Animal Microbiome

RESEARCH Open Access

Metagenomics provides broad detection of pathogens, antimicrobial resistance, and virulence genes in pig diarrhoea and complement conventional methods

Mirena Ivanova[1], Birgitta Svensmark[2], Emilie Egholm Bruun Jensen[1], Frank M. Aarestrup[1], Håkan Vigre[1] and Saria Otani[1*]

Abstract

Background Post-weaning diarrhoea (PWD) remains a major cause of morbidity in pig production and is commonly associated with enterotoxigenic *Escherichia coli* (ETEC). Conventional diagnostics rely on culturing and targeted qPCR, which provide limited resolution of pathogen diversity, virulence and antimicrobial resistance. Here, we evaluated Oxford Nanopore Technologies (ONT) metagenomic sequencing as a diagnostic tool for direct detection of pathogens, virulence factors and antimicrobial resistance genes (ARGs) from diarrhoeal pig faeces.

Results Twenty-six diarrhoeal and six healthy pig faecal samples were analysed using culture, qPCR and ONT metagenomics with both high-output and rapid workflows. Culturing recovered 26 haemolytic *E. coli* and nine *Clostridium perfringens* isolates. PromethION metagenomics detected a significantly higher diversity of bacterial species, virulence factors and ARGs compared with GridION. Direct read mapping achieved 71–96% genome coverage for six *E. coli* isolates. Fourteen high- and medium-quality *E. coli* metagenome-assembled genomes (MAGs) were reconstructed, of which seven clustered closely with corresponding cultured isolates. All virulence factors detected in isolates were captured by metagenomics, while metagenomics identified additional fimbrial and enterotoxin genes not recovered by culture. Metagenomic ARG profiling identified resistance to 16 antibiotic classes, compared to eight classes in cultured isolates. No ESBL, carbapenemase or *mcr* genes were detected.

Conclusions Long-read ONT metagenomics enables culture-independent, strain-resolved characterisation of the pig gut microbiome during PWD, capturing pathogen diversity together with virulence and antimicrobial resistance profiles. This approach reveals within-sample strain heterogeneity and functional potential that are not resolved by conventional culturing, supporting its value for studying microbial ecology and dysbiosis in diseased animal microbiomes.

*Correspondence:
Saria Otani
saot@food.dtu.dk
[1]Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Kgs Lyngby, Denmark
[2]LVK, Veterinary Practice, Hobro, Denmark

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Diarrhoea diseases are a major cause of morbidity and mortality in pig production [1]. Diarrhoea in pigs can be caused by different enteropathogens (bacteria, viruses, parasites and microsporidia) [2]. Rapid and accurate detection of infectious agents is essential for managing diseases pigs, including post-weaning diarrhoea (PWD) [3].

PWD occurs a few days after weaning of piglets, with enterotoxigenic *Escherichia coli* (ETEC) as a common pathogen (pathovar) isolated from PWD in newly weaned pigs [4]. ETEC pathogenesis involves adhesion to specific receptors of the pig small intestinal epithelium using fimbrial F4 (K88) or F18 adhesins, followed by colonization and enterotoxin production (heat-stable (STa, STb) and heat-labile (LT)), resulting in increased fluid and electrolyte secretion into the intestinal lumen and reduced absorption [5]. Another agent that is commonly associated with diarrhoea in pigs is *Clostridium perfringens* type C, which causes a highly fatal, necrohemorrhagic enteritis primarily in 1–5 days old piglets [6]. *C. perfringens* pathogenesis involves rapid overgrowth in the pigs small intestine, inhibition of beta-toxin (CPB) degradation by trypsin inhibitors in the pigs colostrum and initial damage to the small intestinal epithelial barrier [7].

Current diagnostic practice to detect infectious agents causing diarrhoea diseases in pigs, in veterinary laboratories typically combines bacterial culturing and qPCR assays [2], which target known virulence genes such as *E. coli* fimbrial adhesins along culture-based methods to isolate and characterise the pathogens. These approaches are user-friendly and cost-effective, yet they are inherently targeted and therefore may not capture the full spectrum of causative agents or provide genome-wide information on antimicrobial resistance genes (ARGs) and virulence factors (VFs) repertoires [8, 9].

