



OPEN ACCESS

EDITED BY

Huiluo Cao,
The University of Hong Kong, Hong
Kong SAR, China

REVIEWED BY

Hongliang Liu,
Shandong University of Technology, China
Dong Yu,
Second Military Medical University, China

*CORRESPONDENCE

Ruibai Wang
✉ wangruibai@icdc.cn

RECEIVED 30 August 2024

ACCEPTED 18 November 2024

PUBLISHED 12 December 2024

CITATION

Yu Z-L and Wang R-B (2024) Revised
taxonomic classification of the
Stenotrophomonas genomes, providing new
insights into the genus *Stenotrophomonas*.
Front. Microbiol. 15:1488674.
doi: 10.3389/fmicb.2024.1488674

COPYRIGHT

© 2024 Yu and Wang. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Revised taxonomic classification of the *Stenotrophomonas* genomes, providing new insights into the genus *Stenotrophomonas*

Ze-Lin Yu and Rui-Bai Wang*

National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Disease, National
Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and
Prevention, Beijing, China

Background: *Stenotrophomonas* strains are important opportunistic pathogens with great potential applications in industry and agriculture. Their significant genetic and phenotypic diversity has led to several changes in their taxonomic localization and was prone to inaccurate species classification based on traditional identification methods.

Methods: All 2,615 genomes of the genus *Stenotrophomonas* were obtained from the NCBI genome database. Genomic methods, including average nucleotide identity (ANI), were used to evaluate the 31 defined species. After evaluating the ANI thresholds applicable to *Stenotrophomonas*, the species classification of all submitted genomes was revised.

Results: Compared to the reference genomes of each species, 41.17% of the submitted *Stenotrophomonas* genomes had ANI values below 95, and 8.58% of the genomes were even below 90%. Moreover, 45.3% (705/1555) of the *S. maltophilia* strains actually belonged to other species within the *S. maltophilia* complex (Smc), or even to distant relatives outside the Smc. Based on the ANI threshold values of 95 and 90% for species and complexes confirmed to be applicable to *Stenotrophomonas*, 2,213 submitted *Stenotrophomonas* genomes were re-divided into 116 ANI genome species.

Conclusion: The results confirmed that 16S rRNA gene sequencing has low discriminability for the closely related *Stenotrophomonas* species. The annotated species of a considerable strain were indeed incorrect, especially since many *S. maltophilia* strains did not belong to this representative pathogenic species of *Stenotrophomonas*. This makes it necessary to reconsider the evolutionary relationship, pathogenicity, and clinical significance of *Stenotrophomonas*.

KEYWORDS

Stenotrophomonas, average nucleotide identity, taxonomy, species classification, genetic diversity

1 Introduction

The genus *Stenotrophomonas* comes under the family *Lysobacteraceae* (*Xanthomonas* group), the order *Lysobacterales*, the class *Gammaproteobacteria*, and the phylum *Pseudomonadota* (Palleroni and Bradbury, 1993; Brooke, 2021). The members of the genus *Stenotrophomonas* are ubiquitous in diverse habitats, with soil and plants serving as their main environmental reservoirs (Ryan et al., 2009). This microorganism is associated with the plant rhizosphere, has an important role in the elemental cycling of sulfur and nitrogen, and engages in beneficial interactions with plants to promote their growth and health (An and Berg, 2018; Pérez-Pérez et al., 2020; Kumar et al., 2023). In addition, these bacteria can degrade complex compounds and pollutants and thus have a potential role in environmental remediation (Venkidesamy and Megharaj, 2016; Xiong et al., 2020). However, a few members of this genus are pathogens that can cause various infections and certainly impact human health worldwide. The representative species is *Stenotrophomonas maltophilia*, the first member of this genus and also a globally emerging multidrug-resistant and opportunistic pathogen that has attracted much attention (Brooke, 2012; Chang et al., 2015). According to global clinical data, the attributed mortality rate of bacteremia caused by this bacterium is as high as 65% (Hua et al., 2013).

Owing to its well-known ecological and phenotypic diversity and intraspecific heterogeneity (Palleroni and Bradbury, 1993), the taxonomic status of the genus *Stenotrophomonas* has undergone several revisions (Tamma et al., 2022) and was once referred to as *Pseudomonas* and *Xanthomonas* (Swings et al., 1983), as well as the composition of the so-called *S. maltophilia* group/complex (Smc) (Ochoa-Sánchez and Vinuesa, 2017; Vinuesa et al., 2018). To date, this genus has different taxonomic structures in different databases. In the National Center for Biotechnology Information (NCBI) taxonomy (<https://www.ncbi.nlm.nih.gov/datasets/taxonomy/>), the genus *Stenotrophomonas* currently comprises 31 classified species, 3,044 unclassified species, and 65 unclassified species from environmental samples. Smc is composed of three classified species, *S. maltophilia*, *S. sepiolia*, and *Pseudomonas hibiscicola*, and 24 unclassified species. While present in Jean Euzéby's List of Prokaryotic Names with Standing in Nomenclature (LPSN, <https://www.bacterio.net/genus/stenotrophomonas>), only 31 classified species are included as child taxa in the genus *Stenotrophomonas* (23 validly published under the ICNP and 21 with correct names), of which 28 species are consistent with NCBI. In the genome taxonomy database structured based on genome phylogeny (Parks et al., 2020) (GTDB, <https://gtdb.ecogenomic.org/>), 1,302 genomes of the genus *Stenotrophomonas* are divided into 112 species, of which 45 species have not been assigned names and many classified species are subdivided. For example, *S. maltophilia* is divided into 34 *S. maltophilia* species (A, AA, G, K, etc.) and *S. rhizophila* is divided into 4 *S. rhizophila* species; moreover, *S. panacihumi* and 3 *S. sp.* are removed from the genus *Stenotrophomonas* and placed under a new genus, *Stenotrophomonas_A* (https://gtdb.ecogenomic.org/tree?r=g__Stenotrophomonas_A).

Taxonomic methods include phenotypic, physiological, and genotypic characteristics, and the way microbiologists describe and classify species has undergone a major revision in the light of genomics (Konstantinidis and Tiedje, 2007; Sentausa and Fournier, 2013; Vandamme and Peeters, 2014). Among these methods, 16S rRNA sequencing is one of the most universal single-gene tools for differentiating bacteria, with 98.65% similarity generally set as the threshold for species differentiation (Ludwig and Klenk, 2001; Kim et al., 2014; Rosselló-Móra and Amann, 2015). However, this threshold does not apply to

Stenotrophomonas. Gröschel MI et al. defined *S. maltophilia* strains in Smc with 16S rRNA gene sequence similarities > 99.0% (Svensson-Stadler et al., 2012; Gröschel et al., 2020), and our study also showed that the similarities between *S. pigmentata* and *S. terrae* (99.24%) as along with the *S. humi* strain (98.96%) were greater than 98.65% (Li et al., 2024). Multilocus sequence typing (MLST) was used as a standard typing tool, enabling the investigation of evolutionary relationships among species classified within the same genus (Rizek et al., 2018). However, only *S. maltophilia* has a well-established combined MLST scheme available in the PubMLST database (Kaiser et al., 2009); other species lack a standardized definition. DNA–DNA hybridization (DDH) has been used as the gold standard for circumscribing prokaryotic species at the genomic level (Goris et al., 2007; Richter and Rosselló-Móra, 2009). Due to its complexity and time-consuming nature, DDH has been gradually replaced by digital DDH (dDDH) and average nucleotide identity (ANI) (Jain et al., 2018; Meier-Kolthoff and Göker, 2019). The recommended cutoff point of 70% DDH for species delineation corresponded to 95% ANI (Gosselin et al., 2022). In addition, a study suggested that 98% ANI is the accurate threshold for distinguishing subspecies (Pearce et al., 2021), although it is not recognized as a standard cutoff value in the taxonomy of different prokaryotes.

The number of *Stenotrophomonas* species has increased rapidly in the past 10 years, and an increasing number of *Stenotrophomonas* genomes have been deposited in the NCBI database (Li et al., 2024). The correct definitions of the genus *Stenotrophomonas*, its species, and the *S. maltophilia* complex and the correct species annotation of the submitted genomes are necessary to obtain accurate bioinformatics analysis results, which are crucial for the identification of new species and pathological epidemiology research and are the basis for a correct understanding of this genus. However, the submitted genomes in different periods were species-annotated based on different levels of strain identification methods. Therefore, we compared the discrimination and operability of several taxonomic methods for *Stenotrophomonas* and revised the species classification of all submitted genomes in this study.

