ARTICLE

Unveiling the genomic landscape: Whole genome draft and phenotypic profiling of *Oceanobacillus* sp. SE10311 isolated from Mount Makiling, Philippines

Noel H. Tan Gana*[1], Cernan P. Ruz[1], Eufrocinio C. Marfori[1], Edwin P. Alcantara[1], Maria Genaleen Q. Diaz[2], Lucille C. Villegas[2], Eldrin DLR. Arguelles[1], Armi R. Creencia[1], Neil H. Tan Gana[3,4], Rosario G. Monsalud[1], and Rina B. Opulencia[2]

[1]National Institute of Molecular Biology and Biotechnology, University of the Philippines Los Baños
[2]Institute of Biological Sciences, College of Arts and Sciences, University of the Philippines Los Baños
[3]Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines Manila
[4]Department of Arts and Sciences, City College of Calamba

ABSTRACT

The genus *Oceanobacillus* remains one of the least understood groups of *Bacillaceae* due to limited information and a scarcity of accessions. With the full genome sequencing and partial phenotypic characterization of the cellulase and brown-pigment-producing *Oceanobacillus* sp. SE10311, an isolate from Mount Makiling, this study provided taxonomic, genomic, and functional insights into this bacterial discovery. The results showed that *Oceanobacillus* sp. SE10311 draft genome contained 3.946 million base pairs assembled from 137 contigs and 39.30 percent G+C. The whole genome phylogeny and pangenome analyses revealed the identity of *Oceanobacillus* sp. SE10311 as an isolate of *Oceanobacillus indicireducens* with greater than 98% genome similarity to the type strain of the species. The confirmed presence of the associated genes from the genome annotation rationalized the pigmentation and cellulose degradation of *Oceanobacillus* sp. SE10311. This research milestone presents *Oceanobacillus* sp. SE10311, with distinct chromogenic and cellulolytic properties as a completed whole genome-sequenced bacterial accession from the Philippines.

INTRODUCTION

Mount Makiling, situated at 14.136389 latitude and 121.194444 longitude, is a biodiversity-rich dormant volcano fifty

KEYWORDS

Oceanobacillus, antimicrobial activity, Illumina sequencing technology, cellulase activity

*Corresponding author
Email Address: nhtangana@up.edu.ph
Date received: January 5, 2024
Date revised: April 16, 2024
Date accepted: June 30, 2024
DOI: https://doi.org/10.54645/202417SupNVF-47

kilometers south of Manila, Luzon Island, The Philippines. A significant part of the mountain, managed as a reserve by the University of the Philippines Los Baños, serves as a habitat for unique flora, fauna, and microorganisms (Magcale-Macandog et al. 2020, Gonzalez et al. 2020). The microenvironment of Mt. Makiling drove the evolution of distinct biological assemblages across various habitats. Interestingly, numerous studies indicated Mount Makiling as a steady source of novel microbial species, including archaea (Lantican et al. 2011, Montecillo et al. 2019), bacteria (Gatpatan et al. 2021), and fungi (Solis et al. 2015). This study is no exception, and it continuously proves Mount Makiling is a vibrant host to rare bacteria, as our group isolated a bacterium we identified as *Oceanobacillus.*

The *Oceanobacillus* is a remarkable genus in the family *Bacillaceae* for its diversity and adaptability in distinctive environments. Reports showed that *Oceanobacillus* may live in fermented products (Namwong et al. 2009, Whon et al. 2010, Akuzawa et al. 2016, Hirota et al. 2013b, Hyun et al. 2015, Akuzawa et al. 2016), plant crops rhizospheres (Albdaiwi et al., 2019, Alhindi and Albdaiwi, 2022), marine ecosystems (Lu et al. 2001, Yu et al. 2014, Seghal Kiran et al. 2014, Yongchang et al. 2015, Yu et al. 2019, Subramanian and Maruthamuthu, 2019), salt lakes (Rohban et al. 2009, Amoozegar et al. 2014, Amoozegar et al. 2016a, Amoozegar et al. 2016b, Didari et al. 2020, Zhu et al. 2020), saline-alkali soils (Gan et al. 2018, Gan et al. 2021, Jin et al. 2022, OuYang et al. 2022), gut microbiota (Kumar Mondal et al. 2017), dairy products (Mayr et al. 2006), food production equipment (Tominaga et al. 2009), mining sites (Mwandira et al. 2019, Mwandira et al. 2020) among others. Despite these reports, the presence of *Oceanobacillus* in the Philippines still awaits confirmation backed up by genomic sequencing and phenotypic data.

While genomic information about *Oceanobacillus* has grown considerably in recent years, as described in 81 NCBI whole genome accessions (before this report), much work remains. The ecological and functional significance of *Oceanobacillus* among microbes represents an area ripe for further discovery and exploration. The current collection of *Oceanobacillus* presents interesting functional properties like antifungal (Alhindi and Albdaiwi 2022), antimicrobial, and remarkable tolerance to extreme environmental conditions.

In this study, we reported the draft genome of *Oceanobacillus* sp. SE10311, notably with brown pigmentation **and** cellulolytic activity. Average Nucleotide Identity (ANI) revealed that the isolate has >98% genome similarity to *Oceanobacillus indicireducens* JCM 17251^T. Pangenome analysis of the isolate indicated close associations of the core and accessory genomes to *Oceanobacillus indicireducens*. The phenotypic characters found to have putative genes concurrently expressed in the draft genome of the *Oceanobacillus* sp. SE10311 offers rationalizations for its cellulase production and brown pigmentation. Other bioactivity-related genes found that are of possible interest include arsenic metallo-regulation and vancomycin resistance activities. This research milestone chronicled the isolation and characterization of *Oceanobacillus* sp. from Mount Makiling, providing insights into its genome architecture, functional properties, and ecology.

MATERIALS AND METHODS

Collection of the Strain Used in the Study

Oceanobacillus sp. SE10311 is a bacterial accession (BIOTECH 10857) from the Philippine National Collection of Microorganisms, National Institute for Molecular Biology and Biotechnology, University of the Philippines Los Baños, College, Los Baños, Laguna, Philippines. It was isolated from the Mt. Makiling Forest Reserve, Philippines (14.102713, 121.201460), particularly from Sipit Creek sediments. Sampling occurred within a five-meter by five-meter plot transect on 10 September 2019. To maintain sample integrity, the collected specimens were transported on ice to the laboratory and stored at 4°C before plating. Analyses were conducted within six hours of sample collection.

Sediment samples, suspended in sterile distilled water, were serially diluted up to 10^{-6}. Aliquots of 100 µl were spread plated in 10 replicate plates of modified peptone yeast extract glucose agar (mPYGA) without phosphate incubated at 30°C for 30 days (Tanaka et al. 2014). The plates were supplemented with 10 ppm cycloheximide. Colonies that grew after three days of incubation were randomly selected and purified on mPYGA. Pure cultures were maintained on tryptic soy agar (TSA).

Phenotypic Characterization and Related Assays

In conjunction with a colorimetric reagent card (BCL), the Biomerieux VITEK 2 Compact Identification System determined the bacterial isolate's partial identity and phenotypic characteristics through an automated growth-based system. The BCL contains 64 wells with an individual test substrate measuring various metabolic activities, for example, acidification, alkalinization, and enzymatic hydrolysis, that specifically determines characteristics of spore-forming Gram-positive bacilli and related genera. The pure bacterial colonies were initially grown TSA for 48 hours, 30. Cells were harvested using sterile swabs, and the cells were suspended in 3.0 mL sterile saline (0.45% NaCl, pH 4.5-7.00) in clear polystyrene tubes (within the 1.80-2.20 McFarland Turbidity Range MTR checked using a DensiCheck® reader. The tube and a BCL card were placed into the designated carousel and loaded into the VITEK 2 Compact Identification System. The test reactions proceeded with calculations comparing raw data and reaction thresholds for each test spanning 11-18 hrs (for BCL cards). The results were run against the installed VITEK 2 databases of well-characterized microbial strains tested in various culture conditions. The analytical workflow tests the data from the unknown culture compared to the database, determining a quantitative value for proximity to a database taxon. Follow-up assays further characterizing the phenotype included pigment production, lipase activity, and cellulase activity. Water-soluble pigment production was monitored in TSA cultures grown for 1-2 weeks at 30°C.

Observations for lipase activity were made on 7-14-day-old streak plates of the isolates grown in Sierra's medium with Tween 80 at 30°C. A positive result for lipolytic activity was indicated by the presence of precipitation zones on the medium around the inoculation area.

Cellulase activity was monitored by flooding with 1% Congo red solution and subsequent washing with 1M NaCl of 7-14 day-old streak plates of the isolates grown in Carboxymethyl Cellulase Agar at 30°C. The presence of clearing zones around the inoculation area indicates cellulase production.

