Proceedings of the Eleventh International Conference on Urban Pests 2025
AF-Borgen, Academic Society, Lund, Sweden
William H Robinson, editor

PATHOGEN IDENTIFICATION IN NORWAY RAT (*RATTUS NORVEGICUS*) FAECES USING LONG-READ SEQUENCING OF BACTERIAL COMMUNITIES

[1]ANA CAROLINA YAMAKAWA, [2]MATTHEW DAVIES, [1,3]HYUN S. GWEON

[1] School of Biological Sciences, University of Reading, Reading, United Kingdom, RG6 6UR
[2] Technical Department, Killgerm Chemicals Ltd, Ossett, United Kingdom, WF5 9AJ
[3] UK Centre for Ecology & Hydrology, Wallingford, Oxfordshire, United Kingdom, OX10 8BB

Abstract Norway rats (*Rattus norvegicus*) have adapted to thrive in association with humans, making them frequent carriers of pathogens and a public health concern. Monitoring rodent populations for pathogens is crucial. Sequencing technologies provide an efficient and resource-effective method to characterise bacterial communities, including potential pathogens. While the 16S ribosomal RNA (rRNA) gene is commonly used for bacterial phylogenetics, its limited taxonomic resolution prompted interest in sequencing longer regions, such as the full *rrn* operon (16S-ITS-23S, excluding the 5S), which may offer enhanced resolution. This study aims to assess the taxonomic resolution of the 16S rRNA V4 region, the full-length 16S rRNA, and the 16S-ITS-23S operon in identifying pathogenic bacterial species in Norway rat droppings. *In silico* analysis was performed on 168 bacterial genomes from foodborne pathogen genera. Additionally, DNA was extracted from 64 rat droppings collected from various locations across the United Kingdom. Oxford Nanopore Technologies' (ONT) new R10.4.1 chemistry was employed to sequence the full-length 16S rRNA and 16S-ITS-23S operon, while Illumina technology was used for the 16S rRNA V4 region. Preliminary results suggest that sequencing longer regions (full-length 16S rRNA and 16S-ITS-23S operon) with ONT's improved R10.4.1 chemistry provides superior species-level resolution, enabling more accurate identification of potential pathogens and comprehensive microbiota profiling. This enhanced accuracy in identifying pathogens holds potential for improving public health monitoring and developing targeted interventions for rat-borne diseases, while also expanding our understanding of microbial diversity in urban, sub-urban and rural environments.

Key words rodent, disease, genomics, surveillance, pest control, one-health.

INTRODUCTION

As commensal species, Norway rats (*Rattus norvegicus*) have evolved to take advantage of human settlements, benefiting from food, water and shelter. Considering the escalating rates of urbanization (United Nations, 2018), coupled with the documented rise in rat population in the UK (Environment Agency, 2021), diseases transmitted by rodents can pose a growing public health threat.

Detection of pathogens can be performed by a wide number of techniques, from conventional techniques such as culturing, enzyme-linked immunosorbent assay (ELISA) and polymerase chain reaction (PCR), to cutting-edge approaches including biosensors and sequencing methodologies (Firth et al., 2014; Janda & Abbott, 2007; Lawson, 2004; Mocan et al., 2017; Overmann et al., 2017; Relman, 1993; Williams et al., 2018). The early detection and

management of possible zoonotic diseases in animals, is the most effective approached to avoid outbreaks, life losses, and reduce the economic burden (Bird & Mazet, 2018; Bisson et al., 2015). In addition, the surveillance of diseases and an accurate diagnosis in possible animal's reservoirs can significantly reduce the costs of a further outbreak (World Bank, 2012).

The bacterial 16S ribosomal RNA (rRNA) gene, or SSU (small subunit) ribosomal RNA, has been widely used as genetic marker for bacterial phylogenetics and taxonomic of microbial communities (Janda & Abbott, 2007; Yarza et al., 2014). Universal in all bacteria, it includes nine hypervariable regions (V1 to V9) divided by preserved areas, with more than one operon may coexist within a genome, exhibiting variations even in the same bacterial genome (de Oliveira Martins et al., 2020; Větrovský & Baldrian, 2013).

Along with the SSU, the entire rrn operon (16S-ITS-23S, excluding the 5S; ~4.5 Kb) encompasses a significant portion of taxonomic diversity, and your potential have been explored for microbiome studies (Kinoshita et al., 2021; Petrone et al., 2023). The combination of 16S-ITS-23S operon can present four times more variability when compared to only the 16S region (Benítez-Páez & Sanz, 2017). *In silico* and experimental analyses of the entire rrn operon, have also highlighted an improvement in taxonomic resolution when using longer sequences, achieving specie and even strain levels (de Oliveira Martins et al., 2020; Kinoshita et al., 2021; Petrone et al., 2023).

