaricidamnosum NBRC 102515^{T}	Mushroom	HG322949
Polynucleobacter acidiphobus MWH-PoolGreenA3^{T}	Freshwater	CP023277
Comamonas soli JC664^{T}	Rhizosphere soil	ON908990
Comamonas koreensis KCTC 12005^{T}	Sediment	AF275377
Comamonas sediminis S3^{T}	Sediment	LT009500
Comamonas piscis CN1^{T}	Korean Rockfish	KM263565
Comamonas suwonensis EJ-4^{T}	Stream water	MN381948
Comamonas guangdongensis CYO1^{T}	Forest sediment	EU515237
Comamonas aquatilis SB30-Chr27-3^{T}	Garden pond	KU355878
Comamonas fluminis CJ34^{T}	River	MW599737
Comamonas testosteroni ATCC 11996^{T}	Soil	MN381948
Comamonas thiooxydans DSM 17888^{T}	Sulphur spring	BBVD01000034
Comamonas odontotermitis Dant3-8^{T}	Gut of termite	DQ453128
Comamonas terrae NBRC 106524T BCN^{T}	Soil	01000031
Comamonas aquatica NBRC 14918^{T}	Water	JQ941713
Comamonas jiangduensis YW1^{T}	Agricultural soil	JQ941713
Comamonas kersterii AF61^{T}	Human	AJ430347
Comamonas terrigena 247^{T}	Hay infusion filtrate	AF0787721
Comamonas composti CC-YY287^{T}	Food waste compost	EF015884
Comamonas granuli Ko03^{T}	Microbial granules	AB187586
Comamonas denitrificans 123^{T}	Activated sludge	AF233877
Comamonas nitrativorans 23310^{T}	Denitrifying reactor	AJ251577
Comamonas faecalis ff42^{T}	Faeces	KF509984
Comamonas serinivorans SP-35^{T}	Compost	JN604116
Comamonas humi GAU11^{T}	Soil	AB907700
Comamonas zonglianii BF-3^{T}	Phenol contaminated soil	GQ245981
Comamonas badia 12^{T}	Activated sludge	AB1644321
Burkholderia cepacian ATCC 25416^{T}	Onion	KU355878
Rhodoferax saidenbachensis ED16^{T}	Drinking water reservoir	AWQR01000064

RESULTS

3. RESULTS

3.1. Analysis of bacterial communities of Loktak lake

3.1.1. Culture-independent study

3.1.1.1. Sampling strategy

Soil samples were collected from four different sites (east, west, north and south) of Loktak lake. The GPS location and description of land use are given in Table. 10. The pH and temperature of the lake ranges between 6-7 and 20-28 ˚C. The samples collected were processed separately for metagenome and culture dependent studies.

Table. 10. Sample name, sample type, GPS location and description of land use

Sl.no.	Sample Name	Sample type	Sample Location and GPS	Description of Land use
1	Lok1	Lake sediment	East-a) Komlakhong Latitude: 24˚51'50'' N Longitude: 93˚86'67'' E b) Thanga: Latitude: 24˚53'22'' N Longitude: 93˚83'25'' E	Area inside the Loktak lake inhabited by local people with mostly aquaculture activities
2	Lok2	Lake sediment	West- a) Ningthoukhong Latitude: 24˚56'67 '' N Longitute: 93˚75'87 'E b) Thinungei Latitude: 24˚55'03'' N Longitute: 93˚76'00'' E c) Phubala Latitude: 24˚53'41'' N Longitute: 93˚75'62'' E	Aquaculture and water runoff from hydroelectric project
3	Lok3	Lake sediment	North-a) Mayang Imphal Latitude: 24˚59'94'' N Longitute: 93˚87'06'' E	Aquaculture and open water area
4	Lok4	Lake sediment	South-a) Moirang Latitude: 24˚49'80'' N Longitude: 93˚77'65'' E b) Ithing Latitute: 24˚51'53'' N Longitute: 93˚79'69'' E c) KBLN Latitute: 24˚47'88'' N Longitute: 93˚83'95'' E	Tourist and national park surrounded by human settlement

3.1.1.2. Metagenomic DNA isolation, sequencing and quality control

DNA was extracted from four distinct composite soil samples of Loktak lake where the V1-V3 region was targeted and amplified. The metagenomic sequences obtained were submitted to the NCBI-SRA (Sequence Read Archive) and GC content (%), bio-sample numbers, and accession numbers of each sample are provided in Table 11. A total of 2,78,450 raw reads were produced by Illumina Miseq500 sequencing, but only 32,953 reads remained after quality and length filtering.

Table. 11. Total no. of reads, GC content (%), Bio-sample no. and NCBI accession numbers and reads of each sample

Sample no.	Sample source	Total no. of reads	GC Content (%)	Bio-sample no.	NCBI Accession no.
Lok1	Lake sediment	64,060	56.1	SAMN21212473	SRX12006901
Lok2	Lake sediment	68,174	57.1	SAMN21212473	SRX12020787
Lok3	Lake sediment	69,680	57.1	SAMN21212473	SRX12020789
Lok4	Lake sediment	76,536	55.8	SAMN21212473	SRX12020801

3.1.1.3. Bacterial abundance and diversity analysis

A total of 4097 operational taxonomic units (OTUs) were obtained from the four samples belonging to forty-four different phyla with 98 % cut-off. Longer unclassifiable sequences were placed in distinct OTUs with fewer sequences at the 98 % cut-off, while shorter, more abundant classifiable sequences were put into other OTUs, increasing the number of recognisable sequences. Rarefaction curves of the four samples were

constructed based on OTUs clustered at 98 % sequence similarity, indicating majority of the taxonomic diversity was covered in all the samples and the curves saturation also indicates proper sampling had been carried out and could be continued for further analysis (Fig. 5).

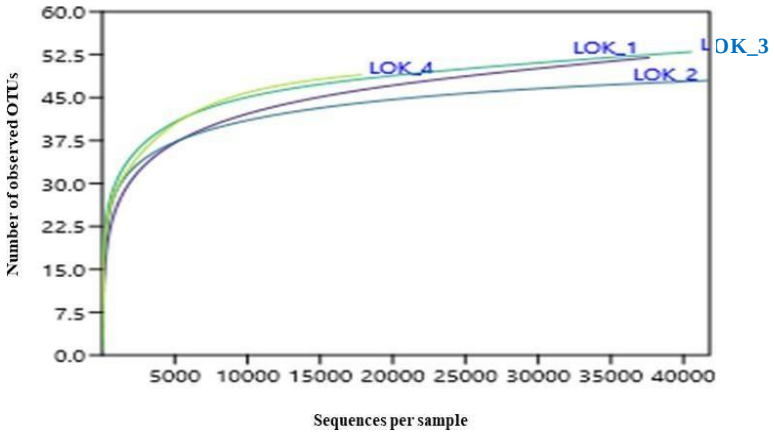
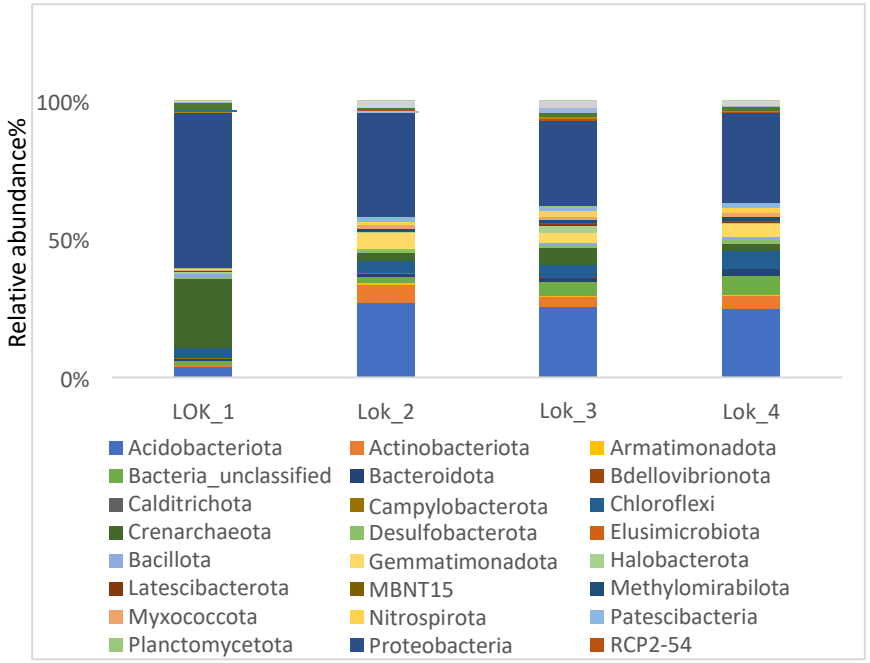


Fig. 5. Rarefaction analysis of the amplicon data obtained from Loktak lake sediment samples. Rarefaction curves were constructed based on OTUs clustered at 98 % sequence identity across different samples.

The taxonomic profiling of the lake revealed that at the phylum level most abundant OTUs belong to the phylum *Pseudomonadota* (32 to 43 %), *Acidobacteriota* (7 to 19 %), Bacteria unclassified (10 to 17 %), *Chloroflexota* (6 to 8 %), *Actinomycetota* (1 to 7 %) (Figure 6). Compared to other phyla, *Verrucomicrobiota* (3 %) and *Planctomycetota* (1 %), *Bacillota* (1.7 %), and *Spirochaetota* (0.6 %) have low percentages of the observed OTUs but are distributed in all four samples. The abundance here refers to the percentage of reads in the total OTUs present in the study. Phylum *Pseudomonadota* was most dominant in LOK1 (54 %), followed by LOK2 (35 %), LOK3 (32.4 %) and LOK4 (32.1 %). Phylum *Acidobacteriota* was dominant in LOK2 (20 %), followed by LOK3 and LOK4 (19.8 %) and LO1 (7 %). Similarly, bacteria unclassified was dominant in LOK3 (17 %), followed by LOK4 (15 %), LOK2 (13 %) and LOK1 (10 %) (Fig. 6)

a)



b)

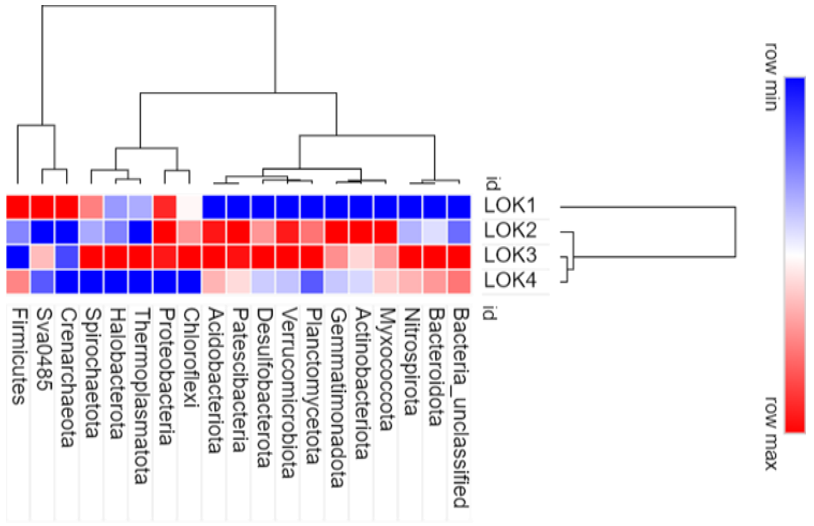


Fig. 6. A. Relative abundance of bacterial taxa of Loktak lake with pre-dominant bacterial phyla as obtained by 16S rRNA gene amplicon-based metagenome analysis

B. Microbial community composition with cluster analysis at the phylum level obtained from amplicon-based metagenome of Loktak lake

Bacterial abundance at the level of phylum is displayed in the form of heat map (Fig. 6. B). Overall, forty-four bacterial phyla (thirty-one recognised bacterial phyla and thirteen "*Candidatus*'' phyla) were obtained from the four samples. The different "*Candidatus*'' phyla are "*Ca*." GAL15, "*Ca*." LCP-89, "*Ca*." FCPU426, NKB15, RCP2-54, SAR324, Sva0485, TA06, WOR-1, WPS-2, WS2, WS4, and MBNT15. The sequences with OTU representations below 98 % were not classifiable by the RDP classifier and were therefore labelled as unclassified. In this study, 10- v 15 % of the total OTUs found were unclassified OTUs at the phylum level which were extracted for more in depth analysis. The distribution of common and unique phyla are given in Venn diagra (Fig. 7) and Table 12.

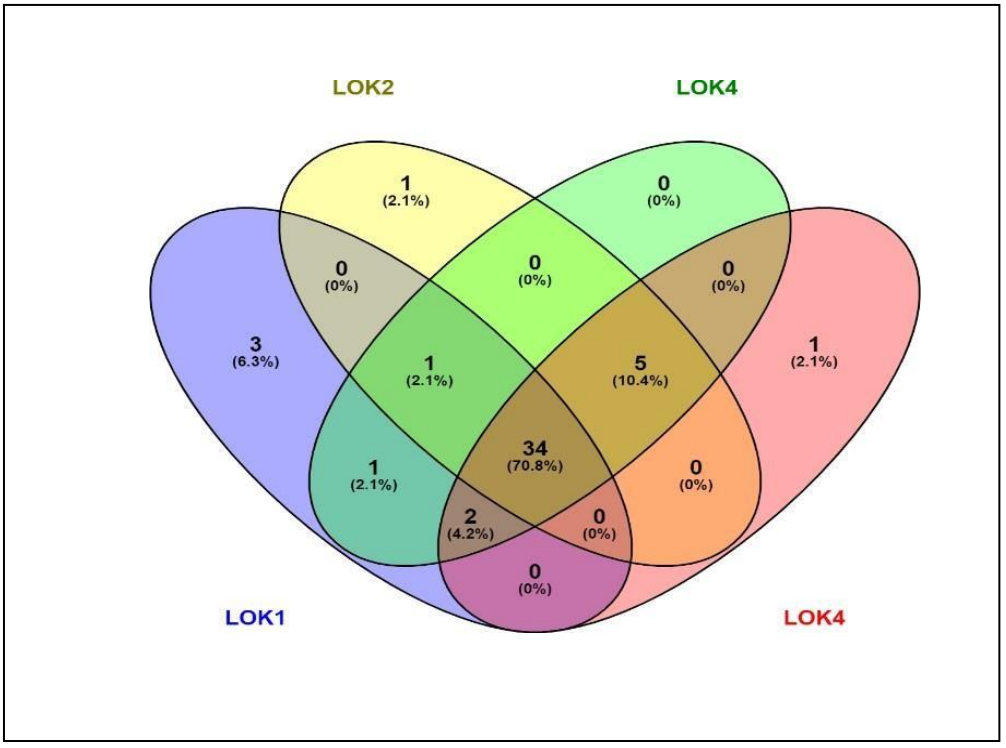


Fig. 7. Distribution of bacterial phyla amongst the four samples of Loktak lake

Table. 12. The distribution of common and unique bacterial phyla from the four soil samples of Loktak lake

Sample no.	Total Phyla	Taxa
LOK1	3	LCP-89, *Modulibacteria*, Uncultured
LOK2	1	*Dadabacteria*
LOK4	1	*Fusobacteriota*
LOK1, LOK2 and LOK4	1	*Campylobacterota*
LOK2, LOK3 and LOK4	5	*Cyanobacteria, Deferribacterota, Deferrisomatota, GAL15, SAR324*_clade
LOK1, LOK3 and LOK4	2	*Calditrichota, Edwardsbacteria*
LOK1, LOK2, LOK3 and LOK4	34	*Acidobacteriota ,Actinobacteriota, Armatimonadota* ,Bacteriaunclassified, *Bacteroidota, Bdellovibrionota, Chloroflexi, Crenarchaeota, Dependentiae, Desulfobacterota, Elusimicrobiota, FCPU426, Fibrobacterota, Firmicutes, Gemmatimonadota, Halobacterota, Hydrogenedentes, Latescibacterota, MBNT15, Methylomirabilota, Myxococcota, Nanoarchaeota, Nitrospirota, Patescibacteria, Planctomycetota, Proteobacteria, RCP254, Spirochaetota, Sumerlaeota, Sva0485, Thermoplasmatota, Verrucomicrobiota,WOR-1,WPS-2 ,Zixibacteria*

The distribution of each sample at the genus level is given in Fig.8. A and B. Bacterial v abundance at the genus level is shown in the form of a heatmap (Figure xx). The most abundant genus included "*Ca*. Solibacter" (11 %), *Bryobacter* (4.7 %), *Cavicella* (4.6 %), *Spirochaeta* (3.3 %), *Acidothermus* (3.3 %), *Curvibacter* (3.1 %), *Haliangium* (2.9 %), *Anaeromyxobacter* and *Pseudolabrys* (2.3 %). Additionally, genera like *Zixibacteria* (1.1 %), *Thiobacillus* (0.9 %), *Syntrophorhabdus* (1.5 %), *Phenylobacterium* (0.74 %), *Alkanindiges* (1.38 %), *Sumerlaea* (0.69 %), *Thiobacillus* (0.91 %),

Desulfobacca (0.69 %) and *Mucilaginibacter* (0.24 %) were also present in all the samples. Unclassified OTUs at the genus level consisted of 73-83 % of the total OTUs observed.

The distribution of common and unique taxa at the genus level are presented in the Venn diagram (Fig. 9) and Table. 13.

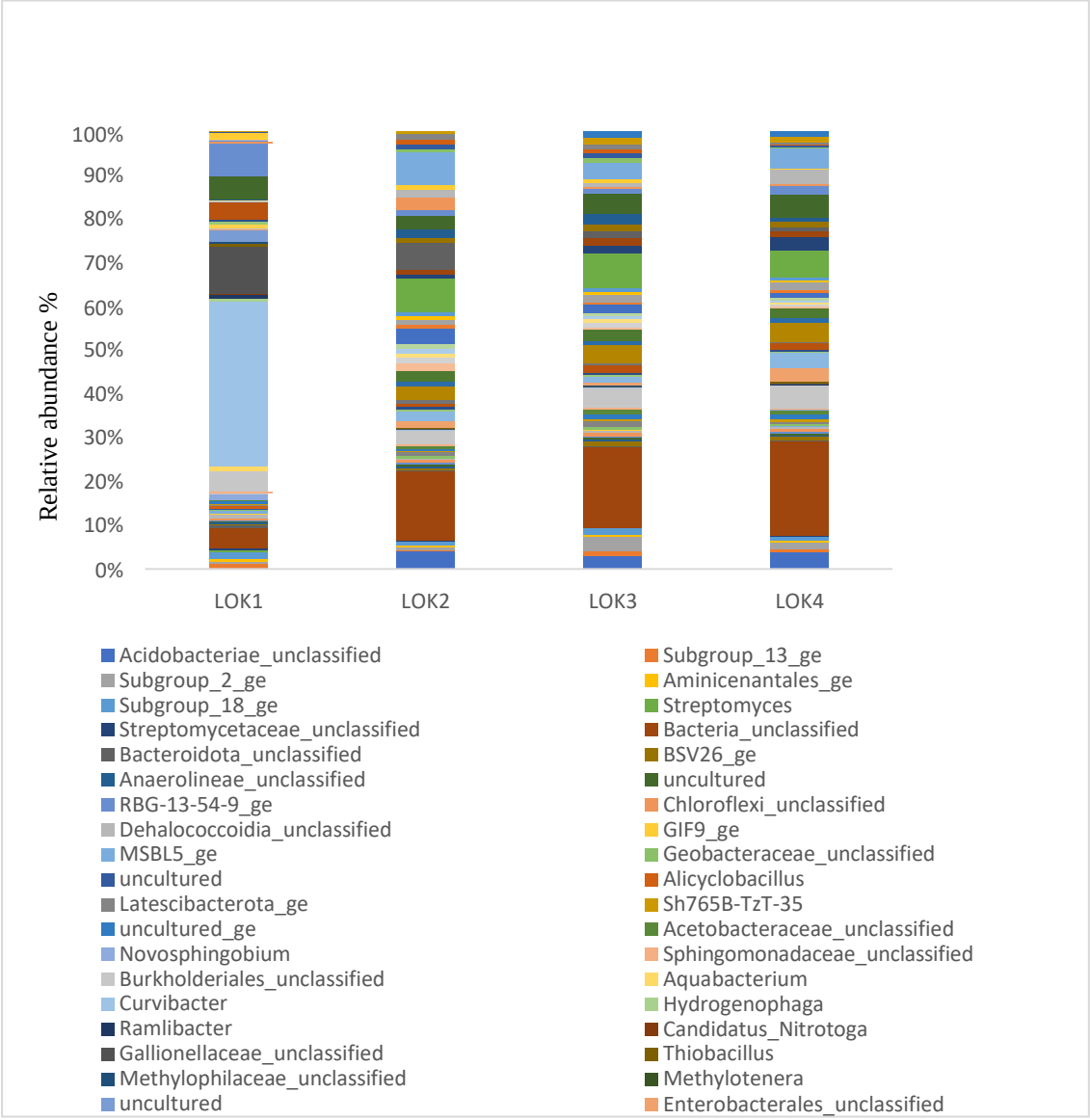


Fig. 8. A. Relative abundance of bacterial taxa of Loktak lake with pre-dominant bacterial genera as obtained by 16S rRNA gene amplicon-based metagenome analysis

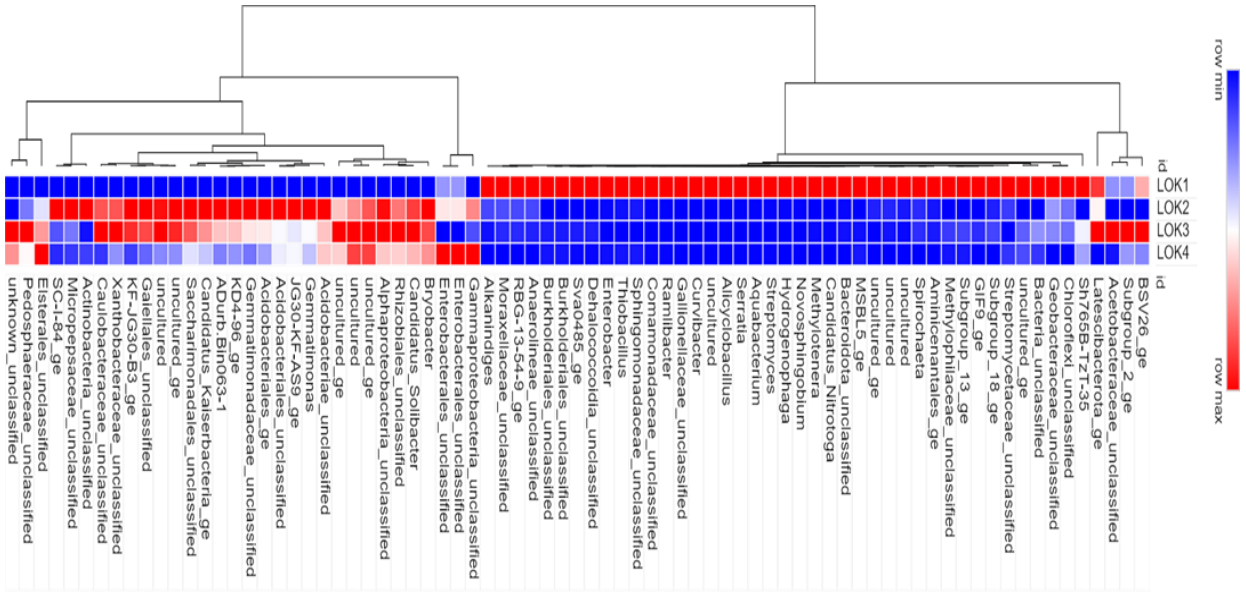


Fig. 8. B. Microbial community composition with cluster analysis at the genera level obtained from amplicon-based metagenome of Loktak lake

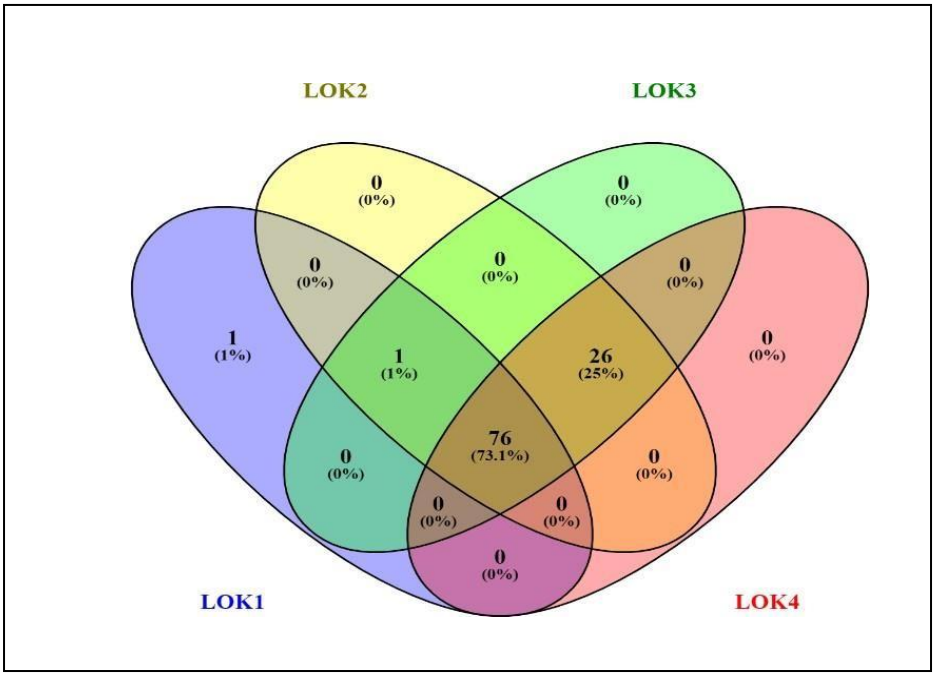


Fig. 9. Distribution of bacterial genera amongst the four samples of Loktak lake

Table. 13. The distribution of common and unique bacterial genera from Loktak lake

Sample no.	Total genera	Taxa
LOK1	1	*Alicyclobacillus*
LOK1, LOK2 and LOK3	1	*Pseudomonadales* unclassified
LOK2, LOK3 AND LOK4	26	*Terracidiphilus, Acidobacteriales ge, Candidatus Koribacte*r, Subgroup 7 ge, *Acidimicrobiia* unclassified, *IMCC26256 ge, Acidothermus, Micromonosporaceae* unclassified, *Actinobacteriota* unclassified, *Anaerolinea, Gemmatimonadaceae ge, Gemmatimonas, Rokubacteriales ge, Polyangia* unclassified, *Polyangiaceae* ,unclassified, *Nitrospira,Candidatus Kaiserbacteria ge, Parcubacteria* ge, LWQ8 ge, *Saccharimonadales*ge, *Elsterales* unclassified, *Micropepsaceae*, unclassified, *Rhizomicrobium, Reyranella, Nitrosomonadaceae* unclassified, *Rhodanobacteraceae* unclassified
LOK1, LOK2, LOK3 and LOK4		*Acidobacteriaceae*(Subgroup1)*unclassified, Acidobacteriales* unclassified, uncultured ge,*Bryobacter*,*Candidatus Solibacter*,Subgroup13 ge,Subgroup 2 ge, *Acidobacteriota* unclassified, *Aminicenantales* ge, Subgroup 18 ge, *Thermoanaerobaculum, Vicinamibacterales* unclassified, *Actinobacteria* unclassified, *Fimbriimonadaceae* ge, Bacteria unclassified, *Bacteroidales* unclassified, *Bacteroidetes* vadinHA17 ge, *Muribaculaceae ge, Bacteroidota unclassified, BSV26 g*e, *Anaerolineae*_unclassified, *Anaerolineaceae* unclassified, uncultured, *Chloroflexi* unclassified, *Dehalococcoidia unclassified, Ktedonobacteraceae* unclassified, *Desulfobacca, Deulfobacterota unclassified, Desulfuromonadia* unclassified, *Geobacteraceae* unclassified, *Geobacterales* unclassified, *Syntrophobacteraceae* unclassified, *Syntrophorhabdus, Clostridia unclassified, Lachnospiraceae unclassified, Oscillospiraceae* unclassified, *Firmicutes* unclassified, *Gemmatimonadaceae* unclassified, S0134 terrestrial group ge, *Latescibacterota* ge, *Anaeromyxobacter, Haliangium, Parcubacteria unclassified, Saccharimonadales* unclassified, *Pla1* lineage ge, *Planctomycetota* unclassified, *Acetobacteraceae* unclassified, *Alphaproteobacteria* unclassified, *Caulobacteraceae* unclassified, *Phenylobacterium, Micropepsaceae ge, Rhizobiales* unclassified, *Xanthobacteraceae* unclassified, *Sphingomonadaceae* unclassified, *Burkholderiales* unclassified, *Comamonadaceae unclassified, Curvibacter, Gallionellaceae* unclassified, *Thiobacillus, Oxalobacteraceae* unclassified, *Rhodocyclaceae* unclassified, *SCI84 ge, Enterobacterales* unclassified, *Enterobacteriaceae* unclassified, *Gammaproteobacteria* unclassified, *Alkanindiges, Cavicella, Moraxellaceae unclassified, Nevskia, Proteobacteria unclassified, Spirochaeta, Sva0485 ge, ADurb.Bin063, Pedosphaeraceae* ge ,*Pedosphaeraceae* unclassified, *Zixibacteria* ge

The community relationship between all four samples was carried out by performing non-metric multidimensional scaling (nMDS) plot based on Bay-Curtis distance using PAST software (Fig.10). The alpha diversity indices were calculated by considering Shannon (H), Inverse Simpson, and Observed Richness diversity indices. Shannon diversity index was highest in LOK4 (H = 7.7 ± 0.0) and lowest in (H = 5.8 ± 0.0). Inverse Simpson and Observed Richness indices were also highest in LOK4 (I = 335.2 ± 0.0) (sobs = 6574 ± 0.0), followed by LOK3 (I = 342.2 ± 5.5) (sobs = 5177.8 ± 39.9), LOK2 (I = 167.1 ± 2.6) (sobs = 4511.1 ± 39.2) and LOK1 (I = 22.1 ± 0.0) (sobs = 3989.6 ± 37.2) (Table 14).

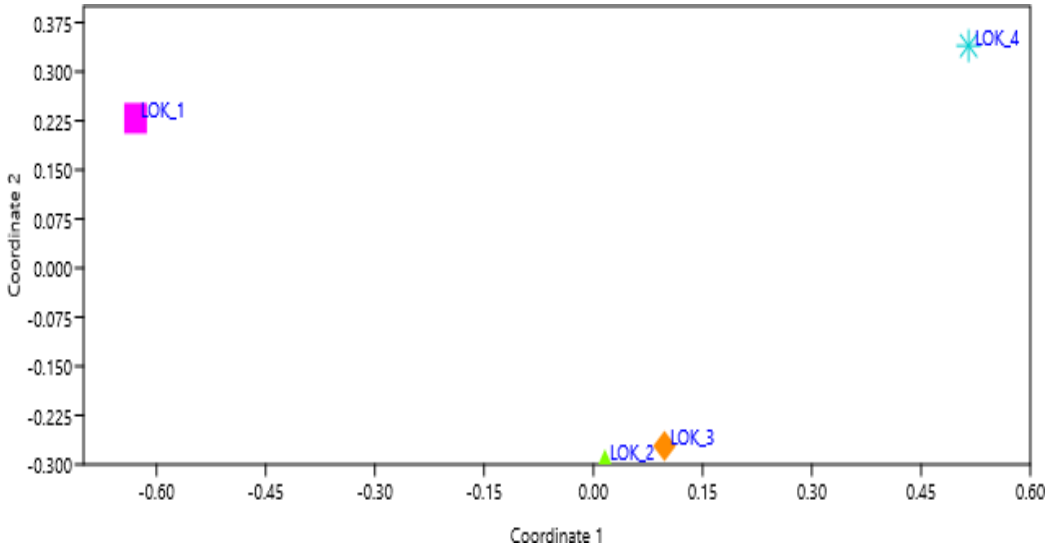


Fig. 10. Non-metric multidimensional scaling (NMDS) plot measured by Bray-Curtis distance between Loktak lake sediment samples; the difference between the samples here demonstrate the differences in bacterial diversity

Table. 14. Comparison of microbial community richness and diversity indices of the four samples of Loktak lake

Cut-off	Sample ID	Shannon(H)	Inverse Simpson	Observed Richness
98 %	Lok1	5.8 ± 0.0	22.1 ± 0.0	3989.6 ± 37.2
	Lok2	6.8 ± 0.0	167.1 ± 2.6	4511.1 ± 39.2
	Lok3	7.3 ± 0.0	342.2 ± 5.5	5177.8 ± 39.9
	Lok4	7.7 ± 0.0	335.2 ± 0.0	6574 ± 0.0

3.1.1.4. Unclassified and unique bacterial Phylotypes

The unclassified OTUs at the phylum level range between 10-15 % of the total OTUs observed suggesting that a certain percentage of unknown bacterial diversity existed. The phylogenetic analysis of these unclassified sequences was conducted using the reference sequences of 16S rRNA genes of validly published or "*Candidatus"* ("*Ca*.") members (IRPCM 2004). LOK3 has the highest number of unique phylotypes i.e sixty-nine ("*Ca*. UH165 - UH237") whereas LOK1 has the lowest number of 46 unique phylotypes ("*Ca*. UH55 - UH103"). The shared distribution of common and unique phylotypes are given in Table 15. The number of overall phylotypes detected and unique phylotypes are 347 and 232 (66.8 %) given in Fig. 11-14.

Table. 15. Common and unique phylotypes shared between the four samples of Loktak lake

Sample name	Total	Unique Phylotypes
LOK1, LOK2, LOK3, LOK4	3	*"Ca*. UH 24", "*Ca*. UH 60", "*Ca*. UH 76"
LOK1, LOK2, LOK3	1	"*Ca*. UH 91"
LOK1, LOK2,	8	"*Ca*. UH 54", "*Ca*. UH 65", "*Ca*. UH 66", "*Ca*. UH 67", "*Ca*. UH 68", "*Ca*. UH 69", "*Ca*. UH 70"
LOK1, LOK3	9	*"Ca*. UH 26", "*Ca*. UH 60", "*Ca*. UH 64", "*Ca*. UH 74", "*Ca*. UH 77", "*Ca*. UH 78", "*Ca*. UH 102", "*Ca*. UH 82", "*Ca*. UH 99"
LOK1, LOK4	2	*"Ca*. UH 75", "*Ca*. UH 90"
LOK2, LOK3, LOK4	2	*"Ca*. UH 59", "*Ca*. UH 117"
LOK2, LOK3	1	"*Ca*. UH 37"
LOK3, LOK4	1	*"Ca*. UH 214"
LOK2, LOK4	2	"*Ca*. UH 25", "*Ca*. UH 26"
LOK1	1	"*Ca*. UH 29"

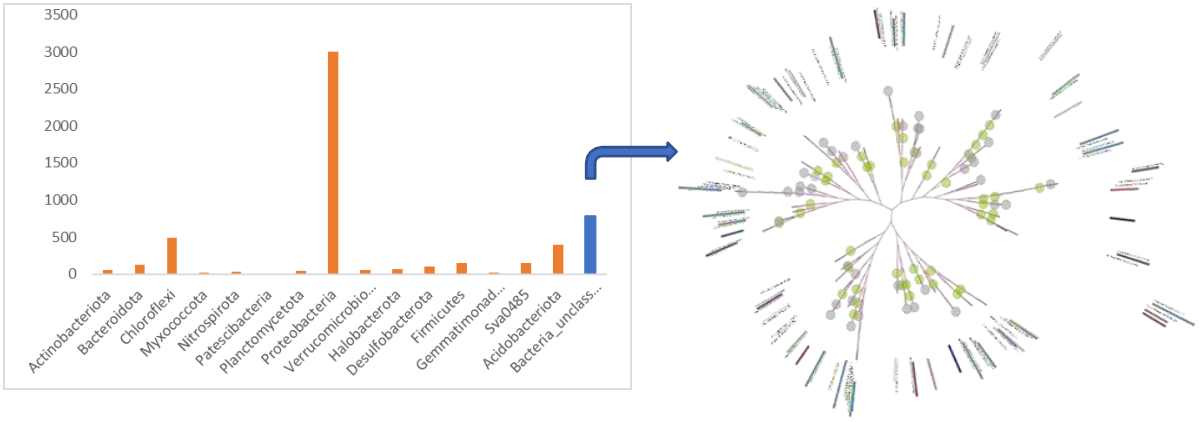


Fig. 11. LOK1: Soil sediment associated bacterial tree of life with with 81 total phylotypes with 46 unique phylotypes ("*Ca*. UH55 - *Ca*. UH103") at the phylum level. Unique phylotypes: green colour circle

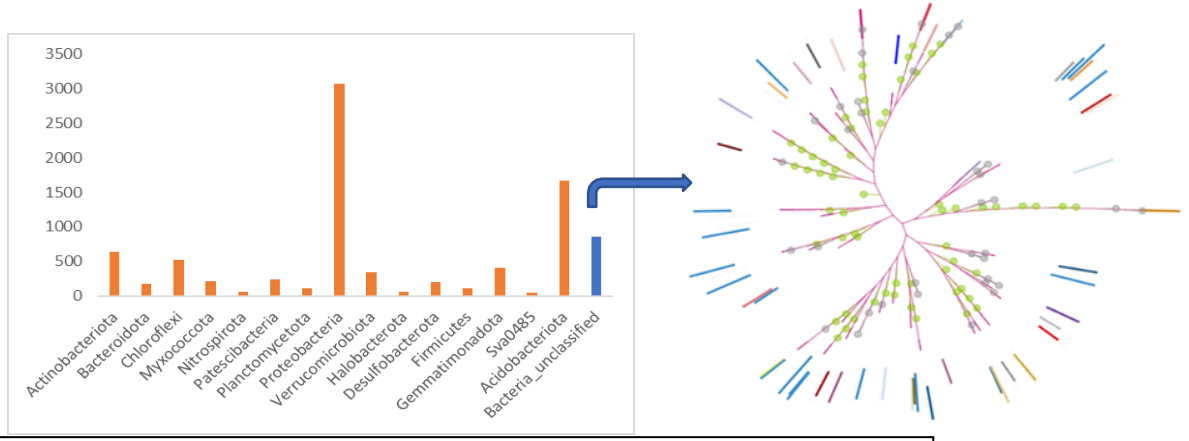


Fig. 12. LOK2: Soil sediment associated with bacterial tree of life with 97 total phylotyptes with 62 unique phylotypes ("*Ca*. UH104 - *Ca*. UH164") at the phylum level. Unique phylotypes: green colour circle

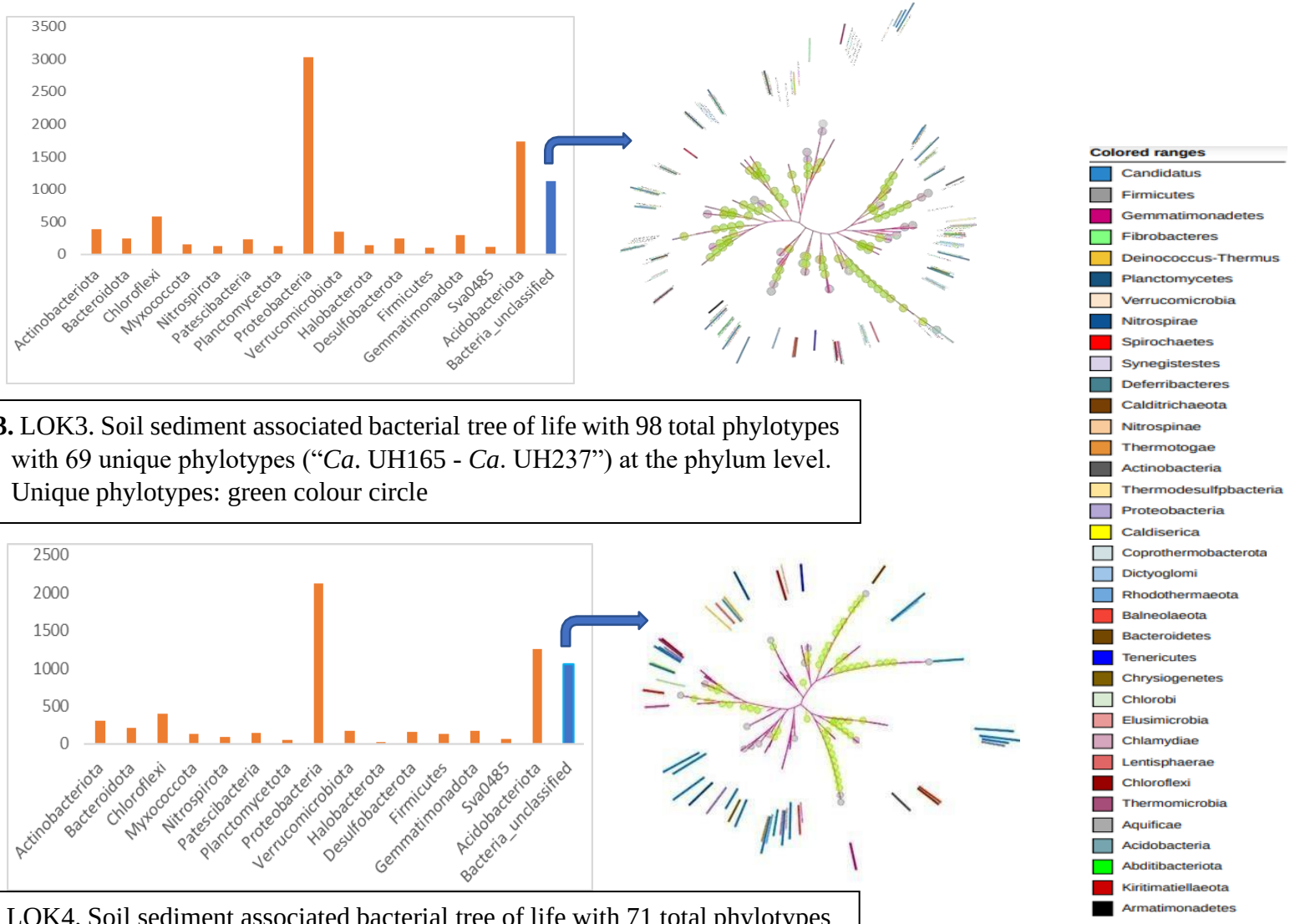


Fig. 13. LOK3. Soil sediment associated bacterial tree of life with 98 total phylotypes with 69 unique phylotypes ("*Ca*. UH165 - *Ca*. UH237") at the phylum level. Unique phylotypes: green colour circle

Fig. 14. LOK4. Soil sediment associated bacterial tree of life with 71 total phylotypes with 55 unique phylotypes ("*Ca*. UH238 - *Ca*. UH 290") at the phylum level. Unique phylotypes: green colour circle

3.1.1.5. Functional profile prediction

The metagenome predicted functions classified using KEGG database in PICRUSt software provides a holistic view of the metabolic pathways of the lake metagenome. KEGG level two and three functional annotations were performed to reveal more information on the functional profile (Fig. 15). Understanding the metabolic capabilities of the microbiomes inferred from amplicon genomic libraries will help in better understanding of how the microbiome affects host metabolism and disease (Steinert et al. 2019). Most of the KEEG Orthologs (KO) at level two are associated with biosynthesis of secondary metabolites, amino acid metabolism, membrane transport, cell growth and death, carbohydrate metabolism, environmental adaptation, signalling molecules and interaction, transcription, lipid metabolism, xenobiotics degradation and metabolism, metabolism of terpenoids and polyketides, cellular process and signalling. *Pseudomonadota*, *Actinomycetota*, *Bacteroidota* were abundantly present in the study and reportedly thought to have important roles in membrane transport, nutrient cycling, and carbon metabolism (Kooh et al. 2017). Various studies have revealed that amino acid metabolism, carbohydrate metabolism and energy metabolism are all related to the main functions of bacteria (Oluseyi et al. 2019, Guan et al. 2019). At a more profound resolution, KEGG level 3 functions also predicted involvement of Loktak lake biotic communities in xenobiotics biodegradation and metabolism (benzoate degradation, aminobenzoate, caprolactam, toluene). Xenobiotics are synthetic chemicals from anthropogenic sources that are not or rarely found in nature (Rieger et al. 2002). They also showed gene related to biosynthesis of amino acid metabolism (arginine, proline, glycine, valine, leucine, isoleucine, cysteine and methionine) and metabolism of terpenoids and polyketides (geraniol, limonene and pinene degradation, phenyl transferase, terpene

backbone biosynthesis). Terpenoids are highly diverse class of organic chemicals and terpene synthases are widely dispersed in bacteria and could be a valuable source for the discovery of novel natural compounds. *Betaproteobacteria* has been reported to serve in bioremediation to anaerobically metabolize benzene, ethylbenzene, xylene and toluene. Functionally related genes for biosynthesis of secondary metabolites like streptomycin, novobiocin, tropane, piperidine and pyridine biosynthesis were also predicted from all the samples. Such a broad secondary metabolite repertory may be essential for the development of microbiome community and host defence.