Genomic DNA Isolation, Library Preparation, and Sequencing

Bacterial total genomic DNA (gDNA) was prepared from a verified single colony isolate grown on liquid tryptic soy broth medium for 72 hours using the NucleoSpin® Microbial DNA Isolation Kit (Macherey-Nagel), following the manufacturer's instructions. The gDNA quantification, quality assessment, and sequencing were done at Macrogen Inc. (Seoul, South Korea). Specifically, the gDNA generated libraries with TruSeq Nano DNA (350) proceeded with sequencing via the Illumina NextSeq® platform, resulting in short sequence paired-end reads at 100 bp length. Initially, the purified DNAs were sent to Macrogen Inc., which met quality control standards, to proceed

to library preparation. The SE10311 DNA underwent random fragmentation and ligation with TruSeq Nano DNA 350 adapters. The SE10311 ligated fragments obtained 70nM concentration with sizes 639 bp (Supplemental Figure 1A) were loaded into the flow cell for cluster generation, and the fragments were amplified into clonal clusters via bridge amplification, which ultimately served as sequencing templates. The Illumina NextSeq® 1000 platform proceeded sequencing by synthesis. This proprietary technology allowed terminator-based sequencing to accurately detect single nucleotide addition into DNA template strands as a massively parallel sequencing procedure. The generation of raw reads underwent a real-time analysis of sequence control and base calling (BCL), which then converted the BCL files into FASTQ files with adapters. The SE10311 yielded 5,204,444 total reads from 525,648,844 bp total bases, with 39.1% GC, 99.3% Q20, and 98.6% Q30 values (Table 1, Supplemental Figure 1B). The Q20 and Q30 values indicate the ratios of bases with PHRED quality scores of over 20 (1/100 error rate) and over 30 (1/1000 error rate).

Downstream Processing of Sequence Reads
The sequence reads underwent downstream processing, which included 1) preprocessing, quality trimming and variant calling; 2) *de novo* assembly; 3) gene calling and annotation; 4) genome mapping and plot construction; and 5) genome completeness auditing. The FastQC (Andrews 2010) and Trim Galore version 0.6.7 (Krueger 2015) programs assessed the short sequence reads, trimmed adaptor sequences, and removed bad reads. The quality reads are then assembled *de novo* assembly using SPAdes V3.13.0 (Bankevich et al. 2012), followed by polishing using the Pilon version 1.23 (Walker et al., 2014), implemented in UniCycler 0.5 (Wick et al. 2017). The software checkM v. V1.2.2 (Parks et al. 2015) and BUSCO (Manni et al. 2021) assessed genome completeness, while QUAST V5.2.0 (Gurevich et al. 2013) generated the assembly statistics. MGCplotter (Shimoyama 2022) developed the genome plot map (Supplemental Figure 3). Gene calling and annotation of the draft genome utilized the Prokka V1.14.5 (Seemann, 2014) for amino acid sequence prediction, followed by mapping into different functional categories like KO, COG, and enzymes using the eggnog mapper V-2.1.3 (Heurta-Cepas et al. 2017) against the eggnog Database 5 (http://eggnog5.embl.de/).

Comparative Genomics Analyses and Functional Annotation
The United States Department of Energy Systems Biology Knowledgebase (KBase) (Arkin et al. 2018), a repository of bioinformatics platforms, generated comparative genomics reports, particularly phylogenetics and pangenomics analyses. The FastANI v0.13 tool (Jain et al. 2018) was used to compute the Average Nucleotide Identity (ANI) of whole genomes, relying on the identities of orthologous genes among genomes to demarcate interspecies relations. Parallel runs of Classify Microbes with GTDB-Tk - v1.7.0 (Chaumeil et al. 2020) revealed the taxonomic identity of *Oceanobacillus* sp. SE10311. Further, an extended phylogenetic tree construction for the whole genomes of *Oceanobacillus* species was carried out using the Species Tree Builder and the Type (Strain) Genome Server (TYGS) (Meier-Kolthoff and Göker, 2019) accessible at https://tygs.dsmz.de.

The pangenome construction utilized the different modules of the KBase (Arkin et al. 2018), which included Build Pangenome with OrthoMCL - v2.0v0.0.8 (Chen at al. 2006). This genome-wide algorithm grouped orthologous protein sequences for complete annotation of genomes. In addition, the Genomes from Pangenome v0.0.7 tool allowed genomic comparison of the genome in question against a given set of reference genomes, considering their functions and sequences. This process organized the homologous protein families with similar putative functions, distinguishing genomes for improved genome characterization. Specifically, the comparison classified proteins into families with potential activities, validating homologous protein groups with inconsistent annotations or identifying consistently annotated proteins lacking substantial sequence similarity. The tool determined the extent of gene conservation among an input set of genomes, comprehended genome evolution and adaptation to their unique environments, and identified novel genes and fascinating biology in the context of related genomes. The antiSmash 7.0 (Blin et al. 2023) web version scanned the genomes for secondary metabolite gene clusters with secondary referencing to published *in silico* predictors of secondary metabolites.

Accessions in NCBI
The genome data, including the short sequence reads and the draft assembly, are accessible at GenBank with accessions BioProject ID PRJNA1045277, BioSample ID SAMN38440221, SRA ID SRR27027422, and Genome ID JAXHDB000000000. Unless specified, this manuscript used the Version 01 of the NCBI data.

RESULTS AND DISCUSSION

The draft genome of *Oceanobacillus* sp. SE10311 contained 137 contigs totaling 3.946 million bp (mbp) and a G+C content of 39.30 percent. The *Oceanobacillus* sp. SE10311 genome size falls within the 3.343 (*Oceanobacillus halotolerans*) (Zhu et al. 2020) to 5.1679 (*Oceanobacillus jeddahense*) (Khelaifia et al. 2016) mbp range as sizes of *Oceanobacillus* species (4.155 mbp average) in the NCBI Genomes Database (accessed August 2023). *Oceanobacillus limi* IBRC-M 10780 (Amoozegar et al. 2014), with a size of 3.914 mbp, is the closest in size to *Oceanobacillus* sp. SE10311. The G + C content also approximates the average for members of *Oceanobacillus*. The projected number of coding genes was 3,880, consisting of 3,822 protein-coding genes. Additionally, the genome is anticipated to include 54 transfer RNAs (tRNAs) and three ribosomal RNAs (rRNAs), comprising seven 5S, one 16S, and one 23S. The primary characteristics of the genome are listed in Table 1.

Table 1: Genomic profiles of *Oceanobacillus sp.* SE10311 by Prokka v. V1.14.5 (Seemann 2014) of the NCBI Prokaryotic Genome Annotation Pipeline (PGAP), annotation date 11/30/2023 20:03:37

Genome Features	Values
Total bases, bp (Mb)	5,204,444 bp (520.4 Mb)
Number of raw reads	525,648,844
Sequence read coverage	250.0x
Number of reads after trimming (Q20/Q30, %)	99.3/98.6
Genome size, bp (Mb)	3,946,157 bp (3.9 Mb)
G + C content (%)*	39.5
N_{50} contig length, bp (kb)	61,765 bp (61.8 kb)
L_{50} contig	18
Number of contigs	137
CheckM analysis v1.2.2 genome completeness	67.96%
CheckM analysis v1.2.2 genome contamination	13.15%
BUSCO genome completeness*	99.19% (123/124)

Number of genes (total)	3,913
Number of CDS (total)	3,851
Number of genes (coding)	3,767
Number of CDS (with protein)	3,767
Number of genes (RNA)	62
rRNAs	1, 1, 1 (5S, 16S, 23S)
tRNAs	54
ncRNAs	5
Pseudo genes	84
CRISPR Arrays	2
BioProject accession	PRJNA1045277
SRA accession	SRX22719256 SRR27027422
BioSample accession	SAMN38440221
GenBank accession	JAXHDB000000000

Notes: * Refer to Supplemental Figure 2. for further details

Many of the predicted genes with annotation are involved in the metabolism of amino acids, namely active arylamidase activities, indicating several catalytic hydrolysis paths of N-terminal amino acids in amides, arylamides, and peptides, contrasting to the isolate's limited abilities for carbohydrate metabolism. Such observations correspond to the affinity of *Oceanobacillus* species to thrive in limiting and extreme environments where alternative energy sources are utilized, such as amino acids. In some instances, *Oceanobacillus* sp. carries carbohydrate-related genes like xylose, fructose, chitin, etc. More limitedly, the *Oceanobacillus jordanicus* (Alhindi and Albdaiwi, 2022) exhibit photosynthetic electron transporters, quinone cofactors, phylloquinone, menaquinone biosynthesis on top of the genes involved in the acquisition of nitrogen, sulfur, phosphate, iron, and plant hormone auxin for marked plant-bacteria symbiotic interaction. These substrate utilization traits demonstrate the diversity of *Oceanobacillus* species. Table 2 presents the phenotypic features of *Oceanobacillus* sp. SE10311 relative to carbon and nitrogen utilization, along with the results of other related assays.

Table 2: Phenotypic profiles of *Oceanobacillus* sp. SE10311.

