Characterization of a novel soil ultramicrobacterium from the rarely cultured *Verrucomicrobiota* phylum

(難培養性の *Verrucomicrobiota* 門に属する新規超微小土壌細菌の特性)

Graduate School of Agriculture, Hokkaido University

Doctoral Dissertation

ISLAM MD SAMIUL

June 2025

TABLE OF CONTENTS

ABSTRACT……... II

ACKNOWLEDGEMENTS……... V

LIST OF FIGURES……..…….. VI

LIST OF TABLES……. VII

LIST OF ABBREVIATIONS…… VIII

CHAPTER I

General Introduction 1

1.1 Microbial diversity and its ecological significance ……..…… 2

1.2 The *Verrucomicrobiota* phylum: current knowledge and challenges……..…… 4

1.3 Significance of ultramicrobacterium and slow growers…… 6

1.4 Importance of *Verrucomicrobiot*a members……. 8

1.5 Dissertation aims and objectives……..……. 13

CHAPTER II

Morphological, physiological, and phylogenetic characterization of a novel *Verrucomicrobiota* strain 14

2.1. Introduction……..…… 15

2.2 Materials and Methods……..…… 15

2.3 Results and Discussion……..…… 25

2.4 Conclusions……..…… 41

CHAPTER III

Genomic characterization and potential distribution of a novel *Verrucomicrobiota* strain 43

3.1. Introduction……..…… 44

3.2 Materials and Methods……..…… 44

3.3 Results and Discussion……..…… 47

3.4 Conclusions……..…… 58

CHAPTER IV

Taxonomic proposal of a novel *Verrucomicrobiota* strain 60

Proposal, submission and accession no. ……..……... 61

GENERAL CONCLUSION……..... 64

REFERENCES ……..……. 66

APPENDIX……..……. 77

ABSTRACT

Verrucomicrobiota is a widespread but poorly recognized bacterial phylum found in soil, freshwater, and host-associated habitats. Despite efforts to cultivate uncultured microorganisms, members of this phylum are rarely isolated, with only 82 strains currently reported. While certain members contribute to biopolymer degradation, nitrogen fixation, and plant symbiosis, their physiological and genetic traits remain unclear. This study focused on the characterization and taxonomic classification of a novel ultramicrobacterium within the *Verrucomicrobiota* phylum.

1. Morphological, physiological, and phylogenetic characterization of a novel *Verrucomicrobiota* strain

This study characterized a recently isolated *Verrucomicrobiota* strain ASA1^T. The strain exhibited no growth on nutrient-rich media but formed colonies on lower nutrient media (R2A and 1/100× PYG) after 8–14 days of incubation. The colonies appeared whitish-yellow and smooth, while the cells were ultramicrobacterial (<0.1 μm³), coccoid to oval in shape (0.05–0.07 μm³), and formed cluster-like aggregates (>10 μm). ASA1^T was Gram-negative, aerobic, motile, and tested negative for catalase and oxidase activities. Growth occurred within the temperature range 20–40°C (optimum 30°C) and pH 7–11 (optimum 8.4), and NaCl tolerance 0–3% (optimum 0%). Chemotaxonomic analysis revealed anteiso-$C_{15:0}$ (53%) as the major fatty acid, followed by $C_{16:0}$ (15%), with menaquinone-7 (MK-7) as the sole respiratory quinone. The strain showed a limited substrate utilization and enzymatic activity spectrum. Phylogenetic analysis indicated that strain ASA1^T represents a novel lineage within the family *Alterococcaceae* of the class *Opitutia*, based on low 16S rRNA gene sequence similarity (<93%).

2. Genomic characterization and potential distribution of a novel *Verrucomicrobiota* strain

Strain ASA1^T exhibited low average nucleotide identity (ANI) values (<77%) with its closest described type of strain. Additionally, average amino acid identity (AAI) values were below 54%, far below the accepted genus-level threshold (60–65%), further supporting its designation as a new genus. Clusters of Orthologous Groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of strain ASA1^T showed gene enrichment in carbohydrate metabolism, cell wall biogenesis, energy production, and signal transduction, alongside a notable proportion of genes with unknown functions. Habitability analysis revealed that strain ASA1^T and its relatives are widely distributed across diverse terrestrial and aquatic environments. Genome mining identified six biosynthetic gene clusters (BGC), including non-ribosomal peptide synthetase (NRPS), polyketide synthase (PKS), and arylpolyene types, all showing low similarity (<1%) to known clusters. These findings suggest the potential of ASA1^T as a source of novel secondary metabolites.

3. Taxonomic Proposal: *Congregicoccus parvus* gen. nov., sp. nov.

Based on comprehensive phylogenetic, genomic, and chemotaxonomic analysis, a novel genus and species is proposed within the family *Alterococcaceae* of the phylum *Verrucomicrobiota*. The type of strain, ASA1^T (= JCM 36569^T = NCIMB 15508^T), represents *Congregicoccus parvus* gen. nov., sp. nov. The complete genome is 5.82 Mbp with a G+C content of 65.6% and is publicly available in DDBJ/ENA/GenBank under accession number AP028972.1 (BioProject: PRJDB16742, BioSample: SAMD00647589, DRA: DRA017385).

This study presents a comprehensive characterization of strain ASA1^T, a novel member of the family *Alterococcaceae* within the class *Opitutia*. Its physiological traits and genomic features provide a valuable understanding of the biology of ultramicrobacteria from this bacterial lineage. Notably, the findings of this study reveal a previously untapped source of bioactive compounds

and offer new insights into the isolation of uncultured microorganisms within the phylum *Verrucomicrobiota*.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Souichiro Kato, for his ongoing support, valuable guidance, and encouragement throughout my PhD journey. I also wish to express my sincere appreciation to Dr. Wataru Kitagawa, who formally became my supervisor following Dr. Kato's move to the University of Tokyo and subsequent retirement from AIST. I am especially thankful to my first supervisor, Dr. Isao Yumoto, for allowing me to pursue this study before he also retired from AIST.

I am deeply grateful to my co-supervisors, Dr. Kensuke Igarashi and Dr. Ryosuke Nakai, for their insightful feedback, kind assistance, and steady support during my research. I also thank the members of my thesis evaluation committee for their thoughtful comments and suggestions.

I sincerely appreciate the generous support provided by Dr. Naoki Morita, especially with fatty acid analysis. Special thanks go to Dr. Teruo Sone, Dr. Kyosuke Yamamoto, and Dr. Yoshitomo Kikuchi for their valuable input and guidance. I am also thankful to Dr. Shusei Kanie and Dr. Maiko Furubayashi for their constant encouragement and support.

I cordially thanks to Shono Munekawa, Kota Ishigami, , Ai Miura, Mika Yamamoto, Ms. Chihiro Misaki, Kasai, Masaru Sakata, Xie Ruoyun, Putcha Jyothi Priya, Antoine-Olivier Lirette, and Nowshin Farjana for their kind assistance with daily lab work. Also, my sincere gratitude to the researchers and personnel at AIST, Hokkaido, for always being so kind.

I thank Almighty Allah for his infinite mercy, guidance, and strength. My heartfelt appreciation goes to my wife for her unwavering support, love, and patience. I also thank my parents, elder sister, and all my friends and colleagues for their encouragement and kindness.

Lastly, I would like to gratefully acknowledge the MEXT scholarship and Hokkaido University, Japan, for supporting my PhD studies.

LIST OF FIGURES

Figure 1.1. The diversity of bacterial and archaeal phyla across terrestrial environments… 3

Figure 1.2. Evolutionary relationships within the PVC superphylum…….. 5

Figure 1.3. Types of filterable microorganisms found in environmental sources……. 7

Figure 1.4. Diversity of soil-derived genomes and their biosynthetic potential…… 10

Figure 1.5. TEM image of *Verrucomicrobium spinosum*……. 12

Figure 2.1. Growth characteristics of ASA1^T on different media types…… 27

Figure 2.2. Colony morphology of ASA1^T on R2A medium…….. 27

Figure 2.3. Phase-contrast, fluorescence, and SEM images of strain ASA1^T…… 28

Figure 2.4. Gram-stain comparison of ASA1^T and reference strains…… 30

Figure 2.5. KOH test results for ASA1^T and reference strains…… 31

Figure 2.6. Catalase test results for ASA1^T and reference strains…… 32

Figure 2.7. Oxidase test results for ASA1^T and reference strains……. 33

Figure 2.8. Motility test results for ASA1^T and reference strains……. 33

Figure 2.9. Quinone profile of ASA1^T determined by UPLC…… 38

Figure 2.10. Genomic DNA quality and 16S rRNA gene PCR products……. 39

Figure 2.11. An NJ tree illustrating the relationship between strain ASA1^T and type strains of class *Opitutia* based on their 16S rRNA gene sequences…….. 41

Figure 2.12. An ML tree was constructed to illustrate the relationship between strain ASA1^T and type strains of the class *Opitutia*……. 41

Figure 3.1. Schematic representation of the circular genome map of the ASA1…….. 49

Figure 3.2. Distribution of protein-coding genes of strain ASA1^T across COG functional categories…….. 50

Figure 3.3. KEGG functional classification of strain ASA1^T based on KO (KEGG Orthology) assignments generated by BlastKOALA…… 51

Figure 3.4. Phylogenomic tree with BGC count……... 52

Figure 3.5. Phylogenomic tree generated using 120 concatenated conserved proteins…… 53

Figure 3.6. Habitat preference prediction of strain ASA1^T and its related *Verrucomicrobiota* members based on ProkAtlas……. 54

Figure 3.7. Correlation between rRNA operon copy number and incubation time in 19 related genomes of *Verrucomicrobiota*…… 55

Figure 3.8. Graphical representation of BGC regions of ASA1 genome…….. 56

LIST OF TABLES

Table 2.1. Differentiation between strain ASA1^T and other closely related representatives of *Opitutales*. …… 34

Table 2.2. Major (>1% of total) fatty acids of strain ASA1^T……... 37

Table 3.1. Summary of predicted BGC in the strain ASA1 genome by antiSMASH (v7.0.1) ……. 57

LIST OF ABBREVIATIONS

PVC	*Planctomycetota*, *Verrucomicrobiota*, and *Chlamydiota*
UMB	Ultramicrobacteria
BGC	Biosynthetic gene clusters
CPR	Candidate phyla radiation
DPANN	An archaeal superphylum
PKS	Polyketide synthase
NRPS	Non-ribosomal peptide synthetase
T3PKS	Type III polyketide synthase
TEM	Transmission electron microscopy
PYG	Peptone yeast glucose
R2A	Reasoner's 2A agar
R2B	Reasoner's broth
LB	Luria-Bertani
TSA	Tryptic soy agar
PS	Phosphate and agar autoclaved separately
NaCl	Sodium chloride
DAPI	4',6-diamidino-2-phenylindole
SEM	Scanning electron microscopy
KOH	Potassium hydroxide
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
H_2O_2	Hydrogen peroxide

PBS	Phosphate-buffered saline
PFA	Paraformaldehyde
OD	Optical density
FAME	Fatty acid methyl ester
GC	Gas chromatograph
UPLC	Ultra-performance liquid chromatography
BLAST	Basic local alignment search tool
MiBIG	Minimum information about a biosynthetic gene cluster
PS	Autoclaving phosphate and agar separately
PCI	Phenol-chloroform-isoamyl alcohol
SDS	Sodium dodecyl sulfate
NCBI	National center for biotechnology information
MEGA	Molecular evolutionary genetics analysis
PCR	Polymerase chain reaction
NJ	Neighbor-joining
ML	Maximum-likelihood
MK	Menaquinone
GTDB	Genome taxonomy database
NCBI	National center for biotechnology information
SMRT	Single molecule real-time sequenceing
HiFi	High fidelity
DFAST	Fast annotation and submission tool
ANI	Average nucleotide identity

AAI	Average amino acid identity
COG	Clusters of orthologous groups
KEGG	Kyoto encyclopedia of genes and genomes
KO	KEGG oethology
BV-BRC	Bacterial and viral bioinformatics resource center
CDSs	Coding sequences
CRISPR	Clustered regularly interspaced short palindromic repeats
MAGs	Metagenome-assembled genomes
IMNGS	Integrated microbial next generation sequencing
MiBIG	Minimum information about a biosynthetic gene cluster
JCM	Japan collection of microorganisms
NCIMB	National collection of industrial, food and marine bacteria

Chapter I

General Introduction

1.1. Microbial diversity and its ecological significance

Microorganisms constitute the vast majority of Earth's biodiversity and are fundamental to the functioning of ecosystems. These include bacteria, archaea, fungi, protists, and viruses, many of which inhabit environments ranging from deep oceans to plant roots. These microbes represent a major component of global biodiversity and are critical to essential functions such as organic matter decomposition, nutrient recycling, and soil structure formation (Powlson et al. 2001; Anthony et al. 2023; Lavergne et al. 2024). They also enhance plant health by promoting nutrient uptake, producing growth-stimulating compounds, and suppressing soil-borne pathogens (Kapoor et al. 2024). Soil, in particular, represents one of the richest microbial habitats (Figure 1.1), harboring up to 10 billion microbial cells per gram and thousands of phylogenetically diverse taxa (Torsvik and Øvreås 2002). This remarkable diversity is shaped by spatial heterogeneity, physicochemical gradients, and complex plant-microbe interactions in microhabitats such as the rhizosphere and soil aggregates. While the role of microbes in maintaining ecosystem health is well accepted, the precise contribution of microbial diversity to these functions remains debated. A better understanding of this relationship is vital for advancing sustainable ecosystem management and conservation efforts.

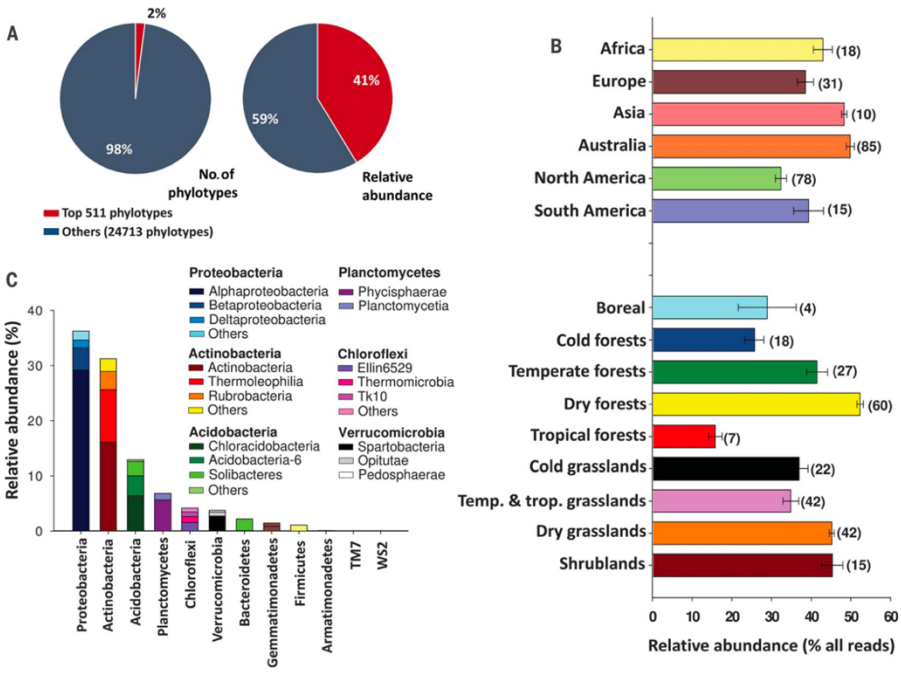

Figure 1.1. The diversity of bacterial and archaeal phyla across terrestrial environments (adapted from (Delgado-Baquerizo et al. 2018)). (A) Proportion and relative abundance of 16S rRNA gene sequences from dominant bacterial phylotypes compared to all others. (B) The mean relative abundance of the most common phylotypes in different ecosystems and continents. (C) A breakdown of the main bacterial phylotypes found by taxonomic group.

However, much of this microbial diversity remains unexplored, as the majority of environmental microorganisms are not readily cultivable under standard laboratory conditions, a phenomenon often referred to as "microbial dark matter" (Lok 2015). Cultivation-independent techniques such as metagenomics, single-cell genomics, and stable isotope probing have begun to unveil the genomic potential and ecological roles of these elusive lineages (Solden et al. 2016; Lloyd et al. 2018). Among the underrepresented phyla is *Verrucomicrobiota*, which is frequently detected in environmental DNA datasets yet poorly understood due to the scarcity of cultured representatives. Gaining access to these rarely cultured microbes is critical for refining the tree of

life and unlocking novel functions with potential applications in biotechnology, agriculture, and environmental science (Hug et al. 2016; Overmann et al. 2017). Recent efforts focusing on the isolation and characterization of such taxa offer a promising path toward expanding our understanding of microbial diversity and its relevance in natural systems.

1.2. The *Verrucomicrobiota* phylum: current knowledge and challenges

Amid the vast diversity of bacterial life on Earth, the phylum *Verrucomicrobiota* emerges as a globally pervasive yet remarkably underexplored lineage (Bergmann et al. 2011). Despite its presence across a variety of habitats, our understanding of this phylum lags far behind that of more extensively studied groups such as *Proteobacteria* or *Actinobacteriota* (previously named as *Actinobacteria*) (Fuerst 2019). This gap largely stems from cultivation challenges, which make them difficult to isolate in laboratory settings (Petroni et al. 2000; Freitas et al. 2012). *Verrucomicrobiota* belongs to the PVC superphylum (Figure 1.2), alongside *Planctomycetota* and *Chlamydiota*, which are noted for having unique cellular features, including compartmentalized cytoplasm, internal membrane structures, and unusual cell wall architecture (Fuerst and Sagulenko 2011; Devos and Ward 2014). These features challenge traditional bacterial taxonomy and reveal the evolutionary transition between prokaryotic and eukaryotic cellular organization, making PVC superphylum members useful for evolutionary microbiology research.

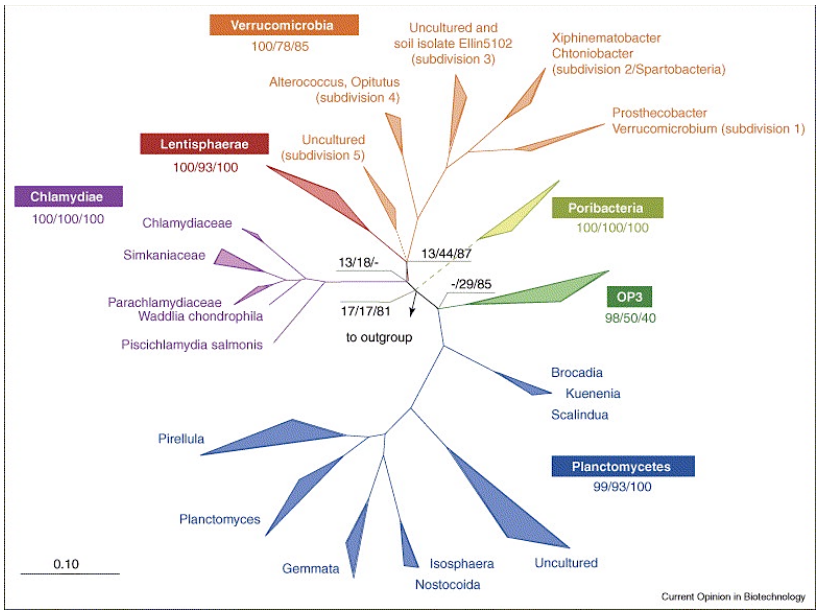

Figure 1.2. Evolutionary relationships within the PVC superphylum, including *Verrucomicrobia*, *Planctomycetes*, and *Chlamydiae* (adapted from (Wagner and Horn 2006)).