2 Materials and methods

2.1 *Stenotrophomonas* genome retrieval and processing

All 2,615 genomes of the genus *Stenotrophomonas* were obtained from the NCBI genome database (<https://www.ncbi.nlm.nih.gov/datasets/genome/>) as of 1 April 2024. Among them, 402 genomes noted as atypical assemblies (including 22 contaminated genomes, 74 too large and too small genomes, and 306 genomes from unverified sources) were excluded. A total of 2,213 genomes were ultimately included in this analysis. The genome size and GC content of the genomes were ascertained using Seqkit v2.8.2 (<https://github.com/shenwei356/seqkit>) (Shen et al., 2024).

2.2 Analysis of reference genomes based on 16S rRNA gene identity, dDDH, and ANI

Currently, there are 31 recognized species of the genus *Stenotrophomonas* in the NCBI Taxonomy database. Notably, *S. indologenes* and *S. detusculanense* do not have available genome sequences. Furthermore, neither *S. aracearum* nor *S. oahuensis* has a

reference genome assigned in the NCBI database; their only available genomes served as reference genomes. Two novel species of the genus *S. pigmentata* and *S. tuberculopleuritidis* (610A2 and 704A1, currently temporarily classified as *S. sp.*) identified in our previous research (Li et al., 2024) were also incorporated as reference genomes for the study. Consequently, a total of 31 reference genomes were included in this study (Supplementary Table S1).

The 16S rRNA gene sequences were extracted from the genomes using barrnap v0.9 (<https://github.com/tseemann/barrnap>). Pairwise similarities between 16S rRNA genes were assessed through alignment using BLAST+ v2.15.0 (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The ANI similarity matrix was constructed using skani v0.2.1 (<https://github.com/bluenote-1577/skani>) (Shaw and Yu, 2023) and FastANI v1.33 (<https://github.com/ParBLiSS/FastANI>) (Jain et al., 2018). dDDH values of the reference genomes were compared using the GGDC 3.0 (<https://ggdc.dsmz.de/ggdc.php>) (Meier-Kolthoff et al., 2022). The visualization of the results was conducted using Matplotlib v3.8.4 (<https://matplotlib.org/>) and Seaborn v0.11.0 (<https://seaborn.pydata.org/>) in Python 3.12.1.

2.3 Phylogenetic analysis based on core-genome alignments

For the phylogenetic analysis of all reference genomes, *Pseudoxanthomonas broegbernensis* DSM 12573 was selected as the outgroup. We used Prokka v1.14.6 (<https://github.com/tseemann/prokka>) (Seemann, 2014) to annotate the genomes. The output GFF-formatted files from Prokka were used as input files for Roary v3.13.0 (<https://github.com/sanger-pathogens/Roary>) (Page et al., 2015) to calculate the genus *Stenotrophomonas* core and pan-genomes with the minimum blastp percentage identity set at 90%. The phylogenetic tree was constructed using FastTree 2 (<https://github.com/PavelTorgashov/FastTree>) (Price et al., 2010) based on the multi-FASTA alignment of all of the core genes and visualized using Tree Visualization By One Table (tvBOT) (Xie et al., 2023).

2.4 Consistency between skani and FastANI

The ANI value matrix between the 31 reference genomes and the ANI of all the *Stenotrophomonas* genomes compared to the 31 reference genomes was generated using both FastANI and skani. For each species, genomes were categorized into three groups based on their ANI compared to the reference genome as follows: ANI \geq 95, 95% > ANI \geq 90%, and ANI < 90%. The difference in classification between the two ANI tools was analyzed using the McNemar–Bowker test, and the consistency was evaluated by the kappa coefficient. The difference between the ANI values calculated by the two tools after stratifying was evaluated using a paired *t*-test. All statistical analyses were conducted using R version 4.4.0 (<https://cran.r-project.org/bin/windows/base/old/4.4.0/>), with $p < 0.05$ indicating statistical significance.

2.5 ANI clustering and evaluation of ANI cutoff values

The ANI matrix for all genomes of the genus *Stenotrophomonas* was generated using skani or FastANI. A helper Python script

provided using skani was employed to visualize and cluster the lower-triangular similarity matrix that skani output. Building upon this script, the fcluster function from Scipy v1.13.1 (<https://docs.scipy.org/doc/scipy/>) was integrated to actually obtain clusters by setting an ANI threshold.

We examined the distribution of pairwise ANI values separately between all genomes of the genus *Stenotrophomonas* ($n = 2,213$) and between all genomes of *S. maltophilia* ($n = 1,555$) using skani. In our evaluation, one genome was taken as the reference genome, and all genomes (including itself) were taken as query genomes. A total of n -squared pairwise comparisons were executed. The density curve of the pairwise ANI values was plotted using ggplot2 (<https://ggplot2.tidyverse.org/>) in R version 4.4.0.

The ANI cutoff value was evaluated through the analysis of clustering results and the density distribution. Based on the ANI cutoff values, genomes within the genus *Stenotrophomonas* were reclassified into the designated species.

3 Results

3.1 Composition and general characteristics of the genomes of the genus *Stenotrophomonas*

A total of 2,213 genomes were ultimately included in this analysis. In this study, the genomes of unclassified species were all included as *Stenotrophomonas sp.* (including uncultured organisms). In addition to *S. sp.*, the study encompasses 29 defined species composed of 1,841 genomes. The number of each species is shown in Figure 1A. *S. maltophilia* is the most abundant, with a total of 1,555 genomes, accounting for 70.27% of all genomes. The second most abundant species is *S. sp.*, with a total of 372 genomes. The remaining species all have less than 50 genomes, including 12 species with only one genome. Genomic characterization of different species was shown in Figures 1B,C. The size range of all *Stenotrophomonas* genomes is from 1.77 to 6.98 Mb, while that of *S. maltophilia* is 3.80 to 5.54 Mb. The average genome size of each species is similar, except for *S. koreensis*, whose genomes are significantly smaller than other species, ranging from 2.07 to 3.09 Mb. *S. sp.* contains various unclassified strains, so its relatively wide genome size range is expected. but the performance of *S. nitritireducens* is beyond expectation. *S. nitritireducens* only has 6 genomes, but there is one largest (GCA_037190265.1, 6.98Mb) and one smallest (GCA_017304365.1, 1.92Mb) genome of genus *Stenotrophomonas*, although the average size of the other 4 genomes (4.22 Mb) is similar to other species.

The GC content of *Stenotrophomonas* genomes ranges from 62.54 to 70.17%. Apart from *S. maltophilia*, *S. sp.*, and *S. nitritireducens*, the GC content distribution of each species is relatively concentrated (fluctuations do not exceed 2%). Although the GC content of over 95% of *S. maltophilia* genomes concentrates at 65.5–67.0%, it ranges from 63.6 to 68.17%. *S. sp.* strains are more dispersed, ranging from 62.54 to 69.87%. The GC content of *S. nitritireducens* genomes exhibited a bimodal distribution, with four genomes ranging from 68.08 to 68.62% and the other two genomes ranging from 63.83 to 63.98%. The GC content of *S. acidaminiphila* is higher than that of other species, ranging from 68.06 to 69.75%.

