Highly accurate metagenome-assembled genomes from human gut microbiota using long-read assembly, binning, and consolidation methods

Daniel M. Portik[1], Xiaowen Feng[2,3], Gaetan Benoit[4], Daniel J. Nasko[1], Benjamin Auch[5], Samuel J. Bryson[5], Raul Cano[6], Martha Carlin[6], Anabelle Damerum[7], Brett Farthing[7], Jonas R. Grove[5], Moutusee Islam[5], Kyle W. Langford[5], Ivan Liachko[5], Kristopher Locken[7], Hayley Mangelson[5], Shuiquan Tang[7], Siyuan Zhang[1], Christopher Quince[8,9,10,11], and Jeremy E. Wilkinson[1]

1. PacBio, 1305 O'Brien Dr, Menlo Park, CA 94025 USA
2. Department of Data Sciences, Dana-Farber Cancer Institute, Boston, MA 02215 USA
3. Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115 USA
4. Sequence Bioinformatics, Department of Computational Biology, Institut Pasteur, Paris, France
5. Phase Genomics, 1617 8th Ave, Seattle, WA 98109 USA
6. The BioCollective LLC., 5650 Washington St., Denver, CO 80216 USA
7. Zymo Research Corporation, 17062 Murphy Ave, Irvine, CA 92614 USA
8. Organisms and Ecosystems, Earlham Institute, Norwich, UK
9. Gut Microbes and Health, Quadram Institute, Norwich, UK
10. School of Biological Sciences, University of East Anglia, Norwich, UK
11. Warwick Medical School, University of Warwick, Coventry, UK

ABSTRACT

Long-read metagenomic sequencing is a powerful approach for cataloging the microbial diversity present in complex microbiomes, including the human gut microbiome. We performed a deep-sequencing experiment using PacBio HiFi reads to obtain metagenome-assembled genomes (MAGs) from a pooled human gut microbiome. We performed long-read metagenome assembly using two methods (hifiasm-meta, metMDBG), used improved bioinformatic and proximity ligation binning strategies to cluster contigs and identify MAGs, and developed a novel framework to compare and consolidate MAGs (pb-MAG-mirror). We found proximity ligation binning yielded more MAGs than bioinformatic binning, but our novel comparison framework resulted in higher MAG yields than either binning strategy individually. In total, from 255 Gbp of total HiFi data we produced 595 total MAGs (including 175 high-quality MAGs) using hifiasm-meta, and 547 total MAGs (including 277 high-quality MAGs) with metaMDBG. Hifiasm-meta assembled almost twice as many strain-level MAGs as metaMDBG (246 vs. 156), but both assembly methods produced up to five strains for a species. Approximately 85% of the MAGs were assigned to known species, but we recovered >35 high-quality MAGs that represent uncultured diversity. Based on strict similarity scores, we found 125 MAGs were unequivocally shared across the assembly methods at the strain level, representing ~22% of the total MAGs recovered per method. Finally, we detected more total viral sequences in the metaMDBG assembly versus the hifiasm-meta assembly (~6,700 vs. ~4,500). Overall, we find the use of HiFi sequencing, improved metagenome assembly methods, and complementary binning strategies is highly effective for rapidly cataloging microbial genomes in complex microbiomes.

INTRODUCTION

The human gut microbiome contains a diversity of microbes that potentially impact health and disease (Lynch & Pederson 2016; Wang & Jia 2016; Duvallet et al. 2017; Gilbert et al. 2018). Despite considerable efforts to catalog the microbial diversity present in the human gut, large numbers of species and strains remain uncultured and undetected (Lagier et al. 2018; Nayfach et al. 2019; Almeida et al. 2019; Forster et al. 2019; Pasolli et al. 2019; Zou et al. 2019; Almeida et al. 2021). The pace of biodiversity discovery is limited using traditional isolation and culturing methods, but it can be greatly accelerated using metagenomic sequencing. Metagenome assembly is a powerful approach for reconstructing the genomic contents of species contained in

a microbiome sample (Tyson et al. 2004). Historically, the contigs produced by metagenome assemblies often represent small fragments of microbial genomes, and binning methods are used to group contigs into putative genomes. The genomes obtained through assembly and binning are commonly referred to as metagenome-assembled genomes (MAGs). Metagenome assembly based on short-read sequencing generally requires substantial effort to produce high-quality MAGs (Bowers et al. 2017; Chen et al. 2020). For example, repetitive genomic regions and interspecies genomic overlaps are difficult to resolve using short reads, and this generally leads to highly fragmented assemblies (Arumugam et al. 2019). Furthermore, binning can introduce major errors by grouping contigs belonging to different species or strains into MAGs (Chen et al. 2020). This can result in chimeric MAGs or even MAGs with human contamination, which can have severe negative effects on downstream analyses (see Gihawi et al. 2023).

Long-read sequencing can overcome many of the challenges associated with metagenome assembly (Kolmogorov et al. 2020; Feng et al. 2022; Albertsen 2023; Benoit et al. 2023; Agustinho et al. 2024). The most popular long-read sequencing platforms are those produced by Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT). PacBio HiFi sequencing produces highly accurate consensus reads (>Q20, median Q30) that are 10–20 kb in length (Wenger et al. 2019). Several studies have demonstrated that HiFi sequencing generally produces more total MAGs and higher quality MAGs than short-read sequencing (Priest et al. 2021; Gehrig et al. 2022; Meslier et al. 2022; Eisenhofer et al. 2023; Orellana et al. 2023; Tao et al. 2023; Zhang et al. 2023), and that HiFi sequencing outperforms ONT for metagenome assembly (Meslier et al. 2022; Sereika et al. 2022). The high accuracy of HiFi reads has led to the development of two new metagenome assembly methods, hifiasm-meta (Feng et al. 2022) and metaMDBG (Benoit et al. 2023), which perform better than previous methods such as HiCanu (Nurk et al. 2020) and metaFlye (Kolmogorov et al. 2020). Hifiasm-meta phases reads based on single nucleotide variants and constructs the string graph using only intra-haplotype read overlaps. Heuristics to keep contained reads with ambiguous phasing and to resolve tangles using either unitig read coverages or topology are implemented in graph cleaning steps. By contrast, metaMDBG uses de-Bruijn graph assembly in a mimizer-space and a progressive abundance-based filtering strategy to simplify strain complexity. Performing metagenome assembly with these HiFi-specific methods routinely produces complete, circular MAGs, and sometimes in large numbers (Bickhart et al. 2022; Feng et al. 2022; Kato et al. 2022; Kim et al.

2022; Zhang et al. 2022; Benoit et al. 2023; Jiang et al. 2023; Saak et al. 2023; Schaerer et al. 2023; Masuda et al. 2024). For example, using hifiasm-meta Kim et al. (2022) recovered 102 complete, circular MAGs from five human gut microbiome samples (including 39 MAGs that remain uncultured), and a reanalysis of data from Bickhart et al. (2022) with metaMDBG recovered 447 high-quality MAGs from a single sheep gut microbiome, including 266 circular MAGs (Benoit et al. 2023).

Although metagenome assembly with HiFi sequencing data can produce complete MAGs, additional genomes will be represented by two or more contigs. Reconstructing these additional MAGs requires the use of binning algorithms. A majority of bioinformatic binning methods perform clustering using tetranucleotide frequencies, depth of coverage, or deep learning methods (Sczyrba et al. 2017; Meyer et al. 2022), and most were designed for short-read assemblies. When applied to long-read assemblies, these methods can produce unexpected or suboptimal results. To address this, new long-read binning algorithms have been developed, including MetaBCC-LR (Wickramarachchi et al. 2020), LRBinner (Wickramarachchi & Lin 2022), GraphMB (Lamurias et al. 2022), and most recently SemiBin2 (Pan et al. 2023). Metagenomic assembly workflows often include multiple bioinformatic binning methods, and the results from multiple binning algorithms are typically merged. This key step de-replicates the bin sets from alternative methods, resulting in improvements to MAG quality and yield (Olm et al. 2017; Uritskiy et al. 2018). In addition to bioinformatic approaches, proximity ligation information (e.g., Hi-C) can also be used to bin contigs (Burton et al. 2014; Press et al. 2017; DeMaere & Darling 2019). The ProxiMeta platform (Press et al. 2017), MetaCC (Du & Sun 2023), and metaBAT-LR (Ho et al. 2023) can perform binning using Hi-C data and a set of assembled contigs. Proximity ligation also allows mobile elements to be associated with their respective host genomes, offering a unique advantage. Proximity ligation binning methods have been successfully applied to long-read metagenome assemblies, presumably improving MAG yields relative to bioinformatic binning (Bickhart et al. 2022; Saak et al. 2023). However, no studies have performed a systematic comparison of the MAG sets produced from bioinformatic and proximity ligation binning using long reads.

Here, we performed a deep-sequencing experiment on a pooled human gut microbiome to produce a catalog of highly resolved MAGs. This pooled gut sample is particularly challenging for metagenome assembly, as the number of species and strains is higher than what is typically

found in a gut microbiome. We obtained HiFi data from the PacBio Sequel IIe and Revio systems. We assembled the combined sequencing dataset using hifiasm-meta and metaMDBG, and performed binning using bioinformatic and proximity ligation approaches. For bioinformatic binning, we created a new version of a workflow designed to process long-read metagenome assemblies (HiFi-MAG-Pipeline), and we used the ProxiMeta platform to perform proximity ligation binning and association of mobile elements. We also developed a new algorithm (pb-MAG-mirror) which compares the contents of MAGs obtained from two alternate binning approaches (based on the same set of starting contigs), and consolidates the MAGs into a single non-redundant set. In an effort to characterize tradeoffs in performance, we compared the number, quality, and taxonomy of MAGs obtained from the various combinations of assembly and binning methods. Finally, we performed a downsampling experiment to understand the effects of total data on MAG recovery. Overall, our study illustrates how HiFi sequencing can be used to obtain highly complete MAGs from the human gut microbiome, accelerating our ability to catalog the microbial diversity in this complex system.

RESULTS

Sequencing datasets

Our combined PacBio HiFi sequencing dataset consists of 34.7 million HiFi reads and 255.9 Gb of total data, with a mean QV of 42.5 (Table 1). We obtained 17.1 million HiFi reads (120.8 Gb) from six 8M SMRT Cells on the Sequel IIe system, and 17.5 HiFi reads (135.0 Gb) from two 25M SMRT Cells on the Revio system. The Sequel IIe HiFi reads display a mean length of 7.0 kb and mean QV of 40.5 (Supplemental Fig. S1); yields across cells ranged from 2.30–3.33 million reads and 14.1–27.8 Gb total data. The Revio HiFi reads had a similar mean length of 7.7 kb but a higher mean QV of 45.8. Importantly, predicted read QV scores were higher in the Revio dataset despite a fewer mean number of passes (18 versus 24; Supplemental Fig. S1). Both Revio runs were highly consistent in yielding 8.7–8.8 million HiFi reads and 67.0–67.9 Gb total data.

Metagenome assembly and binning

We performed metagenome assembly using two HiFi-specific methods: hifiasm-meta and metaMDBG (Fig. 1). The assembly with hifiasm-meta resulted in 34,117 contigs with an N50 of 458 kbp and total assembly size of 2.76 Gbp. We found that metaMDBG produced 66,952 contigs with an N50 of 306 kbp and total assembly size of 2.75 Gbp. The runtime performance differed between the assembly methods. Using hifiasm-meta took 5 days and 5.8 hours wall time using 64 threads (7767 CPU hrs) with peak memory of 713 GB, whereas metaMDBG required 21.13 hours wall time (1352 CPU hrs) and 12.7 GB peak memory (using 64 threads).