Metagenomic sequencing offers a powerful, culture-independent *complement* that is capable of profiling the entire microbial community in a single assay [10]. By capturing both pathogens and mutualistic microbes, metagenomics can simultaneously detect bacterial, viral and eukaryotic pathogens while also profiling the gene content; ARGs and VFs [11]. Long-read sequencing platforms, such as Oxford Nanopore Technologies (ONT), further enable quicker than before, real-time sequencing data generation and strain-level resolution, making them quickly attractive for diagnostic applications [12]. However, the practical implementation of ONT in veterinary diagnostics requires balancing sequencing depth, turnaround time, and cost against the resolution needed for actionable results.

ONT platforms can be used in different strategies depending on the diagnostic goals. PromethION devices with high-output flow cells support deep sequencing for comprehensive microbiome and pathogen characterisation. Whereas GridION devices allow multiplexing of many samples for faster and more affordable surveillance, albeit with lower sequencing depth [13].

The aim of this study is to evaluate whether various ONT sequencing can detect the causative agents of diarrhoea in animals (e.g., pigs in this study) directly from faecal samples, and how results compare with conventional diagnostics. We combined culture-based isolation, qPCR and WGS (whole genome sequencing) of *E. coli* and *C. perfringens* with ONT metagenomic sequencing using both PromethION (research-based) and GridION (rapid diagnostic) workflows. We assessed the ability of each approach to detect bacterial pathogens, ARGs, virulence genes, and intestinal parasites, and critically compared their performance, resolution, taking into consideration their turnaround time.

Materials and methods

Clinical sample collection and experimental set-up

Twenty six diarrhoeal pig faecal samples were obtained from Landbrugets Veterinære Konsulenttjeneste (LVK) clinics (Fig. 1), Denmark between January 29th and April 3rd, 2024. All samples came from individual pigs exhibiting visible bloody diarrhoea. Pig ages ranged from 3 weeks (newly weaned pigs) to 16 weeks old (slaughtering pigs about 40–50 kg) pigs. Faecal samples were collected aseptically in sterile tubes and stored at 4 °C for maximum one day until processing. For healthy microbiomes, six faecal samples from clinically healthy age-matched pigs were included from a separate research project [14] (Table 1). All included diarrhoeal faecal samples were investigated using conventional diagnostics (culture-based and qPCR) and metagenomics diagnostics (Fig. 1).

Isolation of *Escherichia coli* and *Clostridium perfringens* and species identification

Serial 10-fold dilutions from 1 gr faeces were prepared in LB broth and inoculated on CHROMagar Orientation (Sigma Aldrich, Denmark) to isolate *Escherichia coli* and other Enterobacteriaceae, on MacConkey agar (Sigma Aldrich, Denmark) to confirm isolation of gram-negative enteric bacteria such as *E. coli*, and on blood agar plates (BAP, Thermo Scientific, Denmark) to isolate haemolytic *E. coli*. All plates were incubated for 24 h at 37 °C. To isolate *Clostridium perfringens*, samples were plated on BAP or Reinforced Clostridial AgarNutriSelect (Merck Life Science A/S, Denmark) and incubated in BD GasPack™ EZ jars under anaerobic conditions using AnaeroGen™ (Thermo Scientific, Denmark) for 48 h at 37 °C. BAP were examined for colonies with *Clostridium*-specific double zone haemolysis, or various types of colonies from the Reinforced Clostridial AgarNutriSelect plate. All positive