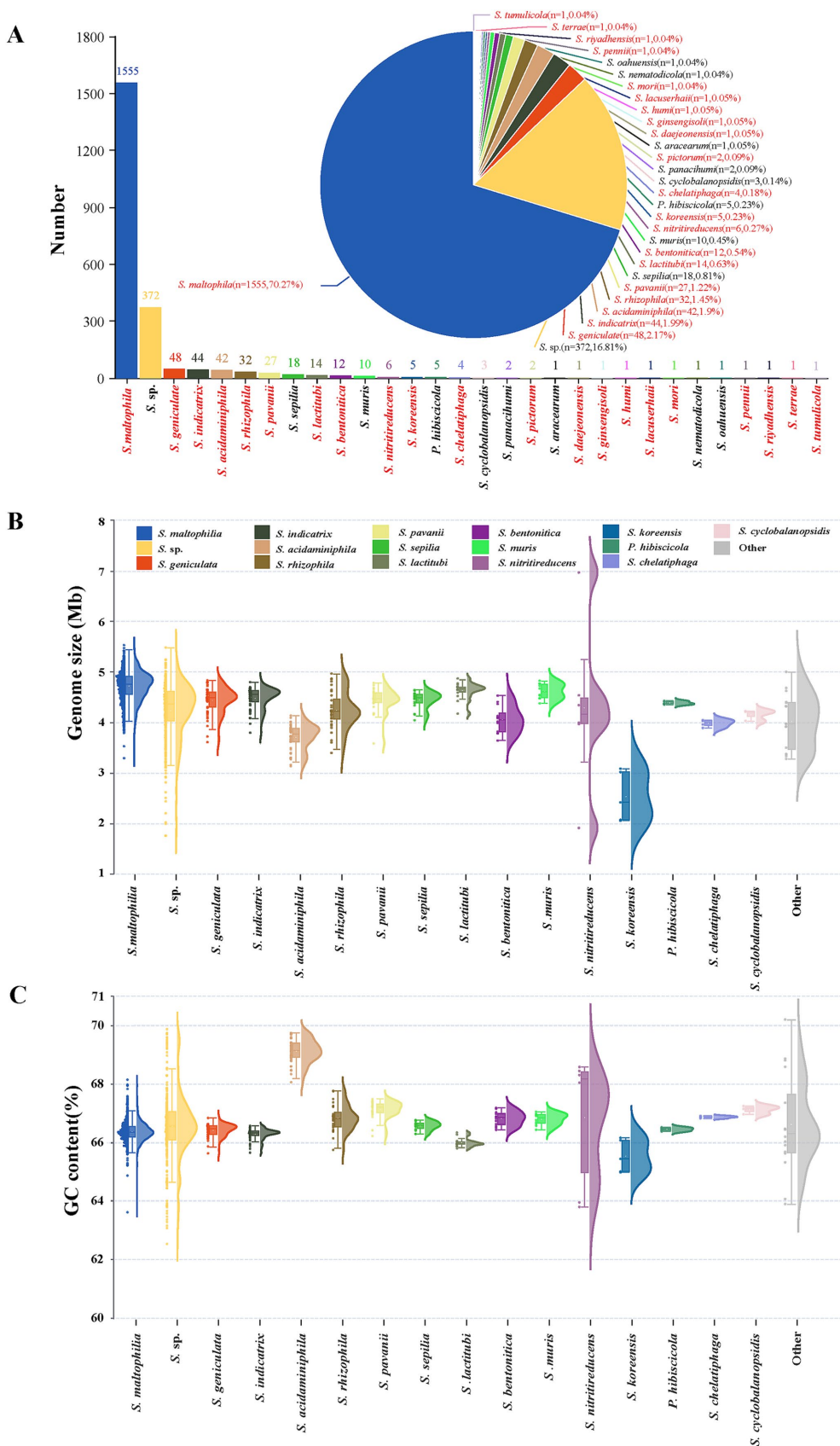


FIGURE 1 *Stenotrophomonas* species and general genomic characteristics of genomes analyzed in this study. (A) *Stenotrophomonas* species, arranged in descending order from left to right based on their submitted genome numbers in the NCBI genome database. In total, 21 species with validly published

(Continued)

FIGURE 1 (Continued)

and correct names are marked in red. (B,C) show the genome size and GC content of each *Stenotrophomonas* species using raincloud plots, respectively. The left beeswarm plots depict the distribution of individual data points with 'raindrops', the central box plots indicate the median, interquartile range, and potential outliers, and the right kernel density plots illustrate the probability density function of the continuous variable. Other refers to species with only one or two genomes, including *S. aracearum* ($n = 1$), *S. daejeonensis* ($n = 1$), *S. ginsengisoli* ($n = 1$), *S. humi* ($n = 1$), *S. lacuserhaii* ($n = 1$), *S. mori* ($n = 1$), *S. nematodocola* ($n = 1$), *S. panacihumi* ($n = 2$), *S. pennii* ($n = 1$), *S. pictorum* ($n = 2$), *S. riyadhensis* ($n = 1$), *S. terrae* ($n = 1$), and *S. tumulicola* ($n = 1$).

One reason for the large fluctuations in genome size and GC content is that 20.83% (461/2213) of the submitted genomes of *Stenotrophomonas* are metagenome-assembled genomes (Supplementary Table S2), and the majority of these outlier genomes belong to this type, limiting the accuracy of these genomes. The other speculated reason is that many isolates are from diverse sources, leading to genetic variability and diversity. There are nine strains that have been sequenced more than once and have genomes with different levels of completeness in the database. For example, *S. maltophilia* strain NCTC10498 has two complete genomes (GCA_011386925.1 and GCA_900475685.1) and one scaffold genome (GCA_014235265.1). The ANI values of genomes from different-levels of the same strain were similar to the reference genomes (Supplementary Table S6), indicating that the genome sequencing method and genomic assembly level do not affect the species classification based on the ANI.

3.2 Analysis of the reference genomes

The 16S rRNA gene sequences were extracted, and pairwise 16S rRNA gene similarities were calculated (Figure 2A). The 16S rRNA gene similarity between many species is $\geq 98.65\%$, and even greater than 99%. Obviously, the 16S rRNA gene sequence has low discriminability for *Stenotrophomonas* species.

As the gold standard for the delineation of bacterial species, the trends of pairwise ANI and dDDH were essentially consistent across the reference genomes analyzed (Figures 2B–D). The dDDH and ANI values between the defined species are all below the commonly used cutoff values for species determination. The only exceptions were *S. pennii* Sa5BUN4 and *S. lacuserhaii* K32. Their 16S rRNA gene sequence similarity, dDDH, and ANI values calculated using skani and FastANI between these two species were 100, 74.5, 97.32, and 97.22%, respectively, all of which exceeded the species cutoffs. Although *S. pennii* was first reported in 2021 (Gilroy et al., 2021), it was not validly published. *S. lacuserhaii* was reported in 2022 (Deng et al., 2022) and validly published under the International Code of Nomenclature of Prokaryotes (ICNP). *S. pennii* and *S. lacuserhaii* should be merged into one species.

In addition, we observed that the ANI values between *S. panacihumi* GSS15 and other genomes were low, with values calculated using skani being below 82% and values calculated using FastANI ranging between 81 and 84%. *S. panacihumi* is a non-validly described and poorly characterized species (Yi et al., 2010). It only has two assembled genomes. The study by Vinuesa et al. also demonstrated the inconsistency of *S. panacihumi* in the phylogeny of the genus *Stenotrophomonas* (Vinuesa et al., 2018). This could be the reason why the GTDB classified *S. panacihumi* into the new genus *Stenotrophomonas_A*. We also discovered the

anomaly of *S. koreensis* and *S. ginsengisoli* in ANI comparisons, although both were validly described species (Yang et al., 2006; Kim et al., 2010). In addition, the paired ANI between *S. koreensis* DSM 17805 and *S. ginsengisoli* DSM 24757 was 86.76% (skani) and 86.92% (FastANI). However, their ANI with the other 29 reference genomes is very low, with values below 82% (skani) or ranging between 81 and 82% (FastANI).

Furthermore, we compared the concordance between the ANI clustering heatmap (Figure 3A) and the phylogenetic tree constructed based on the core genes (Figure 3B). Considering that many pairwise ANI values between reference genomes are below 82% and output as "0" in skani as the default parameter setting (Shaw and Yu, 2023), we utilized FastANI to calculate accurate ANI values for clustering. The results showed that the classification of the four clades of the genus *Stenotrophomonas* based on these two methods and the composition of the species in each clade were similar. The relatively simple ANI calculation can replace the complex core gene analysis for identifying closely related species and classifying them on the phylogenetic tree. In the phylogenetic tree, we also observed inconsistencies between *S. panacihumi* GSS15 and other genomes of the genus *Stenotrophomonas*. While *S. koreensis* DSM 17805 and *S. ginsengisoli* DSM 24757 had lower ANI values than other genomes, they still belonged to the genus *Stenotrophomonas* on the phylogenetic tree.

3.3 Consistency between skani and FastANI

We compared the ANI values between all 1,841 genomes of 29 defined species and their reference genomes using two ANI tools to assess their consistency. The results indicated that when using 90 and 95 as thresholds, the ANI of each genome calculated by these two ANI tools was completely consistent in classification (Table 1 and Supplementary Table S3). The McNemar–Bowker test showed no significant difference (kappa coefficient = 1).

We conducted paired comparisons for ANI values across each range (Figure 4) and observed that for ANI $\geq 95\%$, ANI values calculated using skani were higher than those calculated using FastANI, exhibiting a notable statistical difference ($p < 0.05$). For ANI values between 90 and 95%, the ANI values computed by both tools were essentially consistent ($p > 0.05$). For ANI $< 90\%$, the ANI values calculated using skani were slightly lower than those using FastANI. However, the difference was not statistically significant ($p > 0.05$). Despite a few differences in the ANI values calculated by the two ANI tools, they demonstrated identical discriminatory capabilities at thresholds of 90% or 95%. Considering that skani's computations were several to tens of times faster than FastANI, we primarily utilized skani for large-scale paired ANI calculations in this study.