The term "*Verrucomicrobiota*" comes from the wart-like surface that was seen on the early members, but not all species have this feature (Fuerst 2019). Since then, phylogenomic research has shown that this phylum represents a deep-branching lineage within bacteria. It includes a wide range of classes, such as *Verrucomicrobiae*, *Opitutia*, *Spartobacteria*, and *Methylacidiphilae*. Each lineage has a lot of different roles in the environment and metabolic features. *Verrucomicrobiota* that live in soil are especially interesting since they can live in environments with limited nutrients. They usually grow slowly, possess small cell sizes, and have genomes that are streamlined (Bergmann et al. 2011; Nixon et al. 2019). Genomic studies show that they can break down complex polysaccharides, which shows their functions in nutrient recycling and breaking down organic matter (Camargo et al. 2023). But the majority of these findings come from metagenomic and single-cell research because most of them have not been cultivated yet. Cultivation is well-known for being hard, and it often needs media with low nutrients, a pH buffer,

or chemicals that are well-defined, as well as long incubation times. If colonies do form, they are usually small or weak. This challenge exemplifies the “great plate count anomaly,” where the observed microbial diversity in environmental samples far exceeds what can be cultured in the laboratory (Reguera 2016). These drawbacks have made researchers interested in other ways of culturing cells, such as microfluidics, *in situ* diffusion chambers, and co-culture approaches that better resemble natural circumstances (Ben-Dov et al. 2009; Vartoukian 2016; Dai et al. 2025).

Even with increased interest, there are still a number of things that make it hard to fully describe *Verrucomicrobiota*. Taxonomic classification is often changed as new genomes are found. Many environmental sequences are still grouped as "uncultured," which makes it hard to understand how species evolved and how they interact with each other (Chuvochina et al. 2019). Functional annotation is another issue; many genes are still called hypothetical proteins since there are no homologs, which makes it hard to assess their metabolic potential. Recent work has found biosynthetic gene clusters (BGC) and carbohydrate-active enzymes (Crits-Christoph et al. 2018), but experimental proof is still hard to come by. Also, the lack of genetic tools makes functional studies harder, which makes it hard to change genes and look at regulatory pathways. To unlock the ecological and biotechnological potential of this phylum, integrative approaches and cultivation methods are required.

1.3. Significance of ultramicrobacteria and slow growers

Ultramicrobacteria (UMB) are cells that are less than 0.1 μm^3 in size (Duda et al. 2012). These tiny organisms are often found in places with few nutrients, like glacial ice, underground aquifers, marine gyres, and the rhizosphere. Their small genomes and energy-efficient metabolisms help them stay alive (Giovannoni et al. 2005). One interesting thing about UMB is that it can go through

regular 0.2 µm filters. This is why they were left out of microbial surveys that looked at bigger cells in the past (Nakai 2020). As a result, these microbes remained underexplored for decades.

The word "ultramicrobacteria" was first reported in the early 1980s to describe cells with a diameter of less than 0.3 µm (Torrella and Morita 1981). However, the current definition is based on cells with a volume of less than 0.1 µm³ (Duda et al. 2012). Filterable UMB are not a uniform group, and they can be broadly classified into five categories (Figure 1.3): obligate ultramicrobacteria, which maintain small size throughout their life cycle; facultative UMB, which can reduce their size under stress; stress-induced ultramicrocells, formed in response to environmental pressure; slender filamentous forms that pass through filters due to their shape; and ultra-small representatives of the CPR (Candidate Phyla Radiation) and DPANN (an archaeal superphylum) lineages (Nakai 2020). It is nevertheless important to tell the difference between the real UMB and forms that are only temporarily smaller.

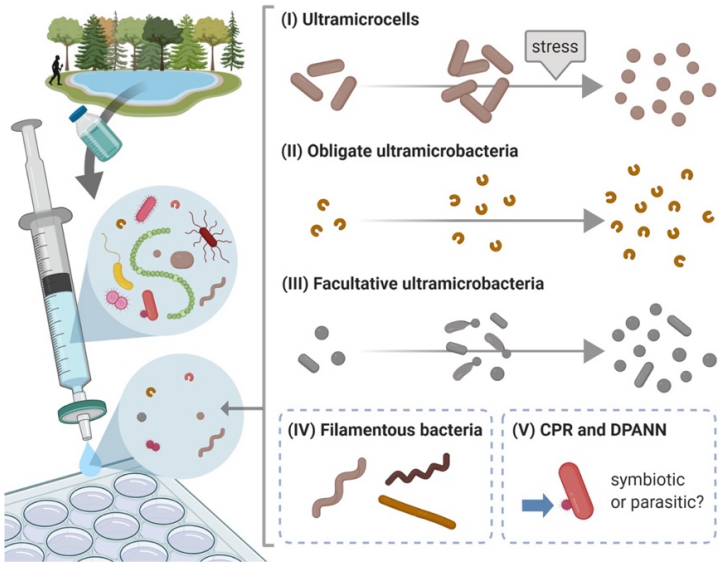

Figure 1.3. Types of filterable microorganisms commonly found in environmental samples (adapted from (Nakai 2020)). These include: (I) stress-induced ultramicrocells, (II) obligate

ultramicrobacteria, (III) facultative ultramicrobacteria, (IV) thin filamentous forms, and (V) ultra-small cells from CPR bacteria and DPANN archaea (indicated by arrow).

Slow-growing bacteria are characterized not by their size, but by extended lag phases and incubation times. They are adapted to habitats that are oligotrophic, stable, and have few resources. They value metabolic efficiency and persistence over quick replication. Many of them work as specialists, breaking down complicated organic substances or building syntrophic partnerships. For example, acidophilic methanotrophs oxidize methane in geothermal systems (Hou et al. 2008), while soil bacteria help keep ecosystems stable over time by slowly cycling nutrients and going dormant (Lennon and Jones 2011). Despite their ecological importance, slow growers and ultramicrobacteria remain underrepresented in culture collections.

However, recent advances are changing this landscape. Cultivation strategies such as dilution-to-extinction, microfluidic separation, and diffusion-based systems are facilitating the proliferation of slow-growing microorganisms formerly deemed unculturable. These methods elucidate the behavioral characteristics, ecological roles, and symbiotic relationships of these hidden organisms.

1.4. Importance of *Verrucomicrobiota* members

1.4.1 Medical relevance of *Akkermansia muciniphila* in the human gut

Among *Verrucomicrobiota*, *A. muciniphila* stands out for its pivotal role in the gut microbiome and its growing connection to host health. Commonly found in the intestines of humans and other animals, it thrives in the mucosal layer lining the gut, where it specializes in breaking down mucin, a glycoprotein secreted by goblet cells (Johansson et al. 2013; Derrien et al. 2017). This niche

provides both nutrients and a mildly oxygenated habitat that *A. muciniphila* is well-adapted to exploit using enzymes like sialidases, fucosidases, and sulfatases (Ottman et al. 2017a). It is a crucial species in keeping microbial balance because its metabolic byproducts feed other gut microorganisms (Chia et al. 2018). Interestingly, its benefits may go beyond action in living cells. Studies have revealed that heat-killed *A. muciniphila* can also enhance the metabolic health of obese mice. This is probably because of surface proteins like Amuc_1100 that affect immune responses and the intestinal barrier (Ottman et al. 2017b; Plovier et al. 2017). There is still a lot of clinical study to be done, but early evidence suggests that it could help treat obesity and related illnesses (Sheng et al. 2018). Overall, this single species illustrates how a well-adapted microbe can exert broad effects on host physiology, underscoring the biomedical potential hidden within the *Verrucomicrobiota* phylum.

1.4.2 Biotechnological and biosynthetic potential of *Verrucomicrobiota*

Verrucomicrobiota are getting more interest in biotechnology since they have unique enzymes and may produce beneficial compounds. The tyrosinase from *Verrucomicrobium spinosum* is a great example. It has an uncommon eukaryote-like C-terminal extension, which is something that bacteria do not usually have (Fairhead and Thöny-Meyer 2010). This enzyme has shown potential in making bioadhesives, biosensors, and protein cross-linking (Axambayeva et al. 2018). Genomic investigations have also shown that many uncultured verrucomicrobial lineages can break down carbohydrates. Enzymes like glycoside hydrolases and sulfatases show potential for processing biomass and making bioenergy (Fuerst 2019). Beyond enzymes, *Verrucomicrobiota* harbor BGC for secondary metabolites (Figure 1.4). For example, Candidatus *Didemnitutus mandela*, a symbiont of marine tunicates, encodes polyketide synthases (PKS) thought to produce cytotoxic

mandelalides for host defense (Murray et al. 2021). Besides, broader metagenomic analysis has also suggested that *Verrucomicrobiota* from marine environments may be rich sources of unexplored polyketide and nonribosomal peptide pathways (Vollmers et al. 2017). Taken together, these findings highlight the untapped metabolic and biosynthetic potential of *Verrucomicrobiota*.

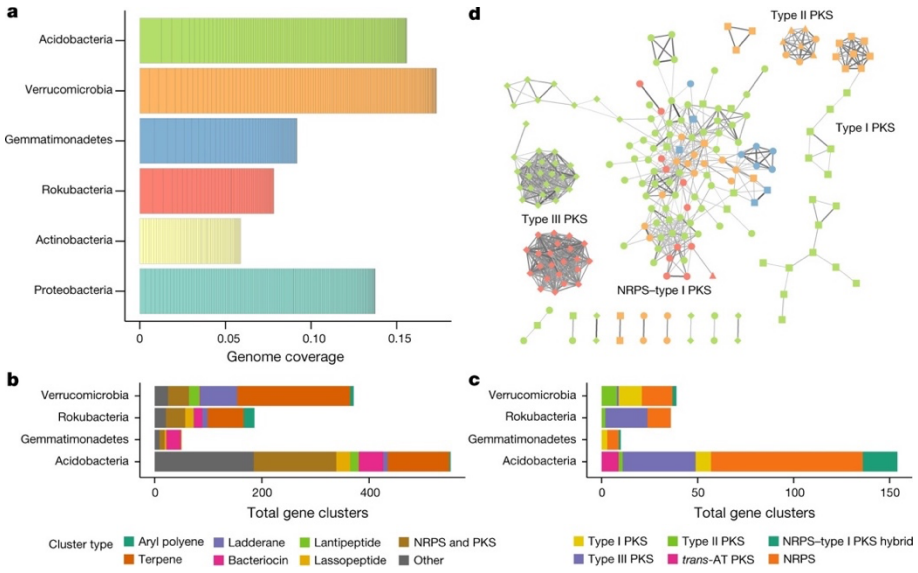

Figure 1.4. Diversity of soil-derived genomes and their biosynthetic potential (adapted from (Crits-Christoph et al. 2018)). (A) Average genome abundance across 60 soil samples, based on sequencing coverage. (B) BGC identified on contigs over 10 kb, categorized by predicted product types using antiSMASH. (C) Distribution of NRPS (non-ribosomal peptide synthetase) and PKS clusters from contigs >10 kb, grouped by phylum. (D) BGC similarity network showing connections between clusters that share genes.

1.4.3. Cell biology and evolutionary significance of *Verrucomicrobiota*

As members of the PVC superphylum, they are different from most prokaryotic models because they have internal membrane systems and surface appendages. Early-described genera like *V.*

spinosum and *Prosthecobacter* make prosthecae, which are cytoplasmic extensions that are hypothesized to help with surface adhesion and nutrient uptake. These have even been employed for selective isolation (Schlesner 1987; Hedlund et al. 1997). Electron microscopy has shown internal compartments in species such as *V. spinosum*, *Prosthecobacter dejongeii*, and *Chthoniobacter flavus*, resembling the "pirellulosome" structures seen in *Planctomycetes*, hinting at early forms of cellular compartmentalization (Figure 1.5) (Lee et al. 2009). These traits suggest that there are evolutionary similarities with eukaryotic cells, although more imaging is needed (Fuerst and Sagulenko 2011). The chemical composition of the cell wall can also be variable. Most members are Gram-negative and have normal peptidoglycan layers, but some marine species, such as *Pelagiococcus* and *Cerasicoccus*, do not seem to have any detectable peptidoglycan. This could be because they have very little or highly modified peptidoglycan (Yoon et al. 2007b, a; Anders et al. 2015). But the fact that they have peptidoglycan-related genes in their genomes suggests that they still have some of the old cell wall parts. These properties make *Verrucomicrobiota* a very interesting group to study the evolution of bacterial cell complexity.

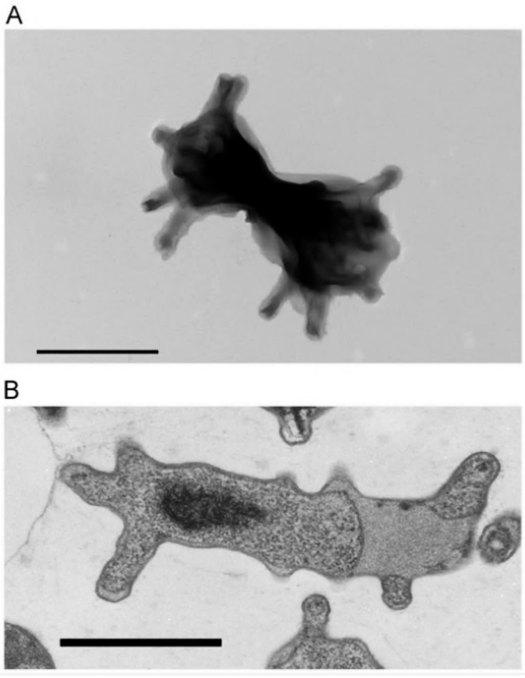

Figure 1.5. (A) The transmission electron microscopy (TEM) image of *V. spinosum* displays numerous tubular prosthecae extending from all areas of the cell surface (adapted from Lee et. al., 2009). (B) The sectioned TEM image of cryo-fixed *V. spinosum* emphasizes internal, membrane-bound compartments abundant in ribosomes, ribosome-free zones, and the structural continuity of prosthecae with the cytoplasm and cell envelope. Scale bar: 1 μm.

1.4.4. Ecological roles in nutrient and carbon cycling

Verrucomicrobiota are emerging as important players in global biogeochemical cycles, particularly in carbon and nitrogen transformation across soil and aquatic ecosystems. In soils, they thrive in oligotrophic habitats like boreal forests, tundra, and acidic peatlands. For example, *C. flavus*, a *Spartobacteria* member, contributes to plant polysaccharide breakdown through a wide array of CAZymes, influencing carbon turnover and soil health (Kant et al. 2011). In aquatic systems, genera such as *Luteolibacter* and *Rubritalea* bloom during phytoplankton die-offs, helping recycle algal glycans and drive microbial succession (Herlemann et al. 2013). In geothermal soils, thermoacidophilic methanotrophs like *Methylacidiphilum infernorum* and *Methylacidimicrobium tartarophylax* oxidize methane under extreme pH and temperature conditions, reducing greenhouse gas emissions (Schmitz et al. 2021). Together, these examples showcase the ecological versatility of *Verrucomicrobiota*, from soil to sea to volcanic hotspots, and their growing relevance in maintaining ecosystem balance.

1.5. Dissertation aims and objectives

This study aims to characterize and formally classify a novel soil-derived *Verrucomicrobiota* and is divided into three studies:

- **Study 1:** To characterize a novel soil-derived *Verrucomicrobiota* strain through morphological, physiological, and phylogenetic analysis.
- **Study 2:** To elucidate the genomic features of the newly isolated strain and clarify its taxonomic position through whole-genome sequencing and habitability analysis.
- **Study 3:** To establish the formal taxonomic status of the isolated strain using a polyphasic approach, leading to the proposal of a novel genus and species.

Chapter II

Study I

Morphological, physiological, and phylogenetic characterization of a novel *Verrucomicrobiota* strain

2.1. Introduction

Characterizing novel bacterial isolates is critical for understanding their ecological activities, physiological characteristics, and biotechnology potential. While fast-growing bacteria are extensively studied, many slow-growing species are overlooked due to cultivation challenges. The phylum *Verrucomicrobiota*, in particular, remains underrepresented in culture collections and genome databases. This chapter focuses on the morphological, physiological, and phylogenetic characterization of a recently isolated soil-derived *Verrucomicrobiota* strain.

2.2. Materials and Methods

2.2.1. Strain ASA1^T and culture conditions

Strain ASA1^T, a member of the phylum *Verrucomicrobiota*, was recently isolated from garden soil collected at a depth of 5–15 cm at the National Institute of Advanced Industrial Science and Technology (AIST) Hokkaido Center, Japan (43.0201° N, 141.4182° E). The isolation procedure involved suspending 1 g of soil in 9 mL of sterile 0.9% NaCl, followed by serial dilutions (10^{-1} to 10^{-6}), and plating 100 µL aliquots onto 100× diluted PYG agar (peptone, yeast extract, and glucose; phosphate and agar autoclaved separately [PS method]) at pH 8.4 (Kato et al. 2018). Plates were incubated at 30°C in the dark for 14 days, and small colonies were selected to target slow-growing bacteria. These colonies were further transferred through three rounds of streaking on R2A medium (DAIGO; Nihon Pharmaceutical, Tokyo). To determine optimal growth, various media, including Luria-Bertani (LB) and tryptic soy agar (TSA) were tested. The compositions of all media used are listed in Appendix 2.1–2.4.

2.2.2. Physiological and biochemical characterization

Strain ASA1^T was maintained on solid R2A medium for routine cultivation. For liquid cultures, the strain was grown in R2B medium (without agar). Unless specified otherwise, all incubations were conducted at 30°C and pH 8.4. For microscopy, exponentially growing ASA1^T cells were collected from liquid cultures (7 days) and placed on a glass slide under a coverslip. Light microscopic images were obtained using an Axio Imager microscope (Carl Zeiss). For DAPI (4',6-diamidino-2-phenylindole) staining, bacterial cells were washed twice with phosphate-buffered saline (PBS, pH 7.4) and fixed with 4% paraformaldehyde (PFA) at room temperature for 10 min. The fixed cells were incubated with DAPI solution (1 µg/mL final concentration) in the dark for 10 min. After staining, cells were washed twice with PBS, mounted on glass slides, and observed under the fluorescence mode of the microscope. Digital images were captured and processed using the associated imaging software.

2.2.2.1 Scanning electron microscopy (SEM) analysis

For SEM, strain ASA1^T cells were collected from 7-day cultures grown in R2B medium. The cells were initially fixed in a mixture of 2% PFA and 2% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7.4) at 4°C for 3 days. Post-fixation steps included treatment with 1% tannic acid at 4°C for 2 h and several washes (four times) in cacodylate buffer, followed by further fixation using 2% osmium tetroxide at 4°C for 2 h. The samples were then dehydrated through a graded ethanol series (50%, 70%, 90%, and multiple washes in 100%), before transitioning to tert-butyl alcohol and freeze-drying. Dried specimens were coated with a 30 nm osmium layer using an osmium plasma coater (NL-OPC80A, Nippon Laser & Electronics Laboratory, Japan). Imaging was

conducted with a JSM-7500F SEM (JEOL Ltd., Tokyo) operated at 3.0 kV, and the resulting images were digitally saved for analysis.

2.2.2.2. Gram-staining test

Gram staining of strain ASA1^T was performed following the manufacturer's instructions using a Sigma-Aldrich Kit (St. Louis, MO, USA). A single colony of ASA1^T was taken from an 8-day R2A culture plate and smeared with one drop of distilled water onto a clean glass slide and spread evenly. The smear was air-dried and subsequently heat-fixed by slightly passing the slide through a Bunsen burner flame, so attaching it to the slide. The smear was flooded with crystal violet stain and left for 1 min, then rinsed gently with water. Gram-iodine solution was applied to the smear and left for 1 min to form a complex with the crystal violet, then rinsed with water. The slide was washed with 95% ethanol for 10–30 seconds to remove the stain, then rinsed with water immediately. The prepared smear was counterstained with safranin for approximately 30 seconds, followed by a gentle rinse with water. After blotting the slide dry, it was examined under an oil immersion lens (100× magnification). *Bacillus subtilis* served as the Gram-positive control, while *Escherichia coli* was used as the Gram-negative reference. Gram-positive cells retained crystal violet, resulting in a purple appearance, while Gram-negative cells absorbed safranin, leading to a pink appearance.