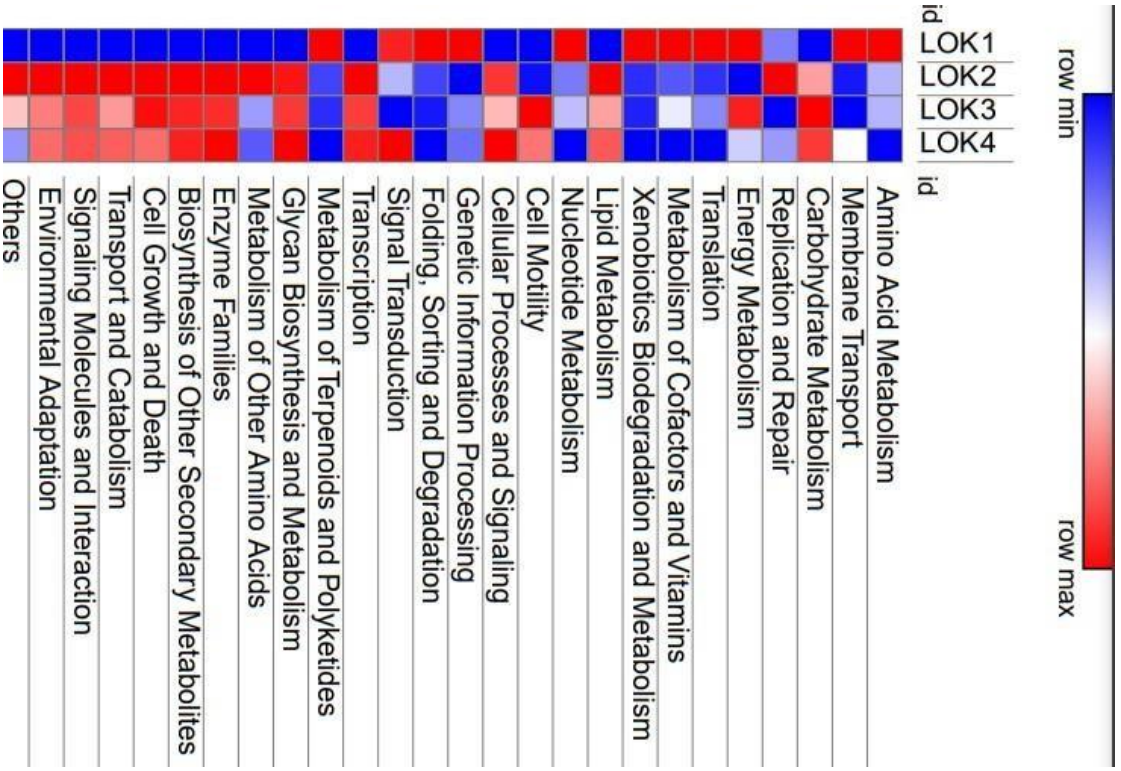
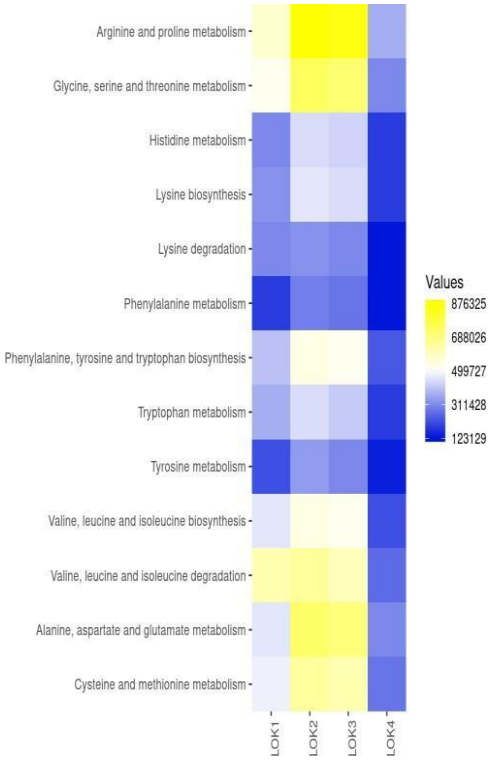
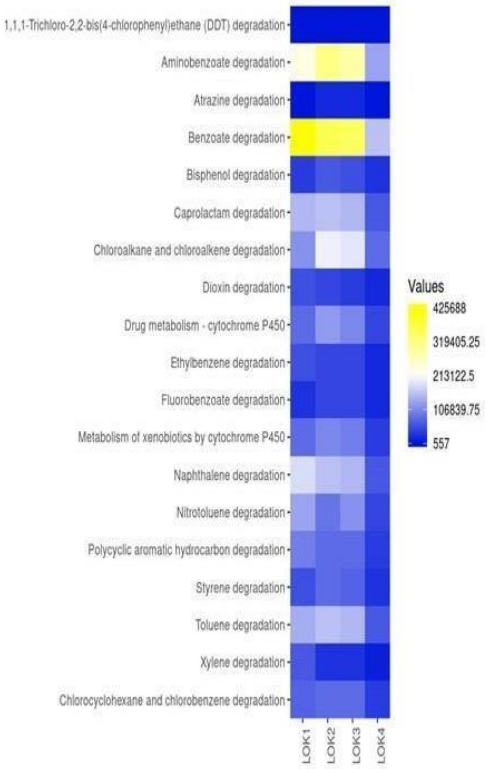


Fig. 15. An overview of the predicted functional analysis of the bacterial microbiome based on Kyoto Encyclopedia of Genes and Genomes (KEGG) between the four samples of Loktak lake

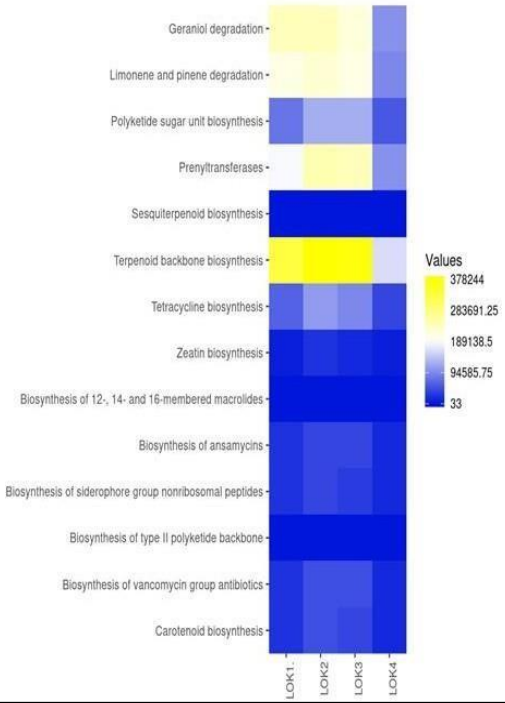
A. Amino acid metabolism



B. Xenobiotics biodegradation and metabolism



C. Metabolism of terpenoids and polyketides



D. Biosynthesis of secondary metabolites

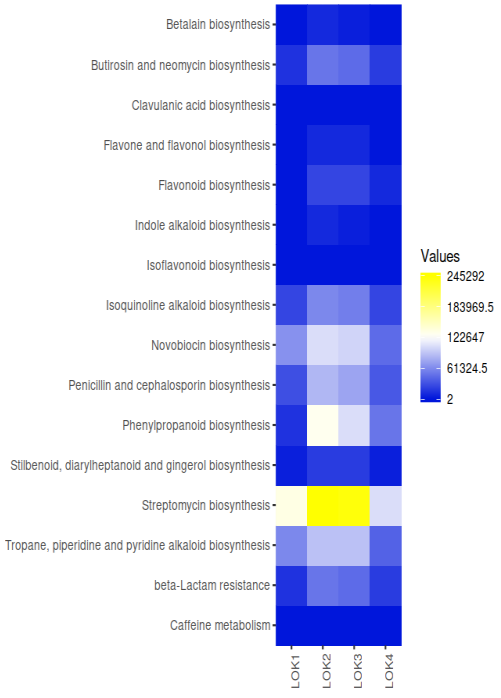


Fig. 16. Detailed analysis of the predicted metabolic functions of the microbiome of the four samples of Loktak lake. A. amino acid metabolism B. xenobiotics biodegradation and metabolism C. metabolism of terpenoids and polyketides D. biosynthesis of secondary metabolites

3.1.1.6. Archaeal abundance and diversity analysis

Archaea are ubiquitous with microbial communities and coexists with other microbes in a niche environment (Kangabam et al. 2020). The three most dominant OTUs belongs to the phylum *Crenarchaeota* (62.3-88.3 %), followed by *Halobacterota* (2.8-22.3 %) and *Thermoplasmatota* (1.9-15 %) (Fig. 17). Phylum like *Euryarchaeota*, *Nanoarchaeota*, were also present in less abundance. Phylum *Aenigmarchaeota* were present only in LOK1 and LOK3, while phylum *Asgardarchaeota* and *Hadarchaeota* were present only in LOK1 and LOK2.

The distribution of genus at the taxonomic level is given in Figure 18. At the genera level, OTUs belonging to *Crenarchaeota* unclassified, *Thermoplasmata* unclassified were most abundant in LOK1 (56.7 %), LOK2 (14.2 %), LOK3 (18.7 %) and LOK4 (14.2 %). Also, genera like *Methanosaeta*, *Methanocella*, and *Methanolinea* were also present in all the samples.

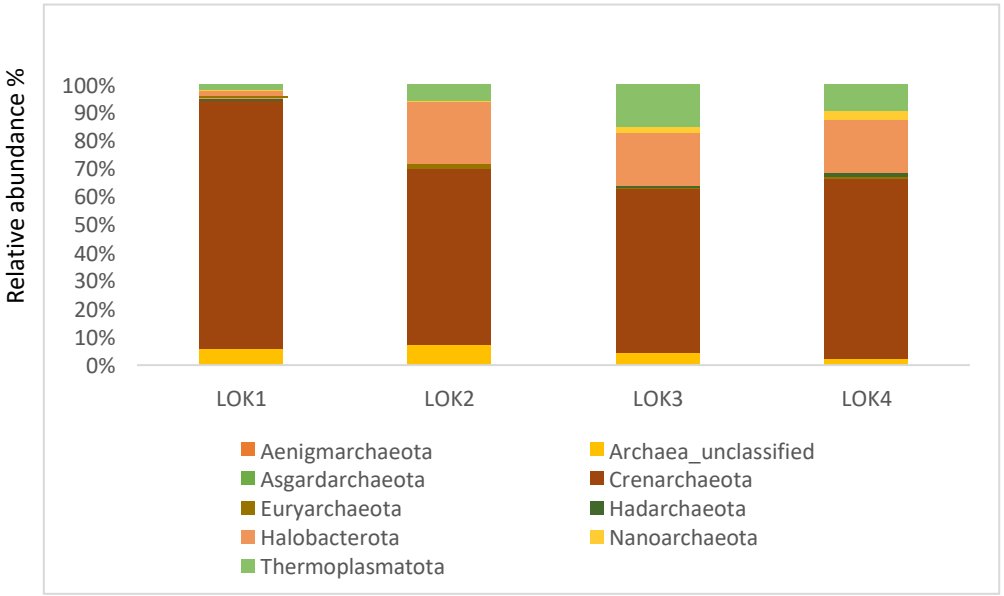


Fig. 17. Relative abundance of archaeal taxa distribution from the four soils samples of Loktak lake with pre-dominant archaeal phyla as obtained by 16S rRNA gene amplicon-based metagenome analysis

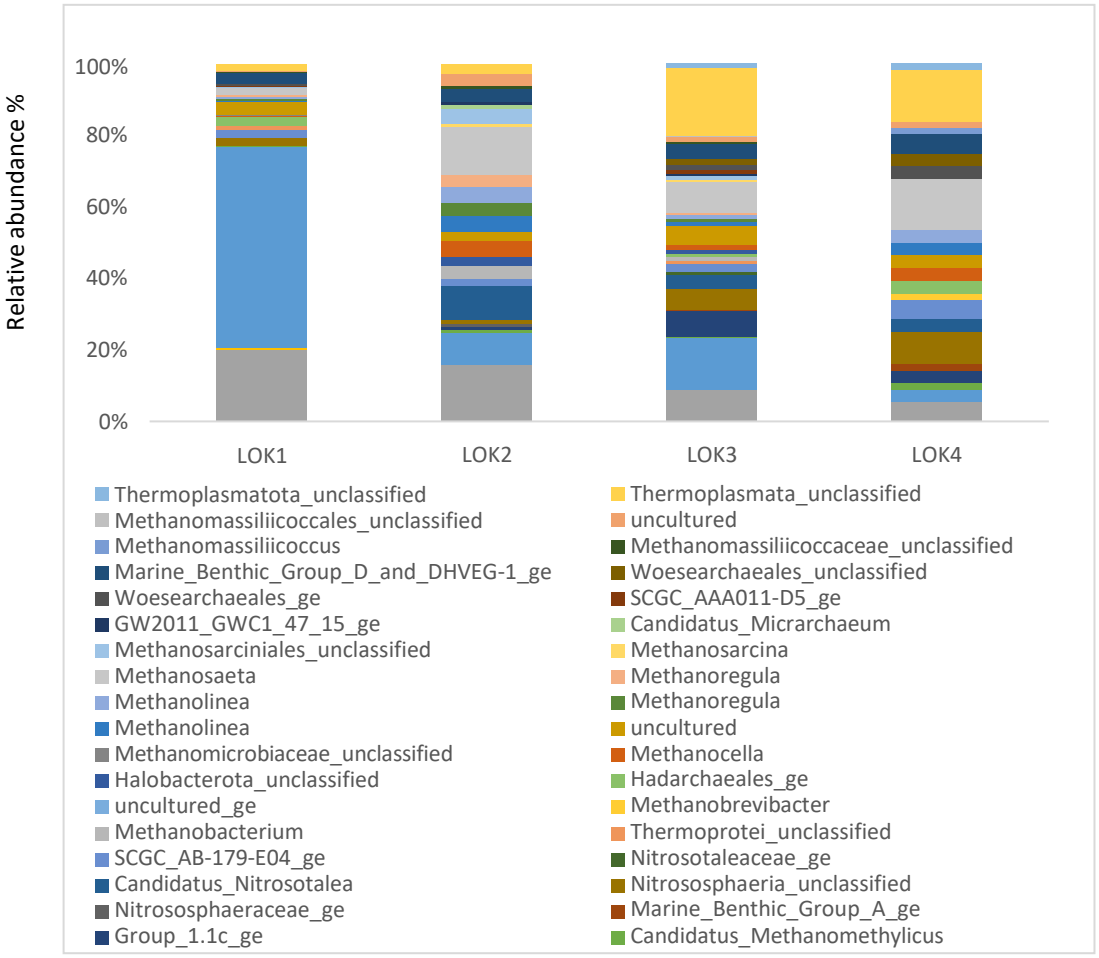


Fig. 18. Relative abundance of archaeal taxa distribution from the four soil samples of Loktak lake with pre-dominant archaeal genera as obtained by 16S rRNA gene amplicon-based metagenome analysis

An overview illustrating the combined archaeal tree of life consisting of eleven phylotypes with eight unique phylotypes has been depicted in Figure 19. Also, the unexplored and novel phyla of Loktak lake along with the existing phyla in the LPSN database is illustrated in the pie graph (Fig. 20). A total of 232 and 8 represents the bacterial and Archaeal ‘‘*Candidatus*’’ phyla obtained from Loktak lake, 53 represents the bacterial ‘‘*Candidatus*’’ phyla obtained from UoH, 89 and 27 represents the known bacterial and archaeal phyla available in the LPSN database.

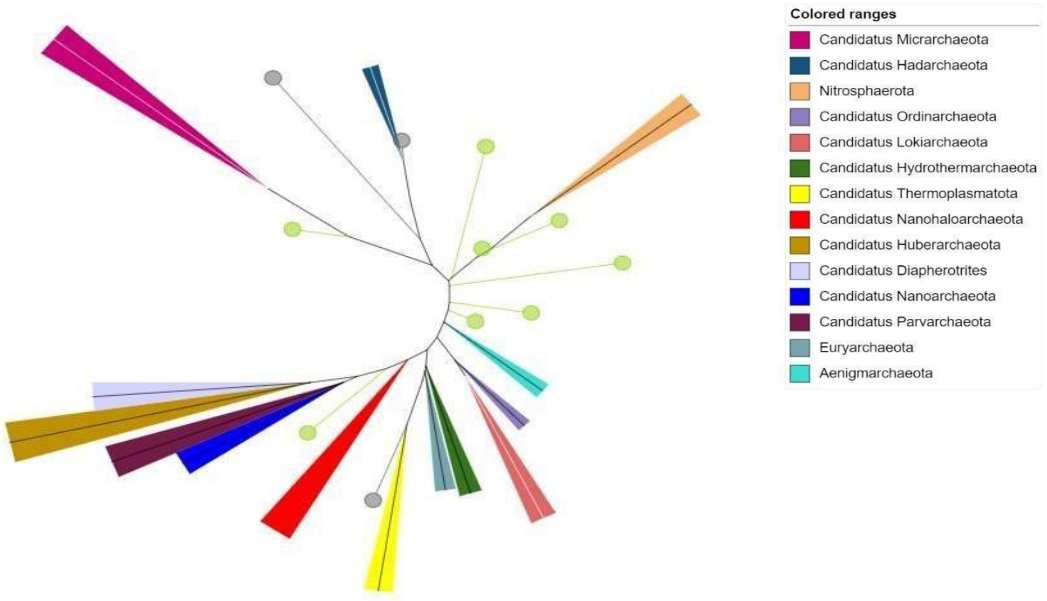


Fig. 19. Archaeal tree of life with 11 phylotypes with 8 unique phylotypes at the Phylum level (LOK1, LOK2, LOK3, LOK4)

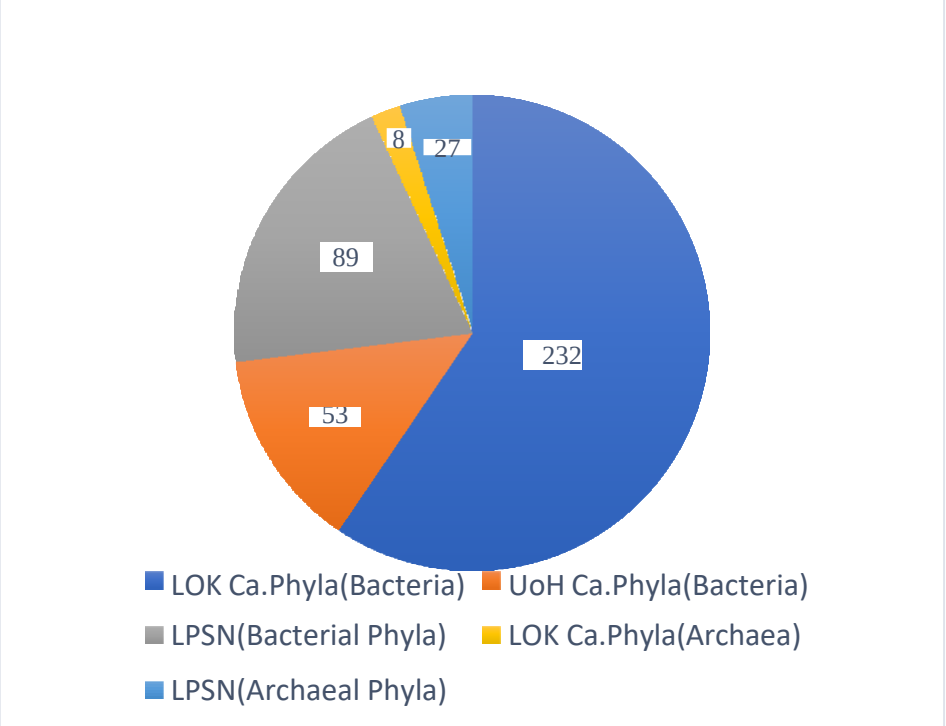


Fig. 20. Unexplored and novel phyla of Loktak lake

3.1.2. Culture dependent study of Loktak lake

3.1.2.1. Identification of bacteria isolated from Loktak lake

Bacteria were isolated from soil sediments and rhizosphere soils of different plants in and around Loktak lake by using standard cultivation techniques. A total of forty-eight bacteria were isolated among which, ten were from LOK1, eleven from LOK2, eleven from LOK3, and sixteen from LOK4 (Table. 16). The 16S rRNA sequences of all the forty-eight bacterial strains were deposited with the NCBI GenBank database whose accession numbers are shown in Table 16. The GPS positioning of the sample collection site ranged between 24˚30' 94'' N - 24˚59' 94'' N, 93˚47' 43 ' E - 93˚87' 06'' E. The temperature of the lake at the time of sampling was between 16-28 ˚C with a pH of 7.0.

Table. 16. The phylogenetic affiliation of the isolated strains along with its accession number, isolation source, and 16S rRNA gene sequence identity with nearest type strain

LOK1

NCBI Accession no.	Organism and strain Number	Isolation source	16S rRNA gene sequence identity with nearest type strain (%)
ON908987	***Glutamicibacter* sp. JC586**	**Rhizopheric soil of *Alocasia cucullata***	***Glutamicibacer halophytocola* KLBMP 5180^T (99.3)**
ON795159	*Cellulosimicrobium* sp. JC802	Rhizospheric soil of *Alpina nigra*	*Cellulosimicrobium cellulans* LMG 16121 (100)
ON908990	***Comamonas* sp. JC664**	**Rhizospheric soil of *Colocasia esculenta***	***Comamonas koreensis* KCTC 12005^T (98.5)**
ON795157	*Pseudomonas* sp. JC7800	Lake sediment	*Pseudomonas hibiscicola* ATCC 19867^T (100)
OP412379	*Aquisphaera* sp. JC813	Rhizospheric soil of *Phragmites karka*	*Aquisphaera giovannonii* OJF2^T (96.4)
MT093437	*Paludisphaera* sp. JC675	Rhizospheric soil of *Dioscorea alata*	*Paludisphaera borealis* PX4^T (94.5)
ON795130	*Rhodococcus* sp. JC773	Rhizospheric soil of *Arundo donax*	*Rhodococcus aetherivorans* 10bc312^T (99.36)
ON795131	*Krasilnikoviella* sp. JC774	Lake sediment	*Krasilnikoviella flava* DSM 21481^T (100)

ON7951691	*Pelagibacterium* sp. JC812	Rhizospheric soil of *Alternanthera phitoxiroides*	*Pelagibacterium lixinzhangensis* H642^T (99.4)
ON795144	*Microbacterium* sp. JC787	Roots of *Ipomoea aquatica*	*Microbacterium resistens* NBRC 103078^T (100)

LOK2

NCBI Accession no.	Organism and strain Number	Isolation source	16S rRNA gene sequence identity with nearest type strain (%)
ON908989	***Janthinobacterium* sp. JC611**	**Lake sediment**	***Janthinobacterium lividum* DSM 1522^T (98.8)**
ON795158	*Brevibacterium* sp. JC801	Lake sediment	*Brevibacterium epidermidis* NBRC 14811^T (99.46)
ON795160	*Bosea* sp. JC803	Rhizospheric soil of *Hedychium coronarium*	*Bosea thiooxidans* DSM9653^T (98.8)
ON795161	*Burkholderia* sp. JC804	Rhizospheric soil of *Brachairia mutica*	*Burkholderia anthina* (R-4183^T (99.8)
ON795162	*Microvirga* sp. JC805	Roots of *Eichhornia crassipes*	*Micro virga soli* R491^T (98.6)
ON795163	*Paludisphaera* sp. JC806	Rhizospheric soil of *Saccharum spontaneum*	*Paludisphaera borealis* PX4^T (94.3)
ON795164	*Glutamicibacter* sp. JC807	Rhizospheric soil of *Lersia hexandra*	*Glutamicibacter arilaitensis* Re117^T (99.3
ON795165	*Leucobacter* sp. JC808	Rhizospheric soil of *Echinochloa stagnina*	*Leucobacter albus* IAM 14851^T (99.7)
ON795168	*Streptomyces* sp. JC811	Rhizospheric soil of *Desmodium motorium*	*Streptomyces ardesiacus* NRRL B-1773^T (99.6)
ON795167	*Williamsia* sp. JC810	Rhizopheric soil of *Alocasia cucullata*	*Williamsia herbipolensis* ARP1^T (100)
ON795166	*Deinococcus* sp. JC809	Rhizospheric soil of *Ranunculus sceleratus*	*Deinococcus grandis* ATCC 43672^T (99.8)

LOK3

NCBI Accession no.	Name and strain Number	Isolation source	16S rRNA gene sequence identity with nearest type strain (%)
ON908988	***Sinomonas* sp. JC656**	**Rhizospheric soil of *Alisma plantago-aquatica***	***Sinomonas notoginsengisoli* SYP-B575^T (98.5)**
ON795145	*Sphingomonas* sp. JC788	Roots of *Hydrilla verticillata*	*Sphingomonas floccifaciens* FQM01^T (99.4)
ON795146	*Microbacterium* sp. JC789	Rhizospheric soil of *Saccharum munja*	*Microbacterium resistens* 103078^T (100)
ON795147	*Pseudomonas* sp. JC790	Rhizospheric soil of *Carex indica*	*Pseudomonas hibiscicola* ATCC 19867^T (99.6)
ON795148	*Micromonospora* sp. JC791	Rhizospheric soil of *Gynura cusimbua*	*Micromonospora aurantiaca* ATCC 27029^T (99.5)
ON795149	*Variovorax* sp. JC792	Rhizospheric soil of *Jussiaea ripens*	*Variovorax guangxiensis* DSM 27352^T (99.6)
ON795150	*Chromobacterium* sp. JC793	Rhizospheric soil of *Crotalaria juncea*	*Chromobacterium piscinae* CCM3329^T (99.8)

ON795151	*Sphingomonas* sp. JC794	Rhizospheric soil of *Colocasia esculenta*	*Sphingomonas pruni* NBRC 15498^{T} (99.8)
ON795152	*Agromyces* sp. JC795	Lake sediment	*Agromyces italicus* DSM 16388^{T} (99.49)
ON795154	*Humibacillus* sp. JC797	Rhizospheric soil of *Mikania scandens*	*Humibacillus xanthopallidus* KV-663^{T} (99.8)
ON795155	*Stenotrophomonas* sp. JC798	Rhizospheric soil of *Cuscuta reflexa*	*Stenotrophomonas maltophilia* MTCC434^{T} (99.5)

LOK4

NCBI Accession no.	**Name and strain Number**	**Isolation source**	**16S rRNA gene sequence identity with nearest type strain (%)**
ON908991	***Paludisphaera* sp. JC665**	**Rhizospheric soil of *Erianthus ravennae***	***Paludisphaera borealis* PX4^{T} (94.5)**
L5782144	*Aquisphaera* sp. JC669	Rhizospheric soil of *Zizania latifolia*	*Aquisphaera giovannonii* (OJF2^{T} (95.2)
ON795156	*Rhodococcus* sp. JC799	Rhizospheric soil of *Centella asiatica*	*Rhodococcus ruber* DSM 43338^{T} (98.9)
ON795153	*Rhodococcus* sp. JC796	Rhizospheric soil of *Cynodon dactylon*	*Rhodococcus coprophilus* NBRC 100603^{T} (99.3)
ON795131	*Krasilnikoviella* sp. JC774	Lake sediment	*Krasilnikoviella flava* DSM 21481^{T} (100)
ON795138	*Methylobacterium* sp. JC775	Rhizospheric soil of *Zizania latifolia*	*Methylobacterium tardum* RB677^{T} (99.8)
ON795132	*Exiguobacterium* sp. JC776	Rhizospheric soil of *Sagittaria sagittifolia*	*Exiguobacterium acetylicum* DSM 20416^{T} (99.7)
ON795133	*Methylobacterium* sp. JC77	Rhizospheric soil of *Ludwigia octovalvis*	*Methylobacterium goesingense* EII3^{T} (99.2)
ON795134	*Brevundimonas* sp. JC778	Rhizospheric soil of *Neptunia oleracea*	*Brevundimonas vesicularis* NBRC 12165^{T} (100)
ON795135	*Mucilaginibacte*r sp. JC779	Lake sediment	*Mucilaginibacter oryzae* DSM 19975^{T} (100)
ON795136	*Glutamicibacter* sp. JC779	Rhizospheric soil of *Marsilea minuta*	*Glutamicibacter uratoxydans* NBRC 15515^{T} (100)
ON795137	*Microbacterium* sp. JC781	Rhizospheric soil of *Oenanthe javanica*	*Microbacterium enclense* NIO-1002^{T} (99.8)
ON795139	*Methylobacterium* sp. JC782	Rhizospheric soil of *Pomea aquatica*	*Methylobacterium komagatae* 002-079^{T} (99.6)
ON795140	*Rhodopseudomonas* sp. JC783	Rhizospheric soil of *Marsilea minuta*	*Rhodopseudomonas telluris* TUT3615^{T} (99.3)
ON795141	*Methylobacterium* sp. JC784	Rhizospheric soil of *Persicaria posumbu*	*Methylobacterium longum* 440^{T} (99.7)
ON795142	*Mycolicibacterium* sp. JC785	Rhizospheric soil of *Polygonium barbatum*	*Mycolicibacterium anyangense* QIA-38^{T} (99.76)

The number of taxa isolated from the four sampling sites is as follows: from LOK1 four taxa were isolated, LOK2 ten taxa, LOK3 nine taxa, and from LOK4 thirteen taxa were isolated. Two overlapping taxa were isolated from LOK 2 and LOK3, while from LOK1, LOK2, and LOK4 only one common taxon was isolated (Fig. 21) (Table. 17).

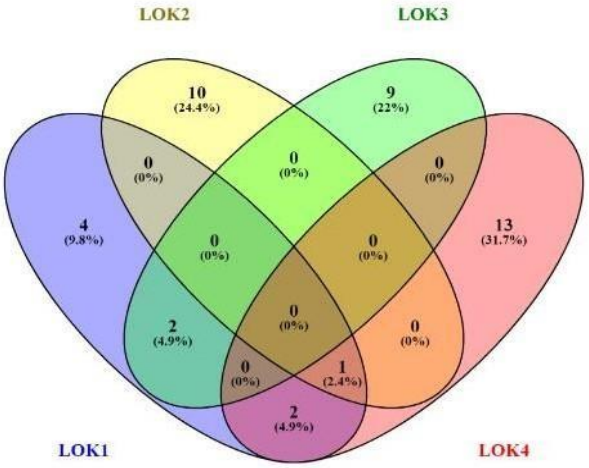


Fig. 21. Distribution of isolated bacterial taxa representing core and shared taxa within the four soil samples of Loktak lake. The number in the overlapping zone indicates the number of taxa shared between the samples, and the number in non-overlapping zone indicates the taxa exclusively isolated from each sample

Table. 17. Distribution of isolated bacterial taxa representing core and shared taxa within the four soil samples of Loktak lake

Sample no.	Total genera	Taxa with highest sequence similarity based on 16S rRNA gene identity
LOK1	4	*Glutamicibacer halophytocola, Cellulosimicrobium cellulans, Comamonas koreensis, Pelagibacterium lixinzhangensis*
LOK2	10	*Janthinobacterium lividum, Brevibacterium epidermidis, Bosea thiooxidans, Burkholderia anthina, Microvirga soli, Glutamicibacter arilaitensis, Leucobacter albus, Streptomyces ardesiacus, Williamsia herbipolensis, Deinococcus grandis*
LOK3	9	*Sinomonas notoginsengisoli, Sphingomonas floccifaciens, Micromonospora aurantiaca, Variovorax guangxiensis, Chromobacterium piscinae, Sphingomonas pruni, Agromyces italicus, Humibacillus xanthopallidus, Stenotrophomonas maltophilia*
LOK4	13	*Rhodococcus ruber, Rhodococcus coprophilus, Methylobacterium tardum, Exiguobacterium acetylicum, Methylobacterium goesingense, Brevundimonas vesicularis, Mucilaginibacter oryzae, Glutamicibacter uratoxydans, Microbacterium enclense, Methylobacterium komagatae, Rhodopseudomonas telluris, Methylobacterium longum, Mycolicibacterium anyangense*
LOK1, LOK3	2	*Pseudomonas hibiscicola, Microbacterium resistens*
LOK1, LOK4	2	*Aquisphaera giovannonii, Krasilnikoviella flava*
LOK1, LOK2, LOK4	1	*Paludisphaera borealis*

According to a thorough floristic survey, numerous plant species are present in and around Phumdi (Devi et al. 2022). Our study also reveals that most of the bacterial strains were isolated from the rhizosphere soil and roots of plants growing in Phumdis and on the surface of the lake. A total of forty-one bacteria were isolated from forty-one different plant species. Among them, twenty-one species were from herbs (*Alocasia cucullata*, *Alternanthera phitoxiroides*, *Lersia hexandra*, *Gynura cusimbua*, *Jussiaea ripens*, *Centella asiatica*, *Sagittaria sagittifolia*, *Oenanthe javanica*, *Marsilea minuta*, *Polygonium barbatum*, *Alpina nigra*, *Colocasia esculenta*, *Alisma plantago-aquatica*, *Saccharum munja*, *Carex indica*, *Gynura cusimbua*, *Sagittaria sagittifolia*, *Neptunia oleracea*, *Pomea aquatica*, *Persicaria posumbu*, *Polygonium barbatum*), eight grass species (*Phragmites karka*, *Brachairia mutica*, *Lersia hexandra*, *Zizania latifolia*, *Cynodon dactylon*, *Erianthus ravennae*, *Saccharum spontaneum*, *Arundo donax*), five shrubs (*Hedychium coronarium*, *Ludwigia octovalvis*, *Desmodium motorium*, *Crotalaria juncea*, *Ranunculus sceleratus*) , three climbers (*Dioscorea alata*, *Mikania scandens*, *Cuscuta reflexa*) and three macrophytes (*Eichhornia crassipes*, *Hydrilla verticillate*, *Ipomoea aquatica*). Further, seven bacterial cultures were also isolated from the sediment of the lake (Fig. 22).

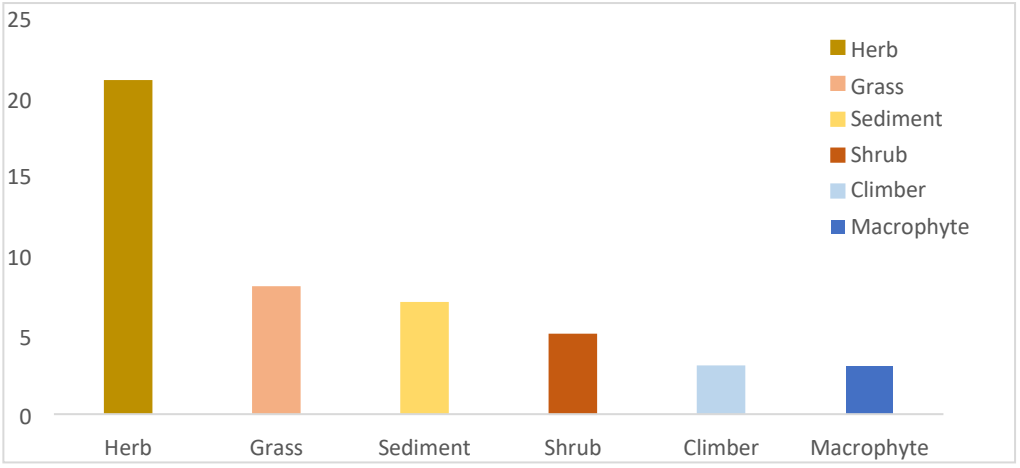


Fig. 22. Total number of taxa isolated from different sources of Loktak lake

3.1.2.2. BLAST analysis and phylogenetic inference

Soil sediments and rhizosphere soils of different plants in and around Loktak lake were collected and subjected for cultivation and isolation of bacteria. A total of 48 bacterial strains were isolated, of which 10 were isolated from LOK1, 11 from LOK2, 11 from LOK3, and 16 from LOK4 (Table. 16). Out of the forty-eight isolates, phylogenomic studies of 5 strains (JC665^T, JC656^T, JC586^T, JC11^T, JC664^T) were carried out (Table. 16). This is primarily based on the16S rRNA nucleotide sequence species delineation (98.7 % cut-off). Of the five strains selected for further characterization, four strains (JC665^T, JC656^T, JC664^T, JC611^T) exhibited 16S rRNA sequence similarity values lower than 98.7 % (except strain JC611^T showed 98.8 %) while strain JC586^T showed sequence similarity higher than the recommended cut-off. Though strain JC586^T shares highest 16S rRNA gene sequence identity (99.3 %) with *Glutamicibacter halophytocola* KLBMP 5180^T and other members of this genus < 99.3 %, the members of the genus *Glutamicibacter* are more ecologically adapted (Table. 26) and hence it was of interest to characterise strain JC586^T. All strains were preserved in 50 % glycerol stock at -20 ˚C. The 16S rRNA gene identity affiliation of the isolated axenic strains along with their source and accession numbers are given in Table 16.

3.1.2.3. Phylogenomic and polyphasic studies of novel strains, if any

3.1.2.3.1. Characterization of strain JC665^T

3.1.2.3.1.1. Habitat, isolation and nucleotide accession numbers

Strain JC665^T was isolated from the rhizosphere soil of *Erianthus ravennae* (commonly known as “Plume grass”) of the “phumdis” (floating island) of Loktak lake

located in the Northeast part of India, Manipur (precise location: 24°30'21" N 93°47'43" E). The pH and temperature of the lake at the time of sampling was pH 7.0 and 22 °C. Strain JC747 was isolated from a sediment sample of a wetland located in Pallikkara village, Kerala, India (12° 23' 02" N and 75° 02'33" E). The GenBank accession numbers of the 16S rRNA gene sequence and genome sequence of strains JC65^T and JC747 are LR746340, OU374731 and JAALCR000000000 and JAHPZK000000000, respectively.

3.1.2.3.1.2. BLAST analysis and phylogenetic inference

BLAST analysis of 16S rRNA gene sequence of strains JC665^T (1521 nt) and strain JC747 (1521 nt) in EzBioCloud server showed identity of 94.6 % and 96.7 % with *P. borealis* PX4^T and "*P. soli*" JC670^T, respectively (Fig. 23). Comparison of *d*DDH, *g*ANI, and AAI values of strains JC665^T and JC747 with *Paludisphaera* spp. yielded similarity of (19.4-20.3) %, (62.4-68.6) %, and (75.1-77.9) %, respectively. Also, the AAI, OrthoANI and *d*DDH values fall below the recommended cut-off of 80 %, 95-96 % and 70 %, respectively for prokaryotic species delineation (Rodriguez and Konstantinidis 2014; Meier-Kolthoff et al. 2014; Chun et al. 2018). Thus, both the newly isolated strains represent a novel species of the genus *Paludisphaera*. However, high values of 16s rRNA (100 %), *d*DDH (100 %), *g*ANI (100 %), and AAI (99.9 %) between strains JC665^{T}and JC747 indicate that these strains are members of the same species from two distinct and distantly located ecosystems. The 16S rRNA gene sequence based phylogenetic tree with combined bootstrap values obtained from NJ, ME, ML trees (Fig. 24) and 92 core genes-based phylogenomic tree (Fig. 25) confirmed that strains JC665^T and JC747 forms a distinct monophyletic clustering with *Paludisphaera* members within the family *Isosphaeraceae* suggesting that strain JC665^T belongs to a novel species within the genus *Paludisphaera*.