Phenotypic Test	SE10311
β-Xylosidase	-
L-Lysine-Arylamidase	+
L-Aspartate Arylamidase	+
Leucine Arylamidase	(-)
Phenylalanine Arylamidase	+
L-Proline Arylamidase	-
β-Galactosidase	-
L-Pyrrolidonyl-Arylamidase	+
α-Galactosidase	-
Alanine Arylamidase	+
Tyrosine Arylamidase	+
β-N-Acetyl-Glucosaminidase	-
Ala-Phe-Pro Arylamidase	+
Cyclodextrine	-
D-Galactose	-
Glycogene	-
myo-Inositol	-
Methyl-α-D-Glucopyranoside Acidification	-
Ellman	+
Methyl-D-Xylosidase	-
α-Mannosidase	-
Maltotriose	-
Glycine Arylamidase	-
D-Mannitol	-
D-Mannose	-
D-Melezitiose	-
N-Acetyl-D-Glucosamine	-
Palatinose	-
L-Rhamnose	-
β-Glucosidase	-
β-Mannosidase	-
Phosphoryl Choline	-
Pyruvate	-
α-Glucosidase	-
D-Tagatose	-
D-Trehalose	-
Inulin	-
D-Glucose	-
D-Ribose	-
Putrescine Assimilation	-
Growth in 6.5% NaCl	-
Kanamycin Resistance	-
Oleandomycin Resistance	-
Esculin Hydrolysis	-
Tetrazolium Red	-
Polymyxin B Resistance	-
Additional Tests	
Cellulase*	+
Lipase	-
Pigment**	brown

Notes:

(-) = weak negative, reaction slightly above detection threshold
* Genome annotation auditing conducted for the presence of cellulolytic genes
** Genome annotation auditing conducted for the presence of chromogenic genes

As shown in Figure 2, the whole-genome phylogenetic analysis demonstrates the close relationship of *Oceanobacillus* sp. SE10311 with *Oceanobacillus indicireducens* JCM 17251^T (Hirota et al. 2013a). Genomic signatures corroborated the evidence that placed *Oceanobacillus* sp. SE10311 as a close relative of *Oceanobacillus indicireducens* JCM 17251^T and *Oceanobacillus* sp. AG. The FAST Average Nucleotide Identity (ANI) computation (Table 3) showed the same result. The whole phylogenomic alignments and the FAST ANI reports differ in the most closely identified relative of *Oceanobacillus* sp. SE10311. Such variations may stem from differences in the algorithms employed by the two methods. With less than one percent differences in the ANI values, the relational proximity of the three strains is indistinguishable. Further, the Type (Strain) Genome Server (TYGS) digital DNA: DNA hybridization (dDDH) values confirmed the identity of *Oceanobacillus* sp. SE10311 as strain of *Oceanobacillus indicireducens* at 79.4% D_4, dDDH. Figure 3 shows the generated GBDP trees related to the TYGS analysis. The similarities in taxonomic genome features, such as size and G+C content, led us to conclude that the identified isolate represents a new strain of *Oceanobacillus indicireducens*.

Table 3: FAST ANI (Arkin et al. 2018) computed Average Nucleotide Identity (ANI) estimation of the *Oceanobacillus* sp. SE10311 genome compared to the most closely associated *Oceanobacillus* sp. as identified in the whole genome phylogenetic analysis.

Query	Reference	ANI Estimate	Matches	Total
Oceanobacillus alkalisoli APA_J-2 (6-2)	*Oceanobacillus* sp. AG	85.2697	830	1217
Oceanobacillus sp. AG	*Oceanobacillus alkalisoli* APA_J-2 (6-2)	85.3008	830	1230
Oceanobacillus alkalisoli APA_J-2 (6-2)	Oceanobacillus sp. SE10311	85.3299	840	1217
Oceanobacillus_sp._SE10311	*Oceanobacillus alkalisoli* APA_J-2 (6-2)	85.3468	850	1245
Oceanobacillus alkalisoli APA_J-2 (6-2)	*Oceanobacillus indicireducens* JCM17251	85.4871	855	1217
Oceanobacillus indicireducens JCM17251	*Oceanobacillus alkalisoli* APA_J-2 (6-2)	85.4955	853	1248
Oceanobacillus indicireducens JCM17251	*Oceanobacillus* sp. AG	97.2239	1024	1248
Oceanobacillus sp. AG	*Oceanobacillus indicireducens* JCM17251	97.2887	1020	1230
Oceanobacillus indicireducens JCM17251	Oceanobacillus sp. SE10311	97.3568	1017	1248
Oceanobacillus sp. SE10311	*Oceanobacillus indicireducens* JCM17251	97.3642	1022	1245
Oceanobacillus sp. AG	Oceanobacillus sp. SE10311	98.1694	1020	1230
Oceanobacillus sp. SE10311	*Oceanobacillus* sp. AG	98.1814	1028	1245

The pangenome analysis performed in two steps showed the collection of all genes found among related organisms grouped by sequence homology. In the first step, the construction focused on identifying genes that are the most common among the genomes of *Oceanobacillus* sp., including the *Oceanobacillus* sp. SE10311 (Figure not shown). Specifically, when *Oceanobacillus* sp. SE10311 and its clonal isolate, *Oceanobacillus* sp. MO10714-A were compared with the 37 NCBI *Oceanobacillus* reference genomes, and a total of 27,068 families were identified. Among these, 17,213 were homolog families, and 9,855 were singleton families. The analysis encompassed a total of 147,594 translatable genes, with 137,739 being classifiable into homolog families.

In the second step, the pangenome construction is composed of *Oceanobacillus* sp. SE 10311 and closely related species, namely, *Oceanobacillus* sp. AG, *Oceanobacillus indicireducens* JCM 17251^{T}, *Oceanobacillus alkalisoli* APJA-2^{T} with *Oceanobacillus iheyensis* HTE831 as an outer group identified during the whole-genome alignment stage (Figure 4). This comparison highlighted the distinct features of *Oceanobacillus* sp. SE 10311 against its closely identified strains. Specifically for *Oceanobacillus* sp. SE10311, the analysis revealed 3,843 homologous genes out of 3,870 genes. Further comparison of the *Oceanobacillus* sp. SE10311 genome against the *Oceanobacillus* sp. pangenome identified 52 singletons (unique sequences, Table 4), providing additional evidence suggesting *Oceanobacillus* sp. SE10311 as a distinct isolate of *Oceanobacillus indicireducens*. The pangenome analysis further confirmed that the earlier whole genome alignment result was closely associated with *Oceanobacillus* sp. SE10311 with *Oceanobacillus indicireducens* JCM 17251^{T} and *Oceanobacillus* sp. AG. The reported high number of singletons in *Oceanobacillus* sp. SE10311 implies interesting biological and functional importance. The singletons suggest horizontal gene transfer events, gene acquisition, or loss outcomes. The *Oceanobacillus* sp. SE10311, as isolated from volcanic soil, may have to expand genomes due to increased opportunities for gene acquisition to find new traits enhancing strain flexibility and fitness (Cummins et al. 2022).

Table 4: Unique protein functions found in *Oceanobacillus* sp. SE10311.

Representative Protein Function	Count of Representative Function	Count of SE10311 (153310/2/1)
Transporter, LysE family	0.05%	2
Site-specific recombinase	0.03%	2
Transcriptional regulator, Xre family	0.03%	2
Type III restriction-modification system DNA endonuclease res (EC 3.1.21.5)	0.05%	2
Putative amino-acid transporter YisU, LysE family	0.05%	2
FIG01233583: hypothetical protein	0.03%	2
Acetylornithine deacetylase (EC 3.5.1.16)	0.03%	2
β-ureidopropionase (EC 3.5.1.6)	0.03%	2
Transketolase, N-terminal section (EC 2.2.1.1)	0.02%	1
Vancomycin (or other glycopeptides) histidine kinase VanS => VanF/M-type	0.02%	1
Streptogramin B lyase (EC 4.2.99.-) => Vgb(B)	0.02%	1
Transposase subunit	0.02%	1
transposase, IS605 family	0.02%	1
Sulfite reductase, assimilatory-type (EC 1.8.-.-)	0.02%	1
Tagatose-6-phosphate kinase (EC 2.7.1.144); 1-phosphofructokinase (EC 2.7.1.56)	0.02%	1
ThiJ/PfpI family protein	0.03%	1

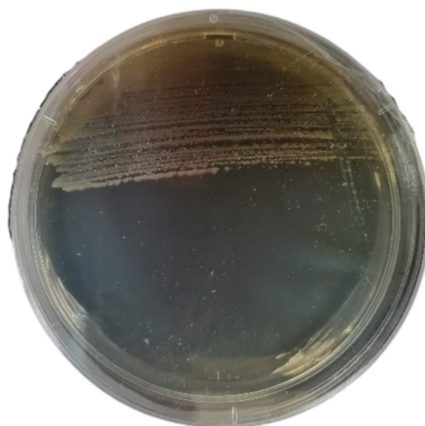
Uncharacterized membrane protein YmcC	0.02%	1
Serine/threonine protein phosphatase (EC 3.1.3.16)	0.02%	1
YoeB toxin protein	0.02%	1
Ribulose-5-phosphate 4-epimerase and related epimerases and aldolases	0.02%	1
Vancomycin (or other glycopeptides) response regulator VanR => VanF/M-type	0.02%	1
Zinc ABC transporter, ATP-binding protein AdcC	0.02%	1
Zinc ABC transporter, permease protein AdcB	0.02%	1
Putative EsaC protein analog (Listeria type 3)	0.02%	1
putative lantibiotic ABC transporter, permease protein	0.02%	1
Repressor CsoR of the copZA operon	0.02%	1
Prophage Clp protease-like protein	0.02%	1
Putative glycosyltransferase	0.02%	1
Putative sodium-dependent transporter	0.02%	1
Permease	0.02%	1
Lipid A export ATP-binding/permease protein MsbA	0.02%	1
Integrase	0.02%	1
Integrase/recombinase, RitC	0.02%	1
FIG01114076: hypothetical protein	0.02%	1
FIG01228292: hypothetical protein	0.02%	1
FIG01240914: hypothetical protein	0.02%	1
FIG01247172: hypothetical protein	0.02%	1
FIG131328: Predicted ATP-dependent endonuclease of the OLD family	0.02%	1
hypothetical cytosolic protein	0.02%	1
D-Ala-D-Ala dipeptidase/carboxypeptidase (EC 3.4.16.4)(EC 3.4.13.22) => VanXY-unclassified	0.02%	1
Carbohydrate kinase, FGGY family	0.02%	1
Acetyltransferase, GNAT family	0.02%	1
Acetylxylan esterase related enzyme	0.02%	1
α-xylosidase (EC 3.2.1.177)	0.02%	1
Anhydro-N-acetylmuramic acid kinase (EC 2.7.1.170)	0.02%	1
Arsenical resistance operon repressor	0.02%	1
Arsenite/antimonite pump-driving ATPase ArsA (EC 3.6.3.16)	0.02%	1
Arsenic metallochaperone ArsD, transfers trivalent metalloids to ArsAB pump	0.02%	1
Class II aldolase/adducin domain protein	0.02%	1
COGs COG2843	0.02%	1
Cyanophycin synthetase (EC 6.-.-.-)	0.02%	1
Cadmium efflux system accessory protein	0.02%	1