The Illumina next-generation sequencing technology is one of the mostly used platform for 16S rRNA sequencing. However, its short-read limitations impede the fully sequence of the entire gene, which typically spans approximately 1500 bp (Johnson et al., 2019; Yarza et al., 2014). Consequently, short-read sequencing is generally used for sequencing one or more 16S regions, which is considered to be insufficient for accurately classifying sequences at lower taxonomic resolution, such as species and genera (Schloss et al., 2016).

In contrast, long-reads sequencing approaches, such as the Oxford Nanopore Technologies (ONT), provides the capability to sequence the full-length of the 16S rRNA gene (Zhang et al., 2023). While ONT has been known for its high error rate of up to 25% (Buck et al., 2017; Jain et al., 2017; Laver et al., 2015; Rang et al., 2018; Van Dijk et al., 2018; Wick et al., 2018), which previously limited its application in microbiome research, advancements in flow cell chemistry over time resulted in improved performance (Jain et al., 2017; Rang et al., 2018; Sereika et al., 2022). More recently, the ONT R10.4.1 technology, through the use of the Q20+ chemistry and the SQK-LSK114 kit, can achieve a reads accuracy of approximately 99%.

With recent improvements in ONT long-read sequencing, its use for sequencing the full-length 16S and even the 16S-ITS-23S operon has starting to be explored for microbiome-focused studies (de Oliveira Martins et al., 2020; Kinoshita et al., 2021; Seol et al., 2022). Meanwhile, its application for pathogen identification remains undocumented. To evaluate the potential use and taxonomic resolution of the V4 region, full-length 16S and the 16S-ITS-23S operon for identifying pathogens species, or even strain, an *in silico* analyses was performed. Additionally, the applicability of ONT R10.4.1, sequencing of the full-length 16S and the 16S-ITS-23S, as well as Illumina sequencing of the V4 region, for microbial characterisation and pathogen screening in rat faecal samples, is being accessed in this study.

MATERIAL AND METHODS

In silico analyses. To compare the use of the 16S V4 region, the full-length 16S, and the operon 16S-ITS-23S in identifying species, an *in silico* analysis was performed. We have selected 168 bacterial genomes from common foodborne pathogens genera, that can be potentially present in

Norway rat droppings. Including seven *Bacillus cereus* genomes strains, one *B. anthracis*, 49 genomes from 14 *Campylobacter* species, 14 genomes from five *Clostridium* species, 41 genomes from 20 *Escherichia coli* serovars, eight genomes from four *Listeria* species, five genomes from 3 *Proteus* species, 25 genomes from 18 *Salmonella enterica* serovars and 17 genomes from four *Shigella* species.

The genomes were downloaded from the RefSeq database (accessed and downloaded on October 2023), and analysed using the software Geneious (v2023.2.1). Target sequences (16S V4 region, the full-length 16S, and the operon 16S-ITS-23S) were extracted from the genomes, after alignments using the primers described in the experiments methods below. Alignments were performed within the same genera and between 50 sequences from 26 species, through the function MUSCLE 5.1 with default settings. Information regarding pairwise identity was calculated plotted using R (v 4.2.1) with package ggplot2 (v 3.4.3).

Samples and DNA extraction. Norway rat (*R. norvegicus*) faecal samples were collected from different regions and environments in the UK by pest controllers (figure 1). The droppings were stored in conical tubes (15 mL) containing 2/3 of volume of silica beads, and if same day postage wasn't possible, they were stored at -20°C. The silica beads remove the humidity from the samples, improving preservation during transport (Allison et al., 2021). The DNA were extracted using the HigherPurity™ Stool DNA Isolation Kit (Canvax Biotech), following the manufacturers protocol. The extraction was performed on a single dropping per location/conical tube and the extracted DNA were stored at -20°C until process.

