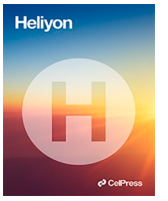
Heliyon 11 (2025) e41938

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Evaluating the potential of assembler-binner combinations in recovering low-abundance and strain-resolved genomes from human metagenomes

Hajra Qayyum [a,b], Muhammad Sarfraz Talib [a], Amjad Ali [a,**], Masood Ur Rehman Kayani [c,*]

[a] Integrative Biology Laboratory, Department of Microbiology and Biotechnology, Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Srinagar Highway, Sector H-12, Islamabad, Pakistan
[b] Capital University of Science & Technology, Islamabad Expressway, Kahuta Road Zone-V Sihala, Islamabad, Pakistan
[c] Metagenomics Discovery Lab, School of Interdisciplinary Engineering and Sciences (SINES), National University of Sciences and Technology (NUST), Srinagar Highway, Sector H-12, Islamabad, Pakistan

ARTICLE INFO

Keywords:
Gut microbiome
De novo assembly
Genome binning
Metagenome assembled genomes (MAGs)
Low abundance species
Strain-resolved analysis

ABSTRACT

Human-associated microbial communities are a complex mixture of bacterial species and diverse strains prevalent at varying abundances. Due to the inherent limitations of metagenomic assemblers and genome binning tools in recovering low-abundance species (<1 %) and strains, we lack comprehensive insight into these communities. Although many bioinformatics approaches are available for recovering metagenome-assembled genomes, their effectiveness in recovering low-abundance species and strains is often questioned. Moreover, each tool has its trade-offs, making selecting the right tools challenging. In this study, we investigated the combinatory effect of various assemblers and binning tools on the recovery of low-abundance species and strain-resolved genomes from real and simulated human metagenomes. We evaluated the performance of nine combinations of metagenome assemblers and genome binning tools for their potential to recover genomes of useable quality. Our results revealed that the metaSPAdes-MetaBAT2 combination is highly effective in recovering low-abundance species, while MEGAHIT-MetaBAT2 excels in recovering strain-resolved genomes. These findings highlight the significant variation in the performance of different combinations, even when aiming for the same objective. This suggests the profound impact of selecting the right assembler-binner combination for metagenome analyses. We believe this study will be a cornerstone for the scientific community, guiding the choice of tools by highlighting their complementary effects. Furthermore, it underscores the potential of existing tools to address the current challenges in the field improving the recovery of information from metagenomes.

1. Introduction

Humans are colonized by trillions of microbes, representing a diverse and coexisting community of microorganisms [1]. These

* Corresponding author.
** Corresponding author.
E-mail addresses: amjad.ali@asab.nust.edu.pk (A. Ali), m.kayani@sines.nust.edu.pk (M.U.R. Kayani).

https://doi.org/10.1016/j.heliyon.2025.e41938
Received 20 August 2024; Received in revised form 8 January 2025; Accepted 13 January 2025
Available online 14 January 2025

communities, ranging from a complex gut microbiome to a simpler skin microbiome, play a crucial role in maintaining human health [2]. In recent years, genome-resolved metagenomics has emerged as a powerful approach to unveil the significant impact of the microbiome on human health and diseases [3]. These analyses are conducted using two key computational processes: metagenomic assembly and genome binning. Metagenomic assembly commonly requires specialized de Bruijn graph (DBG) based assemblers to recover information about the genomic context. Whereas the resulting contigs are disentangled by grouping them into separate bins (further referred to as metagenome-assembled genomes (MAGs)) using dedicated genome binning tools [4].

Recovery of MAGs facilitates the unprecedented analysis of genetic diversity among microbiomes, the identification of unculturable bacterial species, and the unraveling of the etiology of complex diseases [5]. However, the recovered MAGs usually represent the most abundant taxa in the metagenomes [6]. Similarly, due to the inherent limitations of assemblers or genome binners, recovery of strains also becomes difficult [7]. Thus, there is a reasonable probability that a major proportion of information is usually lost from the metagenomes yielding limited insights [4]. This is particularly concerning considering the crucial role of low-abundance genomes in human health, for example, lactic acid bacteria constitute <1 % of the human gut microbiome [8]. Likewise, skin can be colonized by multiple strains of a species such as *C. acnes*, where individual pores contain a different clonal strain suggesting the spatial significance of strain variations [7].

Multiple bioinformatics tools have been developed to facilitate the metagenome assembly and binning process, but each has its limitations and challenges [9]. Therefore, selecting suitable assembly and binning tools remains crucial for optimal downstream analysis [6]. The choice of the right tool for genome-resolved analyses can be challenging, especially for less-experienced researchers, due to the numerous comparative studies on these tools [10–14]. These studies conclude that there is no optimal solution, and the selection of tools depends on the nature of the dataset and the research objectives [15]. They also deduced that the selection of any metagenomic tools presents trade-off effects. For instance, though MEGAHIT is a computationally efficient option, it achieves this efficiency at the cost of increased misassemblies and reduced contiguity relative to the MetaVelvet and metaSPAdes, respectively [15]. This makes MEGAHIT a good choice for resource-limited settings but not an optimal solution. Likewise, for recovering the genomes of low-abundance taxa, the choice of optimal assembler depends on whether the accuracy or genomic context is valued the most [16]. Furthermore, the Critical Assessment of Metagenome Interpretation (CAMI) challenge's (https://data.cami-challenge.org/) Assessment of Metagenome BinnERs (AMBER) benchmarked various MAG recovery methods in terms of global completeness and bin purity [17]. Similarly, a recent study used the AMBER approach to evaluate 15 different binning tools with a common metaSPAdes assembly [18]. However, to the best of our knowledge, no study has specifically tested assembler-binner combinations for the recovery of low abundance species or strain-resolved MAGs, despite their biological significance.

Considering the significant role of the right tools for optimal results, we hypothesized that pairing the assembler and binning tool in the right combination can facilitate the optimal recovery of information from the metagenomes. Therefore, this study aimed to evaluate various widely used assembler-binner combinations for their effectiveness in recovering low-abundance species or strain-resolved MAGs. We believe our findings will guide the scientific community in selecting the most appropriate combination for their specific objectives. Moreover, this study will serve as a cornerstone in showcasing the potential of various assembler-binner combinations. Additionally, our study will pave the way for a more precise and comprehensive recovery of information from metagenomes, ultimately enhancing our understanding of microbial communities and their roles in human health and disease.

2. Materials and methods

2.1. Selection of metagenome assemblers and genome binning tools

For the recovery of MAGs from metagenomes of varying complexity, we selected the top three DBG-based *de novo* assemblers and three widely adopted genome binners for their efficiency with both high and low-complexity metagenomes as reported elsewhere, [14–19]. Specifically, we recruited MEGAHIT (v1.2.9) [20] metaSPAdes (v3.15.3) [21], and IDBA-UD (v1.1.3) [22] as assemblers and MetaBAT2 (v1.7) [23], MaxBin2.0 (v2.2.4) [24] and CONCOCT (v1.1) [25] as genome binners. This approach resulted in nine assembler-binner combinations: MEGAHIT-MetaBAT2, MEGAHIT-MaxBin2.0, MEGAHIT-CONCOCT, metaSPAdes-MetaBAT2, metaSPAdes-MaxBin2.0, metaSPAdes-CONCOCT, IDBA-UD-MetaBAT2, IDBA-UD-MaxBin2.0, and IDBA-UD-CONCOCT.

2.2. Metagenome dataset and preprocessing

For the assessment of the assembler-binner combinations, we utilized publicly accessible real and simulated datasets. The real metagenomes were recruited from the Human Microbiome Project Phase III (Supplementary Table 1). It included 70 gut and 30 anterior nares samples (further referred to as skin samples), representing high- and low-complexity metagenomes, respectively. These metagenomes were retrieved from the NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra) using the prefetch (v3.0.2) and fasterq-dump (v3.0.2) tools from the SRA Toolkit (https://github.com/ncbi/sra-tools). We utilized the mOTUs2 reference MAGs, gold standard taxonomic, and relative abundance profiles [26] for simulating metagenomes using InSilicoSeq 2.0 [27]. Quality assessment of the raw FASTQ reads was conducted using FastQC (v0.12.1) (https://github.com/s-andrews/FastQC), followed by pre-processing with fastp (v0.23.2) [28]. This pre-processing included the removal of low-quality bases ($Q < 20$), adaptor sequences, and duplicate reads (using parameters: *q, –detect_adapter_for_pe,* and *–dedup,* respectively).

2.3. De-novo metagenome assembly and assessment

High-quality FASTQ reads of real and simulated metagenomes were assembled *de-novo* via the KBase platform (https://www.kbase.us/) using a single-sample assembly approach. The assemblies were performed independently across the three assemblers with their default settings, except for a minimum contig length (2000 bp). The quality of metagenome assemblies was evaluated based on assembly size, total number of contigs in the assembly, length of the largest contig, and N50. The pairwise statistical comparison of these results was performed in R studio (v4.3.0) using the Wilcoxon Rank Sum test [29]. However, we did not evaluate the assemblies for their accuracy and coverage as these aspects have already been comprehensively assessed in multiple studies [10,15,30–32].