In the second step, the pangenome construction is composed of *Oceanobacillus* sp. SE 10311 and closely related species, namely, *Oceanobacillus* sp. AG, *Oceanobacillus indicireducens* JCM 17251^{T}, *Oceanobacillus alkalisoli* APJA-2^{T} with *Oceanobacillus iheyensis* HTE831 as an outer group identified during the whole-genome alignment stage (Figure 4). This comparison highlighted the distinct features of *Oceanobacillus* sp. SE 10311 against its closely identified strains. Specifically for *Oceanobacillus* sp. SE10311, the analysis revealed 3,843 homologous genes out of 3,870 genes. Further comparison of the *Oceanobacillus* sp. SE10311 genome against the *Oceanobacillus* sp. pangenome identified 27 singletons (unique sequences, Table 4), providing additional evidence suggesting *Oceanobacillus* sp. SE10311 as a distinct isolate of *Oceanobacillus indicireducens*. The pangenome analysis further confirmed that the earlier whole genome alignment result was closely associated with *Oceanobacillus* sp. SE10311 with *Oceanobacillus indicireducens* JCM 17251^{T} and *Oceanobacillus* sp. AG.

The phenotypic assays presented two noteworthy properties of *Oceanobacillus* sp. SE10311 (Table 2), prompting further validation of the genome annotation to confirm the existence of genes expressing these phenotypes. The first distinctive functional property observed in SE10311 is the production of a soluble brown pigment (Figure 1A). The annotated genes linked to the brown pigmentation of *Oceanobacillus* sp. SE10311 include cyanophycin synthetase, naphthoate synthase, dihydroxynaphthoic acid synthetase, menA: 1,4-dihydroxy-2-naphthoate octaprenyltransferase and MenG_heptapren: demethylmenaquinone methyltransferase with 20 occurrences in the genome assembly. As previously reported, *Oceanobacillus longus* sp. nov. exhibited brown pigmentation (Amoozegar et al. 2016a). Chromogenicity, as observed in *Bacillaceae* (Drewnowska et al. 2015) and other microorganisms (Rao et al. 2017), suggests several possibilities of bioactivities, for example, UV protection, stress tolerance, and antimicrobial action.

Second is the cellulase activity of *Oceanobacillus* sp. SE10311 (Figure 1B). An audit of the glycoside hydrolases in the *Oceanobacillus* sp. SE10311 annotated genome confirmed their presence in multiple copies. Found in the annotation is 6-phospho-β-glucosidase, a gateway cellulolytic gene classified from Carbohydrate-Active enZYmes (CAZy) database. Though other primary cellulolytic genes like β-1,4-endoglucanase, and cellobiohydrolases were absent, the presence of cryptic 6-phospho-β-glucosidase, 6-phospho-β-galactosidase, etc. may indicate the existence of an alternative cellulose degradation pathway. Other minor genes with possible affinity to cellulase activity include cyclic β-1,2-glucan ABC transporter, putative bifunctional 4-α-glucanotransferase/malto-oligosyltrehalose synthase. Generally, the genomic and phenotypic evidence on *Oceanobacillus* sp. SE10311 cellulase activity concurs with observations that the *Bacillaceae* have inherent cellulolytic properties (Fatani et al. 2021).

A)



B)

Figure 1: *Oceanobacillus* sp. SE10311 cultures showing **A)** brown pigmentation formation and **B)** cellulase activity.

Figure 2: The whole genome sequence phylogenetic tree of *Oceanobacillus* sp. SE10311 generated by the KBase Species Tree Builder (Arkin et al. 2018).

A)

B)

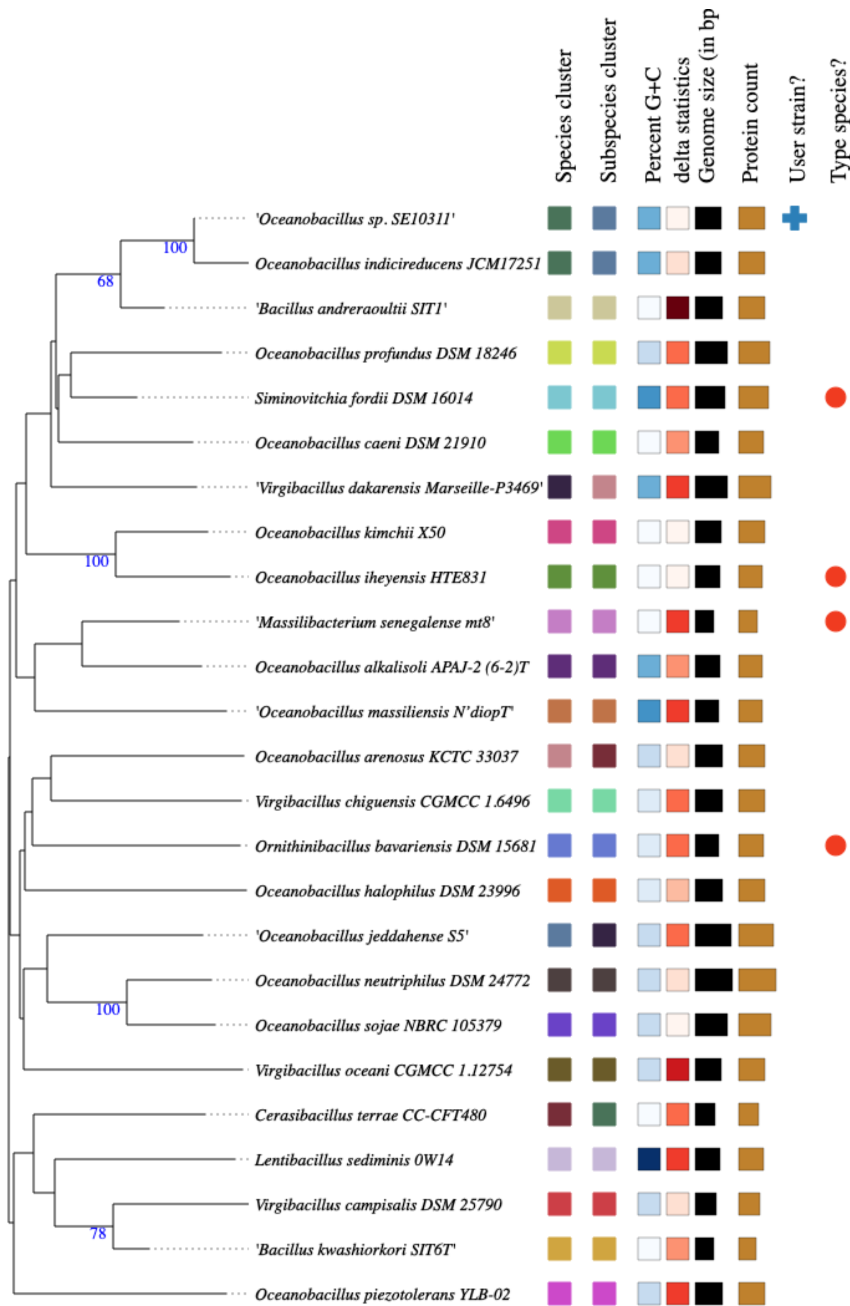
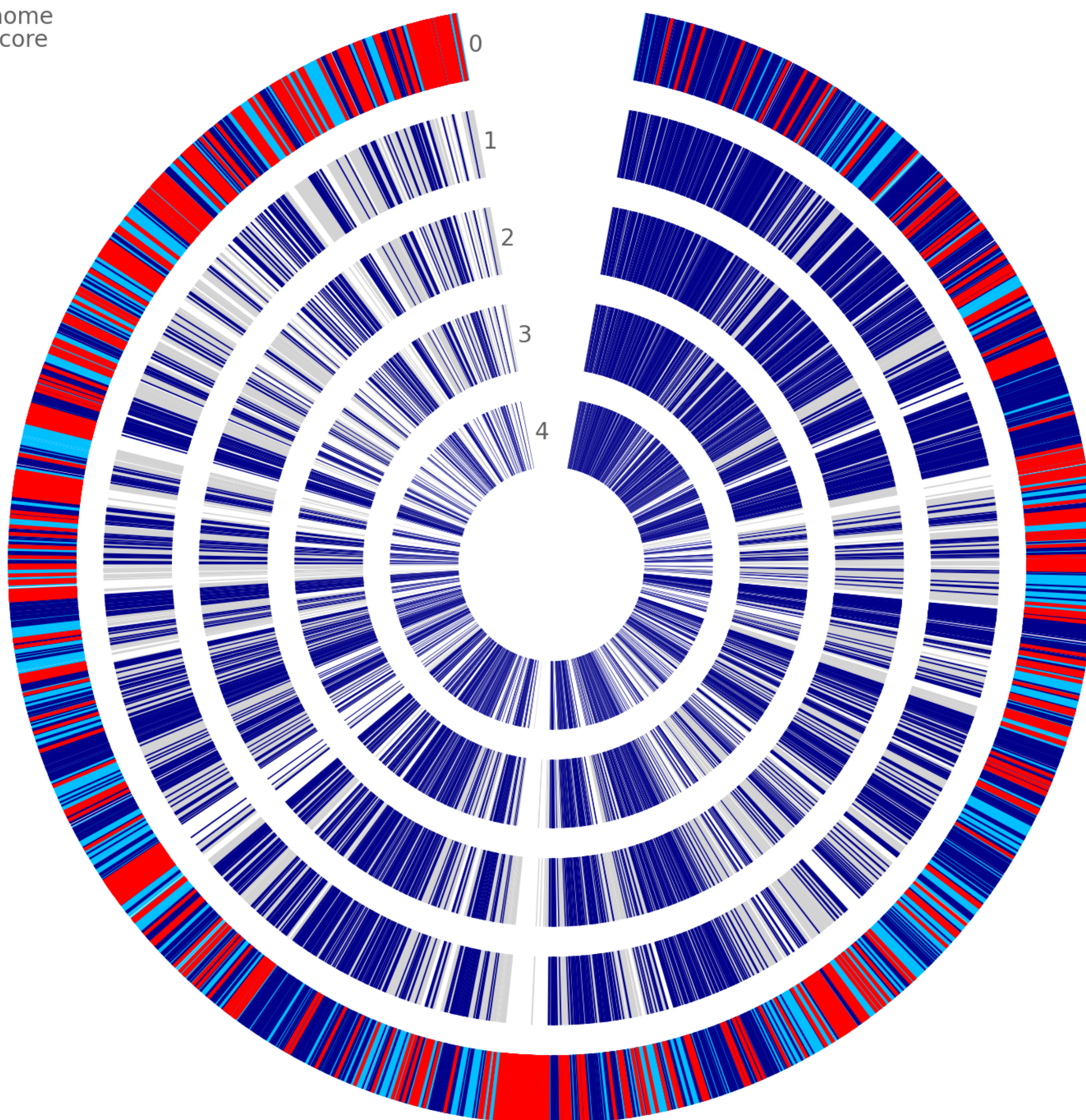