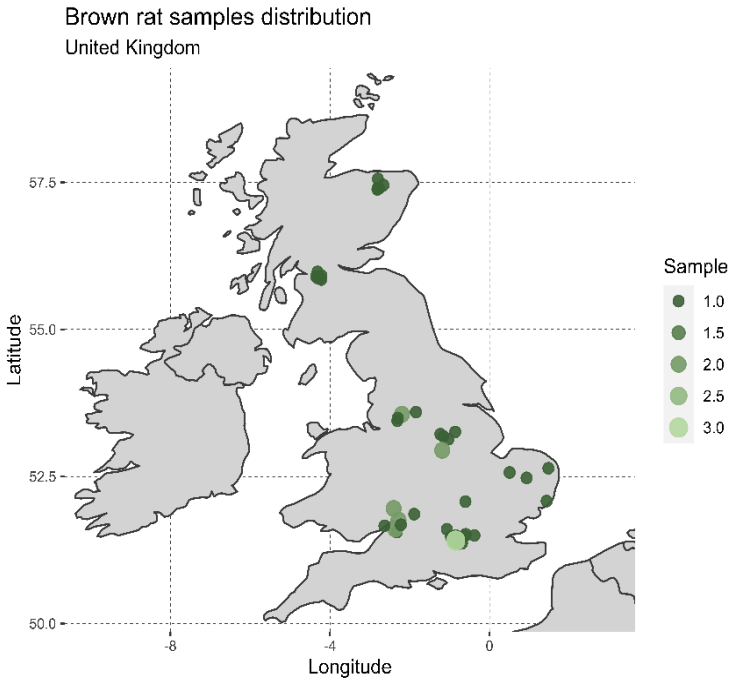

Figure 1. Map of sampling points distribution across the United Kingdom.

Nanopore R10.4 sequencing. For the long-read sequencing a single PCR was performed to amplify the target and attach the barcodes. To amplify the full-length 16S, a PCR was conducted using the pair of primers: 16S_27F (forward: barcode+AGRGTTYGATYMTGGCTCAG) and

16S_1391R (reverse: barcode+GACGGGCGGTGWGTRCA) (Stackebrandt E & Goodfellow M, 1991). For the 16S-ITS-23S region, the same forward primer was used with the reverse primer: 23S_2490R (reverse: barcode+CGACATCGAGGTGCCAAAC) (Hunt et al., 2006).

The PCR for the full-length 16S was performed with 0.4 µL of Platinum SuperFi II (Thermo Fisher Scientific Inc., MA, USA), 4 µL of SuperFi II PCR buffer (Thermo Fisher Scientific Inc., MA, USA), 0.4 µL of dNTPs mix (10 nM), 1 µL of each primer (10 nM), 12.2 µL of nuclease-free water and 1 µL of sample, totalizing 20 µL. The thermocycle conditions were a first denaturation of 98℃ for 3 min, followed by 30 cycles of 98℃ for 30 seg, 60℃ for 30 seg, 72℃ for 2 min, and a final extension of 72℃ for 5 min.

To amplified the 16S-ITS-23S operon the PCR was performed with the following reagents: 0.4 µL of Platinum SuperFi II (Thermo Fisher Scientific Inc., MA, USA), 4 µL of SuperFi II PCR buffer (Thermo Fisher Scientific Inc., MA, USA), 0.4 µL of dNTPs mix (10 nM), 1.5 µL of each primer (10 nM), 11.2 µL of nuclease-free water and 1 µL of sample, totalizing 20 µL. The amplification cycle condition was 98℃ for 3 min, followed by 30 cycles of 98℃ 45 seg, 60℃ for 45 seg, 72℃ for 3 min, and a final extension of 72℃ for 6 min.

A high-fidelity polymerase was used to minimize amplification errors during the process. Both PCR amplification were confirmed through electrophoresis, and subsequently, a purification step using magnetic beads (1x v/v) was carried out. Samples were quality checked with a DeNovix DS-11 spectrophotometer (DeNovix Inc, DE, USA) and quantified with a Qubit™ dsDNA high sensitivity (Thermo Fisher Scientific Inc., MA, USA). Up to 6 samples were pooled, and 50-100 fmol used for library preparation with the Ligation Sequencing Kit V14 (SQK-LSK114). Each library was sequenced using Flongle flow cells (R10.4.1) in a MinION Mk1C up to pores were exhausted.

Illumina sequencing. The V4 region of the 16S rRNA gene was amplified with the following pair of primers: 515F (forward: GTGYCAGCMGCCGCGGTAA) (Parada et al., 2016) and 806R (reverse: GGACTACNVGGGTWTCTAAT) (Apprill et al., 2015). The reaction consisted of 12.5 µL of JumpStart™ Taq ReadyMix™ (Sigma-Aldrich), 0.5 µL of each primer (10 µM), 8.5 µL of Nuclease-free water and 3 µL of DNA. The cycle conditions were a first denaturation of 94°C for 3 minutes, following 30 cycles of 94°C/45 seconds, 50°C/60 s, 72°C/90s, and a final extension of 72°C for 10 min. The PCR products (~300 bp) were visualised on a 1% agarose gel with references ladders, after electrophoresis.