2.4. Genome binning and quality assessment

To recover MAGs from the metagenomes across all the assembler-binner combinations, we independently binned the three types of metagenome assemblies using the single-sample approach across the selected genome binners with the default settings. Quality assessment of the recovered MAGs, in terms of completeness and contamination levels, was conducted using the lineage workflow (lineage_wf) from CheckM (v1.0.18) [33]. The recovered MAGs were classified into high, medium, or low by adopting the completeness and contamination thresholds as defined by Minimum Information about Metagenome-Assembled Genome (MIMAG) [34]. However, we did not consider the proportions of 5S, 16S, 23S rRNA, and tRNA during the quality evaluation. Consequently, a high-quality (HQ) MAG was defined to have a completeness level of >90 % and contamination <5 %. In contrast, medium-quality (MQ) MAGs were defined as those with completeness ≥50 % and contamination <10 %. Otherwise, the MAGs were classified as

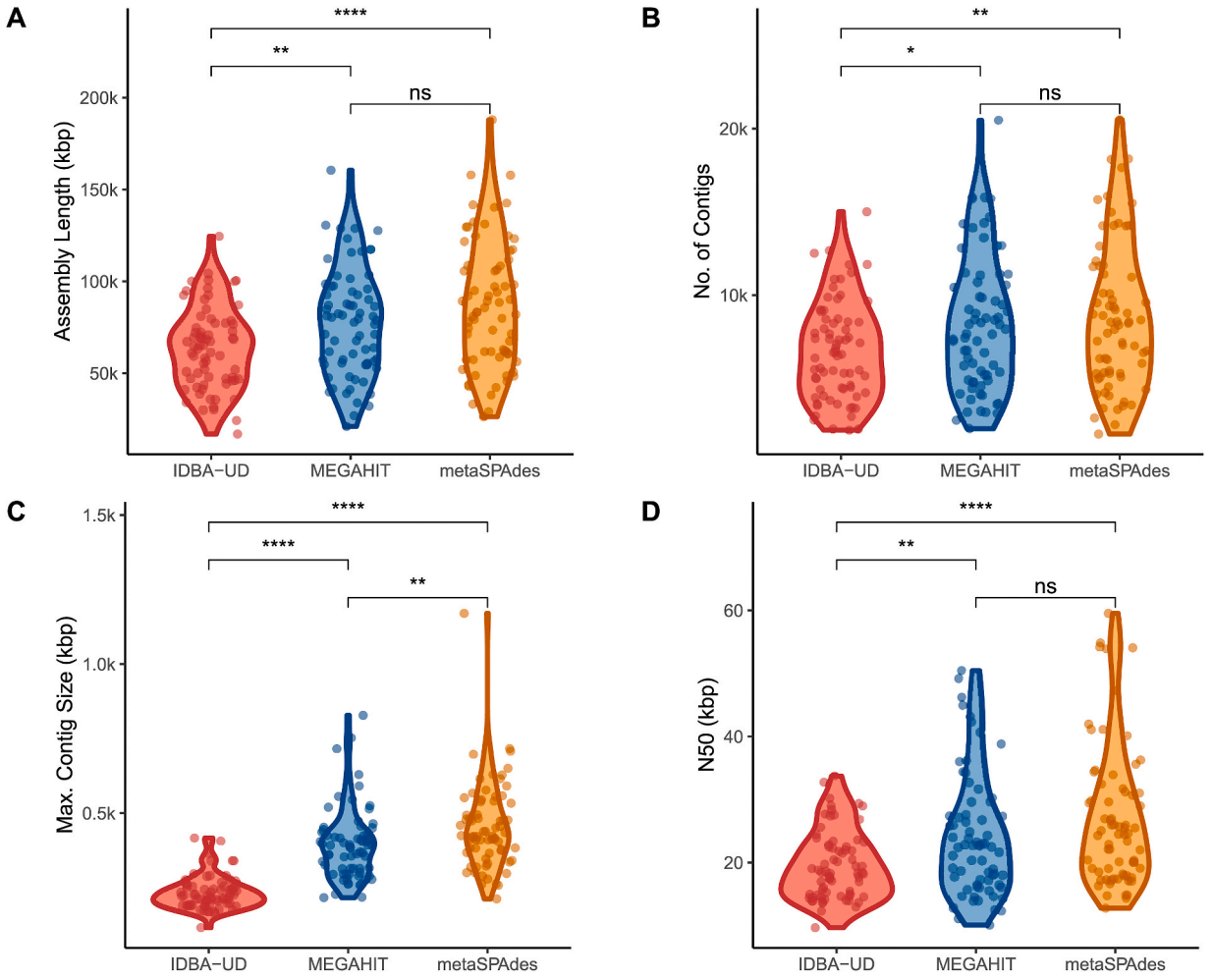


Fig. 1. Comparative performance analysis of assemblers across the defined metrics for real human gut metagenomes. (A) Assembly size produced by three assemblers. (B) Number of contigs contained in assemblies produced by the assemblers. (C) Maximum contig length generated by the assemblers. (D) N50 lengths of the assembly attained by different assemblers. Note: In the figure, asterisks represent the p-value as determined by the Wilcoxon rank sum test (* = p-value <0.1, ** = p-value <0.01, **** = p-value <0.0001).

low quality (LQ). It is important to note that we excluded the evaluation of proportions related to 5S, 16S, 23S rRNA, and tRNA.

2.5. Taxonomic classification and low-abundance taxa

We determined the composition of real metagenomes and the relative abundances of taxa utilizing a read-based approach using Kaiju (v1.9.0) [35]. The relative abundance of species is calculated using the number of reads mapped to a specific taxon relative to the total number of reads classified in that metagenome. Reads were classified in greedy mode using default settings, with the RefSeq database as the reference. Here, species with <1 % abundance in at least 5 % of metagenome samples were defined as low abundance. Additionally, to evaluate the combinations' potential to recover very low abundance taxa, we set the low abundance filter to 0.05 %. Next, for the taxonomic classification of the nine sets of MAGs recovered across the combinations from real as well as simulated metagenomes, we employed GTDB-Tk (v2.3.2) [36] with its database version r214 with default settings. Gold-standard relative abundance and phylotype profiles were used to compare results obtained for simulated MAGs.

2.6. Recovery of strain-resolved MAGs

To evaluate the potential of recovering strain-resolved MAGs across all the combinations, we used whole genome Average Nucleotide Identity (ANI) scores calculated using FASTANI [18]. These values were inferred from GTDB-Tk results obtained during the taxonomic classification of MAGs. Here strains were defined as the MAGs assigned to the same species and shared an ANI value that is >95 % and ≤99 % with the reference genome [37,38].

2.7. Evaluation criteria for the assessment of assembler-binner combinations

The performance of the nine assembler-binner combinations was evaluated based on the following criteria: (i) efficient recovery of microbial species from metagenomes, as assessed by the total number of recovered MAGs and the fraction of useable (HQ + MQ) MAGs. (ii) high-resolution genome binning measured by the ability to recover microbial species present at low abundance in metagenomes; and (iii) improved potential to recover strains from assemblies in terms of the number and quality of the recovered strains.

3. Results

3.1. De novo assembly across the metagenomes of varying complexities

Evaluation of the real metagenome assemblies revealed metaSPAdes producing a larger assembly length, a maximum size of the largest contig, and a higher N50 length, identifying it as the best assembler for gut metagenomes. In contrast, MEGAHIT produced comparable assemblies. However, IDBA-UD performed significantly worse than the other two assemblers, indicating that it was a less suitable option for assembling complex metagenomes such as the gut (Fig. 1A–D). When compared with MEGAHIT (77451.2 ± 29098.3 kbp), metaSPAdes produced a larger assembly length i.e. 86707.6 ± 34752.6 kbp with an insignificant statistical difference ($p>0.05$) as shown in Fig. 1A. Whereas, IDBA-UD performed the worst, with a total assembly length of 63477.2 ± 22156.1 kbp, which was significantly smaller than MEGAHIT ($p<0.001$) and metaSPAdes ($p < 0.0001$). In terms of contiguity, metaSPAdes outperformed by yielding a considerably larger maximum contig size and better N50 lengths with an average of 454.4 ± 144.5 kbp and 26.9 ± 10.9 kbp respectively. MEGAHIT yielded comparably contiguous assembly with an average maximum contig size of 392.3 ± 118.9 kbp and an average N50 length of 24.1 ± 9.6 kbp. In contrast, IDBA-UD assemblies were the most fragmented and least contiguous. It was also the only assembler to generate an N50 length of less than 20 kbp, which was significantly lower than those produced by metaSPAdes ($p<0.0001$) and MEGAHIT ($p<0.01$). For the largest contig size, IDBA-UD underperformed, with a statistically significant difference of $p<0.0001$ compared to both assemblers (Fig. 1C–D and Supplementary Table 2). However, the differences in assembly metrics among the three assemblers were statistically insignificant for skin metagenomes. Our results indicated that MEGAHIT is the most optimal solution, metaSPAdes is the second, and IDBA-UD is the least efficient assembler for assembling skin metagenomes (Supplementary Figs. 1A–1D). These results underscore the differences in the compatibilities of different assemblers with datasets of varying complexities.

For the simulated metagenome, metaSPAdes demonstrated its superior performance achieving the largest assembly size (mean: 200902.3 ± 11603.0 kbp), the maximum length of the largest contig (mean: 671.3 ± 148.7 kbp), and the highest N50 value (mean: 25399.8 ± 3655.9 kbp). MEGAHIT ranked second with a mean assembly size of 177265.1 ± 12327.7 kbp, and a maximum contig size of 530.6 ± 150.8 kbp. Despite producing the smallest assembly and contig size, IDBA-UD outperformed MEGAHIT in N50 values, achieving an average of 22127.4 ± 2933.9 kbp compared to MEGAHIT's 21905.9 ± 3380.9 kbp (Supplementary Table 3).

3.2. Genome binning and recovery of useable MAGs

The total number of MAGs recovered from the real gut metagenomes using different combinations varied from 1178 to 4724. The highest number of MAGs was obtained using CONCOCT, followed by MetaBAT2 and MaxBin2.0, respectively. Our results demonstrated that CONCOCT performed the best, recovering the highest number of MAGs (n = 4724 ± 15.2) from metaSPAdes assemblies, followed by MEGAHIT (n = 4535 ± 12.4), and IDBA-UD (n = 4272 ± 11.5). Contrarily, MetaBAT2 produced a total of 2256 ± 14.4 MAGs with metaSPAdes assembly, followed by MEGAHIT (n = 2023 ± 11.9) and IDBA-UD (n = 1624 ± 10.0) assemblies. Conversely,

MaxBin2.0 yielded the lowest count, recovering a total of 1605 ± 10.5, 1508 ± 9.5, and 1178 ± 7.3 MAGs recovered from metaSPAdes, MEGAHIT, and IDBA-UD assemblies, respectively (Fig. 2, Supplementary Table 4).