2.2.2.3. Potassium hydroxide test (KOH) test

The Gram reaction of the strain was further assessed using the 3% KOH solubility test (Halebian et al. 1981). A fresh KOH solution was prepared by dissolving 1.5 g of KOH in distilled water to a final volume of 50 mL. A drop of this solution was placed on a glass slide, and cells from an 8-

day R2A culture were mixed into the drop. The reaction was observed within 1 min for viscosity changes. *B. subtilis* and *E. coli* served as the Gram-positive and Gram-negative controls, respectively.

2.2.2.4. Catalase test

Catalase activity of the strain was assessed by placing a loopful of bacterial biomass, collected from an 8-day R2A culture, onto a clean glass slide (Reiner 2010). The working solution was made by diluting 1 mL of 30% stock H_2O_2 with 9 mL of sterile distilled water and used immediately due to the unstable and reactive nature of H_2O_2. A freshly prepared 3% H_2O_2 solution was then added dropwise to the cell mass to observe the reaction. The presence of bubbles indicated a positive reaction, while no bubbling signified a negative result. *B. subtilis* and *E. coli* were used as positive and negative controls, respectively.

2.2.2.5. Oxidase test

Oxidase activity of the strain was determined using a commercially available oxidase reagent (Sigma-Aldrich, St. Louis, MO, USA). A colony grown on 8-day R2A medium was gently smeared onto filter paper pre-soaked with the reagent, and any immediate color change was monitored. The appearance of a dark blue or purple coloration within 30 seconds was considered a positive reaction, indicating cytochrome c oxidase activity, while no color change signified a negative result. *Oligoflexus tunisiensis* (Nakai et al. 2014) and *E. coli* served as positive and negative controls, respectively.

2.2.2.6. Motility test

Motility of strain was evaluated using semi-solid soft agar prepared with 0.3% R2A medium. To prepare 100 mL of medium, 0.33 g of R2B powder and 0.3 g of agar were dissolved in distilled water and autoclaved at 121°C for 20 min. After cooling to around 50–55°C, 5 mL of the medium was dispensed into sterile test tubes. A colony from a 7-day R2B culture was stab-inoculated into the center of the agar and incubated at 30°C for 8 days. To validate the assay, *B. subtilis* was used as a positive control. *Mycobacterium austroafricanum* was employed as a negative control. Diffuse, outward (extending) growth indicated motility, while growth restricted to the stab line signified a non-motile phenotype.

2.2.2.7. Cell volume measurement

Cell volume was estimated following a previously established method (Andersson et al. 1986), where cell dimensions were measured microscopically, and the volume was calculated based on a standard geometric formula for bacterial cell shapes.

2.2.2.8. Growth optimization by temperature, pH, and salt

Cells were cultured in R2B medium, harvested during the exponential growth phase, and then processed for the following growth optimization tests.

1. The temperature range for growth was evaluated by inoculating ASA1^{T} onto R2A medium and incubating at 15, 20, 25, 30, 35, 37, 40, and 45°C. Growth was recorded based on colony formation after 7–14 days of incubation.

2. The pH tolerance range was determined by cultivating ASA1^T in R2A medium adjusted to pH 5.0, 7.0, 8.4, 9.0, and 11.0 after autoclaving. Growth was recorded based on colony formation after 7–14 days of incubation.
3. Salt tolerance was assessed by cultivating ASA1^T on R2A medium supplemented with NaCl at concentrations of 0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, and 5.0 (%, w/v). Growth was recorded based on colony formation after 7–14 days of incubation.

2.2.2.9. Substrate utilization test

The substrate utilization assay was performed to determine the carbon source preferences of ASA1^T.

1. Exponential-phase cultures 1% (v/v) grown in R2B medium for 7 days were inoculated into 5 mL R2B in 50 mL culture tubes.
2. These tubes were subsequently incubated at 30°C and 100 rpm. The filter-sterilized substrates were added so that the final concentration was 10 mM, except for starch, which was added at a final concentration of 0.1%.
3. The test substrates included D-glucose, D-fructose, D-galactose, D-mannose, sucrose, D-ribose, D-xylose, L-arabinose, D-mannitol, D-sorbitol, D-maltose, D-cellobiose, and starch.
4. Cell growth was measured with a spectrophotometer set to 600 nm

2.2.2.10. Enzyme activity test

Enzyme activity profiles and physiological characteristics were examined using API 20 NE and API ZYM test strips (bioMérieux, Marcy-l'Étoile, France) following the manufacturer's instructions. The API test strips were inoculated with an overnight culture of ASA1^T adjusted to

McFarland standard 2.0, incubated at 30°C, and examined for color changes or precipitate formation indicative of enzymatic activity.

2.2.3. Chemotaxonomic profiling

2.2.3.1. Fatty acid analysis

Cells were grown in R2B medium at 30°C for 7 days, then harvested and directly subjected to fatty acid methyl ester (FAME) preparation. Methylation was performed by treating the biomass with 10% (v/v) acetyl chloride in methanol at 100°C for 3 h. After cooling, the resulting FAMEs were extracted three times using n-hexane to ensure efficient recovery of the methylated compounds. The extracted FAMEs were analyzed using a gas chromatograph (GC) (Shimadzu GC-1700 system, Kyoto, Japan), equipped with a BPX70 capillary column (0.22 mm × 50 m, SGE Analytical Science) and a flame ionization detector. Helium served as the carrier gas. The injection and detector ports were maintained at 260°C, while the column oven was programmed to rise from 160°C to 240°C at a rate of 10°C per/min, then held at 240°C for 7 min (Kato et al. 2019). The fatty acid composition was determined by comparing the retention times of the detected FAMEs with known standards.

2.2.3.2. Quinone analysis

The respiratory quinone profile of strain ASA1^{T} was examined using ultra-performance liquid chromatography (UPLC) to support chemotaxonomic identification (Nakai et al. 2021). Quinones were extracted from the bacterial biomass following an established protocol, ensuring high recovery and minimal degradation (Tamaoka et al. 1983). The extracted compounds were analyzed using an Acquity UPLC H-Class system (Waters, Milford, MA, USA) equipped with a photodiode

array detector (UPPDA-E, Waters) and an Eclipse Plus C18 reverse-phase column (2.1 × 150 mm; 1.8 µm pore size, Agilent Technologies, USA). Chromatographic separation and compound detection were monitored using MassLynx (v4.2) software. Identification of quinone species was based on retention times and spectral features compared to reference standards.

2.2.4. DNA extraction by using the phenol-chloroform method

The genomic DNA of the bacterial isolate was extracted using the phenol-chloroform method, following the protocol described by Sambrook et al. (1989) (Sambrook et al. 1989), with slight modifications to enhance DNA purity. The procedure is briefly described below.

Reagents

- TEN solution

  [For 30 mL, add 0.18 g of NaCl into Tris-EDTA (TE) buffer]

- Lysis solution

  [For 30 mL, add 0.06 g of lysozyme (2 mg/mL) and 6 µL of RNase A (20 µg/mL) directly into the previously prepared TEN solution]

- 10% sodium dodecyl sulfate (SDS)

  [For 30 mL, add 3 g of SDS into Milli-Q water and dissolve completely]

- Proteinase K
- PCI (Phenol:chloroform:isoamyl alcohol- 25:24:1)
- CI (Chloroform:isoamyl alcohol- 24:1)
- 2-propanol
- 70% and 100% Ethanol

- TE buffer

Procedure

1. Bacterial cells were harvested from an overnight culture by centrifugation at 12,000 × g for 5 min at 4°C, and the supernatant was discarded.
2. 4 ml of lysis solution was added to the pellet, gently resuspended, and incubated for 30 min at 37°C. 200 µL of proteinase K was added, followed by incubation for 30 min at 37°C.
3. 400 µL of 10% SDS solution was added, gently mixed, and incubated for 60–90 min at 37°C.
4. The pellet was ensured to be completely dissolved, and the mixture became transparent.
5. The mixture was then transferred into eight 1.5 mL tubes, each containing at least 500 µL.
6. 500 µL of PCI solution was added to each tube, mixed well, and centrifuged for 10 min at 15,000 rpm at 4°C.
7. 500 µL of the supernatant was carefully transferred into another eight tubes, avoiding the lower phase. The PCI steps were repeated at least twice.
8. The supernatant from each tube was then collected into two tubes, ensuring the volume did not exceed 700 µL per tube.
9. One-tenth volume of CI was added, mixed well, and centrifuged for 10 min at 15,000 rpm at 4°C.
10. The supernatant from each tube was collected into a new tube, and an equal volume of 2-propanol was added and gently mixed.
11. A white precipitate of DNA became visible, and the solution was centrifuged for 10 min at 15,000 rpm at 4°C. The supernatant was carefully discarded using a pipette.

12. 1 mL of 70% ethanol was added, gently mixed, and centrifuged for 10 min at 15,000 rpm at 4°C. The supernatant was discarded, and this step was repeated at least twice.
13. The supernatant was again discarded, followed by the addition of 1 mL of 100% ethanol, gentle mixing, and centrifugation for 10 min at 15,000 rpm at 4°C.
14. The supernatant was discarded, and the tubes were allowed to dry for 10 min.
15. 100 µL of TE buffer was added, and the tubes were left to stand for several min to allow the DNA to dissolve. The tubes were stored at 4°C overnight.
16. DNA solubilization was facilitated by incubating at 37°C for 5 min, and the purity was checked using NanoDrop spectrophotometry.

Qubit analysis

The concentration of the extracted genomic DNA was measured using a Qubit fluorometer according to the manufacturer's guidelines (Thermo Fisher Scientific). A working solution was freshly prepared by diluting the Qubit dsDNA reagent 1:200 in Qubit buffer (199 µL buffer + 1 µL reagent). For each DNA sample, 198 µL of the working solution was added to a 0.5 mL assay tube, followed by 2 µL of DNA. The mixture was gently vortexed and incubated at room temperature for 2 min to allow optimal dye-DNA binding. Calibration was performed using the Qubit DNA standards: 190 µL of working solution was mixed with 10 µL of standard 1 or 2 in separate tubes. After reading the standards to generate a calibration curve, DNA concentration was measured, and results were expressed in ng/µL. Samples outside the quantifiable range were diluted and remeasured. The integrity of the genomic DNA was assessed by 1% agarose gel electrophoresis, and samples were stored at -20°C for long-term use or at 4°C for immediate downstream experiments.

2.2.5. Phylogenetic tree analysis

To determine the phylogenetic placement of strain ASA1^T within the phylum *Verrucomicrobiota*, 16S rRNA gene sequencing approaches were employed. The amplified 16S rRNA gene (~1.5 kb) (two universal primer pairs 27F and 1492R) was verified by 1% agarose gel electrophoresis and sequenced using an Applied Biosystems BigDye Terminator (v3.1) Cycle Sequencing Kit (Thermo Fisher Scientific) with the 907R sequencing primer (5′-CCGYCAATTCMTTTRAGTTT-3′) on an Applied Biosystems 3500xL Genetic Analyzer (Thermo Fisher Scientific). The resulting sequence was compared with existing databases, including EzBioCloud and NCBI nt/nr, for taxonomic identification. A BLASTn search was performed using the NCBI nucleotide database (nt/nr) to identify the closest related type strains, as well as unclassified and uncultured environmental clones, based on sequence similarity (accessed on September 9, 2024). The 16S rRNA gene was also retrieved from the complete genome (AP028972.1). All sequences were aligned using MEGA11 (Tamura et al. 2021), applying the CLUSTAL method (Higgins and Sharp 1988), and evolutionary distances were estimated using the Kimura two-parameter model (Kimura 1979). Phylogenetic trees were constructed using the neighbor-joining (NJ) (Saitou and Nei 1987) and maximum-likelihood (ML) methods (Yang 2007), with 1,000 bootstrap replicates applied for statistical robustness.

2.3. Results and Discussion

2.3.1. Growth characteristics of strain ASA1^T

Strain ASA1^T exhibited a slow-growing phenotype, with colony formation in low-nutrient media by using PS media. After 14 days of incubation at 30°C, visible colonies were observed on R2A and 100× diluted PYG plates, while no growth was detected on nutrient-rich media such as LB

and TSA. Colonies typically appeared on low-nutrient media after 8 days, with complete development by day 14 (Figure 2.1). In liquid culture, the strain showed growth in R2B medium, with an OD_{600} increase observed after 7 days. The findings of both lower nutrient growth and long incubation times hypothesis that an oligotrophic lifestyle may possess starin ASA1^T lower nutrient requirements. The preference for low-nutrient media and requirement for prolonged incubation align with previous reports on *Verrucomicrobiota* members, which often exhibit slow growth rates and extended lag phases (Janssen et al. 2002; Davis et al. 2005). This is in contrast to fast-growing soil bacteria such as *Pseudomonas aeruginosa*, *Arthrobacter* sp., and *B. subtilis*, which typically form colonies within 24–48 h (Green et al. 1974; Luscombe and Gray 1974; Earl et al. 2008). The use of PS medium, which minimizes reactive oxygen species production during autoclaving, likely contributed to the successful cultivation of ASA1^T (Kato et al. 2018). However, the approach used here may also facilitate the isolation of other slow-growing taxa from environmental samples.

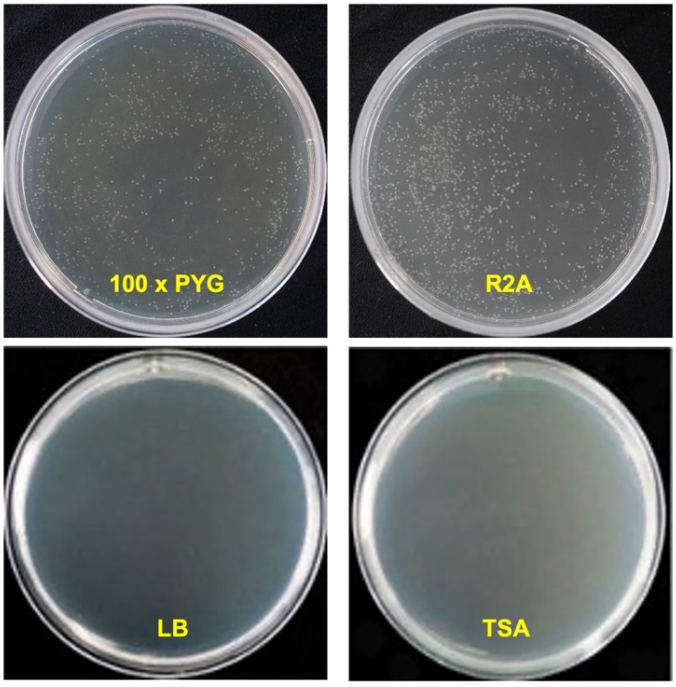

Figure 2.1. Growth characteristics of strain ASA1^T on different media types after 14 days of incubation at 30°C. Colonies were visible on low-nutrient media such as R2A and 100× diluted PYG, while no visible colonies were observed on nutrient-rich media, including LB and TSA.

2.3.2. Morphological, physiological, and biochemical characterization

2.3.2.1. Microscopic analysis

Following incubation on R2A plates at 30°C, colonies of ASA1^T displayed a whitish-yellow coloring and a smooth texture (Figure 2.2). Under optical microscopy, individual cells exhibited a coccus to oval form, measuring approximately $0.45 \times 0.45–0.60$ µm in diameter, which corresponds to estimated cell volumes of 0.05–0.07 µm³ (Figure 2.3 A, B). These measurements categorize ASA1^T as an ultramicrobacterium, according to the defined range of <0.1 µm³ cell volume.

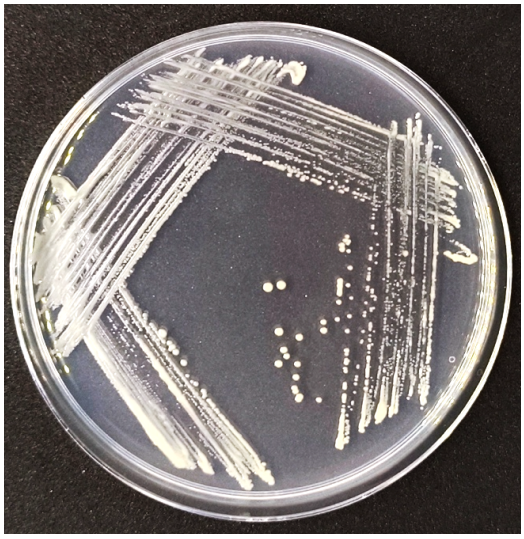

Figure 2.2. Colony morphology of strain ASA1^T on R2A medium. After 8 days of incubation at 30°C, colonies appeared circular, smooth, and whitish-yellow.

DAPI staining corroborated the cellular architecture, exhibiting strong and intact nucleoid structures in cells obtained from exponential-phase cultures (Figure 2.3 B, D). ASA1^{T} cells were often observed forming cluster-like aggregates exceeding 10 µm in diameter (Figure 2.3 C, D). SEM examination validated these structures, demonstrating that the aggregates comprised tightly packed coccoid cells organized in a clustered configuration (Figure 2.3 E, F). However, the precise composition of the matrix could not be determined under current experimental conditions.

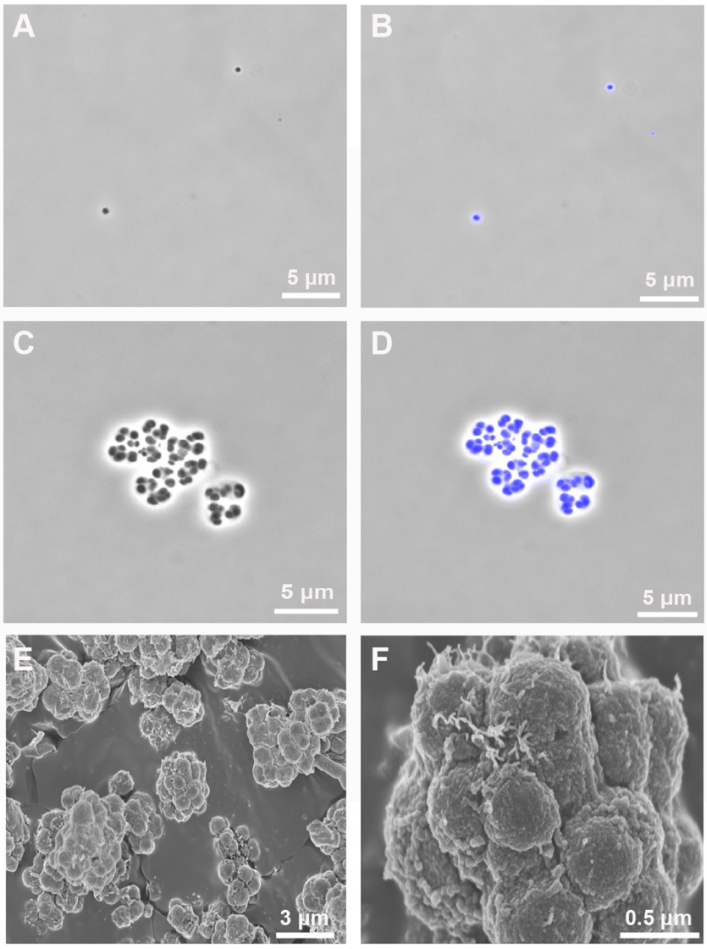

Figure 2.3. Phase-contrast, fluorescence (DAPI-stained), and SEM images of strain ASA1^{T}. Images (A–D) were observed after 7 days of incubation on R2B medium at 30°C. The image in (A) shows small coccus- and oval-shaped cells during the exponential phase (7 days), while that in (C) demonstrates cluster-like cell aggregation under the same conditions. (B) and (D) are

overlays of (A) and (C) with DAPI, respectively. (E) and (F) show SEM images of cluster-like aggregates formed after a longer incubation period (two weeks). Scale bars are 5 μm for images (A–D), 3 μm for image (E), and 0.5 μm for image (F).

