

Metagenome-assembled genomes from urban pigeon feces in Istanbul, Türkiye

İrem Kocakahya,¹ Gül Şahin,¹ Erdem Büyükkahraman,¹ Muzaffer Arkan¹

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT We report herein about 101 metagenome-assembled genomes (MAGs) obtained from pigeon fecal samples collected in 2025 from the Beyazıt, Kadıköy, and Beşiktaş squares of Istanbul. The MAGs were predominantly composed of members of the phyla *Firmicutes*, *Actinobacteria*, and *Proteobacteria*, with a lower representation of *Campylobacterota* and *Patescibacteriota*.

KEYWORDS pigeon, metagenomics, fecal microbiome, DNA sequencing, bacterial genome

Pigeons (*Columba livia domestica*) are ubiquitous in urban ecosystems and serve as important indicators of environmental microbial dynamics. Their feces can act as reservoirs and potential vectors for the dissemination of antibiotic resistance genes (ARGs), posing as a significant public health concern in densely populated cities. In this study, a total of 75 pigeon fecal samples were collected from three major public squares in Istanbul: Beyazıt (41.010266 N, 28.964197 E), Kadıköy (40.992596 N, 29.023987 E), and Beşiktaş (41.042404 N, 29.007358 E). Samples were aseptically transferred to sterile containers, transported on ice, and stored at –20°C until further processing. Total DNA was extracted using the Qiagen DNeasy PowerSoil Pro Kit as previously described (1). Five randomly selected DNA samples from each square were quantified using the Qubit dsDNA HS assay and pooled at equal DNA concentrations, resulting in a total of 15 pooled samples for sequencing. Sequencing libraries were prepared from 100 ng of purified DNA using the Illumina DNA Prep Kit and sequenced on the Illumina NovaSeq 6000 platform using paired-end 2 × 150 bp reads following the manufacturer's default protocol. Library construction failed for one sample, resulting in 14 metagenomic samples.

Unless otherwise specified, all bioinformatics tools were run using their default parameters recommended by the developers. Raw sequencing reads were subjected to quality control using FastQC (v0.12.1) (2), and then processed with Trimmomatic (v0.39) (3) to remove adapters and low-quality bases using the parameters: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10, LEADING:3, TRAILING:3, SLIDINGWINDOW:4:20, and MINLEN:36. The cleaned paired-end reads were mapped to the *Columba livia* reference genome using Minimap2 (v2.24) (4), and unmapped reads were retained using Samtools (v1.18) (5) to exclude host-derived sequences. Metagenomic assemblies were generated using metaSPAdes (v3.15.5) (6), and binning was performed in three stages using MetaBAT2 (7), MaxBin 2.0 (8), and CONCOCT (9), after which the resulting bins were integrated using DAS Tool (10) to produce a consensus set. Quality evaluation was conducted with CheckM2 (v1.0) (11), with bins meeting ≥90% completeness and ≤5% contamination retained as high-quality MAGs based on MIMAG recommendations (12). The final high-quality MAGs were classified taxonomically using the Genome Taxonomy Database (GTDB) (Release 10-RS226) (13).

Editor Frank J. Stewart, Montana State University, Bozeman, Montana, USA

Address correspondence to Muzaffer Arkan, muzafferarikan@istanbul.edu.tr.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 5 December 2025

Accepted 18 March 2026

Published 20 April 2026

Copyright © 2026 Kocakahya et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

The sequencing effort generated a total of 75 Gbp of data, averaging 5.4 Gbp per sample (approximately 18 million read pairs or 36 million reads per sample), enabling high-resolution metagenomic characterization of the urban pigeon fecal microbiome. Quality control analyses showed that, on average, 90% of bases per sample had a quality score of Q30 or higher. Assembly and binning yielded 101 high-quality MAGs, with a median completeness of 95.46% and a median contamination of 0.96% (see Table S1 at <https://doi.org/10.6084/m9.figshare.30801050>). GTDB-based taxonomic assignments revealed *Firmicutes*, *Actinobacteriota*, and *Proteobacteriota* as the dominant phyla, with additional representation from *Campylobacterota* and *Patescibacteriota* (see Table S1 at <https://doi.org/10.6084/m9.figshare.30801050>).

ACKNOWLEDGMENTS

This work was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK 2209A, Project No.: 1919B012420662) and Scientific Research Coordination Unit of Istanbul University (IU BAP, Project No.: 41481).

AUTHOR AFFILIATION

¹Biotechnology Division, Department of Biology, Faculty of Science, Istanbul University, Istanbul, Turkey

AUTHOR ORCIDs

Muzaffer Arıkan  <http://orcid.org/0000-0001-5162-2000>

FUNDING

Funder	Grant(s)	Author(s)
Scientific and Technological Research Council of Turkey	1919B012420662	İrem Kocakahya Gül Şahin Erdem Büyükkahraman
Scientific Research Coordination Unit of Istanbul University	41481	Muzaffer Arıkan

AUTHOR CONTRIBUTIONS

İrem Kocakahya, Formal analysis, Funding acquisition, Methodology, Writing – original draft | Gül Şahin, Formal analysis, Funding acquisition, Methodology, Writing – original draft | Erdem Büyükkahraman, Formal analysis, Funding acquisition, Methodology, Writing – original draft | Muzaffer Arıkan, Conceptualization, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

The shotgun sequence data have been deposited at the NCBI Sequence Read Archive under BioProject accession number [PRJNA1375659](https://doi.org/10.6084/m9.figshare.30801050). The recovered MAGs and the spreadsheet detailing characteristics of these metagenome-assembled genomes can be found on Figshare at <https://doi.org/10.6084/m9.figshare.30801050>.

REFERENCES

- Clasen F, Yildirim S, Arıkan M, Garcia-Guevara F, Hanoğlu L, Yılmaz NH, Şen A, Celik HK, Neslihan AA, Demir TK, Temel Z, Mardinoglu A, Moyes DL, Uhlen M, Shoaie S. 2025. Microbiome signatures of virulence in the oral-gut-brain axis influence Parkinson's disease and cognitive decline pathophysiology. *Gut Microbes* 17:2506843. <https://doi.org/10.1080/19490976.2025.2506843>
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>

4. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
5. Li Heng, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
6. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>
7. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. <https://doi.org/10.7717/peerj.7359>
8. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32:605–607. <https://doi.org/10.1093/bioinformatics/btv638>
9. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. *Nat Methods* 11:1144–1146. <https://doi.org/10.1038/nmeth.3103>
10. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* 3:836–843. <https://doi.org/10.1038/s41564-018-0171-1>
11. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2023. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 20:1203–1212. <https://doi.org/10.1038/s41592-023-01940-w>
12. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Elie-Fadrosh EA, et al. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 35:725–731. <https://doi.org/10.1038/nbt.3893>
13. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P-A, Hugenholtz P. 2022. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 50:D785–D794. <https://doi.org/10.1093/nar/gkab776>