Next, we deployed CheckM to estimate the quality of MAGs by assessing their completeness and contamination levels. For gut metagenomes, our analysis of genome completeness across the three binning tools revealed that MaxBin2 yielded the most complete MAGs with an average completeness of 71.1 ± 31.0 % using metaSPAdes, followed by IDBA-UD (69.4 ± 29.3 %) and MEGAHIT (67.5 ± 31.7 %) assemblies. Next in rank was MetaBAT2 which output the most complete MAGs with an average completeness ranging from 52 to 57 % with all three assemblers. In contrast, CONCOCT demonstrated the lowest genome completeness across all assemblers, with an average completeness level below 30 % (Table 1 and Supplementary Table 5). We then categorized the recovered MAGs as HQ, MQ, and LQ according to the criteria defined in the Methods. The results showed that MetaBAT2 excelled by recovering a substantially higher proportion of nearly complete genomes. In contrast, CONCOCT, despite recovering the highest number of MAGs, failed to produce a significant proportion of useable genomes. Our results revealed that metaSPAdes-MetaBAT2 recovered the highest fraction of HQ + MQ MAGs i.e. 52.6 %, followed by MEGAHIT-MetaBAT2 with 51.1 %. IDBA-UD-MetaBAT2 was ranked third for recovering a proportion of 48.3 % useable MAGs. This proportion drastically decreased to 17%–22 % useable MAGs when the assemblies were binned using CONCOCT. On the other hand, the lowest fraction of LQ MAGs was also recovered by MetaBAT2 with metaSPAdes, MEGAHIT, and IDBA-UD assemblies respectively (Fig. 3, Supplementary Table 5).

A similar trend was observed for skin metagenomes, with CONCOCT producing the highest number of MAGs, followed by MetaBAT2 (Supplementary Figs. 2A–2B). These genome binners performed comparatively well using MEGAHIT assemblies. In contrast, MaxBin2 recovered the lowest number. However, despite producing the least number of MAGs, MaxBin2 achieved the highest average completeness from IDBA-UD assemblies. In contrast, CONCOCT produced the least complete genomes. Regarding the recovery of useable MAGs, metaSPAdes-MetaBAT2 recovered the highest proportion, followed by metaSPAdes-MaxBin2.0. Whereas the fraction of useable MAGs dropped to 4%–6% with CONCOCT (Supplementary Fig. 2C and Supplementary Table 6).

For simulating metagenomes, mOTUs2 used 320 human gut MAGs in different relative abundances. Recovering MAGs from these simulated samples using the nine combinations yielded 316 species. Out of these, the highest number was recovered by metaSPAdes-MetaBAT i.e. 305, followed by metaSPAdes-CONCOCT (n = 300) and MEGAHIT-CONCOCT (n = 294) respectively (Supplementary Tables 7 and 8). Considering the completeness and contamination of the recovered MAGs determined by CheckM, metaSPAdes-MaxBin2.0 outperformed by producing the most complete bins with an average of 81.4 ± 20.2 %, followed by MEGAHIT-

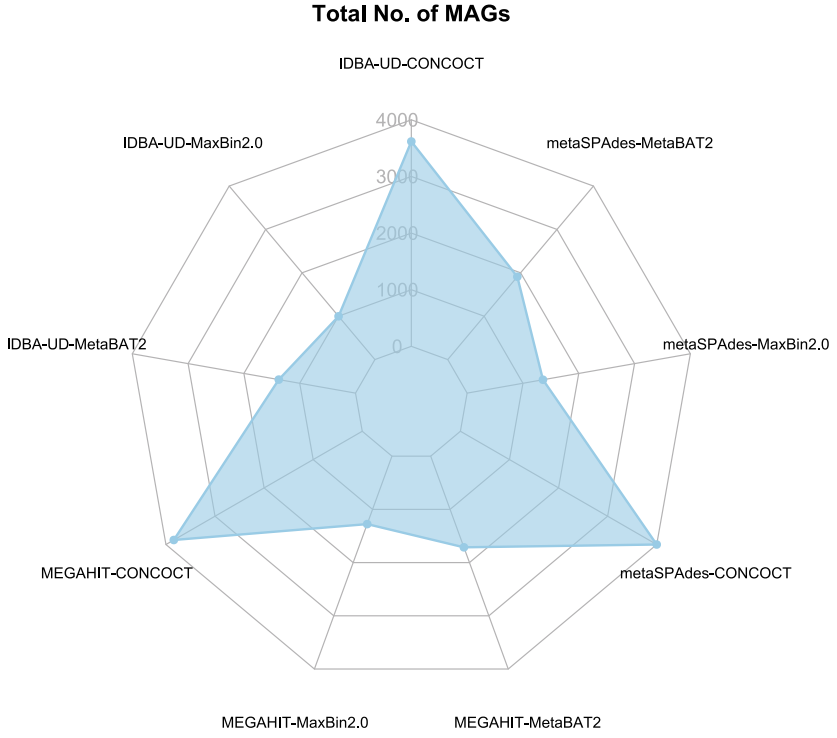


Fig. 2. Total number of MAGs recovered from real gut metagenomes across the nine combinations.

Table 1
Quality assessment of the MAGs recovered from the real gut metagenomes across the combinations.

No.	Combinations	MAGs Count	Average Completeness (%)	Average Contamination (%)
1	MEGAHIT-MaxBin2.0	59 ± 3.8	67.5 ± 31.7	9.3 ± 14.5
2	MEGAHIT-MetaBAT2	79 ± 4.4	55.1 ± 37.9	5.9 ± 23.6
3	MEGAHIT-CONCOCT	631 ± 18.0	25.2 ± 38.6	5.2 ± 23.5
4	metaSPAdes-MaxBin2.0	60 ± 4.2	71.1 ± 31.0	11.3 ± 18.0
5	metaSPAdes-MetaBAT2	83 ± 5.3	57.2 ± 37.6	5.8 ± 23.0
6	metaSPAdes-CONCOCT	622 ± 18.5	27.3 ± 39.7	0.6 ± 14.6
7	IDBA-UD-MaxBin2.0	49 ± 2.6	69.4 ± 29.4	11.9 ± 17.1
8	IDBA-UD-MetaBAT2	59 ± 3.4	52.2 ± 36.2	4.7 ± 17.7
9	IDBA-UD-CONCOCT	533 ± 18.0	21.8 ± 35.7	3.7 ± 17.1

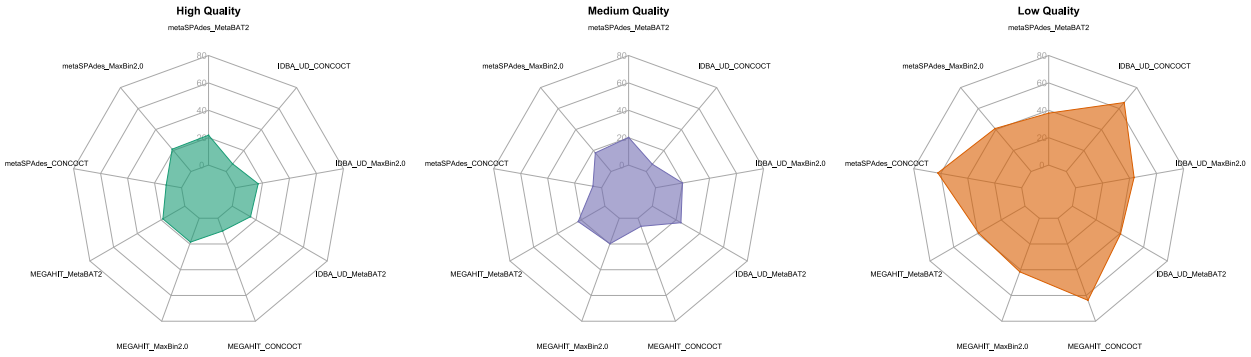


Fig. 3. Number of MAGs classified as high, medium, or low quality as recovered from the real gut metagenomes across the combinations.

MaxBin2.0 (79.8 ± 22.2 %) and MEGAHIT-MetaBAT2 (78.2 ± 25.6 %) (Supplementary Table 8). The purest bins were generated by metaSPAdes-MetaBAT2 with an average contamination rate of 2.9 ± 13.5 %, followed by MEGAHIT-MetaBAT2 (3.0 ± 13.4 %). Regarding the MAGs' quality, all the combinations recovered >70 % useable MAGs except for the IDBA-UD-MaxBin2.0 recovering only an average proportion of 66.7 %. The highest proportion was recovered by MEGAHIT-MetaBAT2 with an average of 82.5 %, followed by metaSPAdes-MetaBAT2 (80.5 %) and IDBA-UD-MetaBAT2 (77.0 %) respectively. Hence, MEGAHIT-MetaBAT2 can be considered the most efficient combination for simulated metagenomes since it presents a reasonable trade-off between the MAGs' completeness and contamination levels while producing the highest proportions of useable MAGs.

3.3. Taxonomic classification and determining the recovery of low-abundance species

Recovering from the real metagenomes across the nine combinations, we subjected a total of 23,725 gut and 2195 skin MAGs to GTDB-Tk for taxonomic classification (Fig. 4A). This process successfully classified 51.7 % (n = 12,281) gut and 21.5 % (n = 473) skin MAGs at the species level (Supplementary Table 9). Comparative analysis of the taxonomic classification of the MAGs recovered across the combinations is illustrated in Fig. 4B. The taxonomic abundances of the classified MAGs (species) were estimated as described in Methods (Supplementary Tables 10–11). These results revealed that the combinations collectively recovered 43 low-abundance species from the gut and 5 from the skin metagenomes. Evaluating the effectiveness of the combinations in recovering these low-abundance species, our results demonstrated that the power of different metagenomic tools diminishes when it comes to the recovery of low-abundance species (Fig. 4C). This finding was in line with the previously published findings [39]. We found that metaSPAdes-MetaBAT2 was able to recover the highest number of low abundance species i.e. 41 (95 %) from gut metagenomes, followed by MEGAHIT-CONCOCT and metaSPAdes-CONCOCT each recovering the second highest number i.e. 38 (88 %). Whereas IDBA-UD-MaxBin2.0 recovered the least number of low-abundance species i.e. 18 (41 %) (Fig. 5A).