To obtain MAGs from the assembled contigs, we performed bioinformatic binning using the HiFi-MAG-Pipeline and proximity ligation binning using ProxiMeta (Fig. 1). We categorized MAGs as medium-quality or high-quality (MQ-MAG, HQ-MAG, respectively) based on standards proposed by the Genomic Standards Consortium (Bowers et al. 2017). The HQ-MAGs require ≥90% completeness (based on universal single-copy genes - SCGs) and ≤5% contamination, whereas MQ-MAGs fall below these criteria but display ≥50% SCG completeness and ≤10% contamination. Overall, we obtained large numbers of total MAGs and HQ-MAGs across all the assembly and binning combinations. Specifically, using bioinformatic binning we obtained 485 and 480 total MAGs for hifiasm-meta and metaMDBG, respectively (Fig. 2, Table 2). Comparing the binning strategies, we found ProxiMeta produced more total MAGs than bioinformatic binning, with 522 and 492 total MAGs for hifiasm-meta and metaMDBG, respectively (Fig. 2, Table 2). This represents an 8% and 3% increase in total MAGs for the two assembly methods. Across both binning strategies, the total number of MAGs produced by hifiasm-meta is higher than metaMDBG. However, we found metaMDBG has a higher proportion of HQ-MAGs relative to hifiasm-meta across both binning strategies (bioinformatic: 239 vs. 157; proximity ligation: 258 vs. 166; Fig 2). Both assembly methods recovered similar numbers of SC-HQ-MAGs, but the numbers differed across binning strategy (Table 2). The first step of the HiFi-MAG-Pipeline was designed to identify highly complete, single-contig MAGs using a “completeness-aware” strategy. Based on this approach, we recovered 98 and 96 SC-HQ-MAGs from hifiasm-meta and metaMDBG, respectively, with 70 and 74 being circular. The ProxiMeta workflow does not include a parallel initial step, and as a result we recovered 79 and 70 SC-HQ-MAGs (from hifiasm-meta and metaMDBG, respectively, both with 59 circular). For ProxiMeta we observed several cases in which one or more small contigs (<50kb) were binned with complete bacterial chromosomes (identified as SC-HQ-MAGs

using HiFi-MAG-Pipeline), and they were subsequently excluded from being counted as SC-HQ-MAGs. The smaller contigs may possibly represent unannotated mobile elements. Overall, ProxiMeta produced higher HQ and MQ-MAG yields than the HiFi-MAG-Pipeline.

We compared and consolidated the bioinformatic and proximity ligation bin sets using our new method, pb-MAG-mirror. Our analysis with pb-MAG-mirror resulted in more MQ- and HQ-MAGs than either binning approach individually (Fig. 2). The analysis assigned MAGs to four comparison categories, including identical, superset/subset, unique, and high-, medium- and low-similarity mixed (Fig. 1). Three of these categories involved identifying highly similar pairs of MAGs across the two binning strategies (e.g., the identical, superset/subset, and high-similarity mixed categories; Fig. 1), whereas the unique, medium- and low-similarity mixed categories identified distinct or poorly overlapping MAGs. These classifications allowed us to identify the proportion of highly similar MAGs shared across binning methods, which guided the consolidation step (see Methods). Our analysis with pb-MAG-mirror produced 595 and 547 total consolidated MAGs for hifiasm-meta and metaMDBG, representing 11–22% more total MAGs than bioinformatic or proximity ligation binning individually. We found that 61–66% of total MAGs were classified as highly similar across the bioinformatic and proximity ligation binning methods (Fig. 2, Supplementary Table S1). This result indicates that a majority of MAGs contain highly similar contents across the binning approaches. By contrast, we found approximately 18–23% of the total MAGs were assigned to the medium and low-similarity mixed categories, which had little to no overlap in MAG composition. These MAGs tended to have lower completeness scores and higher contamination scores relative to other categories (Supplemental Fig. S2). Finally, the number of unique MAGs varied across binning and assembly methods, ranging from 11–20% of the total MAGs. ProxiMeta produced more unique MAGs than HiFi-MAG-Pipeline, and the highest number of unique MAGs occurred for the hifiasm-meta and ProxiMeta combination (Fig. 2, Supplementary Table S1). We investigated the characteristics of the unique MAGs from each binning method, and found the unique MAGs from HiFi-MAG-Pipeline and ProxiMeta had similar completeness scores (~70%, for both assembly methods; Supplementary Fig. S2). However, the unique MAGs from HiFi-MAG-Pipeline displayed a higher average number of contigs (hifiasm-meta: 17 vs. 9 contigs; metaMDBG: 18 vs. 13 contigs) and a lower average depth of coverage (hifiasm-meta: 21x vs. 64x; metaMDBG: 27x vs. 87x), relative to ProxiMeta (Supplementary Fig. S2). In the consolidated MAG set, a major contribution comes

from the inclusion of the unique MAGs from both methods. The combined unique MAGs represent 30% and 22% of the total consolidated MAGs for hifiasm-meta and metaMDBG, respectively.

Effects of total data

To investigate the effects of total data on MAG yield, we downsampled our Sequel IIe dataset to lower data levels (Supplementary Table S2). Overall, we found the patterns from the full dataset were recapitulated in the downsampled datasets. For example, across both assembly methods we found the consolidated MAG set resulted in the greatest numbers of MQ- and HQ-MAGs, followed by proximity ligation binning, then bioinformatic binning (Supplementary Figs. 3, 4; Supplementary Table S3). In the full sequencing datasets, hifiasm-meta produced more total MAGs than metaMDBG (Fig. 2), but this pattern was not consistent across the downsampled datasets. We found metaMDBG produced more total MAGs in 14% of the downsampled comparisons, most often in conjunction with the HiFi-MAG-Pipeline (Supplementary Figs. S3, S4, Table S3). However, across all data levels we found metaMDBG produced more HQ-MAGs, and fewer MQ-MAGs, than hifiasm-meta (Fig. 3). We note our smallest downsampled dataset contained only ~360,000 HiFi reads and 3Gb total data, yet it resulted in up to 73 total MAGs and 10 HQ-MAGs (Supplementary Table S3). Looking across all data levels (including our full dataset), we found a predictable logarithmic relationship between total data and the number of HQ or MQ-MAGs in the consolidated MAG sets (Fig. 3; Supplementary Table S4). These trendlines indicate that although additional sequencing could further increase MAG yields, there are diminishing returns relative to lower data levels.

Taxonomic diversity and strain-level variation

The number of species assigned by GTDB-Tk ranged from 313–385 species across assembly and binning combinations, with the highest number of species occurring in the consolidated MAG sets for hifiasm-meta (n=364) and metaMDBG (n=385; Table 3). The number of MAGs that were not assigned to the species level ranged from 65–96 across the method combinations, representing 13–17% of the total MAGs. The highest numbers again occurred in the consolidated MAG sets, with 96 MAGs not assigned to the species rank for hifiasm-meta (16%) and 75 for metaMDBG (14%). Given the high number of unassigned MAGs, the species counts from

GTDB-Tk represent a conservative estimate of the total species diversity. The dRep results support this notion, as we recovered more species-level clusters across assembly and binning combinations, ranging from 385–455 (Table 3). The highest numbers of dRep species clusters were also found for the consolidated MAG sets, with 449 clusters for hifiasm-meta and 455 for metaMDBG (Fig. 4).

We observed a large proportion of strain-level variation in our MAG sets, particularly for method combinations including hifiasm-meta. Based on dRep species clusters, we found 73–100 species represented by multiple strains across the binning methods for hifiasm-meta, with only 52–64 species displaying multiple strains for metaMDBG (Fig. 4; Table 3). For hifiasm-meta, we found a total of 246 strain-level MAGs in the consolidated MAG set, representing 40% of the total MAGs (Table 3). A large proportion of these strain-level genomes for hifiasm-meta were classified as MQ-MAGs (68%). By contrast, we only found 156 strain-level MAGs for metaDBG, representing 28% of the total MAGs (Table 3). Of these, 52% were classified as MQ-MAGs. In the consolidated MAG sets there were many species containing two strains (hifiasm-meta: 68; metaMDBG: 46), several with three strains (hifiasm-meta: 20; metaMDBG: 10), but only a few displaying four (hifiasm-meta: 10; metaMDBG: 6) or five strains (hifiasm-meta: 2; metaMDBG: 2). Across all species with 2–5 strains, approximately ~65% contained at least one HQ-MAG, ~25% contained two or more HQ-MAGs, ~15% contained three HQ-MAGs, and none contained four or more HQ-MAGs (Supplementary Table S5). These results indicate that assembling multiple HQ-MAGs at the strain level is currently possible, but that strain-level variation is typically captured as one HQ-MAG plus one or more incomplete genomes. Among the species clusters displaying four to five strains, there are 12 for hifiasm-meta and 8 for metaMDBG (Supplementary Table S6). There are 15 high-strain-diversity species in the combined set: *Adlercreutzia celatus_A/equolifaciens*, *Agathobacter faecis*, *Agathobaculum butyriciproducens*, *Bacteroides uniformis*, *Blautia_A massiliensis*, *CAG-41 sp900066215*, *Copromonas sp000435795*, *Dorea_A longicatena*, *Dysosmobacter sp001916835*, *Faecalibacterium prausnitzii_D*, *Faecalibacterium prausnitzii_G*, *Faecalibacterium sp900539945*, *Lachnospira eligens_A/sp003451515*, *Phocaeicola dorei/vulgatus*, and *Ruminococcus_B gnavus*. We note that in addition to these high-strain-diversity species, there are an additional 26 species which contain three strains each.

Our main phylogenetic and taxonomic analyses were focused on the consolidated MAG sets for hifiasm-meta and metaMDBG (Fig. 4). At the phylum level, the MAG assignments were dominated by Firmicutes_A (hifiasm-meta: 450; metaMDBG: 416), followed by Bacteroidota and Firmicutes (hifiasm-meta: 50 and 30, respectively; metaMDBG: 46 and 26). Within Firmicutes_A, representation was highest among the families Lachnospiraceae (hifiasm-meta: 184; metaMDBG: 153), Oscillospiraceae (hifiasm-meta: 61; metaMDBG: 68), Ruminococcaceae (hifiasm-meta: 54; metaMDBG: 57), and Acutalibacteraceae (hifiasm-meta: 33; metaMDBG: 31), but included 33 families in total. Within Bacteroidota seven families were represented, with Bacteroidaceae (hifiasm-meta: 24; metaMDBG: 19) and Rikenellaceae (hifiasm-meta: 13; metaMDBG: 12) displaying the greatest number of MAGs. The phyla with the least amount of representation included Methanobacteria, Cyanobacteria, Desulfobacterota_I, Firmicutes_B, Firmicutes_C, Firmicutes_G, and Verrucomicrobiota (hifiasm-meta: 2, 6, 6, 5, 4, 0, and 5 MAGs, respectively; metaMDBG: 1, 4, 4, 5, 4, 1, and 5 MAGs). Based on GTDB-Tk, we determined 96 of the consolidated MAGs were not assigned to the species level for hifiasm-meta, and 75 were not assigned for metaMDBG. Of these, 21 from hifiasm-meta and 16 from metaMDBG were not assignable to the genus level. Of the set of MAGs not assigned at the species level, we found 16 and 36 were categorized as HQ-MAGs for hifiasm-meta and metaMDBG, respectively (Supplementary Table S7). A majority of these novel HQ-MAGs occur in the phylum Firmicutes_A (hifiasm-meta: 10; metaMDBG: 28), and within this phylum most occur in the orders Lachnospirales and Oscillospirales (Supplementary Table S7). In total, the novel HQ-MAGs recovered from hifiasm-meta and metaMDBG represent 33 distinct genera distributed across thirteen orders and six phyla.

Comparison of assembly methods

We performed comparisons of hifiasm-meta and metaMDBG at the contig and MAG-level, using several approaches. We performed an alignment of the contigs across assemblers, and found 69% of bases were shared at the 98% identity threshold, and 63% of bases were shared at the 99% identity threshold. We also compared large contigs (>500kb) using Mash and FracMinHash to understand contig similarity and containment (Supplementary Table S8). At a 99% mash similarity threshold (e.g., strain-level), we found 13.8% and 16.3% of large contigs were matched across hifiasm-meta and metaMDBG, with 13.2% and 14.4% being contained. At

95% for mash and 80% for FracMinHash, the number of contig matches for hifiasm-meta and metaMDBG was 39% and 53%, and 44% and 33% were contained. Finally. for 90% for mash and 70% for FracMinHash, the number of matched long contigs for hifiasm-meta and metaMDBG was 51% and 64%, with 47% and 33% contained.

We mapped the HiFi reads to the contigs and to the consolidated MAGs for both assembly methods, and found similar numbers of reliably mapped reads in both cases (Supplementary Table S9). Approximately 78% of HiFi reads were reliably mapped to the contigs (for both assembly methods), and 61% and 62% were mapped to the consolidated MAGs (hifiasm-meta and metaMDBG, respectively). These results indicate a large proportion of the information contained in the HiFi reads was represented in the contigs, with a somewhat smaller proportion represented in the final MAGs.

To understand how much species diversity was assembled into genomes, we searched for 16S genes in the HiFi reads and compared them to the 16S genes contained within MAGs. We found 108,830 reads contained full length 16S rRNA genes, and greedy clustering yielded 3,772 species-level OTUs (with a minimum of two reads each). We found 1,645 OTUs had more than 10 reads of support. Of the 1,645 OTUs with high coverage, 436 (26.5%) were also found in the consolidated HQ-MAG set (Fig. 5). When considering the MQ- and HQ-MAGs, we found a total of 778 OTUs represented (47.3%).