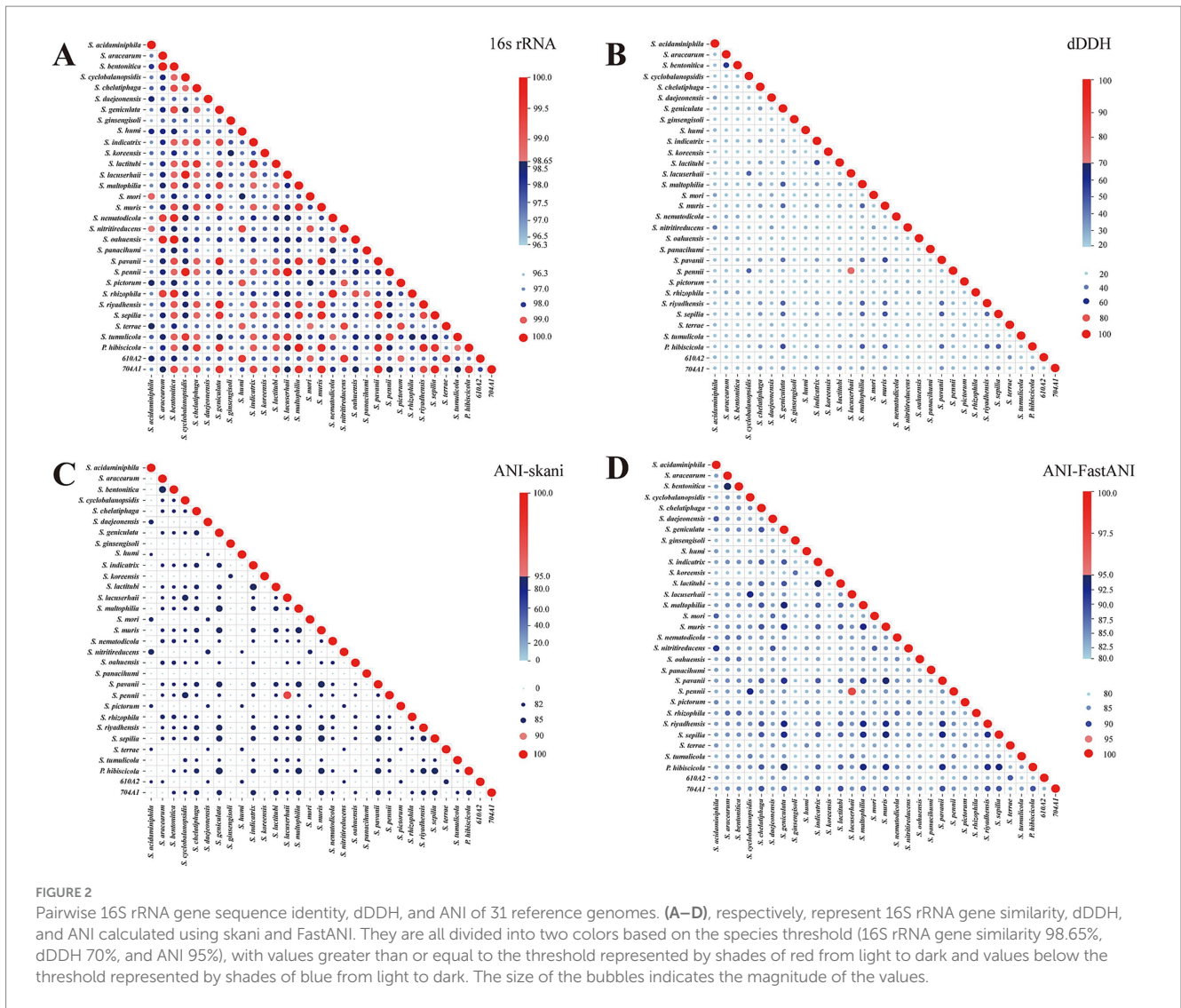


FIGURE 2 Pairwise 16S rRNA gene sequence identity, dDDH, and ANI of 31 reference genomes. (A–D), respectively, represent 16S rRNA gene similarity, dDDH, and ANI calculated using skani and FastANI. They are all divided into two colors based on the species threshold (16S rRNA gene similarity 98.65%, dDDH 70%, and ANI 95%), with values greater than or equal to the threshold represented by shades of red from light to dark and values below the threshold represented by shades of blue from light to dark. The size of the bubbles indicates the magnitude of the values.

3.4 ANI clustering and evaluation of the ANI cutoff values for *Stenotrophomonas*

We conducted paired ANI comparisons of all 2,213 *Stenotrophomonas* genomes and performed clustering. Out of a total of 4,897,369 comparisons of the genus *Stenotrophomonas* genomes, 4,196,423 yielded ANI values greater than 82%. For *S. maltophilia*, out of a total of 2,418,025 comparisons, 2,403,945 yielded ANI values in the 82–100% range. The distribution of paired ANI values of the genus *Stenotrophomonas* and *S. maltophilia* species is quite similar to the multimodal distribution (Figure 5A), which may be partially due to the fact that the genomes annotated as *S. maltophilia* account for the majority of the genus *Stenotrophomonas*.

We found that the peak of the ANI distribution is most abundant between 90 and 94.7%, accounting for approximately half of all ANI values. This indicated that numerous genomes have been mistakenly annotated as *S. maltophilia*, which actually belonged to other species closely related to *S. maltophilia* in the ANI clustering. The ANI values between the species classified in *S. mauri*, *S. sepilia*, *P. hibiscicola*, and *S. maltophilia* also fell in this range. Therefore, we recommended that a 90% ANI value could potentially serve as

the threshold for distinguishing complex/group species within the genus *Stenotrophomonas*.

In our study, despite a minor peak between 95.5 and 96.5%, the curve representing ANI values between 94.7 and 96.5% remained flat and very low, not exceeding 0.45% of all pairs, especially around 95%. Therefore, we suggested that the interspecies ANI threshold for the genus *Stenotrophomonas* might fall within this range. We examined the number of ANI clusters formed at different ANI cutoffs from 94 to 96% with every 0.1% increment and did not observe a stable plateau in the number of clusters (Figure 5B). Considering that 95% is the nadir of the kernel density distribution curve, as the cutoff for species delineation in prokaryotes, ANI ≥ 95% can correctly classify the genus *Stenotrophomonas* at the species level.

3.5 Correct classification based on the ANI cutoff

In the ANI classification, 758 (41.17% of 1,841) genomes had ANI values below 95%, with 158 genomes (8.58%) even below 90% (Table 1). Table 2 provides a detailed list of species with ANI values

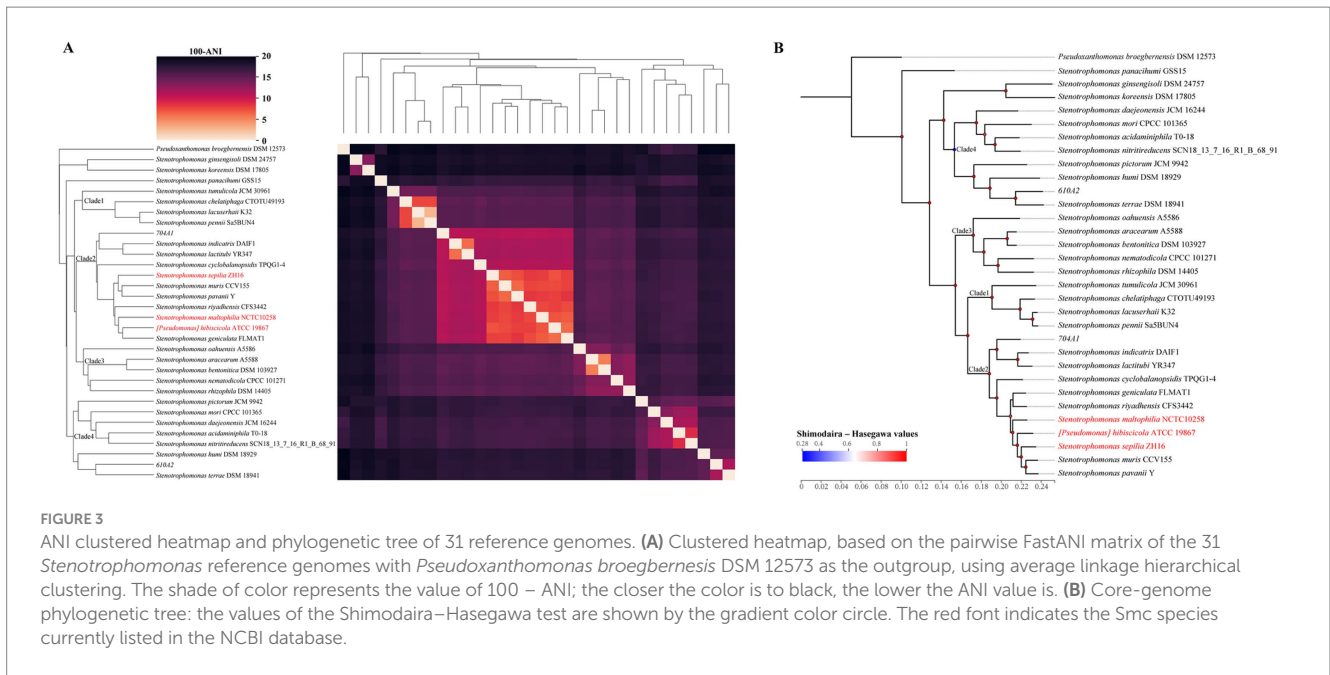


TABLE 1 Consistency between ANI calculated using skani and FastANI.