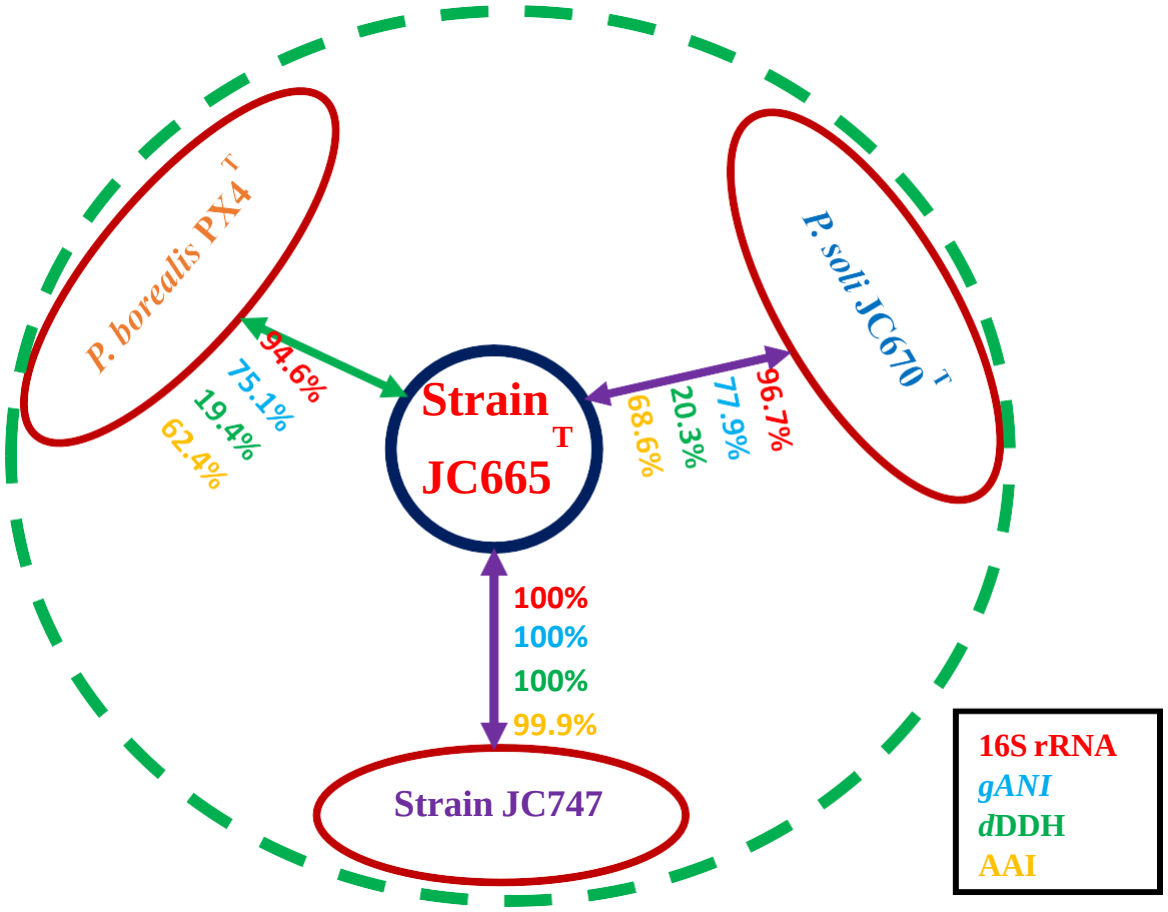


Fig. 23. Analysis of phylogenetic markers for delineation of the novel isolate strain JC665 T. Methods used: 16S rRNA gene identity (16S), average amino acid identity (ANI), average amino acid identity (AAI), digital DNA-DNA hybridization (*d*DDH)

3.1.2.3.1.3. Genomic characteristics

Genome size of the strains JC665T and JC747 are 8.05 Mb and 8.04 Mb with an N_{50} value of 238,467 and 226,135, respectively. The genome of the strain JC665T has 6,431genes of which 6,307 are protein-coding genes, 80 genes code for RNAs (3 genes for encoding r-RNAs, 74 genes for t-RNAs and 3 for other RNAs) and 44 genes are pseudogenes (Table. 18). The genome of the strain JC747 has 6,420 genes of which 6, are protein-coding genes, 80 genes code for RNAs (3 genes for encoding r-RNAs, 74 genes for t-RNAs and 3 for other RNAs) and 41 genes

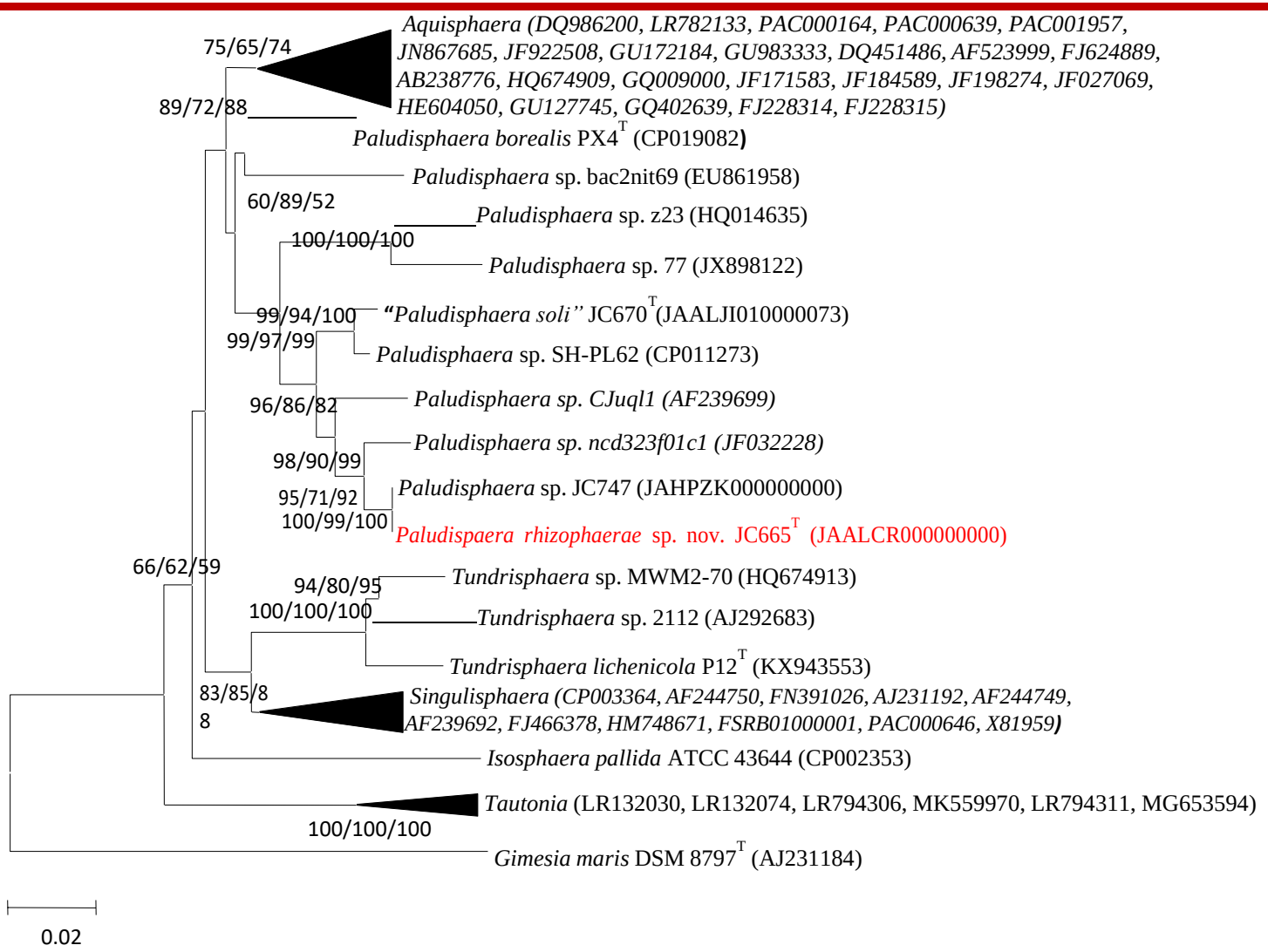


Fig. 24. Phylogenetic tree based on 16S rRNA gene sequences showing the phylogenetic relationship of strains JC665^T, JC747 and other closely related species of *Isosphaeraceae*. The tree was constructed using MRGA 7 software and rooted using *Gimesia maris* DSM 8797^T as the outgroup. Bootstrap percentage refers to NJ/ME/ML analysis. Bar, 0.02 nucleotide substitution per positon

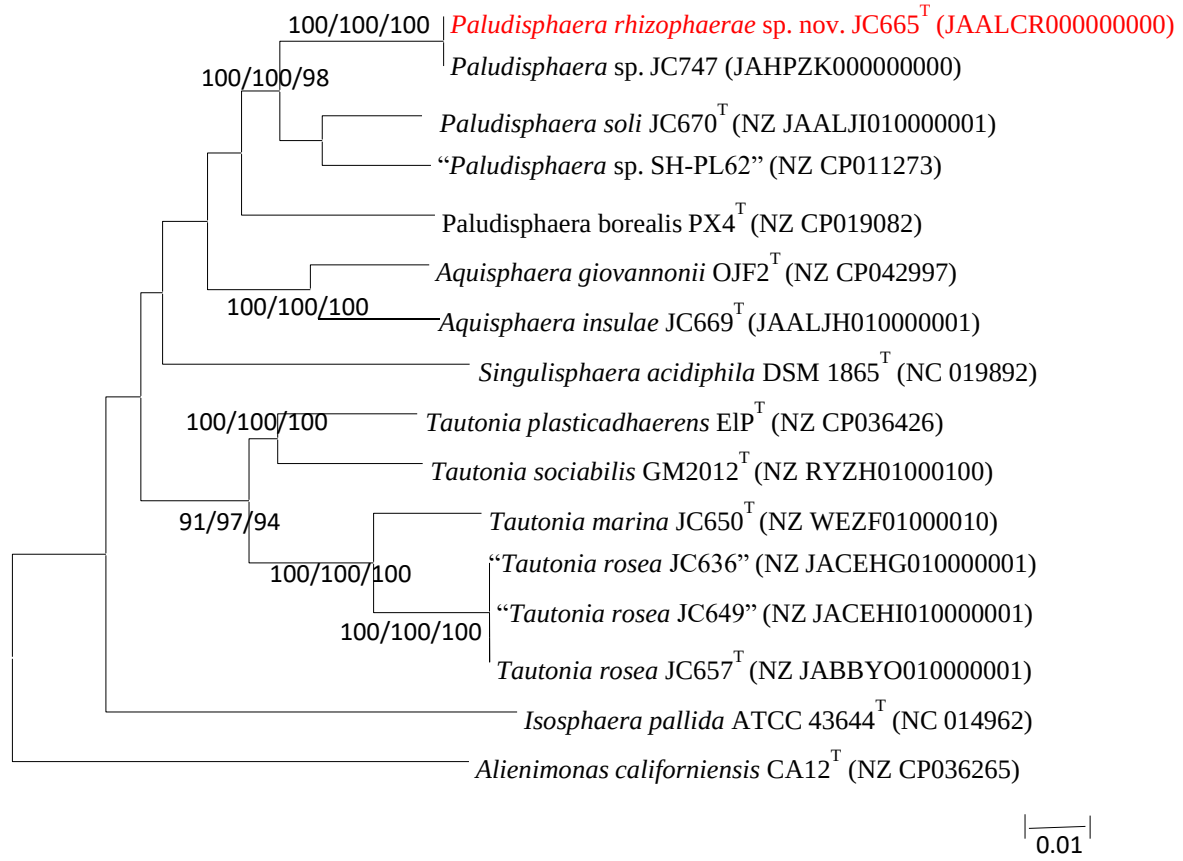


Fig. 25. RAxML Phylogenomic tree of strains JC665^T and JC747 along with publicly available genome sequences of *Isosphaeraceae* family. The GenBank accession numbers of genome sequences are shown in parentheses. The tree was computed with MEGA7 software and rooted with *Alienimonas californiensis* CA12^T as the out- group. Bootstrap percentage refers to ML/ME/NJ analysis. Bar, 0.1 nucleotide substitution per position.

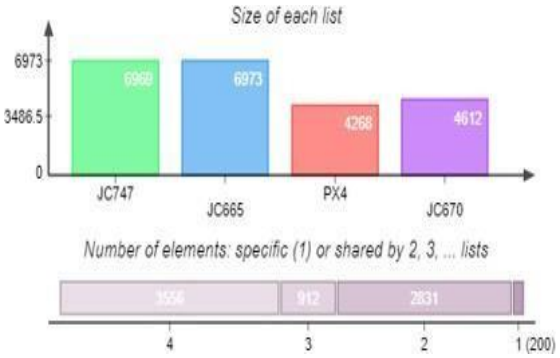
are pseudogenes (Table. 18). The genome-wide annotation predicted, 7106, 7102, 7373 and 6643 proteins for strains JC665^T, JC747, "*P. soli*" JC670^T and *P. borealis* PX4^T respectively. The predicted protein-coding gene of strains JC665^T and JC747 are 6973 and 6969 orthologous clusters and 103 singletons for which no orthologs could be found in other species (Fig. 26A). "*P. soli*" JC670^T showed 4612 clusters and 2464 singletons whereas *P. borealis* PX4^T showed 4268 clusters and 2160 single tons (Fig. 26B). The

comparative analysis of orthologous gene clusters performed shows that these species formed 7499 clusters, 4032 orthologous clusters (at least contains two species) and 3467 single-copy gene clusters (Fig. 26C). A total of 14738 proteins were present in orthologous clusters in all the strains, however, 314 and 190 proteins were present in strains JC670^{T} and PX4 314, specific to them (Fig. 26C). Further comparison of shared orthologous gene clusters revealed that all strains had 3556 clusters, however, "*P. soli*" JC670^{T} and *P. borealis* PX4^{T} showed 127 and 73 unique clusters respectively (Fig. 26D). Genomic DNA G+C content of both strains (JC665^{T}, JC747) is 66.4 mol % (Table 18).

Table. 18. Comparative genomic features among strains belonging to genus *Paludisphaera* including strain JC665^{T}, strain JC747, strain PX4^{T} and strain JC670^{T}

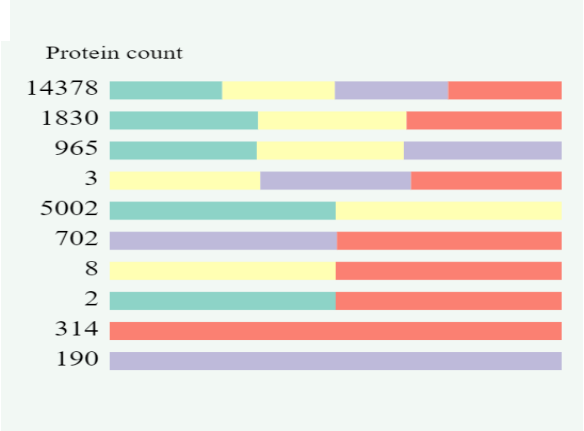
Accession No	Sequencing Technology	Genome coverage	Genome size (Mb)	No. of contigs	GC %	N50 (bp)	Genes (total)	CDS (Total)	CDS (Protein)	rRNA	tRNA	Other RNA	Pseudo gene	CRISPRs
NZ_JAHPZK000000000	Illumina HiSeqX10	100.0x	8	115	66.4	226,135	6,420	6,340	6,299	3	74	3	41	1
NZ_JAALCR000000000	Illumina HiSeqX10	100.0x	8	110	66.4	238,467	6,431	6,351	6,307	3	74	3	44	1
NZ_CP019082	PacBio RSII	60X	7.7	3	66.3	7,497,544	5,870	5,785	5,699	9	73	3	86	1
NZ_JAALJI010000001	Illumina HiSeq	100.0x	8	197	70.4	217,617	6,453	6,392	6,213	3	55	3	179	0

A)



B)

C)



D)

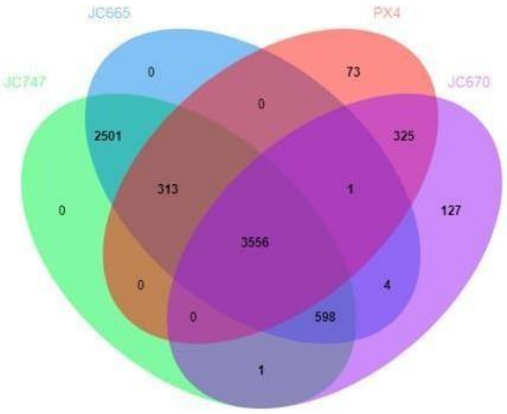


Fig. 26. A) Predicted protein-coding gene of strains JC665^T, JC747, PX4^T and JC670^T
B) Heat map showing the comparison of orthologous gene clusters among members of genus *Paludisphaera*
C) Total number of proteins present in orthologous clusters in all the strains i.e. JC747, JC665^T, PX4^T and JC670^T, and
D) Venn diagram depicting the comparison and distribution of proteins among the strains JC747, JC665^T, PX4^T and JC670^T

3.1.2.3.1.4. *In-silico* metabolic characterisation

COGs annotation of strains, JC665^T and JC747 showed similar results as that of the other members of the genus *Paludisphaera.* Most of the genes predicted were of unknown function followed by genes involved in energy production and conversion (Fig. 27). The CAZy annotation of genomes also shows that the strains JC665^T and JC747 contains more

genes encoding glycoside hydrolases followed by glycosyl transferases compared to that of other members of the genus *Paludisphaera.* The analysis also shows that 80-90 % of the enzymes belonged to families of glycoside hydrolases and glycosyl transferases (Fig. 28). *In silico* metabolic characterisation revealed that strains JC665^{T} and JC747 have the 1-deoxy-D-xylulose 5-phosphate/2-C-methyl-Derythritol 4-phosphate pathway for the biosynthesis of five carbon isoprene units (isopentenyl pyrophosphate), which serves as a precursor for the synthesis of quinones and carotenoids The putative genes encoding for enzymes like phytoene desaturase [EC:1.3.99.26 1.3.99.28 1.3.99.29 1.3.99.31] , 15-cis-phytoene synthase [EC:2.5.1.32], which helps in formation of lycopene as an end product with intermediary products like phytoene, zeta carotene and neurosporene were predicted in the genomes of both strains. The genes for assimilatory nitrate reduction were observed only in *P. borealis* PX4^{T}. All the strains belonging to genus *Paludisphaera* along with JC665^{T} and JC747 showed putative gene clusters for synthesis of Type I, Type III polyketide synthases and terpene biosynthesis. However, genes for production of indole were predicted exclusively in JC665^{T} and JC747 only (Fig. 29). *In-silico* genome analysis of all the four strains of *Paludisphaera* also showed the presence of the putative hopanoid biosynthesis pathway genes like squalene/phytoene desaturase (*hopC*), squalene synthase (*hpnC),* radical S-adenosyl-L-methionine (SAMe) required for addition of adenosyl group to hopane skeleton (*hpnH*), squalene hopene cyclase (*Shc*; codes for the key enzyme of hopanoid biosynthesis), hopanoid associated sugar epimerase (*hpnA),* acetylornithine aminotransferase/amino-bacteriohopanetriol synthase (*hpnO*), and sterol desaturase family proteinein (*erg32*).

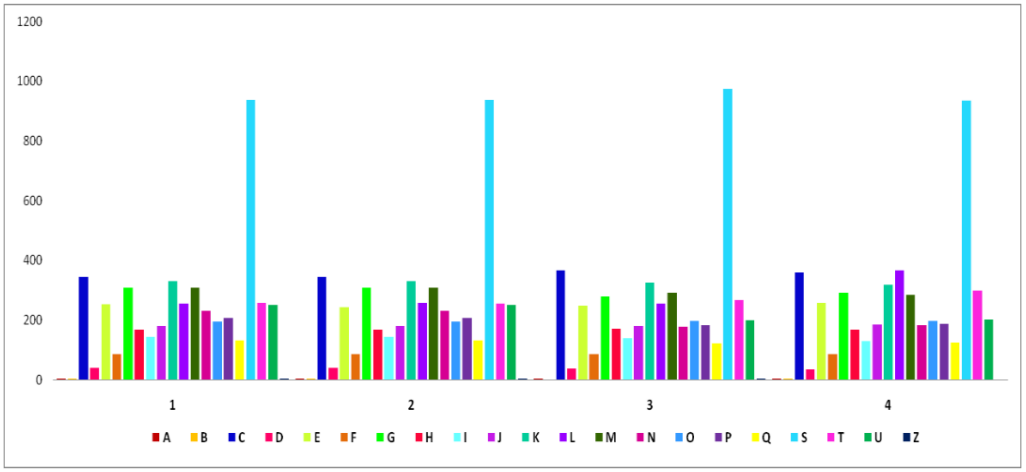


Fig. 27. Functional classification COGs clusters of orthologs of two isolates and related strains in the genus of *Paludisphaera* A: RNA processing and modification; B: Chromatin Structure and dynamics; C: Energy production and conversion; D: Cell cycle control and mitosis; E: Amino Acid metabolis and transport; F: Nucleotide metabolism and transport; G: Carbohydrate metabolism and transport; H: Coenzyme metabolism; I: Lipid metabolism; J: Tranlsation; K: Transcription; L: Replication and repair; M: Cell wall/ membrane/ envelop biogenesis; N: Cell motility; O: Post-translational modification, protein turnover, chaperone functions; P: Inorganicion transport and metabolism; Q: Secondary Structure; S: Function Unknown; T: Signal Transduction; U: Intracellular trafficing and secretion; V: Defense mechanisms; Z: Cytoskeleton; 1: strain JC665^{T}, 2: strain JC747, 3*: P. soli* JC670^{T}, 4: *P. borealis* PX4^{T}

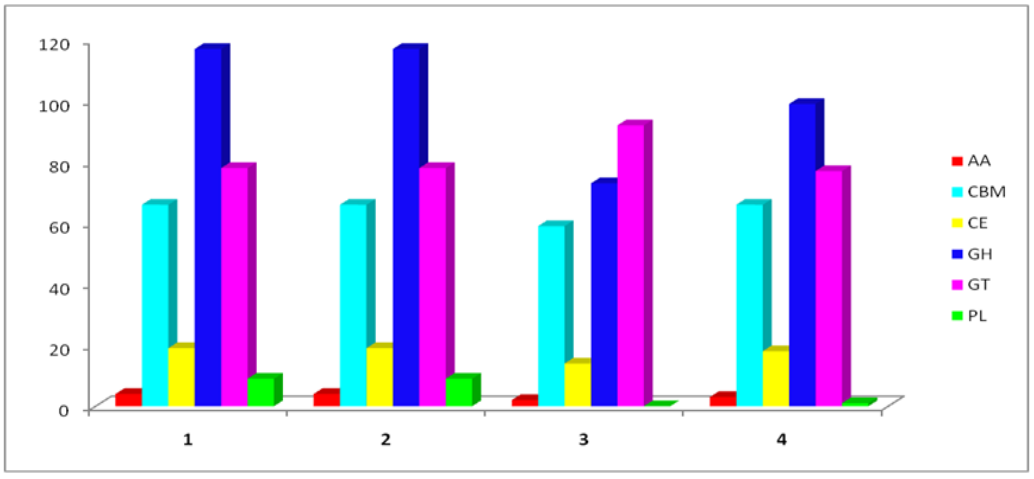


Fig. 28. Comparative analysis of CAZymes of strains JC747, JC665^{T}and other related members from genus *Paludisphaera* 1. Strain JC747 2. strain JC665^{T} 3. *Paludisphaera borealis* PX4^{T} 4. *Paludisphaera soli* JC670^{T}. Abbreviations: AA, Auxiliary Activities (AAs); CBM, Carbohydrate-Binding Modules (CBMs); CE, Carbohydrate Esterases (CEs); GH, Glycoside Hydrolases (GHs); GT,GlycosylTransferases (GTs); PL, Polysaccharide

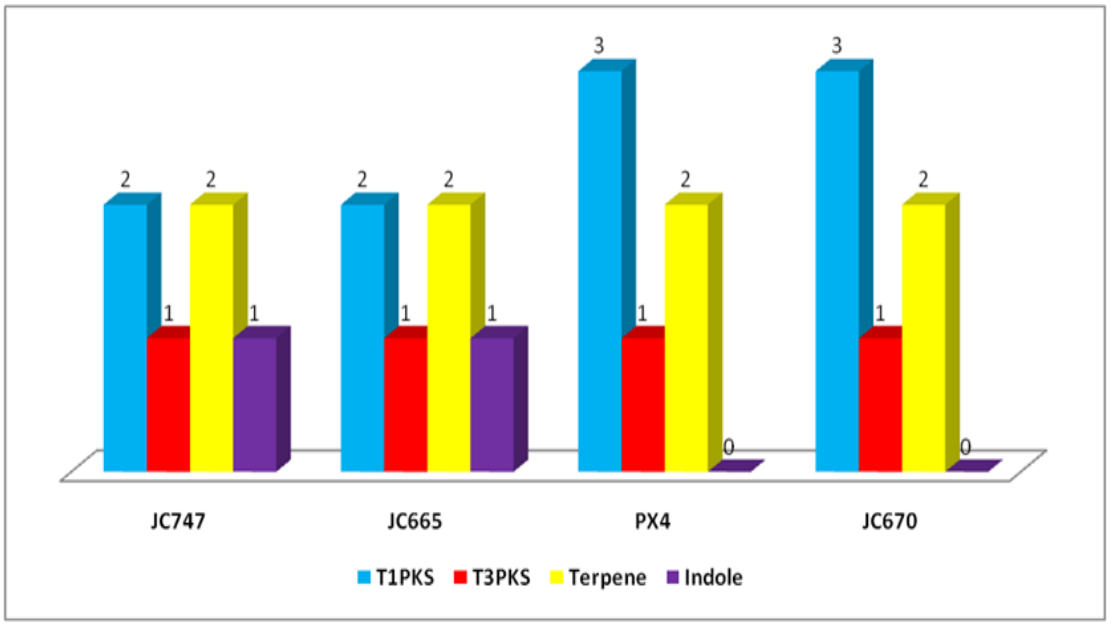


Fig. 29. Number of genetic clusters putatively involved in the production of secondary metabolites. The analysis is based antismash analysis of genome sequences 1. Strain JC747 2. Strain JC665^T 3. *Paludisphaera borealis* PX4^T 4. *Paludisphaera soli* JC670^T

3.1.2.3.1.5. Morphological and physiological analysis

SEM images show that cells of strains JC665^T and JC747 are spherical to oval-shaped (1.7-1.8x1.3-1.5 µm; Fig. 30A) and have a well distribution of crateriform structures (CR) all over the cell surface. SEM image of strain JC665^T also shows the presence of two distinct morphological i.e. mother cells and daughter cells with crateriform structures which indicate the dimorphic life cycle of strain JC665^T (Fig. 31).

TEM image of the cells shows the presence of cytoplasmic membrane (CM), cytoplasm (CP), outer membrane (OM), invagination of the cytoplasmic membrane (ICM), ribosomes (RB) nucleoid region (N) and cell reproduction by budding (BD) where daughter cell is protruding from mother cell (Fig. 30B).

NAG is not obligate for the growth of all four strains i.e. JC665^T, JC747, "*P. soli*" JC670^T and *P. borealis* DSM 28747^T. All the four strains utilize the following organic carbon sources: sucrose, α-D-glucose, D-galactose, Na-pyruvate, mannose, rhamnose and trehalose. Neither of the strains (strains JC665^T, JC747) utilizes the following organic carbon sources: starch, mannitol, acetone, ascorbate, inulin, acetate, malate, Na-succinate, benzoate and citrate. Lactose and maltose are utilized by the strains JC665^T, JC747, and *P. borealis* DSM 28747^T. Fructose and D-xylose are utilized by the strains "*P. soli"* JC670^T and *P. borealis* DSM 28747^T. Cellobiose and ribose are solely utilized by the strain *P. borealis* DSM 28747^T. Fumarate and propionate are exclusive for the strains JC665^T (including JC747) and "*P. soli"* JC670^T, respectively. The following nitrogen sources are utilized by all the strains for their growth: yeast extract, ammonium sulphate, peptone, L-arginine, casamino acid, DL-alanine and sodium nitrate. The following nitrogen sources are not utilized by any of the strains: urea, L-aspartic acid and valine. The following nitrogen sources are exclusively utilized by the strains JC665^T, JC747 and "*P. soli"* JC670^T: L-lysine, L-glycine, L- glutamine, L-phenylalanine, L-proline, L-leucine, L-isoleucine, DL-ornithine and DL-threonine. However, cysteine and L-methionine are solely utilized by the strains JC665^T and JC747. L-serine and L-tyrosine is exclusive for the strain "*P. soli"* JC670^T (Table 20)**.** Strain JC665^T can hydrolyse phytagel only in the absence of N-acetylglucosamine in the medium (Fig. 32).

All the four strains showed positive for esterase (C4), valine arylamidase and leucine arylamidase. However, all the strains show negative for cysteine arylamidase, lipase (C14), trypsin, α-galactosidase, α-chymotrypsin, α-glucosidase, β-glucuronidase, α-mannosidase, β- glucosidase, and α-fucosidase. Esterase lipase (C8), Alkaline phosphatase, and acid phosphatase are exclusively positive for the strains "*P. soli"* JC670^T

and *P. borealis* DSM 28747^{T}. Naphthol-AS-BI-phosphohydrolase shows positive for the strains JC665^{T}, JC747 and "*P. soli"* JC670^{T}. N-acetyl- β-glucosaminidase and β-galactosidase are exclusively positive for the strain *P. borealis* DSM 28747^{T}.

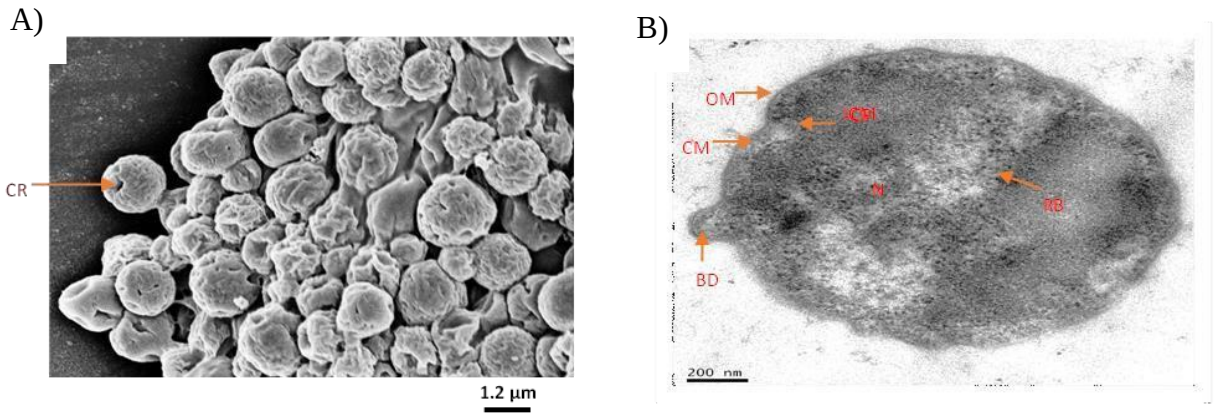


Fig. 30. Scanning (A) and transmission (B) electron micrographs of the cell of strain JC665^{T}. A. Cells having crateriform structures (CR). Bar, 0.3 µm., B. Ultrathin section showing invagination of cytoplasmic membrane (ICM), outer membrane (OM), cytoplasmic membrane, cytoplasm (CP), Ribosome (RB), nucleoid region (N) and cell of strain JC665^{T} multiply by budding (BD) wherein daughter cells are smaller than parent cells. Bar, 0.2 µm.

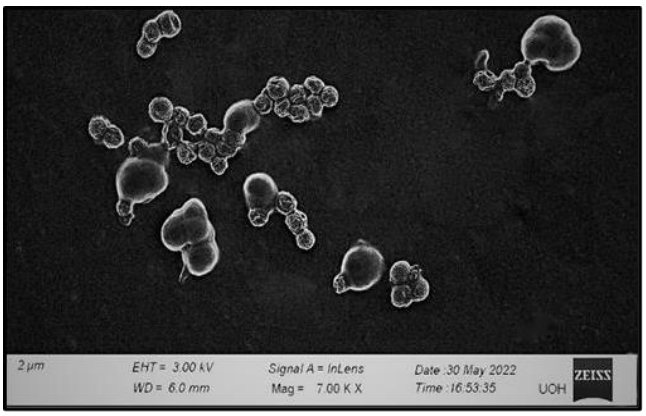


Fig. 31. Scanning electron micrograph showing the presence of two distinct morphological i.e. mother cells and daughter cells with crateriform structures which indicate the dimorphic life cycle of strain JC665^{T}

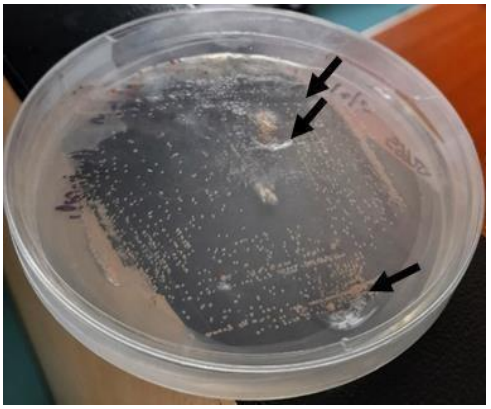


Fig. 32. Hydrolysis of phytagel by strain JC665^{T}and crateriforms are indicated in black arrows

3.1.2.3.1.6. Chemotaxonomic Characterisation

3.1.2.3.1.6.1. Cellular fatty acid, polar lipid, polyamine, and quinone

$C_{18:1}\omega9c$, $C_{18:0}$ and $C_{16:0}$ are the major fatty acids in strains JC665^T, JC747, "*P. soli*" JC670^T and *P. borealis* DSM 28747^T (Table. 19).

Table. 19. Fatty acids composition of *Paludisphaera* strains including JC665^T, JC747, *P. soli* JC670^T and *P. borealis* PX4^T

Fatty acid (%)	**JC665^T**	**JC747**	***P.soli* JC670^T**	***P. borealis* PX4^T**
anteiso $C_{11:0}$	6.8	-	-	-
anteiso $C_{12:0}$	1.42	-	-	-
$C_{13:0}$	4.3			
$C_{14:0}$	0.4	-	0.80	3.4
anteiso $C_{15:0}$	3.70	-	-	-
$C_{15:2}$ OH	1.0	-	-	-
$C_{15:1}$ ω5c	1.0	-	-	-
$C_{16:0}$	**16.6**	**27.65**	**15.51**	**17.4**
$C_{16:0}$ N alcohol	1.4	-	-	-
$C_{16:1}\omega9c$	-	-	1.8	1.8
$C_{17:0}$	1.12	-	0.27	-
$C_{17:1}\omega8c$	1.0	1.02	1.43	0.4
anteiso-$C_{17:0}$	2.44	-	5.5	-
$C_{18:0}$	**9.23**	**10.68**	**6.78**	**12.4**
$C_{18:1}\omega9c$	**45.30**	**59.53**	**67.63**	**47.9**
$C_{18:1}\omega7c$/ $C_{18:1}\omega6c$		-	2.3	-
$C_{18:3}\omega6c,9c,12c$	1.2	-	1.88	-

The polar lipids of strains JC665^T and JC747 comprise of phosphatidylcholine (PC), two unidentified glycolipids (GL1, 2), six unidentified lipids (UL1-7) and two unidentified phospholipid (PL1, 2) (Fig. 33). The polar lipids of strain *P. borealis* DSM 28747^T comprise of phosphatidylethanolamine (PE), phosphatidylcholine (PC), one unidentified choline lipid (CL1), two unidentified lipids (UL1, 2), two unidentified glycolipids (GL1, 2), two unidentified amino lipids (AL1, 2) and four unidentified phospholipids (PL3-6) (Fig. 33). The polar lipids of strain "*P. soli*" JC670^T include phosphatidylcholine, two unidentified phospholipids and six unidentified lipids (Fig. 33) Putrescin and sym-

homospermidine are the major polyamine of strains JC665^T and "*P. soli"* JC670^T whereas spermidine and two unidentified polyamines (1, 3) are present in strain JC747. Strain *P. borealis* DSM 28747^T consists of spermidine and two unidentified polyamines (2, 3). (Fig. 34). MK6 is the predominant quinone for all the strains.

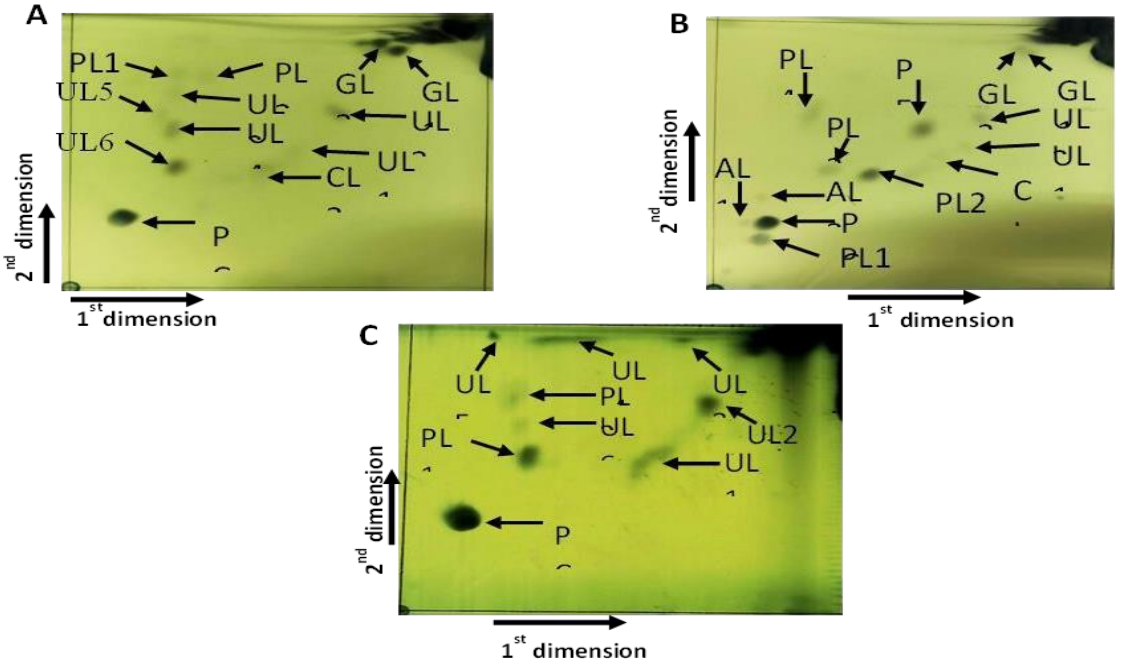


Fig. 33. Two dimensional chromatograms showing polar lipids of strains (A) JC747 and JC665^T (B) *P. borealis* PX4^T (C) *P. soli* JC670^T
PC, Phosphatidylcholine; CL, unidentified choline lipids
AL, unidentified amino lipid; UL, unidentified lipids;
PL, unidentified Phospholipid; GL, unidentified Glycolipids;
PE, Phosphatidylethanolamine

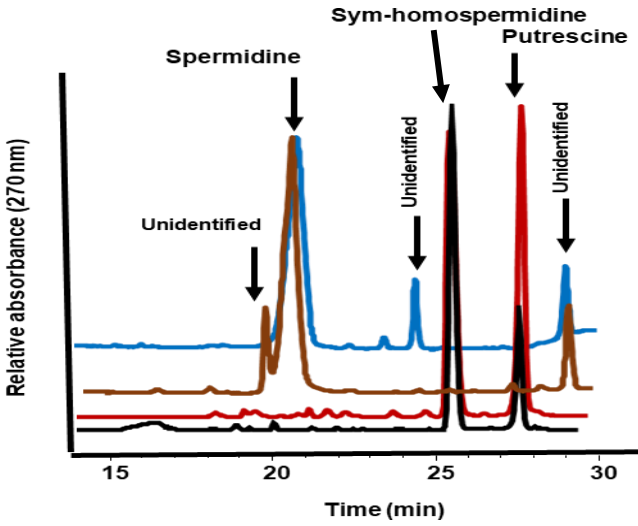


Fig. 34. HPLC chromatograms showing polyamines of strains JC665^T, JC747, *P. soli* JC670^T and *P.borealis* PX4^T

Table. 20. Differences in the characteristics of strains JC665^T, JC747, *P. soli* JC670^T (Kaushik et al. 2020) and *P. borealis* PX4 (Kulichevskkaya et al. 2016)

Characteristics	***Strain JC665^T**	***JC747**	***P. soli* JC670^T**	***P. borealis* PX4^T**
Cell shape	Spherical to oval	Spherical to oval	Spherical to oval	Spherical
Cell size (LxW,in µm)	1.8-1.7 x 1.5-1.3	1.8-1.7 x 1.5-1.3	1.6-1.7 x 1.3-15	1.5 to 2.5
Arrangement of cell	Single or in tissue-like aggregates	Single or in tissue-like aggregates	Single or in tissue-like aggregates	Single, in pairs or short chains
pH range(optima)	6-9 (7.0)	6-8 (7.0)	7.0-8.0 (7.0)	3.8-8.0 (5.0-5.5)
NaCl range(%w/v) (Optimum)	0-2 (0)	0-2 (0)	0-2 (0)	0-3 (0)
Temperature range(optimum)	4-34 (26-28)	8-30 (26-28)	4-30 (22-25)	4-37 (15-25)
Nitrogen sources utilization				
L-Phenylalanine	+	+	+	-
L-Lysine	+	+	+	-
DL-Threonine	+	+	+	-
Glycine	+	+	+	-
L-Isoleucine	+	+	+	-
L-Glutamine	-	-	+	-
L-Proline	+	+	+	-
Carbon sources utilization				
Maltose	+	+	-	+
Succinate	-	+	-	+
Propionate	-	+	+	-
Activity of enzymes				
Alkaline phosphatase	-	+	+	-
β-Galactosidase	-	-	-	+
N-acetyl-β-glucosaminidase	-	+	-	+
Fatty acids compositions				

$C_{12:0}$ *anteiso*	+	-	-	-
$C_{13:0}$	+	-	-	-
$C_{16:0}$ N alcohol	+	-	-	-
$C_{16:1}\omega9c$	-bn	-	+	+
$C_{17:0}$	+	-	+	-
Long chain [a]hydroxy fatty acid	-	-	-	+
Major Polar lipids				
Phosphatidylcholine	+	+	+	+
Phosphatidylethanolamine	-	-	-	+
Major Polyamines				
Spermidine	-	+	-	+
Sym homospermidine	+	-	+	-
Putrescine	+	-	+	-
Genomic features				
G+C content(mol%)	66.4	66.4	70.4	66.3
Genome size (Mb)	8.05	8.04	7.97	7.65
Coding sequences	6351	6340	6392	5785
RNAs	80	80	61	85
CRISPRs	1	1	0	1

*Data presented were performed in the author's laboratory

3.1.2.3.2. Characterization of strain JC656^T

3.1.2.3.2.1. Habitat and Nucleotide sequence accession numbers

Strain JC656^T was isolated from the rhizosphere soil of *Alisma plantago-aquatica*, an indigenous plant found in Loktak lake, Manipur, India (24°30'21" N 93°47'43" E). The GenBank accession number for the 16S rRNA gene sequence of strain JC656^T is ON908988 and the genome sequence has been deposited in GenBank under the accession number JAERRC010000000.

3.1.2.3.2.2. BLAST analysis and phylogenetic inference

The EzBioCloud analysis of the 16S rRNA gene sequence (1344 nt) of strain JC656^T extracted from the draft genome has the highest identity with the members of the genus *Sinomonas*; *S. notoginsengisoli* (98.58 %) being the closest. The OrthoANI and *d*DDH values of strain JC656^T with members of the genus *Sinomanas* fall below the recommended cut-off values i.e. 95-96 % and 70 % for prokaryotic species delineation (Rodriguez and Konstantinidis 2014; Meier-Kolthoff et al. 2014; Chun et al. 2018) (Fig. 35). The 16S rRNA gene based phylogenetic neighbour joining tree with combined bootstrap values obtained from NJ, ME, ML trees (Fig. 36) and 92 core genes based phylogenomic tree (Fig. 37) confirmed the clustering of strain JC656^T with its closest related species of the genus *Sinomonas* suggesting that strain JC656^T belongs to a new species of the genus *Sinomonas*.

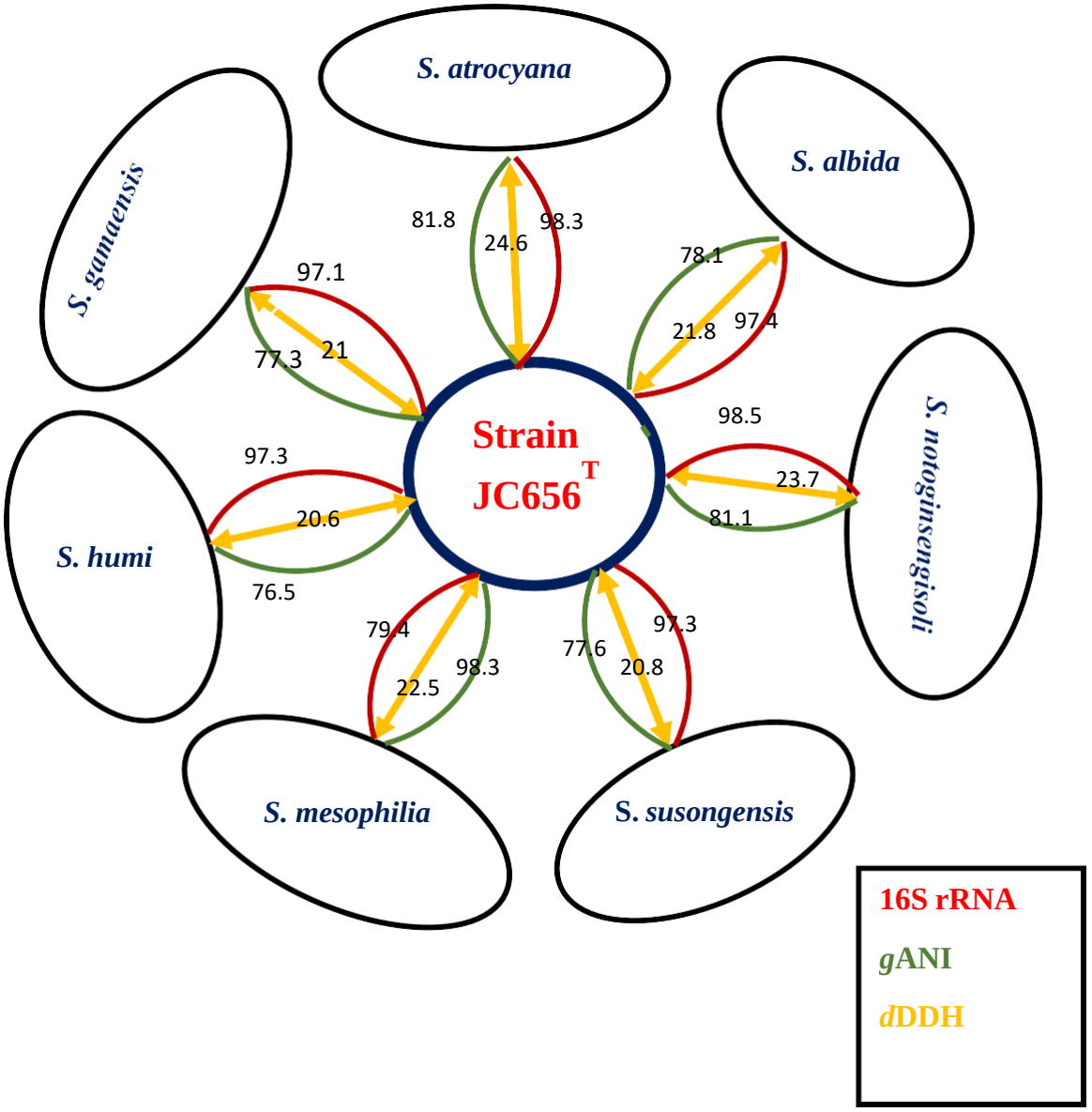


Fig. 35. Analysis of phylogenetic markers for delineation of the novel strain JC656[T] along with reference to the strains of the genus *Sinomonas*
Strains: 1. JC656[T] 2. *S. albida* LC13[T] 3. *S. notoginsengisoli* KCTC 29237[T] 4. *S. susongensis* A31[T] 5. *S. mesophilia* MPKL 26[T] 6. *S. humi* MUSC 117[T] 7. *S. gamaensis* NEAU-HV1[T] 8. *S. atrocyana* KCTC 3377[T]
Methods used: 16S rRNA gene identity (16S), average amino acid identity (ANI), digital DNA-DNA hybridization (*d*DDH)

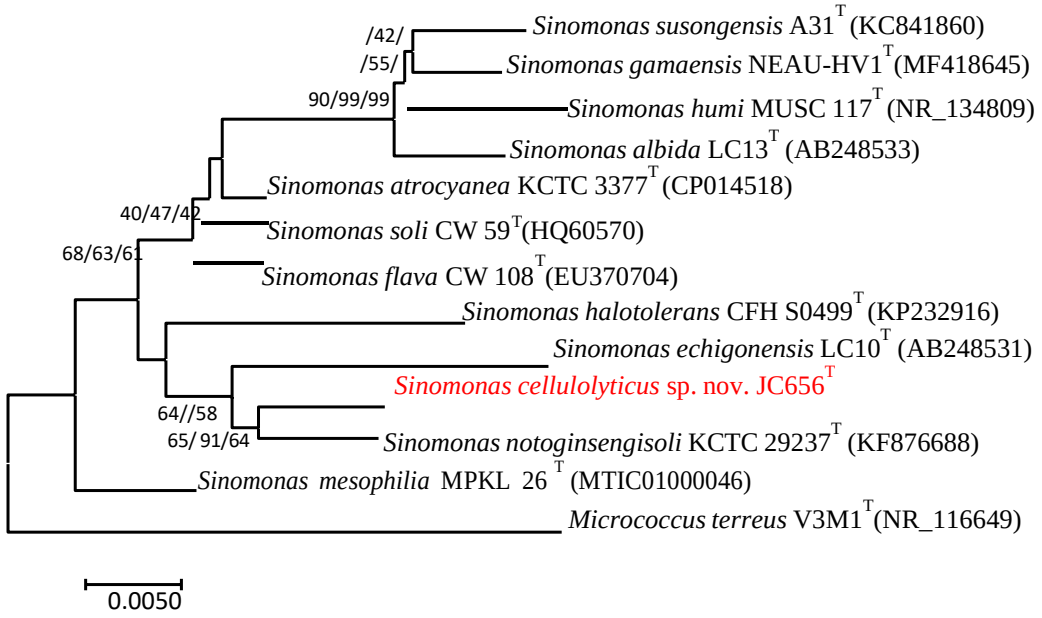


Fig. 36. Phylogenetic tree (NJ/ML/ME) based on 16S rRNA gene sequences showing the phylogenetic relationship of strain JC656T with their closely related species of the genus *Sinomonas.* Only values above 50 % are shown. Bar,0.005 substitutions per nucleotide position

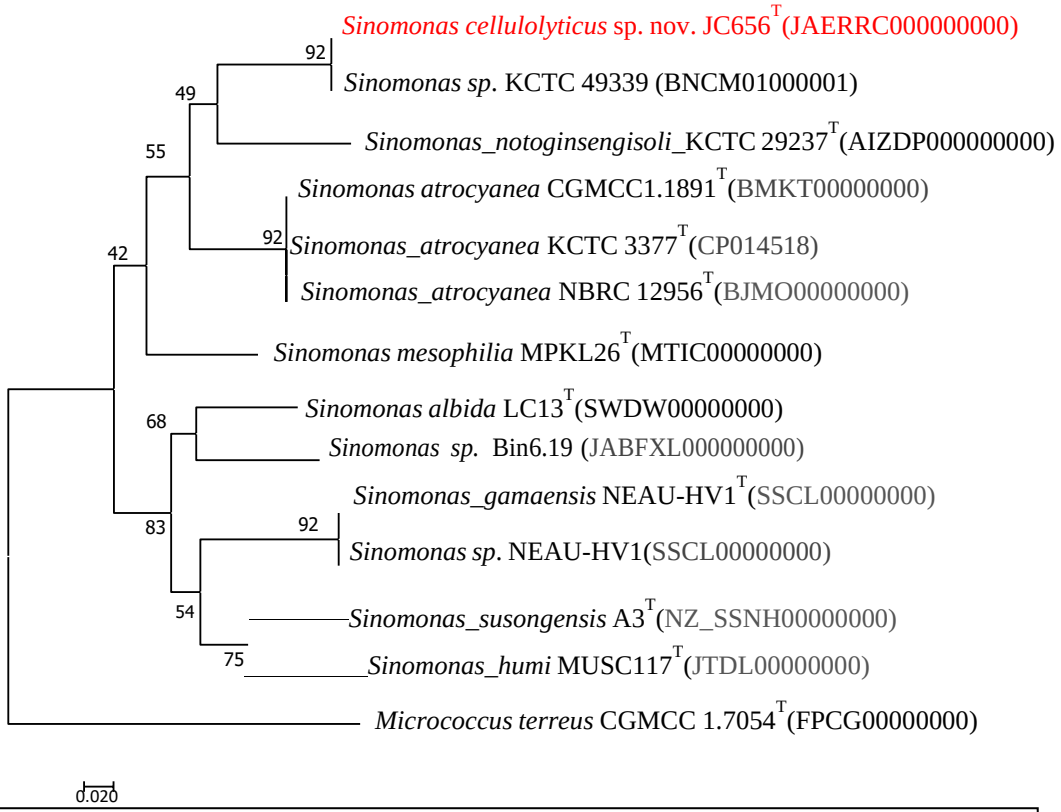


Fig. 37. RAxML Phylogenomic tree of strain JC656T along with publicly available genome sequences of the genus *Sinomonas.*The tree was constructed using 92 core genes tool based on the Up-to-date Bacterial Core Gene (UBCG) and rooted with *Micrococcus terreus* CGMCC1.7054T as the out-group. Bar. 0.02 substitutions per nucleotide

3.1.2.3.2.3. Genomic characteristics

The genome size of strain JC656^T was 3.9 Mb with N_{50} value 269315, while the genome of *S. notoginsengisoli* was 3.5 Mb with N_{50} value 261769. The genomic G + C content (mol %) of strain JC656^T and *S. notoginsengisoli* was 69.9 % and 68.8 %. The ANI and *d*DDH value between strain JC656^T and *S. notoginsengisoli* was 81.1 % and 23.7 % which were less than the recommended values for species delineation (Rosselló-Móra and Amann 2015; Richter and Rosselló-Móra 2009). The genome data analysis through RAST of strain JC656^T and other members of the genus *Sinomonas* are given in Table 21. The genome sequences of strain JC656^T and members of the genus *Sinomonas* were aligned using PATRIC software to identify the multiple maximal matches and local collinear blocks (LCBs; www.patricbrc.org). The genome sequence of strain JC656^T was used as a reference. The alignment of LCBs in the genus *Sinomonas* varied greatly with one another (Fig. 38). The genomic arrangement of strain JC656^T appeared to be rearranged and inverted when compared to the other genomes as a result of DNA rearrangements, recombination, and horizontal transfer (Fig. 39). The comparison of genome map based on protein sequence identity also revealed that majority of the proteins of strain JC656^T shared 20-80 % similarity with those of the clade members of *Sinomonas*. This indicated a clear-cut dissimilarity between strain JC656^T and other *Sinomonas* strains.