Figure 3: Type (Strain) Genome Server (TYGS) generated GBDP Trees. A) 16S rDNA gene sequence-based tree inferred with FastME 2.1.6.1 (Lefort et al. 2015) from GBDP distances calculated from 16S rDNA gene sequences. The branch lengths are scaled in terms of the GBDP distance formula d_5. The numbers above branches are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 59.3 %. The tree was rooted at the midpoint (Farris 1972); **B)** whole-genome sequence-based tree inferred with FastME 2.1.6.1 (Lefort et al. 2015) from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of the GBDP distance formula d_5. The numbers above branches are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 33.4 %. The tree was rooted at the midpoint (Farris 1972).



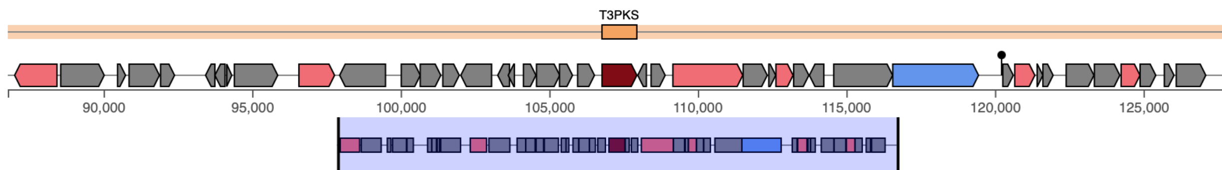
Note: **0** – *Oceanobacillus* sp. SE10311; **1** – *Oceanobacillus* sp. AG; **2** – *Oceanobacillus indicireducens* JCM 17251; **3** – *Oceanobacillus alkalisoli* APA-J2 6-2; **5** – *Oceanobacillus iheyensis* HTE831

Figure 4: KBase (Arkin et al. 2018) pangenome plot analysis of *Oceanobacillus* sp. SE10311 with closely associated *Oceanobacillus* strains. The outgroup is *Oceanobacillus iheyensis* HTE831.

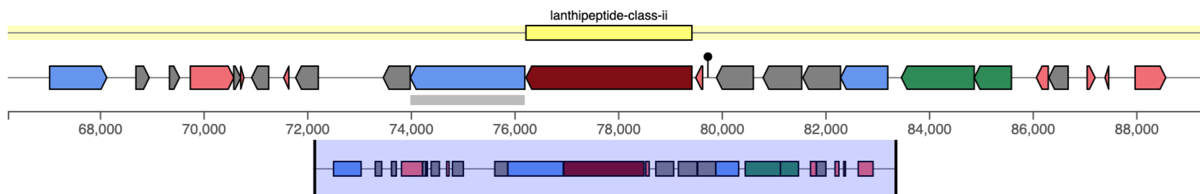
The AntiSMASH (Seemann 2014) secondary metabolite detection algorithm predicted two putative type-III polyketide synthase sites and a lanthipeptide class-II, mersacidinbiosynthetic type, and linear azol(in)e-containing peptides (LAP). Figure 5 illustrates the gene flanking regions of the antimicrobial genes present in *Oceanobacillus* sp. SE10311. These annotation features were accounted as validated polyketide β-ketoacyl: acyl carrier protein synthase and polyketide biosynthesis enol-CoA hydratase distributed in 19 locations of the genome assembly. In addition, the annotation identified lanti_2_LanM: type 2 lantibiotic biosynthesis protein LanM, and a putative lantibiotic ABC transporter, permease protein both located at contig 5, suggesting peptide lantibiotic production capability of *Oceanobacillus* sp. SE10311. These lantibiotics may include the synthesis of mesolanthionine, 3-methyllanthionine, dehydroalanine, and dehydrobutyrine, which are identifiable in some Gram-positive bacteria, including *Bacillaceae*. Also, antibiot_sagB: SagB-type dehydrogenase domains exist in contigs 5 and 39 of the *Oceanobacillus* sp. SE10311 genome assembly. We hypothesize that these genomic features can work synergistically to produce antimicrobial actions against other bacteria.

This result concurs with previous reports highlighting the broad-spectrum antimicrobial properties observed in *Oceanobacillus* species. As an example, *Oceanobacillus oncorhynchi* subsp. *incaldanensis* (J18TS1) was identified with four putative macrolide resistance genes and exhibited >256 MIC to LS and LCM (Okamoto et al., 2021). Another instance is *Oceanobacillus iheyensis,* the source of a highly proficient class A β-lactamase OIH-1 (Toth et al. 2010, Toth et al. 2009).

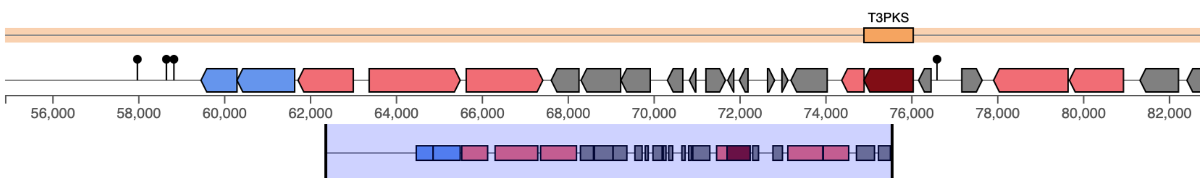
Location: 54,901 - 83,027 nt. (total: 28,127 nt); type-III polyketide synthase (T3PK)



Location: 66,225 - 89,428 nt. (total: 23,204 nt); lanthipeptide class-II



Location: 54,901 - 83,027 nt. (total: 28,127 nt); mersacidin biosynthetic, type-III polyketide synthase (T3PK)



Location: 1 - 20,894 nt. (total: 20,894 nt); linear azol(in)e-containing peptides (LAP)

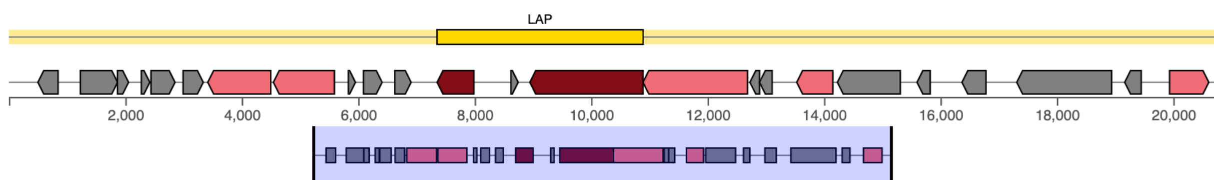


Figure 5: antiSMASH (Blin et al. 2023) Prediction of Secondary Metabolite Genes Putatively Conferring Antimicrobial Activity of *Oceanobacillus* sp. SE10311.