Skani	FastANI			Total
	ANI ≥ 95%	95%>ANI ≥ 90%	ANI<90%	
ANI ≥ 95%	1,083	0	0	1,083
95%>ANI ≥ 90%	0	600	0	600
ANI<90%	0	0	158	158
Total	1,083	600	158	1841

below 95% compared to the reference genomes. These results indicated that a large proportion of genomes had been incorrectly identified and annotated. Meanwhile, only 135 out of 372 *S. sp.* strains had ANI > 95% compared to the reference genomes and could be classified into the defined species.

Gröschel et al. classified Smc into 23 monophyletic lineages based on MALDI-TOF MS and 98.8% 16S rRNA gene sequence identity as the strain inclusion criteria (Gröschel et al., 2020). For consistency, we used and amended the naming convention of lineages from previous reports. In our study, 850 genomes, including the *S. maltophilia* reference genome NCTC10258, belonged to the lineage Sm6, which is strictly defined as *S. maltophilia* species by 95% ANI. Through ANI comparison, we discovered that the ANI values of several Smc lineages described by Gröschel et al. with species within or outside Smc were ≥ 95%, for example, Sm9 and Sm16, Sm3 and *S. seipilia*, and Sm5 and *P. hibiscicola* within Smc; Sm2 and *S. pavanii*, Sm4a and *S. muris*, Sm10 and *S. geniculata*, Sm15, and *S. riyadhensis*, and Sgin3 and *S. indicatrix* outside Smc. Moreover, it is worth mentioning that when the ANI threshold for the complex/group is set to 90%, only lineages Sm1–Sm18 should be kept in Smc, while lineages Sgn1–Sgn4 with more distant relationships should be removed.

The clustering heatmap of paired ANI comparisons of all 2,213 *Stenotrophomonas* genomes indicated that there are a considerable number of clusters within the genus *Stenotrophomonas* (Supplementary Figure S1). Using 95 and 90% ANI as species and

complex/group thresholds, 2,213 *Stenotrophomonas* genomes in the NCBI database were re-clustered into 126 ANI genomic species and 51 ANI genomic complex groups (Supplementary Tables S4, S5). The clustering result in our study suggested that Smc should be redefined. It may consist of 1,668 genomes from 33 species, including 18 previously described species and lineages (*P. hibiscicola*/Sm5, *S. geniculata*/Sm10, *S. maltophilia*/Sm6, *S. muris*/Sm4a, *S. pavanii*/Sm2, *S. riyadhensis*/Sm15, *S. seipilia*/Sm3, Sm1, Sm4b, Sm7–8, Sm9(Sm16), Sm11–14, and Sm17–18) and 15 novel ANI species.

The 10 genomes closely related to *S. panacihumi* were clustered into 4 complexes composed of *S. panacihumi* and other 7 ANI species (ANI species 117, 119–124). The ANI values between these genomes and other species in the genus *Stenotrophomonas* were below 82%. Whether plotted on the ANI clustering tree or the core gene phylogenetic tree, *S. panacihumi* is a clade that is far apart from other *Stenotrophomonas* species. Moreover, *Stenotrophomonas* sp. MMGLT7 (ANI species 125) should also be removed from the genus *Stenotrophomonas* as its genome (GCA_021168555.1) showed the highest identity with *Xanthomonas massiliensis* SN8 (GCA_900018785.1), with an ANI of 86.1%. The classification of *S. koreensis* and *S. ginsengisoli* was challenging. Their ANI values for other *Stenotrophomonas* species were even lower than those of *S. panacihumi*. However, on the core gene phylogenetic tree (Figure 3B), they were related to the clade including *S. terrae*, *S. nitritireducens*, and *S. acidaminiphila*. If *S. koreensis* and

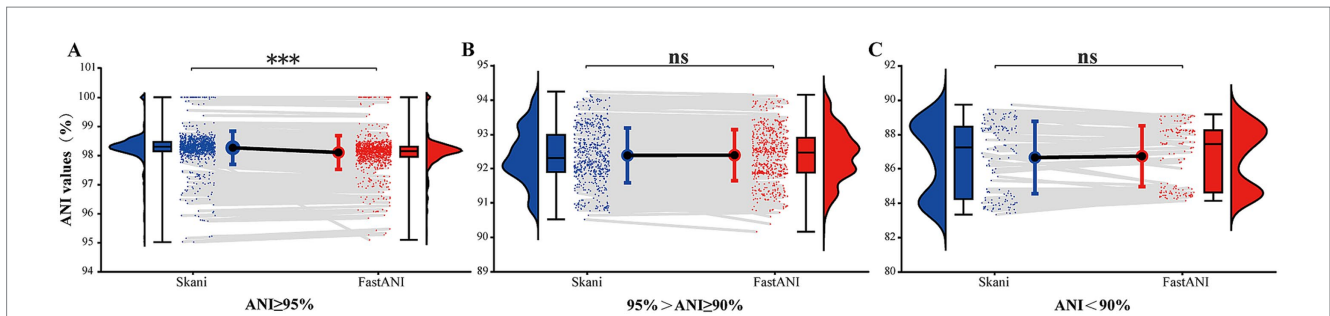


FIGURE 4 Raincloud plots of paired ANI comparisons calculated by two ANI tools. (A–C), respectively, represent 1,083 ANI values greater than 95%, 600 ANI values between 90 and 95%, and 152 ANI values below 90% (excluding 6 pairs of values for which the skani output was 0). The plots were created using raincloud plots v0.2.0 (Allen et al., 2019) in R 4.4.0. Blue represents skani, and red represents FastANI. “***” indicates $p < 0.001$, “ns” indicates $p > 0.05$.

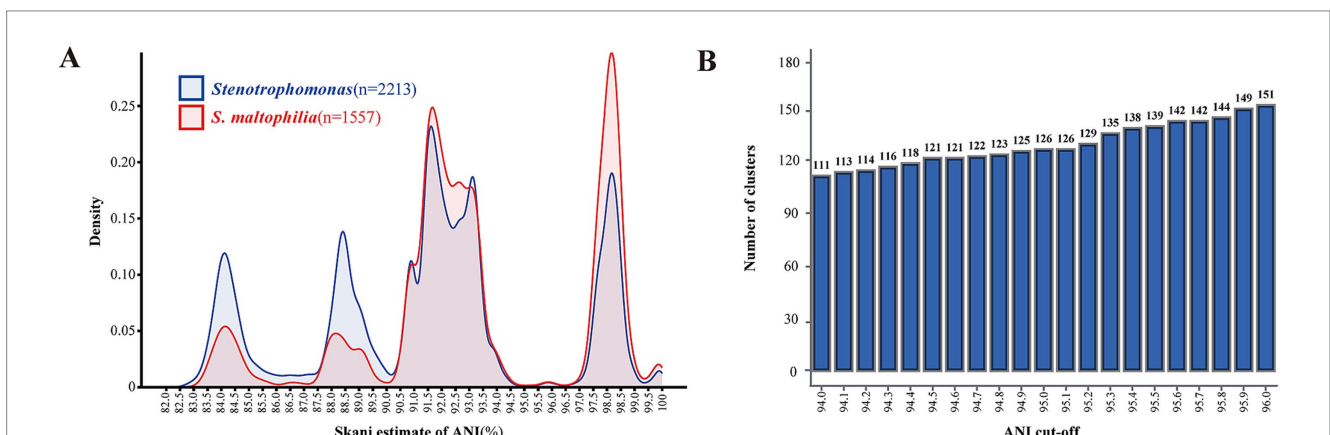


FIGURE 5 Clustering and distribution of ANI values. (A) Cluster heatmap of pairwise ANI comparison calculated for 2,213 *Stenotrophomonas* genomes, with colors ranging from light to dark, indicating nucleotide identity from high to low. (B) Kernel density plots of the distribution of paired ANI values of all *Stenotrophomonas* genomes and genomes annotated as *S. maltophilia*. (C) Clustering of 2,213 genomes based on pairwise comparison of ANI dissimilarity. The x-axis shows the ANI cutoffs from 94.0 to 96.0%, and the y-axis is the number of ANI clusters generated based on the ANI cutoffs.

TABLE 2 Species inconsistent with the reference genome.