We estimated the number of MAGs that were unequivocally shared across the assembly methods, based on the consolidated MAG sets. We used a combination of aligned sequence lengths and ANI to measure MAG similarity, and explored the effects of different values on the number of matches (Supplemental Fig. S5). Ultimately, we defined an unequivocal match as requiring ≥90% of the total bases per MAG to be aligned with ≥99% ANI (Fig. 6), which is representative of a strain-level match. Based on these criteria, we found a total of 125 MAGs that were unequivocally shared across the assembly methods, representing 21–23% of the total MAGs for each method (Fig. 6, Supplementary Table S10). These unequivocally shared MAGs had higher average percent completeness scores (97% vs. 73–84%), lower average numbers of contigs (3–7 vs. 9–12), and higher average depths of coverages (186–192x vs. 72–78x) relative to the unmatched MAGs from each assembly method (Fig. 6). We determined 115 of the shared MAGs (92%) were classified as HQ-MAGs, but we also found matched MAGs with as low as 71% completeness.

Mobile element association and viral detection

We used ProxiMeta, geNomad, and VirSorter2 to detect viral, proviral, and plasmid sequences in the contigs from both assembly methods. For both assemblers geNomad detected the most total viral contigs (hifiasm-meta 4,551, metaMDBG 6,679), followed closely by VirSorter2 (hifiasm-meta 4,264, metaMDBG 6,203), and then ProxiMeta (hifiasm-meta 2,175, metaMDBG 2,501; Supplemental Table S11). For a given assembly, a large proportion of viral and proviral contigs were annotated by both VirSorter2 and geNomad (62–80%; Supplementary Fig. S6). In addition to viral and proviral categories, ProxiMeta and geNomad annotated plasmid sequences. We found geNomad recovered more plasmids for metaMDBG (11,983) relative to hifiasm-meta (6,827). ProxiMeta identified fewer total plasmid sequences than geNomad but differentiated between plasmids and integrated plasmids. For metaMDBG ProxiMeta detected 129 plasmids and 1,878 integrated plasmids, whereas for hifiasm-meta it found 164 plasmids and 2,268 integrated plasmids (Supplementary Table S11). In addition to identifying mobile element sequences, ProxiMeta's ProxiPhage algorithm predicts if a given mobile sequence is interacting with a microbial bin as a host. We found that ProxiMeta identified more viral-host and plasmid-host associations in metaMDBG (333 viral-host and 32 plasmid-host associations) than hifiasm-meta (123 viral-host and 25 plasmid-host associations). Taken together, these results indicate that we consistently detected more viral sequences in the metaMDBG contigs, but more proviral sequences in the hifiasm-meta contigs.

DISCUSSION

Cataloging the microbial diversity of the human gut presents major challenges, and large numbers of species and strains remain unknown. HiFi metagenomic sequencing provides a rapid and scalable alternative to culturomics, and also offers substantial improvements over short-read sequencing. Here, we performed a deep-sequencing experiment on a pooled human gut microbiome - a challenging sample containing an inflated number of species and strains. We generated large numbers of MAGs using different combinations of long-read assembly, binning, and consolidation methods (Fig. 2). Our study demonstrates that metagenome assembly with HiFi reads can produce large numbers of highly complete MAGs, corroborating the findings of previous studies (Bickhart et al. 2022; Feng et al. 2022; Kato et al. 2022; Kim et al. 2022; Zhang

et al. 2022; Benoit et al. 2023; Jiang et al. 2023; Saak et al. 2023; Schaerer et al. 2023; Masuda et al. 2024). Across all method combinations, we found up to 595 total MAGs, 277 HQ-MAGs, and 98 single-contig HQ-MAGs (including 70 circular). These MAGs represent up to 455 species clusters (Fig. 4), with 15% of the MAGs representing uncultured species (including >35 HQ-MAGs). Based on our results, we discuss the relative performance and trade-offs for different methods, which overall we find to be complementary.

We used bioinformatic and proximity ligation binning methods to group contigs into putative MAGs, and developed a new framework to explicitly compare MAGs across binning strategies (Fig. 1). From a simple performance perspective, we found ProxiMeta generally produced more total MAGs than the HiFi-MAG-Pipeline (Fig. 2, Supplementary Figs. S3, S4). However, our framework with pb-MAG-mirror allowed us to move beyond counting MAGs to better understand how the contents of MAGs were distributed across both methods. In an ideal scenario, two orthogonal methods would produce MAG sets with a large proportion of highly similar MAGs (e.g., MAGs with near-identical contig contents), increasing our confidence in the biological reality of those genomes. In the worst case scenario, the methods could produce MAG sets with no sensible overlap - the contig contents of individual MAGs are highly mixed across MAGs in the other method. It is also possible to recover unique MAGs (e.g., those with contig contents not appearing in the other method), which may highlight strengths or blindspots for a given method. In our study, we found ~65% of total MAGs were classified as highly similar across the binning methods, with ~20% being highly mixed and another ~15% being unique to each method (Fig. 2). These proportions were consistent across hifiasm-meta and metaMDBG, suggesting they are robust to the starting contigs. The large proportion of highly similar MAGs is reassuring, as the same core set of MAGs (roughly ~320) can be recovered using HiFi-MAG-Pipeline or ProxiMeta (Fig. 2). The 20% of highly mixed MAGs (~100) is somewhat disconcerting, highlighting major disagreements across the binning methods. Furthermore, to perform consolidation a decision must be made about which MAGs are more accurate. As previous studies have demonstrated higher accuracy for bins using proximity ligation information (Burton et al. 2014; Ho et al. 2023), we selected these in our consolidation step. Finally, we found each method produced ~15% unique bins each (ranging from 55–100 MAGs), highlighting information captured by only one method. Unique bins from HiFi-MAG-Pipeline had significantly lower coverages than those from ProxiMeta (which requires a minimum

coverage threshold), suggesting bioinformatic binning may perform better at recovering very low coverage MAGs. Conversely, ProxiMeta can use contact information to associate contigs and refine bins, potentially grouping contigs otherwise missed by tetranucleotide frequencies and coverage distributions. By comparing and consolidating the MAGs from ProxiMeta and HiFi-MAG-Pipeline, we identified overlap and incorporated complementary aspects of both methods to improve MAG yields. For example, ProxiMeta provided a 3–8% increase in total MAGs over the HiFi-MAG-Pipeline, but pb-MAG-mirror resulted in an 11–22% increase in total MAGs over both binning methods. Overall, our comparison revealed that high numbers of MAGs can be obtained using both binning methods, a large core set of MAGs is shared across the methods, and that consolidation can incorporate the unique information captured by each.

We used hifiasm-meta and metaMDBG for metagenome assembly, and based on performance we found different strengths for each method. At the highest level, we were interested in which assembly method produced the most MAGs. In general, hifiasm-meta produced more total MAGs than metaMDBG (Fig. 2), with some exceptions in the downsampled datasets. However, regarding genome quality we found that metaMDBG consistently produced more HQ-MAGs than hifiasm-meta (Figs. 2, 3). Specifically, HQ-MAGs make up 50% of the total MAGs for metaMDBG in the consolidated set, but only 30% of the total for hifiasm-meta. We also investigated how well each method assembled strain-level variation. Here, hifiasm-meta produced strains for 56% more species and assembled 57% more total strain-level MAGs relative to metaMDBG, for the consolidated MAGs (Fig. 4, Table 3). Of the species with strains assembled, we wanted to understand whether strain-level variation was represented as two or more near-complete genomes (e.g., HQ-MAGs), as one near-complete genome and several incomplete genomes, or as many fragmented genomes. Despite their differences in the total number of strains assembled, hifiasm-meta and metaMDBG recovered two or more strain-level, near-complete genomes for a similar number of species in the consolidated MAGs (21 and 24, respectively; Supplementary Table S5). For both assembly methods, strains were most often represented by one near-complete genome and multiple partial genomes. We observed that for hifiasm-meta, strains were also commonly represented by sets of partial genomes, whereas this was infrequent for metaMDBG. These results highlight that both assembly methods are capable of assembling strain-level variation into two or more near-complete genomes, resolving strains remains challenging, and that hifiasm-meta currently produces more strain-level information.

We investigated how many MAGs were shared across the assembly methods in the consolidated sets. We used strict criteria to identify unequivocal matches across the assembly methods, which resulted in 125 high-confidence matches at the strain level (Fig. 6). These unequivocally shared MAGs tended to be very high-quality, with greater completeness scores, fewer contigs (including many single-contig), and higher depth of coverage relative to MAGs not shared across the methods (Fig. 6). Although this only represents ~22% of the total MAGs in each method, we emphasize that even slight differences in the assembly or binning of contigs could disqualify a potential match. It is also possible that partial strain-collapsing in some MAGs could impact this type of analysis, which would be exacerbated in this high-strain-diversity sample. Remarkably, there are 17 matched MAGs which both consist of a single-contig, are within 200 bp of the total length of one another (across a 1.42–6.09 Mb size range), and display only 2–100 nucleotide differences (Supplementary Table S10). These results highlight that the use of highly accurate long reads can lead to the recovery of nearly identical MAGs, despite the use of disparate assembly and binning methods.

Beyond bacterial and archaeal genomes, viral and plasmid genomes pose a different set of assembly challenges (Antipov et al. 2020). For long reads specifically, their size range can fall below the length of a typical HiFi read (10–20 kb), and their structure is quite variable. We used ProxiMeta, VirSorter2, and geNomad to identify and annotate viral sequences, and in most cases recovered more total viral sequences for metaMDBG (Supplementary Table S11). With the ProxiMeta platform, we detected more viral and plasmid host associations for metaMDBG versus hifiasm-meta. These results indicate metaMDBG may have an advantage for assembling smaller, viral sequences, but that additional work is required to improve assembly for these highly variable genomes.

The routine production of single-contig HQ-MAGs from long-read metagenome assemblies is a recent phenomenon, and we propose it requires some important considerations. First, a majority of binning algorithms operate with an implicit assumption that the assembled contigs represent fragmented genomes, and this assumption is often violated with HiFi metagenome assembly. Performing naïve binning on HiFi assemblies can result in the mis-binning of single-contig MAGs with additional contigs, inflating their contamination score and causing removal during quality filtering (Feng et al. 2022; Benoit et al. 2023). For this reason, we built a “completeness-aware” strategy into the HiFi-MAG-Pipeline, and propose it is an

essential step for any long-read binning workflow. This strategy involves screening all contigs longer than 500kb to obtain completeness and contamination scores, and moving all highly complete, single-contig MAGs into individual bins. The subsequent binning steps therefore only include incomplete contigs, preventing the mis-binning and removal of any SC-HQ-MAGs. Second, the quality standards proposed by the Genomic Standards Consortium (Bowers et al. 2017) have proven useful for comparing short-read MAGs, but they do not capture additional information relevant for long-read MAGs. Specifically, contiguity (i.e., the number of contigs per MAG) is becoming increasingly important for evaluating quality. For example, a short-read MAG can be defined as high-quality based on the presence of single-copy genes despite containing hundreds of contigs (and potentially missing large portions of the accessory genome). An ideal HQ-MAG is composed of a single contig per chromosome, and these are often distinguished in long-read studies (Bickhart et al. 2022; Feng et al. 2022; Kim et al. 2022; Zhang et al. 2022; Benoit et al. 2023; Jiang et al. 2023; Masuda et al. 2024), including ours (labeled as SC-HQ-MAGs). However, there is currently no distinction between HQ-MAGs composed of a few long contigs (e.g., hundreds of kilobases, or megabases) versus those containing hundreds of short contigs (e.g., tens of kilobases or shorter), despite the former being much higher quality. We propose this idea should be reflected in future terminology, allowing distinctions to be made along the contiguity spectrum. Beyond contiguity, circularity is sometimes also highlighted as an important quality metric (Bickhart et al. 2022; Kim et al. 2022; Zhang et al. 2022; Jiang et al. 2023; Masuda et al. 2024). However, Kim et al. (2022) demonstrated that circular MAGs can contain large gaps, and circularity alone should not be taken as a reliable indicator of completeness. Furthermore, linear contigs can sometimes be closed using additional read-mapping information. Therefore, information about linear vs. circular SC-HQ-MAGs can be useful, but may not be as meaningful as other quality categories. Although defining new standards is beyond the scope of this study, we suggest these criteria should be re-evaluated by the community as long-read metagenomics continues to become more mainstream.