The PCR product were clean with magnetic beads in a 1.2x proportion. A second PCR was performed to add custom barcodes sequences on the amplicons both ends. The reaction mix consisted of 25 µL of JumpStart™ Taq ReadyMix™ (Sigma-Aldrich), 10 µL of each primer (10 µM), 10 µL of Nuclease-free water and 5 µL of PCR product. The cycle conditions were a first denaturation of 95°C for 2 minutes, following by 8 cycles of 95°C/15 seconds, 55°C/30 s, 72°C/30s, and a final extension of 72°C for 10 min. After electrophoresis the PCR products were visualised on a 1% agarose gel with references ladders to confirm the barcodes attachment. Products were clean with magnetic beads and quality was assessed using the DeNovix DS-11 FX + spectrophotometer (DeNovix Inc., DE, USA) and quantified with a Qubit Fluorometers (Thermo Fisher Scientific Inc., MA, USA). The sequencing was performed in a MiSeq (Illumina, San Diego, CA).

Sequences analysis and classification. Nanopore sequencing data were basecalled using Dorado v0.8.1 on the super high accuracy basecalling model “SUP”. Resulting fastq files were

quality filtered for ≥Q15 reads, which represents ~3% of error rates, using nanoq v0.10.0. and chimeras sequences were removed with yacrd 1.0.0 (Marijon et al., 2020). Filtered reads were demultiplexed with seqdemu (https://github.com/hsgweon/seqdemu) allowing one mismatch and with now mismatch. Primers were removed from sequences ends using cutadapt v4.9 (Martin, 2011). Reads were filtered by length using SeqKit v2.8.2. (Shen et al., 2024), within an interval of 1,300 to 1,950 for the full length 16S and 3,500 to 5,000 for the 16S-ITS-23S. To enable diversity comparison between samples and targets, a random subsampling of 10,000 reads per sample was performed using vsearch v2.21.1 (Rognes et al., 2016).
The Illumina short-read sequences were quality-filtered and demultiplexed using TrimGalore (https://github.com/FelixKrueger/TrimGalore). For both taxonomic classification and relative abundance estimation Emu v3.5.0 (Curry et al., 2022) was used, along with the GROND databases (Walsh et al., 2024). The databases were formatted for Emu. Resulting phylogenic files are being process using R (v 4.2.1) with packages Vegan (v2.6-4) and Phyloseq (v 1.42.0), and for data visualisation the ggplot2 (v 3.4.3). Analyses are still on performance.

RESULTS AND DISCUSSION

***In silico* analyses.** For the analyses comparing species within the same genera for the V4 region, six of the analysed genera presented more than 99.7% of pairwise identity for this target, with *Bacillus* spp. and *Listeria* spp. achieving 100%. The pairwise identity represents how much the analyse sequences are similar. So, if the pairwise identity is 100%, the sequences analysed present the same sequence of nucleotides. The full-length 16S produced similar results, with at least 99.3% of identity for six genera. *Clostridium* spp. presented the smallest (92.2%) identity within the same genera. Alignments for the 16S-ITS-23S had a similarity under 98%, with maximum of 97.9% for *Listeria* spp. and a minimum of 76.9% for *Campylobacter* spp. The use of the operon presented the lowest pairwise identity, consequently, the best results to distinguish species and even strains.

The analyses of sequences between different genera for the three target (16SV4, full-length 16S and 16S-ITS-23S) presented similarities under 90%, confirming that all targets would be enough to differentiate between genera level for the select bacteria. As expected, the 16S-ITS-23S operon presented the smallest pairwise identity average with 67.2%, followed by the full-length 16S with 83.3% and 16S V4 region with 87.2%. Comparisons between pairwise identity for the three targets across the analyses genera and species are represented in Figure 2.

In silico analyses of the target regions 16SV4, full length 16S and 16S-ITS-23S indicates that the 16S-ITS-23S is the best target to distinguish species level for most of the analysed genera and species. The results are in agreeing with that larger gene region will result in a better taxonomic resolution (de Oliveira Martins et al., 2020; Drengenes et al., 2021; Yarza et al., 2014), and with previous researches focusing on rrn operon for microbiome (de Oliveira Martins et al., 2020; Kinoshita et al., 2021; Petrone et al., 2023).