Next, we sought the assessment of the potential of these combinations in recovering useable genomes for low-abundance species. To this end, we categorized species into three hypothetical levels based on their relative abundances: very low (0.05 %–0.49 %), low (0.5 %–0.99 %), and high (≥1 %). Ultimately, a combination striking a balance between the maximum proportion of useable MAGs and the highest number of recovered species in each category was considered the best. Assessing the quality of the MAGs corresponding to species at very low abundance, we found MEGAHIT-MaxBin2.0 outperforming by recovering the highest fraction i.e. 61.4 % of useable MAGs from real gut metagenomes followed by metaSPAdes-MaxBin2.0 with 58.9 % MAGs. For MAGs corresponding to low abundance species, metaSPAdes-CONCOCT appeared to be the most suitable option with a total of 69.0 % useable MAGs, followed by MEGAHIT-CONCOCT with a proportion of 67.3 %. Whereas all the designed combinations performed well for high-abundance species by recovering >55 % of useable MAGs. However, metaSPAdes-CONCOCT achieved the highest proportion of 62.6 % useable MAGs, followed by MEGAHIT-CONCOT with a fraction of 61.8 % (Fig. 5B and Supplementary Table 11).

For skin metagenomes, MEGAHIT-MetaBAT2 recovered the highest number i.e. 3 (75 %) of low abundance species, followed by

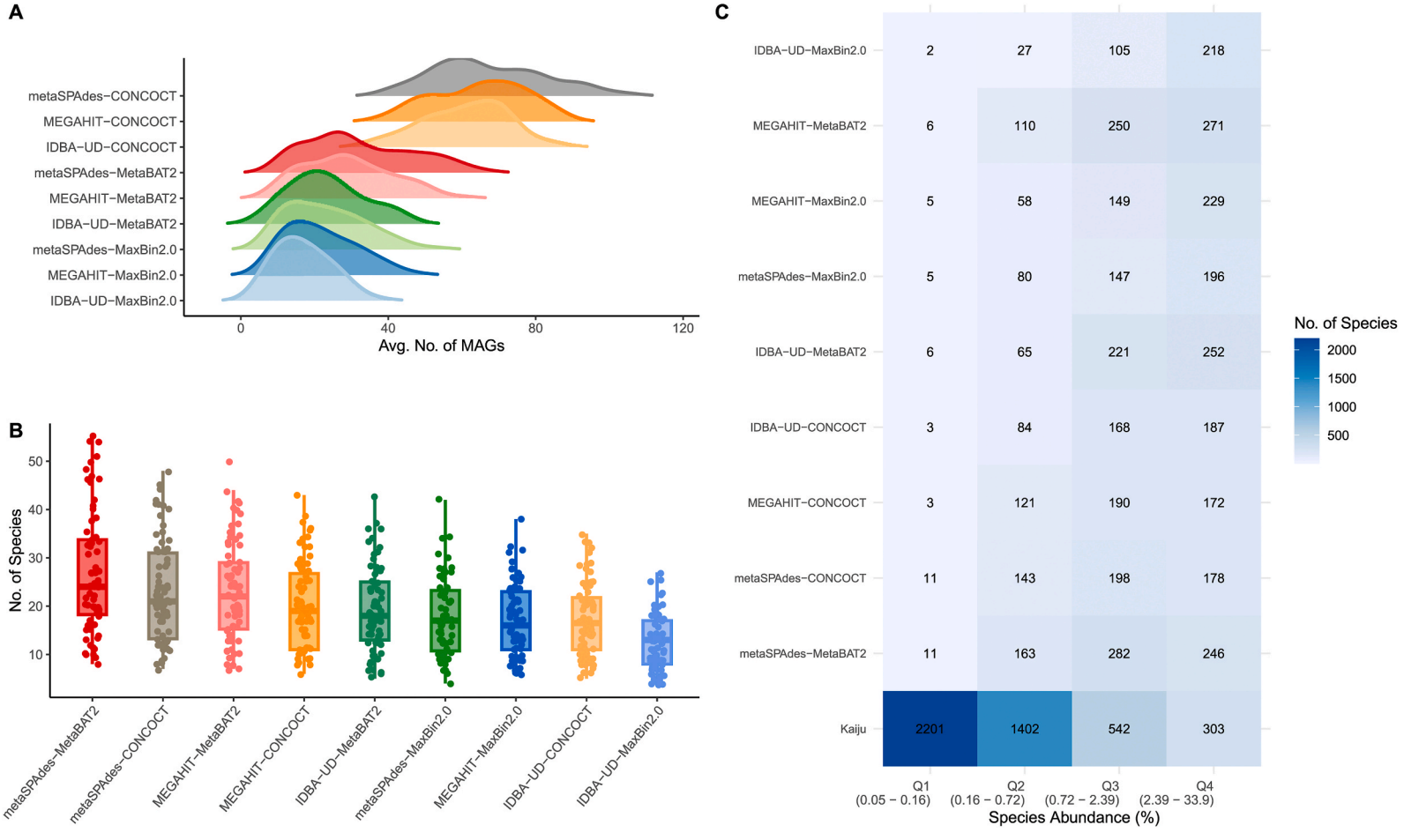


Fig. 4. Taxonomic classification of the nine sets of MAGs recovered from the real gut metagenomes using combinations. (A) Average number of recovered MAGs to be classified against each combination. (B) Number of species identified corresponding to the MAGs recovered from different combinations. (C) Number of classified species recovered by each combination at varying abundances is presented here as quartiles. (*species were divided into four quartiles based on their relative abundance and the number of species recovered by each combination was calculated within each quartile).

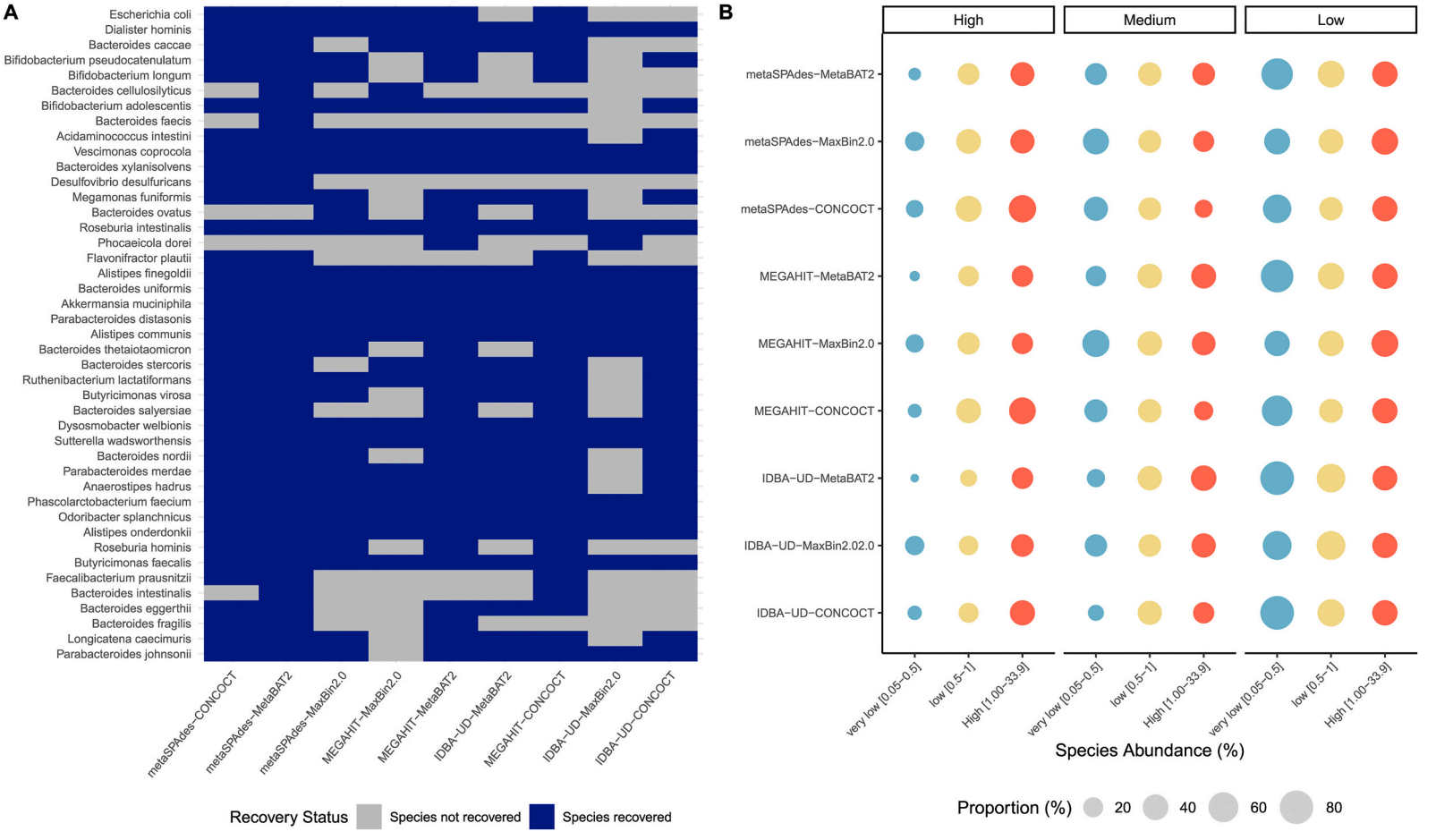


Fig. 5. Performance assessment of the combinations for recovering low-abundance species from the real gut metagenomes. (A) Species recovered by each combination. (B) Representation of quality of MAGs recovered by combinations corresponding to the low abundance species.

metaSPAdes-MetaBAT2 and IDBA-UD-MetaBAT2, each recovering 2 species (50 %). However, the remaining combinations recovered only 1 species (25 %) each (Supplementary Fig. 3A). Regarding genome quality, our findings revealed that all the combinations performed similarly and could not recover any HQ MAG corresponding to the low-abundance species (Supplementary Fig. 3B). Meanwhile, for the high abundance category, the highest number of HQ MAGs was recovered by metaSPAdes-CONCOCT (Supplementary Table 12).