METHODS

Sample

The ZymoBIOMICS™ Fecal Reference with TruMatrix™ Technology [D6323] (Zymo Research – CA, USA) was made by collecting multiple fecal samples from healthy human

485 volunteers following ethical guidelines and informed consent procedures. Samples were
486 collected using sterile collection tubes and stored at -80°C until processing. Collections from
487 multiple donors were pooled and homogenized in one large suspension together with the
488 microbiome preservative reagent DNA/RNA Shield™ (Zymo Research – CA, USA) to prevent
489 outgrowths or depletions of microbial taxa. The fecal suspension was then distributed to
490 containers to be stored at -80°C until aliquoting to individual tubes. Each stored container was
491 tested individually by sequencing to confirm consistency prior to aliquoting to individual tubes.
492 All aliquots have been demonstrated to have consistent metagenomic and metatranscriptomic
493 profiles.
494
495 **Sequencing**
496 DNA was extracted using the ZymoBIOMICS DNA Miniprep Kit (D4300), which involved
497 either mechanical (Vortex Genie 2) or enzymatic lysis. Fragment sizes were measured using a
498 Femto Pulse (Agilent). A total of eight SMRTbell libraries were prepared from four DNA
499 extractions. Libraries for the PacBio Sequel IIe system were created using the SMRTbell Express
500 Template Prep Kit 2.0 (n=3) or SMRTbell Prep Kit 3.0 (n=3), and libraries were prepared for the
501 Revio system using the SMRTbell Prep Kit 3.0 (n=2). Size selection was performed using 3.7X
502 35% Ampure SPRIselect (Beckman Coulter), which targeted the removal of fragments <3kb. Six
503 libraries were sequenced on the Sequel IIe system (using 40–60 ng library input) and two on the
504 Revio system (using 130–140 ng library input), with each library sequenced on an individual 8M
505 or 25M SMRT Cell, respectively. For the Sequel IIe system, HiFi reads were generated using ccs
506 v6.2, whereas for the Revio system HiFi reads were generated on-instrument using Google
507 DeepConsensus (Baid et al. 2023). For downstream data analysis, we combined the HiFi data
508 obtained from Sequel IIe and Revio. To investigate the effects of lower data levels on assembly,
509 we performed downsampling on the Sequel IIe dataset. Our downsampling design includes the
510 equivalent of six 8M SMRT Cells down to one cell, along with a 2plex, 4plex, and 8plex on one
511 cell (Supplemental Table S1).
512 A proximity ligation (Hi-C) library was prepared from an aliquot of the Fecal Reference
513 standard using the ProxiMeta Hi-C v4.0 Kit from Phase Genomics according to the manufacturer
514 provided protocol (Lieberman-Aiden et al. 2009). Briefly, intact cells were crosslinked using a
515 formaldehyde solution, simultaneously digested using the *Sau3AI* and *MlucI* restriction enzymes,

and proximity ligated with biotinylated nucleotides to create chimeric molecules composed of fragments from different regions of microbial genomes that were physically proximal *in vivo*. Proximity ligated DNA molecules were pulled down with streptavidin beads and processed into an Illumina-compatible sequencing library. Sequencing was performed on an Illumina NovaSeq with PE150 read pairs, producing 181.7M reads in total.

Metagenome assembly and binning

We performed assemblies with hifiasm-meta r74 (Feng et al. 2022) and metaMDBG v0.3 (Benoit et al. 2023) using default parameter settings. The resulting contigs were processed using two distinct methodologies, including bioinformatic binning and proximity ligation (Hi-C) binning. A visual overview of our analysis workflow is shown in Figure 1.

For bioinformatic binning, we developed a new version of the HiFi-MAG-Pipeline (available from: https://github.com/PacificBiosciences/pb-metagenomics-tools), which uses a completeness-aware binning strategy. The workflow begins by identifying all contigs longer than 500kb which display >90% completeness and <10% contamination, as measured by CheckM2 v1.0.1 (Chklovski et al. 2023). These highly complete, single-contig MAGs are removed from the contig set and placed in individual bins. The remaining set of incomplete contigs is then subjected to binning using the long-read mode of SemiBin2 v1.5 (Pan et al. 2022, 2023) and with MetaBAT2 v2.15 (Kang et al. 2019). To obtain coverage scores per contig, minimap2 v2.17 (Li 2018, 2021) is used to map reads to the contig set, and the jgi_summarize_bam_contig_depths script of MetaBAT2 is used to summarize depth per contig. Settings for MetaBAT2 included -m 30000, whereas SemiBin2 used the single_easy_bin module and included the --self-supervised, --sequencing-type=long_reads, and --environment=human_gut flags. The two bin sets are de-replicated using DAS_Tool v1.1.6 (Sieber et al. 2018), and the de-replicated bin set is assessed with CheckM2. The bins which pass minimum criteria for single-copy gene (SCG) completeness and contamination are added to the set of highly complete, single-contig MAGs identified in the first step. Finally, the Genome Taxonomy Database Toolkit (GTDB-Tk) v2.1.1 (Chaumeil et al. 2019, 2022) is used to assign taxonomy to all filtered MAGs.

MAGs were categorized as medium-quality or high-quality (MQ-MAG, HQ-MAG, respectively): HQ-MAGs require ≥90% SCG completeness and ≤5% contamination, whereas MQ-MAGs fall below these criteria but display ≥50% SCG completeness and ≤10%

contamination. We note that the total number of MAGs reported in our results is equal to the number of HQ-MAGs plus MQ-MAGs. We also include a third category for the highest level of quality, which we label as a single-contig, high-quality MAG (SC-HQ-MAG). This category includes the same criteria as an HQ-MAG, plus the presence of a single contig (circularity optional).

In addition to bioinformatic binning, we used proximity ligation to generate bins (Press et al. 2017). Proximity ligation sequencing files and assembled contigs were uploaded to the cloud-based ProxiMeta platform (Uritskiy et al. 2021). Proximity ligation reads were aligned using BWA-MEM (Li & Durbin 2010) with the -5SP options specified, and all other options default. SAMBLASTER (Faust & Hall 2014) was used to flag PCR duplicates, which were later excluded from analysis. Alignments were then filtered with samtools (Li et al. 2009) using the -F 2304 filtering flag to remove non-primary and secondary alignments. Metagenome deconvolution was performed with ProxiMeta (Press et al. 2017; Stewart et al. 2018), resulting in the creation of putative genome and genome fragment clusters, as well as viral, plasmid, and AMR gene host annotations. For consistency, we ran the bins inferred by the ProxiMeta workflow through CheckM2, filtered based on the above quality criteria, and assigned taxonomy using GTDB-Tk v2.1.1. This allowed a more direct comparison to the bins recovered by HiFi-MAG-Pipeline.

Binning method comparison

We compared the HiFi-MAG-Pipeline and ProxiMeta binning results using a new approach we developed, called pb-MAG-mirror. We compared the contig content of MAGs across methods to identify four categories: identical bins, superset/subset bins, mixed bins, and unique bins (Fig. 1). Identical bins occur when a bin from each method contains the exact same contig set. A subset bin occurs when the contig set of a bin is fully contained in a bin of the alternate method (i.e., the superset bin). We further require that any additional contigs in the superset bin do not occur in any other bin (e.g., any additional contigs present in the superset bin were unique to the superset bin). Lastly, a unique bin occurs when it contains a set of contigs that does not occur in any bin of the alternative method. If a single contig of the set can be found in a bin of the alternative method, the bin cannot be classified as unique. All bins not falling into these categories (identical, subset/superset, unique) are considered mixed, as they contain two or more contigs

that occur in two or more bins of the alternative method. For these mixed bins, we perform cross-method pairwise comparisons. We examine the shared contig content of each comparison to identify the best match, which is selected based on the percentage of total shared bases. We consider high-similarity (HS) matches as those with ≥80% shared bases in each bin, medium-similarity (MS) as those with ≥50% shared bases in each bin, and low-similarity (LS) as those with <50% shared bases in each bin. Based on the category results, pb-MAG-mirror can be used to consolidate the two bin sets. The default behavior is to include one representative bin for each of the identical bins, all superset bins and unique bins from both sets, and for cross-set mixed matches the bin with the highest completeness score is selected. However, for cross-set mixed matches the bins from one set can be preferentially selected instead. For our comparisons, for the cross-set mixed matches we selected ProxiMeta bins. We refer to the resulting MAGs as the consolidated MAG set (Fig. 1). To identify potential differences in characteristics of unique bins between HiFi-MAG-Pipeline and ProxiMeta, we compared completeness scores, number of contigs, and average depth of coverage for the unique bins. pb-MAG-mirror is available as an open-source workflow at: https://github.com/PacificBiosciences/pb-metagenomics-tools.

Taxonomic and phylogenetic analyses

We quantified the number of species contained in the various MAG sets using two approaches. First, we used the taxonomy assigned by GTDB-Tk to count the number of unique species assignments. However, several genomes were not assigned to the species level. These unassigned genomes could represent one or more species groups. To address this issue, we used dRep v3.4.3 (Olm et al. 2017) to cluster MAGs into species-level groups based on an average nucleotide identity (ANI) of 95% (Jain et al. 2018; Olm et al. 2020). The default settings were used, which involved the ANImf algorithm for genome comparisons, 90% ANI for creating primary clusters, and 95% ANI for creating secondary clusters. We regarded each resulting cluster as a unique species. For strain-level counts, we aggregated counts based on genomes with the same species (GTDB-TK) or cluster (dRep) assignment.

We reconstructed the phylogenetic relationships of the MAGs from the consolidated MAG sets for each assembly method. We obtained protein alignments for core bacterial and archaeal genes (Pfam, TIGRFAM) using the identify and align modules of GTDB-Tk. The core gene sets differ in number between bacteria and archaea (n=120 and n=53, respectively), with

five genes shared across both sets. We created a concatenated supermatrix of all genes and analyzed the supermatrix using RAxML v8.2.12 (Stamatakis 2014). Tree searches used the gamma model of rate heterogeneity and JTTDCMUT amino acid substitution model. We visualized the phylogeny using the Interactive Tree of Life (iTOL) v6.8.1 online tool (Letunic & Bork 2021).

Comparison of assembly methods

We compared contigs from hifiasm-meta and metaMDBG using two approaches. First, we aligned the contig sets using minimap2, and calculated the percent of bases covered based on 98 or 99% identity. Second, for contigs >500kb we calculated Mash similarities (MinHash) using mash (Ondov et al. 2016) and FracMinHash-based similarities using sourmash (Pierce et al. 2019; Irber et al. 2022). There were 996 contigs >500kb for hifiasm-meta and 867 for metaMDBG. Concordantly high values for both Mash and FracMinHash similarities were taken as strong evidence for a contig match across assemblers, whereas a high value for Mash and lower value for FracMinHash indicated a contig was likely contained within the other contig. A range of minimum cutoffs for each metric was explored.

We also investigated the number of HiFi reads that could be reliably mapped to the contigs or consolidated MAGs, per assembly method. We performed read mapping using minimap2 with HiFi settings (-x map-hifi -k19 -w19 --secondary=no), and subsequently filtered the alignments. We kept primary alignments in which ≥90% of the HiFi read was aligned to the reference, and with ≥95% identity (defined as number of matched query bases divided by total query bases in the alignment). To identify potential differences in MAG characteristics across the assembly methods, we compared the average depth of coverage and number of contigs for the consolidated HQ- and MQ-MAGs.

In an effort to understand how representative the MAGs are of the total species diversity in the sample, we compared 16S rRNA sequences obtained from the HiFi reads with those contained in MAGs (Feng & Li 2024). Methods were highly similar to those described in Feng and Li (2024). In brief, HiFi reads containing 16S regions were discovered by mapping to the SILVA database (Quast et al. 2013), rRNA genes were identified using barrnap v0.9 (https://github.com/tseemann/barrnap), annotated using the RDP classifier (Cole et al. 2014), and OTUs were defined using greedy incremental clustering and 99% mismatch identity. The MAG

16S sequences were identified and extracted from the consolidated MAG sets from hifiasm-meta and metaMDBG, and mapped to the OTU set from the HiFi reads. A MAG could have more than one OTU assignment if it contained multiple, distinct 16S copies.