Table. 21. Genome comparison of strain JC656^{T} with related strains of the genus *Sinomonas*

Metric	**1***	**2***	**3**	**4**	**5**	**6**	**7**	**8**
Genome Size	3.9Mb	3.5Mb	3.7Mb	4.4Mb	4.3Mb	4.4Mb	4Mb	3.7Mb
mol.% G+C	69.9	68.8	69.5	71.4	67.6	67.2	71.1	68.6
No. of contigs	48	32	38	24	39	156	62	39
L50(bp)	6	3	7		6	13	6	5
N50(bp)	269315	261,796	200,527	206,486	233,829	111,730	111,730	199,011
rRNAs	4	3	9	3	4	5	5	6
tRNAs	55	53	53	50	54	54	52	53
Genes (coding)	3,351	3,368	3,401	4,023	4,162	4,011	3,536	3,465
Pseudo genes	32	59	48	83	70	109	79	49
Biosample accession no.	SAMN17245872	SAMN22059025	SAMN1137094	SAMD00097228	SAMN11370741	SAMN03070057	SAMN06212359	SAMN11370879
Genbank WGS[a] accession no.	NZ_JAERRC010000000.1	NZ_JAIZDP000000000.1	NZ_SWDW010000035.1	NZ_BJM000000000.1	NZ_SSCL01000000.1	NZ_JTDL00000000	NZ_MTIC00000000.1	NZ_SSNH00000000.1
GenBank assembly accession no.	GCF016735455.1	GCA_021554725.1	GCF004919045.1	GCF_006539205.1	GCA_05876955.1	GCF_000802235.	GCA_001999765.1	GCA_005877005.1

Strains: *1. JC556^{T}; 2. *S. notoginsengisoli* KCTC 29237^{T} (Zhang e al. 2014); 3. *S. albida* LC13^{T} (Ding et al. 2009); 4. *S. atrocyana* KCTC 3377^{T} (Ding et al. 2009) (Guo et al. 2015); 5. *S. flava* CW 108^{T} (Zhou et al. 2009); 6. *S. humi* MUSC 117^{T} (Lee et. al 2015); 7. *S. mesophilia* MPKL 26^{T} (Prabhu et al. 2015); 8. *S. susongensis* A31^{T} (Bao et al. 2015); 9. *Sinomonas gamaensis* NEAU-HVI^{T} (Fu et al. 2019); 10. *Sinomonas soli* CW 59^{T} (Zhou et al. 2012).

***Data presented were performed in the author's laboratory**

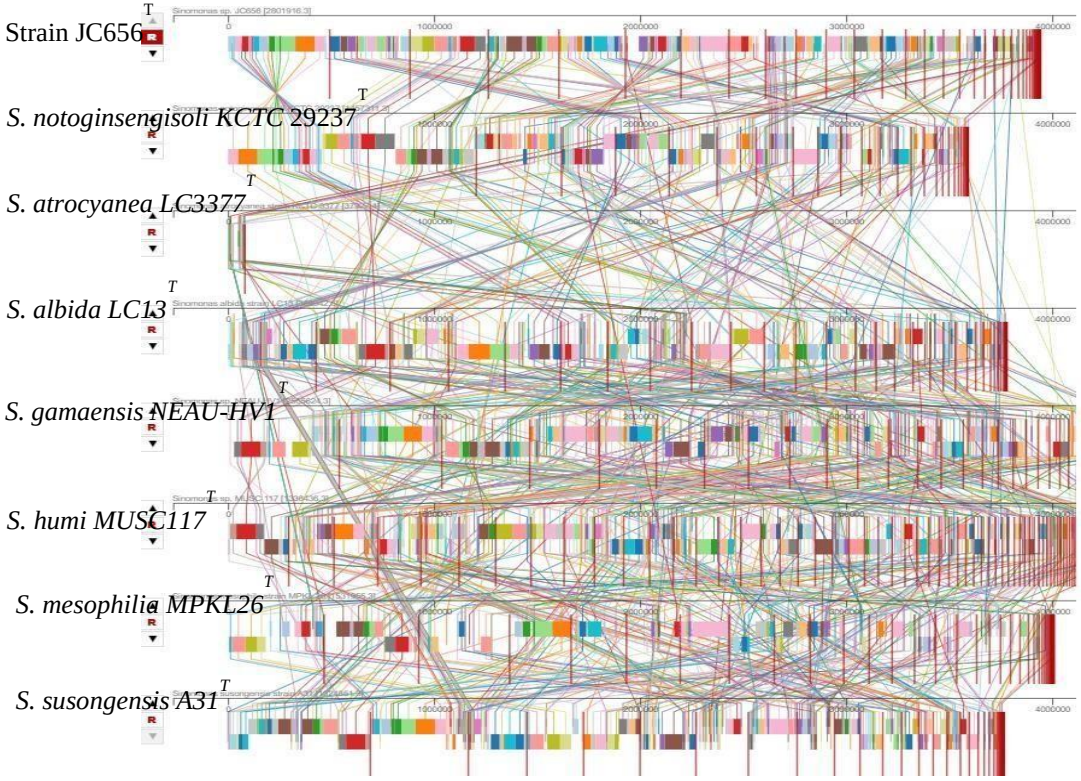


Fig. 38. Genome map based on protein sequence identity. The genome of strain JC656[T] was used as a reference. Lists of scale from outside to inside:1. Scale of contig length. Strains 2. JC656[T]; 3. *S. notoginsengisoli* KCTC 29237[T]; 4. *S. atrocyanea* LC3377[T]; 5. *S. albida* LC13[T]; 6. *S. gamaensis* NEAU-HV1[T] 7. *S. humi* MUSC117[T] 8. *S. mesophilia* MPKL 26[T]; 9. *S. susongensis* A31[T]

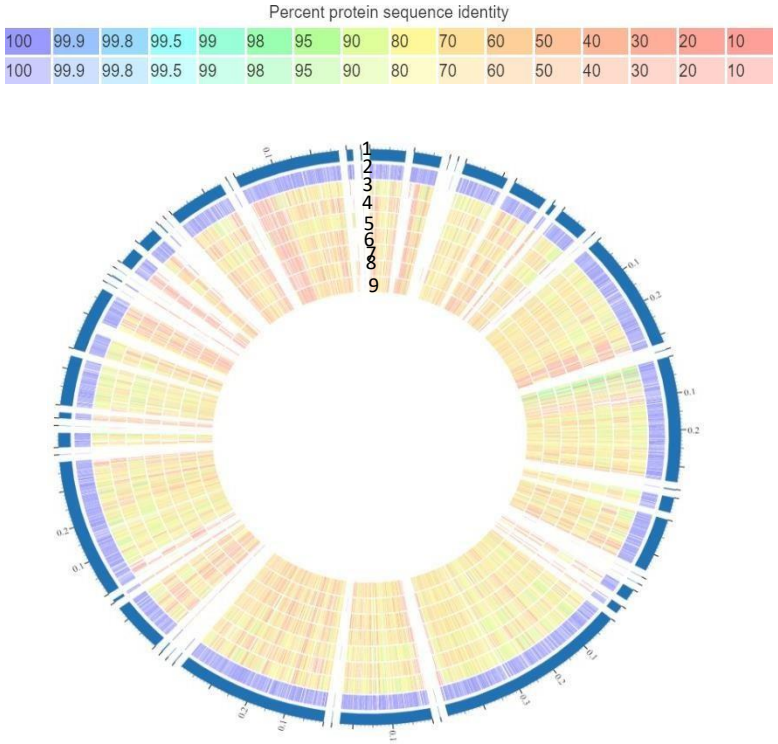


Fig. 39. Genome organization of strains:
1. JC656[T]; 2. *S. notoginsengisoli* KCTC 29237[T]; 3. *S. atrocyanea* LC3377[T]; 4. *S. albida* LC13[T]; 5. *S. gamaensis* NEAUHV1[T]; 6. *S. humi* MUSC117[T]; 7. *S. mesophilia* MPKL 26[T]; 8. *S. susongensis* A31[T]

3.1.2.3.2.4. Analysis of core and pan-genome

The distribution of genes for the genus *Sinomonas* as given by BPGA is depicted in table xx. Members of the genus *Sinomonas* share 1647 core genes (9.31 %), 10,785 accessory genes (60.9 %), and 5,258 strain-specific genes (29.7 %) (Fig. 40). The core-pan plot revealed that genus *Sinomonas* has an open pan-genome since it did not plateau and expanded with the increase in the number of genomes (Fig. 41).

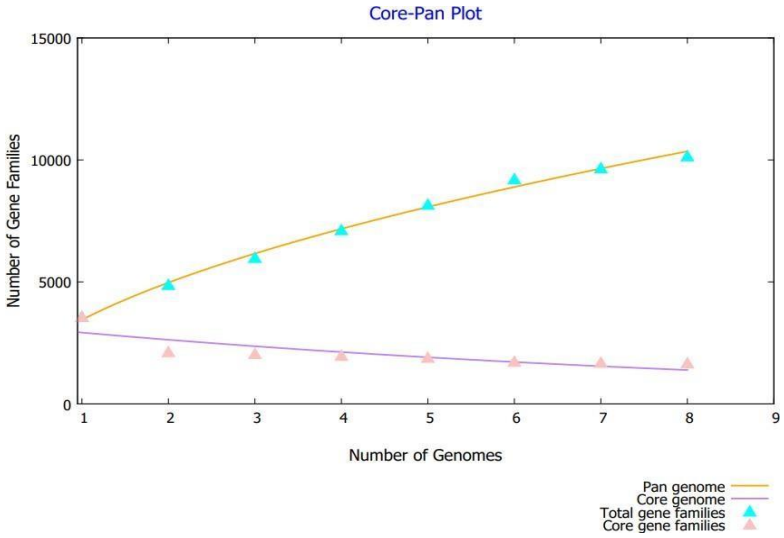


Fig. 40. Pan-genome analysis of strain JC656^T with closely related strains of *Sinomonas*. Inner grey circle represents the core genes, the outer blue represents the accessary genes and the buds represent the unique genes present in each species.

Fig. 41. Core-pan plot of the members of genus *Sinomonas* shows that it is an open pan genome

Strains: 1. JC656^T ; 2. *S. albida* LC13^T 3.*S. atrocyanea* KCTC 3377^T 4. *S. gamaensis* NEAU-HV1^T 5. *S. humi* MUSC 117^T 6. *S. mesophilia* MPKL 26^T 7. *S. notoginsengisoli* KCTC 29237^T 8. *S. susongensis* A31^T

3.12.3.1.5. *In-silico* metabolic characterization

CAZy annotation of genome of the strain JC656^T reveals number of genes encoding for carbohydrate-active enzymes (CAZymes). Strain JC656^T contain more genes encoding for glycoside hydrolases (GH), followed by glycosyl transferases (GT), and carbohydrate esterases (CE). Table xx also depicts the presence of enzymes like Endo β-1,4 gluconisidase (3.2.1.2.1) and β-glucosidase involved in cellulose degradation in the members of the genus *Sinomonas* (Fig. 42).

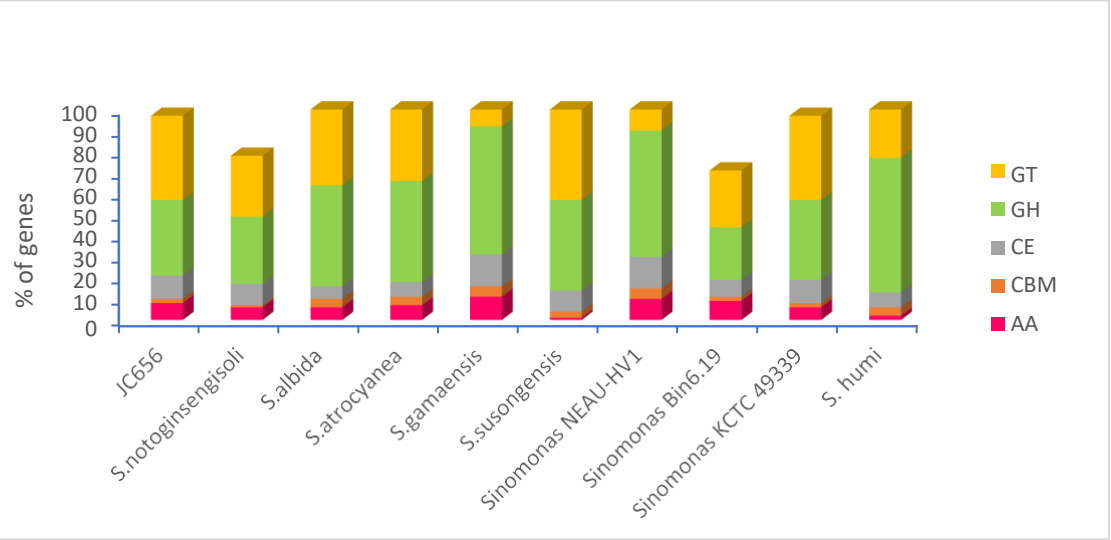


Fig. 42. Comparative analysis of CAZymes of strain JC656^T with the reference strains. Abbreviations: AA, Auxiliary Activities (AAs); CBM, Carbohydrate-Binding Modules (CBMs); Carbohydrate Esterases (CEs); Glycoside Hydrolases (GHs); GT, GlycosylTransferases (GTs)

3.1.2.3.1.5. Morphological, physiological and biochemical analysis

Strain JC656^T was found to be gram-positive, aerobic, non-motile and rod shaped. Colonies were observed to be round, yellow, convex in nutrient broth medium. Both strains JC656^T and *S. notoginsengisoli* SYP-B575^T could grow in nutrient broth at 28 ˚C for 3 days. Strain JC656^T grows well in the temperature range of 14-37 ˚C (optimum 28 ˚C), whereas *S. notoginsengisoli* SYP-B575^T could grow well in temperature range of 17-41

˚C (optimum 30 ˚C). Both the strains can grow well in pH range of 6.0-8.0 (optimum pH 7). Strain JC656^T does not require NaCl for its growth, however it can tolerate up to 0-6.0 % (w/v) whereas *S. notoginsengisoli* can tolerate up to 0-7.0 % (w/v). Both strains utilize the following carbon sources for their growth: glucose, D-mannitol, D-lactose, D-raffinose, D-maltose, D-galactose and citric acid. Mannose, D-sorbitol, D-fructose and D-cellubiose were solely utilized by *S. notoginsengisoli* SYP-B575^T (Table 22) (Fig. 43).

Strain JC656^T showed positive for easterase (C4), easterase lipase (C8) and α-monosidase while *S. notoginsengisoli* showed positive for easterase (C4) and easterase lipase (C8). Both strains were negative for lipase, valine arylamidase, cysteine arylamidase, α-galactosidase and *N*-acetyl-β-glucosaminidase (Table 22). Antibiotic susceptibility also revealed that strain JC656^T was resistance to ampicillin, vancomycin and amoxiclav whereas *S. notoginsengisoli* SYP-B575^T was resistant to ampicillin and methicillin. The growth and physiological characteristics differences of strain JC656^T and its closest relatives with different organic substrates are given in the Table 22.

3.1.2.3.1.6. Chemotaxonomic characterization

3.1.2.3.1.6.1. Cellular fatty acid, polar lipid, quinone and cell wall amino acid analysis

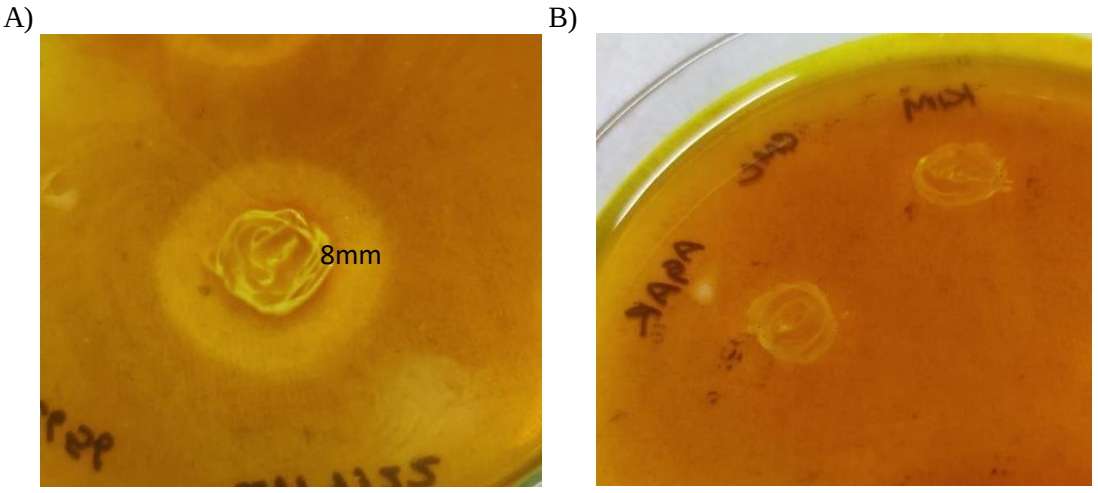
The major fatty acids identified in strain JC656^T were anteiso-$C_{15:0}$, iso-$C_{16:0}$, anteiso-$C_{17:0}$. The differences in fatty acids profile of strain JC656^T along with the type strains of the genus *Sinomonas* are given in Table 23. The polar lipids of strain JC656^T contains diphosphatidylglycerol, phosphatidylinositol, phosphatidylglycerol, phosphatidylmonomethylethanolamine, glycolipid (GL1-GL5) and unidentified lipid (Fig. 44). Both strains have Menaquinone-9 (H2) as the predominant quinone. Lysine, alanine, glutamine, DAP, and two unidentified amino acids were detected as the diagnostic cell wall amino acids of strain JC656^T and *S. notoginsengisoli* (Fig. 45) (Table 23)

Table. 22. Physiological characters that distinguish strain JC656^{T} from its closely related species of the genus *Sinomonas*

Characteristics	***1**	***2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Morphological characters											
Colony shape	round, convex	round	round, convex	circular, convex	circular, convex	circular, convex	circular, convex	circular, convex	Circular, convex	Circular, convex	Circular, convex
Colony colour	yellow	pale yellow	whitish	grey-white	plae yellow	yellowish to white	creamy white	pale yellow	yellow	creamy white	pale yellow
Physiological characters											
Growth media	Biebl and Pfennig	YDC	Luria broth	TYB	TYB	TSB	TYB	PYES	Luria broth	Luria broth	Nutrient broth
pH range (optima)	6-8 (7)	6-8(7)	6-10(7)	4-8(ND)	5-9(6-8)	5-8(6-7)	6-8 (7)	5-9(8)	5-9(6-8)	5-9(7)	5-10(8)
Temp range© (optima)	14-37(28)	17-41(28)	10-40(30)	14-43(ND)	15-42(30-37)	24-40(28-36)	2-40(30)	15-37(30)	15-42(30-37)	10-37(28)	10-45(30)
NaCl Tolerance (w/v) %	0-6	0-5	3-7	ND	0-3	0-4	0-4	0-5	0-3	0-7	0-4
Hydrolysis of											
Casein	-	-	ND	ND	-	ND	ND	ND	-	-	-
Starch	-	-	ND	ND	ND	+	-	-	-	-	-
Tween 80	-	+	+	-	+		-	-	-	-	-
Gelatin	-	-	-	-	-	-	-	-	-	+	-
Biochemical characters											
Catalase	+	+	-	+	+	+	+	+	+	-	
Urease	-	-	-	-	-	-	ND	+	-	-	+
Voges Proskauer	+	+	ND	+	+	ND	+	-	+	-	+
Nitrate reduction	-	-	-	+	-	+	-	-	+	-	+
Carbon sources utilization											
D-Mannitol	+	+	w	+	+		+	+	+	+	-
D-lactose	+	w	+	w	w	-	-	-	-	+	-
D-raffinose	+	w	+	w	w	+	+	-	-	-	+

D-sorbitol	-	+	w	+	w	-	-	+	+	-	-
D-maltose	+	+	+	+	+	+	+	+	-	+	-
D-fructose	-	+	+	+	+	+	+	+	+	-	+
D-cellubiose	-	+	+	+	+	-	ND	+	+	+	-
D-galactose	+	+	+	w	+	+	+	-	w	W	+
Inositol	+	+	+	+	+		+	-	+	-	-
Citric acid	+	w	w	+	w	ND	ND	ND	+	-	+
Mannose	-	w	+	+	+	-	+	-	+	W	-
Glucose	+	+	+	+	-	+	+	+	+	+	-
Enzyme activities (API ZYM)											
Easterase (C4)	+	+	+	w	+	+	+	-	+	+	+
Easterase lipase (C8)	+	+	-	w	+	+	+	-	+	+	+
Lipase	-	-	-	-	-	+	-	-	-	-	+
Valine-arylamidase	-	-	-	-	+	+	-	-	+	-	+
Cysteine-arylamidase	-	-	-	-	-	+	+	-	-	-	-
α-Galactosidase	-	-	-	-	-	+	-	+	-	-	+
N-acetyl-β-glucosaminidase	-	-	+	-	-	+	-	-	-	-	+
α-Mannosidase	+	-	-	-	-	+	-	+	w	-	+
Antibiotic susceptibility											
Ampicillin	+	+	ND	ND	ND	ND	ND	ND	ND	ND	ND
Methicillin	-	+	ND	ND	ND	ND	ND	ND	ND	ND	ND
Amoxyclav	+	-	ND	ND	ND	ND	ND	ND	ND	ND	ND
Vancomycin	+	W	ND	W	W	ND	+	ND	W	ND	ND
Rifamycin	-	W	ND	W	W	ND	+	ND	W	ND	ND

Strains: 1*. JC556^T; 2. ****S. notoginsengisoli* **KCTC 29237^T (Zhang e al. 2014); 3. *S. albida* LC13^T (Ding et al. 2009); 4. *S. atrocyana* KCTC 3377^T (Ding et al. 2009) (Guo et al. 2015); 5. *Sinomonas soli* CW 59^T (Zhou et al. 2012); 6. *S. humi* MUSC 117^T (Lee et. al 2015); 7. *S. mesophilia* MPKL 26^T (Prabhu et al. 2015); 8. *S. susongensis* A31^T (Bao et al.2015); 9. *S. flava* CW 108^T (Zhou et al. 2009); 10. *S. halotoelrans* CFH S0499^T (Guo et al. 2015); 11. S. *gamaensis* NEAU-HVI^T (Fu et. al 2019). *Data presented were performed in the author's laboratory**



Strain JC656^{T} ***S. notoginsengisoli*** **KCTC 29237^{T}**

Fig. 43. A) Zone of clearance on carboxymethylcellulose (CMC) agar plates for strain JC656^{T} after 48 h of incubation. The formation of clearing zone around the colonies confirm the secretion of extracellular cellulase.
B) No formation of zone in *S. notoginsengisoli* KCTC 29237^{T}

Table. 23. Cellular fatty acid compositions (%) of strain JC656^{T} and its closely related species of the genus *Sinomonas*

	1	2	3	4	5	6	7	8	9	10
iso-$C_{14:0}$	4.11	0.5	1.2	1.2	0.9	-	-	-	-	-
$C_{14:0}$	-	-	-	0.8	0.6	-	-	-	-	-
iso-$C_{15:0}$	5.04	8.6	6.8	18.8	20.7	8.2	10.6	8.8	14.5	7.9
anteiso-$C_{15:0}$	45.37	51.6	60.4	51.6	51.3	56.8	57.6	41.4	52.6	44.6
iso-$C_{16:0}$	24.69	3.3	6.8	5.3	4.2	3.8	4.8	12.7	3.3	-
$C_{16:0}$	1.60	2.0	1.3	2.5	2.1	2.1	1.5	1.8	-	20.5
iso-$C_{17:0}$	-	2.3	1.0	1.9	2.1	1.6	1.3	-	3.3	-
anteiso-$C_{17:0}$	16.14	27.0	20.5	15.9	16.6	25.8	21.3	16.2	24.6	26.9
$C_{18:1}\omega 7c1$	-	-	-	-	-	-	-	16.2	-	-

Strains: 1. JC656^{T}; 2. *S. notoginsengisoli* KCTC 29237^{T} (Zhang e al. 2014); 3. *S. atrocyanea* DSM 20127^{T} (Ding et al. 2009) (Guo et al. 2015); 4. *S. flava* CW 108^{T} (Zhou et al. 2009); 5. *S. soli* CW 59^{T} (Zhou et al. 2012); 6. *S. susongensis* A31^{T} Bao et al. 2015); 7. *S. albida* LC13^{T} (Ding et al. 2009); 8. *S. humi* 117^{T} (Lee et. al 2015); 9. *S. mesophila* MPKL26^{T} (Prabhu et al. 2014); 10. *S. gamaensis* NEAU-HVI^{T} (Fu et. al 2019. Not detected or <1% of fatty acid content

Table.24. Chemotaxonomic and molecular profiling of strain JC656^T with closely related species of the genus *Sinomonas*

Characteristics	1*	2	3	4	5	6	7	8	9	10
Cell wall amino acids	Lysine, Alanine, Glutamine, DAP, UN	Lysine, Alanine, Glutamine, Glycine, Serine	Lysine, Alanine, Serine	Lysine, Alanine, Glutamine, Glycine, DAP	Lysine, Alanine, Glutamine, Glycine	Lysine, Serine, Glycine, Alanine, Glutamine, Muramic	Lysine, Alanine, Glycine, Glutamine	Lysine, Alanine, Glutamine, Glycine, Serine	ND	Lysine, Alanine, Glutamine
Polar lipids	DPG, PG, GL, PME, UN	DPG, PG, PI, PL, GL, UN	DPG, PG, PI, PL	DPG, PG, PI, PL, PME	DPG, PG, P1, PME	DPG, PG, PI, UN	DPG, PG, PI, UN	DPG, PG, PI, PME, GL	DPG, PG, PI, GI	DPG, PG, PI, GL, PME
Fatty acids	anteiso-$C_{15:0}$, iso-$C_{16:0}$, anteiso-$C_{17:0}$	anteiso-$C_{15:0,}$ anteiso-$C_{17:0,}$	iso-$C_{15:0}$, aanteiso-$C_{15:0}$, anteiso-$C_{17:0}$	anteiso-$C_{15:0}$, anteiso-$C_{17:0}$, iso-$C_{15:0}$	iso-$C_{15:0}$, anteiso-$C_{15:0}$, anteiso-$C_{17:0}$	anteiso-$C_{15:0}$, anteiso-$C_{16:0}$, anteiso-$C_{17:0}$	iso-$C_{15:0,}$ anteiso-$C_{15:0}$, anteiso-$C_{17:0}$	anteiso-$C_{15:0,}$ anteiso-$C_{17:0,}$	anteiso-$C_{15:0,}$ anteiso-$C_{17:0,}$ $C_{16:0,}$ iso-$C_{15:0}$	iso-$C_{15:0}$, anteiso-$C_{15:0}$, anteiso-$C_{17:0}$,
Respiratory quinones	MK-9, MK-8	MK-9, MK-8, MK-6	MK-9 MK-10	MK-9, MK-8	MK9, MK8, MK-6	MK-9, MK8, MK10	MK-9, MK-10, MK-8	MK-9, MK-10, MK-8	Mk-9, MK-10, MK-8	MK-9

Strains: 1*. JC556^T; 2. *S. notoginsengisoli* KCTC 29237^T (Zhang e al. 2014); 3. *S. albida* LC13^T (Ding et al. 2009); 4. *S. atrocyana* KCTC 3377^T (Ding et al. 2009) (Guo et al. 2015); 5. *S. flava* CW 108^T (Zhou et al. 2009); 6. *S. humi* MUSC 117^T (Lee et. al 2015); 7. *S. mesophilia* MPKL 26 ^{T} (Prabhu et al. 2014); 8. *S. susongensis* A31^T (Bao et al. 2015); 9. *S. gamaensis* NEAU-HVI^T (Fu et. al 2019); 10. *S. soli* CW 59^T (Zhou et al. 2012).

***Data presented were performed in the author's laboratory;**

DPG: Diphosphatidylglycerol;
PI: Phosphatidylinositol
PG: Phosphatidylglycerol
PME: Phosphatidylmonomethylethanolamine
GL: Glycolipid
UN: Unidentified
DAP: Diaminopimelic acid

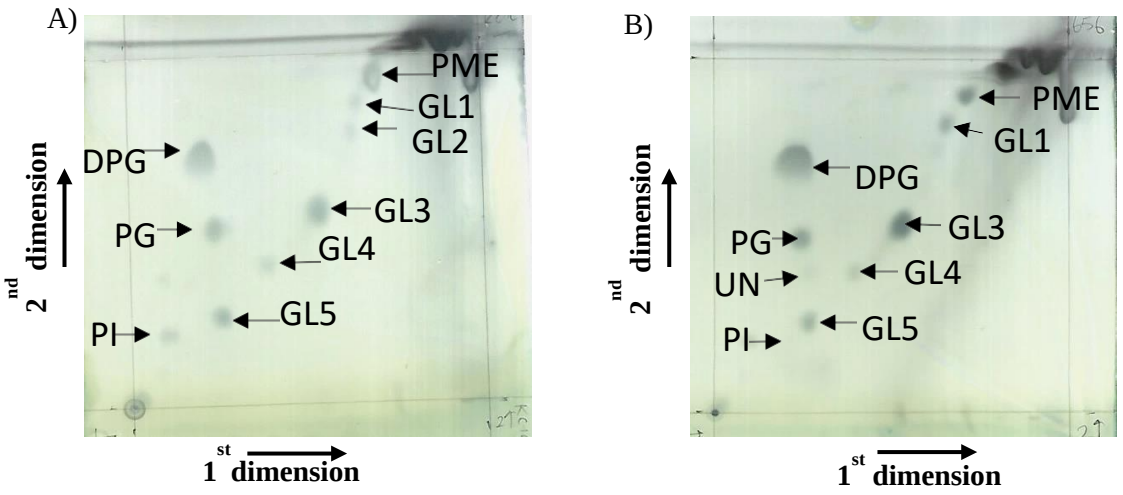


Fig. 44. Two-dimensional thin-layer chromatogram showing polar lipids of strains
A) *S. notoginsengisoli* KCTC 29237^T
B) Strain JC656^T
DPG: Diphosphatidylglycerol
PI: Phosphatidylinositol
PG: Phosphatidylglycerol
PME: Phosphatidylmonomethylethanolamine
GL: Glycolipid
UN: Unidentified

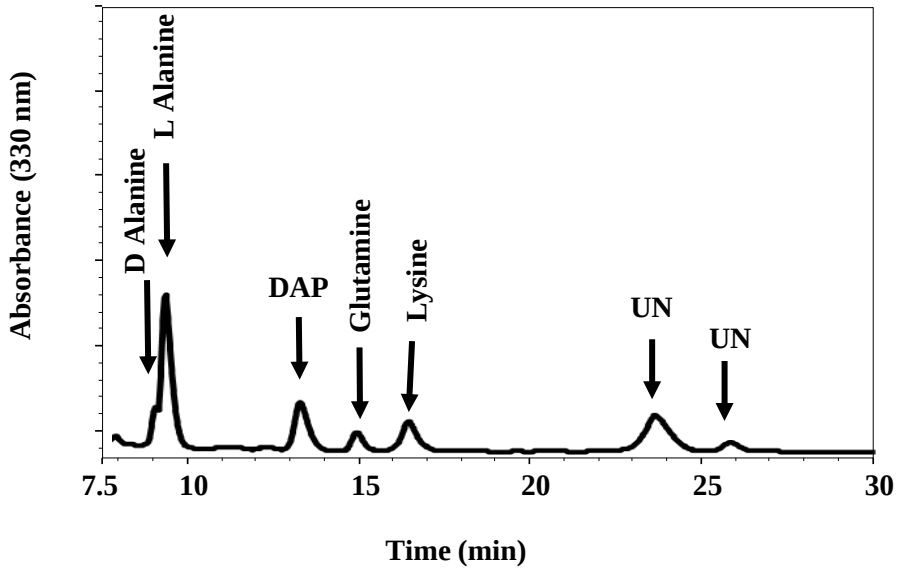


Fig. 45. HPLC chromatogram showing cell wall amino acid peaks of strain JC656^T
(Strain JC656^T contains alanine, lysine, glutamine, diaminopimelic acid (DAP) and two unidentified amino acids)

3.1.2.3.3. Characterization of strain JC586^T

3.1.2.3.3.1. Habitat and nucleotide accession numbers

Strain JC586^T was isolated from the rhizosphere soil of *Alocasia cucullata* collected from floating islands or Phumdis (composed of heterogeneous masses of vegetation, soil and organic matter formed at different stages of decomposition) of Loktak Lake, India (24°30'21" N/93°47'43" E). The yellow color colonies were purified by repeated streaking on nutrient agar (NA) and preserved in 50 % glycerol stock at -20 ˚C. The GenBank accession number for the 16S rRNA gene sequence of strain JC586^T is ON908987, and the whole genome sequence has been deposited at GenBank under the accession number NZVHIN00000000.

3.1.2.3.3.2. BLAST analysis and phylogenetic inference

The EZBioCloud BLAST analysis of the 16S rRNA gene sequence of strain JC586^T shows that it shares the highest sequence similarity with the members of the genus Glutamicibacter; *Glutamicibacter halophytocola* KLBMP^T (99.3 %) and with *Glutamicibacter mishrai* S5-52^T (99.2 %) being the closest members. The OrthoANI and *d*DDH values of strain JC586^T with members of the genus *Glutamicibacter* fall below the recommended cut-off values i.e. 95-96 % and 70 % for prokaryotic species delineation (Goris et al. 2007; Richter and Rosello-Mora 2009) (Fig. 45). The 16S rRNA gene based phylogenetic neighbour joining tree with combined bootstrap values obtained from NJ, ME, ML trees formed a separate cluster with its closely related species of the genus *Glutamicibacter* (Fig. 46) This clustering was also furthered confirmed by the 92 core genes RAxML-based UBCG phylogenomic tree suggesting that strain JC586^T is a new species of the genus *Glutamicibacter* (Fig. 47).

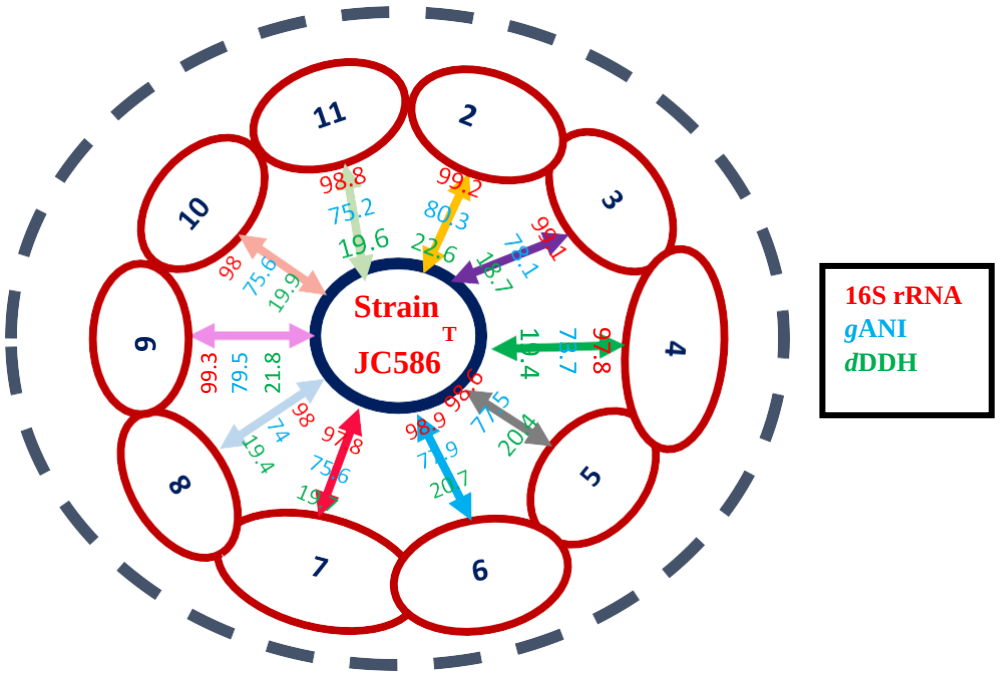


Fig. 46. Analysis of phylogenetic markers for delineation of novel isolate strain JC586[T] along with reference strains of *Glutamicibacter*. Strains **1**. JC586[T]; **2.** *G. mishrai* LMG 29155[T]; **3.** *G. arilaitensis* DSM 16368[T]; **4** *G. creatinolyticus* DSM 15881[T] **5**; *G. mysorens* DSM 12798[T]; **6**. *G. nicotianae* DSM 20123[T]; **7**.*G. protophormiae* DSM 20168[T]; **8**. *G. ardleyensis* DSM 17432[T]; **9**. *G. halophytocola* DSM 101718[T]; **10**. *G. soli* DSM 19449[T]; **11**.*G. uratoxydans* DSM 20647[T]. Methods used: 16S rRNA gene identity (16S), average amino acid identity (ANI), digital DNA-DNA hybridization (*d*DDH)

3.1.2.3.3.3. Genomic characterization

The genome of strain JC586[T] consists of one circular contig of 3,524,842 bp, 28 contigs, N_{50} value of 558,740 bp, with 100 x coverage and a GC content of 60.04 %. Genes were predicted using the NCBI Prokaryotic Genome Annotation Pipeline version 5.0 (PGAP). There were 3,294 predicted genes; 3,177 coding genes; 75 RNAs, including 62 tRNAs, 10 rRNAs, and 3 non-coding RNAs (ncRNAs). The ANI and *d*DDH value between strain JC586[T] and *G. mishrai* was 80.3 % and 22.6 %, which were less than the recommended values (95-96 % cut off for ANI and 70 % cut off for *d*DDH) for species delineation (Rosselló-Móra and Amann 2015; Richter and Rosselló-Móra 2009). Table 24 gives the summary of the genomic information data of strain JC586[T] and other genus *Glutamicibacter* members.