The *Oceanobacillus* species collectively features diverse genes involved in detoxification and stress protection against heavy metals like cobalt, zinc, cadmium, and chromium (Sami et al. 2021, Diba et al. 2021, Menasria et al. 2019). Specifically, *Oceanobacillus* sp. SE10311 carries two types of metallo-regulator genes targeting arsenic and cadmium. The identified genes include the arsenical resistance operon repressor, arsenic metallochaperone ArsD, which transfers trivalent metalloids to the ArsAB pump, arsenite/antimonite pump-driving ATPase ArsA (EC 3.6.3.16), and the cadmium efflux system accessory protein. Hypothetically, *Oceanobacillus* sp. SE10311 is more efficient in detoxifying arsenate, arsenite, and antimonite (Table 4).

As earlier indicated, *Oceanobacillus* species harbor genes for antibiotic resistance. Specifically, *Oceanobacillus* sp. SE10311 has a singleton gene, vancomycin (or other glycopeptides) histidine kinase VanS => VanF/M-type for vancomycin resistance. VanS is a Class-I HK that is membrane-bound and homodimeric and consists of a periplasmic domain, a transmembrane (TM) domain with two transmembrane helices, a linker region/HAMP domain, a dimerization and histidine phospho-acceptor (DHp) domain, and a catalytic ATP-binding (CA) domain (Guffey and Loll 2021). These domains contribute to the HK's signal detection, signal transduction, and catalytic activity. The VanS protein crystallographic structures are currently unavailable.

CONCLUSION

In summary, this study reported the whole genome sequence draft of *Oceanobacillus* sp. SE10311, subsequently identified as a strain of *Oceanobacillus indicireducens*. Phenotyping, genome annotation, and comparative analyses identified antimicrobial genes, resistance genes to β-lactamase and vancomycin, and arsenic and cadmium responsive regulators in the *Oceanobacillus* sp. SE10311. Notably, this study represents the first attempt at comparative genomics between strains of *O. indicireducens*.

ACKNOWLEDGMENT

The lead author extends his gratitude to the Department of Science and Technology of the Philippines and the National Research Council of the Philippines for partially funding this study.

CONFLICT OF INTEREST

Authors declare no conflict of interests.

CONTRIBUTIONS OF INDIVIDUAL AUTHORS

NHTG* – Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Reviewing and Editing
CPR – Investigation
ECM – Writing – Review & Editing
EPA – Writing – Review & Editing
MGQD – Writing – Review & Editing
LCV – Writing – Review & Editing
EDLRA – Investigation, Supervision, Writing – Review & Editing
ARC – Investigation, Supervision, Writing – Review & Editing
NHTG – Data Curation, Formal Analysis, Software, Validation, Reviewing and Editing
RGM – Project
Administration, Resources, Supervision, Writing – Review & Editing
RBO – Conceptualization, Methodology, Supervision, Reviewing and Editing

*Corresponding author

REFERENCES

Akuzawa S, Nagaoka J, Kanekatsu M, Kaanesake Y, Suszuki T. 2016. Draft genome sequence of *Oceanobacillus picturae* Heshi-B3, isolated from fermented rice bran in a traditional Japanese seafood dish. Genome Announc 2010; 4.

Albdaiwi RN, Khyami-Horani H, Ayad JY, Alananbeh KM, Al-Sayaydeh R. Isolation and characterization of halotolerant plant growth promoting rhizobacteria from durum wheat (*Triticum turgidum* subsp. *durum*) cultivated in saline areas of the Dead Sea region. Front Microbiol 2019; 10:1639.

Alhindi T, Albdaiwi R. Draft genome sequence of *Oceanobacillus jordanicus* strain GSFE11, a halotolerant plant growth-promoting bacterial endophyte isolated from the Jordan Valley. Evol Bioinform Online 2022; 18:11769343211071114.

Amoozegar MA, Bagheri M, Makhdoumi A, Mehrshad M, Didari M, Schumann P, Spröer C, Sánchez-Porro, Ventosa, A. *Oceanobacillus longus* sp. nov., a moderately halophilic bacterium isolated from a salt lake. Int J Syst Evol Microbiol 2016a; 66:4225-4230.

Amoozegar MA, Bagheri M, Makhdoumi A, Nikou MM, Fazeli SAS, Schumann P, Spröer C, Sánchez-Porro, Ventosa, A. *Oceanobacillus halophilus* sp. nov., a novel moderately halophilic bacterium from a hypersaline lake. Int J Syst Evol Microbiol 2016b; 66:1317-1322.

Amoozegar MA, Bagheri M, Makhdoumi-Kakhki A, Didari M, Schumann P, Spröer C, Sánchez-Porro, Ventosa, A. *Oceanobacillus limi* sp. nov., a moderately halophilic bacterium from a salt lake. Int J Syst Evol Microbiol 2014; 64:1284-1289.

Andrews S. FastQC: A quality control tool for high throughput sequence data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ 2010.

Arkin AP, Cottingham R, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia JM, Chia JM, Colasanti R,Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, W. Gerlach W, Gerstein M, A. Greiner A, J. Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 2018; 36:566-569.

Bankevich A, Nurk AS, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N,Tesler G,Alekseyev MA, Pevzner PA. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012; 19:455-477.

Blin K, Shaw S, Augustinj HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, van Wezel GP, Medema MH, Weber T. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Research 2023. 51:W46-W50.

Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Oxford University Press. 2022.

Chen F, Mackey AJ, Stoeckert Jr CJ, Roos DS. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Research 2006:34, D363-D368.

Cummins EA, Hall RJ, Mcinerney JO, Mcnally A. Prokaryote pangenomes are dynamic entities. Current Opinion in Microbiology 2022:66, 73-78.

Diba H, Cohan RA, Salimian M, Mirjani R, Soleimani M, Khodabakhsh F. Isolation and characterization of halophilic bacteria with the ability of heavy metal bioremediation and nanoparticle synthesis from Khara salt lake in Iran. Arch Microbiol 2021; 203:3893-3903.

Didari M, Bagheri M, M. A. Amoozegar MA, S. Bouzari S, Babavalian H,Tebyanian H, Hassanshahian M, and Ventosa A. Diversity of halophilic and halotolerant bacteria in the largest seasonal hypersaline lake (Aran-Bidgol-Iran). J Environ Health Sci Eng 2020; 18:961-971.

Farris JS. Estimating phylogenetic trees from distance matrices. The American Naturalist 1972;106:645-668.

Fatani S, Saito Y, Alarawi M, Gojobori T, Mineta K. Genome sequencing and identification of cellulase genes in *Bacillus paralicheniformis* strains from the Red Sea. BMC Microbiol 2021; 21:254.

Gan L, Zhang H, Long X, Tian J, Z. Wang Z, Y. Zhang Y, Y. Dai Y, Tian Y. *Ornithinibacillus salinisoli* sp. nov., a

moderately halophilic bacterium isolated from a saline-alkali soil. Int J Syst Evol Microbiol 2018; 68:769-775.

Gatpatan M, Amoranto MB, Ballesteros AJ, Sabino N, Zarate J, M. Bautista M, L. Villegas L. Diversity of bacteria in ants (Hymenoptera: Formicidae) from canopy and understory of selected trees at Mount Makiling Forest Reserve, Laguna, Philippines. Philippine Journal of Science 2021; 150:755-765.

Gonzalez JCT, de Guia APO, Dimalibot JC, Pantua KV, Gustilo WO, Bantayan NC. Understorey to canopy vertebrate fauna of a lowland evergreen forest in Mt. Makiling Forest Reserve, Philippines. Biodiversity Data Journal 2020: 8.

Guffey AA and Loll PJ. Regulation of resistance in vancomycin-resistant Enterococci: The VanRS two-component system. Microorganisms 2021: 9.

Gurevich A, Saveliev V, Vyahhi, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics 2013; 29**:**1072-5.

Hirota K, Aino K, NodasakaY, Yumoto I. *Oceanobacillus indicireducens* sp. nov., a facultative alkaliphile that reduces an indigo dye. Int J Syst Evol Microbiol 2013a; 63:1437-1442.

Hirota K, Hanaoka Y, Nodasaka Y, Yumoto I. *Oceanobacillus polygoni* sp. nov., a facultatively alkaliphile isolated from indigo fermentation fluid. Int J Syst Evol Microbiol 2013b; 63:3307-3312.

Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C, Bork P. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. Mol Biol Evol 2017:34**,** 2115-2122.

Hyun DW, Whon TW, Kim JY, Kim PS, Shin NR, Kim MS, Bae JW. Genomic analysis of the moderately haloalkaliphilic bacterium *Oceanobacillus kimchii* strain X50T with improved high-quality draft genome sequences. J Microbiol Biotechnol 2015; 25:1971-1976

Jain C, Rodriguez RLM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nature Communications 2018:9**,** 5114.

Jin PJ, Sun L, Rao MPN, Jiao JY, Sun XJ, Li WJ, Wang S. *Oceanobacillus alkalisoli* sp. nov., an alkaliphilic bacterium isolated from saline-alkaline soil. Int J Syst Evol Microbiol 2022; 72.

Khelaifia S, Lagier JC, Bibi F, Azhar EI, Croce O, Padmanabhan R, Jiman-Fatani AA, Yasir M, Robert C, Andrieu C, Fournier PE, Raoult D. Microbial culturomics to map halophilic bacterium in human gut: genome sequence and description of *Oceanobacillus jeddahense* sp. nov. Omics 2016; 20:248-258.

Krueger F. Trim Galore!: A wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data. Babraham Institute. 2015.