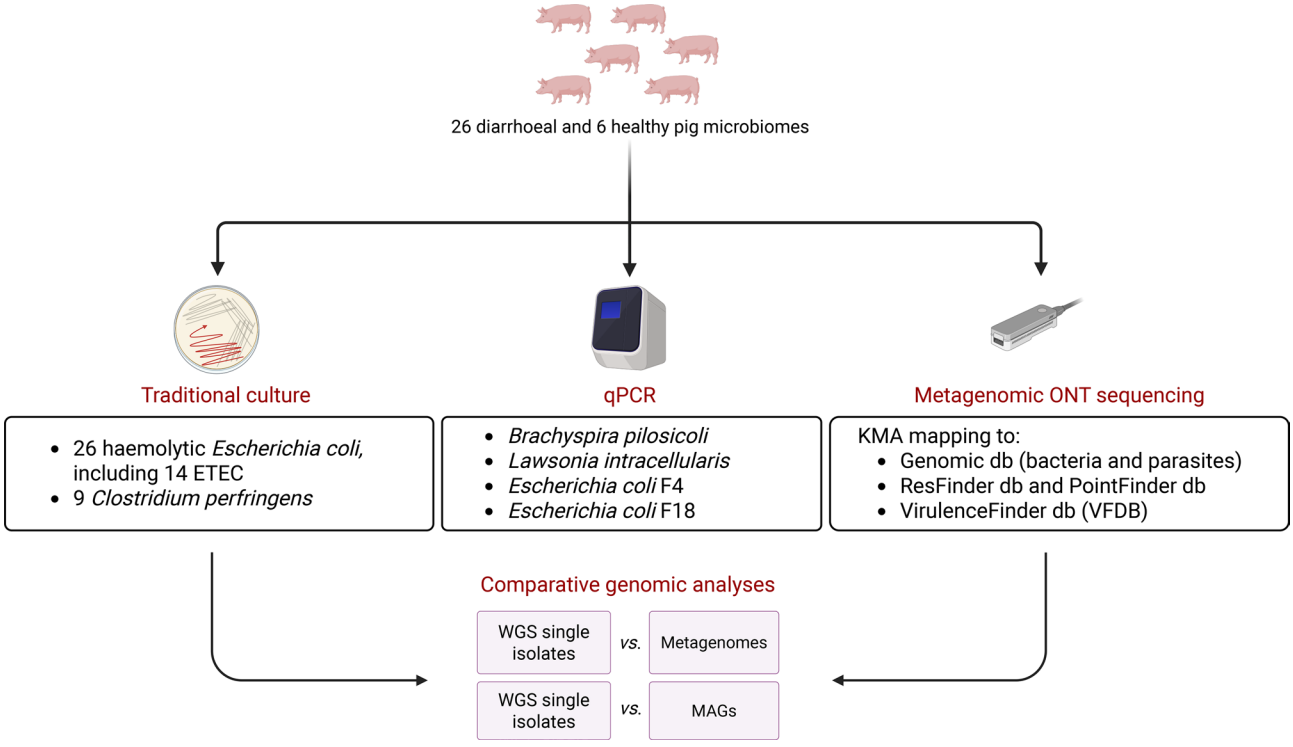

Fig. 1 Experimental study design of both the diarrhoeal and healthy pigs

isolates on the selective media were selected for identification of presumptive *E. coli* and *C. perfringens* using MALDI-TOF MS (Bruker, Daltonik GmbH, Bremen, Germany) before performing genomic analyses.

Detection of *E. coli* with qPCR

qPCR for selected targets was performed as part of the routine diagnostic at the LVK laboratory. Detection of enterotoxigenic *Escherichia coli* (ETEC) was done by detecting fimbrial adhesin genes F4 (K88) and F18 endotoxin genes using quantitative PCR (qPCR) targeting the corresponding virulence genes [4]. DNA was extracted from faecal samples using QIAamp DNA Stool Mini Kit (Qiagen, GmbH, Germany) according to the manufacturer's protocol. qPCR assays were performed using TaqMan qPCR targeting the fimbrial adhesin genes F4 (*faeG*) and F18 (*fedA*), following the assay described by Ståhl et al. [15] using the same volumes, primers and probes. Negative controls (no-template controls) and positive controls (DNA from reference ETEC F4 and F18 strains) were included in each run to ensure assay validity.