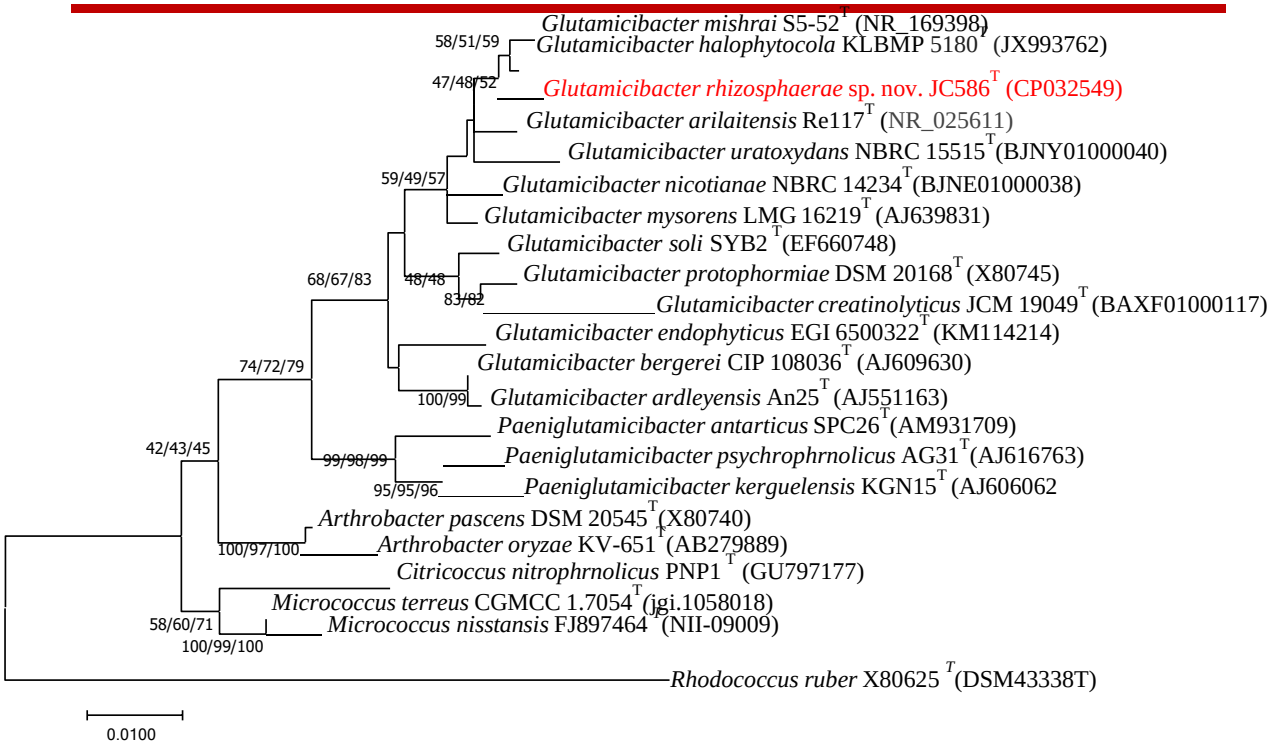


Fig. 47. Phylogenetic tree (NJ/ML/ME) showing the position of strain JC586^T among the related taxa based on 16S rRNA gene sequences. Bootstrap values expressed as percentages of 1000 replications are given at branch points. Only values above 40% are shown. Bar, 0.01 substitution per 1000 nucleotide position. The tree was constructed using MEGA 7 software and *Rhodococcus ruber* X80625^T was used as outgroup

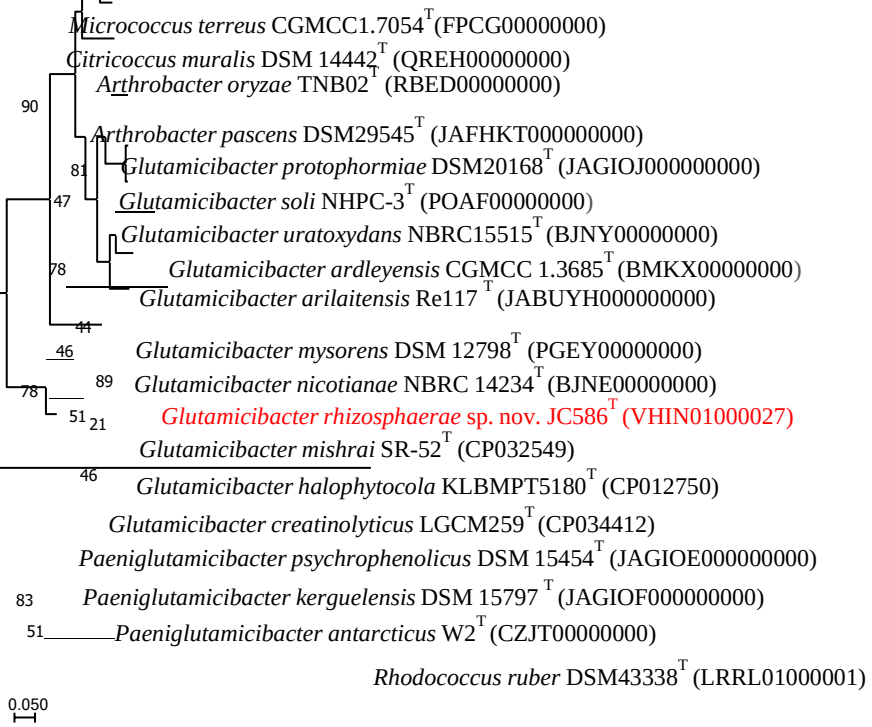


Fig. 48. Phylogenomic tree of members of the genus *Glutamicibacter* constructed using 92 core genes tool based on the Up-to-date Bacterial Core Gene (UBCG). The tree was generated using the MEGAX software (NJ) with *Rhodococcus ruber* DSM43338^T as an outgroup

Table. 25. Differential genomic characteristics between strain JC586^{T} and the type strains of related *Glutamicibacter* species

Characteristics	1	2*	3*	4*	5*	6*	7*	8 *	9*	10 *	11*
Biosample id	SAMN038 4076	SAMN123 68439	SAMN1005 3380	SAMN1005 2990	SAMEA227 2213	SAMB0004 6490	SAMN0832 5510	SAMN1767 3561	SAMD0024 512	SAMD0009 7580	SAMN10502625
Accession no.	VHIN000 00000	CP042260	CP012750	CP033081	NC_014550	BCQO0000 0000	POAF00000 000	AGIOJ0000 00000	BMKX0000 0000	BJNY00000 000	CP034412
Genome coverage	100x	100x	700x	390x	ND	139x	100x	253x	130x	152x	286x
No. of contigs	28	2	2	3	3	1	25	2	34	49	1
Genome size (Mb)	3.52	3.57	3.91	3.6	3.8	3.4	3.8	3.8	3.9	3.7	3.3
GC%	60.4	59.4	60	61.9	59.2	62	64.8	64	56.5	61.1	65.6
N50(bp)	558,740	3,570,647	3,911,798	3,643,989	2,610,692	489,832	563,821	3,825,410	271,783	175,646	3,309
CDS (total)	3219	3346	3246	3482	3593	3167	3536	3552	3663	3539	2979
CDS (protein)	3177	3274	3215	3428	3490	3128	3481	3427	3614	3501	2915
rRNA	10	19	19	19	19	7	7	19	3	5	12
tRNA	62	64	64	67	64	63	67	66	61	57	58
Pseudogene	42	31	18	54	103	39	55	5	49	38	64
CRISPRs	2	0	1	0	0	0	1	1	1	1	0

***Data taken from literature; Strains:1. JC586^{T} 2. *G. mishrai* LMG 29155^{T} (Das et al. 2019); 3. *G. halophytocola* DSM 101718^{T} (Feng et al. 2017); 4 *G. nicotianae* DSM 20123^{T} (Das et al. 2019); 5. *G. arilaitensis* DSM 16368^{T} (Das et al. 2019); 6 *G. mysorens* DSM 12798^{T} (Das et al. 2019); 7 *G. soli* DSM 19449^{T} (Busse et al. 2016); 8. *G. protophormiae* DSM 20168^{T} (Busse 2016); 9. *G. ardleyensis* DSM 17432^{T} (Busse 2016); 10. *G. uratoxydans* DSM 20647^{T} (Busse 2016); 11. *G. creatinolyticus* DSM 15881^{T} (Hou et al. 1998).**

3.1.2.3.3.4. Analysis of core and pan-genome

The distribution of genes for the genus *Glutamicibacter* as given by BPGA, is depicted in table xx. Members of the genus *Glutamicibacter* share 1769 core genes (8.43 %), 13,105 accessory genes (64.7 %), and 5,378 strain-specific genes (26.5 %) (Fig. 48). The core-pan plot revealed that the genus *Glutamicibacter* has an open pan-genome since it did not plateau, and the gene pool size expanded with the increase in the number of genomes incorporated in the analysis (Fig. 49).

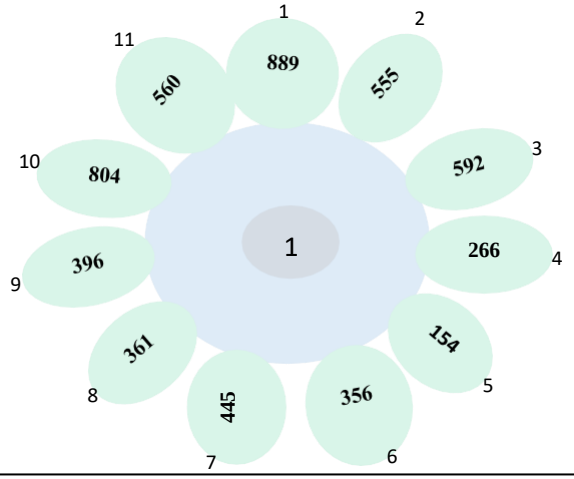


Fig. 49. Pan-genome analysis of strain JC586[T] with closely related strains of *Glutamicibacter*. The inner grey circle represents the core genes, the outer blue represents the accessary genes and the buds represent the unique genes present in each species

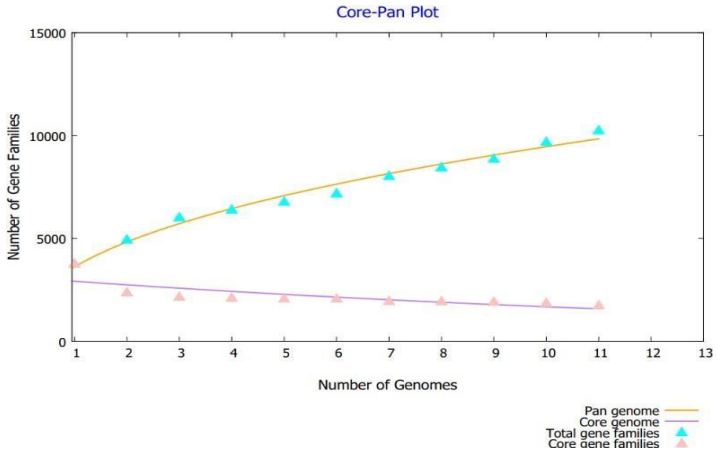


Fig. 50. Core-pan plot of the members of genus *Glutamicibacter* shows the pan-genome is open

Strains: 1. *G. ardleyensis* DSM 16368[T] 2. *G. arilaitensis* DSM 16368[T] 3. *G. halophytocola* DSM101718[T] 4. *G. mishrai* LMG 29155[T] 5. *G. mysorens* DSM 12798[T] 6. *G. nicotianae* DSM 20123[T] 7. *G. protomotphiae* DSM 20168[T] 8. *G. soli* DSM 19449[T] 9. JC586[T] 10. *G. uratoxydans* DSM 20647[T] 11. *G. creatinolyticus* DSM15881[T]

3.1.2.3.3.5. Genomic properties linked with environmental stress response

In the genome of strain JC586^T, several genes related to ecological adaptation were identified, such as genes responsible for oxidative stress response (*sodA, soxR, bcp, katE, katG, osmC, trx, ohr, trxA, trxB*), osmo-protectant (trehalose metabolism, glycogen metabolism, and glycine betaine/proline ABC transporter) and cold shock response (*nusA, csd, pnp, infA, infB, rbfA*). Reactive oxygen species (ROS) are produced in higher abundance due to the increased solubility of oxygen at low temperatures (Dsouza et al. 2015). Moreover, several studies have shown that ROS cause oxidative damage in bacteria. Thus, the presence of genes like *sodA*, *katE*, etc. in the genome of strain JC586^T protects from free radical damage and enables ROS detoxification. It has also been reported that trehalose and glycogen metabolism serve as a carbon source under nutrient-restricted conditions and protect cells from cold shock, osmotic stress and desiccation (Zevenhuizen, 1992). Studies also show that the presence of cold shock proteins like *nusA, csd, infA, infB* might explain the survival of bacteria in low temperatures (Melissa et al. 2015). Therefore, based on the genomic data, it can be concluded that genes involved in environmental stress-related adaptation may be essential for the growth and survival of strain JC586^T under extreme conditions. Comparative list of genes linked to environmental stress response in the genome of *Glutamicibacter* sp. JC586^T are given in Table 26.

Table. 26. List of genes linked to environmental stress response in the genome of strain

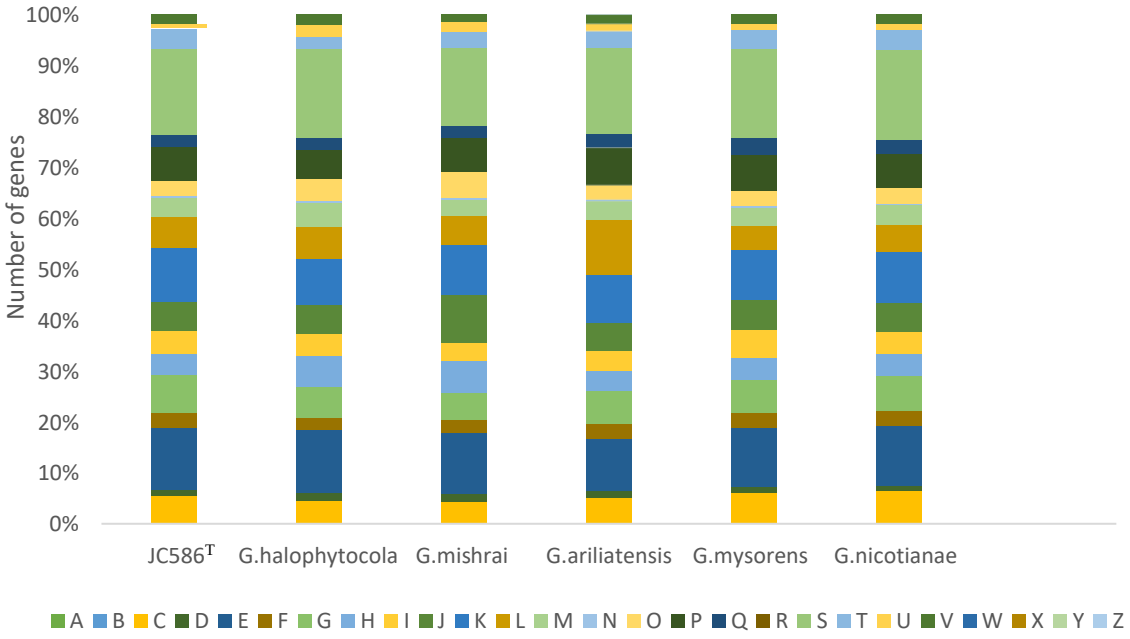
Gene name	Gene symbol	Cog ids	1	2*	3*	4*	5*	6*
I. **Oxidative stress response**								
a. Superoxide dismutase	*sodA*	COG0605	+	+	+	+	+	+
b. Catalase	*katE*	COG0753	-	+	+	+	+	+
c. Catalase (peroxidase I)	*katG*	COG0376	-	+	+	+	-	+
d. Peroxiredoxin	*bcp*	COG1225	-	+	+	+	+	+
e. Predicted redox protein, regulator of disulfide bond formation	*osmC, ohr*	COG1764	-	+	+	+	+	+
f. Thioredoxin domain-containing protein	*trx*	COG3118	-	+	+	+	+	+
g. Thioredoxin reductase	*trxB*	COG0492	+	+	+	+	+	+
h. Thiol-disulfide isomerase and thioredoxins	*trxA*	COG0526	-	+	+	+	+	+
i. Predicted transcriptional regulators	*soxR*	COG0789	+	+	+	+	+	+
II. **Osmoprotection**								
a. Glycogen metabolism								
i. Glycosyltransferase	*glgA*	COG0438	-	+	+	+	+	+
ii. 1,4-alpha-glucan branching enzyme	*glgB*	COG0296	+	+	+	-	-	-
iii. ADP-glucose pyro phosphorylase	*glgC*	COG0448	+	+	+	-	-	-
iv. Glucan phosphorylase	*glgP*	COG0058	+	+	-	+	-	-
v. Type II secretory pathway, pullulanase PulA and related glycosidases	*glgX*	COG1523	-	+	+	+	-	+
b. Trehalose metabolism								
i. Trehalose-6-phosphate synthase	*otsA*	COG0380	-	+	+	+	+	+
ii. Trehalose-6-phosphatase	*otsB*	COG1877	+	+	+	+	+	+
c. Glycine betaine/proline ABC transporter								
i. ABC-type proline/glycine betaine transport systems, ATPase components	*proV*	COG1125	-	+	+	-	+	+
ii. ABC-type proline/glycine betaine transport systems, permease component	*proW*	COG1174	-	+	+	-	+	+
iii. Periplasmic glycine betaine/choline-binding (lipo) protein of an ABC-type transport system (osmoprotectant binding protein)	*proX*	COG1732	-	+	+	-	+	+
III. **Cold shock response**								
a. Cold shock DNA binding domain	*csd*	COG1278	-	+	+	+	+	+
b. Transcription elongation factor	*nusA*	COG0195	+	+	+	+	+	+
c. Polyribonucleotide nucleotidylt ransferase (polynucleotide phosphorylase)	*pnp*	COG1185	-	+	+	+	+	+
d. Ribosome-binding factor A	*rbfA*	COG0858	+	+	+	+	+	+
e. Translation initiation factor 1 (IF-1)	*infA*	COG0361	+	+	+	+	+	+
f. Translation initiation factor 2 (IF-2; GTPase)	*infB*	COG0532	+	+	+	+	+	+

***Data taken from literature. 1. Strain JC586^T along with the reference genomes of 2. *G. mishrai* LMG 29155^T 3. *G. halophytocola* DSM 101718^T 4. *G. nicotianae* DSM 20123^T 5. *G. arilaitensis* DSM 16368 ^{T} 6. *G. mysorens* DSM 12798^T**

3.1.2.3.3.6. *In-silico* metabolic characterization

The COGs annotation was carried out to have a better understanding of the metabolic functions. The results revealed that strain JC586^{T} displayed similar results as that of the other members of the genus *Glutamicibacter*. The COG functional analysis exhibit that most of the genes were associated with unknown function, followed by amino acid metabolism, transcription and carbohydrate metabolism (Fig. 50).

Fig. 51. Functional classification COGs clusters of orthologs of strain JC586^{T} and related strains of the genus *Glutamicibacter*



***Data taken from literature. 1. Strain JC586^{T}; 2. *G. halophytocola* DSM 101718^{T}; 3. *G. mishrai* LMG 29155^{T*}; 4. *G. arilaitensis* DSM 16368^{T*}; 5. *G. mysorens* DSM 12798^{T*}; 6. *G. nicotianae* DSM 20123^{T*}**

A: RNA processing and modification; B: Chromatin Structure and dynamics; C: Energy production and conversion; D: Cell cycle control and mitosis; E: Amino Acid metabolism and transport; F: Nucleotide metabolism and transport; G: Carbohydrate metabolism and transport; H: Coenzyme metabolism; I: Lipid metabolism; J: Tranlsation; K: Transcription; L: Replication and repair; M: Cell wall/ membrane/ envelop biogenesis; N: Cell motility; O: Post-translational modification, protein turnover, chaperone functions; P: Inorganic ion transport and metabolism; Q: Secondary Structure; S: Function Unknown; T: Signal Transduction; U: Intracellular trafficing and secretion; V: Defense mechanisms; Z: Cytoskeleton.

3.1.2.3.3.7. Morphological, physiological and biochemical analysis

Strain JC586^T was found to be gram-positive, aerobic, non-motile, rod-coccus cycle. Colonies appeared smooth, light-yellow, round, non-sporulating, and convex in nutrient broth. Growth occurred at 16-37 ˚C with optimal growth conditions at 28 ˚C, pH range of 5-11 with an optimum pH range of 7, and NaCl range of 0-11 % with an optimum range required for growth was 2-4 %. Enzyme activity for the strain JC586^T was determine along with the other members of the genus *Glutamicibacter* according to the API ZYM kit (BioMerieux). The results are display in the Table 27 along with physicochemical properties that differentiate strain JC586^T from the type strains of closely related members of the genus *Gliutamicibacter*.

3.1.2.3.3.8. Chemotaxonomic characterization

3.1.2.3.3.8.1. Cellular fatty acid, polar lipid, quinone and cell wall amino acid analysis

anteiso-$C_{15:0}$, iso-$C_{15:0}$, anteiso-$C_{17:0}$ were the major fatty acids found in strain JC586^T (>8 % of the total fatty acids). Table 28 depicts the differences in fatty acid profiles between strain JC586^T and related strains of the genus *Glutamicibacter*. DPG: diphosphatidylglycerol, phosphatidylglycerol, glycolipid, dimannosylglyceride, trimannosyldiacylglycerol are the composition of polar lipids in strain JC586^T (Fig. 51) The major isoprenoid quinones are MK-8 and MK-9 and lysine, alanine, glutamine and DAP were detected as the peptidoglycan composition in the cell wall of strain JC586^T (Fig. 52). Details of the chemotaxonomic characteristics of strain JC586^T along with the reference strains are given in Table 29.

Table. 27. Differential physiochemical characteristics between strain JC586^T and the type strains of related *Glutamicibacter* species.

Characteristics	**1.**	**2.***	**3.***	**4.***	**5.***	**6.***	**7.***	**8.***	**9.***	**10.***	**11.***	**12.***	**13.***
Colony colour	Light yellow	Light yellow	Pale yellow	Pale yellow	Pale yellow	White	Yellow	Yellow	Light Yellow	Pale yellow	Yellow	White	Yellow
pH range	5-11	5-11	4-10	6-10	5-10	5-10	5-12	5-11	7-8.5	7-8	6-11	5-10	6-11
NaCl (%) range	0-11	0-12	0-13	0-7	0-10	0-7	0-15	0-12	0-8	0-10	0-12	0-13	0-7.5
Temp range ˚C	16-37	28-37	4-37	15-37	15-30	15-37	16-40	15-37	15-30	4-37	16-37	5-30	10-30
API ZYM													
Easterase	+	+	-	+	-	+	+	+	+	+	+	+	+
Lipase	-	-	+	+	-	+	+	+	-	-	-	-	+
Valine-acrylamidase	+	-	-	+	+	-	+	+	+	+	-	+	+
Esterase lipase	+	+	+	+	-	+	+	+	+	+	+	+	+
Cystine-acrylamidase	-	-	-	-	-	-	+	+	-	+	-	-	+
Acid-phosphatase	+	+	-	+			+	+	+	+	-	+	+
α-Galactosidase	+	+	+	-	-	-	-	-	-	-	-	-	-
β-Galactosidase	-	+	-	-	+	-	-	-	-	-	-	-	+
β-Glucosidase	+	+	-	-	-	-	-	-	-	-	-	-	+
α-Mannosidase	+	+	+	+	+	+	+	+	+	+	-	+	+
API 50CH													
D-arabinose	-	-	-	-	-	+	-	-	-	-	-	+	-
L-arabinose	-	+	+	+	+	+	+	_-	+	-	-	+	+
D-xylose	-	+	+	+	+	+	-	-	+	-	-	+	+
L-xylose	-	-	-	-	-	+	+	-	-	-	-	+	+
D-galactose	-	+	-	+	-	+	+	-	+	-	-	-	+
D-glucose	+	+	+	+	-	+	+	-	+	+	-	+	+
D-mannose	+	+	-	-	-	-	+	-	+	-	-	+	-
L-rhamnose	+	+	+	-	-	+	-	-		-	-	-	-
Inositol	+	+	+	-	-	+	-	-	-	-	-	-	-
D-sorbitol	-	-	-	-	-	+	-	-	-	-	-	+	-
D-cellobiose	+	+	+	+	+	+	+	+	+	-	-	+	+
D-lactose	-	-	+	+	+	+	+	-	+	-	-	+	+
D-melibiose	-	+	+	-	-	-	-	-	-	-	-	-	-

***Data taken from literature**

Strains:1. JC586^T 2. *G. mishrai* LMG 29155^T; 3. *G. halophytocola* DSM 101718; 4 *G. nicotianae* DSM 20123^T; 5. *G. arilaitensis* DSM 16368^T; 6 *G. mysorens* DSM 12798^T; 7 *G. soli* DSM 19449^T; 8. *G. protophormiae* DSM 20168^T; 9. *G. ardleyensis* DSM 17432^T; 10. *G. uratoxydans* DSMZ 20647^T; 11. *G. creatinolyticus* DSM 15881^T 12. *G. endophyticus* JCM 30091^T 13. *G. bergerei* DSM 16367^T; + positive; - negative

Table. 28. Fatty acid profiles of strain JC586^{T} and the type strains of the related species of the genus *Glutamicibacter*

Fatty acid	**1.**	**2.***	**3.***	**4.***	**5.***	**6.***	**7.***	**8.***	**9.***	**10.***	**11.***	**12.***	**13.***
iso-$C_{14:0}$	1.7	4.3	1.6	1.7	2.4	3.2	1.2	1.3	0.5	2.9	2.5	1.7	1.8
$C_{14:0}$	1.11	3.7	2.5	2.8	1.0	2.4	0.5	0.9	1.1	1.1	1.5	0.4	0.4
iso-$C_{15:0}$	14.3	10.2	13.6	11.3	9.8	9.9	7.6	12.9	3.0	15.4	6.6	17.7	6.1
anteiso-$C_{15:0}$	50.2	47.6	41.5	46.1	69.5	57.2	50.0	41.6	67.3	51.4	77.8	47.2	53.1
iso-$C_{16:0}$	4.62	11.1	5	8.9	5.7	9.4	11.3	6.0	1.8	9.4	2.5	9.5	13.8
$C_{16:0}$	2.07	10.9	9.4	11.1	2.7	7.0	5.4	5.5	4.7	4.4	2.7	3.0	1.7
iso-$C_{17:0}$	1.7	1.2	2.8	2.2	0.6	0.9	3.0	4.1	0.3	2.2	0.3	4.4	1.3
anteiso-$C_{17:0}$	8.6	7.4	9	13.7	5.4	7.4	14.8	11.6	7.4	8.6	3.3	13.8	14.4
$C_{18:1}$ ω9c	0.5	0.5	0.6	0.2	0.2	0.2	0.6	2.1	8.9	0.3	0.1	0.1	0.4

*Data taken from literature

Strains: 1. JC586^{T} 2. *G. halophytocola* DSM 101718^{T}; 3.*G. mishrai* LMG 29155^{T}; 4. *G. nicotianae* DSM 20123^{T}; 5. *G. arilaitensis* DSM 16368^{T}; 6. *G. mysorens* DSM 12798^{T}; 7. *G. endophyticus* JCM 30091^{T}; 8. *G. soli* DSM 19449^{T}; 9. *G. bergerei* DSM 16367^{T}; 10. *G. protophormiae* DSM 20168^{T}; 11. *G. ardleyensis* DSM17432^{T}; 12. *G. uratoxydans* DSM 20647^{T}; 13. *G. creatinolyticus* DSM 15881^{T}

Table. 29. Chemotaxonomic characteristics of strain JC586^{T} and the type strains of related *Glutamicibacter* species

Strain	**1**	**2***	**3***	**4***	**5***	**6***	**7***	**8***	**9***	**10***	**11***	**12***	**13***
Peptidoglycan type	Lys, Ala, Glu, DAP	Lys, Ala, Glu	Lys, Ala, Glu	Lys, Ala, Glu	Lys, Ala, Glu	Lys, Ala, Glu	Lys, Ala, Glu	ND	Lys, Ala, Glu	Lys, Ala, Glu	Lys, Ala, Glu	Lys, Ala, Glu	Lys, Ala, Glu
Polar lipids	DPG, PG, GL, DMG, TMDG	DPG, PG, PL, GL, DMG, TMDG	DPG, PG, PL, GL, DMG, TMDG	DPG, PG, PL, DMG, MGDG, TMDG	DPG, PG, PL, GL, DMG, TMDG	DPG, PG, PL, GL, DMG, TMDG	DPG, PG, PL, GL, DMG, DGDG, TMDG	DPG, PG, PL, DMG, TMDG	DPG, PG, PL, DMG, TMDG	DPG, PG, DMG, DGDG, TMDG	DPG, PG, PL, GL, DMG, TMDG	DPG, PG, PL, GL, DMG, TMDG	DPG, PG, DMG, TMDG
Respiratory Quinone	MK-8, MK-9	MK-9	MK-7, MK-8, MK-9	MK-8, MK-9	ND	ND	MK-9	MK-7, MK-8, MK-9, MK-10	ND	MK-8, MK-9	MK-8, MK-9	ND	MK-8, MK-9

Strains:1. JC586^{T}; 2. *G. halophytocola* DSM 101718^{T} (Feng et al. 2017); 3. *G. mishrai* LMG 29155^{T} (Das et al. 2019); 4. *G. nicotianae* DSM 20123^{T} (Hou et al. 1998); 5. *G. arilaitensis* DSM 16368^{T} (Irlinger et al. 2005); 6. *G. mysorens* DSM 12798^{T} (Stackbrandt et al. 1983); 7. *G. endophyticus* JCM 30091^{T} (Wang et al. 2015); 8. *G. soli* DSM 19449^{T} (Busse et al. 2016) 9. *G. bergerei* DSM 16367^{T} (Irlinger et al. 2005); 10. *G. protophormiae* DSM 20168^{T} (Hou et al. 1998); 11. *G. ardleyensis* DSM 17432^{T} (Wang et al. 2015); 12. *G. uratoxydans* DSM 20647^{T} (Hou et al. 1998); 13. *G. creatinolyticus* DSM 15881^{T} (Hou et al. 1998).

*** Data taken from literature; N.D: Not Determined, DAP: Diaminopimelic acid, Lys: Lysine, Ala: Alanine, Glu: Glutamine, DPG: diphosphatidylglycerol, PG: phosphatidylglycerol, GL: glycolipid, DMG: dimannosylglyceride, PL: phospholipid (PL), TMDG: trimannosyldiacylglycerol**

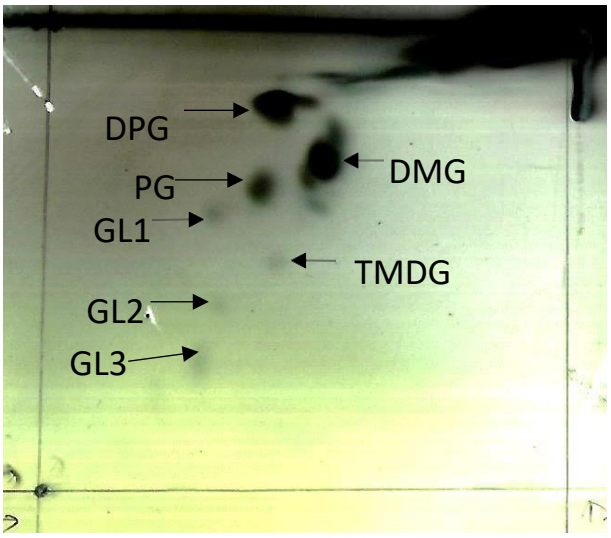


Fig. 52. Two-dimensional thin-layer chromatogram showing the polar lipid profile of strain JC586^T
DPG: Diphosphatidylglycerol;
DMG: Dimannosylglyceride;
TMDG: Trimannosyldiacylglycerol;
PG: Phosphatidylglycerol;
GL: Glycolipid;

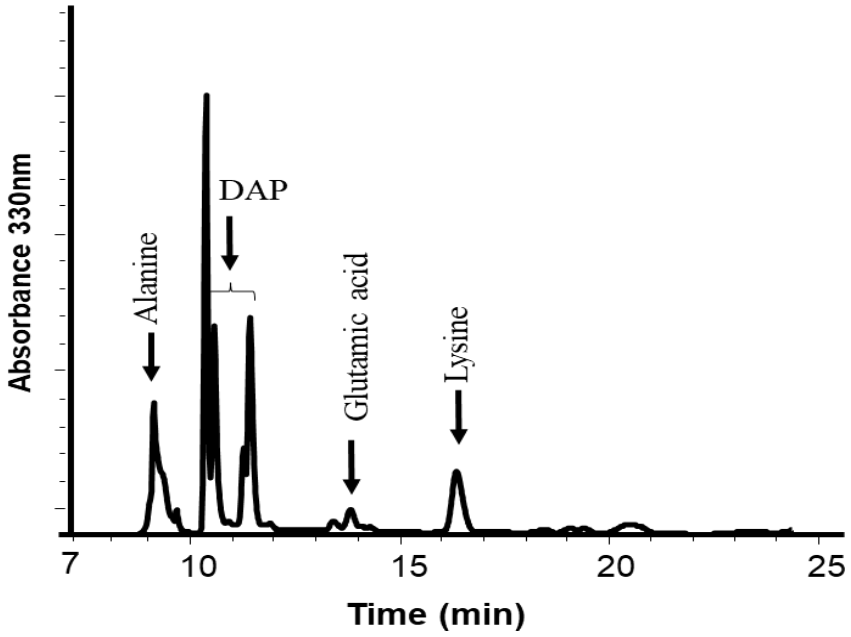


Fig. 53. HPLC chromatogram showing cell wall amino acid peaks of strain JC586^T

3.1.2.3.4. Characterization of strain JC611^T

3.1.2.3.4.1. Habitat and nucleotide accession numbers

Strain JC611^T was isolated from the lake sediment of Loktak lake, located in the north-eastern part of India, Manipur (precise location 24°30'21" N 93°47'43" E). At the time of sampling, the lake's pH was 7. After a few days of streaking, violet colour colonies were purified on nutrient agar (NA) and preserved in 50 % glycerol stock at -20 ˚C. The 16S rRNA gene and whole genome sequences of strain JC611^T have been deposited under the accession number ON908989 and JAFLJN000000000.

3.1.2.3.4.2. BLAST analysis, phylogenetic inference and genomic characteristics

BLAST analysis of the 16S rRNA gene sequences (1436 nt) of strain JC611^T in the EZBioCloud shows that it shares the highest sequence similarity of 98.8 % with the species of the genus *Janthinobacterium*; *Janthinobacterium lividum*. The 16s rRNA gene sequence-based phylogenetic tree based on NJ, ML, ME (Fig. 55) and 92 core genes based phylogenomic tree (Fig. 56) confirmed that strain JC611^T forms a distinct or monophyletic clade within the genus *Janthinobacterium*, suggesting that strain JC611^T forms a novel species within the genus *Janthinobacterium*.

Strain JC611^T has a genome e size of 3.52 Mb with an N_{50} value of 558,740, and a genomic G + C content (mol %) of 60.4 %. The ANI and *d*DDH value between strain JC611^T and *J. rivuli* was 95 % and 62.7 % which were less than the recommended values for species delineation (95-96 % for ANI and 70 % cut off for *d*DDH) (Fig. 54), thus supporting the proposal of a novel species within the genus *Janthinobacterium*.

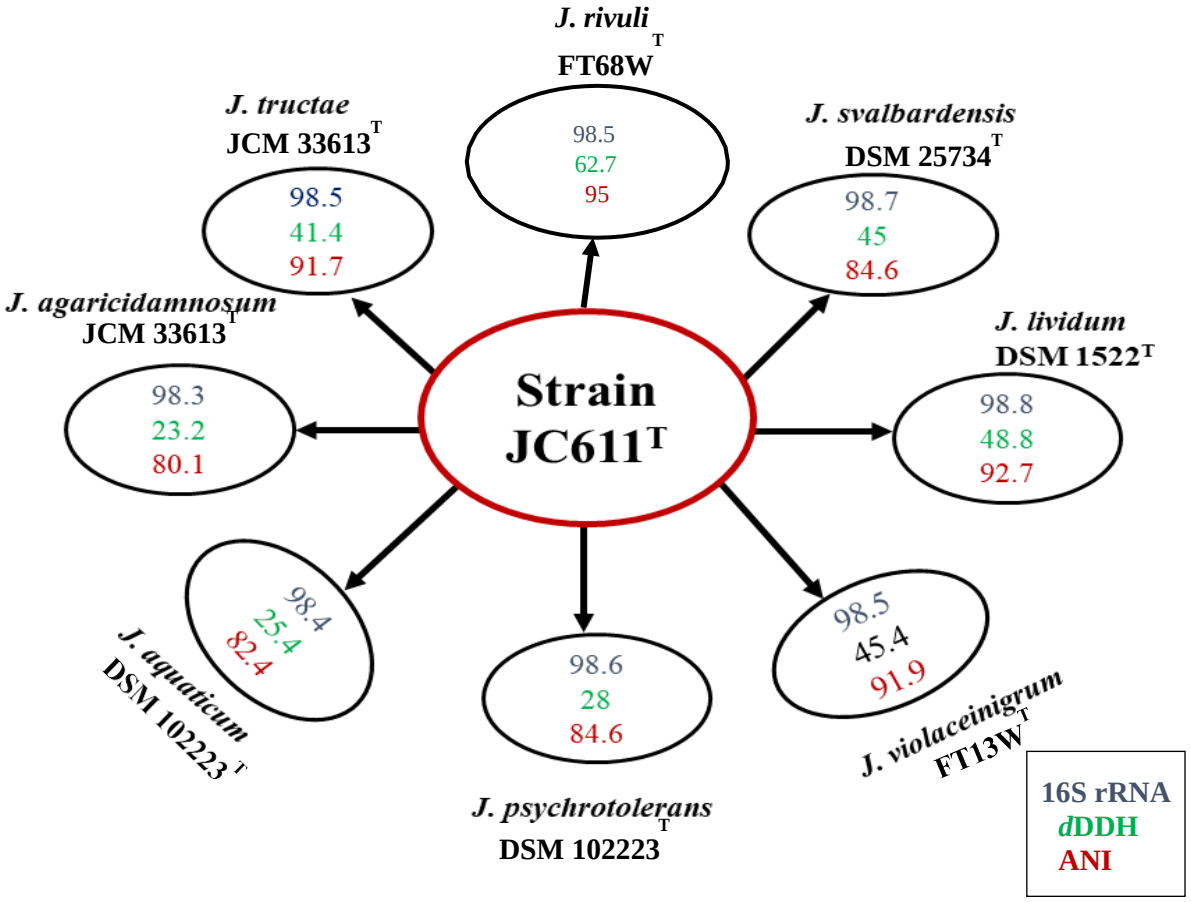


Fig. 54. Analysis of phylogenetic markers for delineation of the strain JC611^{T} along with reference strains of the genus *Janthinobacterium*. Methods used: 16S rRNA gene identity (16S), average amino acid identity (ANI), digital DNA-DNA hybridization (*d*DDH)

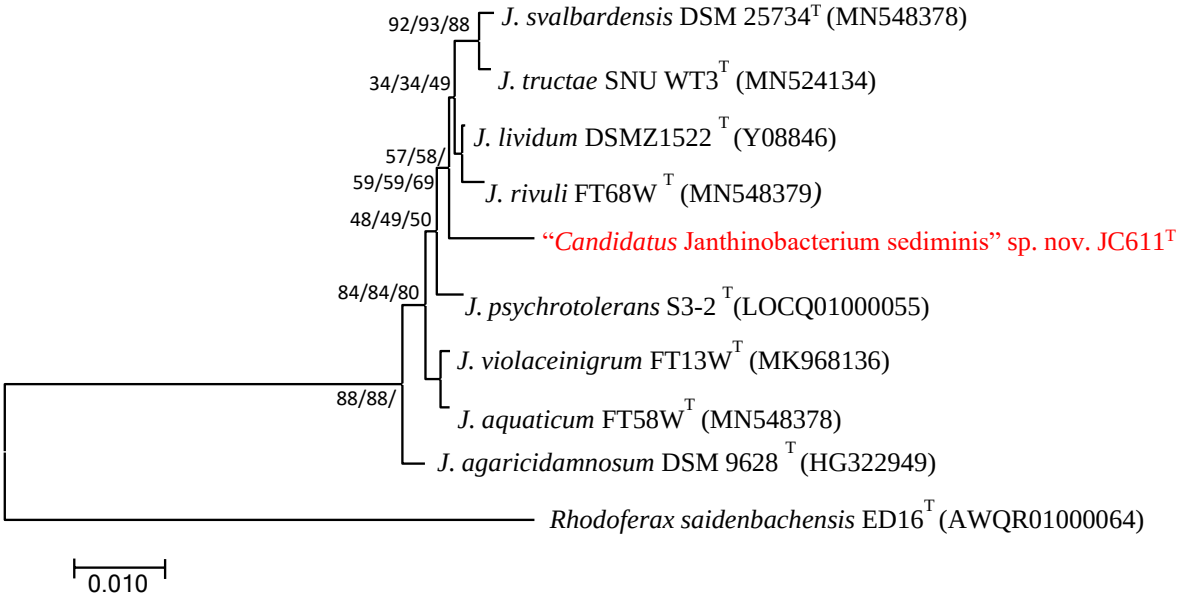


Fig. 55. Phylogenetic tree of strain JC611^{T} based on 16S rRNA gene sequence comparisons, using the Neighbour-joining algorithm (NJ) and the additional algorithms like ME and ML. *Rhodoferax saidenbachensis* ED16^{T} was used as an outgroup. Numbers at the nodes represent the bootstrap values. On the nodes Bar, 0.01 accumulated changes per nucleotide

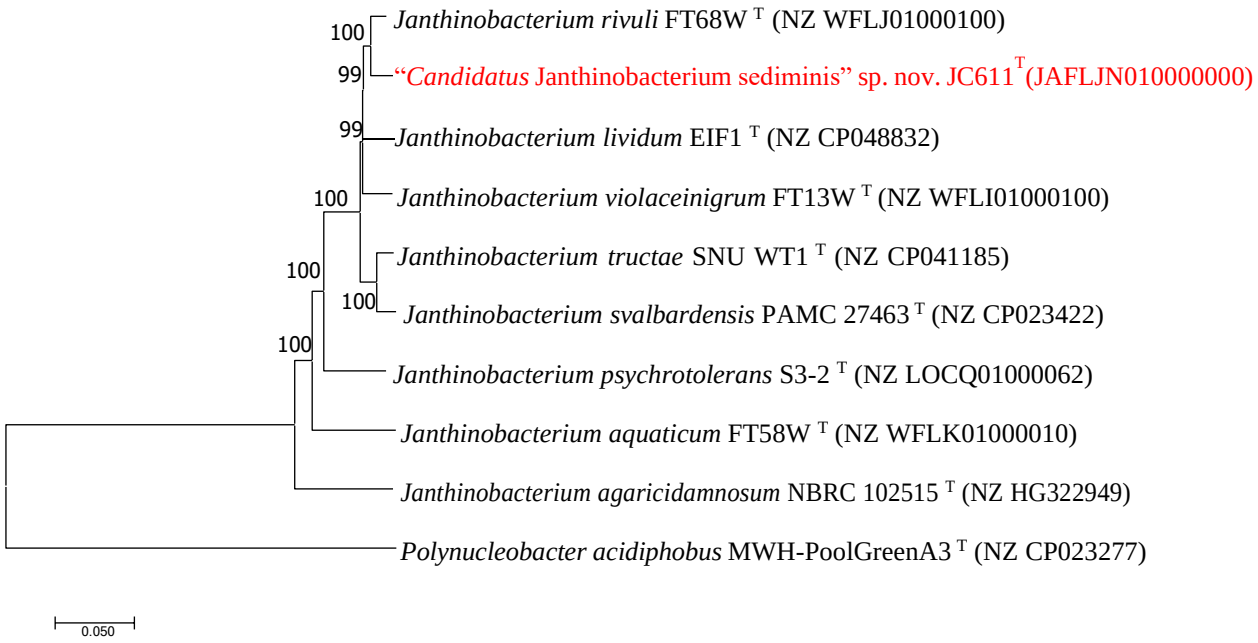


Fig. 56. Phylogenomic tree of strain JC611^{T} based on 92 core genes tool based on the Up-to-date Bacterial Core Gene (UBCG) (Na et al. 2018). The tree was generated using the MEGAX software (NJ) with *Rugamonas rubra* ATCC43154 as an outgroup

3.1.2.3.4.3. *In-silico* metabolic characterization

Functional prediction of gene cluster with BLASTKOALA according to the Kyoto Encyclopedia of Genes was conducted with the genomes of strain JC611^{T} along with reference strains of the genus *Janthinobacterium*. Among all categories, genes involved in signalling and cellular process, environmental information process, genetic information process were most abundant (Fig. 57). *Janthinobacterium* members have the ability to produce significant secondary metabolites (SMs) with remarkable antibacterial, antifungal, antiviral, and antiprotozoal capabilities, making them a prospective source for new pharmaceutical compounds (Ines et al. 2020). Genome analysis with AntiSMASH v5.1.2 (Blin et al. 2019) also reveals that strain JC611^{T} comprised of secondary metabolites clusters where genes involve in synthesis of violacein, terpenes and bacteriocins were

detected (Fig. 58). *In-silico* predictions also indicate the fermentation capability of the novel strain JC611^T, utilization of substrates like lactate, degradation of aromatic hydrocarbons like benzoate, catechol, phenylacetate, capability to utilize urea and reduce nitrate.