The aggregation pattern noted in ASA1^T resembles the morphology of the acidobacterial strain SBC82^T, which develops saccular chambers and cellulose-rich extracellular matrix (Belova et al. 2022). These aggregated structures may represent an adaptive morphological trait that facilitates cellular organization or protection from environmental stresses. In contrast to previously characterized anaerobic verrucomicrobial ultramicrobacteria, strain ASA1^T was cultivated under aerobic conditions, indicating it as the first known aerobic ultramicrobacterium within *Verrucomicrobiota* to be isolated and morphologically characterized in a laboratory environment.

2.3.2.2. Gram-staining

Strain ASA1^T is Gram-negative, as displayed by its pink to reddish appearance under the microscope (Figure 2.4). The Gram-positive control, *B. subtilis*, exhibited a robust retention of the crystal violet, resulting in densely packed, rod-shaped cells that appeared deep purple under the microscope. In contrast, *E. coli*, the Gram-negative control, was observed to have a staining pattern that was consistent with the known Gram-negative by the faintly pink, short rods. These comparisons verified the Gram-negative status of strain ASA1^T and substantiated its classification based on cell envelope features.

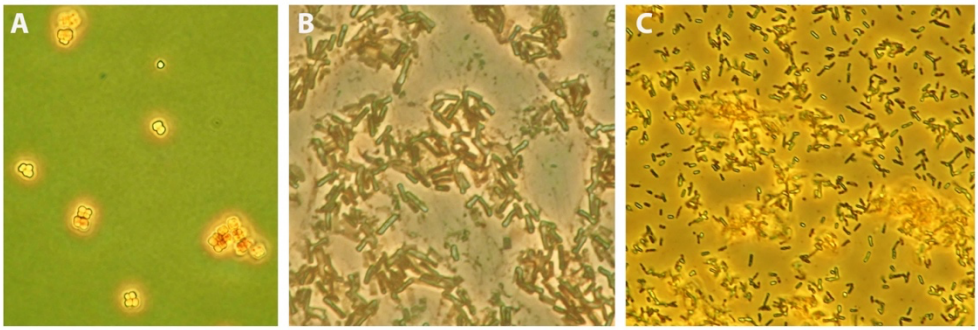

Figure 2.4. Comparison of the Gram-stain patterns of ASA1^T and the reference strains. (A) The pink/red appearance of ASA1^T cells indicates a Gram-negative reaction. (B) The Gram-positive phenomenon of *B. subtilis* is confirmed by its strong crystal violet retention. (C) *E. coli* exhibits the anticipated pink coloration of Gram-negative cells as a result of the loss of crystal violet during decolorization.

2.3.2.3. KOH test

After mixing with 3% KOH, no obvious viscosity production was detected, indicating a negative KOH response (Figure 2.5), which is associated with Gram-positive bacteria that contain a thick peptidoglycan coating. Control strains verified the test: *B. subtilis* (Gram-positive) reacted negatively, while *E. coli* (Gram-negative) generated a viscous appearance with the cells, confirming a positive result. (Figure 2.5). Interestingly, the ASA1^T KOH result reversed its Gram stain results, which consistently revealed pink/red cells, indicating a Gram-negative profile. Previous studies have revealed discrepancies between the KOH test and Gram staining, with certain bacterial strains producing inconsistent results due to differences in cell envelope structure (Gregersen 1978; Beveridge 1990). The KOH test is highly dependent on recent cell wall integrity, whereas Gram staining is influenced by the cell's physiological condition. Furthermore, the majority of *Verrucomicrobiota* are morphologically characterized as Gram-negative (Parte et al.

2011; Devos 2014), supporting this interpretation. Regardless of the KOH-negative result, the Gram-negative identification of ASA1^T was accepted based on consistent Gram staining.

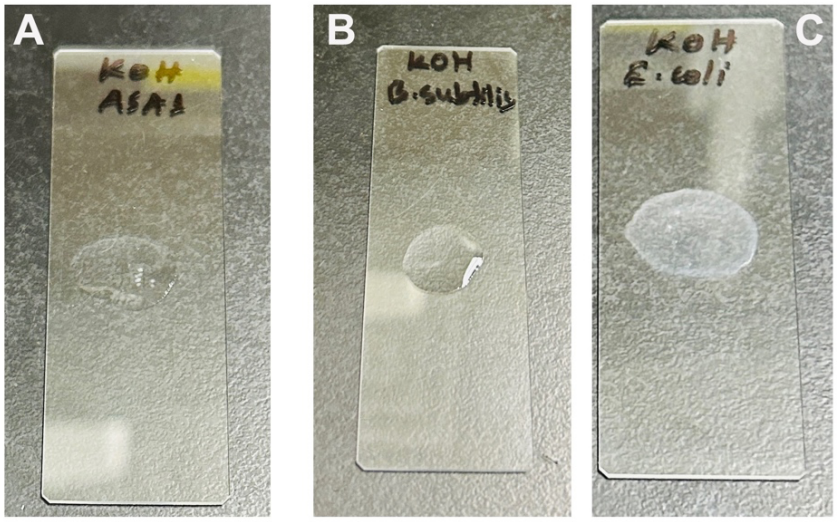

Figure 2.5. KOH test results for ASA1^T and reference strains. (A) ASA1^T shows no visible viscosity, interpreted as a negative KOH reaction. (B) *B. subtilis* also shows a negative reaction, consistent with its Gram-positive cell wall. (C) *E. coli* shows a clear viscosity, confirming its Gram-negative status.

2.3.2.4. Catalase test

No bubbles formed when the colony touched the slide surface, indicating a negative catalase reaction (Figure 2.6). This implies that ASA1^T lacks catalase, the enzyme that converts H_2O_2 into water and oxygen. For validation, *B. subtilis* (positive control) produced immediate, vigorous bubbling, while *E. coli* (negative control) showed no reaction, confirming the assay conditions.

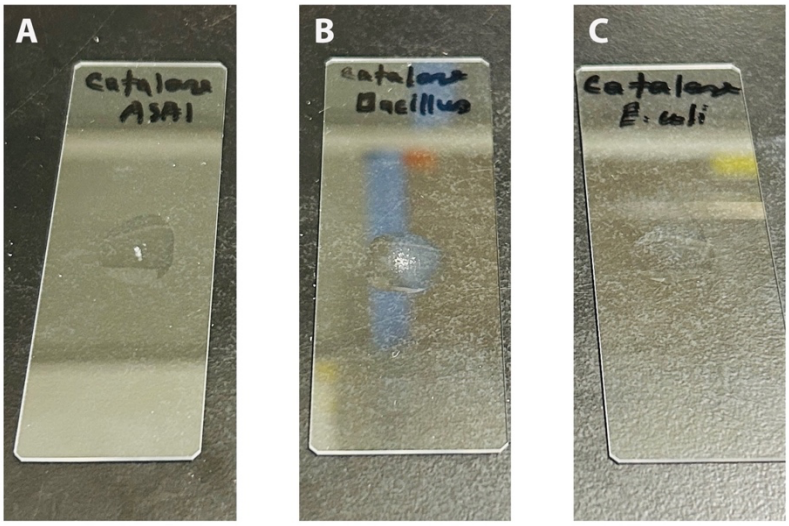

Figure 2.6. Catalase test results for ASA1^T and reference strains. (A) ASA1^T displays no bubbling, indicating a negative catalase reaction. (B) *B. subtilis* reacts positively and forms bubbles immediately. (C) *E. coli* has no bubbling, indicating a negative result under the same conditions.

2.3.2.5. Oxidase test

Cells applied to the oxidase reagent-soaked filter paper did not change color (Figure 2.7). *O. tunisiensis*, an oxidase-positive bacterium, became purple/blue, indicating the reagent reactions. *E. coli*, the negative control, exhibited no changes in color.

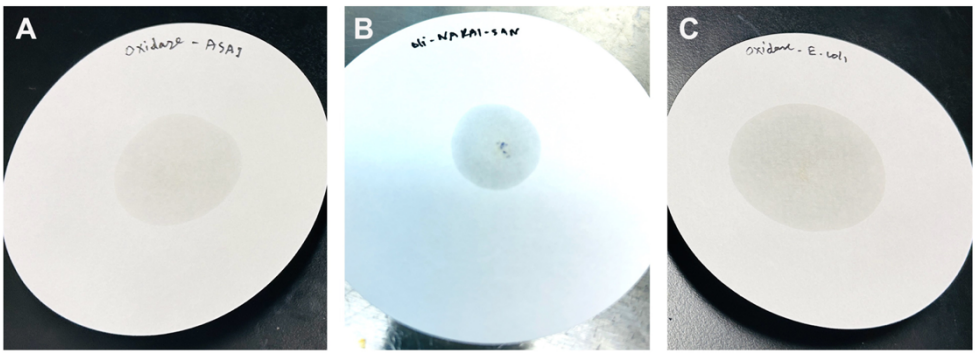

Figure 2.7. Oxidase results for ASA1^T and reference strains. (A) ASA1^T displays no color change, indicating a negative oxidase reaction. (B) *O. tunisiensis* (oxidase-positive control) turns purple/blue, indicating cytochrome c oxidase. (C) As an oxidase-negative organism, *E. coli* is colorless.

2.3.2.6. Motility

ASA1^T exhibited diffuse turbidity from the stab line, indicating motility (Figure 2.8). Like the positive control, *B. subtilis*, this pattern showed lateral spreading. *M. austroafricanum*, the negative control, grew only along the stab line. However, the microscopic observation of ASA1^T confirmed active movement, consistent with flagella-mediated motility.

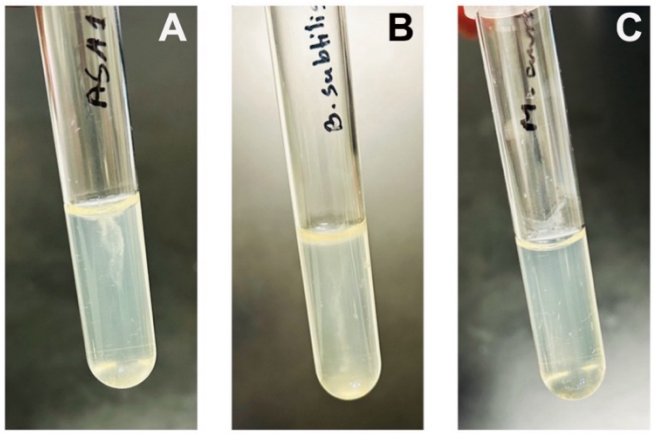

Figure 2.8. Motility test of strain ASA1^T and reference strains. (A) ASA1^T shows diffuse growth away from the stab line, indicating motility. (B) *B. subtilis* (positive control) exhibits similar spreading. (C) *M. austroafricanum* (negative control) shows restricted growth along the stab line.

2.3.2.7. Growth analysis by temperature, pH, and NaCl tolerance

The growth temperature of strain ASA1^T ranged from 20–40°C, with optimal growth at 30°C. The pH range for its growth was 7.0–11.0, with the highest growth observed at pH 8.4. The soil from which the ASA1^T strain was obtained had a pH of around 8. Additionally, bacterial growth was observed in R2B medium containing 0–3.0% NaCl, with maximum growth at 0% NaCl and no growth above 3.0% NaCl.

2.3.2.8. Substrate utilization

The carbon utilization profile of strain ASA1^T was limited to a narrow range of substrates. Positive growth was observed with D-glucose, sucrose, and D-maltose, whereas no utilization was detected for the rest of the substrates. The other phenotypic features of strain ASA1^T and the other phylogenetically related type strains are presented in Table 2.1.

Table 2.1. Differentiation between strain ASA1^T and other closely related representatives of *Opitutales*.

Features	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Isolation source	Garden soil	Hot springs	Orchard soil	Rice paddy soil	Termite gut	Sea polychaete	Freshwater	Oil sands tailing ponds	Hyporheic freshwater	Mangrove sediment
DNA G+C content (mol%)	65.6	65.8	64.3	74.0	60.5	52.1	60.6	66.1	62.2	56.0
Cell shape	Coccus or oval	Coccus or spherical	Coccus or short rod	Coccus	Diplococcus	Coccus	Coccus	Coccus or diplococcus	Coccus	Coccus
Cell diameter (μm)	Coccus	0.8–0.9	0.4–0.7	0.4–0.6	0.5–0.6	0.6–1.0	0.35	0.5–1.0	0.4–0.5	ND

	or oval (0.4–0.6) and aggregate (>10 μm)									
Colony color	Whitish yellow	White	Yellow	Colorless	Colorless	Pale reddish	Colorless	White	Yellow	ND
Metabolism	Obligate aerobe	Facultative anaerobe and aerobe	Obligate aerobe	Obligate anaerobe	Obligate aerobe	Facultative anaerobe	Obligate aerobe	Obligate aerobe	Obligate aerobe	Obligate aerobe
Motility	+	+	-	+	-	-	+	-	-	-
Catalase	-	+	+	-	-	-	-	+	-	+
Oxidase	-	+	-	-	-	-	-	-	-	+
Temp. (°C)	20–40 (30)	40–56	15–37 (25)	10–37	15–35 (30)	10–40 (30–35)	6–30	15–40 (25–37)	15–37 (25–30)	15–40 (30)
Growth pH	7.0–11.0 (8.4)	7.0–8.5 (8.0)	5.0–11.0 (6.0)	5.5–9.0 (7.5–8.0)	5.5–7.5 (7.0)	5.5–9.5 (7.0–7.5)	ND	5.5–11.0 (6.0–8.0)	5.5–9.0 (7.0)	ND
NaCl tolerance (%, w/v)	0–3.0 (0)	1.0–3.5 (2.0–2.5)	0–3.0 (0)	3.0	1.5	1.0–10.0 (2.5–3.0)	0–0.7	0–3.0	0–0.5 (0)	ND
Substrates										
Arabinose	-	-	+	+	-	W	ND	W	+	+
Cellobiose	-	+	+	+	+	+	ND	-	+	ND
Fructose	-	ND	-	+	-	+	+	+	+	ND
Glucose	+	+	+	+	+	+	W	+	+	+
Galactose	-	+	-	+	+	+	ND	+	+	ND
Maltose	+	ND	-	+	+	+	ND	-	ND	ND
Mannitol	-	-	ND	+	-	+	ND	+	-	w
Mannose	-	W	-	+	-	+	ND	-	+	ND

Ribose	-	ND	ND	-	-	ND	ND	ND	ND	ND
Starch	-	ND	-	+	+	ND	ND	-	-	-
Sorbitol	-	-	ND	ND	-	+	ND	+	ND	ND
Sucrose	+	+	-	+	-	+	ND	-	-	ND
Xylose	-	+	+	-	-	ND	ND	+	ND	ND
Major fatty acids (%)	anteiso-$C_{15:0}$ (53.0), $C_{16:0}$ (15.0)	anteiso-$C_{15:0}$ (51.5)	iso-$C_{15:0}$ (34.9), anteiso-$C_{15:0}$ (29.0)	ND	ND	anteiso-$C_{15:0}$ (30.9), $C_{18:0}$ (24.7), $C_{16:0}$ (7.9), and $C_{17:0}$ (7.0)	anteiso-$C_{15:0}$ (31.1), $C_{16:0}$ (26.7), $C_{16:1}\omega5c$ (16.9), iso-$C_{14:0}$ (8.3), and $C_{14:0}$ (6.0)	anteiso-$C_{15:0}$ (28.4), iso-$C_{15:0}$ (17.1)	iso-$C_{14:0}$ (19.9), anteiso-$C_{15:0}$ (19.3), $C_{16:0}$ (16.8), and iso-$C_{16:0}$ (14.5)	iso-$C_{14:0}$ (23.1) and $C_{18:1}$ $\omega9c$ (19.0)
Respiratory quinones	MK-7	ND	MK-7 and MK-6	ND	ND	MK-7	MK-7	MK-7 and MK-6	MK-7	MK-7

Strains: 1. ASA1^T; 2. *Alterococcus agarolyticus* ADT3^T (Shieh and Jean 1998); 3. *Horticcoccus luteus* KSB-15^T (Chung et al. 2022); 4. *Opitutus terrae* PB90-1^T (Chin et al. 2001); 5. *Geminisphaera colitermitum* TAV2^T (Wertz et al. 2012, 2018); 6. *Puniceicoccus vermicola* IMCC1545^T (Choo et al. 2007); 7. *Rariglobus hedericola* 53C-WASEF^T (Pitt et al. 2020); 8. *Oleiharenicola alkalitolerans* NVT^T (Rochman et al. 2018); 9. *Nibricoccus aquaticus* HZ-65^T (Baek et al. 2019); 10. *Coraliomargarita parva* WMMB3^T (Min et al. 2023); (+) present or utilized; (-) absent or not utilized; (w) weakly positive; (ND) not determined. Values in parentheses indicate the optimal conditions.

2.3.2.9. Enzyme activity profile

The enzymatic activity of ASA1^T was assessed using API ZYM and API 20 NE strips. Strong activity was detected for alkaline phosphatase and esterase (C4). Weak but detectable activities

were observed for esterase lipase (C8), α-galactosidase, β-galactosidase, and napthol-AS-BI-phosphohydrolase. All other enzyme reactions tested negative. This enzymatic profile suggests selective metabolic capacities for lipid and phosphate ester hydrolysis.

2.3.3. Chemotaxonomic profiling

The FAME analysis of strain ASA1^T revealed that anteiso-$C_{15:0}$ was the most abundant cellular fatty acid, accounting for approximately 53% of the total fatty acid content (Table 2.2).

Table 2.2. Major (>1% of total) fatty acids of strain ASA1^T

Fatty acids	Percentage (%)
3-hydroxyl Tridecanoic acid (3-OH 12:0)	1
3-hydroxyl Tetradecanoic acid (3-OH 13:0)	1
12-metyl Tetradecanoic acid (anteiso-15:0)	53
Pentadecanoic acid (15:0)	2
14-metyl Pentadecanoic acid (iso-16:0)	1
Hexadecenoic acid (16:1(9))	9
Hexadecanoic acid (16:0)	15
14-metyl Hexadecanoic acid (anteiso-17:0)	5
Heptadecanoic acid (17:0)	4
Octadecenoic acid (18:1)*	7
Others	3

*It contains two isomers.

The prevalence of anteiso-branched fatty acids, especially anteiso-$C_{15:0}$, is significant, as these constituents are frequently linked to improved membrane fluidity, aiding adaptation to low temperatures and extreme conditions (Sun et al. 2012; De Carvalho and Caramujo 2018). While the functional role in ASA1^T remains to be confirmed, similar patterns have been reported in other environmental *Verrucomicrobiota* from variable habitats. Additionally, menaquinone-7 (MK-7) was identified as the sole respiratory quinone (Figure 2.9), consistent with members of the order *Opitutales* and supporting the taxonomic placement of ASA1^T within this lineage.