Long and short sequencing. DNA was extracted from 64 samples of droppings. The full-length 16S sequencing was successfully performed on 60 samples using 14 Flongle flow cells, resulting in a total of 2,712,913 reads with an average quality score of Q19. Regarding the 16S-ITS-23S operon, a total of 56 samples were sequenced using 15 Flongle flow cells, generating 3,710,192 reads with an average quality score of Q19. After filtering and removing samples with less than 10,000 reads, 1,255,987 reads remained for the full length 16S and 1,047,576 reads for the 16S-

ITS-23S operon. Additionally, the V4 region of all 64 samples were sequenced using the Illumina technology, producing a total of 1,544,941 reads.

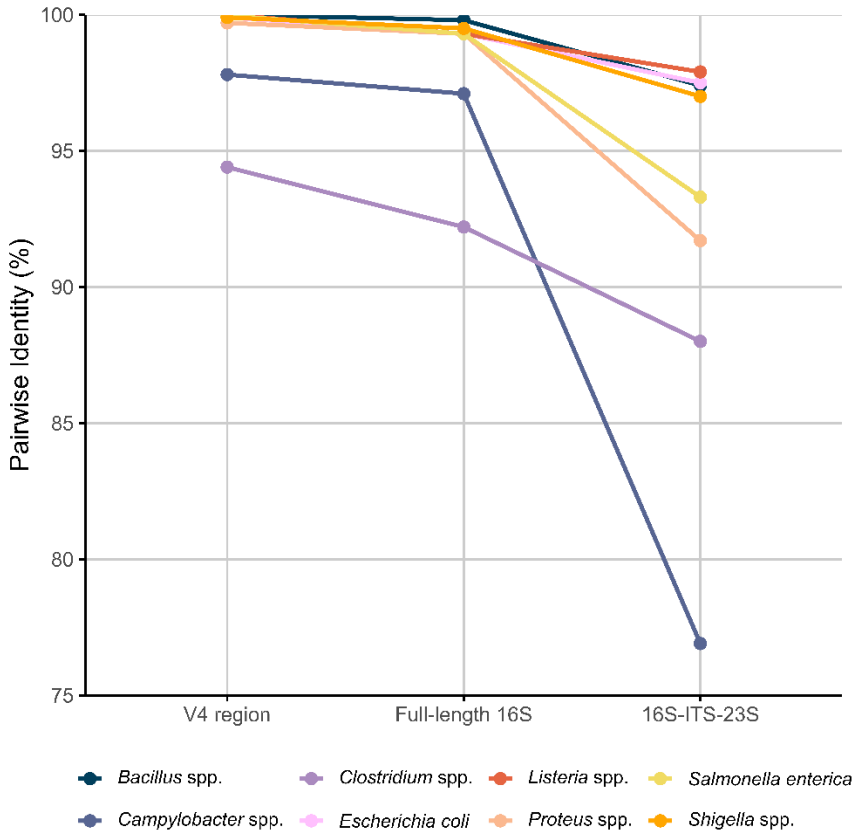

Figure 2. Line graph comparing the pairwise identity (%) between V4 region, full-length 16S and the operon 16S-ITS-23S, within different species from the same genera. Excluding the *S. enterica* and *E. coli*, which are comparing different serovars.

Further taxonomic analyses are currently in progress. However, the remarkable read accuracy achieved with the long-read sequencing is promising. While ONT has been associated with error rates ranging from 5% to 38%, limiting its suitability for microbiome and diversity analyses (Jain et al., 2017; Laver et al., 2015; Van Dijk et al., 2018), recent advancements in ONT chemicals and flow cells have significantly improved performance (Jain et al., 2017; Sereika et al., 2022). The current ONT R10.4.1 flow cells, combined with Q20+ chemistry, can achieve a model read accuracy of approximately 99% for whole-genome sequencing. In this study, we achieved an average of Q19, representing a read accuracy of approximately 98.74%, highlighting the potential of these advancements. Further bioinformatic analyses are ongoing.