Assessment of the combinations for the recovery of 58 low-abundance species from simulated metagenomes once again highlighted the superior performance of metaSPAdes-MetaBAT2 which recovered 51 ± 1.0 species. This was closely followed by MEGAHIT-MetaBAT2 (n = 44 ± 2.5) and metaSPAdes-CONCOCT (n = 43 ± 4.9) species. In contrast, IDBA-UD-MaxBin2.0 recovered the fewest low-abundance species (n = 17 ± 4.0) (Fig. 6A and Supplementary Tables 13–14). Interestingly, all combinations recovered species with relative abundance as low as 0.017 %. The evaluation of the genome quality of the recovered low-abundance species demonstrated that metaSPAdes-CONCOCT produced the highest proportion of useable MAGs for very low-abundance species achieving 65.7 % followed by metaSPAdes-MaxBin2.0 (65.5 %) and IDBA-UD-MetaBAT2 with 62.5 % (Fig. 6B and Supplementary Table 14). For low abundance levels, metaSPAdes-MetaBAT2 outperformed with 92.0 % of useable MAGs followed by metaSPAdes-MaxBin2.0 (80.0 %) (Fig. 6B). The least effective combinations for recovering very low and low abundance species were found to be IDBA-UD-CONCOCT and IDBA-MaxBin2.0.

3.4. *Evaluating the combinations for their potential for strain-resolved MAG recovery*

We comprehensively evaluated the combinations to assess their ability to recover strains from species with varying abundances within metagenomes of different complexities. For real metagenomes, our results demonstrated that the utilized assembler-binner combinations successfully recovered strain-resolved MAGs for species at as low as 0.06 % and as high as 28.7 % abundance within the gut metagenomes. In terms of the number of strains recovered, MEGAHIT-MetaBAT2 outperformed other combinations with a total of 244 strains of 54 gut species, followed by metaSPAdes-MetaBAT2 with 214 strains of 42 species and IDBA-UD-MetaBAT2 with 152 strains of 35 species. In contrast, metaSPAdes-CONCOCT and IDBA-UD-CONCOCT recovered only 4 strains (each) from 2 to 3 species respectively (Fig. 7A–B and Supplementary Tables 15–16). MEGAHIT-MetaBAT2 and metaSPAdes-MetaBAT2 recovered the highest proportion of useable MAGs (Fig. 7C and Supplementary Table 16). In contrast, none of the strains were recovered for species present at low abundance in the skin metagenomes. However, the greatest number of strains were recovered by MEGAHIT-MetaBAT2, followed by IDBA-UD-MetaBAT2. While MEGAHIT-MaxBin2.0 and IDBA-UD-CONCOCT were found least effective for recovering the fewest strains (Supplementary Figs. 4A–4C and Supplementary Table 17).

For the simulated dataset, with 10 strains for 5 species, the assembler-binner combinations collectively recovered 12 strains for 4 species (Supplementary Table 18). The MEGAHIT-MetaBAT2 again performed best while recovering the maximum number of strains i.

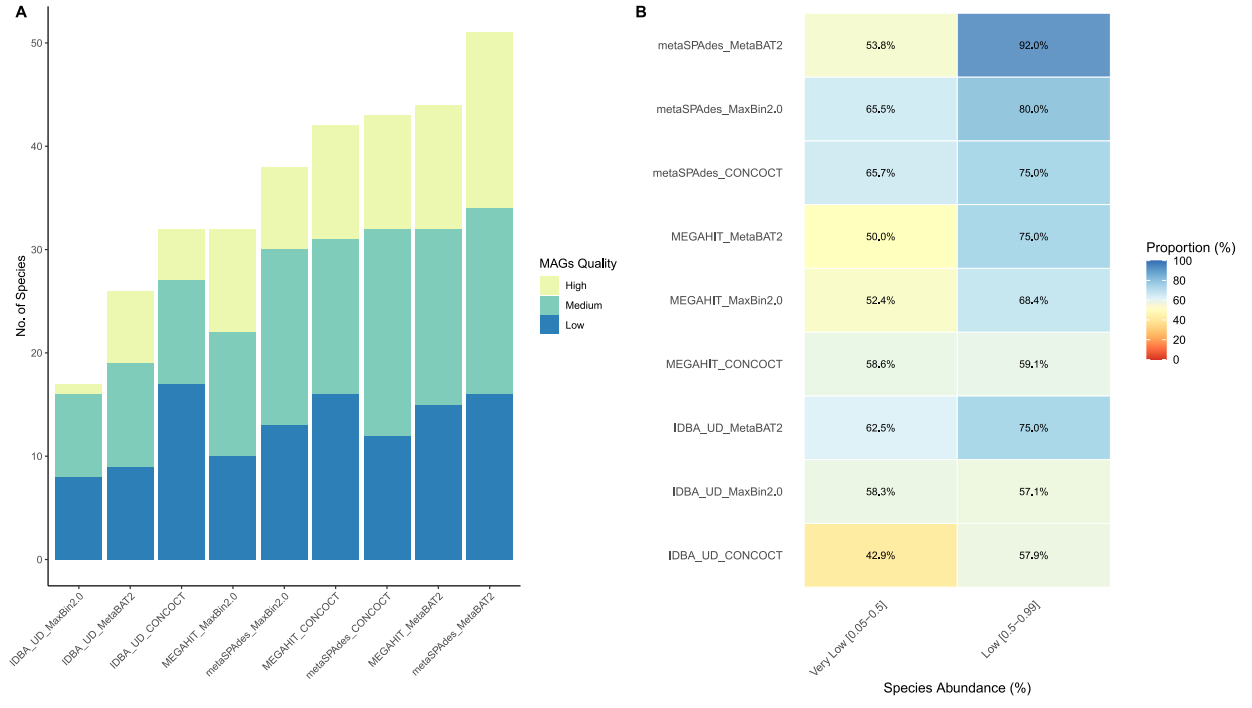


Fig. 6. Evaluation of the combinations for recovering low-abundance species from the simulated gut metagenomes. (A) Low abundance-species recovered by each combination. (B) Proportion of useable MAGs recovered across combinations corresponding to the low abundance species.

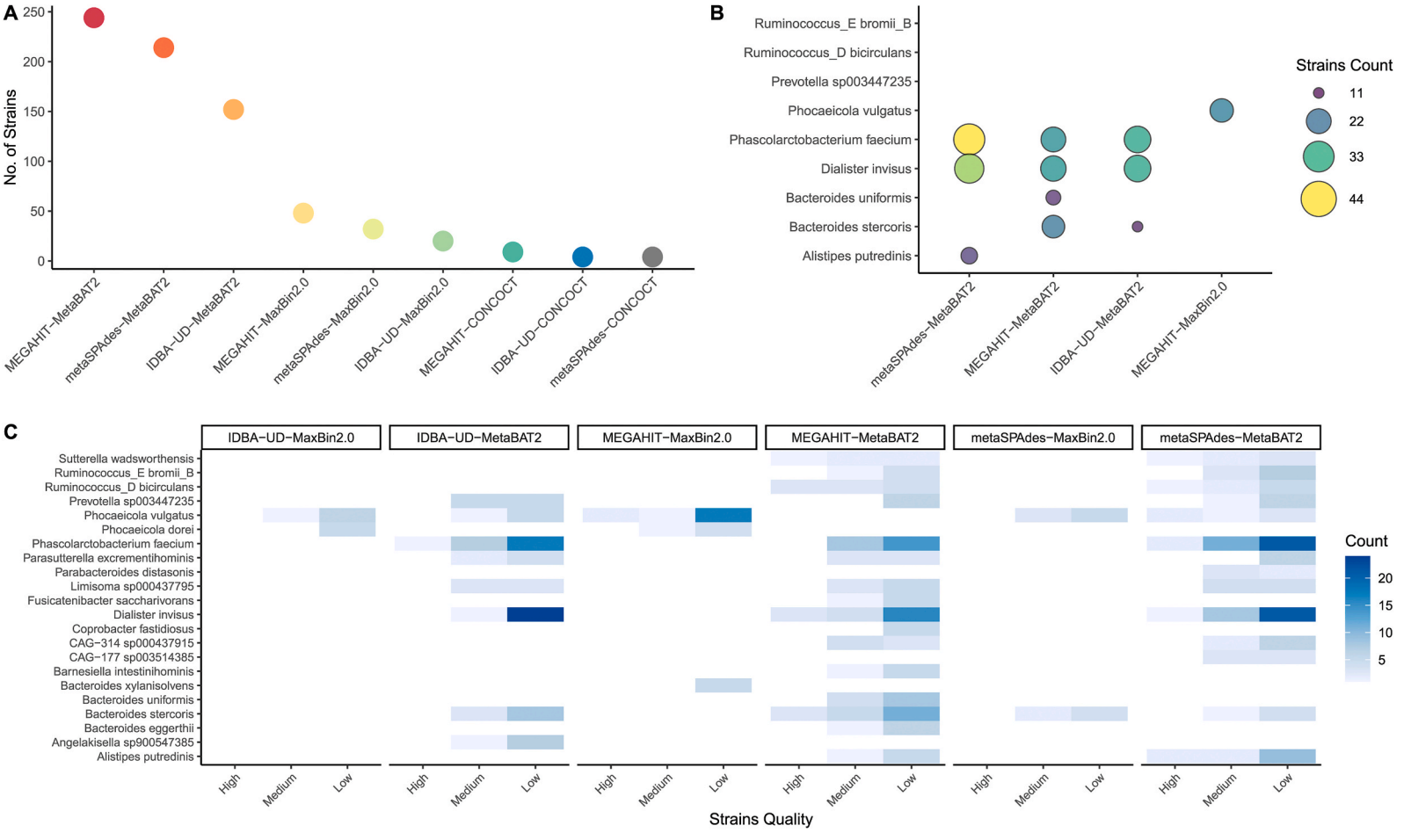


Fig. 7. Performance comparison of combinations for recovering strain-resolved MAGs from the real gut metagenomes. (A) Number of strains recovered by each combination. (B) Species diversity captured by each combination to recover >10 strains. (C) Quality assessment of MAGs recovered for the strains by each combination.

e. 6 from 3 species, followed by metaSPAdes-MetaBAT2 which recovered 4 strains from 2 species. Notably, MEGAHIT-MetaBAT2 successfully recovered these strains from species with a relative abundance as low as 0.002. In contrast, IDBA-UD-MetaBAT2 recovered the fewest strains, recovering only 2 strains from a single species at a relative abundance of 0.03. All other combinations remained unsuccessful in recovering any strains from the simulated dataset.