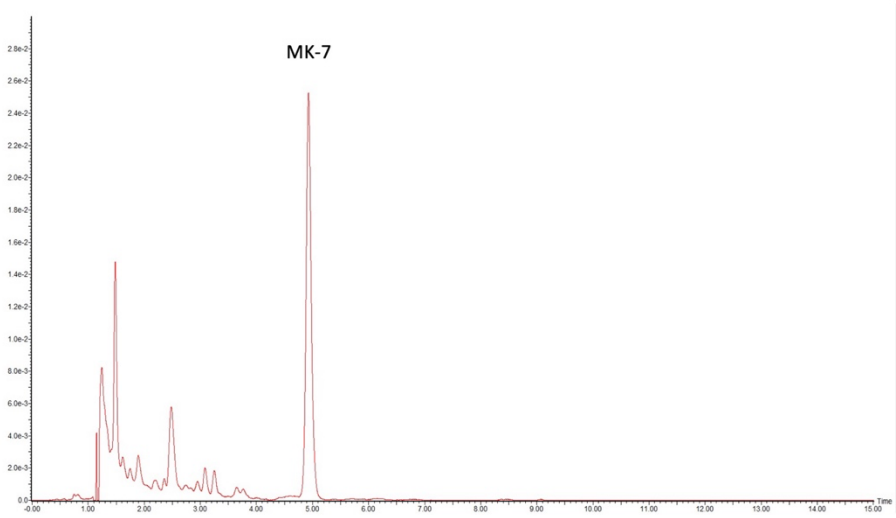

Figure 2.9. Quinone profile of strain ASA1^T determined by UPLC. The dominant peak corresponds to MK-7, confirming it as the major isoprenoid quinone present in the strain.

2.3.4. Genomic DNA extraction and quality assessment

The NanoDrop test indicated a concentration of 355.53 ng/µL and A_{260}/A_{280} and A_{260}/A_{230} ratios of 1.95 and 2.00, respectively, indicating excellent purity with minimal protein or phenol contamination. To validate this, DNA quantification was also performed using a Qubit 4 fluorometer with the dsDNA high sensitivity assay, yielding a concentration of 385 ng/µL. This

result supported the NanoDrop reading and showed that there was enough DNA for sequencing and polymerase chain reaction (PCR).

Integrity was assessed by 1% agarose gel electrophoresis, which showed a single, high-molecular-weight DNA band with no smearing or degradation. The lack of following bands shows that the DNA was neither sheared or broken down during extraction. The nearly full-length 16S rRNA gene (~1500 bp) of strain ASA1^T was successfully amplified using universal primers 27F and 1492R, producing clear bands in all lanes (Figure 2.11). Overall, the high concentration, purity, and intact structure of the genomic DNA confirm the effectiveness of the extraction protocol and its suitability for downstream applications. The DNA was stored at -20°C for long-term use.

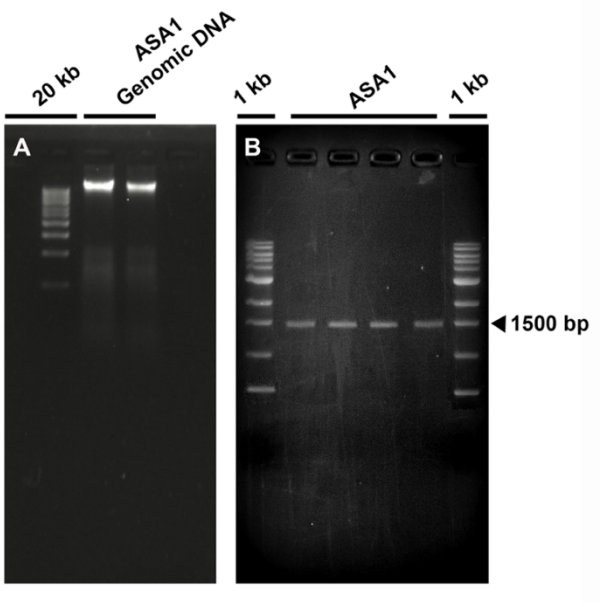

Figure 2.10. Agarose gel electrophoresis of strain ASA1^T genomic DNA and 16S rRNA gene PCR products. (A) High-molecular-weight genomic DNA of strain ASA1^T, showing intact, unsheared DNA bands alongside a 20 kb size marker. (B) PCR amplification of the ~1500 bp 16S rRNA gene of strain ASA1^T yielded clear bands in all lanes. A 1 kb DNA ladder was used as a size marker.

2.3.6. Phylogenetic study

The phylogenetic position of strain ASA1^T was determined based on the comparative analysis of its 16S rRNA gene sequence. The complete sequence was later retrieved from the whole-genome assembly (GenBank accession number: AP028972.1). Sequence similarity analysis using the NCBI BLASTn database revealed that strain ASA1^T shared <93% sequence identity with the closest type strain, *Alterococcus agarolyticus* ADT3^T (accession no. NR_036763), the only currently described species within the family *Alterococcaceae* of the class *Opitutia* (phylum *Verrucomicrobiota*). The 16S rRNA gene sequence of strain ASA1^T also displayed <93% identity with multiple uncultivated environmental clones, including sequences retrieved from biofilms (e.g., FJ516822), tall grass (FJ479434), and soil (JQ716335). Phylogenetic analysis was carried out using both NJ (Figure 2.12) and ML (Figure 2.13) algorithms with 1,000 bootstrap replications. In both trees, ASA1^T consistently clustered within the *Alterococcaceae* lineage of the class *Opitutia*.

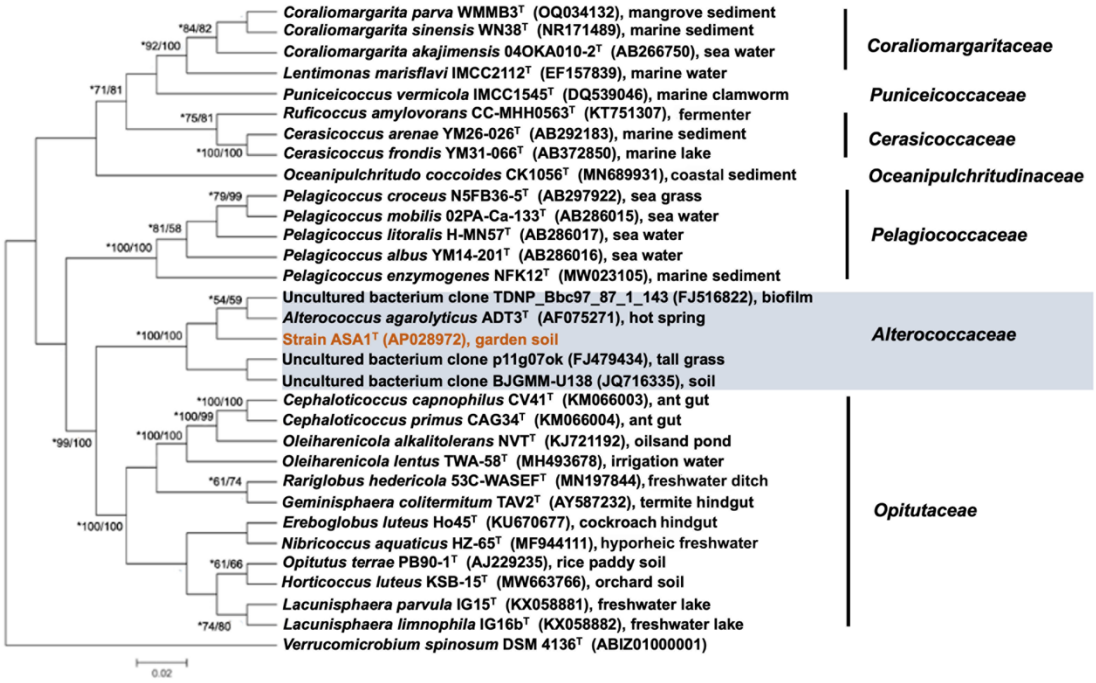

Figure 2.11. An NJ tree illustrating the relationship between strain ASA1^T and type strains of class *Opitutia* based on their 16S rRNA gene sequences. Bootstrapping was carried out with 1,000 replicates; only values exceeding 50% are displayed. Asterisks (*) denote the distance score in the ML method. *Verrucomicrobium spinosum* DSM 4136^T served as an outgroup. The bar set at 0.02 signifies the nucleotide substitutions per position.

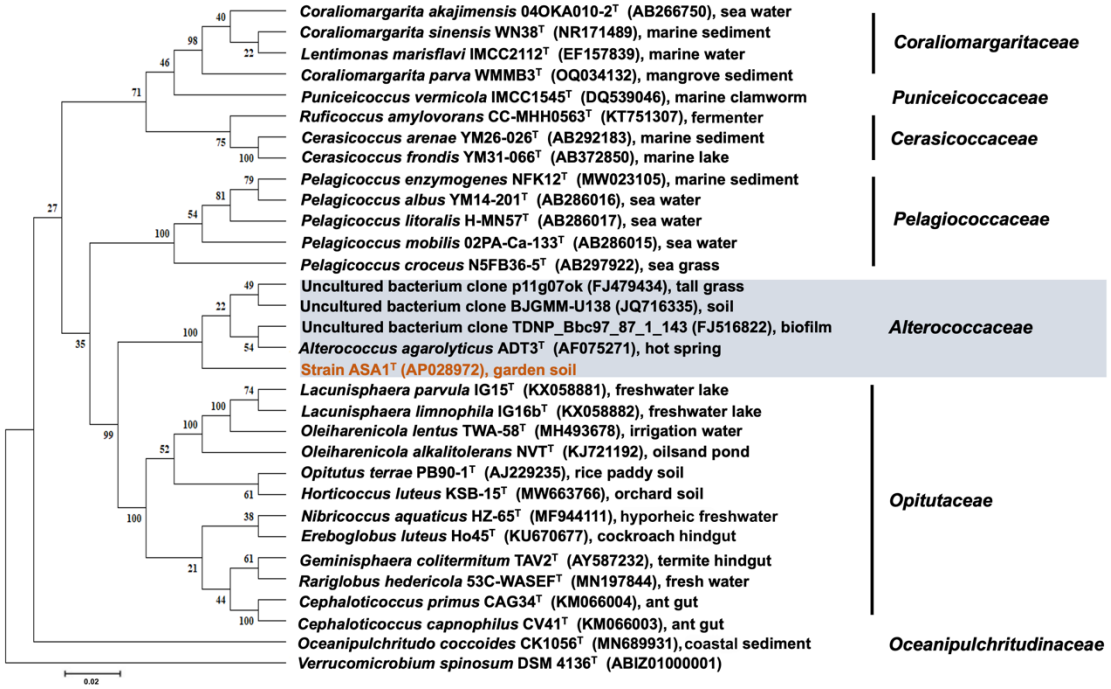

Figure 2.12. An ML tree was constructed to illustrate the relationship between strain ASA1^T and type strains of the class *Opitutia*. Bootstrap values (n = 1,000) exceeding 50% are denoted at the internodes of the tree. The scale bar represents a sequence difference of 0.02. *Verrucomicrobium spinosum* DSM 4136^T was employed as the outgroup.

2.4. Conclusion

This study provides an in-depth characterization of strain ASA1^T, a novel member of the phylum *Verrucomicrobiota*, previously isolated from garden soil using a modified PS medium. The strain

displayed coccus to oval-shaped morphology, with cell volumes <0.1 μm³, and formed distinct cluster-like aggregates (>10 μm). Physiological and biochemical characterization revealed that ASA1^T is an aerobic, motile, Gram-negative bacterium, and a cell volume well below 0.1 μm³, qualifying it as an ultramicrobacterium. Although the KOH test result suggested a Gram-positive feature, consistent Gram-staining results supported a Gram-negative identity, consistent with most *Verrucomicrobiota*. Biochemical assays demonstrated negative catalase and oxidase activity, but confirmed moderate metabolic flexibility through selective substrate utilization. Chemotaxonomic analysis revealed anteiso-$C_{15:0}$ as the predominant fatty acid (53.0%) and MK-7 as the sole quinone, aligning with features typical of the order *Opitutales*. Phylogenetic analysis based on both nearly full-length 16S rRNA gene data placed ASA1^T within the family *Alterococcaceae* of the class *Opitutia*, sharing <93% sequence identity with its closest type strain, *Alterococcus agarolyticus* ADT3^T. Both NJ and ML trees showed strong bootstrap support for its affiliation within *Alterococcaceae*, confirming its distinct phylogenetic position within the *Verrucomicrobiota*.

Chapter III

Study II

Genomic characterization and potential distribution of a novel *Verrucomicrobiota* strain

3.1. Introduction

Understanding the genetic framework of microorganisms is essential for revealing their identification, ecological functions, and evolutionary relationships. Whole-genome sequencing facilitates the precise classification of novel strains while elucidating their distinct genes, metabolic processes, and prospective ecological roles. Resources such as the Genome Taxonomy Database (GTDB) and marker gene-based phylogenomic analysis offer a reliable foundation for placing novel strains within the bacterial tree of life (Parks et al. 2018). This chapter explores the genomic characteristics and potential habitability of strain ASA1^{T}.

3.2. Materials and Methods

3.2.1. Whole-genome sequencing, genome annotation, and functional analysis

3.2.1.1. Whole genome sequencing

The extracted DNA was then fragmented into 10–20 Kbp lengths using g-TUBE (Covaris) to prepare for single-molecule real-time (SMRT) sequencing. Library preparation was carried out using the SMRTbell gDNA Sample Amplification Kit and the SMRTbell Express Template Prep Kit 2.0 (PacBio), followed by whole-genome sequencing on the PacBio Sequel IIe platform. High-fidelity (HiFi) reads were generated using SMRT Link (v12.0) under default parameters. Raw data were processed by removing adapters, duplicates, and short reads (≤1000 bp) using tools including lima (v2.7.1), pbmarkdup (v1.0.3), and Filtlong (v0.2.1) (https://github.com/rrwick/Filtlong) with parameters "--min_length 1000" and "--keep_percent 90". After filtering, 45,742 HiFi reads were obtained with an average read length of 6,486 bp, totaling approximately 296,678,282 bp. The genome was de novo assembled using Flye (v2.9.1) (with --pacbio-hifi and -i 3 options), and the circular genome completeness was confirmed using Bandage (v0.8.1). Genome annotation was

performed using DFAST (v1.2.11), which provided gene prediction and functional assignments. The graphical genome map was created using CGView tools (Grant and Stothard 2008) within the KBase platform (https://www.kbase.us/) (Arkin et al. 2018). Taxonomic comparisons were conducted by calculating the average nucleotide identity (ANI) against type strain genomes using DFAST's ANI tool. The average amino acid identity (AAI) was further computed using the Enveomics AAI Calculator (https://enveomics.scigap.org/), focusing on comparisons with representatives of the class *Opitutia* (Rodriguez-R and Konstantinidis 2016).

3.2.1.2. Genome annotation and functional analysis

The annotated genome of ASA1 was investigated using the zDB platform (Comparative Bacterial Genomics Made Easy) to elucidate its functional and metabolic capabilities (Marquis et al. 2024). This platform assigned coding sequences to clusters of orthologous groups (COG) based on conserved evolutionary relationships. Furthermore, metabolic pathway reconstruction was conducted with BlastKOALA from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, facilitating the identification of essential pathways associated with carbon metabolism, biosynthesis, and energy production. Together, these analysis enabled the functional interpretation of the ASA1 genome.

3.2.2. Phylogenomic study

Whole-genome-based phylogenomic analysis were conducted using the BV-BRC pipeline (Bacterial and Viral Bioinformatics Resource Center) and GTDB-Tk (v2.3.2), which included type strains, metagenome-assembled genomes (MAGs), and uncultured *Verrucomicrobiota*

representatives. All tree visualizations were generated using the interactive tree of life (iTOL) platform for detailed evolutionary interpretation.

3.2.3. Habitability analysis

To investigate the ecological distribution and environmental adaptability of strain ASA1^T and its phylogenetically related taxa, a metagenomic survey was performed using the integrated microbial next-generation sequencing (IMNGS) platform (https://www.imngs.org/) (Lagkouvardos et al. 2016). This platform facilitates high-throughput screening of 16S rRNA gene amplicon data from various environmental metagenomic datasets. The nearly complete 16S rRNA gene sequence of ASA1^T was utilized as a query, employing a 97% sequence identity threshold to identify and map closely comparable sequences from diverse environmental sources. To evaluate habitat preference, ProkAtlas (http://prokatlas.bs.s.u-tokyo.ac.jp/) was utilized, a comprehensive resource comprising over 360,000 annotated 16S rRNA gene sequences with curated habitat metadata (Mise and Iwasaki 2020). Employing the identical 97% identity threshold, the ASA1^T sequence was matched with the ProkAtlas database to assess its relative abundance and distribution across diverse ecological niches. This integrated method facilitated a comprehensive assessment of the environmental distribution and possible habitat adaptability of strain ASA1^T.

3.2.4. Genome-based study of rRNA operon number for assessing bacterial growth rate

The number of rRNA operons was studied across genomes to assess the potential growth rate and protein synthesis capacity of strain ASA1^T and similar taxa. Genomic datasets for ASA1^T and phylogenetically affiliated species were obtained from the EzBioCloud database (https://www.ezbiocloud.net/), and gene annotation, encompassing 5S, 16S, and 23S rRNA genes,

was performed utilizing the DFAST annotation pipeline (https://dfast.ddbj.nig.ac.jp/). The final dataset, including incubation period and rRNA copy number, was illustrated utilizing GraphPad Prism version 9.5.0.

3.2.5. Identification and comparative BGC analysis of strain ASA1^{T}

To investigate the biosynthetic potential of strain ASA1^{T}, its genome was analyzed using antiSMASH (v7.0.1). This tool integrates multiple features, including ClusterBlast, Pfam analysis, SubClusterBlast, ActiveSiteFinder, and TIGRFam, to identify and annotate biosynthetic gene clusters (BGC). The analysis was associated with the minimum information about a biosynthetic gene cluster (MiBIG v3.0) database, facilitating comparison with established, experimentally verified pathways. Specific focus was directed to NRPS, PKS, and hybrid clusters to assess their capacity for generating specific metabolites.

3.3. Results and Discussion

3.3.1. Whole-genome sequencing and genome assembly

The complete genome of ASA1 was successfully sequenced using the PacBio Sequel IIe platform employing SMRT sequencing technology. This long-read sequencing approach enabled the generation of HiFi reads, which are essential for accurate and contiguous genome assembly. A total of 45,742 high-quality HiFi reads were obtained following rigorous library preparation, size selection (10–20 kb using g-TUBE), and quality filtering steps. These reads had an average length of 6,486 base pairs, producing approximately 297 megabase pairs (Mbp) of sequencing data. The sequencing depth achieved was approximately 51×, providing sufficient coverage for reliable assembly. The final genome assembly of strain ASA1 comprises a single chromosome of

5,821,695 base pairs with a G+C content of 65.6%. The completeness of the genome was visually confirmed using Bandage software, which showed no fragmentation or unresolved regions, further confirming the high quality of the assembly.

Functional annotation was performed using the DFAST pipeline, which identified a total of 4,690 genes. These included 4,638 protein-coding sequences (CDSs), 48 transfer RNA (tRNA) genes, 3 ribosomal RNA (rRNA) genes (5S, 16S, and 23S), and a single clustered regularly interspaced short palindromic repeats (CRISPR) array. The presence of a complete set of core genes, along with the identified regulatory and defense-related features, suggests the genomic robustness and metabolic versatility of strain ASA1^T. To investigate the taxonomic placement of strain ASA1^T, genome-wide comparisons were conducted using ANI and AAI analysis. The ANI values between ASA1^T and its closest available genomes, including *Rariglobus hedericola* 53C-WASEF^T (MN197844) and the MAG VE3 (GCA_014879845.1), were below 77%. These values are significantly lower than the widely accepted threshold for species-level similarity (>95%) (Jain et al. 2018), indicating that ASA1^T represents a distinct species. Further, AAI comparisons with representative genomes of the class *Opitutia* revealed even more pronounced divergence. The highest AAI was 54.0%, shared with HZ-65^T, followed by 53.8% with the closest described genus PB90-1^T (Appendix 5). These AAI values fall below the generally accepted genus-level threshold of 60–65% (Konstantinidis et al. 2017; Riesco and Trujillo 2024), strongly supporting the designation of strain ASA1^T as a representative of a novel genus within the class *Opitutia*. At the time of this analysis, no genome sequence was available for *Alterococcus agarolyticus*, the only other described member of the family *Alterococcaceae*. Therefore, ASA1^T remains the sole genome-sequenced, cultivated representative of this family, emphasizing its importance in advancing phylogenomic resolution of the *Verrucomicrobiota* lineage.