Notably, ONT R10.4.1 flow cells for full-length 16S sequencing exhibit better species identification than PacBio sequencing (T. Zhang et al., 2023). Sequencing longer targets, such as the rrn operon, has also proven advantageous in taxonomic resolution, offering more accurate classifications at lower levels (Karst et al., 2021; Kinoshita et al., 2021; Petrone et al., 2023; Shin et al., 2016). Focusing on the 16S-ITS-23S operon sequences enhances the exploration of taxonomic diversity, presenting four times more variability than the 16S region alone (Benítez-

Páez & Sanz, 2017). Both, *in silico* analyses, including this study, and experimental analyses have confirmed improved taxonomic resolution with longer sequences, reaching species and even strain levels (de Oliveira Martins et al., 2020; Kinoshita et al., 2021; Petrone et al., 2023).
One limitation of this study is the use of only droppings to estimate microbiota composition. Bacterial composition can vary between different areas from the gastrointestinal tract (D. Li et al., 2017; Matsumoto et al., 2023; Staubach et al., 2012; Suzuki & Nachman, 2016), however, faeces have been suggested to be a good representative of the lower gastrointestinal microbiota (Suzuki & Nachman, 2016). Furthermore, samples with a higher bacterial load seems to be less influence by protocol bias, such as number of PCR steps or contamination from laboratory reagents, when compared to low biomass samples (Drengenes et al., 2019, 2021; Pollock et al., 2018; Salter et al., 2014).

Another limitation is the potential presence of intragenomic variation within 16S-ITS-23S operons. Pei et al. (2009) analysed 184 prokaryotic genomes and identified intragenomic variations in the 23S in 61.4%. However, around 4% of the genomes had a significant intragenomic variation (approximately 1.17 – 4.04%) between 23S regions, suggesting a minimal overall impact (Pei et al., 2009). A further limitation is the need of an accurate and comprehensive 16S-ITS-23S database. As the use of the rrn operon for microbiome analysis is relatively recent, only a few databases are currently available (Curry et al., 2022; Kinoshita et al., 2021; Seol et al., 2022; Walsh et al., 2024). However, efforts are being made to enhance and maintain these resources, with two databases actively compromised to incorporate new releases (Seol et al., 2022; Walsh et al., 2024).

CONCLUSION

Norway rats are well adapted to live in human environments and often act as reservoirs for pathogens, posing a threat for public health. Effective monitoring strategies for pathogens related risks are essential. The possibility to screen bacterial communities and identify pathogenic species within a sample (e.g. droppings), is highly promising. Herein, we present an efficient and early-stage protocol that compares three potential target gene regions (16SV4, full length 16S and 16S-ITS-23S) in their potential for confirming bacterial taxonomic with accuracy, aiming to achieve the species level.

In silico analyses indicates that the 16S-ITS-23S operon is the best target for distinguishing species and strains across the analysed genera and species. Preliminary results suggest that sequencing the full-length 16S and the 16S-ITS-23S operon with ONT's improved R10.4.1 chemistry generates high quality sequences, enabling more accurate pathogen identification and comprehensive microbiota profiling. Bioinformatics analyses are ongoing to ensure the reliability of this pipeline for pathogens identification, with a focus on improving public health monitoring and developing targeted interventions for rat-borne diseases.

REFERENCES CITED

Allison, M. J., Round, J. M., Bergman, L. C., Mirabzadeh, A., Allen, H., Weir, A., and Helbing, C. C. 2021. The effect of silica desiccation under different storage conditions on filter-immobilized environmental DNA. BMC Research Notes, 14(1). https://doi.org/10.1186/s13104-021-05530-x

Apprill, A., Mcnally, S., Parsons, R., and Weber, L. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquatic Microbial Ecology, 75(2), 129–137. https://doi.org/10.3354/AME01753

Benítez-Páez, A., and Sanz, Y. 2017. Multi-locus and long amplicon sequencing approach to study microbial diversity at species level using the MinIONTM portable nanopore sequencer. GigaScience, 6(7), 1–12. https://doi.org/10.1093/gigascience/gix043

Bird, B. H., and Mazet, J. A. K. 2018. Detection of Emerging Zoonotic Pathogens: An Integrated One Health Approach. Annual Review of Animal Biosciences, 6, 121–139. https://doi.org/10.1146/ANNUREV-ANIMAL-030117-014628

Bisson, I.-A., Ssebide, B. J., and Marra, P. P. 2015. Early Detection of Emerging Zoonotic Diseases with Animal Morbidity and Mortality Monitoring. EcoHealth, 12(1). https://doi.org/10.1007/s10393-014-0988-x

Buck, D., Weirather, J. L., de Cesare, M., Wang, Y., Piazza, P., Sebastiano, V., Wang, X. J., and Au, K. F. 2017. Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. F1000Research, 6(100). https://doi.org/10.12688/F1000RESEARCH.10571.2/DOI