3.1.2.3.4.4. Morphological, physiological and biochemical analysis

Strain JC611^T colonies were violet coloured, gram-stain negative, motile and aerobic. Cultures lose viability on refrigeration. Cells of strain JC611^T were non-viable, rod shaped with 0.5-0.7 µm in length and 0.3-0.5 µm in width (Fig. 59). Strain JC611^T could grow well in a temperature between 16-28 ˚C with an optimum temperature of 28 ˚C, pH range between 6-9 (optimum pH 7) and NaCl tolerance between 1-2 % (optimum 0.5 %). Strain JC611^T shows positive for catalase, urease and oxidase activity, however, it did not hydrolyse starch and gelatin. Table 30 and 31 shows the differential phenotypic and genotypic characters of strain JC611^T and the members of the genus *Janthinobacterium.*

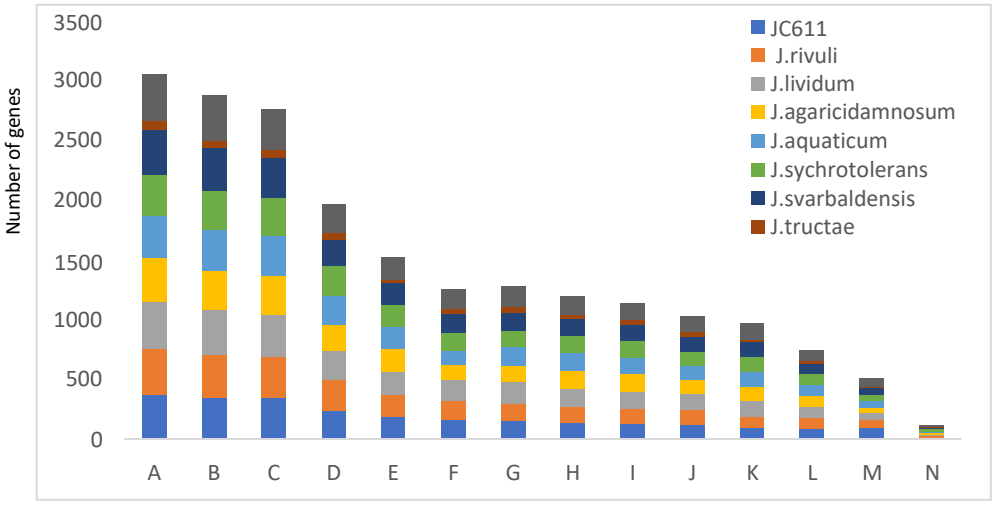


Fig. 57. Functional prediction of gene clusters was conducted with BLASTKOALA according to the Kyoto Encyclopedia of Genes and Genomes of straJC611^T along with the reference strains of the genus *Janthinobacterium*

A. Signalling and cellular process; B. Environmental information process; C. Protein families; D. Genetic information process; E. Genetic information process; F. Unclassified; G. Unclassified metabolism; H. Amino acid metabolism; I. Cellular process; J. Metabolism of cofactors and vitamins; K. Energy metabolism; L. Neucleotide metabolism; M. Lipid metabolism; N. Terpenoids and polyketides metabolism

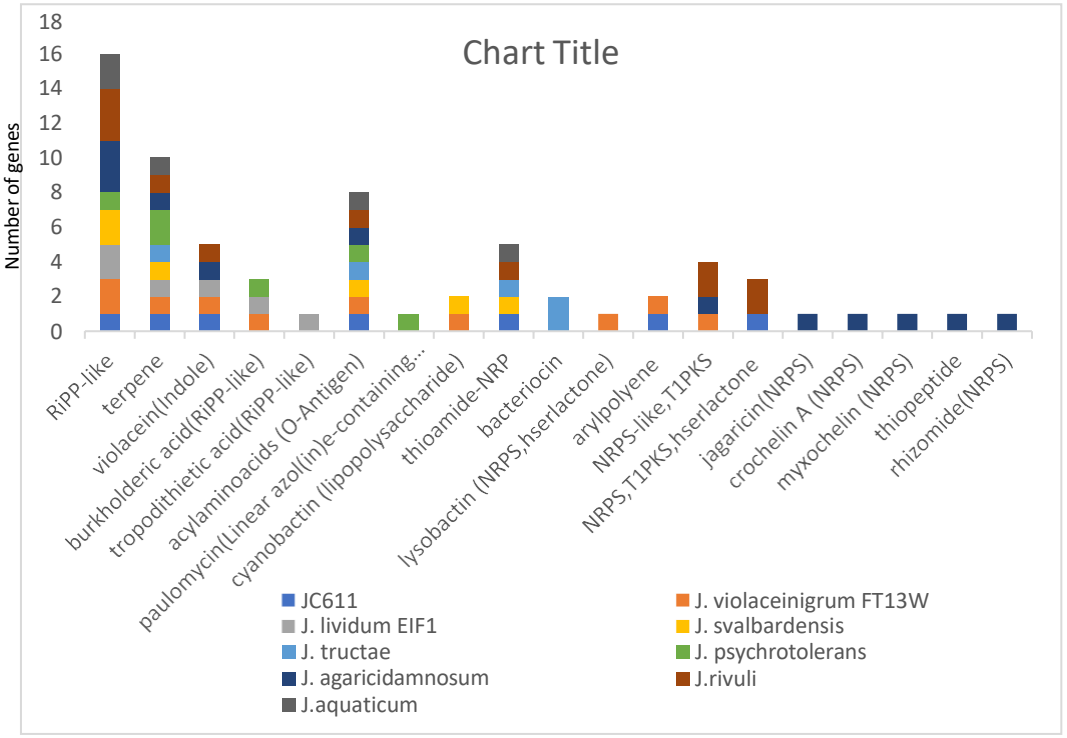


Fig. 58. Genome analysis of strain JC611^T with AntiSMASH v5.1.2 showing the putative SM (Secondary Metabolites) gene clusters

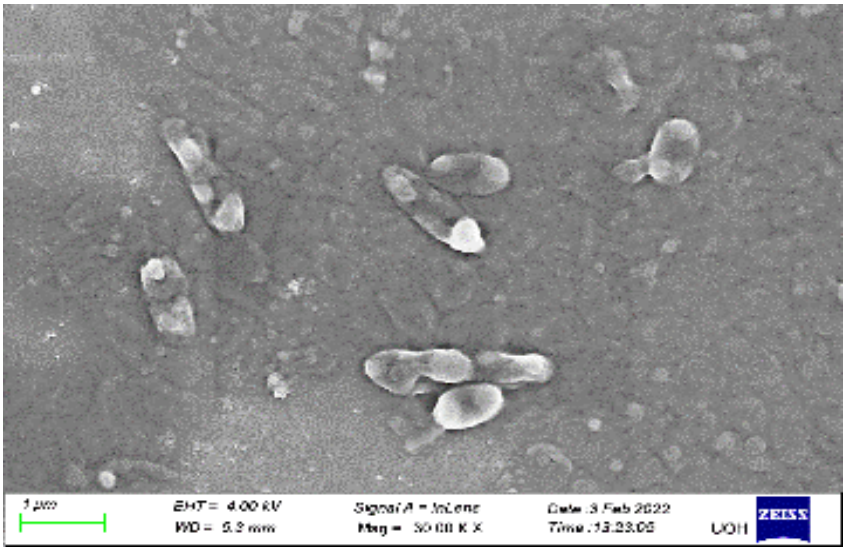


Fig. 59. SEM images showing rod-shaped cell morphologies of strain JC611^T grown in NA for 3 days at 30˚C

Table. 30. Differential physiochemical characteristics of strain JC611^{T} along with its related strains of the genus *Janthinobacterium*

	Strain JC611^{T}	***J. lividum* DSM 1522^{T*}**	***J.rivuli* FT68W^{T*}**	***J. agaricidamnosum* DSM 9628^{T*}**	***J. violaceinigrum* FT13W^{T*}**	***J. aquaticum* FT58W^{T*}**	***J. psyhrotolerans* DSM 102223^{T*}**	***J. tructae* JCM 33613^{T*}**	***J. svarbaldensis* DSM 25734^{T*}**
Physiological characters									
Colony colour	Violet	Violet	shallow violet	Beige	Purple,white	White	Pale yellow, mucoid	Whitish	Aubergine
Colony shape	Round, convex pigmented	Circular, convex	convex	Round, convex	Circle, convex	Convex, circle,	Convex, circle	Convex, circle	Convex, circle
Violacein pigment	+	+	+	-	+	-	-	-	+
pH range (optimum)	6-9 (7)	4-9 (7-8)	4-9.5 (7-7.5)	4-10 (7-7.5)	4.5-10.5 (7-8)	4.5-9.5 (7-7.5)	6-8 (7)	4-7	4.5-9.5 (7-8)
Temp˚C range (optimum)	16-30 (28)	2-30 (24)	4-30 (24)	4-30 (24)	4-34 (24)	4-30 (24)	3-30 (25)	2-30	4-24 (20)
NaCl range	1-2 (0.5)	1-2 (0-0.5)	0-2 (0-0.5)	0-1.5 (0.5)	0-2 (0.5)	0-2.5 (0-0.5)	0-3 (2)	0-6	0-0.5 (0.5)
Starch-hydrolysis	-	ND	-	ND	-	-	-	ND	ND
Gelatin- hydrolysis	-	+	+	ND	+	-	-	-	ND
Catalase-activity	+	+	+	+	+	+	+		+
Urease activity	+	-	+	ND	+	+	-	-	ND
Oxidase activity	+	+	+	+	+	+	+	ND	ND
Respiratory quinone(Q-8)	+	+	+	+	+	+	+	+	+
Source	Soil	Pond	Stream	Fresh water	Stream	Stream	Freshwater	kidney	freshwater

*** Data taken from literature; Strains:1. JC611^{T} ; 2. *J. lividum* DSM 1522^{T*} (Lu et al. 2020); 3. *J. rivuli* FT68W^{T}* (Lu et al. 2020); 4. *J. agaricidamnosum* DSM 9628^{T*} (Lu et al. 2020); 5. *J. violaceinigrum* FT13W^{T*} (Lu et al. 2020); 6. J. *aquaticum* FT58W^{T*} (Lu et al. 2020); 7. *J. psyhrotolerans* DSM 102223^{T*} (Gong et al. 2017); 8. *J. tructae* JCM 33613^{T*}(Jung et al. 2021); 9. *J. svarbaldensis* DSM 25734^{T*} (Lu et al. 2020); +, positive; -, negative; ND, not determined**

Table. 31. Differential genotypic characteristics of strain JC611^{T} along with its related strains of the genus *Janthinobacterium*

	1	**2***	**3***	**4***	**5***	**6***	**7***	**8***	**9***
Genomic characters									
Genome size (mb)	3.52	6.3	6.3	5.95	6.45	5.7	5.8	6.3	6.2
mol. % G+C	60.4	62.4	63.1	61.1	63.4	61.6	63	62.4	62.1
No. of contigs	28	1	116	1	152	83	61	1	1
N_{50}(bp)	558,740	6,373,589	278,928	5,949,001	136,055	283,907	261,276	6,314,370	6,274,078
L_{50} (bp)	3	1	10	1	15	7	8	1	1
rRNAs	10	25	43	7	31	44	25	25	25
tRNAs	62	93	86	74	81	86	78	93	91
Genes (coding)	3,177	5,535	5,670	5077	5,665	4,947	5,167	5,450	5,457
Pseudo genes	42	42	92	110	75	86	119	58	84
CRISPRs	6	1		3	0	9	3	0	1
Biosample accession Number	SAMN12096126	SAMN14082727	SAMN13015371	SAMEA3139017	SAMN13015368	SAMN13015370	SAMN04229034	SAMN12084176	SAMN07638148
Genbank WGS[a] accession Number	NZ_VHIN01000000	NZ_CP048832	NZ_WFLJ00000000	NZ_HG322949	NZ_WFLI00000000	WFLK01000000	NZ_WFLK00000000	NZ_CP041185	NZ_CP023422
GenBank assembly accession no.	GCF_009805585.1	GCA_013372045.1	GCF_009208735.1	GCF_000723165.1	GCF_009208555	GCF_009208565.1	GCF_001677885.1	GCF_006517255.1	GCF_002327145.1

*** Data taken from literature; Strains:1. JC611^{T}; 2. *J. lividum* DSM 1522^{T*} (Lu et al. 2020); 3. *J. rivuli* FT68W^{T}* (Lu et al. 2020); 4. *J. agaricidamnosum* DSM 9628^{T*} (Lu et al. 2020); 5. *J. violaceinigrum* FT13W^{T*} (Lu et al. 2020); 6. *J. aquaticum* FT58W^{T*} (Lu et al. 2020); 7. *J. psyhrotolerans* DSM 102223^{T*} (Gong et al. 2017); 8. *J. tructae* JCM 33613^{T*} (Jung et al. 2021); 9. *J. svarbaldensis* DSM 25734^{T*} (Lu et al. 2020)**

3.1.2.3.4.5. Characterization of violacein-like pigment

The violacein-like pigment (in methanol) showed a peak with absorption at 580-750 nm (Fig. 60), this absorption peak is mainly due to the presence of chromophores such as carbonyl and alkene groups which are responsible for electron absorption.

The Fourier transform infrared (FTIR) analysis results of the pigment are as follows: 3,700–3,000 cm−1(OH), 3,256.53 cm−1(N–H), 1,635.13 cm−1(C= O amide), 1,543.93 (C–N), 1,223.69 cm−1(C–O phenol) (Fig. 61).

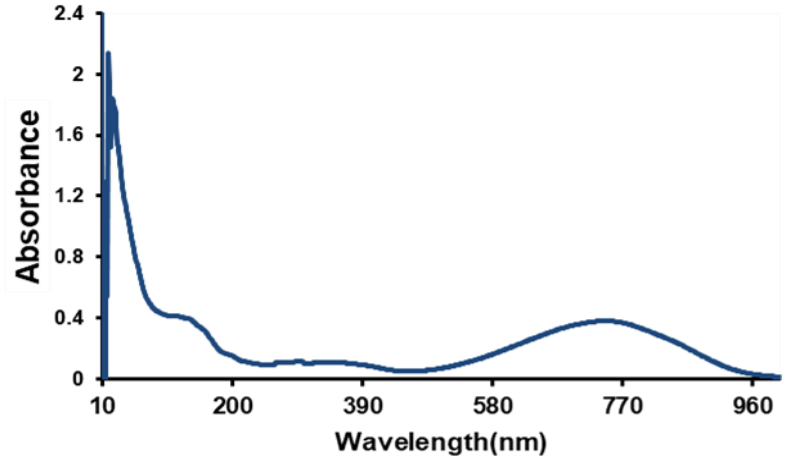


Fig. 60. UV-vis spectrum of violacein like pigment of strain JC611 [T]

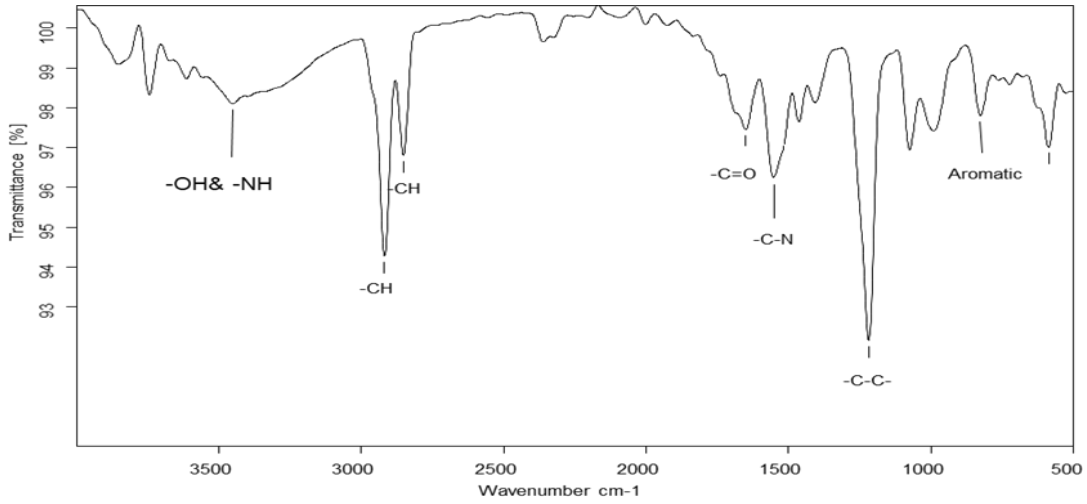


Fig. 61. FTIR spectra of violacein like pigment of strain JC611 [T]

3.1.2.3.5. Characterization of strain JC664^T

3.1.2.3.5.1. Habitat and nucleotide accession numbers

Strain JC664^T was isolated from soil sediment collected from Loktak lake, Manipur, India (GPS positioning of the sample collection site: 24˚30' 94''N and 93˚47'43 E). Translucent pale-yellow colonies which appear after few days of streaking were purified by streaking on nutrient agar (NA) and preserved in 50 % glycerol stock stored at -20 ˚C. The 16S rRNA gene sequences of strain JC664^T have been deposited in GenBank under the accession number ON908990.

3.1.2.3.5.2. BLAST analysis and phylogenetic inference

BLAST analysis of the 16S rRNA gene sequences (1346 nt) of strain JC664^T in the EZBioCloud shows that it shares the highest sequence similarity of 98.5 % with the species genus *Comamonas*; *Comamonas koreensis* KCTC 12005^T. The 16S rRNA gene sequence-based phylogenetic tree with combined bootstrap values obtained from NJ, ME, and ML trees (Fig. 62) confirmed that strain JC664^T forms a distinct monophyletic clustering within genus *Comamonas* indicating that strain JC664^T represents a putative novel species within the genus *Comamonas*.

3.1.2.3.5.3. Morphological, physiological and biochemical analysis

Strain JC664^T was found to be aerobic, non-pigmented, pale translucent yellow, and rod-shaped in a nutrient broth medium. Cells were Gram-stain-negative, non-motile, and non-endospore forming. Growth occurs at a temperance range between 18-37 ˚C (optimum 28 ˚C) and a pH range between pH 6-7 (optimum pH 7). NaCl is not obligate for the growth of the strain and can tolerate up to 3 % (w/v). Casein, starch, tween 80 were hydrolysed whereas tween 20 and gelatin were not hydrolysed. Strain JC664^T showed

positive for utilisation of carbon sources like glucose, alanine, citrate, sodium propionate and nitrogen sources like glutamic acid, phenylalanine, leucine and threonine. Strain JC664^{T} also showed positive for catalase, oxidase, indole production, α-Glucosidase, α-Mannosidase, Alkaline phosphatase, Trypsin, Leucine arylamidase, β - Glucosidase and negative for *N*-acetyl-β-glucosaminidase, Cystine arylamidase, α-Fucosidase, Acid phosphatase, Cystine arylamidase, and β - Galactosidase. The following antibiotics showed sensitivity to strain JC664^{T}: chloramphenicol and nalidixic acid. Tabel 32 shows the differential physiological and biochemical analysis of strain JC664^{T} along with close members of the genus *Comamonas*.

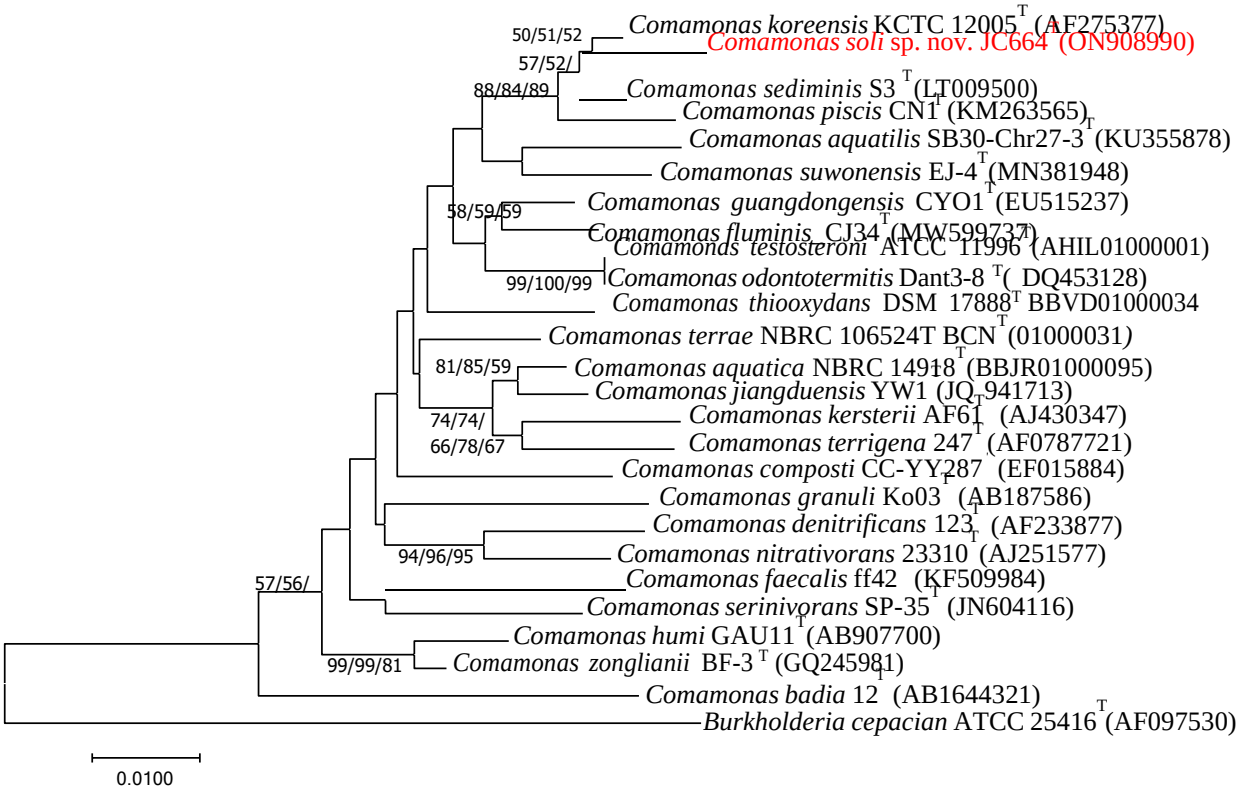


Fig. 62. Phylogenetic tree based on 16S rRNA gene sequences showing the affiliation of strain JC644^{T} and the nearest members of the genus *Comamonas*. The tree was constructed by the neighbour-joining method using the MEGA7 software and rooted by using *Burkholderia cepacian* ATCC 25416^{T} as the out-group. Numbers at nodes denote bootstrap values (based on 1000 resamplings). Bootstrap percentages refer to NJ/ML/MP analysis. Bar 1nucleotide substitutions per 100 nucleotides.

Table. 32. Differentiating characteristics of the species of the genus *Comamonas*

	# Strain JC664^T	***C. koreensis***	***C. sediminis***
Source	sediment	sediment	sediment
Colony colour	Translucent pale yellow	Translucent pale yellow	transparent
Colony shape	Circular rod	curved rod	rod shaped
pH range (optimum)	6-7 (7)	6.0-8.9 (7-7.5)	6.1-8.8 (6.5-7.7)
Temp range(˚C) (optimum)	18-37 (28)	10-40 (28-30)	12-45 (30-35)
NaCl tolerance % (w/v)	0-3 (0-0.5)	0-3.6 (0-0.6)	0-3.5 (0-1.5)
Indole production from Tryptophan	+	+	-
Hydrolysis of			
Casein	+	+	+
Starch	+	-	+
Tween 20	-	-	+
Tween 80	+	+	-
Biochemical characters			
Catalase	+	+	+
Oxidase	+	+	+
Utilisation of			
Glucose	+	+	+
Mannitol	-	+	+
Rhamnose	-	+	-
Sorbitol	-	+	+
Sucrose	-	+	+
Fructose	-	-	-
Inositol	-	+	ND
Na. propionate	+	-	ND
Citrate	+	-	ND
Glutamic acid	+	ND	ND
Leucine	-	+	ND
Threonine	+	-	ND
Phenylalanine	+	+	ND
Serine	-	-	ND

Susceptibility to			
Chloramphenicol	+	+	-
Ampicillin	-	-	-
Rifampicin	-	+	-
Nalidixic acid	+	+	+
API enzyme activity			
Acid phosphatase	-	+	-
α-Glucosidase	+	+	-
N-acetyl-β-glucosaminidase	-	-	+
α-Mannosidase	+	+	-
α-Fucosidase	-	-	+
Cystine arylamidase	-	-	+
Alkaline phosphatase	+	-	-
Trypsin	+	-	-
Leucine arylamidase	+	-	-
β - Galactosidase	-	+	-
β - Glucosidase	+	+	+

 # Data from our study; ND: Not determined

1. Strain JC664[T], 2. *C. koreensis* KCTC 12005[T] (data taken from Subhash et al. 2016), 3. *C. sediminis* S3[T] (data taken from Subhash et al. 2016)

3.1.2.3.5.4. Chemotaxonomic characterization

3.1.2.3.5.4.1. Cellular fatty acid, polar lipid and quinone analysis

Strain JC664^T had $C_{16:1}\omega6c$ / $C_{16:1}\omega7c$, $C_{18:1}\omega6c$ / $C_{18:1}\omega7c$, $C_{17:0}$ cyclo and $C_{16:0}$ as the major predominant fatty acids with minor amounts of $C_{12:0}$, $C_{14:0}$, $C_{18:0}$, iso-$C_{17:0}$ 3OH, $C_{10:0}$ 3OH and $C_{16:0}$ 2OH. The differences in fatty acid profiles of strain JCC4^T along with its closest members are given in Table 33. The major polar lipids of strain JC664^T was phosphatidylglycerol, phosphatidylethanolamine, unidentified phospholipid, and unidentified lipid (L1). One unidentified lipid (L1) helps in the differentiation of strain

JC664^T from its nearest strain *C. koreensis* KCTC 12005^T (Fig. 63). Menaquinone 8 (MK-8) was the major isoprenoid quinone in strain JC664^T.

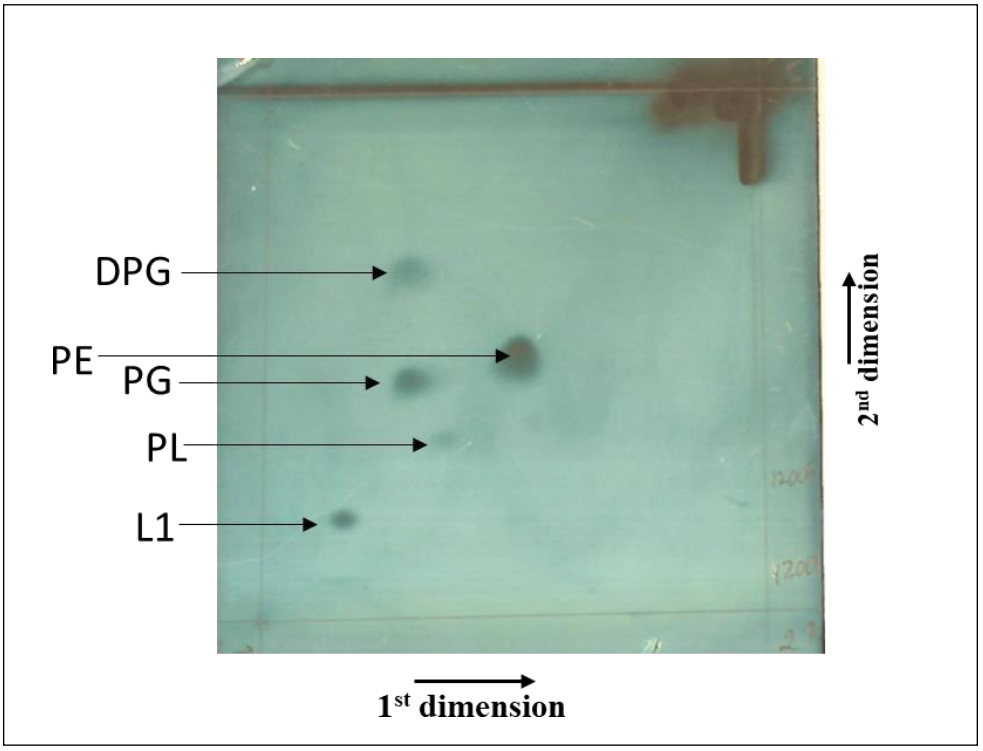


Fig. 63. Two - dimensional thin-layer chromatogram of strain JC644^T
DPG: Diphosphatidylglycerol;
PE: Phosphatidylethanolamine;
PG: Phosphatidylglycerol;
PL: Unidentified phospholipid and
L1: Unidentified lipid

Table 33. Cellular fatty acid composition, polar lipids and quinones of strain JC664^T and nearest members of the genus *Comamonas*. # Data from our study; *C. koreensis* KCTC 12005^T; *C. sediminis* S3^T (Data taken from Subhash et al. 2016)

Composition	**# Strain JC664^T**	***C. koreensis* KCTC 12005^T**	***C. sediminis* S3^T**
Fatty acids %			
$C_{12:0}$	4.69	3.66	2.66
$C_{14:0}$	1.57	0.89	0.61
$C_{16:1}\omega6c/C_{16:1}\omega7c$	21.8	24.58	27.46
$C_{16:0}$	30.42	33.03	32.51
$C_{17:0}$ cyclo	11.94	10.83	13.73
$C_{18:0}$	1.48	-	-
$C_{18:1}\omega6c$ /C_{18}: 1$\omega7c$	13.5	10.22	15.01
iso-$C_{17:0}$ 3OH	3.59	-	-
M $C_{10:0}$ 3OH	4.84	3.5	3.60
$C_{16:0}$ 2OH	3.11	2.19	2.29
Polar lipids	Diphosphatidyl-glycerol, phosphatidyl-glycerol, Phosphatidyl-ethanolamine, Unidentified phospholipid, Unidentified lipid (L1)	Diphosphatidyl-glycerol, Phosphatidyl-ethanolamine, Unidentified phospholipid, Unidentified lipid (L3)	Diphosphatidyl-glycerol, Phosphatidyl-ethanolamine, Unidentified phospholipid, Unidentified lipid (L1, L2)
Quinones	Q-8	Q-8	Q-8

DISCUSSION

4. DISCUSSION

4.1. Diversity of bacterial communities of Loktak lake

4.1.1. Culture independent assessment of bacterial communities

The metagenome analysis of Loktak lake revealed a plethora of bacterial communities using amplicon-based sequencing, which targets different groups of bacteria present in the lake. It provides in-depth knowledge of the taxonomic distribution and functional potential of the lake. From the results of taxonomic profiling of the lake, bacteria were the major domain detected along with few archaeal reads. At the phylum level, among the bacteria, *Pseudomonadota* was the most dominant phylum, followed by, *Acidobacteriota*, *Chloroflexota* and *Actinomycetota* (Fig. 6). *Pseudomonadota* has been reported to be the most diverse group of microbes in freshwater lakes, regulating water quality and pollution (Yannarell and Kent 2009, Kurilkina et al. 2016). Moreover, they are also involved in the metabolism of sulphur, methane, and nitrogen fixation (Kaushal et al. 2022). The predominance of *Pseudomonadota* in this study is consistent with the previous metagenomic findings of lake water systems in India, where they revealed the dominance of *Pseudomonadota*, *Bacteroidota*, *Chlorofexota*, *Bacillota*, *Actinomycetota*, *Cyanobacteria*, *Verrucomicrobiota*, and *Planctomycetota* (Puranik et al. 2016; Kangabam et al. 2020; Ahmad et al. 2021; Ghosh et al. 2021; Rathour et al. 2017; Raina et al. 2019; Chakraborty et al. 2020).

Other important phyla such as *Verrucomicrobiota*, *Planctomycetota*, *Bacillota* and *Spirochaetota* were also identified in less (<3 %) abundance in the metagenome analysis (Fig. 6). *Verrucomicrobiota* are associated with the degradation of high-molecular-weight polysaccharides found in organic matter (Tran et al. 2018). *Planctomycetota* plays an

important role in the global carbon and nitrogen cycle (Pushpakumara et al. 2023). *Bacillota* are associated with carbohydrate polymers decomposition and biomass degradation (Govil et al. 2021). They also serve as a possible marker of anthropogenic intervention (Vijayan et al. 2023). The present study also identified the phylum *Spirochaetota*, which had not been reported in earlier metagenomic analyses of Loktak Lake. *Spirochaetota* are fermenting bacteria that have the potential to depolymerise EPS, which liberates $Ca^{2+,}$ which may further contribute to microbialite formation (Schneider et al. 2013; Monteiro et al. 2020).

Novel "*Candidatus*" phyla (GAL15, LCP-89, FCPU426, NKB15, RCP2-54, SAR324, Sva0485, TA06, WOR-1, WPS-2, WS2, WS4, MBNT15) were also found in the present study, which were not reported in previous metagenomic studies of Loktak lake suggesting that the lake ecosystem has a reservoir of potentially undiscovered novel biodiversity. These "*Candidatus*" members are part of the "Candidate Phyla Radiation" (CPR) which is an expansion of the tree of life that has grown as a result of metagenomic and single-cell analysis (Danczak et al. 2017). The lake metagenome also revealed unique phylotypes of bacterial communities of whose metabolic and ecological functions remain unknown. These unique phylotypes were given the name "*Ca*. UH''.

This study also provided information on the taxonomic diversity on Archaea phyla such as *Crenarchaeota*, *Halobacterota* and *Thermoplasmatota* which were the most abundant at the phylum level (Fig. 17). Reports on freshwater systems have demonstrated the role of *Crenarchaeota* in ammonia-oxidation and nutrient cycling in both lentic and lotic habitats (Ghai et al. 2011). Studies also show that *Crenarchaeota* and *Halobacterota* were found to be responsible for dissimilatory sulfate reduction in subtropical mangrove wetland ecosystems (Mo et al. 2020). *Euryarchaeota* present in this study are also reported

to play crucial role biogeochemical cycles such as methanogenesis (Chakraborty et al. 2020).

Among the four samples, LOK4 has the highest bacterial diversity indices (Shannon, Inverse Simpson and observed richness), indicating that it has the highest species distribution compared to the other three samples, i.e. LOK1, LOK1, and LOK3. Currently, about ninety- two bacteria and thirty Archaea phyla are recognised by 16S rRNA databases even though the true phyla count might be higher and could range up to 1500 bacterial phyla (Solden et al. 2016). Therefore, updating the Candidate Phyla (CP) database is needed as it would help in understanding the microbial diversity in nature and expanding the tree of life as a result of metagenomic-based sequencing techniques (Oren 2022). The present study highlights the enormous and rich bacterial diversity and the high abundance of cultured and uncultured candidate phyla.

From the point of view of the function of microbiota, the most abundant functional classification was amino acid metabolism, followed by membrane transport, carbohydrate metabolism, replication and repair, and energy metabolism. The functional analysis clearly reflects the diversity of metabolic pathways, suggesting that such ecologically and functionally wealthy ecosystems should be conserved. It also opens up a wide range of opportunities for the investigation of secondary metabolites of industrial importance and bioremediation agents. In addition to this, culture-dependent approaches can be employed for studying the metabolism of important compounds such as carotenoids, terpenoids, polyketides, vitamins, and other secondary metabolites with potential food and pharmaceutical applications.

4.1.2. Culture-dependent analysis of bacterial communities

The culture-dependent method based on 16S rRNA gene sequences revealed the presence of bacteria belonging to 37 genera across different phyla, i.e., *Actinomycetota* (41 %), *Pseudomonadota* (37.5 %), *Panctomyceota* (10 %), *Bacteroidota* (2 %) and *Bacillota* (2 %). The results of the culture-dependent approach revealed significant similarities with the metagenome data (Table 16). Out of the total isolated strains, strains JC665^T, JC656^T, JC586^T, JC611^T and JC664^T were identified as novel species based on 16S rRNA gene identity (%) and overall genome related index (OGRI) like *d*DDH, *g*ANI, and AAI.

4.1.2.1 Proposal of *Paludisphaera rhizosphaerae* sp. nov.

Strains JC665^T and JC747 have distinct phylo-genomic differences with "*P. soli"* JC670^T and *P. borealis* DSM 28747^T, however based on the phylogenetic and phylogenomic inferences between them, strains JC665^T and JC747 belong to the same species of the genus *Paludisphaera* (Fig. 23; Fig. 24, Fig. 25, Table. 18). The phylo-genomic differences are well supported by chemotaxonomic and phenotypic differences (Table. 20), which support strain JC665^T as a novel species of the genus *Paludisphaera.* For this, we propose the name of the type strain JC665^T as *Paludisphaera rhizosphaerae* and strain JC747 as its non-type strain.

4.1.2.1.1. Descriptions of *Paludisphaera rhizosphaerae* sp. nov.

Paludisphaera rhizosphaerae (rhi.zo. sphae'rae. Gr. n. *rhiza*, root; L. n. *sphaera*, sphere; N.L. gen. n. *rhizosphaerae*, from the rhizosphere)

Cells are oval to spherical shaped, strictly aerobic and cell division is through budding. Colonies are pale pink coloured, NaCl is not obligate for growth and can tolerate

up to 2 % (w/v). The optimum pH and temperature for growth are 7.0 (range 6.0-9.0) and 25 ºC (range 4-34 °C), respectively. N-acetylglucosamine (NAG) is not obligate for growth. Utilized carbon substrates like D-glucose, pyruvate, sucrose, D-galactose, rhamnose, mannose, inositol, maltose, lactose, fumarate, trehalose and sorbitol. Fructose, starch, acetate, Na-propionate, D-xylose, malic acid, mannitol, ascorbate, succinate, inulin, citrate and benzoic acid are not utilised. Utilized nitrogen sources like Ammonium sulphate, yeast extract, peptone, sodium nitrate, casamino acid, L-cysteine, L-histidine, L-glutamic acid, L-methionine, DL-alanine, L-arginine, L-glycine, L-proline, L-isoleucine, L-ornithine, L-glutamine, and DL-threonine. L-serine, L-valine L-tyrosine, L-tryptophan, L-aspartic acid, and Urea are not utilised as nitrogen sources. Hydrolyse phytagel. The predominant fatty acids include $C_{18:1}\omega9c$, $C_{18:0}$ and $C_{16:0}$ with minor amount of $C_{18:3}\omega6c,9c,12c$, anteiso-$C_{17:0,}$ anteiso-$C_{17:0}$, $C_{17:1}\omega8c$, $C_{17:0}$, $C_{15:1}\omega5c$, $C_{15:2}$OH, anteiso-$C_{15:0}$, $C_{13:0}$, $C_{I4:0}$, anteiso-$C_{12:0}$ and anteiso-$C_{11:0}$. Major polyamines includes putrescine and sym- homospermidine. The major polar lipids are phosphatidylcholine, two unidentified glycolipids (GL1, 2), two unidentified phospholipid (PL1, 2) and seven unidentified lipids (UL1-7). The predominant quinone is MK6 and nitrate is not reduced. API ZYM shows positive for leucine arylamidase, esterase (C4), Naphthol-AS-BI-phosphohydrolase and valine arylamidase. Negative for lipase (C14), lipase (C8), trypsin, α-chymotrypsin, cysteine arylamidase, β-glucuronidase, α-galactosidase, α-glucosidase, α-mannosidase, α-fucosidase and β- glucosidase, alkaline phosphatase, β-galactosidase, esterase acid phosphatase and N-acetyl- β-glucosaminidase.

The type strain JC665^T (= NBRC 114305 = KCTC 72671^T) was isolated from the rhizosphere soil of *Erianthus ravennae* (commonly known as “Plume grass”) collected from Loktak lake located in the Northeast part of India, Manipur (exact location:

24°30’21” N 93°47’43” E). JC747 is an additional strain isolated from a wetland located (village: Pallikkara) in the southwest part of India, Kerala (12° 23’ 02” N 75° 02’ 33” E). The GenBank accession numbers of the 16S rRNA gene sequence and genome sequence of strain JC65^T and JC747 are LR746340, OU374731 and JAALCR000000000 and JAHPZK000000000, respectively.

4.1.2.2. Proposal of *Sinomonas cellulolyticus* sp. nov.

The genome-based and phylogenetic delineation of strain JC656^T from its closest phylogenetic neighbour, *S. notoginsengisoli* KCTC 29237^T is well demonstrated with the differences phylogenetic, chemotaxonomic, phenotypic and genotypic properties presented in this study (Fig. 36, Fig. 37, Table 21, Table 22, Table 23). Therefore, we suggest the placement of strain JC656^T as a novel species in the genus *Sinomonas*, for which we propose the name *Sinomonas cellulolyticus* sp. nov.

4.1.2.2.1 Description of *Sinomonas cellulolyticus* sp. nov.

Sinomonas cellulolyticus (cel.lu.lo. ly’ti.cus. M. L. n. *cellulosum*, cellulose; Gr. adj. *lyticus*, dissolving; M. L. adj. *cellulolyticus*, decomposing cellulose)

Cells are gram-positive, aerobic, non-motile. Colonies are yellow coloured, circular, convex and 2.0-3.0 mm in diameter after 2 days cultivation at 28 ˚C on nutrient agar. Optimum temperature and pH for growth are 28 ˚C (range 14-37 ˚C) and 7.0 (6.0-8.0). NaCl is not essential for growth (tolerate up-to 6 % v/w). Hydrolysed tween 40 and cellulose while tween 80, starch, gelatin, casein are not hydrolysed. Catalase, Vogues-Proskauer are positive, urease, nitrate reduction are negative. Glucose, inositol, citric acid, D- galactose, D-mannitol, D-lactose, D-lactose, D-raffinose, D-maltose are utilised as carbon source. L-glutamic acid, L-alanine, L-aspartic acid are utilised as nitrogen source.

Major fatty acids are anteiso-C15:0, iso-C16:0, anteiso-C17:0. Major cell wall amino acids are alanine, glutamine, lysine, DAP and two unidentified amino acids. Menaquinone-9 (H2) is the predominant respiratory quinone and the polar lipid consists of diphosphatidylglycerol, phosphatidylinositol, phosphatidylglycerol, phosphatidylmonomethylethanolamine, glycolipid (GL1- GL5) and unidentified lipid. The genome size is 3.9 Mb with genomic DNA G + C content of 69.9 %.

The type strain JC656^{T} (=KCTC49339 = NBRC114142) was isolated from the soil sediment of floating island (Phumdis) of Loktak lake, India. The GenBank accession number for the 16S rRNA gene sequence of strain JC656^{T} is ON908988 and the genome sequence has been deposited in GenBank under the accession number JAERRC010000000.

4.1.2.3. Proposal of *Glutamicibacter rhizosphaerae* sp. nov.

The phylogenomic-based species delineation of strain JC586^{T} shows a clear distinction with *Glutamicibacter mishrai* based on genotypic, phenotypic and chemotaxonomic analysis justifying the creation of new species of the genus *Glutamicibacter* (Fig. 45, Fig. 46, Fig. 47, Table 25, Table 27, Table 29). As a result, we propose the name *Glutamicibacter rhizosphaerae* sp. nov. for strain JC586^{T} in the genus *Glutamicibacter*.

4.1.2.3.1. Description of *Glutamicibacter rhizosphaerae* sp. nov.

Glutamicibacter rhizosphaerae (rhi.zo. sphae'rae. Gr. n. *rhiza*, root; L. n. *sphaera*, sphere; N.L. gen. n. *rhizosphaerae*, from the rhizosphere)

Colonies are light yellow, strictly aerobic, gram stain positive, rod-coccus growth cycle, non-sporulating, up to 1.0 mm in diameter. Optimum pH and temperature for growth

occurs at 7 pH (5-11 pH) and 28 ˚C (16-27 ˚C). NaCl is required for growth with optimum at 2-4 % (range 0-11 %). Positive for oxidase, catalase, citrate, gelatin, starch and negative for urease, cellulose, indole, Voges Proskauer, casein and cellulase. Utilises D-glucose, D-rhamnose, D-mannose, inositol, D-cellubiose whereas D-arabinose, L-arabinose, D-sorbitol, D-lactose, D-melibiose, D-xylose, D-galactose does not support the growth. API enzyme kit analysis gives positive results for easterase, valine acrylamidase, easterase lipase, acid phosphatase, α-galactosodase, β-glucosidase, β-galactosidase, α-mannosidase and negative for lipase, cystine acrylamidase. The predominant fatty acids are anteiso-$C_{15:0}$, iso-$C_{15:0}$, anteiso-$C_{17:0}$. Major respiratory quinone include MK-8 and MK-9. Lysine, alanine, glutamine and DAP were detected as the cell wall's peptidoglycan composition. Major polar lipids were DPG: diphosphatidylglycerol, phosphatidylglycerol, glycolipid, dimannosylglyceride, trimannosyldiacylglycerol. The genomic DNA G+C content of the type strain is 60.4 mol %.

The type strain, JC586^T (=NBRC 113657= KACC 21453) was isolated from the rhizosphere soil of *Alocasia cucullate* (commonly called as Budhha's palm) collected from Loktak lake located in Manipur, India (exact location: 24°30'21" N 93°47'43" E). The 16S rRNA gene sequence and the whole genome sequence are deposited to GenBank under the accession number ON908987 and NZ_VHIN00000000.

4.1.2.4. Proposal of "*Candidatus* Janthinobacterium sediminis" sp. nov.

The phylogenomic based species delineation of strain JC611^T from *Janthinobacterium rivuli* (Fig. 56) is well supported by the differences in phenotypic characters such as colony and cell morphology, pH and salt tolerance, temperature, organic carbon substrate utilisation and respiratory quinone (Table 30). Based on the phylogenomic inference supported by the differences in morphology (Fig. 59) and

genomic (Table 31), we propose a new species in the genus *Janthinobacterium*. However, due to Covid-19 pandemic lockdown, the cultures had to be preserved under refrigeration, and they could not be revived once the pandemic was over. Hence, we propose to describe the new strain giving the status of "*Candidatus*" as "*Candidatus* Janthinobacterium sediminis" sp. nov.

4.1.2.4.1. Description of "*Candidatus* Janthinobacterium sediminis" sp. nov.

"*Ca*. Janthinobacterium sediminis" (se.di'mi.nis. L. gen. n. *sediminis*, of sediment)

Colonies are violet coloured, round, convex and pigmented. Cells are gram- stain-negative, motile and aerobic. Cultures lose viability on refrigeration. Non-viable cells are rod-shaped with 0.5-0.7 µm in length and 0.3-0.5 µm in width. Optimum growth occurs at temperature range between 16-28 ˚C, pH range between 6-9, and NaCl range between 1-2 %. Shows positive for oxidase, catalase, urease activities and negative for starch and gelatin hydrolysis. *In-silico* predictions indicate the fermentation capability of the novel species, utilise substrates like lactate, degrade aromatic hydrocarbons like benzoate, catechol, phenylacetate, capability to utilise urea and reduce nitrate. The genome size of the strain is 6.18 Mb and the DNA G+C content is 62.7 mol %.

The GenBank accession number of 16S rRNA gene sequence is ON908989 and the draft genome has been deposited in GenBank under the accession number JAFLJN000000000.

4.1.2.5. Proposal of *Comamonas soli* sp. nov.

The phylogenetic analysis based on 16S rRNA gene sequence revealed that strain JC664^T was affiliated with the genus *Comamonas* with the highest sequence similarity to *Comamonas koreensis* 12005^T (98.5 %). Based on the phylogenetic inference (Fig. 62) supported by the differences in phenotypic traits (morphology, temperature, salt tolerance,

carbon/nitrogen substrate utilisation, API enzyme activity (Table 32), fatty acid composition (Table 33) and polar lipids (Fig. 63), we propose strain JC664^T as a putative new species in the genus *Comamonas*.

