

6 Supplementary Figures

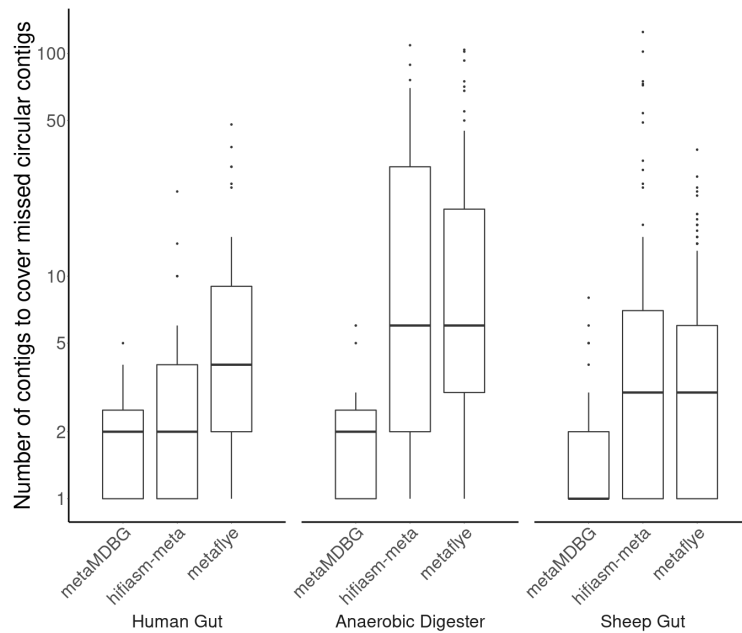


Figure S1: Number of contigs required to cover a near-complete circular MAG reconstructed successfully by an alternative assembler. In order to estimate the degree of fragmentation of assemblers, we aligned the contigs of one assembler against the near-complete circular contigs (cMAGs) recovered by the other assemblers. The fragmentation is then represented as the number of contigs required to cover these cMAGs (see section ‘Assessment of completeness and fragmentation of assemblies using reference sequences’ for details). The boxplot elements are the median (horizontal bar), 25th and 75th percentiles (box limits Q1 and Q3), $Q1-1.5*IQR$ and $Q3+1.5*IQR$ (whiskers, $IQR=Q3-Q1$) and outliers. Summary statistics (min, median, mean, max): Human— metaMDBG (1, 2, 2.1, 5); hifiasm-meta (1, 2, 4, 24); metaFlye (1, 4, 7.5, 48) : AD— metaMDBG (1, 2, 2.3, 6); hifiasm-meta (1, 6, 19.8, 109); metaFlye (1, 6, 15.1, 104) : Sheep— metaMDBG (1, 1, 1.8, 8); hifiasm-meta (1, 3, 10.7, 125); metaFlye (1, 3, 5, 37). The data to generate this boxplot have been extracted from the columns ‘AssemblyStatus’ of Supplementary Table S4.