We sought to determine the number of identical or nearly-identical MAGs that were unequivocally shared across the assembly methods. To accomplish this, we estimated the intersection of the consolidated MAG sets from each assembly method, using minimum thresholds for aligned sequence lengths and ANI. We used dRep to generate clusters for all consolidated MAGs from hifiasm-meta and metaDMBG, using the ANImf algorithm with 90% ANI for primary clusters and 99% ANI for secondary clusters. For each resulting cluster, we first determined if it contained at least one MAG per assembly method. If so, we then performed pairwise comparisons of all MAGs from hifiasm-meta to metaMDBG to identify unequivocal matches. To understand the effects of parameter values on the number of matches, we explored using 95% or 99% ANI and a range of values for the minimum percent of bases per MAG required (70–100%). Based on these results, we defined an unequivocal match as requiring ≥90% of the total bases to be aligned (per genome) with ≥99% ANI (e.g., a nearly identical match at the strain level)

Mobile element association and viral annotation

Mobile element host-association was performed using the cloud-based ProxiMeta platform (Uritskiy et al. 2021). Briefly, viral contigs were identified using VIBRANT (Kieft et al. 2020), and plasmid contigs were identified using BLAST and the NCBI Plasmid Database. Alignments of long-range Hi-C linkage data were used to identify viral-host and plasmid-host linkages. A combination of the Hi-C link count, mobile element read depth, and MAG read depth were then used to estimate the average copy count of each mobile element in each MAG. The density of Hi-C links per kb^2 of sequence between the mobile element and the MAG was then compared to the connectivity of the MAG to itself and normalized to the estimated copy count to compute the normalized connectivity ratio. Mobile-host linkages were then filtered to keep only connections with at least 2 Hi-C read links between the mobile and host MAG, a connectivity ratio of 0.1, and intra-MAG connectivity of 10 links to remove false positives. For the final threshold value, a receiver operating characteristic (ROC) curve is used to determine the optimal copy count cut-off value, which is the value that removes the maximum number of virus-host links while still

finding at least one host for the maximum number of mobile elements. Additional filtering removed linkages with an average copy count less than 80% of the highest copy count value for the given mobile element sequence.

As an alternative to ProxiMeta, we also predicted viral and proviral sequences using geNomad (Camarago et al. 2023) and VirSorter2 (Guo et al. 2021). geNomad (ver 1.8.0) was run in "relaxed" mode. VirSorter2 (v2.2.4) was run with all steps and results were verified with CheckV (v1.0.1) using "end-to-end" mode (Nayfach et al. 2021). VirSorter2 viral counts include all quality types reported from CheckV (complete, high quality, medium quality, low quality, and not determined).

DATA ACCESS

HiFi sequencing data are publicly available on NCBI (PRJNA1110296) and from:
https://downloads.pacbcloud.com/public/revio/2023Q3/ZymoTrumatrix/
https://downloads.pacbcloud.com/public/sequelii/2023Q3/ZymoTrumatrix/

The contigs, MAGs, and add sets for each binning (HiFi-MAG-Pipeline, ProxiMeta, consolidated) and assembly (hifiasm-meta, metaMDBG) combination are publicly available on the Open Science Framework: https://osf.io/cwqzr/

pb-MAG-mirror is publicly available at: https://github.com/PacificBiosciences/pb-metagenomics-tools

COMPETING INTEREST STATEMENT

DMP, DJN, JEW, and SZ are employees of Pacific Biosciences of California, Inc. KL, ST, BF, and AD are employees of Zymo Research. MC and RC are the CEO and CSO of The BioCollective, respectively. IL is the CEO of Phase Genomics, and BA, HM, KWL, MI, JRG, and SJB are employees of Phase Genomics.

ACKNOWLEDGMENTS

CQ acknowledges the support of the Biotechnology and Biological Sciences Research Council (BBSRC), part of UK Research and Innovation; Earlham Institute Strategic Programme Grant (Decoding Biodiversity) BBX011089/1 and its constituent work package BBS/E/ER/230002C; the Core Strategic Programme Grant (Genomes to Food Security) BB/CSP1720/1 and its constituent work packages BBS/E/T/000PR9818 and BBS/E/T/000PR9817; and the Core Capability Grant BB/CCG2220/1. GB was supported by the ERC IndexThePlanet grant (UE ERC-2022-COG INDEXTHEPLANET). Work by members of Phase Genomics were supported by NIH/NIAID grants R44AI172703 and R44AI162570 as well as funding from the Bill and Melinda Gates Foundation.

REFERENCES

Agustinho DP, Fu Y, Menon VK, Metcalf GA, Treangan TJ, and FJ Sedlazeck. 2024. Unveiling microbial diversity: harnessing long-read sequencing technology. *Nature Methods*, https://doi.org/10.1038/s41592-024-02262-1

Albertsen M. Long-read metagenomics paves the way toward a complete microbial tree of life. 2023. *Nautre Methods,* 20: 30–31.

Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, Lawley TD, and RD Finn. 2019. A new genomic blueprint of the human gut microbiota. *Nature*, 568: 499–504.

Almeida A, Nayfach S, Boland M, Strozzi F, Beracochea M, Shi ZJ, Pollard KS, Sakharova E, Parks DH, Hugenholtz P, Segata N, Kyrpides NC, and RD Finn. 2021. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nature Biotechnology*, 39: 105–114.

Antipov D, Raiko M, Lapidus A, and PA Pevzner. 2020. MetaviralSPAdes: assembly of viruses from metagenomic data. *Bioinformatics,* 36: 4126–4129.

Arumugam K, Bağcı C, Bessarab I, Beier S, Buchfink B, Górska A, Qiu G, Huson DH, and RBH Williams. 2019. Annotated bacterial chromosomes from frame-shift-corrected long-read metagenomic data. *Microbiome,* 7: 61.

Baid G, Cook DE, Shafin K, Yun T, Llinares-López F, Berthet Q, Belyaeva A, Töpfer A, Wenger AM, Rowell WJ, Yang H, Kolesnikov A, Ammar W, Vert J-P, Vaswani A, McLean CY, Nattestad M, Chang P-C, and A Carroll. 2023. DeepConsensus improves the accuracy of sequences with a gap-aware sequence transformer. *Nature Biotechnology*, 41: 232–238.

Benoit G, Raguideau S, James R, Phillippy AM, Chikhi R, and C Quince. 2023. Efficient high-quality metagenome assembly from long accurate reads using minimizer-space de Bruijn graphs. *bioRxiv*, https://doi.org/10.1101/2023.07.07.548136

Bickhart DM, Kolmogorov M, Tseng E, Portik DM, Korobeynikov A, Tolstoganov I, Uritskiy G, Liachko I, Sullivan ST, Shin SB, Zorea A, Andreu VP, Panke-Buisse K, Medema MH, Mizrahi I, Pevzner PA, and TPL Smith. 2022. Generating lineage-resolved, complete metagenome-assembled genomes from complex microbial communities. *Nature Biotechnology,* 40: 711-719.

Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A,

745 Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM,
746 Dodsworth JA, Yooseph S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ,
747 Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu W-T, Baker BJ, Rattei T, Eisen
748 JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I,
749 Tyson GW, Rinke C, The Genome Standards Consortium, Lapidus A, Meyer F, Yilmaz P,
750 Parks DH, Murat Eren A, Schriml L, Banfield JF, Hugenholtz P amd T Woyke. 2017.
751 Minimum information about a single amplified genome (MISAG) and a metagenome-
752 assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology*, 35: 725–731.
753 Burton JN, Liachko I, Dunham MJ, and J Shendure. 2014. Species-level deconvolution of
754 metagenome assemblies with Hi-C-based contact probability maps. *G3 (Bethesda)*, 4: 1339–
755 1346.
756 Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, Chain PSG, Nayfach S, and NC
757 Kyrpides. 2023. Identification of mobile genetic elements with geNomad. *Nature*
758 *Biotechnology*, https://doi.org/10.1038/s41587-023-01953-y
759 Chaumeil P-A, Mussig AJ, Hugenholtz P, and DH Parks. 2019. GTDB-Tk: a toolkit to classify
760 genomes with the Genome Taxonomy Database. *Bioinformatics*, 35: 1925–1927.
761 Chaumeil P-A, Mussig AJ, Hugenholtz P, and DH Parks. 2022. GTDB-Tk v2: memory friendly
762 classification with the genome taxonomy database. *Bioinformatics*, 38: 5315–5316.
763 Chen L-X, Anantharaman K, Shaiber A, Murat Eren A, and JF Banfield. 2020. Accurate and
764 complete genomes from metagenomes. *Genome Research*, 30: 315–333.
765 Chklovski A, Parks DH, Woodcroft BJ, and GW Tyson. 2023. CheckM2: a rapid, scalable and
766 accurate tool for assessing microbial genome quality using machine learning. *bioRxiv*,
767 https://doi.org/10.1101/2022.07.11.499243
768 Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske
769 CR, and JM Tiedje. 2014. Ribosomal database project: data and tools for high throughput
770 rRNA analysis. Nucleic Acids Research, 42: D633–D642.
771 DeMaere MZ, and AE Darling. 2019. bin3C: exploiting Hi-C sequencing data to accurately
772 resolve metagenome-assembled genomes. *Genome Biology*, 20: 46.
773 Du Y, and F Sun. 2023. MetaCC allows scalable and integrative analyses of both long-read and
774 short-read metagenomic Hi-C data. *Nature Communications*, 14: 6231.

775 Duvallet C, Gibbons SM, Gurry T, Irizarry RA, and EJ Alm. 2017. Meta-analysis of gut
776 microbiome studies identifies disease-specific and shared responses. *Nature*
777 *Communications*, 8: 1784.
778 Eisenhofer R, Nesme J, Santos-Bay L, Koziol A, Sorenson SJ, Alberdi A, and O Aizpurua. 2023.
779 A comparison of short-read, HiFi long-read and hybrid strategies for genome-resolved
780 metagenomics. *bioRxiv*, https://doi.org/10.1101/2023.10.04.560907
781 Faust GG, and IM Hall. 2014. SAMBLASTER: fast duplicate marking and structural variant
782 read extraction. *Bioinformatics*, 17: 2503–2505.
783 Feng X, Cheng H, Portik D, and H Li. 2022. Metagenome assembly of high-fidelity long reads
784 with hifiasm-meta. *Nature Methods*, 19: 671–674.
785 Feng X, and H Li. 2024. Evaluating and improving the representation of bacterial contents in
786 long-read metagenome assemblies. Genome Biology, 25: 92.
787 Forster SC, Kumar N, Anonye BO, Almeida A, Viciani E, Stares MD, Dunn M, Mkandawire
788 TT, Zhu A, Shao Y, Pike LJ, Louie T, Browne HP, Mitchell AL, Neville BA, Finn RD, and
789 TD Lawley. 2019. A human gut bacterial genome and culture collection for improved
790 metagenomic analyses. *Nature Biotechnology*, 37: 186–192.
791 Gehrig JL, Portik DM, Driscoll MD, Jackson E, Chakraborty S, Gratalo D, Ashby M, and R
792 Valladares. 2022. Finding the right fit: evaluation of short-read and long-read sequencing
793 approaches to maximize the utility of clinical microbiome data. *Microbial Genomics*, 8:
794 000794.
795 Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, and R Knight. Current
796 understanding of the human microbiome. *Nature Medicine*, 24: 392–400.
797 Gihawi A, Ge Y, Lu J, Puiu D, Xu A, Cooper CS, Brewer DS, Pertea M, and SL Salzberg. 2023.
798 Major data analysis errors invalidate cancer microbiome findings. *mBio*, 9: e0160723.
799 Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, Pratama AA,
800 Gazitúa MC, Vik D, Sullivan MB, and S Roux. 2021. VirSorter2: a multi-classifier, expert-
801 guided approach to detect diverse DNA and RNA viruses. *Microbiome*, 9: 37.
802 Ho H, Chovatia M, Egan R, He G, Yoshinaga Y, Liachko I, O'Malley R, and Z Wang. 2023.
803 Integrating chromatin conformation information in a self-supervised learning model
804 improves metagenome binning. *PeerJ*, 11: e16129.

Irber L, Brooks PT, Reiter T, Pierce-Ward NT, Hera MR, Koslicki D, and CT Brown. 2022. Lightweight compositional analysis of metagenomes with FracMinHash and minimum metagenome covers. bioRxiv, https://doi.org/10.1101/2022.01.11.475838

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

Jiang F, Li Q, Wang S, Shen T, Wang H, Wang A, Xu D, Yuan L, Lei L, Chen R, Yang B, Deng Y, and W Fan. 2023. Recovery of metagenome-assembled microbial genomes from a full-scale biogas plant of food waste by Pacific Biosciences high-fidelity sequencing. *Frontiers in Microbiology*, 13: 1095497.

Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, and Z Wang. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, 7: e7359.

Kato S, Masuda S, Shibata A, Shirasu K, and M Ohkuma. 2022. Insights into ecological roles of uncultivated bacteria in Katase hot spring sediment from long-read metagenomics. *Frontiers in Microbiology*, 13: 1045931.

Kieft K, Zhou Z, and K Anantharaman. 2020. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome*, 8: 90.

Kim CY, Ma J, and I Lee. 2022. HiFi metagenomic sequencing enables assembly of accurate and complete genomes from human gut microbiota. *Nature Communications*, 13: 6367.

Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, Kuhn K, Yuan J, Polevikov E, Smith TPL, and PA Pevzner. 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nature Methods*, 17: 1103–1110.

Lagier J-C, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, Levasseur A, Rolain J-M, Fournier P-E, and D Raoult. 2018. Culturing the human microbiota and culturomics. *Nature Reviews Microbiology*, 16: 540–550.

Lamurias A, Sereika M, Albertsen M, Hose K, and TD Nielsen. 2022. Metagenomic binning with assembly graph embeddings. *Bioinformatics*, 38: 4481–4487.

Letunic I, and P Bork. 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49: W293–W296.

836 Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I,
837 Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M,
838 Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, and J Dekker. 2009.
839 Comprehensive mapping of long-range interactions reveals folding principles of the human
840 genome. *Science*, 326: 289–293.
841 Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, 34: 3094–
842 3100.
843 Li H. 2021. New strategies to improve minimap2 alignment accuracy. *Bioinformatics*, 37: 4572–
844 4574.
845 Li H, and R Durbin. 2010. Fast and accurate long-read alignment with Burrows-Wheeler
846 transform. *Bioinformatics*, 26: 589–595.
847 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R
848 and 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map
849 (SAM) format and SAMtools. *Bioinformatics*, 25: 2078–2079.
850 Lynch SV, and O Pedersen. 2016. The human intestinal microbiome in health and disease. *The*
851 *New England Journal of Medicine*, 375: 2369–2379.
852 Masuda S, Gan P, Kiguchi Y, Anda M, Sasaki K, Shibata A, Iwasaki W, Suda W, and K Shirasu.
853 2024. Uncovering microbiomes of the rice phyllosphere using long-read metagenomic
854 sequencing. *Communications Biology*, 7: 357.
855 Meslier V, Quinquis B, Da Silva K, Plaza Onate F, Pons N, Roume H, Podar M, and M Almeida.
856 2022. Benchmarking second and third-generation sequencing platforms for microbial
857 metagenomics. *Scientific Data*, 9: 694.
858 Meyer F, Fritz A, Deng Z-L, Koslicki D, Lesker TR, Gurevich A, Robertson G, Alser M,
859 Antipov D, Beghini F, Bertrand D, Brito JJ, Brown CT, Buchmann J, Buluç A, Chen B,
860 Chikhi R, Clausen PTLC, Cristian A, Dabrowski PW, Darling AE, Egan R, Eskin E,
861 Georganas E, Goltsman E, Gray MA, Hansen LH, Hofmeyr S, Huang P, Irber L, Jia H,
862 Jørgensen TS, Kieser SD, Klemetsen T, Kola A, Kolmogorov M, Korobeynikov A, Kwan J,
863 LaPierre N, Lemaitre C, Li C, Limasset A, Malcher-Miranda F, Mangul S, Marcelino VR,
864 Marchet C, Marijon P, Meleshko D, Mende DR, Milanese A, Nagarajan N, Nissen J, Nurk S,
865 Oliker L, Paoli L, Peterlongo P, Piro VC, Porter JS, Rasmussen S, Rees ER, Reinert K,
866 Renard B, Robertsen EM, Rosen GL, Ruscheweyh H-J, Sarwal V, Segata N, Seiler E, Shi L,

867 Sun F, Sunagawa S, Sørensen SJ, Thomas A, Tong C, Trajkovski M, Tremblay J, Uritskiy G,
868 Vicedomini R, Wang Z, Wang Z, Wang Z, Warren A, Willassen NP, Yelick K, You R, Zeller
869 G, Zhao Z, Zhu S, Zhu J, Garrido-Oter R, Gastmeier P, Hacquard S, Häußler S, Khaledi A,
870 Maechler F, Mesny F, Radutoiu S, Schulze-Lefert P, Smit N, Strowig T, Bremges A, Sczyrba
871 A, and AC McHardy. 2022. Critical assessment of metagenome interpretation: the second
872 round of challenges. *Nature Methods*, 19: 429–440.
873 Nayfach S, Shi ZJ, Seshadri R, Pollard KS and NC Kyrpides. 2019. New insights from
874 uncultivated genomes of the global human gut microbiome. *Nature*, 568: 505–510.
875 Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, and NC Kyrpides. 2021. CheckV
876 assesses the quality and completeness of metagenome-assembled viral genomes. *Nature*
877 *Biotechnology*, 39: 578–585.
878 Nurk S, Walenz BP, Rhie A, Vollger MR, Logsdon GA, Grothe R, Miga K, Eichler EE,
879 Phillippy AM, and S Koren. 2020. HiCanu: accurate assembly of segmental duplications,
880 satellites, and allelic variants from high-fidelity long reads. *Genome Research*, 30: 1291–
881 1305.
882 Olm MR, Brown CT, Brooks B, and JF Banfield. 2017. dRep: a tool for fast and accurate
883 genomic comparisons that enables improved genome recovery from metagenomes through
884 de-replication. *The ISME Journal*, 11: 2864–2868.
885 Olm MR, Crits-Cristoph A, Diamond S, Lavy A, Matheus Carnevali PB, and JF Banfield. 2020.
886 Consistent metagenome-derived metrics verify and delineate bacterial species boundaries.
887 *mSystems*, 5: e00731-19.
888 Orellana LH, Kruger K, Sidhu C, and R Amann. 2023. Comparing genomes recovered from
889 time-series metagenomes using long- and short-read sequencing technologies. *Microbiome*,
890 11: 105.
891 Pan S, Zhu C, Zhao XM and LP Coelho. 2022. A deep Siamese neural network improves
892 metagenome-assembled genomes in microbiome datasets across different environments.
893 *Nature Communications*, 13: 2326.
894 Pan S, Zhao X-M, and LP Coelho. 2023. SemiBin2: self-supervised contrastive learning leads to
895 better MAGs for short- and long-read sequencing. *Bioinformatics*, 39: i21–i29.
896 Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A,
897 Ghensi P, Collado MC, Rice BL, DuLong C, Morgan XC, Golden CD, Quince C,

898 Huttenhower C, and N Segata. 2019. Extensive unexplored human microbiome diversity
899 revealed by over 150,000 genomes and metagenomes spanning age, geography, and lifestyle.
900 *Cell*, 176: 649–662.
901 Pierce NT, Irber L, Reiter T, Brooks P, and CT Brown. 2019. Large-scale sequence comparisons
902 with sourmash. F1000Research, 8: 1006.
903 Press MO, Wiser AH, Kronenberg ZN, Langford KW, Shakya M, Lo C-C, Mueller KA, Sullivan
904 ST, Chain PSG, and I Liachko. 2017. Hi-C deconvolution of a human gut micriobiome yields
905 high-quality draft genomes and reveals plasmid-genome interactions. *bioRxiv*,
906 https://doi.org/10.1101/198713.
907 Priest T, Orellana LH, Huettel B, Fuchs BM, and R Amann. 2021. Microbial metagenome-
908 assembled genomes of the Fram Strait from short and long read sequencing platforms. *PeerJ*,
909 9: e11721.
910 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, and FO Glöckner. 2013.
911 The SILVA ribosomal RNA gene database project: improved data processing and web-based
912 tools. Nucleic Acids Research, 41: D590–D596.
913 Saak CC, Pierce EC, Dinh CB, Portik D, Hall R, Ashby M, and RJ Dutton. 2023. Longitudinal,
914 multi-platform metagenomics yields a high-quality genomic catalog and guides an in vitro
915 model for cheese communities. *mSystems*, 8:e00701-22.
916 Schaerer L, Putman L, Bigcraft I, Byrne E, Kulas D, Zolghadr A, Aloba S, Ong R, Shonnard D,
917 and S Techtmann. 2023. Coexistence of specialist and generalist species within mixed plastic
918 derivative-utilizing microbial communities. *Microbiome*, 11: 224.
919 Sczyrba A, Hofmann P, Belmann P, Koslicki D, Janssen S, Dröge J, Gregor I, Majda S, Fiedler
920 J, Dahms E, Bremges A, Fritz A, Garrido-Oter R, Jørgensen TS, Shapiro N, Blood PD,
921 Gurevich A, Bai Y, Turaev D, DeMaere MZ, Chikhi R, Nagarajan N, Quince C, Meyer F,
922 Balvociūtė M, Hansen LH, Sørensen SJ, Chia BKH, Denis B, Froula JL, Wang Z, Egan R,
923 Kang DD, Cook JJ, Deltel C, Beckstette M, Lemaitre C, Peterlongo P, Rizk G, Lavenier D,
924 Wu Y-W, Singer SW, Jain C, Strous M, Klingenberg H, Meinicke P, Barton MD, Lingner T,
925 Lin H-H, Liao Y-C, Silva GGZ, Cuevas DA, Edwards RA, Saha S, Piro VC, Renard BY, Pop
926 M, Klenk H-P, Göker M, Kyrpides NC, Woyke T, Vorholt JA, Schulze-Lefert P, Rubin EM,
927 Darling AE, Rattei T, and AC McHardy. 2017. Critical assessment of metagenome
928 interpretation – a benchmark of metagenomics software. *Nature Methods*, 14: 1063–1071.

929 Sereika M, Kirkegaard RH, Karst SM, Michaelsen TY, Sorensen EA, Wollenberg RD, and M
930 Albertsen. 2022. Oxford Nanopore R10.4 long-read sequencing enables the generation of
931 near-finished bacterial genomes from pure cultures and metagenomes without short-read or
932 reference polishing. *Nature Methods*, 19: 823-826.
933 Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, and JF Banfield. 2018.
934 Recovery of genomes from metagenomes via a dereplication, aggregation and scoring
935 strategy. *Nature Microbiology*, 3: 836–843.
936 Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
937 large phylogenies. *Bioinformatics*, 30: 1312–1313.
938 Stewart RD, Auffret MD, Warr A, Wiser AH, Press MO, Langford KW, Liachko I, Snelling TJ,
939 Dewhurst RJ, Walker AW, Roehe R, and M Watson. 2018. Assembly of 913 microbial
940 genomes from metagenomic sequencing of the cow rumen. *Nature Communications*, 9: 870.
941 Tao Y, Xun F, Zhao C, Mao Z, Li B, Xing P, and QL Wu. 2023. Improved assembly of
942 metagenome-assembled genomes and viruses in Tibetan saline lake sediment by HiFi
943 metagenomic sequencing. *Microbiology Spectrum*, 11: e03328-22.
944 Tyson GQ, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin
945 EM, Rokhsar DS, and JF Banfield. 2004. Community structure and metabolism through
946 reconstruction of microbial genomes from the environment. *Nature*, 428: 37–43.
947 Uritskiy G, DiRuggiero J, and J Taylor. 2018. MetaWRAP – a flexible pipeline for genome-
948 resolved metagenomic data analysis. *Microbiome*, 6: 158.
949 Uritskiy G, Press M, Sun C, Huerta GD, Zayed AA, Wiser A, Grove J, Auch B, Eacker SM,
950 Sullivan S, Bickhart DM, Smith TPL, Sullivan MB, and I Liachko. 2021. Accurate viral
951 genome reconstruction and host assignment with proximity-ligation sequencing. *bioRxiv*,
952 https://doi.org/10.1101/2021.06.14.448389.
953 Wang J, and H Jia. 2016. Metagenome-wide association studies: fine-mining the microbiome.
954 *Nature Reviews Microbiology*, 14: 508–522.
955 Wenger AM, Peluso P, Rowell WJ, Chang P-C, Hall RJ, Concepcion GT, Ebler J,
956 Fungtammasan A, Kolesnikov A, Olson ND, Töpfer A, Alonge M, Mahmoud M, Qian Y,
957 Chin C-S, Phillippy AM, Schatz MC, Myers G, DePristo MA, Ruan J, Marschall T,
958 Sedlazeck FJ, Zook JM, Li H, Koren S, Carroll A, Rank DA, and MW Hunkapiller. 2019.