Kumar Mondal A, Kumar J, Pandey R, Gupta S, Kumar M, Bansal G, Mukerji M, Dash D, Singh Chauhan N. Comparative genomics of host-symbiont and free-living Oceanobacillus species. Genome Biol Evol 2017; 9:1175-1182.

Lantican NB, Diaz MGQ, Cantera JJL, de los Reyes FL, Raymundo AK. Microbial community of a volcanic mudspring in the Philippines as revealed by 16S rDNA sequence analysis and fluorescence in situ hybridization. World Journal of Microbiology and Biotechnology 2011; 27:859-867.

Lefort V, Desper R, Gascuel O. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. Mol Biol Evol 2015; 32:2798-2800.

Lu J, Nogi Y, Takami H. *Oceanobacillus iheyensis* gen. nov., sp. nov., a deep-sea extremely halotolerant and alkaliphilic species isolated from a depth of 1050 m on the Iheya Ridge. FEMS Microbiol Lett 2001; 205:291-297.

Magcale-Macandog DB, Lapitan FR, Laruya JM, Valerio JIF, Aguila JCD, Meşină C, Santos TMF, Cuevas ANT, Baylon KD, Silapan IM, Dimalibot R, Edrial JD, Larida NJA, Natuel FA, Perez MGLD, Quiñones SGL. Updating of Makiling Biodiversity Information System (MakiBIS) and analysis of biodiversity data. Biodiversitas Journal of Biological Diversity. 2010.

Manni M, Berkeley MR, Seppey M, Zdobnov EM. BUSCO: assessing genomic data quality and beyond. Curr Protoc, 2021; 1**:**e323.

Mayr R, Busse HJ, Worliczek HL, Ehling-Schulz M, Scherer S. *Ornithinibacillus* gen. nov., with the species *Ornithinibacillus bavariensis* sp. nov. and *Ornithinibacillus californiensis* sp. nov. Int J Syst Evol Microbiol 2006; 56:1383-1389.

Meier-Kolthoff JP and Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nature Communications 2019: 10**,** 2182.

Menasria, T., M. Monteoliva-Sánchez, L. Benammar, M. Benhadj, A. Ayachi, H. Hacène, A. Gonzalez-Paredes, and M. Aguilera. Culturable halophilic bacteria inhabiting Algerian saline ecosystems: A source of promising features and potentialities. World J Microbiol Biotechnol 2019; 35:132.

Montecillo A, Ilag L, Kohli A, Lantican N, and Raymundo A. Archaeal community profiling of a mixed culture from mud and water slurry from a solfataric mudspring as revealed by a community proteome approach. 2019.

Mwandira W, Nakashima K, Kawasaki S, Arabelo A, Banda K, Nyambe I, Chirwa M, Ito M, Sato T, Igarashi T, Nakata H, Nakayama S, Ishizuka M. Biosorption of Pb (II) and Zn (II) from aqueous solution by *Oceanobacillus profundus* isolated from an abandoned mine. Sci Rep 2020; 10:21189.

Mwandira W, Nakashima K, Kawasaki S, Ito M, Sato T, Igarashi T, Chirwa M, Banda K, Nyambe I, Nakayama S, Nakata H, Ishizuka M. Solidification of sand by Pb(II)-tolerant bacteria for capping mine waste to control metallic dust: Case of the abandoned Kabwe Mine, Zambia. Chemosphere 2019; 228:17-25.

Namwong S, Tanasupawat S, Lee KC, Lee JS. *Oceanobacillus kapialis* sp. nov., from fermented shrimp paste in Thailand. Int J Syst Evol Microbiol 2009; 59:2254-2259.

Okamoto M, Kumagai M, Kanamori H, Takamatsu D. Antimicrobial resistance genes in bacteria isolated from Japanese honey, and their potential for conferring macrolide and lincosamide resistance in the American foulbrood

pathogen Paenibacillus larvae. Front Microbiol 2021; 12:667096.

OuYang YT, Li MM, Lv AP, Jin PJ, Liu L, Liu ZT, Jiao JY, Li WJ, Cao LX, Wang S. *Oceanobacillus saliphilus* sp. nov., isolated from saline-alkali soil in Heilongjiang Province, China. Curr Microbiol 2022; 79:301.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 2015; 25:1043-55.

Rohban R, Amoozegar MA, Ventosa A. Screening and isolation of halophilic bacteria producing extracellular hydrolyses from Howz Soltan Lake, Iran. J Ind Microbiol Biotechnol 2009; 36:333-340.

Sami A, Elimairi I, Patangia D, Watkins C, Ryan CA, Ross RP, Stanton C. The ultra-structural, metabolomic, and metagenomic characterisation of the sudanese smokeless tobacco 'Toombak'. Toxicol Rep 2012; 8:1498-1512.

Seghal Kiran G, Nishanth LA, Kennedy J, Dobson AD, Selvin J. A halotolerant thermostable lipase from the marine bacterium *Oceanobacillus* sp. PUMB02 with an ability to disrupt bacterial biofilms. Bioengineered 2014; 5:305-318.

Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014; 30: 2068-9.

Shimoyama Y. MGCplotter: Microbial Genome Circular plotting tool. 2022.

Solis MJL, Dela Cruz TE, Schnittler M, Unterseher M. The diverse community of leaf-inhabiting fungal endophytes from Philippine natural forests reflects phylogenetic patterns of their host plant species *Ficus benjamina*, *F. elastica,* and *F. religiosa*. Mycoscience 2015; 57:96-106.

Subramanian M, Maruthamuthu M. Draft genome sequences of two *Bacillus* spp. and an *Oceanobacillus* sp. strain isolated from marine macroalgae. Microbiol Resour Announc 2019; 8.

Tanaka T, Kawasaki K, Daimon S, Kitagawa W, Yamamoto K, Tamaki H, Tanaka M, Nakatsu CH, Kamagata Y. A hidden pitfall in the preparation of agar media undermines microorganism cultivability. Appl Environ Microbiol 2014; 80:7659-7666.

Tominaga T, An SY, Oyaizu H, Yokota A. Oceanobacillus soja sp. nov. isolated from soy sauce production equipment in Japan. J Gen Appl Microbiol 2009; 55:225-232.

Toth M, Smith C, Frase H, Mobashery S, Vakulenko S. An antibiotic-resistance enzyme from a deep-sea bacterium. J Am Chem Soc 2010; 132:816-823.

Toth M, Vakulenko SB, Smith CA. Purification, crystallization, and preliminary x-ray analysis of the beta-lactamase Oih-1 from *Oceanobacillus iheyensis*. Acta Crystallogr Sect F Struct Biol Cryst Commun 2009; 65:582-585.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 2014; 9:e112963.

Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 2017; 1:e1005595.

Whon TW, Jung MJ, Roh SW, Nam YD, Park EJ, Shin KS, Bae JW. *Oceanobacillus kimchii* sp. nov. isolated from a traditional Korean fermented food. J Microbiol 2010; 48:862-866.

Yongchang O, Xiang W, Wang G. *Oceanobacillus bengalensis* sp. nov., a bacterium isolated from seawater of the Bay of Bengal. Antonie Van Leeuwenhoek 2015; 108:1189-1196.

Yu C, Yu S, Zhang Z, Li Z, Zhang XH. *Oceanobacillus pacificus* sp. nov., isolated from a deep-sea sediment. Int J Syst Evol Microbiol 2014. 64:1278-1283.

Yu L, Tang X, Wei S, Qiu Y, Xu X, G. Xu G, Wang Q, Yang Q. Two novel species of the family Bacillaceae: *Oceanobacillus piezotolerans* sp. nov. and *Bacillus piezotolerans* sp. nov., from deep-sea sediment samples of Yap Trench. Int J Syst Evol Microbiol 2019; 69:3022-3030.

Zhu WY, Yang L, Shi YJ, Mu CG, Wang Y, Kou YR, Yin M, Tang SK. *Oceanobacillus halotolerans* sp. nov., a bacterium isolated from salt lake in Xinjiang Province, north-west China. Arch Microbiol 2020; 202:1545-1549.

Zhu WY, Yang L, Zhang ZT, Mu CG, Wang Y, Kou YR, Jiang GQ, Yin M, Tang SK. *Oceanobacillus salinisoli* sp. nov., a bacterium isolated from saline soil of Turpan City in Xinjiang Province, north-west China. Arch Microbiol 2021; 203:2919-2924.