DNA extraction and Oxford Nanopore (ONT) sequencing

A- Single isolate whole genome sequencing (WGS)

From the bacterial cultures, 35 isolates were sequenced: 26 hemolytic *E. coli* and 9 *C. perfringens* (Table 1). Total DNA was extracted using the Quick-DNA Fungal/Bacterial Miniprep Kit (Cat. No. D6005, Zymo Research) following the manufacturer's recommendations. DNA concentration was measured using the Qubit 4 Fluorometer (Cat. No. Q33238, Invitrogen), and DNA was stored at 4 °C until library preparation. From each sample, 1 µg DNA input was used for library preparation using the Native Barcoding Kit 96 V14 (SQK-NBD114.96). All samples were multiplexed and loaded on one FLO-PRO114M flowcell (R10.4 M chemistry) and sequenced on Oxford Nanopore (ONT) PromethION platform. Sequencing was performed for 72 h and reads were basecalled using GuppyBasecaller (v7.2.13) with super-accurate basecalling option. Low-quality reads were filtered out using the default setting on MinKNOW with reads below a score of < 9 and read lengths < 200 bp to be omitted.

B- Metagenomic sequencing of faecal samples

To compare metagenomic sequencing with rapid diagnostic workflows, metagenomic DNA from the faecal samples (Table 1) was sequenced using two approaches:

Research-based approach

This approach aimed to maximize sequencing depth and genome recovery for comprehensive microbiome profiling. DNA extraction and ONT sequencing of all 26 metagenomic samples were carried out as described in Ivanova et al. [13] using the Quick-DNA HMW Magbead Kit (Cat. No. D6060, Zymo Research) and Ligation

Table 1 Diarrhoeal pig samples included in this study and haemolytic *E. coli* and *C. perfringens* isolates recovered from them. qPCR Ct values for detection of *E. coli* F4 and F18, *Brachyspira pilosicoli* and *Lawsonia intracellularis* are included in the table

Sample ID	Pig condition/research project	Cultured isolates		qPCR (Ct value)				Age of pig*
		Haemolytic E. coli	C. perfringens	*E. coli*F4	*E. coli*F18	B. pilosicoli	L. intracellularis	
LVK 8	Diarrhoeal/LVK	-	-	-	29	23	22	slaughtering pig
LVK 11	Diarrhoeal/LVK	x (l)	x	21	21	-	32	weaner pig
LVK 17	Diarrhoeal/LVK	x (l)	-	24	28	23	24	weaner pig
LVK 18	Diarrhoeal/LVK	-	-	26	29	23	23	weaner pig
LVK 42	Diarrhoeal/LVK	x (d)	-	33	34	26	-	slaughtering pig
LVK 43	Diarrhoeal/LVK	x (l)	x	19	17	-	25	weaner pig
LVK 48	Diarrhoeal/LVK	x (d) ×(l)	-	26	22	-	-	weaner pig
LVK 52	Diarrhoeal/LVK	x (d)	x	-	-	-	27	slaughtering pig
LVK 56	Diarrhoeal/LVK	x (d)	-	15	18	-	-	weaner pig
LVK 60	Diarrhoeal/LVK	x (d) ×(l)	x	21	15	-	-	slaughtering pig
LVK 61	Diarrhoeal/LVK	x (l)	x	31	31	-	-	weaner pig
LVK 73	Diarrhoeal/LVK	-	-	-	25	27	22	slaughtering pig
LVK 74	Diarrhoeal/LVK	x (l)	-	26	28	-	-	weaner pig
LVK 76	Diarrhoeal/LVK	x (d)	-	-	22	-	21	slaughtering pig
LVK 79	Diarrhoeal/LVK	x (d)	-	-	-	27	24	slaughtering pig
LVK 80	Diarrhoeal/LVK	x (d) ×(l)	-	20	20	23	-	weaner pig
LVK 82	Diarrhoeal/LVK	x (d)	-	20	15	26	24	weaner pig
LVK 83	Diarrhoeal/LVK	x (d) ×(l)	x	23	16	22	29	weaner pig
LVK 88	Diarrhoeal/LVK	x (d)	-	-	-	27	-	slaughtering pig
LVK 94	Diarrhoeal/LVK	x (l)	x	21	20	-	-	weaner pig
LVK 97	Diarrhoeal/LVK	x (l)	x	20	17	-	-	weaner pig
LVK 98	Diarrhoeal/LVK	x (l)	-	-	30	-	24	slaughtering pig
LVK 101	Diarrhoeal/LVK	x (d)	-	31	24	22	-	weaner pig
LVK 102	Diarrhoeal/LVK	x (l)	x	27	-	27	-	weaner pig
LVK 103	Diarrhoeal/LVK	x (l)	-	-	31	25	-	weaner pig
LVK 107	Diarrhoeal/LVK	-	-	30	-	29	28	slaughtering pig
Farm 1_6	Healthy/DANMAP	NA	NA	NA	NA	NA	NA	slaughtering pig
Farm 1_11	Healthy/DANMAP	NA	NA	NA	NA	NA	NA	slaughtering pig
Farm 2_2	Healthy/DANMAP	NA	NA	NA	NA	NA	NA	weaner pig
Farm 2_10	Healthy/DANMAP	NA	NA	NA	NA	NA	NA	weaner pig
Farm 3_12	Healthy/DANMAP	NA	NA	NA	NA	NA	NA	slaughtering pig
Farm 3_21	Healthy/DANMAP	NA	NA	NA	NA	NA	NA	slaughtering pig