To further visualize the genome architecture and distribution of functional features, a circular genome map was generated using the CGView tool integrated into the KBase platform. The circular representation (Figure 3.1) displays the arrangement of CDSs, tRNA, and rRNA genes, and key genome statistics including GC content and GC/AT skew. The AT and GC skew profiles shown in the inner rings of the map reflect strand-specific nucleotide asymmetries, which are commonly associated with the origin and terminus of replication. Together, this comprehensive genomic analysis confirms the high quality and completeness of the ASA1^{T} genome within *Verrucomicrobiota*, warranting its classification as a novel genus.

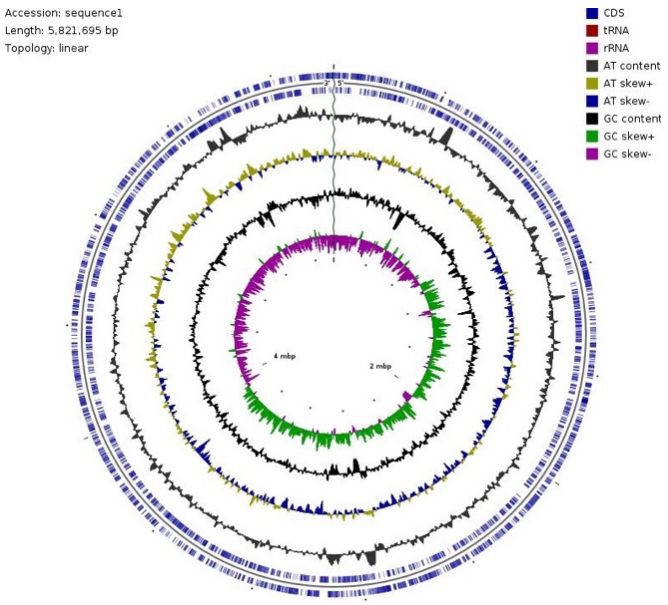

Figure 3.1. Schematic representation of the circular genome map of the ASA1 generated using the CGView server online tool (http://cgview.ca/). Circle 1 (outermost) displays the coding sequence. Circles 2 and 3 display the AT content plot. Circles 4 and 5 (innermost) display the GC content and GC skew, respectively.

3.3.2. Genome annotation and functional analysis

Using the zDB platform, a total of 3,820 coding sequences were assigned to COG categories. The most highly represented COG categories included: G (carbohydrate transport and metabolism), M (cell wall/membrane/envelope biogenesis), and T (signal transduction mechanisms). A substantial number of genes were also assigned to category E (amino acid transport and metabolism), P (inorganic ion transport and metabolism), J (translation, ribosomal structure and biogenesis), and K (transcription), indicating functional gene presence across a broad range of physiological activities (Figure 3.2). Notably, a significant portion of genes also fell under "function unknown (S)," reflecting the novelty of the ASA1^{T} genetic repertoire.

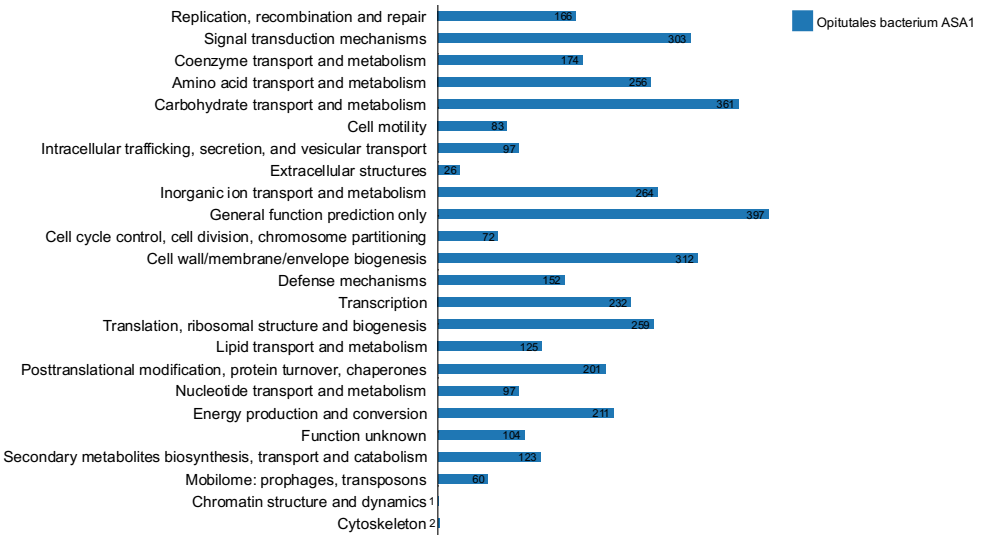

Figure 3.2. Distribution of protein-coding genes of strain ASA1^{T} across COG functional categories.

KEGG Orthology (KO) analysis of the ASA1 genome revealed that cofactor and vitamin metabolism was the most abundant (81 genes) (Figure 3.3). ATP synthesis genes account for approximately 32% of all assigned KO categories (48 genes). Amino acid metabolism pathways were also prominent, including arginine and proline metabolism (14 genes) and aromatic amino

acid metabolism (15 genes). Genes related to central carbohydrate metabolism (43 genes) and other carbohydrate metabolism (57 genes) were detected. The genome also contained genes involved in secondary metabolite biosynthesis, including five for antibiotic biosynthesis and one for aromatic compound degradation. These findings suggest that strain ASA1^T possesses genetic potential for producing bioactive compounds.

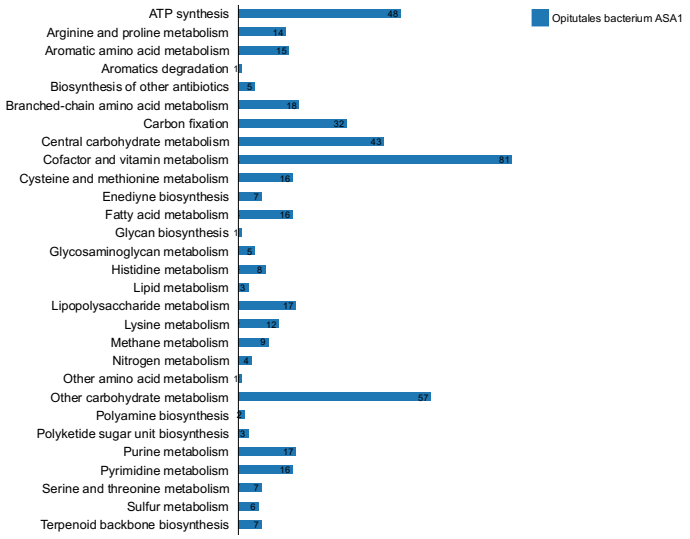

Figure 3.3. KEGG functional classification of strain ASA1^T based on KO (KEGG Orthology) assignments generated by BlastKOALA.

3.3.3. Phylogenomic study

Whole-genome-based phylogenomic trees were constructed using two independent platforms. The first tree was generated using the BV-BRC tool, which constructs species trees based on 100 universal single-copy marker genes from the Bac100 set. In this tree, strain ASA1^T clustered within the class *Opitutia* but formed a distinct branch that did not overlap with any described genus-level clades (Figure 3.4).

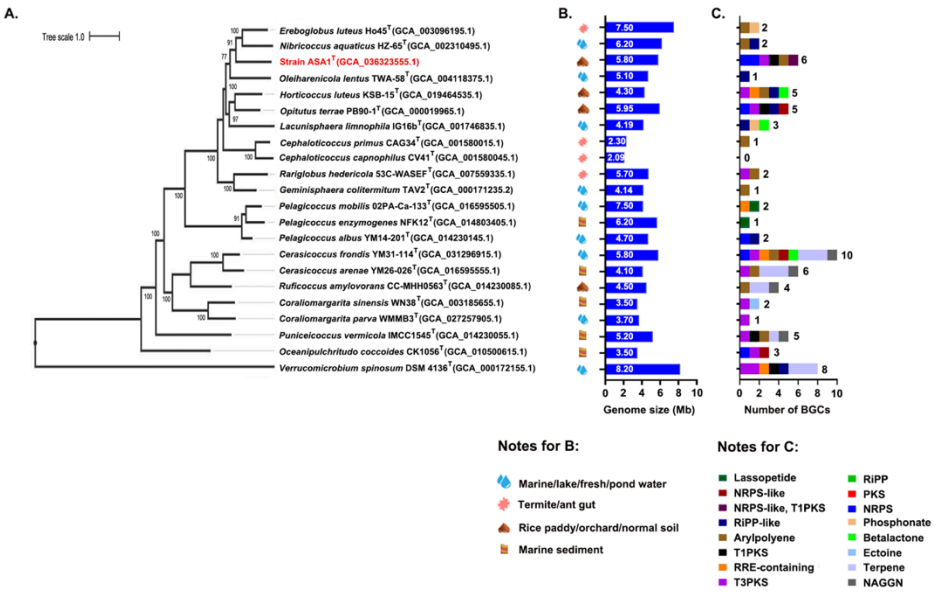

Figure 3.4. Phylogenomic tree with BGC count. (A) A phylogenomic tree coupled with the BGC count was created using 100 conserved concatenated proteins from bacteria, illustrating the evolutionary relationships between strain ASA1^{T} and the type strains belonging to the class *Opitutia*. The root of the tree was established using *V. spinosum* DSM 4136^{T} (GCA_000172155.1). Bootstrap values for nodes exceeding 80 are depicted on the tree. The scale bar set at 1.0 represents the amino acid substitutions per position. (B) The isolation source and genomic size of each taxon are also displayed. (C) Bar chart providing an in-depth analysis of the BGC identified within the genomes of the species in the phylogenomic tree. BGC were determined using the antiSMASH online tool (v7.0.1). The detailed features of the genome and BGC profiling of the corresponding species are listed in Appendix 5.

A second phylogenomic analysis was performed using GTDB-Tk (v2.3.2), which follows the GTDB classification system. The GTDB-Tk tree was constructed based on 120 conserved bacterial marker genes and incorporated both cultured genomes and a wide range of MAGs (Figure 3.5). The placement of strain ASA1^{T} within the GTDB-based tree corroborated the results from

the BV-BRC analysis. ASA1^{T} formed a separate lineage within the class *Opitutia*, distinct from existing genera.

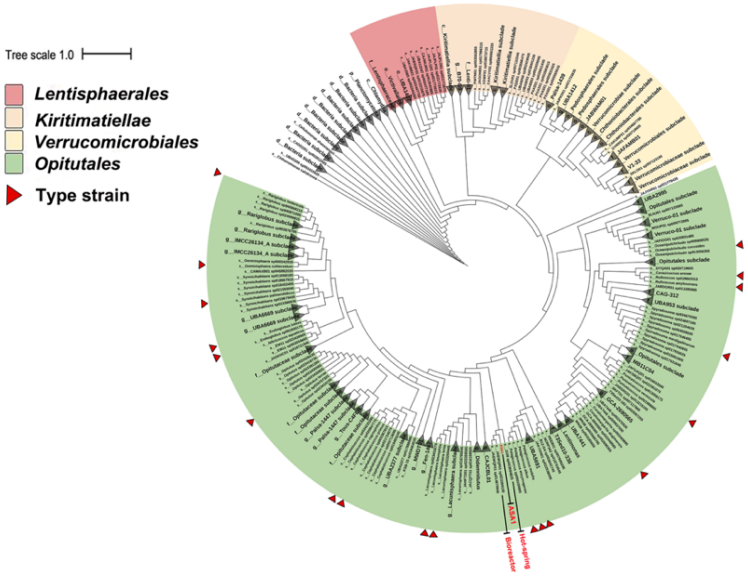

Figure 3.5. Phylogenomic tree generated using 120 concatenated bacterial conserved proteins to elucidate the relationships among genomes and MAGs associated with the class *Opitutia*. Type strains are highlighted in red boxes. Clades lacking cultivated representatives were condensed for clarity.

Both phylogenetic and phylogenomic analysis consistently indicated that strain ASA1^{T} occupies a unique lineage within the class *Opitutia*. The observed divergence from all currently described species and genera supports its distinct taxonomic positioning based on 16S rRNA sequence identity and genome-wide evolutionary markers.

3.3.4. Habitat distribution and environmental preferences of strain ASA1^{T}

The IMNGS-based search, using a similarity threshold of ≥97%, revealed that the 16S rRNA gene of ASA1^{T} shared sequence identity with 229 environmental amplicon datasets. Among these, the

majority originated from soil environments (88 datasets), followed by aquatic habitats (32 datasets) and activated sludge systems (21 datasets), suggesting a widespread environmental presence of ASA1^{T} relatives. To further assess the environmental niches associated with strain ASA1^{T}, a habitat prediction was performed using the ProkAtlas tool. At the 97% similarity threshold, this analysis indicated predominant associations with freshwater environments (20.82%), followed by peat (15.10%), rhizosphere (11.12%), soil (10.88%), general aquatic sources (9.62%), lake water (8.13%), groundwater (7.12%), wetlands (6.69%), freshwater sediments (6.41%), and marine environments (4.07%) (Figure 3.6). These findings suggest that strain ASA1^{T} and its phylogenetic relatives are not restricted to a single ecosystem but instead occupy a diverse range of terrestrial and aquatic environments.

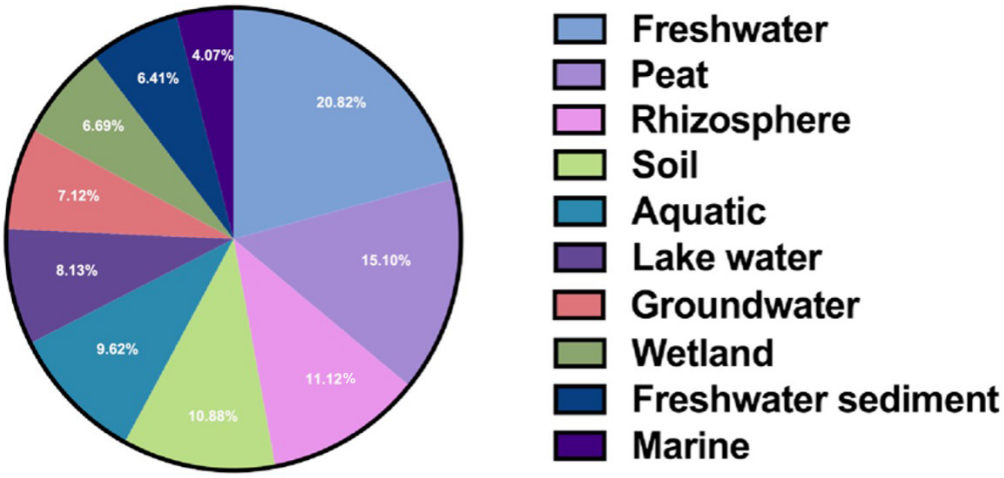

Figure 3.6. Habitat preference prediction of strain ASA1^{T} and its related *Verrucomicrobiota* members based on ProkAtlas.

3.3.5. Estimation of rRNA operon number for predicting bacterial growth conditions

To explore genomic factors potentially contributing to the slow-growing nature of strain ASA1^{T}, the rRNA operon copy number was analyzed in comparison with 19 closely related

Verrucomicrobiota genomes. Approximately 90% of the strains harbored a single rRNA operon (Figure 3.7), a trait generally associated with slower ribosome biogenesis and limited protein synthesis capacity. Among these, only a few strains displayed relatively faster growth (visible colonies within 3–4 days), while most required prolonged incubation. This observation aligns with the established relationship between rRNA operon number and bacterial growth rate, where fast-growing taxa such as *Bacillus*, *Arthrobacter*, and *Pseudomonas* typically possess 3–5 operons (Klappenbach et al. 2000; Roller et al. 2016). The prevalence of low operon counts among *Verrucomicrobiota* suggests an ecological strategy adapted to lower-nutrient environments.

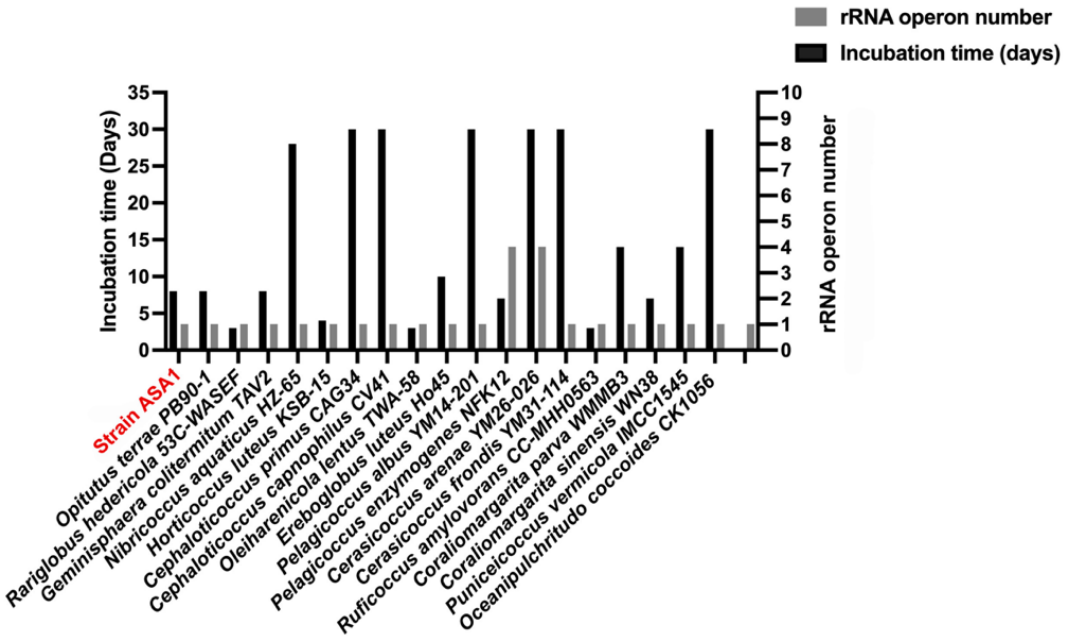

Figure 3.7. Correlation between rRNA operon copy number and incubation time in 19 related genomes of *Verrucomicrobiota*. Approximately 90% of species exhibited a single rRNA operon (1-set), correlating with the prolonged incubation times required for optimal growth. The estimation of the mean average operon numbers and incubation times was based on data provided in Appendix 5, which includes the incubation time for each isolate, with references.

3.3.6. Biosynthetic gene cluster prediction and comparative analysis

Genome mining of strain ASA1^{T} using antiSMASH (v7.0.1) identified six BGC regions. These included three clusters encoding NRPS, along with one T3PKS (type III polyketide synthase), one NRPS-like cluster, one arylpolyene, and one T1PKS. Each cluster contained essential biosynthetic, regulatory, and transport-related domains, and their predicted modular arrangements were visualized using antiSMASH analysis (Figure 3.8).