SUPPLEMENTAL FIGURES

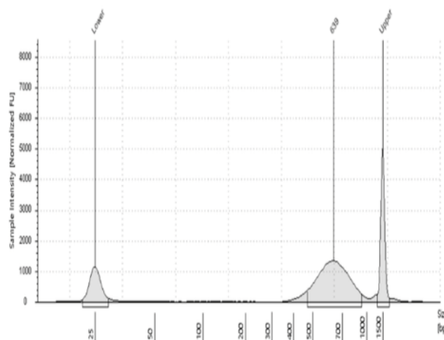
A)

Arrival Date	2022-11-17	Experiment Date	2022-12-02	Tested by	JYH
Comment	LightCycle qPCR				

#	Library Name	Library Type	Conc. (ng/ul)	Conc. (nM)	Size (bp)	Result*	
1	SE10311	TruSeq Nano DNA (350)	29.28	70.5	639	Pass	

Experiment Condition	TapeStation D1000 Screen Tape

Click to Enlarge =>1:Library : SE10311



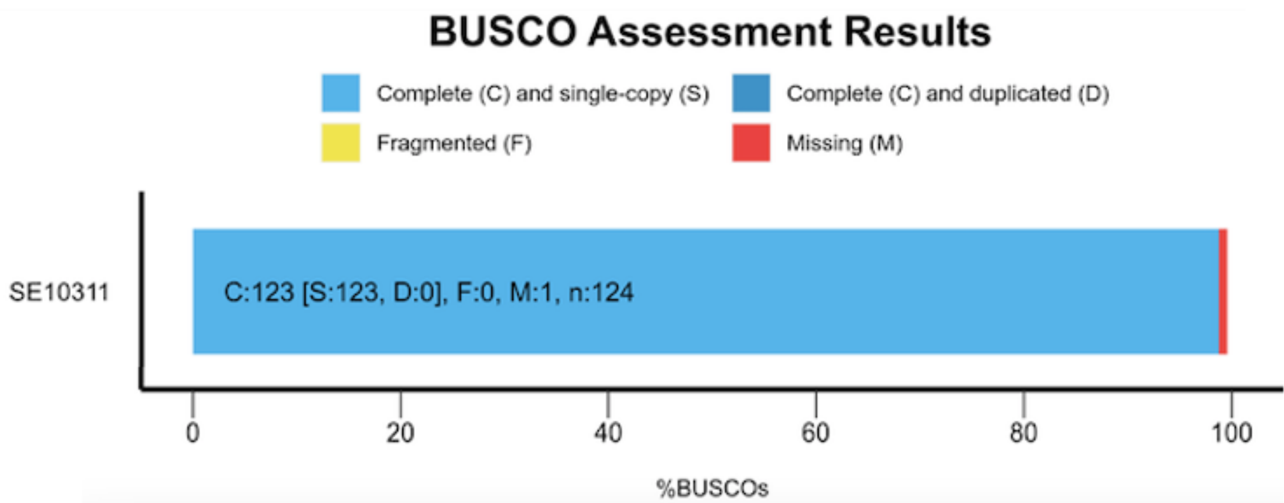
B)

Sample ID	Total bases(bp)	Total reads	GC(%)	AT(%)	Q20(%)	Q30(%)
SE10311	525,648,844	5,204,444	39.1	60.9	99.3	98.6

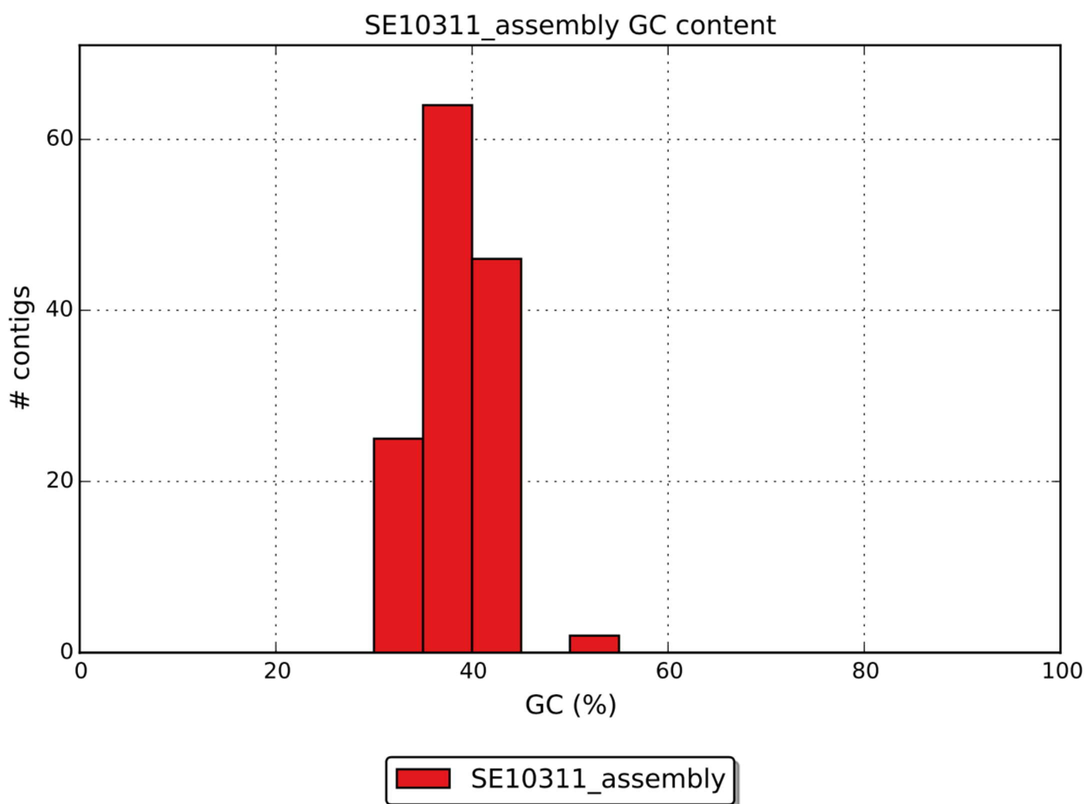
- Sample ID : Sample name.
- Total bases(bp) : Total number of bases sequenced.
- Total reads : Total number of reads. For illumina paired-end sequencing, this value refers to the sum of read1 and read2.
- GC(%) : Ratio of GC content.
- AT(%) : Ratio of AT content.
- Q20(%) : Ratio of bases that have phred quality score of over 20.
- Q30(%) : Ratio of bases that have phred quality score of over 30.

Supplemental Figure 1: Sample SE10311 Reports on **A)** DNA Library Quality Control and **B)** Raw Sequence Reads Data (Macrogen, South Korea)

A)

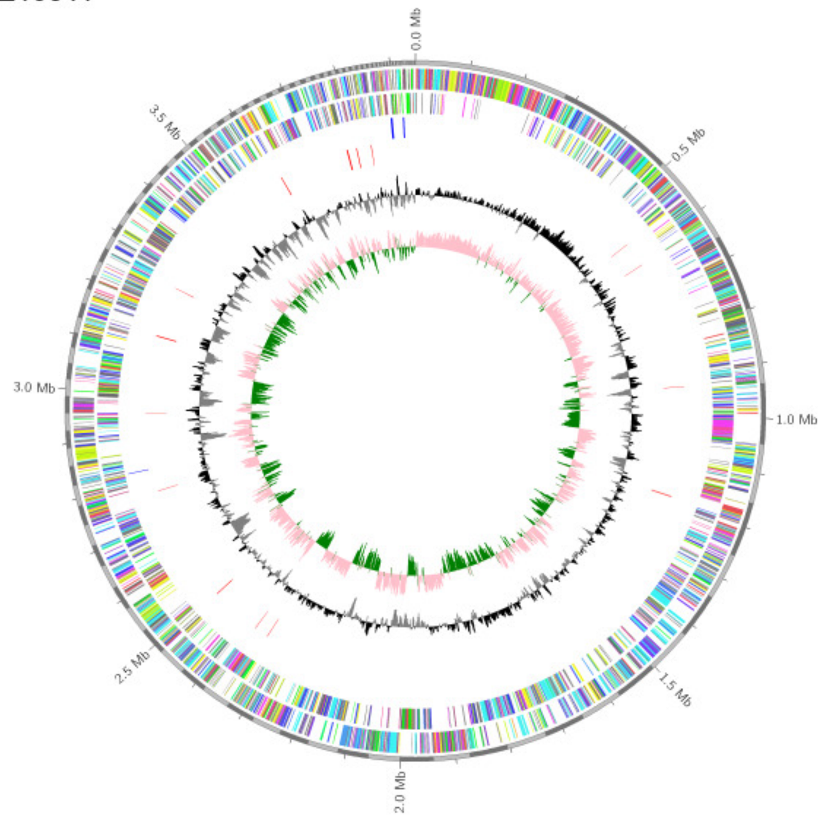


B)



Supplemental Figure 2: *Oceanobacillus* sp. SE10311 Genome Assessment Results, **A)** BUSCO Assessment Result of SE10311 and **B)** Assembly GC Content

SE10311



Track Contents
- Forward CDS
- Reverse CDS
- rRNA
- tRNA
- GC Content (+)
- GC Content (-)
- GC Skew (+)
- GC Skew (-)

- J : Translation, ribosomal structure and biogenesis
- A : RNA processing and modification
- K : Transcription
- L : Replication, recombination and repair
- B : Chromatin structure and dynamics
- D : Cell cycle control, cell division, chromosome partitioning
- Y : Nuclear structure
- V : Defense mechanisms
- T : Signal transduction mechanisms
- M : Cell wall/membrane/envelope biogenesis
- N : Cell motility
- Z : Cytoskeleton
- W : Extracellular structures
- U : Intracellular trafficking, secretion, and vesicular transport
- O : Posttranslational modification, protein turnover, chaperones
- X : Mobilome: prophages, transposons
- C : Energy production and conversion
- G : Carbohydrate transport and metabolism
- E : Amino acid transport and metabolism
- F : Nucleotide transport and metabolism
- H : Coenzyme transport and metabolism
- I : Lipid transport and metabolism
- P : Inorganic ion transport and metabolism
- Q : Secondary metabolites biosynthesis, transport and catabolism
- R : General function prediction only
- S : Function unknown
- \- : No COG classified

Supplemental Figure 3: Genome plot map of *Oceanobacillus* sp. SE10311.