4.1.2.5.1. Description of *Comamonas soli* sp. nov.

Comamonas soli (so'li. L. neut. gen. n. *soli* of soil, referring to the isolation of the type strain from a soil sample)

Cells are Gam-stain-negative, non-motile, rod-shaped and multiply by binary fission. Colonies are translucent pale yellow. NaCl is not required for growth but can tolerate up to 3 %. Optimum growth occurs at temperature 28 ˚C (18-37 ˚C) and pH 7 (pH 6-7). Positive for catalase and oxidase. Starch, casein and tween 80 were hydrolyzed, while urea, tween 20, and gelatin were not hydrolyzed. Glucose, sodium propionate, citrate, glutamic acid, threonine, phenylalanine were utilized for growth, while mannitol, rhamnose, sorbitol, sucrose, fructose, inositol, leucine and serine were not utilized. Sensitive to chloramphenicol and nalidixic acid and resistant to ampicillin and rifampicin. Positive for esterase (C4), esterase lipase (C8), α-glucosidase, β-galactosidase, β-glucosidase and negative for alkaline phosphatase, acid phosphatase, α-mannosidase, *N*-acetyl-β-glucosaminidase and α-fucosidase. $C_{16:1}\omega6c$ / $C_{16:1}\omega7c$, $C_{18:1}\omega6c$ /$C_{18:1}\omega7c$, $C_{17:0}$ cyclo and $C_{16:0}$ are the major fatty acids. Diphosphatidylglycerol, phosphatidylglycerol and phosphatidylethanolamine are the major polar lipids with minor amount of unidentified phospholipid and unidentified lipid (L1). Q-8 is the major isoprenoid quinone.

The type strain, JC664^T (=KCTC 72670 = NBR C114120), was isolated from the rhizosphere soil of *Colocasia esculenta* of Loktak lake, India. The accession number of GenBank for the 16S rRNA gene sequence is ON908990.

4.2. Future perspectives and inferences of Loktak lake

A thorough report on the insight into the bacterial community structure of Loktak lake is described in this study. The findings of this study significantly add to our present understanding of the microbiology of Loktak lake and highlight the abundance of different bacterial taxa present in and around the lake. This study is particularly significant not just from the perspective of microbial taxonomy but also owing to social and economic aspects related to the lake, as it has become a popular tourist destination over the past few decades. To date, the metagenomic work on Loktak lake bacterial taxa has been done at the phylum level which has been found to be useful in assessing the types and abundance of bacteria that may present in the lake. With the advent of environmental DNA approaches, information on the diversity of microorganisms, their coexistence with higher organisms and identification of their novel aspects can be revealed (Vijayan et al. 2023).

Future environmental monitoring of freshwater ecosystems should take into account the microbial community dynamics as they are the missing link in comprehending the rapid changes in the aquatic ecosystem structure and function (Sagova-Mareckova et al. 2021). Loktak lake serves as a source of potable water, a tourist spot, and an opportunity for recreational activities. However, it is crucial to constantly evaluate the lake's biodiversity and water quality without compromising its sustainability. Increased anthropogenic activities such as sewage water runoff near the catchment area, faecal discharge, and the confluence of different waste streams in the lake might lead to the alteration of their microbial community structure, which leads to the prevalence of phyla such as *Pseudomonadota*, *Acidobacteriota*, *Actinomycetota*, *Verrucomicrobiota*, *Bacillota*. The results of this study might serve as a database for future research for the exploitation of microbial taxa, community structures and diversity and also to characterise

the changing pattern of lake biodiversity (Kumar et al. 2022). Future research should also focus on analysing bacterial communities from various lake ecosystem involving different physicochemical parameters, such as pH, temperatures, salinity, pressure and osmotic potential in order to understand the role of microbial community in nutrient cycling, which is a largely untapped reservoir of resources for innovative applications useful to mankind (Yadav et al. 2018). In addition, metagenomic and metatranscriptomics analyses should also be performed to evaluate the changes in gene expression and microbial communities in water and the host microbiome. Moreover, comparing the metagenomic data of Loktak lake with several ecosystems may shed light on microbial interactions and their functional attributes of this habitat. Lastly, our understanding of lake bacterial composition and diversity is still superficial and limited which can be overcome by using phylogenetic framework and employ next-generation sequencing technique which aimed at learning more about the interaction between bacterial phylotypes and taxa, preferred habitats and their functions (Ren et al. 2013).

MAJOR FINDINGS

5. MAJOR FINDINGS

- The study forms the base data on the census of microbiomes of Loktak Lake. From this study we have identified a large number of new phylotypes, a few of these might be endemic to Loktak Lake.
- LOK4 has the highest species richness and bacterial diversity as it falls under the protected and undisturbed area of Keibul Lamjao National Park (KLNP) which is predicted to harbour rich biodiversity in terms of flora and fauna.
- This enormous number of rare, uncultured and unexplored taxa clearly indicates the hidden wealth of Loktak Lake.
- A note on the future perspectives of Loktak lake is proposed in light of the experienced acquired along with the rapid advancement in the field of bacterial taxonomy.
- Phylogenomic including polyphasic taxonomic studies help in establishing 5 new bacterial taxa:
    1. *Paludisphaera rhizosphaerae* sp. nov. JC665^{T}
    2. *Glutamicibacter rhizosphaerae* sp. nov. JC586^{T}
    3. *Sinomonas cellulolyticus* sp. nov. JC656^{T}
    4. ‘‘*Candidatus* Janthinobacterium sediminis’’ sp. nov. JC611^{T}
    5. *Comamonas soli* sp. nov. JC664^{T}

REFERENCES

6. REFERENCES

Abby, S. S., Kazemzadeh, K., Vragniau, C., Pelosi, L., and Pierrel, F. (2020). Advances in bacterial pathways for the biosynthesis of ubiquinone. *Biochim. Biophys. Acta - Bioenerg.* 1861: 148259. doi: 10.1016/j.bbabio.2020.148259

Abram, K., Udaondo, Z., Bleker, C., Wanchai, V., Wassenaar, T.M., Robeson, M.S., and Ussery, D.W. (2021). Mash-based analyses of Escherichia coli genomes reveal 14 distinct phylogroups. *Commun Biol* 26: 117. doi: 10.1038/s42003-020-01626-5

Abril, J.F. and Castellano Hereza, S. (2018). Genome Annotation. In: Ranganathan, S., Gribskov, M., Nakai, K., and Schönbach, C. (eds.) Encyclopedia of Bioinformatics and Computational Biology. 2:195-209. Elsevier: Amsterdam, Netherlands. doi: 10.1016/B978-0-12-809633-8.20226-4

Ahmad, T., Gupta, G., Sharma, A., Kaur, B., El-Sheikh, M. A., and Alyemeni, M. N. (2021). Metagenomic analysis exploring taxonomic and functional diversity of bacterial communities of a Himalayan urban fresh water lake. *PloS one*. 16: e0248116. doi: 10.1371/journal.pone.0248116

Ahmad, T., Singh, R.S., Gupta, G., Sharma, A., and Kaur, B. (2019). Metagenomics in the Search for Industrial. *Adv. Enzyme res. Technol*: Elsevier. pp. 419-451. doi: 10.1016/B978-0-444-64114-4.00015-7

Aliyu, H., Lebre, P., Blom, J., Cowan, D., and De Maayer, P. (2016). Phylogenomic re-assessment of the thermophilic genus *Geobacillus*. *Syst. Appl. Microbiol.* 39: 527-533. doi: 10.1016/j.syapm.2016.09.004

Amrane, S., and Lagier, J. C. (2018). Metagenomic and clinical microbiology. *Hum. Microbiome J.* 9: 1-6. doi: 10.1016/j.humic.2018.06.001

Anto, L., Warykas, S. W., Torres-Gonzalez, M., and Blesso, C. N. (2020). Milk polar lipids: underappreciated lipids with emerging health benefits. *Nutrients*. 12: 1001. doi: 10.3390/nu12041001

Auch, A.F., von Jan, M., Klenk, H-P., and Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genom. Sci*. 2:117-34. doi: 10.4056/sigs.531120

Bahr, M., Hobbie, J. E., and Sogin, M. L. (1996). Bacterial diversity in an arctic lake: a freshwater SAR11 cluster. *Aquat. Microb. Ecol*. 11: 271-277. doi:10.3354/ame011271

Bale, N.J., Sorokin, D.Y., Hopmans, E.C., Koenen, M., Rijpstra, W.I.C., Villanueva, L., Wienk, H. and Sinninghe Damsté, J.S. (2019). New insights into the polar lipid composition of extremely halo (alkali) philic euryarchaea from hypersaline lakes. *Front. Microbiol*. 12:10-377. doi: 10.3389/fmicb.2019.00377

Bao, Y. Y., Huang, Z., Mao, D. M., Sheng, X. F., and He, L. Y. (2015). *Sinomonas susongensis* sp. nov., isolated from the surface of weathered biotite. *Int. J. Syst. Evol. Microbiol*. 65: 1133-1137. doi: 10.1099/ijs.0.000064

Berlin, K., Koren, S., Chin, C.S., Drake, J.P., Landolin, J.M., and Phillippy, A.M. (2015). Assembling large genomes with single-molecule sequencing and locality sensitive hashing. *Nat. Biotechnol*. 33: 623-630. doi: 10.1038/nbt.3238

Blevins, S. M., and Bronze, M. S. (2010). Robert Koch and the 'golden age' of bacteriology. *Int. J. Infect. Dis*. 14: 744-751. doi: 10.1016/j.ijid.2009.12.003

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee, S.Y., Medema, M. H., and Weber,T. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res*. 47: W81-W87. doi: 10.1093/nar/gkz310

Boburque, G., Burns, K.H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H.L., Macfarlan, T.S., and Mager, D.L. (2018). Ten things you should know about transposable elements. *Genome Biol.* 19 :1-12. doi: 10.1186/s13059-018-1577-z

Bosshard, P. P., Santini, Y., Grüter, D., Stettler, R., and Bachofen, R. (2000). Bacterial diversity and community composition in the chemocline of the meromictic alpine Lake Cadagno as revealed by 16S rDNA analysis. *FEMS Microbiol.Ecol. 31*: 173-182. doi: 10.1111/j.1574-6941.2000.tb00682.x

Botstein, D., Cherry, J.M., Ashburner, M., Ball, C.A., Blake, J.A., Butler, H., Davis, A.P., Dolinski, K., Dwight, S.S., and Eppig, J.T. (2000). Gene Ontology: Tool for the unification of biology. *Nat. Genet.*25: 25-9. doi: 10.1038/75556

Brooksbank, C., Bergman, M.T., Apweiler, R., Birney, E., and Thornton, J. (2014). The european bioinformatics institute's data resources 2014. *Nucleic Acids Res*. 42: D18-D25. doi: 10.1093/nar/gkt1206

Brown, M. V., Philip, G. K., Bunge, J. A., Smith, M. C., Bissett, A., Lauro, F. M., Fuhrman, J.A. and Donachie, S.P. (2009). Microbial community structure in the North Pacific Ocean. *ISME. J.* 3: 1374-1386. doi: 10.1038/ismej.2009.86

Busse H. J. (2016). Review of the taxonomy of the genus Arthrobacter, emendation of the genus *Arthrobacter* sensu lato, proposal to reclassify selected species of the genus *Arthrobacter* in the novel genera *Glutamicibacter* gen. nov., *Paeniglutamicibacter* gen. nov., *Pseudoglutamicibacter* gen. nov., *Paenarthrobacter* gen. nov. and *Pseudarthrobacter* gen. nov., and emended description of *Arthrobacter roseus*. *Int. J. Syst. Evol. Microbiol*. 66: 9-37. doi: 10.1099/ijsem.0.000702

Callahan, B.J., Sankaran, K., Fukuyama, J.A., McMurdie, P.J., Holmes, S.P. (2016). Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. *F1000Res*. 5:1492. doi: 10.12688/f1000research.8986.2

Caporaso, J.G, Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C. A., McDonald, D., Muegge, B.D., Pirrung ,M.,

Reeder, J., Sevinsky J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods.* 7: 335-6. doi: 10.1038/nmeth.f.303

Carro, L., Peix, Á., and Velázquez, E. (2021). The Taxonomy of Bacteria in the Genomic Era. In: Villa, T.G., de Miguel Bouzas, T. (eds) Developmental Biology in Prokaryotes and Lower Eukaryotes. Springer, Cham. 289-309. doi. 10.1007/978-3-030-77595-7_12

Castresana, J. (2007). Topological variation in single-gene phylogenetic trees. *Genome Biol.* 8: 1- 4. doi: 10.1186/gb-2007-8-6-216

Chakraborty, J., Sapkale, V., Shah, M., Rajput, V., Mehetre, G., Agawane, S., Kamble, S. and Dharne, M. (2020). Metagenome sequencing to unveil microbial community composition and prevalence of antibiotic and metal resistance genes in hypersaline and hyperalkaline Lonar Lake, India. *Ecol. Indics*. 110: 105827. doi: 10.1016/j.ecolind.2019.105827

Chan, J. Z., Halachev, M. R., Loman, N. J., Constantinidou, C., and Pallen, M. J. (2012). Defining bacterial species in the genomic era: insights from the genus Acinetobacter. *BMC Microbiol.* 12:1-11. doi: 10.1186/1471-2180-12-302

Chandran, A., Varghese, S., Kandeler, E., Thomas, A., Hatha, M., and Mazumder, A. (2011). An assessment of potential public health risk associated with the extended survival of indicator and pathogenic bacteria in freshwater lake sediments. *Int. J. Hyg. Environ. Health*. 214: 258-264. doi: 10.1016/j.ijheh.2011.01.002

Chaudhari, N. M., Gupta, V. K., and Dutta, C. (2016). BPGA-an ultra-fast pan-genome analysis pipeline. *Sci. Rep.* 6: 1-10. doi: 10.1038/srep24373

Chen, X., Zhang, Y., Zhang, Z., Zhao, Y., Sun, C., Yang, M., Wang, J., Liu, Q., Zhang, B., Chen, M. and Yu, J. (2018). PGAweb: a web server for bacterial pan-genome analysis. *Front. Microbiol.* 9: 1910. doi: 10.3389/fmicb.2018.01910

Collins, M. D., and D. Jones. (1981). Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implications. *Microbiol. Rev.* 45:316–354. doi: 10.1128/mr.45.2.316-354.1981

Colwell, R.R. (1970). Polyphasic taxonomy of the genus Vibrio: numerical taxonomy of Vibrio cholerae, Vibrio parahaemolyticus and related Vibrio species. J Bacteriol. 104:410–433. doi: 10.1128/jb.104.1.410-433.1970

Consortium, U. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res*. 2019, 47, D506–D515. doi: 10.1093/nar/gky1049

Costa, S.S., Guimarães, L.C., Silva, A., Soares, S.C., and Baraúna, R.A. (2020). First Steps in the Analysis of Prokaryotic Pan-Genomes. *Bioinform. Biol. Insights*.14. doi:10.1177/1177932220938064

Coughlan, L. M., Cotter, P. D., Hill, C., and Alvarez-Ordóñez, A. (2015). Biotechnological applications of functional metagenomics in the food and pharmaceutical industries. *Front Microbiol*. 6: 672. doi: 10.3389/fmicb.2015.00672

Cowan, D. A., Ramond, J. B., Makhalanyane, T. P., and De Maayer, P. (2015). Metagenomics of extreme environments. *Curr. Opin. Microbiol.* 25: 97-102. doi: 10.1016/j.mib.2015.05.005

Cowan, S. T., and Steel, K. J. (1965). Manual for the identification of medical bacteria. *Manual for the identification of medical bacteria.* pp. 217.

Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., and Kirkegaard, R.H. (2015). Complete nitrification by Nitrospira bacteria. *Nature*. 528: 504-509. doi: 10.1038/nature16461

Danczak, R.E., Johnston, M. D., Kenah, C., Slattery, M., Wrighton, K.C., and Wilkins, M.J. (2017). Members of the Candidate Phyla Radiation are functionally differentiated by carbon- and nitrogen-cycling capabilities. Microbiome. 5:112. doi: 10.1186/s40168-017-0331-1

Darling, A. E., Jospin, G., Lowe, E., Matsen, F. A. 4th, Bik, H. M., and Eisen, J. A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. *Peer J.* 2: e243. doi: 10.7717/peerj.243

Das, S., Bora, S. S., Yadav, R. N. S., and Barooah, M. (2017). A metagenomic approach to decipher the indigenous microbial communities of arsenic contaminated groundwater of Assam. *Genom. Data*. 12: 89-96. doi: 10.1016/j.gdata.2017.03.013

Das, L., Deb, S. and Das, S.K. (2019). *Glutamicibacter mishrai* sp. nov., isolated from the coral *Favia veroni* from Andaman Sea. *Arch Microbiol* 202: 733-745. doi: 10.1007/s00203-019-01783-0

De Carvalho, C. C. C. R., Caramujo, M. J. (2018). The Various Roles of Fatty Acids. *Molecules*. 23: 2583. doi: 10.3390/molecules23102583

De Wever, A., Muylaert, K., Van der Gucht, K., Pirlot, S., Cocquyt, C., Descy, J. P., Plisnier, P. D., and Vyverman, W. (2005). Bacterial community composition in Lake Tanganyika: vertical and horizontal heterogeneity. *Appl. Environ. Microbiol.* 71: 5029-5037. doi: 10.1128/AEM.71.9.5029-5037.2005

Delsuc, F., Brinkmann, H., and Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet*. 6: 361-375. doi: 10.1038/nrg1603

Demko, A. M., Patin, N. V., and Jensen, P. R. (2021). Microbial diversity in tropical marine sediments assessed using culture-dependent and culture-independent techniques. *Environ. Microbiol*. 23: 6859-6875. doi: 10.1111/1462-2920.15798

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol*. 72: 5069-5072. doi: 10.1128/AEM.03006-05

Devi, R. S., Satapathy, K. B., and Kumar, S. (2022). Ethnobotanical Plants of Phumdi, Loktak Lake, Manipur, India.

Dhal, P.K., Kopprio, G.A. and Gärdes, A. (2020). Insights on aquatic microbiome of the Indian Sundarbans mangrove areas. *PLoS One*. 1: p.e0221543. doi: 10.1371/journal.pone.0221543

Insights on aquatic microbiome of the diverse ecological niches in northeast India. (Ed. By Soni, R., Suyal, DC. And Yadav, A.N. doi: 10.1016/j.syapm.2008.07.002.

Ding L., Hirose T., Yokota A. (2009). Four novel *Arthrobacter* species isolated from filtration substrate. *Int J Syst Evol Microbiol*. 59: 856-862. doi: 10.1099/ijs.0.65301-0

Dugail, I., Kayser, B.D., and Lhomme, M. (2017). Specific roles of phosphatidylglycerols in hosts and microbes. *Biochimie* 141:47-53. doi: 10.1016/j.biochi.2017.05.005

Eisen, J. A., and Fraser, C. M. (2003). Phylogenomics: intersection of evolution and genomics. *Science*. 300: 1706-1707. doi: 10.1126/science.1086292

Ejigu, G. F., and Jung, J. (2020). Review on the Computational Genome Annotation of Sequences Obtained by Next-Generation Sequencing. *Biology*. 9: 295. doi: 10.3390/biology9090295

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evol*. 39: 783-791. doi: 10.2307/2408678

Feng, W. W., Wang, T. T., Bai, J. L., Ding, P., Xing, K., Jiang, J. H., Peng, X., and Qin, S. (2017). *Glutamicibacter halophytocola* sp. nov., an endophytic actinomycete isolated from the roots of a coastal halophyte, Limonium sinense. *Int. J. Syst. Evol. Microbiol.* 67: 1120-1125. doi: 10.1099/ijsem.0.001775

Ferraz Helene, L. C., Klepa, M. S., and Hungria, M. (2022). New insights into the taxonomy of bacteria in the genomic era and a case study with rhizobia. *Int. J. Microbiol.* floating islands and their behavioural changes in Loktak Lake with respect to biodiversity. doi: 10.1155/2022/4623713

Friedrich, I., Hollensteiner, J., Schneider, D., Poehlein, A., Hertel, R., and Daniel, R. (2020). First Complete Genome Sequences of *Janthinobacterium lividum* EIF1 and EIF2 and Their Comparative Genome Analysis. *Genome Biol. Evo1*. 2: 1782–1788. doi: 10.1093/gbe/evaa148

Fu, Y., Yan, R., Liu, D., Zhao, J., Song, J., Wang, X., Cui, L., Zhang, J. and Xiang, W. (2019). Characterization of *Sinomonas gamaensis* sp. nov., a Novel Soil Bacterium with Antifungal Activity against *Exserohilum turcicum. Microorganisms*. 7:170. doi: 10.3390/microorganisms7060170

Garrity G. M. (2016). A New Genomics-Driven Taxonomy of Bacteria and Archaea: Are We There Yet? *J. Clin. Microbiol*. 54: 1956-1963. doi: 10.1128/JCM.00200-16

Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., de Peer, Y.V., Vandamme, P., Thompson, F.L., and Swings, J. (2005). Re-evaluating prokaryotic species. *Nat. Rev. Microbiol*. 3: 733-739. doi: 10.1038/nrmicro1236

Ghai R, Rodŕíguez-Valera F, McMahon KD, Toyama D, Rinke R, Cristina Souza de Oliveira, T., Wagner Garcia, J., Pellon de Miranda, F. and Henrique-Silva, F. (2011). Metagenomics of the water column in the pristine upper course of the Amazon River. PloS one. 6: e23785. doi: 10.1371/journal.pone.0023785

Ghosh, S., Sarangi, A. N., Mukherjee, M., Singh, D., Madhavi, M., and Tripathy, S. (2021). Reconstructing Draft Genomes Using Genome Resolved Metagenomics Reveal Arsenic Metabolizing Genes and Secondary Metabolites in Fresh Water Lake in Eastern India. *Bioinform. Biol. Insights*. 15. doi: 10.1177/11779322211025332

Ghuysen, J. M. (1968). Use of bacteriolytic enzymes in determination of wall structure and their role in cell metabolism. *Bacteriol. Rev*. 32: 425-464.

Giri, K., Doley, B., Mishra, G., Suyal, D. C., Baruah, R. C., and Jayaraj, R. S. C. (2022). An *Applied Microbiology for Sustainable Economy*, In. An overview of microbial diversity under diverse ecological niches in northeast India. (Ed. By Soni, R., Suyal, DC. And Yadav, A.N. and Goel, R), Academic press pp. 65-108. doi: 10.1016/B978-0-323-91595-3.00013-6

Glöckner, F. O., Zaichikov, E., Belkova, N., Denissova, L., Pernthaler, J., Pernthaler, A., and Amann, R. (2000). Comparative 16S rRNA analysis of lake bacterioplankton reveals globally distributed phylogenetic clusters including an abundant group of actinobacteria. *Appl. Environ. Microbiol*. 66: 5053-5065. doi: 10.1128/AEM.66.11.5053-5065.2000

Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T, Vandamme, P., Tiedje, J.M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol*. 57:81-91. doi: 10.1099/ijs.0.64483-0

Gong, X., Skrivergaard, S., Korsgaard, B. S., Schreiber, L., Marshall, I. P., Finster, K., and Schramm, A. (2017). High quality draft genome sequence of *Janthinobacterium psychrotolerans* sp. nov., isolated from a frozen freshwater pond. *Stand. Genom. Sci*. 12, 1-7. doi:10.1186/s40793-017-0230-x

Govil, T., Paste, M., Samanta, D., David, A., Goh, K. M., Li, X., Salem, D. R., and Sani, R. K. (2021). Metagenomics and Culture Dependent Insights into the Distribution of Firmicutes across Two Different Sample Types Located in the Black Hills Region of South Dakota, USA. *Microorganisms* 9: 113. doi: 10.3390/microorganisms9010113

Gray, T. N. E., Brook, S. M., McShea, W. J., Mahood, S., Ranjitsingh, M. K., Miyunt, A., Hussain, S. A., and Timmins, R. (2015). *Rucervus eldii*. The IUCN Red List of Threatened Species. Version 2015.2. www.iucnredlist.org.

Guan, Y., Hou, T., Li, X., Feng, L., and Wang, Z. (2022). Metagenomic insights into comparative study of nitrogen metabolic potential and microbial community between primitive and urban river sediments. *Environ. Res.* 212:113592. doi: 10.1016/j.envres.2022.113592

Guo, Q.Q., Ming, H., Meng, X.L., Huang, J.R., Duan, Y.Y., Li, S.H., Li, S., Zhang, J.X., Li, W.J., and Nie, G.X. (2015). *Sinomonas halotolerans* sp. nov., an actinobacterium isolated from a soil sample. *Antonie. Leeuwenhoek. J. Microbiol.* 10: 887-895. doi: 10.1007/s10482-015-0543-y.

Haley, B. J., Grim, C. J., Hasan, N. A., Choi, S. Y., Chun, J., Brettin, T. S., Bruce, D.C., Challacombe, J.F., Detter, J.C., Han, C.S. and Huq, A. (2010). Comparative genomic analysis reveals evidence of two novel Vibrio species closely related to V. cholerae. *BMC Microbiol.* 10: 1-10. doi: 10.1186/1471-2180-10-154

Hamana, K., and Matsuzaki, S. (1992) Polyamines as a chemotaxonomic marker in bacterial systematics. *Crit. Rev. Microbiol*. 18, 261–283. doi: 10.3109/10408419209113518

Handelsman, J., Rondon, M. R., Brady, S.F., Clardy, J., Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5: R245-9. doi: 10.1016/s1074-5521(98)90108-9

Heino, J., Alahuhta, J., Bini, L.M., Cai, Y., Heiskanen, A.S., Hellsten, S., Kortelainen, P., Kotamäki, N., Tolonen, K.T., Vihervaara, P., Vilmi. A., and Angeler, D.G. (2021). Lakes in the era of global change: Moving beyond single-lake thinking in maintaining biodiversity and ecosystem services. *Biol. Rev.* 96: 89-106. doi: 10.1111/brv.12647

Henz, S., Huson, D., Auch, A.F., Nieselt-Struwe, K., and Schuster, S. (2005). Whole-genome prokaryotic phylogeny. *Bioinformatics*. 21: 2329-2335. doi:10.1093/bioinformatics/bth324

Hiorns, W. D., Methé, B. A., Nierzwicki-Bauer, S. A., and Zehr, J. P. (1997). Bacterial diversity in Adirondack Mountain lakes as revealed by 16S rRNA gene sequences. *Appl. Environ. Microbiol. 63*: 2957-2960. doi: 10.1128/aem.63.7.2957-2960.1997

Hoff, K.J., Lange, S., Lomsadze, A., Borodovsky, M., and Stanke, M. (2016). BRAKER1: Unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*. 32: 767-769. doi: 10.1093/bioinformatics/btv661

Holt, C., and Yandell, M. (2011). MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinform*. 12: 491. doi: 10.1186/1471-2105-12-491

Hugenholtz, P., Chuvochina, M., Oren, A., Parks, D. H., and Soo, R. M. (2021). Prokaryotic taxonomy and nomenclature in the age of big sequence data. *the ISME J*. 15. 1879-1892. doi: 10.1038/s41396-021-00941-x

Jiang, P., Green, S.J., Chlipala, G.E., Turek, F.W. and Vitaterna, M.H. (2019). Reproducible changes in the gut microbiome suggest a shift in microbial and host metabolism during spaceflight. *Microbiome* **7**: 113. doi: 10.1186/s40168-019-0724-4

Johnson, J.S., Spakowicz, D.J., Hong, B.Y., Petersen, L.M., Demkowicz, P., Chen. L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., and Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*. 10: 5029. doi: 10.1038/s41467-019-13036-1

Jones, R. T., Robeson, M. S., Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME. J.* 3: 442-453. doi: 10.1038/ismej.2008.127

Jung, W. J., Kim, S. W., Giri, S. S., Kim, H. J., Kim, S. G., Kang, J. W., Kwon, J., Lee, S. B., Oh, W. T., Jun, J. W., and Park, S. C. (2021). *Janthinobacterium tructae* sp. nov., Isolated from Kidney of Rainbow Trout (*Oncorhynchus mykiss*). *Pathogens (Basel, Switzerland)*, 10: 229. https: 10.3390/pathogens10020229

Kämpfer, P., and Glaeser, S. P. (2012). Prokaryotic taxonomy in the sequencing era--the polyphasic approach revisited. *Environ. Microbiol*. 14: 291-317. doi: 10.1111/j.1462-2920.2011.02615.x

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28: 27-30. doi: 10.1093/nar/28.1.27

Kangabam, R. D., Boominathan, S. D., and Govindaraju, M. (2015). Ecology, disturbance and restoration of Loktak Lake in Indo-Burma Biodiversity Hotspot-an overview. *NeBio: j. environ. biodivers*. 6: 9-15.

Kangabam, R. D., Bhoominathan, S. D., Kanagaraj, S., and Govindaraju, M. (2017). Development of a water quality index (WQI) for the Loktak Lake in India. *Appl.Water Sci.* 7: 2907-2918. doi: 10.1007/s13201-017-0579-4

Kangabam, R. D., Selvaraj, M., and Govindaraju, M. (2018). Spatio-temporal analysis of floating islands and their behavioral changes in Loktak Lake with respect to biodiversity using remote sensing and GIS techniques. *Environ. Monit. Assess*. 190: 1-14. doi:10.1007/s10661-018-6485-x

Kangabam, R. D., Silla, Y., Goswami, G., and Barooah, M. (2020). Bacterial operational taxonomic units replace the interactive roles of other operational taxonomic units under strong environmental changes. *Curr. Genomics.* 21:512-524. doi: 10.2174/1389202921999200716104355

Kates, M. (1986). Techniques of lipidology: isolation, analysis and identification of lipids. In *Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 3, part 2, pp. 100-112. Edited by R.H.Burdon & P.H. van Knippenberg, Elsevier, Amsterdam.

Kaushal, G., Thakur, M., Rai, A. K., and Singh, S. P. (2022). A comprehensive metagenomic analysis framework revealing microbiome profile and potential for hydrocarbon degradation and carbohydrate metabolism in a Himalayan artificial lake. *Sustainability*. 14: 11455. doi: 10.3390/su141811455

Kaushik, R., Sharma, M., Gaurav, K., Jagadeeshwari, U., Shabbir, A., Sasikala, C., Ramana, C. V., and Pandit, M. K. (2020). *Paludisphaera soli* sp. nov., a new member of the family *Isosphaeraceae* isolated from high altitude soil in the Western Himalaya. *Leeuwenhoek. J. Microbiol* 113: 1663-1674. doi: 10.1007/s10482-020-01471-w

Kirk, J. L., Beaudette, L. A., Hart, M., Moutoglis, P., Klironomos, J. N., Lee, H., and Trevors, J. T. (2004). Methods of studying soil microbial diversity. *J. Microbiol. Methods*. 58: 169-188. doi: 10.1016/j.mimet.2004.04.006

Klenk, H.P., and Göker. M. (2010). En route to a genome-based classification of Archaea and Bacteria? *Syst Appl Microbiol*. 33 :175-182. doi: 10.1016/j.syapm.2010.03.003

Kodama, Y., Mashima, J., Kosuge, T., Kaminuma, E., Ogasawara, O., Okubo, K., Nakamura, Y. and Takagi, T. (2018). DNA data bank of Japan: 30th anniversary. *Nucleic Acids Res.* 46: D30-D35. doi: 10.1093/nar/gkx926

Konstantinidis, K.T., and Tiedje, J.M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. USA*.102: 2567-72. doi: 10.1073/pnas.0409727102

Konstantinidis, K.T., and Tiedje, J.M. (2005). Towards a genome-based taxonomy for prokaryotes. *J Bacteriol*.187:6258-64. doi: 10.1128/JB.187.18.6258-6264.2005

Konstantinidis, K., Rosselló-Móra, R., and Amann, R. (2017). Uncultivated microbes in need of their own taxonomy. *ISME J.* 11: 2399-2406. doi: 10.1038/ismej.2017.113

Koo, H., Mojib, N., Hakim, J.A., Hawes, I., Tanabe, Y., Andersen, D.T. and Bej, A.K. (2017) Microbial communities and their predicted metabolic functions in growth laminae of a unique large conical mat from Lake Untersee, East Antarctica. Frontiers in microbiology 8: 1347. pmid:28824553. doi: 10.3389/fmicb.2017.01347

Koren, S., Harhay, G.P., Smith, T.P., Bono, J.L., Harhay, D.M., Mcvey, S.D., Radune, D., Bergman, N.H. and Phillippy, A.M. (2013). Reducing assembly complexity of microbial genomes with single molecule sequencing. Genome Biol. 14: 1-16. doi: 10.1186/gb-2013-14-9-r101

Kosygin, L., and Dhamendra, H. (2009). Ecology and conservation of Loktak Lake, Manipur: an overview. In: Wetlands of North East India: ecology, aquatic bioresources and conservation. Akansha Publicing House, New Delhi: 1-20.

Kumar, D., Gaurav, K., Pk, S., Uppada, J., Ch, S., and Ch. V, R. (2020). *Gimesia chilikensis* sp. nov., a haloalkali-tolerant planctomycete isolated from Chilika lagoon and emended description of the genus *Gimesia*. *Int. J. Syst. Evol. Microbiol*. 70: 3647-3655. doi: 10.1099/ijsem.0.004211.

Kumar, S., Stecher, G., Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*. 33:1870-4. doi:10.1093/molbev/msw054

Kumar, V., Kumar, S., and Singh, D. (2022). Metagenomic insights into Himalayan glacial and kettle lake sediments revealed microbial community structure, function, and stress adaptation strategies. *Extremophiles*. 26: 3. doi: 10.1007/s00792-021-01252-x

Kunitsky, C., Osterhout, G., and Sasser, M. (2006). Identification of microorganisms using fatty acid methyl ester (FAME) analysis and the MIDI Sherlock Microbial Identification System. *Encyclopedia of rapid microbiological methods*. 3: 1-18.

Kurilkina, M. I., Zakharova, Y. R., Galachyants, Y. P., Petrova, D. P., Bukin, Y. S., Domysheva, V. M., Blinov, V. V., and Likhoshway, Y. V. (2016). Bacterial community composition in the water column of the deepest freshwater Lake Baikal as determined by next-generation sequencing. *FEMS Microbiol Ecol*. 92: fiw094. doi: 10.1093/femsec/fiw094

Laishram, J., and Dey, M. (2014). Water quality status of Loktak Lake, Manipur, Northeast India and need for conservation measures: a study on five selected villages. *Int. J. Sci. Res*. 4: 1-6

Langelier, B., Linard, A., Bordat, C., Lavialle, M., and Heberden, C. (2010). Long chain-polyunsaturated fatty acids modulate membrane phospholipid composition and protein localization in lipid rafts of neural stem cell cultures. *J. Cell. Biochem*. 110 :1356-1364. doi: 10.1002/jcb.22652

Langille, M. G., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Lechevalier, M.P., and Moss, C.W. Mary. (1977). Lipids in Bacterial Taxonomy - A Taxonomist's view. Crit. Rev. Microbiol. 5: 109-210. doi: 10.3109/10408417709102311

Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R. and Beiko, R.G. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 31: 814-821. doi: org/10.1038/nbt.2676

Lee, L.H., Azman, A.S., Zainal, N., Yin, W.F., Mutalib, N.S.A. and Chan, K.G. (2015). *Sinomonas humi* sp. Nov., an amylolytic actinobacterium isolated from mangrove forest soil. *Int. J. Syst. Evol. Microbiol*. 65: 996-1002. doi: 10.1099/ijs.0.000053

Lee, I., Kim, Y. O., Park, S. C., and Chun, J. (2016). OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* 66: 1100-1103. doi: 10.1099/ijsem.0.000760

Liu, Y., Lai, Q., and Shao, Z. (2017). A multilocus sequence analysis scheme for phylogeny of Thioclava bacteria and proposal of two novel species. *Front Microbiol.* 8: 1321. doi: 10.3389/fmicb.2017.01321

Logan, N.A. (1994). Bacterial systematics. Blackwell Scientific Publications, London. pp.263. doi: 10.1002/9781444313949

Loktak Development Authority (LDA) (2002). Phumdis management, Loktak newsletter, vol 2. Loktak Development Authority, Imphal and Wetland International-South Asia, New Delhi.

Loman, N.J., Constantinidou, C., Chan, J.Z., Halachev, M., Sergeant, M., Penn, C.W., Robinson, E.R., and Pallen, M.J. (2012). High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat. Rev. Microbiol*. 10: 599-606. doi: 10.1038/nrmicro2850

Loman, N.J., Quick, J., and Simpson, J.T. (2015). A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat. Methods*. 12: 733-735. doi: 10.1038/nmeth.3444

Lu, H., Deng, T., Cai, Z., Liu, F., Yang, X., Wang, Y., and Xu, M. (2020). *Janthinobacterium violaceinigrum* sp. nov., *Janthinobacterium aquaticum* sp. nov. and *Janthinobacterium rivuli* sp. nov., isolated from a subtropical stream in China. *Int. J. Syst. Evol. Microbiol.*70: 2719-2725. doi: 10.1099/ijsem.0.004097

Luo, C., Rodriguez-R, L. M., and Konstantinidis, K. T. (2014). My Taxa: an advanced taxonomic classifier for genomic and metagenomic sequences. *Nucleic Acids Res*. 42: e73-e73. doi:10.1093/nar/gku169

Maiden, M.C. (2006). Multilocus sequence typing of bacteria. *Annu. Rev. Microbiol.*60: 561-588. doi: 10.1146/annurev.micro.59.030804.121325

Maiden, M.C., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., and Feavers, I.M. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA*. 17: 3140-5. doi: 10.1073/pnas.95.6.3140

Margos, G., Fingerle, V., Cutler, S., Gofton, A., Stevenson, B., and Estrada-Peña, A. (2020). Controversies in bacterial taxonomy: The example of the genus Borrelia. *Ticks Tick Borne Dis.* 11: 101335. doi: 10.1016/j.ttbdis.2019.101335

Markowitz, V. M., Chen, I. M. A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Ratner, A., Jacob, B., Huang, J., Williams, P., and Huntemann, M. (2012). IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res*. 40(D1): D115-D122. doi: 10.1093/nar/gkr1044

Mateo-Estrada, V., Graña-Miraglia, L., López-Leal, G., and Castillo-Ramírez, S. (2019). Phylogenomics reveals clear cases of misclassification and genus-wide phylogenetic markers for *Acinetobacter*. *Genome Biol. Evolution*. 11: 2531-2541. doi: 10.1093/gbe/evz178

McCarthy, B, J., and Bolton, E.T. (1963). An approach to the measurement of genetic relatedness among organisms. Proc. Natl. Acad. Sci. USA. 50: 156-164.

Meera Krishna, B., Khan, M.A., and Khan, S.T. (2019). Next-Generation Sequencing (NGS) Platforms: An Exciting Era of Genome Sequence Analysis. In: Tripathi, V., Kumar, P., Tripathi, P., Kishore, A., Kamle, M. (eds) *Microbial Genomics in Sustainable Agroecosystems*. Springer, Singapore. doi: 10.1007/978-981-32-9860-6_6

Mehta, P., Yadav, M., Ahmed, V., Goyal, K., Pandey, R., and Chauhan, N. S. (2021). Culture-Independent Exploration of the Hypersaline Ecosystem Indicates the Environment-Specific Microbiome Evolution. *Front Microbiol*. 12: 686549.b doi: 10.3389/fmicb.2021.686549

Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L., and Göker, M. (2022). TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* 50: (D801-D807. doi: 10.1093/nar/gkab902

Meier-Kolthoff, J. P., Klenk, H. P., and Göker, M. (2014). Taxonomic use of DNA G+ C content and DNA–DNA hybridization in the genomic age. *Int. J. Syst. Evol. Microbiol.* 64: 352-356. doi: 10.1099/ijs.0.056994-0

Meitei, M. D., and Prasad, M. N. V. (2015). Phoomdi-a unique plant biosystem of Loktak lake, Manipur, North-East India: traditional and ecological knowledge. *Plant. Biosyst.* 149: 777-787. doi: 10.1080/11263504.2013.870250

Mendes Monteiro, J., Vogwill, R., Bischoff, K., and Gleeson, D. B. (2020). Comparative metagenomics of microbial mats from hypersaline lakes at Rottnest Island (WA, Australia), advancing our understanding of the effect of mat community and functional genes on microbialite accretion. *Limnol. Oceanogr.* 65: S293-S309. doi: 10.1002/lno.11323

Michael, A. J. (2016). Polyamines in eukaryotes, bacteria, and archaea. *J. Biol Chem*. 291: 14896-14903. doi: 10.1074/jbc.R116.734780

Mignard, S., and Flandrois. J.P. (2006). 16S rRNA sequencing in routine bacterial identification: a 30-month experiment. *J. Microbiol. Methods.* 67**:**574-581. doi: 10.1016/j.mimet.2006.05.009

Mo, S., Li, J., Li, B., Yu, R., Nie, S., Zhang, Z., Kashif, M., He, S., Liao, J., Jiang, Q. and Shen, P. (2020). Impacts of *Crenarchaeota* and *Halobacterota* on sulfate reduction in the subtropical mangrove ecosystem as revealed by SMDB analysis. *bioRxiv*. 2020-08. doi: 10.1101/2020.08.16.252635

Moore, E. R., Mihaylova, S. A., Vandamme, P., Krichevsky, M. I., and Dijkshoorn, L. (2010). Microbial systematics and taxonomy: relevance for a microbial commons. *Res. Microbiol.* 161: 430-438. doi: 10.1016/j.resmic.2010.05.007

Mulet, M., Gomila, M., Lemaitre, B., Lalucat, J., and García-Valdés, E. (2012). Taxonomic characterisation of Pseudomonas strain L48 and formal proposal of *Pseudomonas entomophila* sp. nov. *Syst. Appl. Microbiol.* 35: 145-149. doi: 10.1016/j.syapm.2011.12.003

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403: 853-858. doi: 10.1038/35002501

Na, S. I., Kim, Y. O., Yoon, S. H., Ha, S. M., Baek, I., and Chun, J. (2018). UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J. Microbiol*. 56: 280-285. doi: 10.1007/s12275-018-8014-6

Nahar, S., Mizan, M. F. R., Ha, A. J. W., and Ha, S.D. (2018). Advances and future prospects of enzyme-based biofilm prevention approaches in the food industry. *Compr. Rev. Food Sci. Food Saf.* 17:1484-1502. doi: 10.1111/1541-4337.12382

National Wetland Atlas: Wetlands of International Importance under Ramsar Convention (2013). Space Applications Centre, ISRO, Ahmedabad, Loktak Lake, pp 117–123

Newton, R.J., Jones, S.E., Eiler, A., McMahon, K. D., and Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.*75:14-49. doi: 10.1128/mmbr.00028-10

Ning, K., and Tong, Y. (2019). The Fast Track for Microbiome Research. *Genom. Proteom. Bioinform.*17, 1-3. doi: 10.1016/j.gpb.2019.04.001

Nowrousian, M. (2010). Next-generation sequencing techniques for eukaryotic microorganisms: sequencing-based solutions to biological problems. *Eukaryot. Cell*. 9: 1300-1310. doi: 10.1128/EC.00123-10

Nunoura, T., Chikaraishi, Y., Izaki, R., Suwa, T., Sato, T., Harada, T., Mori, K., Kato, Y., Miyazaki, M., Shimamura, S., Yanagawa, K., Shuto, A., Ohkouchi, N., Fujita, N., Takaki, Y., Atomi, H., & Takai, K. (2018). A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile. *Science (New York, N.Y.)*. 359: 559-563. doi: 10.1126/science.aao3407

Oluseyi Osunmakinde, C., Selvarajan, R., Mamba, B. B., and Msagati, T. A. M. (2019). Profiling Bacterial Diversity and Potential Pathogens in Wastewater Treatment Plants Using High-Throughput Sequencing Analysis. *Microorganisms*. 7: 506. doi: 10.3390/microorganisms7110506

Orata, F. D., Meier-Kolthoff, J. P., Sauvageau, D., and Stein, L. Y. (2018). Phylogenomic Analysis of the Gammaproteobacterial Methanotrophs (Order Methylococcales) Calls for the Reclassification of Members at the Genus and Species Levels. *Front. Microbiol*. 9: 3162. doi: 10.3389/fmicb.2018.03162

Oren, A., and Garrity, G.M. (2022). Notification that new names of prokaryotes, new combinations, and new taxonomic opinions have appeared in volume 71, part 10 of the IJSEM. *Int. J. Syst. Evol. Microbiol*. 72: 005165. doi: 10.1099/ijsem.0.005165.