959 Accurate circular consensus long-read sequencing improves variant detection and assembly
960 of a human genome. *Nature Biotechnology*, 37: 1155–1162.
961 Wickramarachchi A, Mallawaarachchi V, Rajan V, and Y Lin. 2020. MetaBCC-LR:
962 metagenomics binning by coverage and composition for long reads. *Bioinformatics*, 36: i3–
963 i11.
964 Wickramarachchi A, and Y Lin. 2022. Binning long reads in metagenomics datasets using
965 composition and coverage information. *Algorithms for Molecular Biology*, 17: 14.
966 Zhang Y, Jiang F, Yang B, Wang S, Wang H, Wang A, Xu D, and W Fan. 2022. Improved
967 microbial genomes and gene catalog of the chicken gut from metagenomic sequencing of
968 high-fidelity long reads. *GigaScience*, 11: 1-12.
969 Zhang Z, Yang C, Veldsman WP, Fang X, and L Zhang. 2023. Benchmarking genome assembly
970 methods on metagenomic sequencing data. *Briefings in Bioinformatics*, 24: 1-17.
971 Zou Y, Xue W, Luo G, Deng Z, Qin P, Guo R, Sun H, Xia Y, Liang S, Dai Y, Wan D, Jiang R,
972 Su L, Feng Q, Jie Z, Guo T, Xia Z, Liu C, Yu J, Lin Y, Tang S, Huo G, Xu X, Hou Y, Liu X,
973 Wang J, Yang H, Kristiansen K, Li J, Jia H, and L Xiao. 2019. 1,520 reference genomes
974 from cultivated human gut bacteria enable functional microbiome analyses. *Nature*
975 *Biotechnology*, 37: 179–185.
976

Table 1. Summary of HiFi sequencing datasets.

Dataset	Sequencing level	HiFi reads (million)	HiFi yield (Gb)	Mean read length	Mean QV
Revio	2 SMRT Cells (25M)	17.56	135.06	7,690	45.8
Sequel IIe	6 SMRT Cells (8M)	17.15	120.85	7,046	40.2
Combined	-	34.71	255.91	7,724	42.5

Table 2. Counts of MAGs by quality for each dataset and binning method combination.

Assembly method	Binning method	SC-HQ-MAGs	HQ-MAGs	MQ-MAGs	Total MAGs
hifiasm-meta	HiFi-MAG-Pipeline	70/98	157	328	485
	ProxiMeta	59/79	166	356	522
	Consolidated	61/83	175	420	595
metaMDBG	HiFi-MAG-Pipeline	74/96	239	241	480
	ProxiMeta	59/70	258	234	492
	Consolidated	61/73	277	270	547

The HQ-MAG category requires >90% completeness and <5% contamination, whereas MQ-MAGs fall below those criteria but display >50% completeness and <10% contamination. Total MAGs = HQ + MQ-MAGs. SC-HQ-MAGs are a subcategory of HQ-MAGs with one contig present. For SC-HQ-MAGs, the first number represents genomes that are circular, whereas the second number represents all genomes (linear or circular).

Table 3. Estimates of species and strain information obtained by reference-based taxonomic assignment (GTDB-Tk) and MAG clustering (dRep).

Assembly method	Binning method	GTDB-Tk			dRep		
		Total species	Species with strains	Total strain-level MAGs	Unique clusters	Clusters with strains	Total strain-level MAGs
hifiasm-meta	HiFi-MAG-Pipeline	313	67	158	385	73	173
	ProxiMeta	329	82	196	399	85	208
	Consolidated	364	97	233	449	100	246
metaMDBG	HiFi-MAG-Pipeline	347	50	115	414	52	118
	ProxiMeta	351	60	137	411	60	141
	Consolidated	385	64	152	455	64	156

Figure 1. **(a)** Visual depiction of analysis steps involved with metagenome assembly, binning, and taxonomy. **(b)** Illustration of the categories used to compare MAGs with pb-MAG-mirror. Within a category, each box represents an individual MAG and the colored lines represent the contigs contained in the MAG. A full description of each category is provided in the Methods.

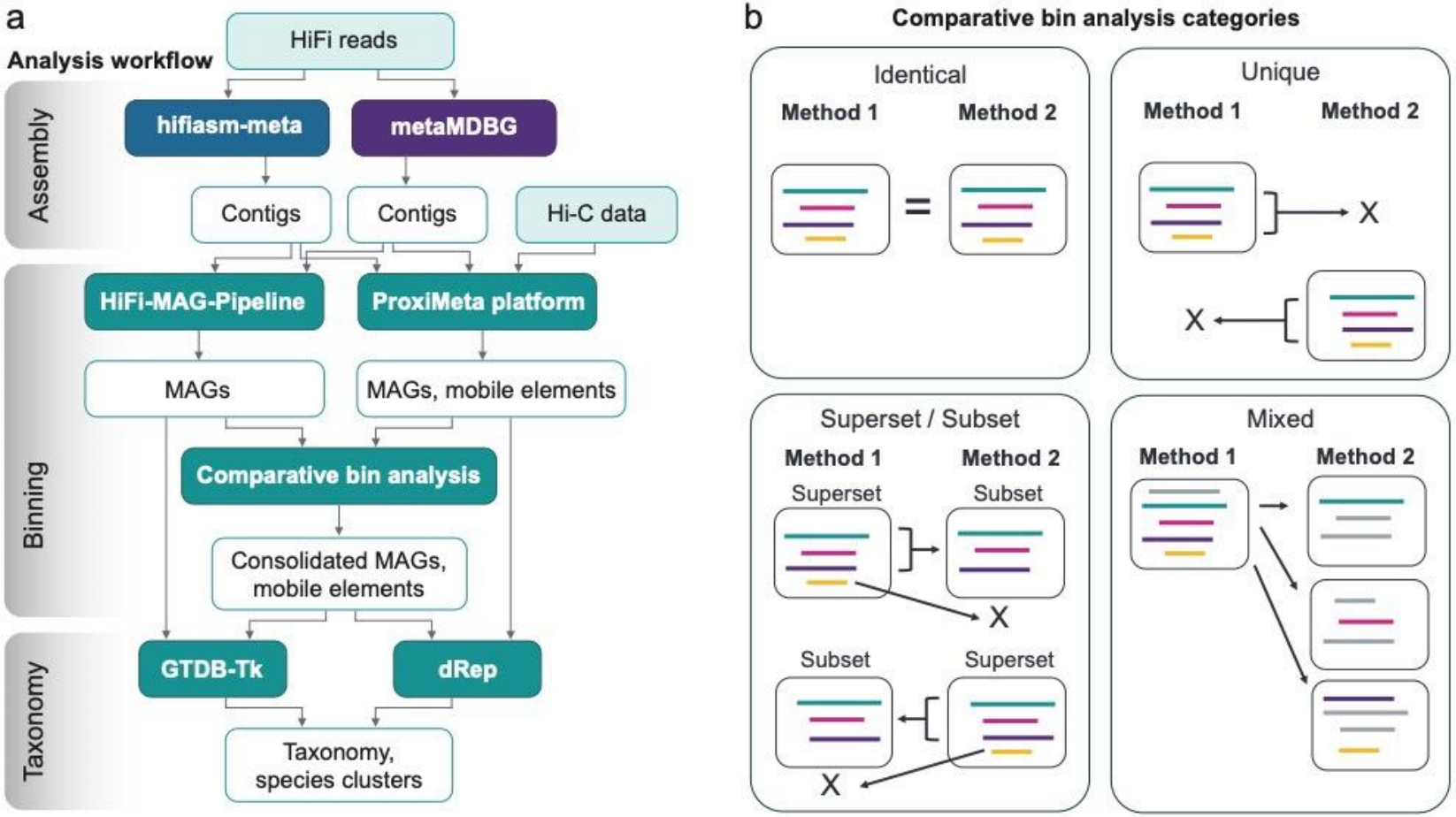

Figure 2. (a) Assignment of MAGs to categories from pb-MAG-mirror across the two assembly methods. Stacked barplot colors represent the number of MAGs occurring in different categories; total MAG numbers are shown on top. Bold outline highlights categories that are considered identical or nearly identical across methods. (**b)** Number of MAGs recovered for each assembly and binning method combination. HQ-MAGs require ≥90% single-copy genes (SCG) completeness and ≤5% contamination, whereas MQ-MAGs fall below the HQ thresholds but display ≥50% SCG completeness and ≤10% contamination. Total MAG numbers are shown above, with colors showing MAG counts exclusive to each category.

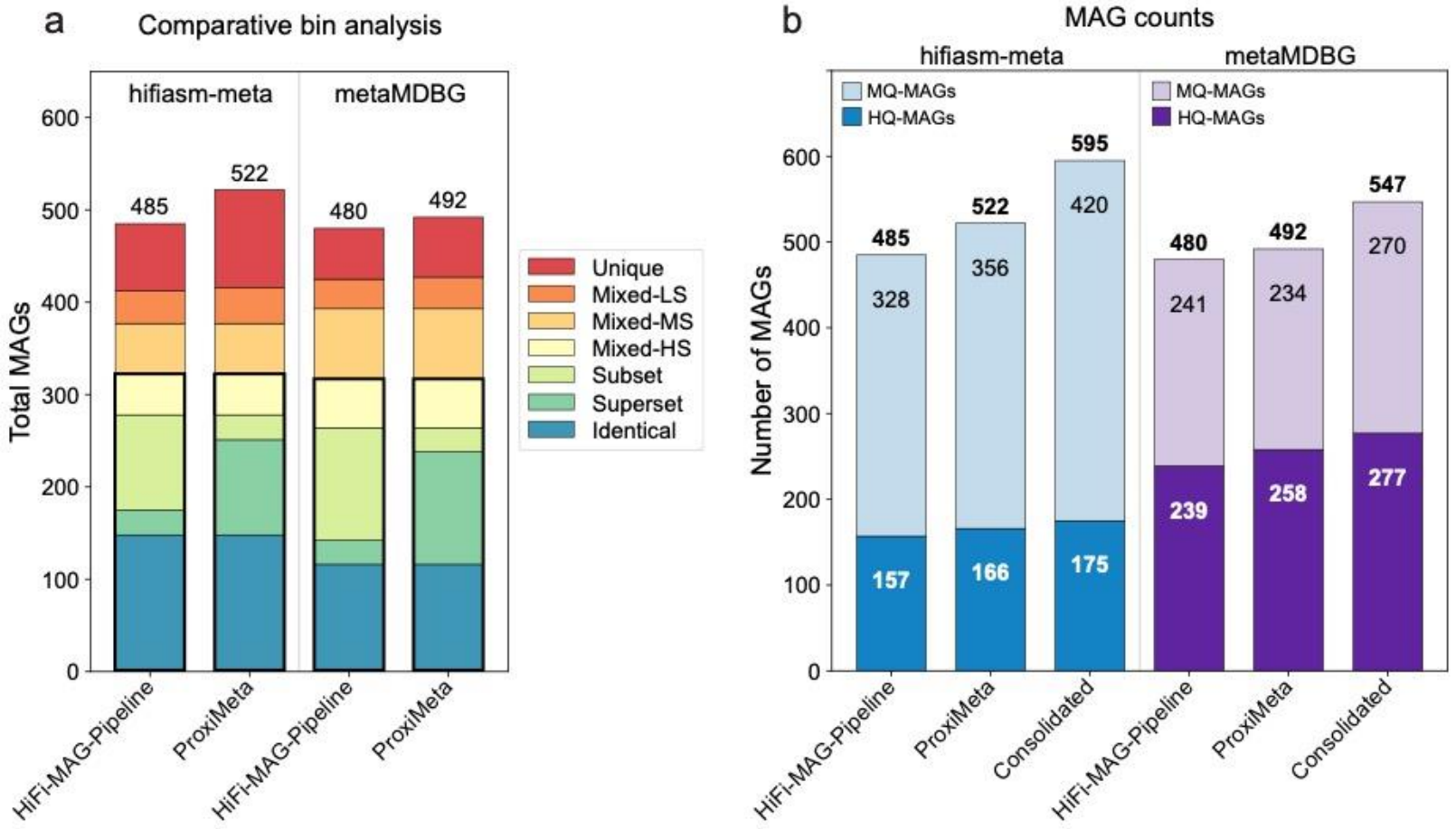

Figure 3. (a) Number of HQ-MAGs for the consolidated MAG set across downsampled datasets. (**b)** Number of HQ-MAGs obtained from hifiasm-meta versus metaMDBG across all binning methods and downsampled datasets. (**c)** Number of MQ-MAGs for the consolidated MAG set across downsampled datasets. (**d)** Number of MQ-MAGs obtained from hifiasm-meta versus metaMDBG across all binning methods and downsampled datasets.