Collectively, these results demonstrated the adaptability of MetaBAT2 in recovering strain-resolved MAGs from the metagenomes of variable complexity. Furthermore, they highlight the ability of MEGAHIT-MetaBAT2 to recover the highest number of strain-resolved MAGs for a wide range of gut species, hence capturing maximum diversity.

4. Discussion

A comprehensive evaluation of metagenomic tools is crucial for optimizing data processing approaches and efficient interpretation of metagenomic data. *De novo* assembly is a key step for metagenomics data analysis and is adopted to recover MAGs and taxon bins from the metagenomes [40]. Many factors including uneven species abundances, uneven sequence coverage, high strain diversity, low recovery rates for taxa, and sequencing errors challenge metagenome assembly [41,42]. To address these challenges, many metagenomic assemblers and genome binners have been developed and employed in genome-resolved metagenomic studies. Previously, many studies have been conducted to evaluate the potential of these *de novo* assemblers and binning tools (independently) to produce high-quality metagenome assemblies and recover nearly complete MAGs from the real and simulated datasets [15,18,43–46]. These studies highlighted the variable performance of different assemblers and genome binners in achieving the same objectives, emphasizing that the choice of tools can profoundly impact results.

In this study, we hypothesized that the effectiveness of MAG recovery varies depending on the combination of assembler and binning tool used. Specifically, certain combinations are more effective than others in recovering low-abundance species and strain-resolved MAGs. By testing different assembler-binner combinations, we aimed to identify the pairs that optimize these recovery metrics, hypothesizing that specific combinations outperform others in achieving targeted objectives.

Subjecting the assemblies produced across the assemblers to the selected genome binning tools, revealed a profound impact of genome binner selection on the recovery rate of useable MAGs. For instance, MetaBAT2 demonstrated a significantly higher recovery rate of useable MAGs, around 50 % for both real (gut and skin) and simulated metagenomes. In contrast, CONCOCT achieved the lowest recovery rate of 20 % for the gut, 5 % for the skin, and approximately 66 % for simulated gut metagenomes. Despite CONCOCT being the only binner to bin all contigs and produce the highest number of MAGs, its recovery rate for nearly complete genomes remained low. Contrarily, MetaBAT2 recovered a lesser number of MAGs than CONCOCT, however, it should be noted that a good estimation of the number of bins is not sufficient to determine the efficiency of a genome binner [19]. On the other hand, though MaxBin2.0 recovered MAGs with the highest average completeness, this higher average was a trade-off with the ~30 % lesser number of MAGs produced as compared to the MetaBAT2. As a result, metaSPAdes-MetaBAT2 was declared as the most effective combination for recovering a higher fraction of nearly complete and draft MAGs, outperforming other combinations for both real and simulated metagenomes.

Assessment of the combinations for the recovery of low-abundance species revealed a significant trade-off between the recovery rate of low-abundance species and the proportion of recovered useable MAGs, concluding that the choice of the tools can greatly impact the results. It also demonstrated that the effectiveness of a combination is influenced by the nature and variability of the dataset being analyzed [47]. For instance, there was a consensus for metaSPAdes-MetaBAT2 recovering the highest number of low-abundance species from both real and simulated metagenomes. Whereas in terms of the MAGs quality, the highest proportion of useable MAGs for low and very low-abundance species within the real gut metagenomes was recovered by metaSPAdes-CONCOCT and MEGAHIT-MaxBin2.0 respectively. In contrast, metaSPAdes-MetaBAT2 and metaSPAdes-CONCOCT exhibited higher rates of useable MAGs for simulated gut metagenomes respectively. The least trade-off was presented by metaSPAdes-CONCOCT as it balances the recovery of the higher number of low-abundance species with a greater fraction of useable MAGs, hence, we declared it as the best approach for the recovery of such species from real metagenomes. Whereas metaSPAdes-MetaBAT2 can be the most suitable option for simulated metagenomes. This divergence suggests that CONCOCT can better handle heterogeneous coverage making it a suitable choice for the real metagenomic data. In contrast, MetaBAT2 excels in situations with uniform coverage and low complexity, common in simulations. Our results however remained limited in the identification of the best combination for the recovery of low-abundance species from skin metagenomes as the evaluated combinations performed similarly by exhibiting the same recovery rate of the useable MAGs.

Microbial strains play a crucial role in adaptation to the host, as they represent versions of a species that have evolved specific traits to thrive in specific conditions within the host environment. This adaptation allows strains to effectively interact with the host, influencing health and disease outcomes [48]. However, the implications of this variation remain underexplored due to the challenges in recovering strain-resolved MAGs from metagenomes [49]. While the whole genome shotgun approach facilitates the less perturbed views of strain diversity, it requires specialized computational tools to disentangle the strains from a metagenomic mix. Though various dedicated tools such as StrainGE [49] latent strain analysis (LSA) [50], and StrainEST [51] can successfully determine the strains present in metagenomes, they are limited to classification purposes only. Whereas the ability of typical MAG recovery approaches to distinguish genomes at the strain level is often questioned [52]. A comprehensive evaluation of the assembler-binner combinations to recover strains from the real and simulated dataset showed that MetaBAT2 outperformed other options in strain recovery across the three assemblers. However, MEGAHIT-MetaBAT2 captured the highest strain diversity and recovered a higher fraction of useable MAGs. Thus, we concluded that MEGAHIT-MetaBAT2 can be a good choice for the recovery of strain-resolved MAGs.

Interestingly, species abundance did not influence the number of recovered strains or the recovery rate of HQ or MQ MAGs corresponding to strains. For example, in gut metagenomes, a low-abundance species like *Phascolarctobacterium faecium* (0.3 %) yielded many strains and a maximum number of useable genomes. Conversely, another low-abundance species, *Sutterella wadsworthensis* (0.9 %), had the fewest HQ strains recovered. Whereas, for a highly abundant species, *Bacteroides stercoris* (28.7 %), combinations recovered the fewest MQ MAGs. Similarly, in skin metagenomes, *Lawsonella clevelandensis,* present at a moderate abundance (6.17 %), showed the highest strain and MQ MAGs recovery, whereas fewer strains were recovered for *Dolosigranulum pigrum*, which is prevalent at 49.6 %. In contrast, the least number of useable MAGs was recovered for *Neisseria* sp*000186165*, which had a relative abundance of less than 0.1 %.

Finally, we assessed the performance of different assemblers in combination with collectively refined bins using the bin-refinement module of MetaWRAP. This aimed to evaluate the impact of the integrated binning approach on the recovery of low-abundance species and microbial strains. As expected, these combinations yielded fewer low-abundance species and failed to recover any MAGs for species with an abundance of $<$0.12 % (Supplementary Table 19). This outcome can be attributed to the fact that the MetaWRAP filters out bins smaller than 50 kb in size [53], and low-abundance species are represented by fewer reads in metagenomic data, resulting in smaller genome bins. In contrast, the number of recovered microbial strains significantly declined across the assemblers (Supplementary Table 20). This reduction is explainable as the integrated binning approach prioritizes merging and refining bins to improve overall genome quality metrics. However, this process often merges closely related strains into composite bins and may not retain the strain-level diversity needed for the strain-resolved MAG recovery [53,54]. Hence, despite being beneficial for general bin refinement, the integrated approach is unsuitable for strain-level resolution.

Collectively, our findings highlighted a few important points to be considered. Firstly, tools present trade-off effects, and thus there is no one-size-fits-all solution. Depending upon the metagenomic dataset and objectives of the study, different tools may perform optimally. Secondly, tools can have complementary effects, where a tool may perform optimally in combination with one but not another. This study provides one such evidence. While we identified some of the best approaches using the existing tools, it is acknowledged that these tools are not yet performing at their best, and there is significant room for improvement. However, the demonstrated potential to recover low-abundance species and strain-resolved MAGs suggests that these combinations can be further optimized to address many current challenges.

Although this study contributes by evaluating the complementary effects of different assemblers and binning tools, providing insights for selecting the right tools for the scientific community, we acknowledge certain limitations. Firstly, the study tested a limited number of assemblers and genome-binning tools, leaving many other potential tools and combinations unexplored. Secondly, differences in underlying databases for taxonomic classification can account for performance differences between combinations for simulated metagenomes. For instance, the gold standard taxonomic profile of the mOTUs2 dataset is based on the mOTUs2 database whereas, we utilized GTDB-Tk to classify the genomes influencing the number of classified species. Thirdly, we exclusively included DBG assemblers due to their wide acceptability in metagenomic analyses, but it would be valuable to test the performance of Overlap Layout Consensus assemblers too in combination with available binning tools.

5. Conclusions

Computational methods play a crucial role in metagenomic studies. Evaluating existing metagenomic analysis tools is essential to understand their advantages and limitations, thus aiding the optimization of metagenomic approaches. This study underscores the significant impact of various assembler-binner combinations on the recovery of MAGs for low-abundance species and strain-resolved MAGs. Despite inherent trade-offs associated with each tool, we have identified combinations that offer a balance of performance with minimal trade-offs. Specifically, the metaSPAdes-CONCOCT and metaSPAdes-MetaBAT2 combination was shown to be particularly effective for recovering low-abundance species from gut metagenomes, whereas the MEGAHIT-MetaBAT2 combination outperformed in recovering strain-resolved MAGs. Our findings are intended to guide the scientific community in selecting the most suitable combination for their specific objectives.

List of Abbreviations

ANI	Average Nucleotide Identity
DBG	de Bruijn Graph
HQ	High Quality
LQ	Low Quality
MAG	Metagenome Assembled Genome
MIMAG	Minimum Information about Metagenome Assembled Genomes
MQ	Medium Quality

CRediT authorship contribution statement

Hajra Qayyum: Writing – original draft, Visualization, Formal analysis. **Muhammad Sarfraz Talib:** Formal analysis. **Amjad Ali:** Writing – review & editing, Writing – original draft, Supervision. **Masood Ur Rehman Kayani:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Conceptualization.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors have read and approved the manuscript.