Species name	ANI ≥ 95%	95% > ANI ≥ 90%	ANI < 90%	Total
<i>S. acidaminiphila</i>	25	17	0	42
<i>S. bentonitica</i>	2	0	10	12
<i>S. maltophilia</i>	850	583	122	1,555
<i>S. nitritireducens</i>	4	0	2	6
<i>S. rhizophila</i>	8	0	24	32
Total	889	600	158	1,647

S. ginsengisoli are retained, there would be 116 ANI genomic species and 47 ANI complexes in the genus *Stenotrophomonas*. In addition to the aforementioned species and their reference genomes for other ANI species or lineages, high-quality genomes were selected to serve as their respective reference genomes. A hierarchical clustering dendrogram was constructed using the ANI matrix of 125 reference genomes and annotated with both ANI genomic species and currently defined species (or lineages) information (Figure 6).

It should be pointed out that ANI clustering is based on all genomes within a cluster, which may result in some

‘transboundary’ occurrences. For example, *S. lactitubi* Bi35 (GCA_918698175.1) exhibited ANI values above 95% when compared to the reference genome of *S. lactitubi* (GCA_900188015.1) and 10 other genomes within this species. However, its ANI values to the remaining 13 genomes of the species were < 95%. According to the genomic similarity to the reference genome, we kept it in *S. lactitubi*. In the 54 paired comparisons between lineage sm4b (18 genomes) and ANI species 88 (3 genomes), 12 values were > 95%, and 42 values were < 95%; therefore, they were divided into separate species.

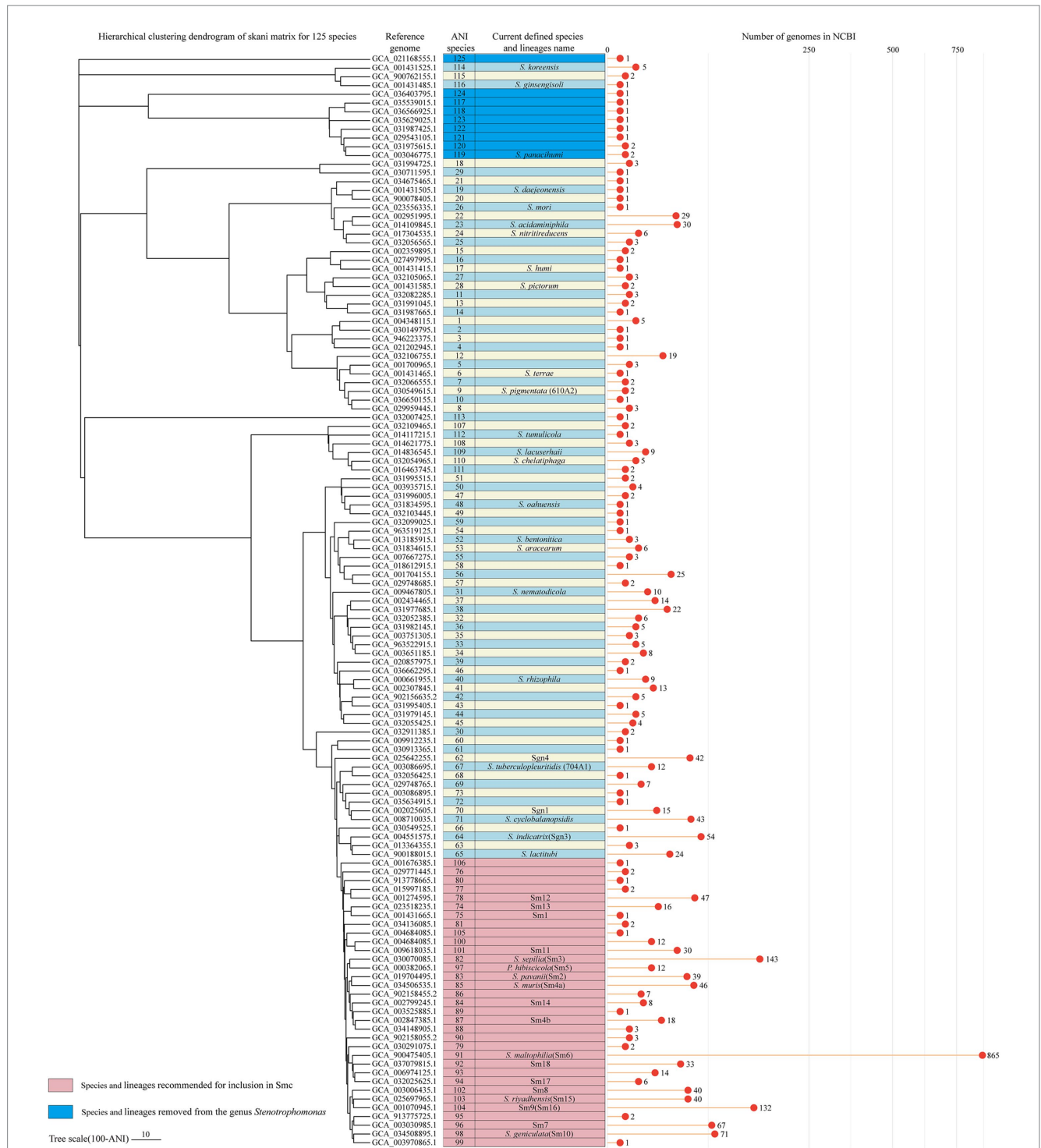


FIGURE 6 Hierarchical clustering dendrogram based on the skani matrix for the reference genomes of the 125 ANI genomic species. Average linkage hierarchical clustering was performed on the ANI matrix using Scipy v1.13.1, and the resulting Newick file was visualized in tvBOT. The first column indicates the corresponding ANI genomic species of each reference genome. The second column indicates the currently defined species and lineage names for these genomes. The lollipop plot shows the number of genomes belonging to each ANI genomic species (2,213 genomes in total).

4 Discussion

The genus *Stenotrophomonas* has been widely studied as its members have both clinical and environmental significance (Ryan et al., 2009; An and Berg, 2018; Pérez-Pérez et al., 2020; Kumar et al., 2023). Currently, species definitions of the genus *Stenotrophomonas*

are based on a variety of methods including traditional physiological and chemotaxonomic properties, 16S rRNA gene, or genome sequence. In recent decades, bacterial taxonomic techniques and standards have undergone a series of changes, shifting from phenotypes that are easily influenced by the environment to stable single/multi-gene sequence analysis with higher resolution and

further to the whole genome level. *Stenotrophomonas* is a genus with significant phenotypic and genomic variability intra- and inter-species. In this study, we proved that the 16S rRNA gene identity of many species within the genus *Stenotrophomonas* exceeded 98.65%, and the taxonomy based on the 16S rRNA gene was not effective and reliable in distinguishing *Stenotrophomonas* species.

The ANI has emerged as a rapid classifier of bacterial species, with a generally accepted 95% cutoff for a separate species (Wu et al., 2024). We assessed the consistency of the ANI and dDDH among the reference genomes of each species and found that they exhibited a consistent trend. Phylogenetic analysis based on core-genome alignments of reference genomes and ANI clustering were similar, especially in the identification of closely related species. We compared the consistency of two rapid ANI tools, skani (Shaw and Yu, 2023) and FastANI (Jain et al., 2018), in paired ANI values between *Stenotrophomonas* reference genomes and strains of their species. We discovered that when ANI values were > 95%, ANI values calculated using skani were slightly higher than those from FastANI, which has not been previously reported. This could be due to skani's key operating regimes being for medium-to-high ANI (>82%) (Shaw and Yu, 2023). Apart from that, the consistency between skani and FastANI is very high. Considering the speed, convenience, and feasibility of large-scale analysis, we used the ANI value reported by skani for species classification.

The distribution of paired ANI values in the genus *Stenotrophomonas* showed a multimodal distribution, which significantly differed from the previously reported wide gap or lack of a gap between 83 and 95% typically observed in prokaryotes (Jain et al., 2018). This valley range suggests genetic discontinuity in prokaryotic genomes, a pattern that appears to be disrupted in the genus *Stenotrophomonas*. This reflects that the genus *Stenotrophomonas* is genetically and phenotypically heterogeneous (Ryan et al., 2009; Vasileuskaya-Schulz et al., 2011; Berg and Martinez, 2015). The genus *Stenotrophomonas* has been shown to possess a number of functional properties, including, but not limited to, plant growth-promoting activity, nitrogen fixation, and rhizoremediation (Kumar et al., 2023). This diversity may be related to their genetic heterogeneity, which may be attributed to horizontal gene transfer. Zhao et al. demonstrated that *Stenotrophomonas* can introduce alien genes by exchanging genetic material with other community members in the habitat, promoting their broader genomic plasticity (Zhao et al., 2024).