Curry, K. D., Wang, Q., Nute, M. G., Tyshaieva, A., Reeves, E., Soriano, S., Wu, Q., Graeber, E., Finzer, P., Mendling, W., Savidge, T., Villapol, S., Dilthey, A., and Treangen, T. J. 2022. Emu: species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. Nature Methods, 19(7), 845–853. https://doi.org/10.1038/s41592-022-01520-4

de Oliveira Martins, L., Page, A. J., Mather, A. E., and Charles, I. G. 2020. Taxonomic resolution of the ribosomal RNA operon in bacteria: implications for its use with long-read sequencing. NAR Genomics and Bioinformatics, 2(1), 1–7. https://doi.org/10.1093/nargab/lqz016

Environment Agency, C. S. Group. **2021**. The state of environment: the urban environment.

Firth, C., Bhat, M., Firth, M. A., Williams, S. H., Frye, M. J., Simmonds, P., Conte, J. M., Ng, J., Garcia, J., Bhuva, N. P., Lee, B., Che, X., Quan, P. L., and Ian Lipkin, W. (2014). Detection of zoonotic pathogens and characterization of novel viruses carried by commensal rattus norvegicus in New York city. MBio, 5(5). https://doi.org/10.1128/MBIO.01933-14/FORMAT/EPUB

Hunt, D. E., Klepac-Ceraj, V., Acinas, S. G., Gautier, C., Bertilsson, S., and Polz, M. F. 2006. Evaluation of 23S rRNA PCR primers for use in phylogenetic studies of bacterial diversity. Applied and Environmental Microbiology, 72(3), 2221–2225. https://doi.org/10.1128/AEM.72.3.2221-2225.2006

Jain, M., Tyson, J. R., Loose, M., Ip, C. L. C., Eccles, D. A., O'grady, J., Malla, S., Leggett, R. M., Wallerman, O., Jansen, H. J., Zalunin, V., Birney, E., Brown, B. L., Snutch,

T. P., and Olsen, H. E. 2017. MinION Analysis and Reference Consortium: Phase 2 data release and analysis of R9.0 chemistry. F1000Research, 6(760). https://doi.org/10.12688/f1000research.11354.1

Janda, J. M., and Abbott, S. L. 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. In Journal of Clinical Microbiology (Vol. 45, Issue 9, pp. 2761–2764). https://doi.org/10.1128/JCM.01228-07

Johnson, J. S., Spakowicz, D. J., Hong, B. Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., and Weinstock, G. M. 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nature Communications, 10(1). https://doi.org/10.1038/s41467-019-13036-1

Kinoshita, Y., Niwa, H., Uchida-Fujii, E., and Nukada, T. 2021. Establishment and assessment of an amplicon sequencing method targeting the 16S-ITS-23S rRNA operon for analysis of the equine gut microbiome. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-91425-7

Laver, T., Harrison, J., O'neill, P. A., Moore, K., Farbos, A., Paszkiewicz, K., and Studholme, D. J. 2015. Assessing the performance of the Oxford Nanopore Technologies MinION. Biomolecular Detection and Quantification, 3, 1–8. https://doi.org/10.1016/j.bdq.2015.02.001

Lawson, A. J. 2004. Discovering new pathogens: culture-resistant bacteria. Methods in Molecular Biology , 266, 305–322. https://doi.org/10.1385/1-59259-763-7:305

Marijon, P., Chikhi, R., & Varré, J. S. (2020). yacrd and fpa: upstream tools for long-read genome assembly. Bioinformatics, 36(12), 3894–3896. https://doi.org/10.1093/BIOINFORMATICS/BTAA262

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal, 17(1), 10–12. https://doi.org/10.14806/EJ.17.1.200

Mocan, T., Matea, C. T., Pop, T., Mosteanu, O., Buzoianu, A. D., Puia, C., Iancu, C., and Mocan, L. (2017). Development of nanoparticle-based optical sensors for pathogenic bacterial detection. In Journal of Nanobiotechnology (Vol. 15, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12951-017-0260-y

Overmann, J., Abt, B., and Sikorski, J. 2017. Present and Future of Culturing Bacteria. Annual Review of Microbiology, 71, 711–730. https://doi.org/10.1146/ANNUREV-MICRO-090816-093449

Parada, A. E., Needham, D. M., and Fuhrman, J. A. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environmental Microbiology, 18(5), 1403–1414. https://doi.org/10.1111/1462-2920.13023