(-) either bacteria were not isolated or Ct values were detected

NA: Not applicable

*Weaner pig: 3–6 weeks old, weight 7–20 kg; slaughtering pig: 10–14 weeks old, weight 30–50 kg

"l" - *E. coli* isolate recovered by LVK

"d" - *E. coli* isolate recovered by DTU

Ct value was considered positive when ≤ 35

gDNA Native Barcoding Kit 24 V14 (SQK-NBD114.24, Oxford Nanopore Technologies, Oxford), respectively, except that the starting material for DNA extraction was increased to 200 mg faeces. DNA concentration was measured using the Qubit 4 Fluorometer and DNA was stored at 4 °C until library preparation. From each sample, 1 µg DNA input was used for library preparation, and four samples were multiplexed and loaded on one FLO-PRO114M flowcell (R10.4 M chemistry), and sequenced on PromethION platform with settings described above for the single isolates.

To assess the effect of sequencing depth due to ONT sequencing platforms, four of the diarrhoeal samples were multiplexed and sequenced on one GridION flowcell (i.e., GridION 4-plex; to keep the same number of samples multiplexed per flowcell as in the PromethION sequencing) to evaluate how the reduced sequencing output would influence pathogen detection. The DNA extraction, library preparation and sequencing settings and bioinformatic workflows for GridION 4-plex were carried out identically as above for PromethION.

Rapid metagenomic approach

This approach was implemented to evaluate the feasibility of rapid, low-complexity ONT sequencing for detecting key pathogens directly from diarrhoeal faeces. Here we employed a rapid DNA extraction and library preparation to detect the hemolytic *E. coli* and *C. perfringens* isolates recovered from the faecal diarrhoeal samples. A subset of 21 diarrhoeal samples (out of the 26 total samples) and two healthy pigs were sequenced using this procedure due to limited faecal leftover from the previous steps. Briefly, total DNA was extracted using Quick-DNA Fecal/Soil Microbe MiniPrep Kit (Cat. No. D6010, Zymo Research). DNA concentration was measured using the Qubit 4 Fluorometer and library preparation was carried out by using the Rapid Barcoding Kit 24 V14 (SQK-RBK114.24, Oxford Nanopore Technologies, Oxford) following the manufacturer's recommendation. The 23 samples were loaded on two FLO-MIN114 flowcells (R10.4 M chemistry) and sequenced on GridION platform for 72 h with super-accurate basecalling. Reads were filtered out using the same settings as described for the PromethION research-based approach above.