We evaluated the ANI thresholds for the genus *Stenotrophomonas*, setting 90 and 95% as the thresholds for complex groups and species, respectively. Although the distribution of ANI values ranging from 94.7 to 96.5% was very low, we did not find a more suitable ANI cutoff value to replace the conventional ANI threshold of 95%. Based on the 95% ANI cutoff, we found many misclassified strains in the genus *Stenotrophomonas*, especially in *S. maltophilia* ($n = 705$, 45.34% of 1,555 strains). This may be the reason why Ciufu et al. chose 88.5% as the cutoff value for *S. maltophilia* (Ciufu et al., 2018). Similarly, this may explain why the genome size and GC content of *S. maltophilia* in our study are broader than those described in previous research (Xu et al., 2023).

Some species of this genus are pathogenic to humans. *S. maltophilia* is an important cause of hospital-acquired drug-resistant infections and is listed by the World Health Organization as one of the leading drug-resistant nosocomial pathogens worldwide (Brooke, 2012; Chang et al., 2015; Brooke, 2021). Lineage Sm6 (strictly defined as *S. maltophilia*) strains are potentially well-adapted to

colonize or infect humans (Gröschel et al., 2020). Unfortunately, a considerable proportion of current *S. maltophilia* strains do not belong to *S. maltophilia*, but rather to other species within the Smc or even more distantly related species. Therefore, the pathogenicity and clinical significance of *S. maltophilia* and other species within the genus *Stenotrophomonas* need to be carefully considered. In other words, some infections caused by other species have been mistakenly identified as *S. maltophilia* infections, for example, Sm2 and *S. pavanii* (same species, ANI $\geq 95\%$). The *S. pavanii* strain obtained from sugarcane is widely used in organic agriculture (Ramos et al., 2011) to improve the yield and quality of many crops. According to the sample details of the genome from NCBI, only a small proportion (9/27) of *S. pavanii* was isolated from humans. However, 39 of 49 Sm2 strains were isolated from humans (Gröschel et al., 2020).

Furthermore, we observed anomalies in the phylogenetic tree of the core genes and ANI clustering of *S. panacihumi*, which is consistent with other studies (Vinuesa et al., 2018). Although *S. koreensis* and *S. ginsengisoli* had significantly lower ANI values, they still belonged to the genus *Stenotrophomonas* on the phylogenetic tree. Therefore, like the GTDB, we also recommend removing the whole *S. panacihumi* clade from the genus *Stenotrophomonas*, but the relationship of *S. koreensis* and *S. ginsengisoli* with the genus requires further biochemical and physiological studies for verification.

In conclusion, this study identified significant misclassified taxa within the genus *Stenotrophomonas* by ANI analysis. We identified an ANI cutoff of 95% for a genomic species delineation and 90% for genomic complexes/groups. After removing outliers, we defined 116 ANI genomic species and 47 ANI genomic complexes. Due to the opportunistic pathogenicity of some species of this genus, a well-defined ANI genomic species classification for *Stenotrophomonas* and Smc is crucial for species identification, diagnosis, clinical treatment, and epidemiological studies.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

Z-LY: Investigation, Visualization, Writing – original draft, Writing – review & editing. R-BW: Conceptualization, Funding acquisition, Investigation, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the project of ICDC (33052).

Conflict of interest

WR conceived the project, review and editing the manuscript.

The remaining author declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2024.1488674/full#supplementary-material>.

References

- Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., van Langen, J., and Kievit, R. A. (2019). Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Research*. 4:63. doi: 10.12688/wellcomeopenres.15191.2
- An, S. Q., and Berg, G. (2018). *Stenotrophomonas maltophilia*. *Trends Microbiol.* 26, 637–638. doi: 10.1016/j.tim.2018.04.006
- Berg, G., and Martinez, J. L. (2015). Friends or foes: can we make a distinction between beneficial and harmful strains of the *Stenotrophomonas maltophilia* complex? *Front. Microbiol.* 6:241. doi: 10.3389/fmicb.2015.00241
- Brooke, J. S. (2012). *Stenotrophomonas maltophilia*: an emerging global opportunistic pathogen. *Clin. Microbiol. Rev.* 25, 2–41. doi: 10.1128/CMR.00019-11
- Brooke, J. S. (2021). Advances in the microbiology of *Stenotrophomonas maltophilia*. *Clin. Microbiol. Rev.* 34:e0003019. doi: 10.1128/CMR.00030-19
- Chang, Y. T., Lin, C. Y., Chen, Y. H., and Hsueh, P.-R. (2015). Update on infections caused by *Stenotrophomonas maltophilia* with particular attention to resistance mechanisms and therapeutic options. *Front. Microbiol.* 6:893. doi: 10.3389/fmicb.2015.00893
- Ciufo, S., Kannan, S., Sharma, S., Badretin, A., Clark, K., Turner, S., et al. (2018). Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int. J. Syst. Evol. Microbiol.* 68, 2386–2392. doi: 10.1099/ijsem.0.002809
- Deng, Y., Han, X.-F., Jiang, Z.-M., Yu, L.-Y., Li, Y., and Zhang, Y.-Q. (2022). Characterization of three *Stenotrophomonas* strains isolated from different ecosystems and proposal of *Stenotrophomonas mori* sp. nov. and *Stenotrophomonas lacuserhaii* sp. nov. *Front. Microbiol.* 13:1056762. doi: 10.3389/fmicb.2022.1056762
- Gilroy, R., Ravi, A., Getino, M., Pursley, I., Horton, D. L., Alikhan, N.-F., et al. (2021). Extensive microbial diversity within the chicken gut microbiome revealed by metagenomics and culture. *PeerJ*. 9:e10941. doi: 10.7717/peerj.10941
- Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijms.0.64483-0
- Gosselin, S., Fullmer, M. S., Feng, Y., and Gogarten, J. P. (2022). Improving phylogenies based on average nucleotide identity, incorporating saturation correction and nonparametric bootstrap support. *Syst. Biol.* 71, 396–409. doi: 10.1093/sysbio/syab060
- Gröschel, M. I., Meehan, C. J., Barilar, I., Diricks, M., Gonzaga, A., Steglich, M., et al. (2020). The phylogenetic landscape and nosocomial spread of the multidrug-resistant opportunist *Stenotrophomonas maltophilia*. *Nat. Commun.* 11:2044. doi: 10.1038/s41467-020-15123-0
- Hua, Z., Li, G., Zhuo, C., Yang, Y., Shi, Y., Chen, B., et al. (2013). Expert consensus on diagnosis, treatment and prevention and control of *Stenotrophomonas maltophilia* infection in China. *Chinese Med J.* 93, 1203–1213. doi: 10.3760/cma.j.issn.0376-2491.2013.16.002
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9
- Kaiser, S., Biehler, K., and Jonas, D. (2009). A *Stenotrophomonas maltophilia* multilocus sequence typing scheme for inferring population structure. *J. Bacteriol.* 191, 2934–2943. doi: 10.1128/JB.00892-08
- Kim, M., Oh, H.-S., Park, S.-C., and Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 64, 346–351. doi: 10.1099/ijms.0.059774-0
- Kim, H.-B., Srinivasan, S., Sathiyaraj, G., Quan, L.-H., Kim, S.-H., Bui, T. P. N., et al. (2010). *Stenotrophomonas ginsengisoli* sp. nov., isolated from a ginseng field. *Int. J. Syst. Evol. Microbiol.* 60, 1522–1526. doi: 10.1099/ijms.0.014662-0
- Konstantinidis, K. T., and Tiedje, J. M. (2007). Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. *Curr. Opin. Microbiol.* 10, 504–509. doi: 10.1016/j.mib.2007.08.006
- Kumar, A., Rithesh, L., Kumar, V., Raghuvanshi, N., Chaudhary, K., and Abhineet Pandey, A. K. (2023). *Stenotrophomonas* in diversified cropping systems: friend or foe? *Front. Microbiol.* 14:1214680. doi: 10.3389/fmicb.2023.1214680
- Li, Y., Yu, Z., Fan, X., Xu, D., Liu, H., Zhao, X., et al. (2024). A novel pathogenic species of genus *Stenotrophomonas*: *Stenotrophomonas pigmentata* sp. nov. *Front. Cell. Infect. Microbiol.* 14:1410385. doi: 10.3389/fcimb.2024.1410385
- Ludwig, W., and Klenk, H.-P. (2001). “Overview: a phylogenetic backbone and taxonomic framework for prokaryotic systematics” in *Bergey's manual*® of systematic bacteriology: Volume one: The Archaea and the deeply branching and phototrophic Bacteria. eds. D. R. Boone, R. W. Castenholz, G. M. Garrity, D. R. Boone, R. W. Castenholz and G. M. Garrity (New York, NY: Springer), 49–65.
- Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L., and Göker, M. (2022). TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* 50, D801–D807. doi: 10.1093/nar/gkab902
- Meier-Kolthoff, J. P., and Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* 10:2182. doi: 10.1038/s41467-019-10210-3
- Ochoa-Sánchez, L. E., and Vinuesa, P. (2017). Evolutionary genetic analysis uncovers multiple species with distinct habitat preferences and antibiotic resistance phenotypes in the *Stenotrophomonas maltophilia* complex. *Front. Microbiol.* 8:1548. doi: 10.3389/fmicb.2017.01548
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421
- Palleroni, N. J., and Bradbury, J. F. (1993). *Stenotrophomonas*, a new bacterial genus for *Xanthomonas maltophilia* (Hugh 1980) Swings et al. 1983. *Int. J. Syst. Bacteriol.* 43, 606–609. doi: 10.1099/00207713-43-3-606
- Parks, D. H., Chuvpochina, M., Chaumeil, P.-A., Rinke, C., Mussig, A. J., and Hugenholtz, P. (2020). A complete domain-to-species taxonomy for bacteria and archaea. *Nat. Biotechnol.* 38, 1079–1086. doi: 10.1038/s41587-020-0501-8
- Pearce, M. E., Langridge, G. C., Lauer, A. C., Grant, K., Maiden, M. C. J., and Chattaway, M. A. (2021). An evaluation of the species and subspecies of the genus *Salmonella* with whole-genome sequence data: proposal of type strains and epithets for novel *S. enterica* subspecies VII, VIII, IX, X and XI. *Genomics* 113, 3152–3162. doi: 10.1016/j.ygeno.2021.07.003
- Pérez-Pérez, R., Oudot, M., Hernández, I., Nápoles, M. C., Pérez-Martínez, S., and Castillo, D. S.-D. (2020). Isolation and characterization of *Stenotrophomonas* associated to maize (*Zea Mays* L.) rhizosphere. *Cultivos Tropicales* 41:e03.
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490
- Ramos, P. L., Van Trappen, S., Thompson, F. L., Rocha, R. C. S., Barbosa, H. R., De Vos, P., et al. (2011). Screening for endophytic nitrogen-fixing bacteria in Brazilian sugar cane varieties used in organic farming and description of *Stenotrophomonas pavani* sp. nov. *Int. J. Syst. Evol. Microbiol.* 61, 926–931. doi: 10.1099/ijms.0.019372-0
- Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. USA* 106, 19126–19131. doi: 10.1073/pnas.0906412106
- Rizek, C. F., Jonas, D., García Paez, J. I., Rosa, J. F., Perdigão Neto, L. V., Martins, R. R., et al. (2018). Multidrug-resistant *Stenotrophomonas maltophilia*: description of new MLST profiles and resistance and virulence genes using whole-genome sequencing. *J. Glob. Antimicrob. Resist.* 15, 212–214. doi: 10.1016/j.jgar.2018.07.009
- Rosselló-Móra, R., and Amann, R. (2015). Past and future species definitions for bacteria and archaea. *Syst. Appl. Microbiol.* 38, 209–216. doi: 10.1016/j.syapm.2015.02.001
- Ryan, R. P., Monchy, S., Cardinale, M., Taghavi, S., Crossman, L., Avison, M. B., et al. (2009). The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. *Nat. Rev. Microbiol.* 7, 514–525. doi: 10.1038/nrmicro2163

- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sentausa, E., and Fournier, P. E. (2013). Advantages and limitations of genomics in prokaryotic taxonomy. *Clin. Microbiol. Infect.* 19, 790–795. doi: 10.1111/1469-0691.12181
- Shaw, J., and Yu, Y. W. (2023). Fast and robust metagenomic sequence comparison through sparse chaining with skani. *Nat. Methods* 20, 1661–1665. doi: 10.1038/s41592-023-02018-3
- Shen, W., Sipos, B., and Zhao, L. (2024). SeqKit2: a Swiss army knife for sequence and alignment processing. *iMeta*, e191 3, e191. doi: 10.1002/imt2.191
- Svensson-Stadler, L. A., Mihaylova, S. A., and Moore, E. R. B. (2012). *Stenotrophomonas* interspecies differentiation and identification by gyrB sequence analysis. *FEMS Microbiol. Lett.* 327, 15–24. doi: 10.1111/j.1574-6968.2011.02452.x
- Swings, J., De Vos, P., Van den Mooter, M., and De Ley, J. (1983). Transfer of *Pseudomonas maltophilia* Hugh 1981 to the genus *Xanthomonas* as *Xanthomonas maltophilia* (Hugh 1981) comb. nov. *Int. J. Syst. Bacteriol.* 33, 409–413. doi: 10.1099/00207713-33-2-409
- Tamma, P. D., Aitken, S. L., Bonomo, R. A., Mathers, A. J., van Duin, D., and Clancy, C. J. (2022). Infectious diseases society of america guidance on the treatment of AmpC β -lactamase-producing Enterobacterales, carbapenem-resistant Acinetobacter baumannii, and *Stenotrophomonas maltophilia* infections. *Clin. Infect. Dis.* 74, 2089–2114. doi: 10.1093/cid/ciab1013
- Vandamme, P., and Peeters, C. (2014). Time to revisit polyphasic taxonomy. *Antonie Van Leeuwenhoek* 106, 57–65. doi: 10.1007/s10482-014-0148-x
- Vasileuskaya-Schulz, Z., Kaiser, S., Maier, T., Kostrzewa, M., and Jonas, D. (2011). Delineation of *Stenotrophomonas* spp. by multi-locus sequence analysis and MALDI-TOF mass spectrometry. *Syst. Appl. Microbiol.* 34, 35–39. doi: 10.1016/j.syapm.2010.11.011
- Venkidesamy, K., and Megharaj, M. (2016). Identification of electrode respiring, hydrocarbonoclastic bacterial strain *Stenotrophomonas maltophilia* MK2 highlights the untapped potential for environmental bioremediation. *Front. Microbiol.* 7:1965. doi: 10.3389/fmicb.2016.01965
- Vinuesa, P., Ochoa-Sánchez, L. E., and Contreras-Moreira, B. (2018). GET_PHYLOMARKERS, a software package to select optimal orthologous clusters for phylogenomics and inferring pan-genome phylogenies, used for a critical genotaxonomic revision of the genus *Stenotrophomonas*. *Front. Microbiol.* 9:771. doi: 10.3389/fmicb.2018.00771
- Wu, R., Payne, M., Zhang, L., and Lan, R. (2024). Uncovering the boundaries of *Campylobacter* species through large-scale phylogenetic and nucleotide identity analyses. *mSystems*. 9, e01218–e01223. doi: 10.1128/mSystems.01218-23
- Xie, J., Chen, Y., Cai, G., Cai, R., Hu, Z., and Wang, H. (2023). Tree visualization by one table (tvBOT): a web application for visualizing, modifying and annotating phylogenetic trees. *Nucleic Acids Res.* 51, W587–W592. doi: 10.1093/nar/gkad359
- Xiong, W., Yin, C., Wang, Y., Lin, S., Deng, Z., and Liang, R. (2020). Characterization of an efficient estrogen-degrading bacterium *Stenotrophomonas maltophilia* SJTH1 in saline-, alkaline-, heavy metal-contained environments or solid soil and identification of four 17 β -estradiol-oxidizing dehydrogenases. *J. Hazard. Mater.* 385:121616. doi: 10.1016/j.jhazmat.2019.121616
- Xu, Y., Cheng, T., Rao, Q., Zhang, S., and Ma, Y. L. (2023). Comparative genomic analysis of *Stenotrophomonas maltophilia* unravels their genetic variations and versatility trait. *J. Hazard. Mater.* 64, 351–360. doi: 10.1007/s13353-023-00752-0
- Yang, H. C., Im, W. T., Kang, M. S., Shin, D. Y., and Lee, S. T. (2006). *Stenotrophomonas koreensis* sp. nov., isolated from compost in South Korea. *Int. J. Syst. Evol. Microbiol.* 56, 81–84. doi: 10.1099/ijs.0.63826-0
- Yi, H., Srinivasan, S., and Kim, M. K. (2010). *Stenotrophomonas panacihumi* sp. nov., isolated from soil of a ginseng field. *J. Microbiol.* 48, 30–35. doi: 10.1007/s12275-010-0006-0
- Zhao, Y., Ding, W. J., Xu, L., and Sun, J. Q. (2024). A comprehensive comparative genomic analysis revealed that plant growth promoting traits are ubiquitous in strains of *Stenotrophomonas*. *Front. Microbiol.* 15:1395477. doi: 10.3389/fmicb.2024.1395477