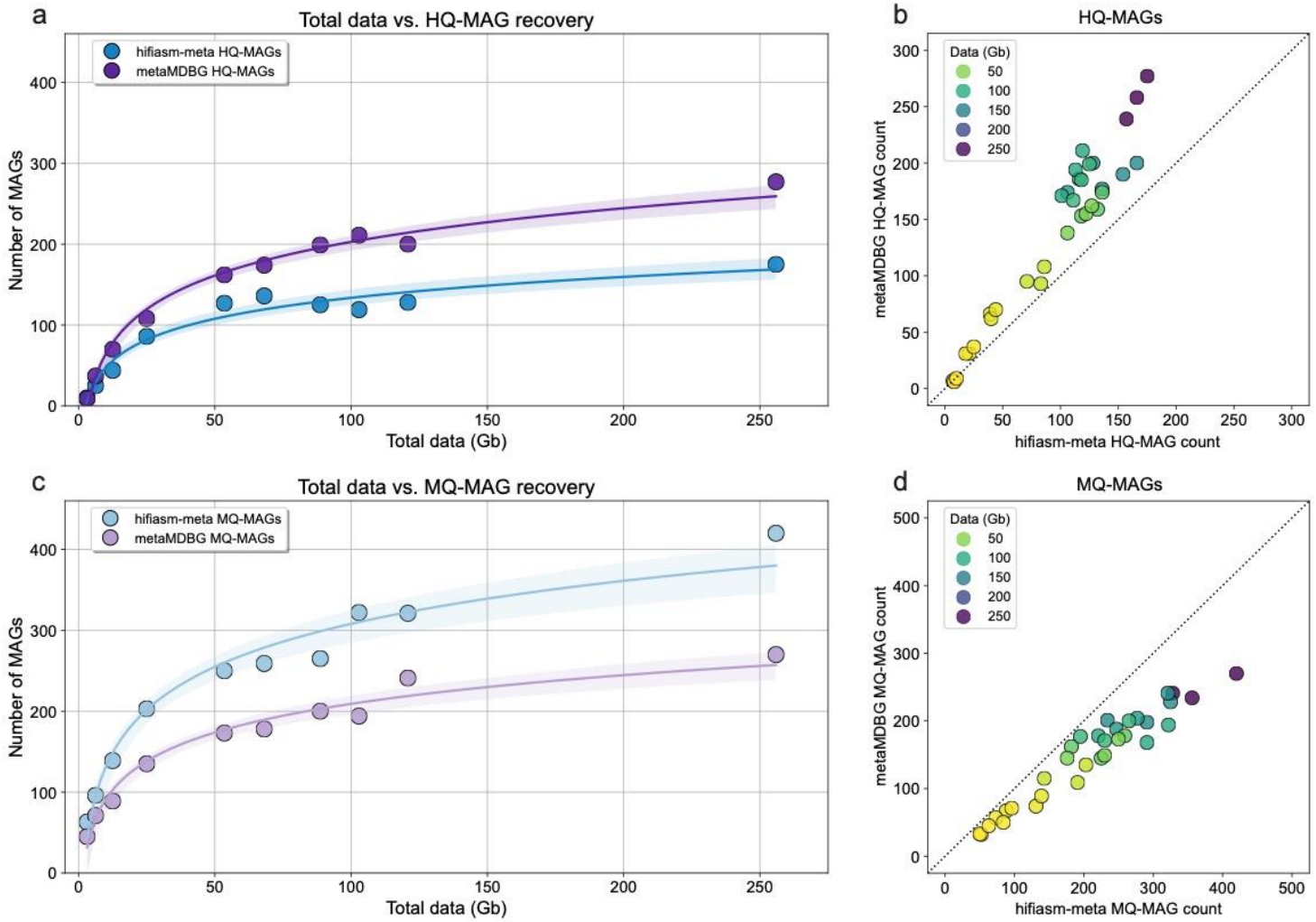

Figure 4. (a) Total number of species and species containing multiple strains, for each assembly and binning combination. Counts are based on dRep clustering results. (**b)** Aggregated taxonomic counts of MAGs assigned at the phylum rank.

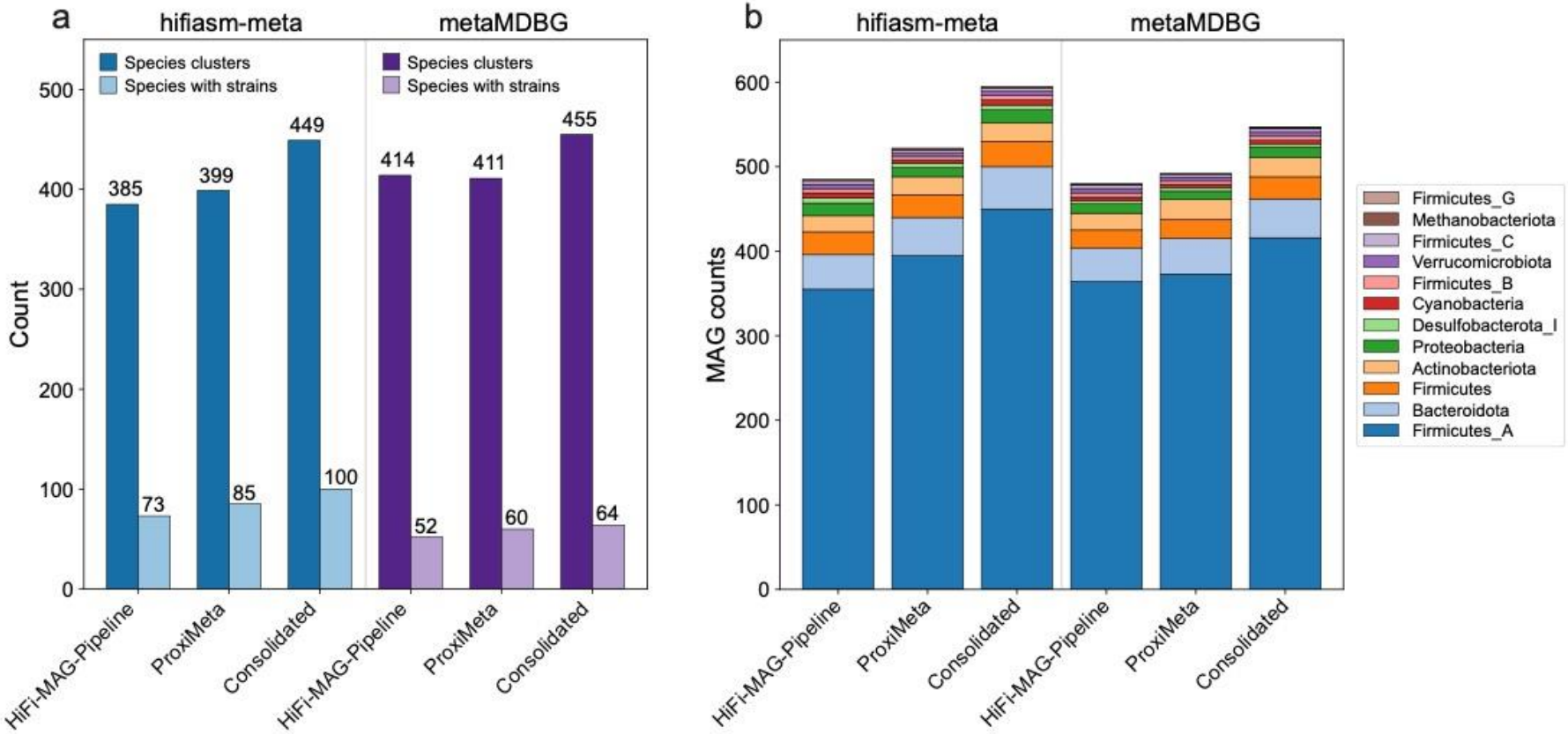

Figure 5. Barplots showing the 16S OTUs identified from the HiFi reads, based on 99% identity clustering. Relative abundance is represented by the height of the bars (log coverage). OTUs only occurring in the HiFi reads are shown in grey, whereas green indicates the OTU was also detected in a MAG. The MAG sets consist of the consolidated MAGs from both hifiasm-meta and metaMDBG, with results shown for **(a)** HQ-MAGs only or (**b)** all MQ- and HQ-MAGs.

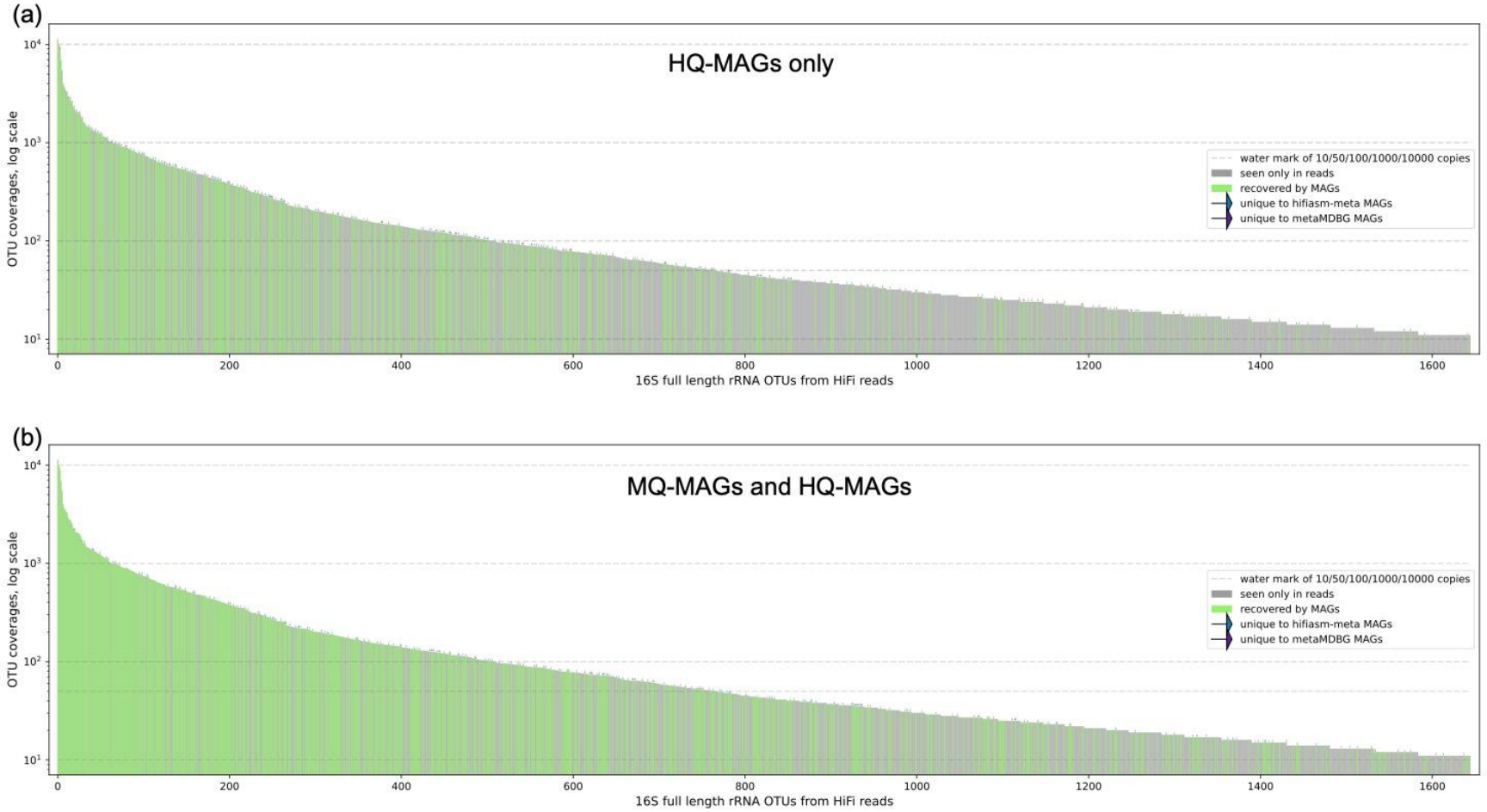

Figure 6. (a) Visual overview of workflow used to identify MAGs unequivocally shared across the assembly methods. Barplots summarizing (**b)** percent completeness, (**c)** average number of contigs per MAG, and (**b)** depth of coverage for MAGs in the shared (grey) and not shared categories (blue: hifiasm-meta, purple: metaMDBG). Outliers not shown.