Data analysis

Genomic annotation of the cultured E. coli and C. perfringens

Genomic classification of the bacterial species was performed using KmerFinder [16]. The ONT sequence data were quality checked using NanoPlot v1.33.0 [17] and assembled using Flye v2.9 [18]. To compare the phylogenetic relationships within each of *E. coli* and *C. perfringens* isolates, SNP-based and core-genome maximum-likelihood (ML) trees were constructed. For the core-genome ML trees, the assembled genomes were annotated using Prokka v1.14.15 [19] and the resulting. gff files were used as input for Roary v3.13.0 [20] to create a core-genome alignment (a core gene is defined as such if it is found in 99% of the genomes). An SNP-based alignment was created by the CSI phylogeny v1.4 [21]. The.aln files for both SNP- and core genes-based phylogenies were used as input to IqTree v2.1.3 [22] to infer ML phylogenetic trees with the GTR+G substitution model and 1000 bootstrap replicates. For the CSI phylogeny, *C. perfringens* ATCC 13,124 (NC_008261.1) and *E. coli* K-12 substr. GM4792 (CP011343.2) were used as reference genomes.

For virulence and ARG annotations of the single isolate bacterial genomes, KMA v1.4.12a was used to map the raw sequence reads (.fastq files) to VFDB (downloaded on Mar 14, 2025) and ResFinder and PointFinder v4.0 (last updated 14 May 2024), and gene depth and coverage were calculated as described previously [13]. Plasmid gene content was searched using BLASTn in ABRicate v1.0.1 (https://github.com/tseemann/abricate) and the PlasmidFinder db (last updated 05 July 2025) [23].

E. coli pathotypes, serotypes and sequence types (STs) were identified by ectyper v2.0.0 [24]. *C. perfringens* isolates were typed following the scheme of Rood et al. [25] using a custom database in ABRicate (https://github.com/tseemann/abricate). Phylogenetic trees of the *E. coli* and *C. perfringens* isolates and the comparison between the *E. coli* isolates and the *E. coli* metagenome-assembled genomes (MAGs) were visualised in Geneious Prime® 2025.2.2. All graphs were generated using the following R packages: "pheatmap v1.0.13" [26], "ggplot2 v3.5.2" [27], "ggtree v3.12.0" [28].

Metagenomics analyses of the faecal samples

For both the rapid and research metagenomic workflows, high-quality reads were taxonomically and functionally classified following procedures adapted from Bogri et al. [29] and Ivanova et al. [13]. In brief, high quality ONT reads were mapped using KMA v1.4.12a to a custom reference genome database (last updated 22 May 2024) previously applied for taxonomic assignments [13, 30]. This database included closed bacterial genomes from NCBI GenBank, archaeal genomes, MetaHitAssembly (PRJEB674–PRJEB1046), HumanMicrobiome genome assemblies, and bacterial draft genomes. For antimicrobial resistance gene (ARG) identification, reads were aligned with KMA v1.4.12a against the ResFinder database v4.0 (last updated 14 May 2024), containing curated sequences of acquired resistance genes [31, 32]. Unlike standard Illumina-based approaches where read counts are directly comparable due to uniform read length, e.g. [30], all abundance estimates from long-read data were calculated taking variable read length into account and similar to Ivanova et al. [13]. High-quality ONT reads were mapped at coverage thresholds of 0.5 similar to previous work [13]. This part of the analysis provides an overall characterisation of all the bacterial taxa and ARGs in the microbiomes.