Availability of data and materials

The real metagenomes used for the assessment of combinations were retrieved from the Human Microbiome Project (III) (NCBI BioProject: PRJNA275349) in June 2023; the corresponding accession numbers are provided in Supplementary Table 1. For simulating metagenomes, the gold standard profiles of the mOTUs2 dataset were downloaded from Zenodo (https://zenodo.org/records/1473645) in Oct 2024, and reads simulated using InSilicoSeq 2.0 for this study can be downloaded from https://zenodo.org/uploads/14354209.

Funding

No specific funding was acquired for this study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The primary author of this study is grateful to the Higher Education Commission (HEC) of Pakistan for sponsoring her Ph.D. studies under the Faculty Development Program (FDP) for Pakistani Universities- Batch III. The authors also acknowledge Mr. Haseeb Manzoor, Research Assistant, Metagenomics Discovery Lab, SINES, NUST, for his assistance in designing the graphics for this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2025.e41938.

References

[1] J. Puschhof, E. Elinav, Human microbiome research: growing pains and future promises, PLoS Biol. 21 (2023), https://doi.org/10.1371/journal.pbio.3002053.
[2] E. Dekaboruah, M.V. Suryavanshi, D. Chettri, A.K. Verma, Human microbiome: an academic update on human body site-specific surveillance and its possible role, Arch. Microbiol. 202 (2020) 2147, https://doi.org/10.1007/S00203-020-01931-X.
[3] M. Dwivedi, F. Shah, P.S. Giri, Role of human gut microbiome in health and disease, Microbiome-Host Interactions (2021) 101–112, https://doi.org/10.1201/9781003037521-11.
[4] F. Maguire, B. Jia, K.L. Gray, W.Y.V. Lau, R.G. Beiko, F.S.L. Brinkman, Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic islands, Microb. Genom. 6 (2020) 1–12, https://doi.org/10.1099/mgen.0.000436.
[5] M. Borderes, C. Gasc, E. Prestat, M. Galvão Ferrarini, S. Vinga, L. Boucinha, M.F. Sagot, A comprehensive evaluation of binning methods to recover human gut microbial species from a non-redundant reference gene catalog, NAR Genom Bioinform 3 (2021), https://doi.org/10.1093/nargab/lqab009.
[6] P. Katara, Recent Trends in "Computational Omics: Concepts and Methodology," n.d.
[7] B.D. Anderson, J.E. Bisanz, Challenges and opportunities of strain diversity in gut microbiome research, Front. Microbiol. 14 (2023), https://doi.org/10.3389/fmicb.2023.1117122.
[8] H. Jin, L. You, F. Zhao, S. Li, T. Ma, L.Y. Kwok, H. Xu, Z. Sun, Hybrid, ultra-deep metagenomic sequencing enables genomic and functional characterization of low-abundance species in the human gut microbiome, Gut Microb. 14 (2022), https://doi.org/10.1080/19490976.2021.2021790.
[9] R. Bharti, D.G. Grimm, Current challenges and best-practice protocols for microbiome analysis, Brief Bioinform 22 (2021) 178–193, https://doi.org/10.1093/bib/bbz155.
[10] A.R. Khan, M.T. Pervez, M.E. Babar, N. Naveed, M. Shoaib, A comprehensive study of de novo genome assemblers: current challenges and future prospective, Evol. Bioinf. Online 14 (2018), https://doi.org/10.1177/1176934318758650.
[11] N.R. Edwin, A.H. Fitzpatrick, F. Brennan, F. Abram, O. O'Sullivan, An in-depth evaluation of metagenomic classifiers for soil microbiomes, Environ Microbiome 19 (2024) 1–17, https://doi.org/10.1186/s40793-024-00561-w.
[12] V. Mallawaarachchi, A. Wickramarachchi, H. Xue, B. Papudeshi, S.R. Grigson, G. Bouras, R.E. Prahl, A. Kaphle, A. Verich, B. Talamantes-Becerra, E.A. Dinsdale, R.A. Edwards, Solving genomic puzzles: computational methods for metagenomic binning, Brief Bioinform 25 (2024), https://doi.org/10.1093/BIB/BBAE372.
[13] G. Goussarov, M. Mysara, I. Cleenwerck, J. Claesen, N. Leys, P. Vandamme, R. Van Houdt, Benchmarking short-, long-and hybrid-read assemblers for metagenome sequencing of complex microbial communities, Microbiology (United Kingdom) 170 (2024) 001469, https://doi.org/10.1099/MIC.0.001469/CITE/REFWORKS.
[14] Z. Zhang, C. Yang, W.P. Veldsman, X. Fang, L. Zhang, Benchmarking genome assembly methods on metagenomic sequencing data, Brief Bioinform 24 (2023) 1–17, https://doi.org/10.1093/BIB/BBAD087.