Petrone, J. R., Rios Glusberger, P., George, C. D., Milletich, P. L., Ahrens, A. P., Roesch, L. F. W., and Triplett, E. W. 2023. RESCUE: a validated Nanopore pipeline to classify bacteria through long-read, 16S-ITS-23S rRNA sequencing. Frontiers in Microbiology, 14. https://doi.org/10.3389/fmicb.2023.1201064

Rang, F. J., Kloosterman, W. P., and de Ridder, J. 2018. From squiggle to basepair: Computational approaches for improving nanopore sequencing read accuracy. In Genome Biology (Vol. 19, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13059-018-1462-9

Relman, D. A. 1993. The Identification of Uncultured Microbial Pathogens. In Source: The Journal of Infectious Diseases (Vol. 168, Issue 1). https://about.jstor.org/terms

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. 2016. VSEARCH: A versatile open source tool for metagenomics. PeerJ, 2016(10), e2584. https://doi.org/10.7717/PEERJ.2584/FIG-7

Schloss, P. D., Jenior, M. L., Koumpouras, C. C., Westcott, S. L., and Highlander, S. K. 2016. Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. PeerJ, 2016(3). https://doi.org/10.7717/peerj.1869

Seol, D., Lim, J. S., Sung, S., Lee, Y. H., Jeong, M., Cho, S., Kwak, W., and Kim, H. 2022. Microbial Identification Using rRNA Operon Region: Database and Tool for Metataxonomics with Long-Read Sequence. Microbiology Spectrum, 10(2). https://doi.org/10.1128/SPECTRUM.02017-21/SUPPL_FILE/SPECTRUM02017-21_SUPP_2_SEQ10.XLSX

Sereika, M., Hansen Kirkegaard, R., Michael Karst, S., Yssing Michaelsen, T., Aarre Sørensen, E., Dam Wollenberg, R., and Albertsen, M. 2022. Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. Nature Methods, 19, 823–826. https://doi.org/10.1038/s41592-022-01539-7

Shen, W., Sipos, B., and Zhao, L. 2024. SeqKit2: A Swiss army knife for sequence and alignment processing. IMeta, 3(3), e191. https://doi.org/10.1002/IMT2.191

Stackebrandt E, and Goodfellow M. 1991. 16S rRNA Sequencing Primers. In Stackebrandt E & Goodfellow M (Eds.), Nucleic Acid Techniques in Bacterial Systematics (Vol. 31, Issue 6, pp. 132–136). John Wiley & Sons, Ltd. https://doi.org/10.1002/JOBM.3620310616

United Nations. **2018**. 68% of the world population projected to live in urban areas by 2050. United Nations Department of Economic and Social Affairs.

Van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., and Thermes, C. **2018**. The Third Revolution in Sequencing Technology A Brief History of Sequencing Technology. CellPress , 34(9). https://doi.org/10.1016/j.tig.2018.05.008

Větrovský, T., and Baldrian, P. 2013. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PloS One, 8(2). https://doi.org/10.1371/JOURNAL.PONE.0057923

Walsh, C. J., Srinivas, M., Stinear, T. P., van Sinderen, D., Cotter, P. D., and Kenny, J. G. 2024. GROND: a quality-checked and publicly available database of full-length 16S-ITS-23S rRNA operon sequences. Microbial Genomics, 10(6). https://doi.org/10.1099/mgen.0.001255

Wick, R. R., Judd, L. M., & Holt, K. E. 2018. Deepbinner: Demultiplexing barcoded Oxford Nanopore reads with deep convolutional neural networks. PLOS Computational Biology, 14(11). https://doi.org/10.1371/journal.pcbi.1006583

Williams, S. H., Che, X., Paulick, A., Guo, C., Lee, B., Muller, D., Uhlemann, A.-C., Lowy, F. D., Corrigan, R. M., & Lipkin, W. I. 2018. New York City House Mice (Mus musculus) as Potential Reservoirs for Pathogenic Bacteria and Antimicrobial Resistance Determinants. https://doi.org/10.1128/mBio

World Bank. **2012**. People, Pathogens and our Planet. www.worldbank.org

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K. H., Whitman, W. B., Euzéby, J., Amann, R., & Rosselló-Móra, R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nature Reviews Microbiology, 12(9), 635–645. https://doi.org/10.1038/nrmicro3330

Zhang, T., Li, H., Ma, S., Cao, J., Liao, H., Huang, Q., & Chen, W. 2023. The newest Oxford Nanopore R10.4.1 full-length 16S rRNA sequencing enables the accurate resolution of species-level microbial community profiling. Applied and Environmental Microbiology, 89(10). https://doi.org/10.1128/aem.00605-23