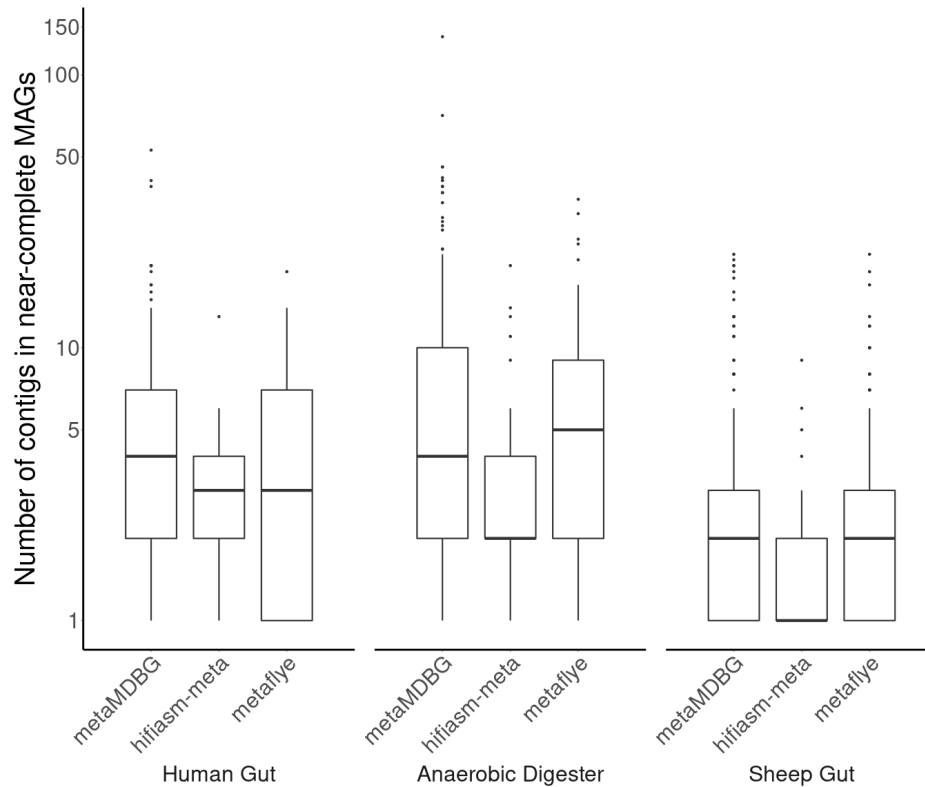


Figure S2: Number of contigs in non-circular near-complete MAGs. The boxplot elements are the median (horizontal bar), 25th and 75th percentiles (box limits Q1 and Q3), $Q1-1.5*IQR$ and $Q3+1.5*IQR$ (whiskers, $IQR=Q3-Q1$) and outliers. Summary statistics (min, median, mean, max): Human— metaMDBG (1, 4, 6.8, 53); hifiasm-meta (1, 3, 3.1, 13); metaFlye (1, 3, 4.6, 19) : AD— metaMDBG (1, 4, 9.2, 138); hifiasm-meta (1, 2, 3.6, 20); metaFlye (1, 5, 7, 35) : Sheep— metaMDBG (1, 2, 3.3, 22); hifiasm-meta (1, 1, 1.5, 9); metaFlye (1, 2, 2.8, 22). The number of contigs per non-circular MAGs have been extracted from columns 'AssemblyStatus' of supplementary File S1.

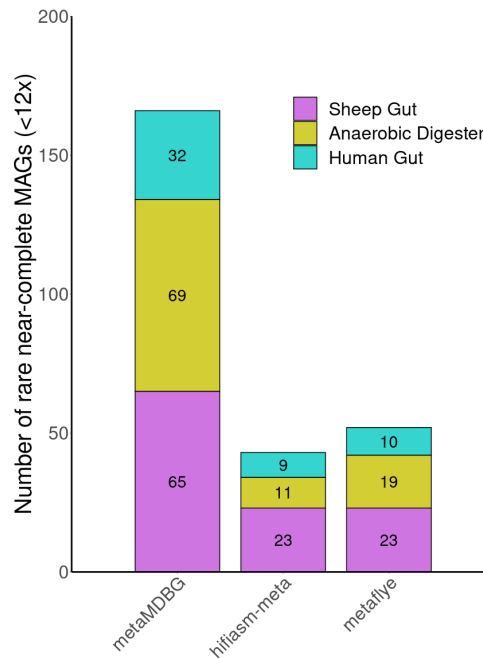


Figure S3: Number of low-coverage non-circular near-complete MAGs recovered by the assemblers. For the three tested datasets, we show the number of non-circular near-complete MAGs with low coverage (< 12x) reconstructed by each assembler.

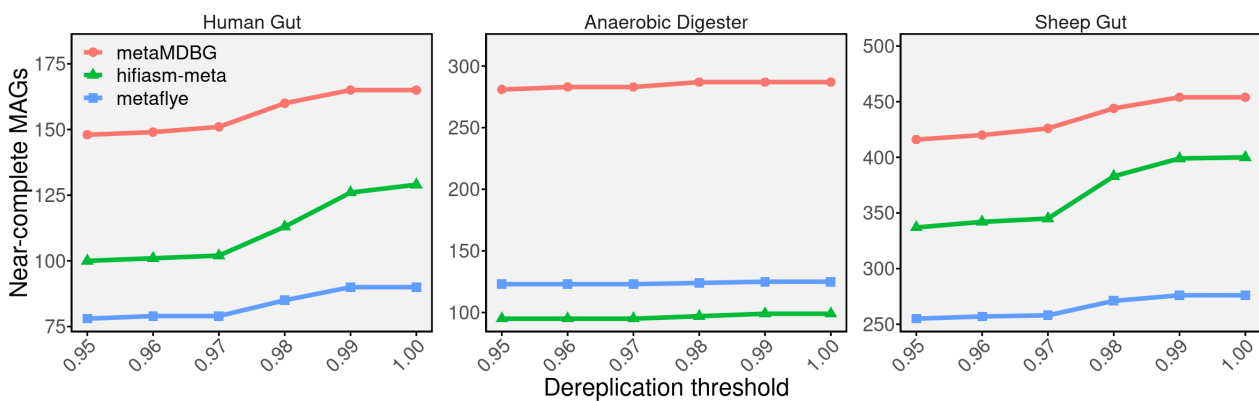


Figure S4: Total number of near-complete MAGs (circular and non-circular) across different dereplication thresholds. We used dRep [21] to cluster MAGs by nucleotide similarity using the parameter -sa from 0.95 to 1.