Oulas, A., Pavloudi, C., Polymenakou, P., Pavlopoulos, G. A., Papanikolaou, N., Kotoulas, G., Arvanitidis, C. and Iliopoulos, L. (2015). Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinformatics and biology insights*, 9: 75-88. doi: 10.4137/BBI.S12462

Panda, A. K., Bisht, S. S., Kaushal, B. R., De Mandal, S., Kumar, N. S., and Basistha, B. C. (2017). Bacterial diversity analysis of Yumthang hot spring, North Sikkim, India by Illumina sequencing. *Big Data Anal*. 2:1-7. doi :10.1186/s41044-017-0022-8

Parker, C.T., Tindall, B. J., and Garrity, G. M. (2019). International Code of Nomenclature of Prokaryotes. *Int J Syst Evol Microbiol* 69: S1-S111. doi: 10.1099/ijsem.0.000778

Peterson, D., Bonham, K. S., Rowland, S., Pattanayak, C. W., RESONANCE Consortium, and Klepac-Ceraj, V. (2021). Comparative analysis of 16S rRNA gene and metagenome sequencing in pediatric gut microbiomes. *Front Microbiol*. 12: 670336. doi: 10.3389/fmicb.2021.670336

Prabhu D.M., Quadri S.R., Cheng J., Liu L., Chen W., Yang Y., Hozzein W. N., Lingappa K., and Li W. J. (2014). *Sinomonas mesophila* sp. nov., isolated from ancient fort soil. *J Antibiot (Tokyo)*. 68:318-321. doi: 10.1038/ja.2014.161

Pradhan, S., Srinivas, T. N. R., Pindi, P. K., Kishore, K. H., Begum, Z., Singh, P. K., Singh,A.k., Pratibha, M.S., Yasala, A.K., Reddy, G.S.N., and Shivaji, S. (2010). Bacterial biodiversity from Roopkund glacier, himalayan mountain ranges, India. *Extremophiles*. 14: 377-395. https://doi.org/10.1038/ja.2014.161

Pramanik, A., Basak, P., Banerjee, S., Sengupta, S., Chattopadhyay, D., and Bhattacharyya, M. (2015). Pyrosequencing based profiling of the bacterial community in the Chilika Lake, the largest lagoon of India. *Genomic. Data*. 4: 112-114. doi: 10.1016/j.gdata.2015.04.003

Prasad, S. N., Ramachandra, T. V., Ahalya, N., Sengupta, T., Kumar, A., Tiwari, A. K., Vijayan, V.S., and Vijayan, L. (2002). Conservation of wetlands of India-a review. *Trop. Ecol. 43*: 173-186.

Priest, F., and Goodfellow, M. (Eds.). (2000). *Applied microbial systematics*. Springer Science & Business Media.

Puranik, S., Pal, R. R., More, R. P., and Purohit, H. J. (2016). Metagenomic approach to characterize soil microbial diversity of Phumdi at Loktak Lake. *Water Sci. Technol*. 74: 2075-2086. doi: 10.2166/wst.2016.370

Pushpakumara, B.L.D.U., Tandon, K., Willis, A., and Verbruggen, H. (2023). Unravelling microalgal-bacterial interactions in aquatic ecosystems through 16S rRNA gene-based co-occurrence networks. *Sci Rep*. 13: 2743. doi: 10.1038/s41598-023-27816-9

Qiang-long, Z., Shi, L., Peng, G., and Fei-shi, L. (2014). High-throughput sequencing technology and its application. *J. Northeast Agric.Univ.* 218: 4-96. doi: 10.1016/S1006-8104(14)60073-8

Qin, Q. L., Xie, B. B., Zhang, X. Y., Chen, X. L., Zhou, B. C., Zhou, J., Oren, A. and Zhang, Y. Z. (2014). A proposed genus boundary for the prokaryotes based on genomic insights. *J. Bacteriol*. 196: 2210-2215. doi: 10.1128/JB.01688-14

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., and Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol*. 35: 833-844. doi: 10.1038/nbt.3935

Rai, P. K., and Singh, M. M (2014). Wetland Resources of Loktak Lake in Bishenpur District of Manipur, India: A Review. *Sci. Technol. J*. 2: 98-103.

Raina, V., Panda, A. N., Mishra, S. R., Nayak, T., and Suar, M. (2019). Microbial biodiversity study of a brackish water ecosystem in eastern India: the Chilika Lake. In *Microbial Diversity in the Genomic Era* (pp. 47-63). Academic Press. doi: 10.1016/B978-0-12-814849-5.00004-6

Ramchander, M., and Rahul, K. (2021). Microbial diversity in wetlands of India. *Res J Chem Environ.* 25: 3.

Ramsar Convention Secretariat, The Ramsar Convention Manual: a guide to the Convention on Wetlands, Ramsar, Iran, 1971, 6th edition Ramsar Convention Secretariat, Gland, Switzerland (2013)

Rathour, R., Gupta, J., Kumar, M., Hiloidhari, M., Mehrotra, A. K., and Thakur, I. S. (2017). Metagenomic sequencing of microbial communities from brackish water of Pangong Lake of the northwest Indian Himalayas. *Genome Announc.* 5: e01029-17. doi: 10.1128/genomeA.01029-17

Ren, L., He, D., Xing, P., Wang, Y. and Wu, Q. (2013). Bacterial diversity and ecological function in lake water bodies. *Biodivers. Sci*. 21: 421. doi: 10.3724/SP.J.1003.2013.12052

Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106:19126-19131. doi: 10.1073/pnas.0906412106

Rieger, P. G., Meier, H. M., Gerle, M., Vogt, U., Groth,T., and Knackmuss, H. J. (2002) Xenobiotics in the environment: Present and future strategies to obviate the problem of biological persistence. *J. Biotechnol*. 94: 101-123. doi: 0.1016/s0168-1656(01)00422-9

Rodriguez, L.M., and Konstantinidis, K.T. (2014). Bypassing cultivation to identify bacterial species. *Microbe*. 9:111-117. doi: 10.1128/microbe.9.111.1

Rong, X., and Huang,Y. (2014). Multi-locus sequence analysis: taking prokaryotic systematics to the next level. *Methods Microbiol*. 41:221–251. doi: 10.1016/bs.mim.2014.10.001

Rosselló-Mora, R., and Amann, R. (2001). The species concept for prokaryotes. *FEMS Microbiol. Rev.* 25: 39-67. doi: 10.1111/j.1574-6976.2001.tb00571.x

Roy, R., and Majumder, M. (2019). Assessment of water quality trends in Loktak Lake, Manipur, India. *Environmental earth sciences*.78: 1-12. doi: 10.1007/s12665-019-8383-0

Rzhetsky, A., & Nei, M. (1992). A simple method for estimating and testing minimum-evolution trees. *Mol Biol Evol*. 9: 945-967. doi: 10.1093/oxfordjournals.molbev.a040771

Sagova-Mareckova, M., Boenigk, J., Bouchez, A., Cermakova, K., Chonova, T., Cordier, T., Eisendle, U., Elersek, T., Fazi, S., Fleituch, T., and Frühe, L. (2021). Expanding ecological assessment by integrating microorganisms into routine freshwater biomonitoring. *Water Res.* 191: 116767. doi: 10.1016/j.watres.2020.116767

Sanapareddy, N., Hamp, T. J., Gonzalez, L. C., Hilger, H. A., Fodor, A. A., and Clinton, S. M. (2009). Molecular diversity of a North Carolina wastewater treatment plant as revealed by pyrosequencing. *Appl. Environ. Microbiol*. 75: 1688-1696. doi: 10.1128/AEM.01210-08

Santamaria, M., Fosso, B., Consiglio, A., De Caro, G., Grillo, G., Licciulli, F., Liuni, S., Marzano, M., Alonso-Alemany, D., Valiente, G., and Pesole, G. (2012). Reference databases for taxonomic assignment in metagenomics. *Brief. Bioinformatics*. 13: 682-695. doi: 10.1093/bib/bbs036

Sasidharan, A., Sasidharan, N. K., Amma, D. B. N. S., Vasu, R. K., Nataraja, A. V., and Bhaskaran, K. (2015). Antifungal activity of violacein purified from a novel strain of *Chromobacterium* sp. NIIST (MTCC 5522). *J. Microbiol*. 53: 694-701. doi: 10.1007/s12275-015-5173-6

Sasser, M. (1990). Identification of bacteria by gas chromatography of cellular fatty acids.

Schildkraut, C.L., Marmur, J. and Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its buoyant density in CsCl. *J. Mol. Biol*. 4: 430-443. doi: 10.1016/s0022-2836(62)80100-4

Schleifer, K. H. (2009). Classification of Bacteria and Archaea: past, present and future. *Syst. Appl. Microbiol.* 32: 533-542. doi: 10.1146/annurev.micro.59.030804.121325

Schleifer, K.H., and Kandler, O. (1972). Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol. Rev*. 36: 407-77. doi: 10.1128/br.36.4.407- 477.1972

Schneider, D., Arp, G., Reimer, A., Reitner, J., and Daniel, R. (2013). Phylogenetic analysis of a microbialite-forming microbial mat from a hypersaline lake of the Kiritimati Atoll, Central Pacific. *PloS one*. 8: e66662. doi: 10.1371/journal.pone.0066662

Schoepp-Cothenet, B., van Lis, R., Atteia, A., Baymann, F., Capowiez, L., Ducluzeau, A. L., Duval, S., ten Brink, F., Russell, M.J., and Nitschke, W. (2013). On the universal core of bioenergetics. *Biochim. Biophys. Acta-Bioenerg.* 1827: 79-93. doi: 10.1016/j.bbabio.2012.09.005

Sekora, N. S., Lawrence, K. S., Agudelo, P., Van Santen, E., and McInroy, J. A. (2009). Using FAME analysis to compare, differentiate, and identify multiple nematode species. *J. Nematol*. 41: 163.

Setubal, J.C., Almeida, N.F., and Wattam, A.R. (2018) Comparative Genomics for Prokaryotes. *Methods. Mol. Biol*. 1704: 55-78. doi: 10.1007/978-1-4939-7463-4_3

Sharpton, T.J. (2014). An introduction to the analysis of shotgun metagenomic data. *Front. Plant Sci*. 5:209. doi: 10.3389/fpls.2014.00209

Shet, S. A., and Garg, S. (2021). Prokaryotic diversity of tropical coastal sand dunes ecosystem using metagenomics. *3 Biotech*. 11: 252. doi: 10.1007/s13205-021-02809-5

Simon, C., and Daniel, R. (2011). Metagenomic analyses: past and future trends. *Appl. Environ. Microbiol.* 77: 1153-1161. doi: 10.1128/AEM.02345-10

Singer, S.J. (1972). A fluid lipid-globular protein mosaic model of membrane structure. Ann. NY. Acad. Sci. 195:16-23. doi: 10.1111/j.1749-6632.1972.tb54780

Singh, A. L., & Khundrakpam, M. L. (2011). Phumdi proliferation: a case study of Loktak lake, Manipur. *Water Environ. J.* 25: 99-105. doi: 10.1111/j.1747-6593.2009.00197.x

Singh, N. K. S., Devi, C. B., Sudarshan, M., Meetei, N. S., Singh, T. B., and Singh, N. R. (2013). Influence of Nambul river on the water quality of fresh water in Loktak lake. *Int. J. Water. Resour. Environ. Eng.* 5: 321–327. doi: 10.5897/ IJWREE12.064

Singh, T. H., and Singh, R. K. S. (1994). *Ramsar sites of India, Loktak lake*, *Manipur* (Vol.2). New -Delhi, India: World Wide Fund for Nature (WWF).

Singh, T.S., J.G. Patel, G.S. Rajkumari, N.S. Singh, C.B. Santhosh and T.H. Manihar. (2003). Wetland information systems for conservation and management of Loktak Lake, a project report.

Skerman, V. B. D., V. McGowan, and P. H. A. Sneath, (editors) (1980). Approved lists of bacterial names. *Int. J. Syst. Bacteriol*. 30: 225-420

Smith, L., Sanders, J., Kaiser, R., and Hughes, P. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature*. 321:674-679. doi: 10.1038/321674a

Sneath, P.H.A. (1989). Numerical taxonomy. In: Bergey's Manual of Systematic Bacteriology, Vol. 4 (Williams, S.T., Sharpe, M.E. and Holt, J.G., Eds.), pp. 2303-2305. Williams and Wilkins, Baltimore, MD.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J.M., and Herndl, G.J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci.* 103: 12115-12120. doi: 10.1073/pnas.0605127103

Solden, L., Lloy, K., and Wrighton, K. (2016) The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr Opin Microbiol.* 31:217-226. doi: 10.1016/j.mib.2016.04.020

Stackebrandt, E. and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiol.Today*. 33: 152-155.

Stackebrandt, E. and Goebel, B.M. (1994). Taxonomic note: A place for DNADNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol*. 44, 846-849. doi: 10.1099/00207713-44-4-846

Stackebrandt, E., Frederiksen, W., Garrity, G. M., Grimont, P. A., Kämpfer, P., Maiden, M. C., Nesme, X., Rosselló-Mora, R., Swings, J., Trüper, H.G., and Vauterin, L. (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int. J. Syst. Evol. Microbiol*. 52: 1043-1047. doi: 10.1073/pnas.95.6.3140

Stackebrant, E. (2011). Molecular taxonomic parameters. *Microbiol. Aust*. 32: 59-61. doi: 10.1071/MA11059

Staley, C., Gould, T. J., Wang, P., Phillips, J., Cotner, J. B., and Sadowsky, M. J. (2014). Core functional traits of bacterial communities in the Upper Mississippi River show limited variation in response to land cover. *Front Microbiol.* 5: 414. doi: 10.3389/fmicb.2014.00414

Staley, C., Unno, T., Gould, T. J., Jarvis, B., Phillips, J., Cotner, J. B., Sadowsky, M. J. (2013). Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *J. Appl. Microbiol*. 115: 1147-1158 doi: 10.1111/jam.12323.

Stefanini, I., and Cavalieri, D. (2018). Metagenomic approaches to investigate the contribution of the vineyard environment to the quality of wine fermentation: potentials and difficulties. *Front Microbiol.* 9: 991. doi: 10.3389/fmicb.2018.00991

Steinert, G., Wemheuer, B., Janussen, D., Erpenbeck, D., Daniel, R., Simon, M., Brinkhoff, T., and Schupp, P.J. (2019). Prokaryotic diversity and community patterns in Antarctic continental shelf sponges. *Front Mar Sci.* 6:297. doi: 10.3389/fmars.2019.00297

Sterner, R. W., Keeler, B., Polasky, S., Poudel, R., Rhude, K., and Rogers, M. (2020). Ecosystem services of Earth's largest freshwater lakes. *Ecosyst. Serv.* 41: 101046.

doi: 10.1016/j.ecoser.2019.101046

Strominger, J.L., Izaki, K.A.Z.U.O., Matsuhashi, M.I.C.H.I.O. and Tipper, D.J. (1967). Peptidoglycan transpeptidase and D-alanine carboxypeptidase. Penicillin-sensitive enzymatic reactions. *Fed. Proc*. 26: 9-22.

Sueoka, N. (1962). On the genetic basis of variation and heterogeneity of DNA base composition. *Proc. Natl. Acad. Sci.* 48: 582-592. doi: 10.1073/pnas.48.4.582

Suzuki, K., Goodfellow, M., and O'Donnell, A.G. (1993). Cell envelopes and classification. In: Goodfellow, M., O'Donnell., editors. *Handbook of New Bacterial Systematics.* London: Academic Press Ltd; 1993. pp. 195-250.

Takeuchi, M., and Yokota, A. (1989). Cell-wall polysaccharides in coryneform bacteria. *J. Gen. Appl. Microbiol.* 35: 233-253. doi: 10.2323/jgam.35.233

Takhelmayum, K., and Gupta, S. (2011). "Distribution of aquatic insects in phumdis (floating island) of Loktak Lake, Manipur, northeastern India." *J. Threat. Taxa*. 3: 1856-1861. doi: 10.11609/JoTT.o2526.1856-61

Tamaki, H., Sekiguchi, Y., Hanada, S., Nakamura, K., Nomura, N., Matsumura, M., & Kamagata, Y. (2005). Comparative analysis of bacterial diversity in freshwater sediment of a shallow eutrophic lake by molecular and improved cultivation-based techniques. *Appl. Environ. Microbiol*. 71: 2162-2169. doi: 10.1128/AEM.71.4.2162-2169.2005

Thibaud-Nissen, F., DiCuccio, M., Hlavina, W., Kimchi, A., Kitts, P. A., Murphy, T. D., Pruitt, K.D. and Souvorov, A. (2016). P8008 the NCBI eukaryotic genome annotation pipeline. *J. Anim. Sci*. 94: 184 doi: 10.2527/jas2016.94supplement4184x

Tindall, B.J., Rosselló-Móra, R., Busse, H.J., Ludwig, W., and Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol*. 60: 249-266. doi: 10.1099/ijs.0.016949-0

Toyama, D., Kishi, L. T., Santos-Júnior, C. D., Soares-Costa, A., de Oliveira, T. C., de Miranda, F. P., and Henrique-Silva, F. (2016). Metagenomics Analysis of Microorganisms in Freshwater Lakes of the Amazon Basin. *Genome Announc*. 4: e01440-16. doi: 10.1128/genomeA.01440-16

Tran, P., Ramachandran, A., Khawasik, O., Beisner, B. E., Rautio, M., Huot, Y., and Walsh, D. A. (2018). Microbial life under ice: Metagenome diversity and in situ activity of Verrucomicrobia in seasonally ice-covered Lakes. *Environ Microbiol.* 20: 2568-2584. doi: 10.1111/1462-2920.14283.

Tripp, H. J., Kitner, J. B., Schwalbach, M. S., Dacey, J. W., Wilhelm, L. J., and Giovannoni, S. J. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature*. 452: 741-744. doi: org/10.1038/nature06776

Trishal, C. I., and Manihar, T. (2004). *Loktak-the atlas of Loktak lake. Wetlands International South Asia (WISA)*. New-Delhi, India: Loktak Development Authority (LDA).

Trüper, H.G., Severin, J., Wohlfarth, A., Müller, E., Galinski, E.A. (1991). Halophily, Taxonomy, Phylogeny and Nomenclature. In: Rodriguez-Valera, F. (eds) General and Applied Aspects of Halophilic Microorganisms. NATO ASI Series, vol 201. Springer, Boston, MA. doi: 10.1007/978-1-4615-3730-4_1

Urdiain, M., López-López, A., Gonzalo, C., Busse, H. J., Langer, S., Kämpfer, P. and Rosselló-Móra, R. (2008). Reclassification of *Rhodobium marinum* and *Rhodobium pfennigii* as *Afifella marina* gen. nov. comb. nov. and *Afifella pfennigii* comb. nov., a new genus of photoheterotrophic *Alphaproteobacteria* and emended descriptions of *Rhodobium*, *Rhodobium orientis* and *Rhodobium gokarnense*. *Syst. Appl. Microbiol*. 31: 339-51. doi: 10.1016/j.syapm.2008.07.002

Vandamme, P., Holmes, B., Vancanneyt, M., Coenye, T., Hoste, B., Coopman, R., Revets, H., Lauwers, S., Gillis, M., Kersters, K., and Govan, J.R.W. (1997). Occurrence of Multiple Genomovars of Burkholderia cepacia in Cystic Fibrosis Patients and Proposal of Burkholderia multivorans sp. nov. *Int. J. Syst. Evol. Microbiol.* 47: 1188-1200. doi: 10.1099/00207713-47-4-1188

Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K., and Swings, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol. Rev.* 60: 407-438. doi: 10.1128/mr.60.2.407-438.1996

Van't Wout, J. W., and Bijlmer, H. A. (2005). Bacteremia due to Streptococcus gallolyticus, or the perils of revised nomenclature in bacteriology. *Clin. Infect. Dis.* 40: 1070-1071. doi: 10.1086/428671

Varjani, S., Srivastava, V. K., Sindhu, R., Thakur, I. S., and Gnansounou, E. (2018). Culture based approaches, dependent and independent, for microbial community fractions in petroleum oil reservoirs.

Vaz-Moreira, I., Egas, C., Nunes, O. C., and Manaia, C. M. (2011). Culture-dependent and culture-independent diversity surveys target different bacteria: a case study in a freshwater sample. *Antonie. Leeuwenhoek. J. Microbiol.* 100: 245-257. doi: 10.1007/s10482-011-9583-0

Vijayan, J., Nathan, V. K., Ammini, P., and Ammanamveetil, A. M. H. (2023). Bacterial diversity in the aquatic system in India based on metagenome analysis-a critical review. *Environ. Sci. Pollut. Res*. 30: 28383-28406. doi: 10.1007/s11356-023-25195-2

Warth, A. D., and Strominger, J. L. (1971). Structure of the peptidoglycan from vegetative cell walls of Bacillus subtilis. *Biochem*.10:4349-4358. doi: 10.1021/bi00800a001

Wayne, L. G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R., Stackebrandt, E.S.M.P. and and Starr, M.P. (1987). Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol*. 37:463-464. doi: org/10.1099/00207713-37-4-463

WISA and LDA (2002). Management of Phumdis in Loktak Lake. 9-23. In: Trisal, C.L & T.H. Manihar (eds.). Proceedings of a Workshop held at Imphal, Manipur.

Woese, C.R. (2004). A new biology for a new century. *Microbiol Mol Biol Rev* 68:173-86. doi: 10.1128/MMBR.68.2.173-186.2004. doi: 10.128/mmbr.68.2.173-186.2004

Wu, X., Xi, W., Ye, W., and Yang, H. (2007). Bacterial community composition of a shallow hypertrophic freshwater lake in China, revealed by 16S rRNA gene sequences. *FEMS Microbiol. Ecol.* 61:85-96. doi: 10.1111/j.1574-6941.2007.00326.x

Xu, L., Dong, Z., Fang, L., Luo,Y., Wei, Z., Guo, H., Zhang, G., Gu, Y. Q., Coleman-Derr, D., Xia, Q., Wang, Y. (2019). OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* 47: W52-W58. doi: 10.1093/nar/gkz33

Xu, X., Liu, X., Chen, S., Li, B., Wang, X., Fan, C., Wang, G., and Ni, H. (2016). Selection of relatively exact reference genes for gene expression studies in flixweed (*Descurainia sophia*) by quantitative real-time polymerase chain reaction. *Pestic Biochem Physiol.* 127: 59-66. doi: 10.1016/j.pestbp.2015.09.007.

Yadav, A. N., Sachan, S. G., Verma, P., and Saxena, A. K. (2015). Prospecting cold deserts of north western Himalayas for microbial diversity and plant growth promoting attributes. *J. Biosci. Bioeng*. 119: 683-693. doi: 10.1016/j.jbiosc.2014.11.006

Yadav, N., Kour, D., and Yadav, A. N. (2018). Microbiomes of freshwater lake ecosystems. *J. Microbiol. Exp*. *6*: 245-248. doi: 10.15406/jmen.2018.06.00223

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci*. 13: 555-556. doi: 10.1093/bioinformatics/13.5.555

Yannarell, A., and Ken,t A.(2009). Bacteria, distribution and community structure. In book: Encyclopedia of Inland Waters (pp.201-210). doi: 10.1016/B978-012370626-3.00123-X

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K. H., Whitman, W.B., Euzéby, J., Amann, R., and Rosselló-Móra, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol*.12: 635-45. doi: 10.1038/nrmicro3330

Yoon, J.H., Kang, S.J., Park, S., and Oh, T.K. (2005). Reclassification of the three species of the genus *Krokinobacter* into the genus *Dokdonia* as *Dokdonia genika* comb. nov., *Dokdonia diaphoros* comb. nov. and *Dokdonia eikasta* comb. nov., and emended description of the genus *Dokdonia* Yoon et al. 2005. *Int. J. Syst. Evo.l Microbiol.* 62:1896-1901. doi: 10.1099/ijs.0.035253-0

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., & Chun, J. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. MicrobioL.* 67: 1613-1617. doi: 10.1099/ijsem.0.001755

Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, D., Juzokaite, L., Vancaester, E., Seitz, K.W., Anantharaman, K., Starnawski, P., Kjeldsen, K.U., and Stott, M.B. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature*. 541: 353-358. doi:10.1038/nature21031

Zhang, M. Y., Xie, J., Zhang, T. Y., Xu, H., Cheng, J., Li, S. H., Li, W.J. and Zhang, Y. X. (2014). *Sinomonas notoginsengisoli* sp. nov., isolated from the rhizosphere of Panax notoginseng. *Antonie. Leeuwenhoek. J. Microbiol*. 106: 827-835. doi: 10.1007/s10482-014-0252-y

Zhao, Y., Sun, C., Zhao, D., Zhang, Y., You, Y., Jia, X., Yang, J., Wang, L., Wang, J., Fu, H., Kang, Y., Chen, F., Yu, J., Wu, J., and Xiao, J. (2018). PGAP-X: extension on pan-genome analysis pipeline. *BMC Genom*. 19: 115-124. doi: 10.1186/s12864-017-4337-7

Zhou, Y., Wei, W., Wang, X., and Lai, R. (2009). Proposal of *Sinomonas flava* gen. nov., sp. nov., and description of *Sinomonas atrocyanea* comb. nov. to accommodate *Arthrobacter atrocyaneus*. *Int. J. Syst Evol. Microbiol*. 5: 259-263. doi: 10.1099/ijs.0.000695-0

Zhou Y., Chen X., Zhang Y., Wang W., and Xu J. (2012). Description of *Sinomonas soli* sp. nov., reclassification of *Arthrobacter echigonensis* and *Arthrobacter albidus* (Ding *et al*. 2009) as *Sinomonas echigonensis* comb. nov. and *Sinomonas albida* comb. nov., respectively, and emended description of the genus *Sinomonas*. *Int. J. Syst. Evol Microbiol* .62, 764-769. doi: 10.1099/ijs.0.030361-0

Zwart, G., Crump, B. C., Kamst-van Agterveld, M. P., Hagen, F., and Han, S. K. (2002). Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat. Microb. Ecol*. 28: 141-155. doi: 10.3354/ame028141

Zwart, G., Hiorns, W. D., Methé, B. A., van Agterveld, M. P., Huismans, R., Nold, S. C., Zehr, J. P., and Laanbroek, H. J. (1998). Nearly identical 16S rRNA sequences recovered from lakes in North America and Europe indicate the existence of clades of globally distributed freshwater bacteria. Syst. Appl. Microbiol. 21: 546-556. doi: 10.1016/S0723-2020(98)80067-2

PUBLICATIONS

7. PUBLICATIONS

1. Kumar, G., **Lhingjakim, K. L**., Uppada, J. Ahamad, S., Kumar, D., Kashif, G.M., Sasikala, C., and Ramana, C.V. (2021). *Aquisphaera insulae* sp. nov., a new member in the family *Isosphaeraceae*, isolated from the floating island of Loktak lake and emended description of the genus *Aquisphaera*. *Antonie. Leeuwenhoek. J. Microbiol.* 114: 1465–1477. https://doi.org/10.1007/s10482-021-01615-6
2. **Lhingjakim, K. L**., Smita, N., Kumar, G., Jagadeeshwari, U., Ahamad. S., Sasikala, C., and Ramana, C.V. (2022). *Paludisphaera rhizosphaereae* sp. nov., a new member of the family *Isosphaeraceae,* isolated from the rhizosphere soil of *Erianthus ravennae*. *Antonie. Leeuwenhoek. J. Microbiol.* 115: 1073–1084. https://doi.org/10.1007/s10482-022-01758-0
3. **Lhingjakim, K. L.**, Sasikala, C. and Ramana, C.V. (2022). Draft Genome Sequence of *Glutamicibacter* sp. Strain JC586, Isolated from Soil Sediment of the Floating Islands of Loktak Lake, India. *Microbiol*. *Resour*. *Announc*. Sep 20: e00746-22

CONFERENCES AND SEMINARS

1. Presented poster at the "59th Annual Conference of Association of Microbiologists of India" (Dec. 2018), University of Hyderabad, Hyderabad.
2. Presented an oral presentation at "National Conference on Frontiers in Plant Biology" (Jan. 2019), University of Hyderabad, Hyderabad.

Antonie van Leeuwenhoek (2022) 115:1073–1084
https://doi.org/10.1007/s10482-022-01758-0

ORIGINAL PAPER

Paludisphaera rhizosphaereae sp. nov., a new member of the family *Isosphaeraceae,* isolated from the rhizosphere soil of *Erianthus ravennae*

Khongsai L. Lhingjakim · Nandardhane Smita · Gaurav Kumar · Uppada Jagadeeshwari · Shabbir Ahamad · Chintalapati Sasikala · Chintalapati Venkata Ramana

Received: 4 March 2022 / Accepted: 14 June 2022 / Published online: 2 July 2022

Abstract Two axenic cultures of *Planctomycetota* were isolated from distinct geographical locations of India. Strain JC665^T was isolated from a rhizosphere soil of Loktak lake, Manipur, whereas strain JC747 was isolated from a soil sediment at Pallikkara village, Kerala, India. The two closely related strains shared the highest 16S rRNA gene sequence identity (94.6%) with *Paludisphaera borealis* PX4^T, while the 16S rRNA gene sequence identity between both strains was 100%. Both strains grow aerobically, stain Gram negative, colonies are light pink-coloured, cells are non-motile, spherical to oval-shaped and tolerate NaCl up to 2% (w/v). While strain JC665^T grows well up to pH 9.0, strain JC747 grows only up to pH 8.0. The respiratory quinone in both strains is MK-6. $C_{16:0}$, $C_{18:1}\omega9c$ and $C_{18:0}$ are the major fatty acids. Phosphatidylcholine, two unidentified glycolipids, seven unidentified lipids and two unidentified phospholipids made up the polar lipid composition of both strains. Both strains have genome sizes of about 8.0 Mb and a DNA G+C content of 66.4 mol%. Both strains contain genes coding for enzymes putatively involved in the production of lycopene-related carotenoids. The phylogenetic position together with the results of the analysis of morphological, physiological and genomic features support the classification of strain JC665^T as a new species of the genus *Paludisphaera*, for which we propose the name *Paludisphaera rhizosphaerae*

Khongsai L. Lhingjakim, Nandardhane Smita, Gaurav Kumar have contributed equally and are considered as equal first authors

The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequences of the strains JC665T and JC747 are LR746340 and OU374731, respectively. The GenBank/EMBL/DDBJ accessions for the whole genome shotgun sequence for strains JC665T and JC747 are JAALCR000000000 and JAHPZK000000000, respectively

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10482-022-01758-0.

K. L. Lhingjakim · N. Smita · G. Kumar · S. Ahamad · C. V. Ramana (✉)
Department of Plant Sciences, School of Life Sciences, University of Hyderabad, P.O. Central University, Hyderabad 500046, India
e-mail: cvr449@gmail.com

Present Address:
G. Kumar
Department of Microbiology, School of Sciences, ITM (SLS) Baroda University, Halol, Vadodara, Gujarat 391510, India

U. Jagadeeshwari · C. Sasikala (✉)
Bacterial Discovery Laboratory, Centre for Environment, Institute of Science and Technology, JNT University Hyderabad, Kukatpally, Hyderabad 500085, India
e-mail: sasi449@yahoo.ie; sasikala.ch@gmail.com

Antonie van Leeuwenhoek (2021) 114:1465–1477
https://doi.org/10.1007/s10482-021-01615-6

SHORT COMMUNICATION

Aquisphaera insulae sp. nov., a new member in the family *Isosphaeraceae*, isolated from the floating island of Loktak lake and emended description of the genus *Aquisphaera*

Gaurav Kumar · Khongsai L. Lhingjakim · Jagadeeshwari Uppada · Shabbir Ahamad · Dhanesh Kumar · Gulam Mohammad Kashif · Chintalapati Sasikala · Chintalapati Venkata Ramana

Received: 9 March 2021 / Accepted: 3 July 2021 / Published online: 14 July 2021

Abstract Strain JC669^T was isolated from a floating island of Loktak lake, Manipur, India and shares the highest 16S rRNA gene sequence identity with *Aquisphaera giovannonii* OJF2^T. The novel strain is aerobic, Gram negative, light pink-coloured, non-motile, NaCl intolerant and spherical to oval-shaped. It grows in the form of single cells or aggregates and possibly forms structures which appear like fruiting bodies. Strain JC669^T grows well up to pH 9.0.The isolate produces MK-6 as respiratory quinone, $C_{18:1}\omega9c$, $C_{16:0}$ and $C_{18:0}$ as major fatty acids and phosphatidylcholine, an unidentified amino lipid, an unidentified choline lipid (UCL) and six additional unidentified lipids (UL1, 2, 3, 4, 5, 6) as polar lipids. Strain JC669^T has a large genome size of 10.04 Mb and the genomic G + C content was 68.5 mol%. The genome contained all genes essential for lycopene related carotenoid biosynthesis. The polyphasic analysis of its phylogenetic position, morphological, physiological and genomic features supports the classification of strain JC669^T as a novel species of the genus *Aquisphaera*, for which we propose the name *Aquisphaera insulae* sp. nov. Strain JC669^T (= KCTC 72672^T = NBRC 114306^T) is the type strain of the novel species.

Keywords *Planctomycetes* · *Isosphaeraceae* · *Aquisphaera* · Fruiting body · Sporotan

Gaurav Kumar and Khongsai L. Lhingjakim contributed equally to this work.

The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequence of the strain JC669^T is LR782133. The NCBI accession for the whole genome shotgun sequence of strain JC669^T is JAALJH000000000.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10482-021-01615-6.

G. Kumar · K. L. Lhingjakim · S. Ahamad · D. Kumar · G. M. Kashif · C. V. Ramana (✉)
Department of Plant Sciences, School of Life Sciences, University of Hyderabad, P.O. Central University, Hyderabad 500046, India
e-mail: cvr449@gmail.com

J. Uppada · C. Sasikala
Bacterial Discovery Laboratory, Centre for Environment, Institute of Science and Technology, J. N. T. University Hyderabad, Kukatpally, Hyderabad 500085, India
e-mail: sasi449@yahoo.ie; sasikala.ch@gmail.com

Abbreviations

NCBI	National Centre for Biotechnology Information
*g*ANI	Genome average nucleotide identity
AAI	Average amino acid identity
*d*DDH	Digital DNA–DNA hybridization
HPLC	High performance liquid chromatography
KCTC	Korean collection for type cultures
NBRC	Biological Resource Centre
PGAP	Prokaryotic genome annotation pipeline

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Glutamicibacter* sp. Strain JC586, Isolated from Soil Sediment of the Floating Islands of Loktak Lake, India

Khongsai L. Lhingjakim,[a] Chintalapati Sasikala,[b] Chintalapati Venkata Ramana[a]

[a]Department of Plant Sciences, School of Life Sciences, University of Hyderabad, P.O. Central University, Hyderabad, India
[b]Bacterial Discovery Laboratory, Centre for Environment, IST, JNT University Hyderabad, Kukatpally, Hyderabad, India

ABSTRACT The 3.52-Mbp whole-genome sequence of a *Glutamicibacter* sp. strain isolated from soil sediment of the floating islands of Loktak Lake is reported. The genomic information here gives insight into the presence of genes linked to oxidative stress, osmo-protection, and cold shock proteins which helps in the survival of the organism under extreme environmental conditions.

The genus *Glutamicibacter* belongs to the phylum *Actinomycetota*, consisting of Gram-positive, high GC, rod-coccus cells. Although members of *Glutamicibacter* have been isolated from different ecological niches, relatively little is known about their physical attributes that allow these organisms to survive in harsh environments (1, 2).

Glutamicibacter sp. strain JC586 was isolated from a soil sample collected from the floating islands or Phumdis (3) of Loktak Lake, India (24°30′21″ N/93°47′43″ E). Phumdis constitute a dense rhizosphere extending down to the sediment of the lake and hence serve as an ecological habitat for several groups of bacteria (4). The lake is an ecological hotspot with a remarkable diversity of flora and fauna and was declared a Ramsar site (a wetland of international significance) in 1990. Briefly, a serially diluted rhizosphere soil sample was plated onto nutrient agar and incubated at 30°C for 7 days. Yellow-colored colonies were selected from the 10^{-4} dilution which were further purified by repeated streaking. Genomic DNA was isolated from a single colony using the Nucleo-pore genomic DNA (gDNA) fungal bacterial mini kit from Genetix Biotech Asia Pvt. Ltd., India, and the genome sequence was outsourced to AgriGenome Pvt. Ltd. (Kochi, India). Library preparation was carried out using the NEBNext Ultra DNA library preparation kit. The Illumina HiSeq 2500 instrument was used for whole-genome sequencing. The total number of raw reads generated was 4,730,390 (forward and reverse strands) with a read length of 100 bp. The fastq files were trimmed by removing the adapter sequences using Cutadapt version 1.11 (5) and by filtering out reads with an average quality score Q of <30 in any of the paired-end reads using Sickle version 1.33 (6). The unique reads were fetched using FastUniq version 1.1 (7), and finally, plasmids were removed from the cleaned reads using Bowtie2 version 2.2.6. *De novo* assembly was performed using ABySS version 2.0.2 with a *k*-mer range of 31 to 95 (8). Default parameters were used for all software unless otherwise specified. The draft genome of *Glutamicibacter* sp. JC586 contains 28 contigs, with 3,524,842 bp, and the N_{50} value was 558,740 bp with 100× coverage and a GC content of 60.04%. Genes were predicted using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 5.0 which shows 3,294 total predicted genes; 3,177 coding genes; and 75 RNAs, including 62 tRNAs, 10 rRNAs, and 3 noncoding RNAs (ncRNAs) (Table 1).

An EzBioCloud BLAST analysis of the 16S rRNA gene sequence (1,514 nucleotides [nt]) of strain JC586 yielded the highest identity (99.3%) with the type strain of *Glutamicibacter halophytocola* KLBMP 5180, with an overall average nucleotide identity (ANI) (9) value of 79.53% and DNA-DNA hybridization (dDDH) (10) of 21%. Annotation of the JC586 genome reveals

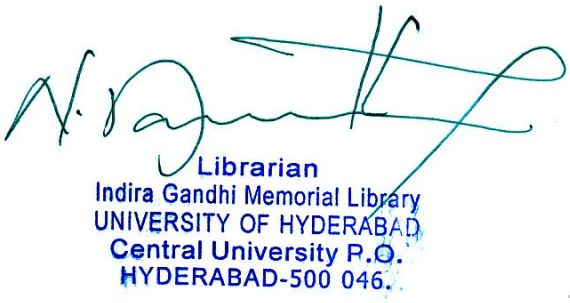
Editor Julia A. Maresca, University of Delaware

Address correspondence to Chintalapati Sasikala, sasi449@yahoo.ie, or Chintalapati Venkata Ramana, cvr449@gmail.com.
The authors declare no conflict of interest.

Received 22 July 2022
Accepted 26 August 2022
Published 20 September 2022

Genetic and Culture Diversity of Bacteria from Loktak Lake and Descriptions of a Few New Taxa

by L Lhingjakim Khongsai



Submission date: 03-Jul-2023 03:49PM (UTC+0530)
Submission ID: 2125968811
File name: L_Lhingjakim_Khongsai.pdf (1.28M)
Word count: 23239
Character count: 128558

Genetic and Culture Diversity of Bacteria from Loktak Lake and Descriptions of a Few New Taxa

ORIGINALITY REPORT

14%	12%	14%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.researchsquare.com Internet Source	5%
2	Submitted to University of Hyderabad, Hyderabad Student Paper	2%
3	link.springer.com Internet Source	1%
4	www.ncbi.nlm.nih.gov Internet Source	1%
5	Rishabh Kaushik, Meesha Sharma, Kumar Gaurav, U. Jagadeeshwari, A. Shabbir, Ch. Sasikala, Ch. V. Ramana, Maharaj K. Pandit. "Paludisphaera soli sp. nov., a new member of the family Isosphaeraceae isolated from high altitude soil in the Western Himalaya", Antonie van Leeuwenhoek, 2020 Publication	<1%
6	www.esp.org Internet Source	<1%

These papers are published by the student herself.

Prof. Ch. Venkata Ramana
Department of Plant Sciences, School of Life Sciences
University of Hyderabad, P.O. Central University,
Hyderabad - 500 046, India
21/7/2023

7	The Prokaryotes, 2014. Publication	<1 %
8	"Bergey's Manual® of Systematic Bacteriology", Springer Nature, 2012 Publication	<1 %
9	Khongsai L. Lhingjakim, Nandardhane Smita, Gaurav Kumar, Uppada Jagadeeshwari et al. "Paludisphaera rhizosphaereae sp. nov., a new member of the family Isosphaeraceae, isolated from the rhizosphere soil of Erianthus ravennae", Antonie van Leeuwenhoek, 2022 Publication	<1 %
10	journals.plos.org Internet Source	<1 %
11	Jasna Vijayan, Vinod Kumar Nathan, Parvathi Ammini, Abdulla Mohamed Hatha Ammanamveetil. "Bacterial diversity in the aquatic system in India based on metagenome analysis—a critical review", Environmental Science and Pollution Research, 2023 Publication	<1 %
12	Dhanesh Kumar, Gaurav Kumar, Jagadeeshwari Uppada, Shabbir Ahmed, Chintalapati Sasikala, Chintalapati Venkata Ramana. "Descriptions of Roseiconus nitratireducens gen. nov. sp. nov. and	<1 %

Roseiconus lacunae sp. nov.", Archives of Microbiology, 2020
Publication

13	www.microbiologyresearch.org Internet Source	<1 %
14	Submitted to MAHSA University Student Paper	<1 %
15	ijcs.ro Internet Source	<1 %
16	www.mdpi.com Internet Source	<1 %
17	Submitted to Vrije Universiteit Brussel Student Paper	<1 %
18	Khongsai L Lhingjakim, Nandardhane Smita, Gaurav Kumar, Uppada Jagadeeshwari et al. "Paludisphaera rhizosphaereaesp. nov., a new member of the familyIsosphaeraceae,isolated from the rhizosphere soil ofErianthus ravennae", Research Square Platform LLC, 2022 Publication	<1 %
19	worldwidescience.org Internet Source	<1 %
20	Ming-Jing Zhang, Xue-Han Li, Li-Yang Peng, Shuai-Ting Yun, Zhuo-Cheng Liu, Yan-Xia Zhou. "Williamsia soli sp. nov., an	<1 %

actinobacterium isolated from soil at a thermal power plant in Yantai, China", Archives of Microbiology, 2021
Publication

21	Neelam Yadav, Divjot Kour. "Microbiomes of freshwater lake ecosystems", Journal of Microbiology & Experimentation, 2018 Publication	<1 %
22	www.frontiersin.org Internet Source	<1 %
23	Submitted to Indian Institute of Technology Guwahati Student Paper	<1 %
24	Dhanesh Kumar, Nandardhane Smita, Gaurav Kumar, Gandham Suresh et al. "Arenibacter lacus sp. nov., Isolated from Chilika Lagoon, India", Current Microbiology, 2020 Publication	<1 %
25	R. K. Sravan. "Shewanella fodinae sp. nov., isolated from a coal mine and from Chilika marine lagoon of India", INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY, 08/28/2009 Publication	<1 %
26	"Systematic Bacteriology", Springer Science and Business Media LLC, 2009 Publication	<1 %

27	Ines Friedrich, Jacqueline Hollensteiner, Dominik Schneider, Anja Poehlein, Robert Hertel, Rolf Daniel. "First Complete Genome Sequences of Janthinobacterium lividum EIF1 and EIF2 and Their Comparative Genome Analysis", Genome Biology and Evolution, 2020 Publication	<1%
28	Raj Kumar Regar, Mohan Kamthan, Vivek Kumar gaur, Satyendra Pratap Singh et al. "Microbiome divergence across four major Indian riverine water ecosystems impacted by anthropogenic contamination: A comparative metagenomic analysis", Research Square, 2020 Publication	<1%
29	www.researchgate.net Internet Source	<1%
30	Chen, Q., L.-N. Sun, X.-x. Zhang, J. He, S.-W. Kwon, J. Zhang, S.-p. Li, and J.-g. Gu. "Roseomonas rhizosphaerae sp. nov., a triazophos-degrading bacterium isolated from soil", INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY, 2014. Publication	<1%
31	Y. Subhash, John J. Bang, Taek H. You, Sang-Seob Lee. "Description of Comamonas	<1%

sediminis sp. nov., isolated from lagoon sediments", International Journal of Systematic and Evolutionary Microbiology, 2016
Publication

32	Bergey's Manual® of Systematic Bacteriology, 2005. Publication	<1 %
33	Girija Kaushal, Monika Thakur, Amit Kumar Rai, Sudhir P. Singh. "A Comprehensive Metagenomic Analysis Framework Revealing Microbiome Profile and Potential for Hydrocarbon Degradation and Carbohydrate Metabolism in a Himalayan Artificial Lake", Sustainability, 2022 Publication	<1 %
34	Javier Edo Varg, Richard Svanbäck. "Multi stress system: Microplastics in freshwater and their effects on host microbiota", Science of The Total Environment, 2022 Publication	<1 %
35	www.genome.jp Internet Source	<1 %
36	assets.researchsquare.com Internet Source	<1 %
37	diaspora.zbmed.de Internet Source	<1 %

38	koreascience.kr Internet Source	<1%
39	D. Clermont. "Microbacterium binotii sp. nov., isolated from human blood", INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY, 05/01/2009 Publication	<1%
40	Nuohan Xu, Jian Kang, Yangqing Ye, Qi Zhang et al. "Machine learning predicts ecological risks of nanoparticles to soil microbial communities", Environmental Pollution, 2022 Publication	<1%
41	Soon Dong Lee, Hong Lim Yang, In Seop Kim. "Rouxiella aceris sp. nov., isolated from tree sap and the emended description of the genus Rouxiella", Antonie van Leeuwenhoek, 2021 Publication	<1%
42	ciencia.ucp.pt Internet Source	<1%
43	Submitted to Higher Education Commission Pakistan Student Paper	<1%
44	Submitted to KYUNG HEE UNIVERSITY Student Paper	<1%