[15] E. Forouzan, P. Shariati, M.S. Mousavi Maleki, A.A. Karkhane, B. Yakhchali, Practical evaluation of 11 de novo assemblers in metagenome assembly, J. Microbiol. Methods 151 (2018) 99–105, https://doi.org/10.1016/j.mimet.2018.06.007.
[16] S. Yorki, T. Shea, C.A. Cuomo, B.J. Walker, R.C. LaRocque, A.L. Manson, A.M. Earl, C.J. Worby, Comparison of long- and short-read metagenomic assembly for low-abundance species and resistance genes, Brief Bioinform 24 (2023), https://doi.org/10.1093/bib/bbad050.
[17] F. Meyer, A. Fritz, Z.L. Deng, D. Koslicki, T.R. Lesker, A. Gurevich, G. Robertson, M. Alser, D. Antipov, F. Beghini, D. Bertrand, J.J. Brito, C.T. Brown, J. Buchmann, A. Buluç, B. Chen, R. Chikhi, P.T.L.C. Clausen, A. Cristian, P.W. Dabrowski, A.E. Darling, R. Egan, E. Eskin, E. Georganas, E. Goltsman, M.A. Gray, L.H. Hansen, S. Hofmeyr, P. Huang, L. Irber, H. Jia, T.S. Jørgensen, S.D. Kieser, T. Klemetsen, A. Kola, M. Kolmogorov, A. Korobeynikov, J. Kwan, N. LaPierre, C. Lemaitre, C. Li, A. Limasset, F. Malcher-Miranda, S. Mangul, V.R. Marcelino, C. Marchet, P. Marijon, D. Meleshko, D.R. Mende, A. Milanese, N. Nagarajan, J. Nissen, S. Nurk, L. Oliker, L. Paoli, P. Peterlongo, V.C. Piro, J.S. Porter, S. Rasmussen, E.R. Rees, K. Reinert, B. Renard, E.M. Robertsen, G.L. Rosen, H. J. Ruscheweyh, V. Sarwal, N. Segata, E. Seiler, L. Shi, F. Sun, S. Sunagawa, S.J. Sørensen, A. Thomas, C. Tong, M. Trajkovski, J. Tremblay, G. Uritskiy, R. Vicedomini, Z. Wang, Z. Wang, Z. Wang, A. Warren, N.P. Willassen, K. Yelick, R. You, G. Zeller, Z. Zhao, S. Zhu, J. Zhu, R. Garrido-Oter, P. Gastmeier, S. Hacquard, S. Häußler, A. Khaledi, F. Maechler, F. Mesny, S. Radutoiu, P. Schulze-Lefert, N. Smit, T. Strowig, A. Bremges, A. Sczyrba, A.C. McHardy, Critical assessment of metagenome interpretation: the second round of challenges, Nat. Methods 19 (2022) 429–440, https://doi.org/10.1038/s41592-022-01431-4.
[18] Y. Yue, H. Huang, Z. Qi, H.M. Dou, X.Y. Liu, T.F. Han, Y. Chen, X.J. Song, Y.H. Zhang, J. Tu, Evaluating metagenomics tools for genome binning with real metagenomic datasets and CAMI datasets, BMC Bioinf. 21 (2020), https://doi.org/10.1186/s12859-020-03667-3.
[19] M. Borderes, C. Gasc, E. Prestat, M. Galvão Ferrarini, S. Vinga, L. Boucinha, M.F. Sagot, A comprehensive evaluation of binning methods to recover human gut microbial species from a non-redundant reference gene catalog, NAR Genom Bioinform 3 (2021), https://doi.org/10.1093/nargab/lqab009.
[20] D. Li, C.M. Liu, R. Luo, K. Sadakane, T.W. Lam, MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph, Bioinformatics 31 (2015) 1674–1676, https://doi.org/10.1093/BIOINFORMATICS/BTV033.
[21] S. Nurk, D. Meleshko, A. Korobeynikov, P.A. Pevzner, MetaSPAdes: a new versatile metagenomic assembler, Genome Res. 27 (2017) 824–834, https://doi.org/10.1101/gr.213959.116.
[22] Y. Peng, H.C.M. Leung, S.M. Yiu, F.Y.L. Chin, IDBA-UD : a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth 28 (2012) 1420–1428, https://doi.org/10.1093/bioinformatics/bts174.
[23] D.D. Kang, F. Li, E. Kirton, A. Thomas, R. Egan, H. An, Z. Wang, MetaBAT 2 : an adaptive binning algorithm for robust and ef fi cient genome reconstruction from metagenome assemblies (2019) 1–13, https://doi.org/10.7717/peerj.7359.
[24] Y. Wu, Y. Tang, S.G. Tringe, B.A. Simmons, S.W. Singer, MaxBin : an Automated Binning Method to Recover Individual Genomes from Metagenomes Using an Expectation-Maximization Algorithm, 2014, pp. 1–18.
[25] J. Alneberg, B.S. Bjarnason, I. De Bruijn, M. Schirmer, J. Quick, U.Z. Ijaz, L. Lahti, N.J. Loman, A.F. Andersson, C. Quince, Binning metagenomic contigs by coverage and composition, Nature Methods 2014 11 (11 11) (2014) 1144–1146, https://doi.org/10.1038/nmeth.3103.
[26] A. Milanese, D.R. Mende, L. Paoli, G. Salazar, H.J. Ruscheweyh, M. Cuenca, P. Hingamp, R. Alves, P.I. Costea, L.P. Coelho, T.S.B. Schmidt, A. Almeida, A. L. Mitchell, R.D. Finn, J. Huerta-Cepas, P. Bork, G. Zeller, S. Sunagawa, Microbial abundance, activity and population genomic profiling with mOTUs2, Nat. Commun. 10 (1 10) (2019) 1–11, https://doi.org/10.1038/s41467-019-08844-4, 2019.
[27] H. Gourlé, O. Karlsson-Lindsjö, J. Hayer, E. Bongcam-Rudloff, Simulating Illumina metagenomic data with InSilicoSeq, Bioinformatics 35 (2019) 521–522, https://doi.org/10.1093/BIOINFORMATICS/BTY630.
[28] S. Chen, Y. Zhou, Y. Chen, J. Gu, Fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics 34 (2018) i884–i890, https://doi.org/10.1093/bioinformatics/bty560.
[29] S.M. Ross, Nonparametric hypotheses tests, Introductory Statistics (2010) 647–697, https://doi.org/10.1016/B978-0-12-374388-6.00014-4.
[30] Z. Wang, Y. Wang, J.A. Fuhrman, F. Sun, S. Zhu, Assessment of metagenomic assemblers based on hybrid reads of real and simulated metagenomic sequences, Brief Bioinform 21 (2020) 777–790, https://doi.org/10.1093/bib/bbz025.
[31] X. Liao, M. Li, Y. Zou, F.X. Wu, Yi-Pan, J. Wang, Current challenges and solutions of de novo assembly, Quantitative Biology 7 (2019) 90–109, https://doi.org/10.1007/S40484-019-0166-9.
[32] F.P. Breitwieser, J. Lu, S.L. Salzberg, A Review of Methods and Databases for Metagenomic Classification and Assembly, vol. 20, 2019, pp. 1125–1139, https://doi.org/10.1093/bib/bbx120.
[33] D.H. Parks, M. Imelfort, C.T. Skennerton, P. Hugenholtz, G.W. Tyson, CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes, Genome Res. 25 (2015) 1043, https://doi.org/10.1101/GR.186072.114.
[34] R.M. Bowers, N.C. Kyrpides, R. Stepanauskas, M. Harmon-smith, D. Doud, J. Jarett, A.R. Rivers, E.A. Eloe-fadrosh, S.G. Tringe, N.N. Ivanova, A. Copeland, A. Clum, E.D. Becraft, R.R. Malmstrom, B. Birren, M. Podar, P. Bork, G.M. Weinstock, G.M. Garrity, J.A. Dodsworth, S. Yooseph, G. Sutton, F.O. Glöckner, J. A. Gilbert, W.C. Nelson, S.J. Hallam, perspective Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea 35 (2017), https://doi.org/10.1038/nbt.3893.
[35] P. Menzel, K.L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju, Nat. Commun. 7 (2016), https://doi.org/10.1038/ncomms11257.
[36] P.A. Chaumeil, A.J. Mussig, P. Hugenholtz, D.H. Parks, GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database, Bioinformatics 36 (2020) 1925–1927, https://doi.org/10.1093/BIOINFORMATICS/BTZ848.
[37] C. Jain, L.M. Rodriguez-R, A.M. Phillippy, K.T. Konstantinidis, S. Aluru, High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries, Nat. Commun. 9 (1 9) (2018) 1–8, https://doi.org/10.1038/s41467-018-07641-9, 2018.
[38] L.M. Rodriguez-R, R.E. Conrad, T. Viver, D.J. Feistel, B.G. Lindner, S.N. Venter, L.H. Orellana, R. Amann, R. Rossello-Mora, K.T. Konstantinidis, An ANI gap within bacterial species that advances the definitions of intra-species units, mBio 15 (2024), https://doi.org/10.1128/MBIO.02696-23/SUPPL_FILE/MBIO.02696-23-S0001.PDF.
[39] H. Jin, L. You, F. Zhao, S. Li, T. Ma, L.Y. Kwok, H. Xu, Z. Sun, Hybrid, ultra-deep metagenomic sequencing enables genomic and functional characterization of low-abundance species in the human gut microbiome, Gut Microb. 14 (2022), https://doi.org/10.1080/19490976.2021.2021790.
[40] F. Meyer, A. Fritz, Z.L. Deng, D. Koslicki, T.R. Lesker, A. Gurevich, G. Robertson, M. Alser, D. Antipov, F. Beghini, D. Bertrand, J.J. Brito, C.T. Brown, J. Buchmann, A. Buluç, B. Chen, R. Chikhi, P.T.L.C. Clausen, A. Cristian, P.W. Dabrowski, A.E. Darling, R. Egan, E. Eskin, E. Georganas, E. Goltsman, M.A. Gray, L.H. Hansen, S. Hofmeyr, P. Huang, L. Irber, H. Jia, T.S. Jørgensen, S.D. Kieser, T. Klemetsen, A. Kola, M. Kolmogorov, A. Korobeynikov, J. Kwan, N. LaPierre, C. Lemaitre, C. Li, A. Limasset, F. Malcher-Miranda, S. Mangul, V.R. Marcelino, C. Marchet, P. Marijon, D. Meleshko, D.R. Mende, A. Milanese, N. Nagarajan, J. Nissen, S. Nurk, L. Oliker, L. Paoli, P. Peterlongo, V.C. Piro, J.S. Porter, S. Rasmussen, E.R. Rees, K. Reinert, B. Renard, E.M. Robertsen, G.L. Rosen, H. J. Ruscheweyh, V. Sarwal, N. Segata, E. Seiler, L. Shi, F. Sun, S. Sunagawa, S.J. Sørensen, A. Thomas, C. Tong, M. Trajkovski, J. Tremblay, G. Uritskiy, R. Vicedomini, Z. Wang, Z. Wang, Z. Wang, A. Warren, N.P. Willassen, K. Yelick, R. You, G. Zeller, Z. Zhao, S. Zhu, J. Zhu, R. Garrido-Oter, P. Gastmeier, S. Hacquard, S. Häußler, A. Khaledi, F. Maechler, F. Mesny, S. Radutoiu, P. Schulze-Lefert, N. Smit, T. Strowig, A. Bremges, A. Sczyrba, A.C. McHardy, Critical assessment of metagenome interpretation: the second round of challenges, Nat. Methods 19 (2022) 429–440, https://doi.org/10.1038/s41592-022-01431-4.
[41] Y. Zhou, M. Liu, J. Yang, Recovering metagenome-assembled genomes from shotgun metagenomic sequencing data: methods, applications, challenges, and opportunities, Microbiol. Res. 260 (2022) 127023, https://doi.org/10.1016/J.MICRES.2022.127023.
[42] X. Liao, M. Li, Y. Zou, F.X. Wu, Yi-Pan, J. Wang, Current challenges and solutions of de novo assembly, Quantitative Biology 7 (2019) 90–109, https://doi.org/10.1007/s40484-019-0166-9.
[43] F. Maguire, B. Jia, K.L. Gray, W.Y.V. Lau, R.G. Beiko, F.S.L. Brinkman, Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic islands, Microb. Genom. 6 (2020) 1–12, https://doi.org/10.1099/mgen.0.000436.
[44] E. Forouzan, P. Shariati, M.S. Mousavi Maleki, A.A. Karkhane, B. Yakhchali, Practical evaluation of 11 de novo assemblers in metagenome assembly, J. Microbiol. Methods 151 (2018) 99–105, https://doi.org/10.1016/j.mimet.2018.06.007.
[45] C. Yang, D. Chowdhury, Z. Zhang, W.K. Cheung, A. Lu, Z. Bian, L. Zhang, A review of computational tools for generating metagenome-assembled genomes from metagenomic sequencing data, Comput. Struct. Biotechnol. J. 19 (2021) 6301–6314, https://doi.org/10.1016/j.csbj.2021.11.028.

[46] Y. Yue, H. Huang, Z. Qi, H.M. Dou, X.Y. Liu, T.F. Han, Y. Chen, X.J. Song, Y.H. Zhang, J. Tu, Evaluating metagenomics tools for genome binning with real metagenomic datasets and CAMI datasets, BMC Bioinf. 21 (2020), https://doi.org/10.1186/s12859-020-03667-3.
[47] S.H. Ye, K.J. Siddle, D.J. Park, P.C. Sabeti, Benchmarking metagenomics tools for taxonomic classification, Cell 178 (2019) 779–794, https://doi.org/10.1016/j.cell.2019.07.010.
[48] Y. Yan, L.H. Nguyen, E.A. Franzosa, C. Huttenhower, Strain-level epidemiology of microbial communities and the human microbiome, Genome Med. 12 (2020), https://doi.org/10.1186/s13073-020-00765-y.
[49] L.R. van Dijk, B.J. Walker, T.J. Straub, C.J. Worby, A. Grote, H.L. Schreiber, C. Anyansi, A.J. Pickering, S.J. Hultgren, A.L. Manson, T. Abeel, A.M. Earl, StrainGE: a toolkit to track and characterize low-abundance strains in complex microbial communities, Genome Biol. 23 (2022), https://doi.org/10.1186/s13059-022-02630-0.
[50] B. Cleary, I.L. Brito, K. Huang, D. Gevers, T. Shea, S. Young, E.J. Alm, Detection of low-abundance bacterial strains in metagenomic datasets by eigengenome partitioning, Nat. Biotechnol. 33 (2015) 1053–1060, https://doi.org/10.1038/nbt.3329.
[51] D. Albanese, C. Donati, Strain profiling and epidemiology of bacterial species from metagenomic sequencing, (n.d.). https://doi.org/10.1038/s41467-017-02209-5.
[52] C. Anyansi, T.J. Straub, A.L. Manson, A.M. Earl, T. Abeel, Computational methods for strain-level microbial detection in colony and metagenome sequencing data, Front. Microbiol. 11 (2020) 538426, https://doi.org/10.3389/FMICB.2020.01925/BIBTEX.
[53] G. V Uritskiy, J. DiRuggiero, J. Taylor, MetaWRAP - a flexible pipeline for genome-resolved metagenomic data analysis. https://doi.org/10.1101/277442, 2018.
[54] K. Arikawa, K. Ide, M. Kogawa, T. Saeki, T. Yoda, T. Endoh, A. Matsuhashi, H. Takeyama, M. Hosokawa, Recovery of strain-resolved genomes from human microbiome through an integration framework of single-cell genomics and metagenomics, Microbiome 9 (2021), https://doi.org/10.1186/s40168-021-01152-4.