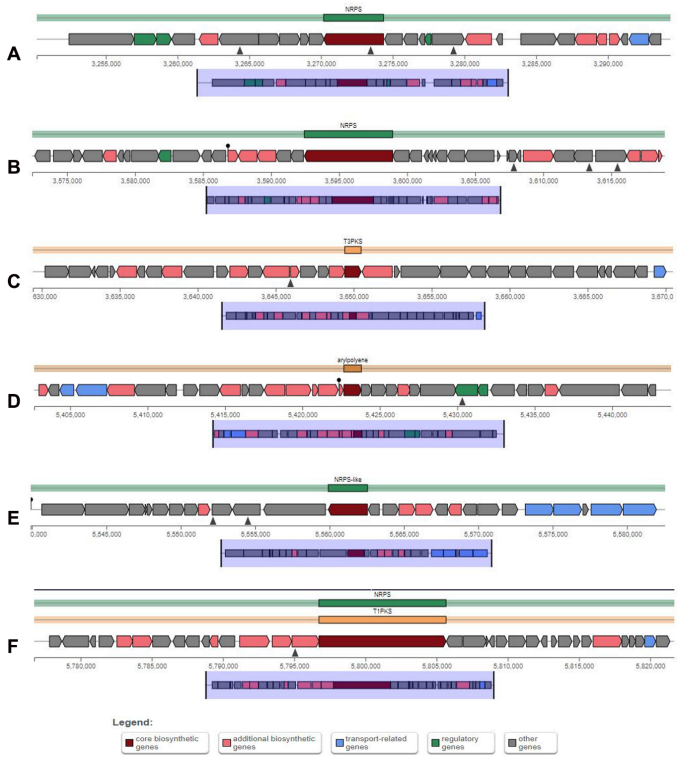

Figure 3.8. Graphical representation of six notable secondary metabolite gene cluster regions identified within the ASA1 genome (generated from antiSMASH v7.0.1). In the visual outputs, each BGC is represented as a linear schematic where individual genes are depicted as colored arrows. These colors correspond to specific functional roles as assigned by antiSMASH. It is important to note that these color codings indicate gene function, not the predicted chemical nature of the final metabolite.

Comparative analysis with the MIBiG (v3.0) database revealed low similarity scores (<1.0) between ASA1^T BGC and experimentally validated clusters, including those tentatively related to enterobactin, xenotetrapeptide, brasilicardin, and lipopolysaccharides (Table 3.1). Furthermore, no cluster in ASA1^T could be linked with known secondary metabolites, supporting the novelty of its BGC profile.

Table 3.1. Summary of predicted BGC in the strain ASA1 genome by antiSMASH v7.0.1.

Genomic region	BGC Type	Gene position		Most similar known cluster	Most similar type in MIBiG comparison (similarity score, %)
		From	To		
Region 1	NRPS	3,252,427	3,293,703	NA	NRP+Alkaloid from *Pseudomonas fluorescens* (0.31)
Region 2	NRPS	3,572,624	3,618,738	NA	NRP+Alkaloid from *Pseudomonas fluorescens* (0.31)
Region 3	T3PKS	3,629,430	3,670,473	NA	Polyketide from *Candidatus* Entotheonella serta (0.32)
Region 4	Arylpolyene	5,402,678	5,443,805	NA	Other from *Planctomycetales* bacterium 10988 (0.33)
Region 5	NRPS-like	5,540,659	5,581,974	NA	NRP+Saccharaide from *Streptomyces hygroscopicus* (0.34)
Region 6	T1PKS, NRPS	5,776,738	5,821,695	NA	NRP from *Pseudomonas* sp. J465 (0.67)

*NRPS, nonribosomal peptide synthase; NRP, nonribosomal peptide; T3PKS, Type III polyketide synthase.

To contextualize the biosynthetic potential of strain ASA1^T within a broader phylogenomic framework, additional genomes from members of the *Verrucomicrobiota* phylum were analyzed, obtained from the NCBI and GTDB databases (accessed September 2024). A total of 22 genomes, including ASA1, were subjected to genome mining using the antiSMASH tool to identify BGC (Appendix 5). In total, 73 BGC spanning 16 biosynthetic classes were predicted (Figure 3.4). Notably, 88% of these clusters showed no significant similarity to any known entries in the MIBiG reference database, suggesting a high potential for novel secondary metabolite production. These findings highlight the untapped metabolic potential of *Verrucomicrobiota* bacteria. Further

functional analysis and heterologous expression studies will be necessary to determine the bioactivity of these gene products, contributing to drug discovery and microbial ecology.

3.4. Conclusion

This chapter presents a comprehensive genomic, functional, and potential distribution analysis of strain ASA1^T. High-quality genome sequencing using PacBio SMRT technology enabled the assembly of a complete, circular chromosome with high GC content and no unresolved regions. Comparative genomic analysis using ANI and AAI placed ASA1^T well below established thresholds for species- and genus-level identity, strongly supporting its designation as a novel genus within the class *Opitutia*. Notably, ASA1 represents the first complete genome from a cultivated member of the family *Alterococcaceae*, filling a critical phylogenetic gap and enhancing the resolution of *Verrucomicrobiota* systematics. Phylogenomic reconstructions using both BV-BRC and GTDB-Tk consistently positioned ASA1^T on an independent lineage. Furthermore, the rRNA operon analysis revealed a single operon set in ASA1^T, consistent with its slow-growing phenotype and typical of other *Verrucomicrobiota* species. Habitat prediction analysis based on 16S rRNA gene comparisons suggests that ASA1^T and its close relatives are broadly distributed across terrestrial and aquatic ecosystems, particularly in soil, rhizosphere, peat, and freshwater environments. Crucially, genome mining uncovered six BGC, including NRPS, PKS, arylpolyene, and T3PKS types, all exhibiting <1% similarity to known clusters in the MIBiG database. Expanding the analysis across 22 *Verrucomicrobiota* genomes revealed a total of 73 BGC representing 16 biosynthetic classes, the majority of which also lacked known homologs. Notably, the findings of this study reveal a previously untapped source of bioactive compounds and offer

new insights into the isolation of uncultured microorganisms within the phylum *Verrucomicrobiota*.

Chapter IV

Study III

Taxonomic Proposal of a Novel

Verrucomicrobiota Strain

4.1. Introduction to taxonomic proposal

Taxonomic characterization is a fundamental aspect of microbiology that enables the classification and naming of newly discovered microbial lineages. Proper taxonomic resolution not only ensures systematic placement but also provides a framework for comparing genomic, physiological, and ecological traits. The present chapter provides the taxonomic status of strain ASA1^T based on a polyphasic approach, culminating in the proposal of a novel genus and species within the family *Alterococcaceae* of the class *Opitutia*.

4.2. Proposal for a new taxon, *Congregicoccus parvus* gen. nov., sp. nov.

Strain ASA1^T, previously isolated from garden soil in Hokkaido, Japan, represents a phylogenetically and phenotypically distinct member of the phylum *Verrucomicrobiota*. Comparative 16S rRNA gene sequence analysis revealed less than 93% similarity to all validly published type strains within the class *Opitutia*, including its closest known relative *Alterococcus agarolyticus* ADT3^T. Furthermore, genome-based metrics reinforced this taxonomic distinction: ANI values were consistently below 77%, and AAI values remained under 54%, both well below the accepted species and genus-level thresholds, respectively. Phenotypically, ASA1^T differed from *A. agarolyticus* in several key aspects. ASA1^T thrives at lower temperatures (optimal growth at 30°C) and displays aerobic metabolism, whereas *A. agarolyticus*, a facultatively anaerobic strain from a hot spring, prefers higher temperatures (40–56°C). Morphologically, ASA1^T cells are ultramicrobacterial (0.4–0.6 μm), aggregating into cluster-like patterns (>10 μm), contrasting with the larger, non-aggregating cells of *A. agarolyticus*. Both strains share a high abundance of anteiso-$C_{15:0}$ fatty acid, yet ASA1^T lacks catalase and oxidase activity. Given its comprehensive combination of genomic, physiological, morphological, and chemotaxonomic traits, strain ASA1^T

merits recognition as the type strain of a new genus and species, for which the name *Congregicoccus parvus* gen. nov., sp. nov. is proposed.

4.2.1. Description of *Congregicoccus* gen. nov.

Congregicoccus (Con.gre.gi.coc'cus. L. masc. adj. *congregus*, united in flocks; N.L. masc. n. *coccus*, from Gr. masc. n. *kokkos*, grain, seed; N.L. masc. n. *Congregicoccus*, a coccus that grows in flocks). Cells are small, motile, and non-spore-forming, showing a cluster-like aggregation. They are Gram-negative, chemoorganotrophic, and obligately aerobic. Major fatty acids include anteiso-$C_{15:0}$ and $C_{16:0}$. The predominant isoprenoid quinone is MK-7. This genus belongs to the family *Alterococcaceae* within the class *Opitutia*. The type species is *Congregicoccus parvus*.

4.2.2. Description of *Congregicoccus parvus* sp. nov.

Congregicoccus parvus (par'vus. L. masc. adj. *parvus*, small, referring to the small cell size of the type strain). Colonies on R2A plates are smooth, whitish-yellow, and <1.0 mm in diameter, becoming slightly yellow with prolonged incubation. Cells are coccoid to oval (0.45 × 0.45–0.60 μm), forming clustered-like aggregates >10 μm in size. The temperature range for growth on R2A medium is 20–40°C (optimal at 30°C). The pH range for growth is 7.0–11.0, with the optimum at 8.4 (potential changes in pH during incubation can affect colony growth). The NaCl concentration for growth ranged from 0–3.0%, with maximum growth at 0% NaCl and no growth above 3.0% NaCl. It is catalase and oxidase negative. Positive enzymatic activities include alkaline phosphatase and esterase (C4), with weak reactions for esterase lipase (C8), naphthol-AS-BI-phosphohydrolase, α-galactosidase, and β-galactosidase. Utilized substrates include D-glucose, sucrose, and D-maltose.

The type strain is ASA1^T (= JCM 36569^T = NCIMB 15508^T), previously isolated from garden soil at AIST Hokkaido Center, Sapporo, Japan. The genome size is approximately 5.8 Mb with a DNA G+C content of 66%. The complete genome sequence is deposited under GenBank accession number AP028972.1 (BioProject: PRJDB16742, BioSample: SAMD00647589; DDBJ Sequence Read Archive: DRA017385).

4.3. Conclusion

The comprehensive genomic analysis confirms that strain ASA1^T represents a novel taxon within the phylum *Verrucomicrobiota*. Its distinct phylogenetic placement and phenotypic properties justify the formal proposal of the novel genus and species *Congregicoccus parvus* gen. nov., sp. nov., within the family *Alterococcaceae*. This taxonomic classification has been validated and certified by the Japan Collection of Microorganisms (JCM) and the National Collection of Industrial, Food and Marine Bacteria (NCIMB).

General Conclusion

This study provides a detailed characterization of *Congregicoccus parvus* gen. nov., sp. nov., a novel slow-growing *Verrucomicrobiota* member. Combining morphological, physiological, and genome analysis led to the formal description of this new taxon, enhancing our understanding of *Verrucomicrobiota* diversity.

Strain ASA1^T was previously isolated from garden soil in the AIST center, Hokkaido, Japan, using PS medium. Colonies form after 8–14 days of incubation at 30°C. The strain displayed characteristic features of ultramicrobacteria (cell volume <0.1 μm), including coccus to oval-shaped cells (0.45 × 0.45–0.60 μm) and cluster-like aggregates (>10 μm), forming a morphology that may serve as an ecological adaptation for survival in oligotrophic environments.

Physiologically, ASA1^T is an obligate aerobe, Gram-negative, motile, and has a narrow substrate utilization range, limited to D-glucose, sucrose, and D-maltose. Biochemical assays confirmed the absence of catalase and oxidase activities, while API ZYM profiling revealed the presence of alkaline phosphatase, esterase (C4), and traces of esterase lipase (C8), α- and β-galactosidase, and naphthol-AS-BI-phosphohydrolase activities. Chemotaxonomic analysis distinguished anteiso-$C_{15:0}$ (53%) as the predominant fatty acid, a feature common among some *Verrucomicrobiota*. The quinone profile exclusively comprised MK-7, consistent with other members of *Opitutales*.

Genomic characterization provided further evidence of novelty and ecological adaptation. The complete genome of ASA1, sequenced using PacBio long-read technology, is 5.82 Mb in size with a high G+C content of 65.6%. Notably, ASA1^T harbors only a single rRNA operon, a trait correlated with slow growth and reduced ribosome synthesis capacity. Phylogenetic analysis based on 16S rRNA and whole-genome data consistently placed ASA1^T within the family

Alterococcaceae of the class *Opitutia* but separated it from the only described genus in the family, *Alterococcus*. ASA1^T exhibited <93% 16S rRNA gene sequence identity and <77% ANI to the closest relatives, well below species-level thresholds, and <54% AAI, below the genus-level cutoffs. These results confirm ASA1^T as a representative of a new genus and species. Functional annotation of the ASA1^T genome revealed a gene repertoire dominated by categories related to carbohydrate transport and metabolism, cell wall biogenesis, and signal transduction.

Ecological distribution analysis using IMNGS and ProkAtlas highlighted the potential cosmopolitan presence of ASA1^T-like sequences in diverse terrestrial and aquatic environments, particularly in freshwater, soil, peat, and rhizosphere ecosystems. This suggests a broad ecological relevance for ASA1^T and its relatives. Importantly, antiSMASH analysis identified six BGC, none of which matched known secondary metabolites, indicating a potential for the production of novel compounds.

Future research should focus on the ecological role and metabolic potential of ASA1^T. Gaining a better understanding of these functions may provide new insights into the properties of rarely cultured *Verrucomicrobiota*.

References

Anders H, Power JF, MacKenzie AD, et al (2015) *Limisphaera ngatamarikiensis* gen. nov., sp. nov., a thermophilic, pink-pigmented coccus isolated from subaqueous mud of a geothermal hotspring. Int J Syst Evol Microbiol 65:1114–1121

Andersson A, Larsson U, Hagström Å (1986) Size-selective grazing by a microflagellate on pelagic bacteria. Mar Ecol Prog Ser 51–57

Anthony MA, Bender SF, van der Heijden MGA (2023) Enumerating soil biodiversity. Proc Natl Acad Sci U S A 120

Arkin AP, Cottingham RW, Henry CS, et al (2018) KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569.

Axambayeva AS, Zhaparova LR, Shagyrova ZS, et al (2018) Unusual stability of a recombinant *Verrucomicrobium spinosum* tyrosinase to denaturing agents and its use for a production of a protein with adhesive properties. Appl Biochem Biotechnol 185:736–754

Baek K, Song J, Cho J-C, et al (2019) *Nibricoccus aquaticus* gen. nov., sp. nov., a new genus of the family *Opitutaceae* isolated from hyporheic freshwater. Int J Syst Evol Microbiol 69:552–557

Belova SE, Naumoff DG, Suzina NE, et al (2022) Building a Cell House from Cellulose: The Case of the Soil Acidobacterium *Acidisarcina polymorpha* SBC82^{T}. Microorganisms 10:2253

Ben-Dov E, Kramarsky-Winter E, Kushmaro A (2009) An in situ method for cultivating microorganisms using a double encapsulation technique. FEMS Microbiol Ecol 68:363–371

Bergmann GT, Bates ST, Eilers KG, et al (2011) The under-recognized dominance of *Verrucomicrobia* in soil bacterial communities. Soil Biol Biochem 43:1450–1455

Beveridge TJ (1990) Mechanism of gram variability in select bacteria. J Bacteriol 172:1609–1620

Camargo AP, de Souza RSC, Jose J, et al (2023) Plant microbiomes harbor potential to promote nutrient turnover in impoverished substrates of a Brazilian biodiversity hotspot. ISME J 17:354–370

Chia LW, Hornung BVH, Aalvink S, et al (2018) Deciphering the trophic interaction between *Akkermansia muciniphila* and the butyrogenic gut commensal *Anaerostipes caccae* using a metatranscriptomic approach. Antonie Van Leeuwenhoek 111:859–873

Chin K-J, Liesack W, Janssen PH (2001) *Opitutus terrae* gen. nov., sp. nov., to accommodate novel strains of the division'*Verrucomicrobia*'isolated from rice paddy soil. Int J Syst Evol Microbiol 51:1965–1968

Choo Y-J, Lee K, Song J, Cho J-C (2007) *Puniceicoccus vermicola* gen. nov., sp. nov., a novel marine bacterium, and description of *Puniceicoccaceae* fam. nov., *Puniceicoccales* ord. nov., *Opitutaceae* fam. nov., *Opitutales* ord. nov. and *Opitutae* classis nov. in the phylum '*Verrucomicrobia*.' Int J Syst Evol Microbiol 57:532–537

Chung J, Lee J-Y, Choi G-H, et al (2022) *Horticoccus luteus* gen. nov., sp. nov., a novel member of the phylum *Verrucomicrobia* isolated from a dichlorodiphenyltrichloroethane (DDT)-contaminated orchard soil. Curr Microbiol 79:340

Chuvochina M, Rinke C, Parks DH, et al (2019) The importance of designating type material for uncultured taxa. Syst Appl Microbiol 42:15–21

Crits-Christoph A, Diamond S, Butterfield CN, et al (2018) Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. Nature 558:440–444

Dai J, Ouyang Y, Gupte R, et al (2025) Microfluidic droplets with amended culture media cultivate a greater diversity of soil microorganisms. Appl Environ Microbiol e01794-24

Davis KER, Joseph SJ, Janssen PH (2005) Effects of growth medium, inoculum size, and

incubation time on culturability and isolation of soil bacteria. Appl Environ Microbiol 71:826–834

De Carvalho CCCR, Caramujo MJ (2018) The various roles of fatty acids. Molecules 23:2583

Delgado-Baquerizo M, Oliverio AM, Brewer TE, et al (2018) A global atlas of the dominant bacteria found in soil. Science (80-) 359:320–325

Derrien M, Belzer C, de Vos WM (2017) *Akkermansia muciniphila* and its role in regulating host functions. Microb Pathog 106:171–181

Devos DP (2014) PVC bacteria: variation of, but not exception to, the Gram-negative cell plan. Trends Microbiol 22:14–20

Devos DP, Ward NL (2014) Mind the PVCs. Environ Microbiol 16:1217–1221

Duda VI, Suzina NE, Polivtseva VN, Boronin AM (2012) Ultramicrobacteria: formation of the concept and contribution of ultramicrobacteria to biology. Microbiology 81:379–390

Earl AM, Losick R, Kolter R (2008) Ecology and genomics of *Bacillus subtilis*. Trends Microbiol 16:269–275

Fairhead M, Thöny-Meyer L (2010) Role of the C-terminal extension in a bacterial tyrosinase. FEBS J 277:2083–2095

Freitas S, Hatosy S, Fuhrman JA, et al (2012) Global distribution and diversity of marine *Verrucomicrobia*. ISME J 6:1499–1505

Fuerst JA (2019) Phylum *verrucomicrobia*. Encycl Microbiol 4:551–563

Fuerst JA, Sagulenko E (2011) Beyond the bacterium: *planctomycetes* challenge our concepts of microbial structure and function. Nat Rev Microbiol 9:403–413

Giovannoni SJ, Tripp HJ, Givan S, et al (2005) Genome streamlining in a cosmopolitan oceanic bacterium. Science (80-) 309:1242–1245

Grant JR, Stothard P (2008) The CGView Server: a comparative genomics tool for circular genomes. Nucleic Acids Res 36:W181–W184

Green SK, Schroth MN, Cho JJ, et al (1974) Agricultural plants and soil as a reservoir for *Pseudomonas aeruginosa*. Appl Microbiol 28:987–991

Gregersen T (1978) Rapid method for distinction of Gram-negative from Gram-positive bacteria. Eur J Appl Microbiol Biotechnol 5:123–127

Halebian S, Harris B, Finegold SM, Rolfe RD (1981) Rapid method that aids in distinguishing Gram-positive from Gram-negative anaerobic bacteria. J Clin Microbiol 13:444–448

Hedlund BP, Gosink JJ, Staley JT (1997) *Verrucomicrobia* div. nov., a new division of the bacteria containing three new species of *Prosthecobacter*. Antonie Van Leeuwenhoek 72:29–38

Herlemann DPR, Lundin D, Labrenz M, et al (2013) Metagenomic de novo assembly of an aquatic representative of the verrucomicrobial class *Spartobacteria*. MBio 4:10–1128

Higgins DG, Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73:237–244

Hou S, Makarova KS, Saw JHW, et al (2008) Complete genome sequence of the extremely acidophilic methanotroph isolate V4, *Methylacidiphilum infernorum*, a representative of the bacterial phylum *Verrucomicrobia*. Biol Direct 3:1–25

Hug LA, Baker BJ, Anantharaman K, et al (2016) A new view of the tree of life. Nat Microbiol 1:1–6

Jain C, Rodriguez-R LM, Phillippy AM, et al (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114

Janssen PH, Yates PS, Grinton BE, et al (2002) Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteria*, *Actinobacteria,*

Proteobacteria, and *Verrucomicrobia*. Appl Environ Microbiol 68:2391–2396

Johansson ME V, Sjövall H, Hansson GC (2013) The gastrointestinal mucus system in health and disease. Nat Rev Gastroenterol Hepatol 10:352–361

Kant R, van Passel MWJ, Palva A, et al (2011) Genome sequence of *Chthoniobacter flavus* Ellin428, an aerobic heterotrophic soil bacterium

Kapoor D, Sharma P, Sharma MMM, et al (2024) Exploring soil microbiota and their role in plant growth, stress tolerance, disease control and nutrient immobilizer. Biocatal Agric Biotechnol 61:103358.

Kato S, Tobe H, Matsubara H, et al (2019) The membrane phospholipid cardiolipin plays a pivotal role in bile acid adaptation by *Lactobacillus gasseri* JCM1131^{T}. Biochim Biophys Acta (BBA)-Molecular Cell Biol Lipids 1864:403–412

Kato S, Yamagishi A, Daimon S, et al (2018) Isolation of previously uncultured slow-growing bacteria by using a simple modification in the preparation of agar media. Appl Environ Microbiol 84:e00807-18

Kimura M (1979) The neutral theory of molecular evolution. Sci Am 241:98–129

Klappenbach JA, Dunbar JM, Schmidt TM (2000) rRNA operon copy number reflects ecological strategies of bacteria. Appl Environ Microbiol 66:1328–1333

Konstantinidis KT, Rosselló-Móra R, Amann R (2017) Uncultivated microbes in need of their own taxonomy. ISME J 11:2399–2406

Lagkouvardos I, Joseph D, Kapfhammer M, et al (2016) IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. Sci Rep 6:33721

Lavergne C, Cabrol L, Cuadros-Orellana S, et al (2024) Rising awareness to improve conservation of microorganisms in terrestrial ecosystems: advances and future directions in soil microbial

diversity from Chile and the Antarctic Peninsula. Front. Environ. Sci. 12

Lee K-C, Webb RI, Janssen PH, et al (2009) Phylum *Verrucomicrobia* representatives share a compartmentalized cell plan with members of bacterial phylum *Planctomycetes*. BMC Microbiol 9:1–10

Lennon JT, Jones SE (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol 9:119–130

Lloyd KG, Steen AD, Ladau J, et al (2018) Phylogenetically novel uncultured microbial cells dominate earth microbiomes. MSystems 3:10–1128

Lok C (2015) Mining the microbial dark matter. Nature 522:270

Luscombe BM, Gray TRG (1974) Characteristics of *Arthrobacter* grown in continuous culture. Microbiology 82:213–222

Marquis B, Pillonel T, Carrara A, Bertelli C (2024) zDB: bacterial comparative genomics made easy. Msystems 9:e00473-24

Min L, Wang W, Oren A, et al (2023) *Coraliomargarita parva* sp. nov., isolated from mangrove sediment and genome-based analysis of the class *Opitutae* revealed five novel families: *Coraliomargaritaceae* fam. nov., *Pelagicoccaceae* fam. nov., *Cerasicoccaeae* fam. nov., *Oceanipulchritudinaceae* fam. Front Microbiol 14:1202141

Mise K, Iwasaki W (2020) Environmental atlas of prokaryotes enables powerful and intuitive habitat-based analysis of community structures. Iscience 23:

Murray AE, Lo C-C, Daligault HE, et al (2021) Discovery of an Antarctic ascidian-associated uncultivated *Verrucomicrobia* with antimelanoma palmerolide biosynthetic potential. Msphere 6:e00759-21

Nakai R (2020) Size matters: ultra-small and filterable microorganisms in the environment.

Microbes Environ 35:ME20025

Nakai R, Naganuma T, Tazato N, et al (2021) Characterization of *Terrihabitans soli* gen. nov., sp. nov., a novel 0.2 μm-filterable soil bacterium belonging to a widely distributed lineage of *Hyphomicrobiales* (*Rhizobiales*). Diversity 13:422

Nakai R, Nishijima M, Tazato N, et al (2014) *Oligoflexus tunisiensis* gen. nov., sp. nov., a Gram-negative, aerobic, filamentous bacterium of a novel proteobacterial lineage, and description of *Oligoflexacea*e fam. nov., *Oligoflexales* ord. nov. and *Oligoflexia* classis nov. Int J Syst Evol Microbiol 64:3353–3359

Nixon SL, Daly RA, Borton MA, et al (2019) Genome-resolved metagenomics extends the environmental distribution of the *verrucomicrobia* phylum to the deep terrestrial subsurface. Msphere 4:10–1128

Ottman N, Geerlings SY, Aalvink S, et al (2017a) Action and function of *Akkermansia muciniphila* in microbiome ecology, health and disease. Best Pract Res Clin Gastroenterol 31:637–642

Ottman N, Reunanen J, Meijerink M, et al (2017b) Pili-like proteins of *Akkermansia muciniphila* modulate host immune responses and gut barrier function. PLoS One 12:e0173004

Overmann J, Abt B, Sikorski J (2017) Present and future of culturing bacteria. Annu Rev Microbiol 71:711–730

Parks DH, Chuvochina M, Waite DW, et al (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996–1004

Parte A, Krieg NR, Ludwig W, et al (2011) Bergey's Manual of Systematic Bacteriology: Volume 4: The *Bacteroidetes*, *Spirochaetes*, *Tenericutes* (*Mollicutes*), *Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae*, and *Planctomycetes*. Springer Science & Business Media

Petroni G, Spring S, Schleifer K-H, et al (2000) Defensive extrusive ectosymbionts of *Euplotidium* (*Ciliophora*) that contain microtubule-like structures are bacteria related to *Verrucomicrobia*. Proc Natl Acad Sci 97:1813–1817

Pitt A, Schmidt J, Koll U, Hahn MW (2020) *Rariglobus hedericola* gen. nov., sp. nov., belonging to the *Verrucomicrobia*, isolated from a temperate freshwater habitat. Int J Syst Evol Microbiol 70:1830–1836

Plovier H, Everard A, Druart C, et al (2017) A purified membrane protein from *Akkermansia muciniphila* or the pasteurized bacterium improves metabolism in obese and diabetic mice. Nat Med 23:107–113

Powlson DS, Hirsch PR, Brookes PC (2001) The role of soil microorganisms in soil organic matter conservation in the tropics. Nutr Cycl Agroecosystems 61:41–51.

Reguera G (2016) "The Great Plate Count Anomaly" That Is No More. Co Microbes Ten Years Small Things Consid 288–291

Reiner K (2010) Catalase test protocol. Am Soc Microbiol 1:1–9

Riesco R, Trujillo ME (2024) Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol 74:6300

Rochman FF, Kim J-J, Rijpstra WIC, et al (2018) *Oleiharenicola alkalitolerans* gen. nov., sp. nov., a new member of the phylum *Verrucomicrobia* isolated from an oilsands tailings pond. Int J Syst Evol Microbiol 68:1078–1084

Rodriguez-R LM, Konstantinidis KT (2016) The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. PeerJ Preprints

Roller BRK, Stoddard SF, Schmidt TM (2016) Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. Nat Microbiol 1:1–7

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold spring harbor laboratory press

Schlesner H (1987) *Verrucomicrobium spinosum* gen. nov., sp. nov.: a fimbriated prosthecate bacterium. Syst Appl Microbiol 10:54–56

Schmitz RA, Peeters SH, Versantvoort W, et al (2021) Verrucomicrobial methanotrophs: ecophysiology of metabolically versatile acidophiles. FEMS Microbiol Rev 45:fuab007

Sheng L, Jena PK, Liu H-X, et al (2018) Obesity treatment by epigallocatechin-3-gallate–regulated bile acid signaling and its enriched *Akkermansia muciniphila*. FASEB J 32:6371

Shieh WY, Jean WD (1998) *Alterococcus agarolyticus*, gen. nov., sp. nov., a halophilic thermophilic bacterium capable of agar degradation. Can J Microbiol 44:637–645

Solden L, Lloyd K, Wrighton K (2016) The bright side of microbial dark matter: lessons learned from the uncultivated majority. Curr Opin Microbiol 31:217–226

Sun Y, Wilkinson BJ, Standiford TJ, et al (2012) Fatty acids regulate stress resistance and virulence factor production for *Listeria monocytogenes*. J Bacteriol 194:5274–5284

Tamaoka J, Katayama-Fujimura Y, Kuraishi H (1983) Analysis of bacterial menaquinone mixtures by high performance liquid chromatography. J Appl Microbiol 54:31–36

Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol 38:3022–3027

Torrella F, Morita RY (1981) Microcultural study of bacterial size changes and microcolony and ultramicrocolony formation by heterotrophic bacteria in seawater. Appl Environ Microbiol 41:518–527

Torsvik V, Øvreås L (2002) Microbial diversity and function in soil: from genes to ecosystems. Curr Opin Microbiol 5:240–245

Vartoukian SR (2016) Cultivation strategies for growth of uncultivated bacteria. J oral Biosci 58:143–149

Vollmers J, Frentrup M, Rast P, et al (2017) Untangling genomes of novel planctomycetal and verrucomicrobial species from monterey bay kelp forest metagenomes by refined binning. Front Microbiol 8:472

Wagner M, Horn M (2006) The *Planctomycetes, Verrucomicrobia, Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. Curr Opin Biotechnol 17:241–249

Wertz JT, Kim E, Breznak JA, et al (2012) Genomic and physiological characterization of the *Verrucomicrobia* isolate *Diplosphaera colitermitum* gen. nov., sp. nov., reveals microaerophily and nitrogen fixation genes. Appl Environ Microbiol 78:1544–1555

Wertz JT, Kim E, Breznak JA, et al (2018) Second Correction for Wertz et al.,"Genomic and Physiological Characterization of the *Verrucomicrobia* Isolate *Geminisphaera colitermitum* gen. nov., sp. nov., Reveals Microaerophily and Nitrogen Fixation Genes." Appl Environ Microbiol 84:e00952-18

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591

Yoon J, Matsuo Y, Matsuda S, et al (2007a) *Cerasicoccus arenae* gen. nov., sp. nov., a carotenoid-producing marine representative of the family *Puniceicoccaceae* within the phylum '*Verrucomicrobia*', isolated from marine sand. Int J Syst Evol Microbiol 57:2067–2072

Yoon J, Oku N, Matsuda S, et al (2007b) *Pelagicoccus croceus* sp. nov., a novel marine member

of the family *Puniceicoccaceae* within the phylum '*Verrucomicrobia*'isolated from seagrass. Int J Syst Evol Microbiol 57:2874–2880

Appendix

Appendix 1. PYG media composition for 1 litre

Components	Amount (g)
Solution A	
Basal salts ($[NH_4)_2SO_4]$, and $MgSO_4$)	15
Bacto agar	
Solution B	
Phosphate buffer ($[KH_2PO_4$), and Na_2HPO_4)	1.36 (KH_2PO_4; 20 mM (pH 7.0)) and 3.58 (Na_2HPO_4; 10 mM)
Solution C	
Bacto peptone	
Bacto yeast extract	0.1 (each)
Glucose	

For the PS medium, each solution (A, B, and C) was autoclaved individually before being combined (pH-7.4; autoclave for 20 min at 12°C).

Appendix 2. R2A media composition for 1 litre

Components	Amount (gm)
Protease peptone	0.50
Yeast extract	0.50
Casamino acids	0.50
Glucose	0.50

Soluble starch	0.50
Na-pyruvate	0.30
K_2HPO_4	0.30
$MgSO_4.7H_2O$	0.05
Agar	15

(pH- 7.4; autoclave for 20 min at 121°C)

Appendix 3. LB media composition for 1 litre

Components	**Amount (gm)**
Peptone	10
Yeast extract	5
NaCl	5
Agar	15

(pH- 7.4; autoclave for 20 min at 121°C)

Appendix 4. TSA media composition for 1 litre

Components	**Amount (gm)**
Casein peptone	15
Soya peptone	5
NaCl	5
Agar	15

(pH- 7.4; autoclave for 20 min at 121°C)

Appendix 5. List of genomes used in this study to investigate genome-based phylogenetic relationships, rrn operons, AAI values, and secondary metabolite gene clusters.

Accession number	Taxon name	Strain/bin name
GCA_036323555.1	*Congregicoccus parvus*	ASA1^T
GCA_000019965.1	*Opitutus terrae*	PB90-1^T
GCA_007559335.1	*Rariglobus hedericola*	53C-WASEF^T
GCA_000171235.2	*Geminisphaera colitermitum*	TAV2^T
GCA_002310495.1	*Nibricoccus aquaticus*	HZ-65^T
GCA_019464535.1	*Horticoccus luteus*	KSB-15^T
GCA_001746835.1	*Lacunisphaera limnophila*	IG16b^T
GCA_001580015.1	*Cephaloticoccus primus*	CAG34^T
GCA_001580045.1	*Cephaloticoccus capnophilus*	CV41^T
GCA_004118375.1	*Oleiharenicola lentus*	TWA-58^T
GCA_003096195.1	*Ereboglobus luteus*	Ho45^T
GCA_014230145.1	*Pelagicoccus albus*	YM14-201^T
GCA_014803405.1	*Pelagicoccus enzymogenes*	NFK12^T
GCA_016595505.1	*Pelagicoccus mobilis*	02PA-Ca-133^T
GCA_016595555.1	*Cerasicoccus arenae*	YM26-026^T
GCA_031296915.1	*Cerasicoccus frondis*	YM31-114^T
GCA_014230085.1	*Ruficoccus amylovorans*	CC-MHH0563^T
GCA_000025905.1	*Coraliomargarita akajimensis*	04OKA010-24^T
GCA_027257905.1	*Coraliomargarita parva*	WMMB3^T
GCA_003185655.1	*Coraliomargarita sinensis*	WN38^T
GCA_014230055.1	*Puniceicoccus vermicola*	IMCC1545^T
GCA_010500615.1	*Oceanipulchritudo coccoides*	CK1056^T
GCA_000172155.1	*Verrucomicrobium spinosum*	DSM 4136^T

Class	Family	Isolation source
Opitutia	*Alterococcaceae*	Garden soil
Opitutia	*Opitutaceae*	Rice paddy soil
Opitutia	*Opitutaceae*	Freshwater ditch
Opitutia	*Opitutaceae*	Termite gut
Opitutia	*Opitutaceae*	Hyporheic freshwater
Opitutia	*Opitutaceae*	Orchard soil
Opitutia	*Opitutaceae*	Freshwater lake
Opitutia	*Opitutaceae*	Ant gut
Opitutia	*Opitutaceae*	Ant gut
Opitutia	*Opitutaceae*	Irrigation water
Opitutia	*Opitutaceae*	Cockroach hindgut
Opitutia	*Pelagicoccaceae*	Sea water
Opitutia	*Pelagicoccaceae*	Marine sediment
Opitutia	*Pelagicoccaceae*	Sea water
Opitutia	*Cerasicoccaceae*	Marine sediment
Opitutia	*Cerasicoccaceae*	Seawater
Opitutia	*Cerasicoccaceae*	Fermenter
Opitutia	*Coraliomargaritaceae*	Sea water
Opitutia	*Coraliomargaritaceae*	Mangrove sediment
Opitutia	*Coraliomargaritaceae*	Marine sediment
Opitutia	*Puniceicoccaceae*	Marine clamworm
Opitutia	*Oceanipulchritudinaceae*	Coastal sediment
Verrucomicrobiia	*Verrucomicrobiaceae*	Lake water

Genome size	No. of contigs	GC content (%)[a]	No. of rRNA genes[a]	No. of rRNA operon[b]	Incubation time (days)
5,821,695	1	65.6	3	1	8-14
5,957,605	1	65.3	3	1	8-14
4,146,955	16	60.6	3	1	3
5,671,497	225	60.9	3	1	8-14
4,730,447	1	62.2	3	1	28
4,320,198	1	64.3	3	1	4-5
4,199,284	1	66.5	3	1	Not specified
2,356,970	72	63.0	3	1	30
2,092,503	66	60.5	3	1	30
4,713,391	2	65.3	3	1	3
4,175,033	1	59.7	3	1	10
4,768,689	13	52.3	12	4	30
6,243,759	61	56.8	12	4	7
7,518,465	168	53.5	3	1	Not specified
4,127,484	129	52.6	2	1	30
5,805,120	69	54.6	2	1	30
4,506,184	64	60.3	3	1	3
3,750,771	1	53.6	6	2	Not specified
4,497,538	27	56.0	2	1	14
3,569,655	36	54.7	3	1	7
5,204,456	143	54.7	3	1	14
3,520,043	52	54.0	3	1	30
8,220,857	1	60.3	12	4	Not specified

Referemces for incubation time	Average amino acid identity (AAI) value against ASA1^{T}
this study	100.0
https://bacdive.dsmz.de/strain/11301	53.8
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.003980	52.0
https://bacdive.dsmz.de/strain/131412	51.2
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.003198	54.0
https://link.springer.com/article/10.1007/s00284-022-03036-8#Sec9	53.1
https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2017.00202/full#h3	53.4
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.001141	52.9
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.001141	53.4
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.004191	53.7
https://www.sciencedirect.com/science/article/abs/pii/S0723202017301753?via%3Dihub	51.1
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.64970-0	48.4
https://www.frontiersin.org/articles/10.3389/fmars.2021.655735/full	49.3
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.64970-0	47.8
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.65102-0	46.8
https://www.jstage.jst.go.jp/article/jgam/56/3/56_3_213/_article	46.0
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.001723	47.1
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.64755-0	46.8
https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2023.1202141/full#h3	47.0
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.003205	47.1
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.64616-0	45.4
https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.004458?crawler=true	47.1
https://www.sciencedirect.com/science/article/pii/S0723202087800103	42.4

Biosynthetic gene clusters (BGC) counts detected by antiSMASH ver 7.0.1[c]

NRPS	PKS	RiPP	T3PKS	RRE-containing	T1PKS	Arylpolyene	RiPP-like
2	0	0	1	0	0	1	0
1	0	0	1	0	1	0	1
0	0	0	1	0	0	1	0
0	0	0	0	0	0	1	0
0	0	0	0	0	0	1	1
0	0	0	1	1	0	1	1
0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	0
1	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0
0	0	0	1	0	0	1	0
1	0	0	1	1	0	1	0
0	0	0	0	0	0	1	0
0	0	0	1	0	0	0	0
0	0	0	0	0	0	1	0
0	0	0	1	0	0	0	0
0	0	0	1	0	1	1	0
1	0	0	1	0	0	0	0
0	0	0	2	1	1	0	1

NRPS-like, T1PKS	NRPS-like	Lassopeptide	Phosphonate	Betalactone	Ectoine	Terpene	NAGGN
1	1	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	0	0	0	0	3	1
0	1	0	0	1	0	3	1
0	0	0	0	0	0	2	1
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0
0	0	0	0	0	0	1	1
0	1	0	0	0	0	0	0
0	0	0	0	0	0	3	0

[a] Genome data not available in EzBioCloud Database (https://www.ezbiocloud.net/) were manually annotated using the DFAST pipeline (https://dfast.ddbj.nig.ac.jp/).

[b] Draft genomes with only two rRNA genes detected were considered to possess one rrn operon.

[c] NRPS, Nonribosomal peptide synthetase; PKS, Polyketide synthases; RiPP, Ribosomally synthesized and post-translationally modified peptide; T3PKS, Type III polyketide synthases; RRE-containing, RiPP recognition element; T1PKS, Type I polyketide synthase; NAGGN, n-acetyl glutaminyl glutamine amide.