Identifying the cultured bacteria in the metagenomic samples

In order to compare traditional culture methods to metagenomics, we aimed to detect our isolated *E. coli* and *C. perfringens* genomes in our metagenomes. The metagenomics were mapped to a database consisting of only the 26 *E. coli* and 9 *C. perfringens* assembled genomes from this study. This was to detect if the single isolates recovered by culture could be identified in the respective metagenomics samples. Additionally, to further identify those targeted *E. coli* and *C. perfringens* in the metagenomes, metagenome-assembled genomes (MAGs) were assembled from the pig microbiomes sequenced using the research-based approach

on PromethION only (as it has the high output required for MAGs assemblies). The MAGs were assembled using Flye version 2.9.3 (–meta –nano-hq) [18]. Afterward, we extracted all *E. coli* and *C. perfringens* MAGs from our metagenomics and compared them to the genomes of *E. coli* and *C. perfringens* from culturing to identify their relatedness. We compared the *E. coli* and *C. perfringens* MAGs and cultured genomes via average nucleotide identity based on BLASTn+ (ANIb) using pyANI-plus [33] and whole-genome alignment using D-GENIES, a standalone tool performing large genome alignments using minimap2 and providing similarity overview between compared genomes, highlighting repetitions, breaks and inversions [34].

Community analyses of the pig microbiomes

Ordination analyses of the microbiome similarities were conducted for bacterial taxa, ARGs and virulence genes. Ordination was calculated using Bray Curtis Dissimilarity and visualised with PCoA (Principal coordinate analysis). Community similarities were done on bacterial species, genera and families, in addition to ARGs and virulence genes at different coverage values. The normalised depth of bacteria, ARGs and virulence genes were CLR-transformed. PCoA was performed and visualised using "vegan v2.8.0" [35], "ggplot2 v3.5.2" [27] and "dplyr v1.1.4" [36] packages in R (details in SI text). Comparison between the various communities (healthy vs. diarrhoeal) was also done using alpha diversity indices (details in SI text).

The comparison between the alpha diversity indices of the healthy vs. diarrhoeal pig microbiomes sequenced on PromethION, and the comparison between diarrhoeal pig microbiomes sequenced on GridION (*rapid diagnostics approach*) vs. PromethION P2 Solo (*research-based approach*) were calculated using the "vegan" R package on the fragment count output from the KMA mapping, checked for normality using the Shapiro-Wilk test and the appropriate statistical test (t-test or Mann-Whitney non-parametric test) between the compared groups was applied. Box plots were generated using the "ggplot2" R package [27]. To compare microbial diversity between diarrhoeal and healthy pigs, alpha diversity indices (observed species, Chao1, ACE, Shannon, Simpson, and evenness) were calculated from the KMA mapping output using the "vegan" R package [35]. Group differences were tested for statistical significance ($p < 0.05$) after assessing normality with the Shapiro–Wilk test, followed by t-test or Mann–Whitney U test as appropriate.

Differential abundance analysis of bacterial taxa and ARGs between diarrhoeal and healthy pigs was performed using ALDEx2 [37]. Beta diversity was assessed using Bray–Curtis dissimilarity, and ordinations were visualised by PCoA. Coverage thresholds of 0.1, 0.5, and 0.9 were tested, with 0.5 giving the highest discrimination between groups. Bacterial taxa, ARGs and virulence genes were included in the analyses, and ordination plots were generated using the "ggplot2" and "ggtree" R packages [27, 28].

Results

Culture-based isolation and qPCR diagnostics

A total of 26 faecal samples from pigs with visible bloody diarrhoea were analysed by culturing and qPCR for *E. coli* and *C. perfringens* (Table 1). Twenty-six hemolytic *E. coli* isolates were recovered from 22 microbiome samples, and nine *C. perfringens* isolates were obtained from nine samples. Both pathogens were isolated from nine faecal samples, while four samples yielded neither bacterial species. Two *E. coli* isolates were recovered from four faecal samples (one at DTU and one at LVK) (Table 1). All isolates were confirmed by MALDI-TOF MS prior to whole-genome sequencing, confirming the 26 *E. coli* and nine *C. perfringens* genomes for downstream analyses (Table 1, Figure S1).

qPCR diagnostics at Landbrugets Veterinære Konsulenttjeneste (LVK) identified *E. coli* F18 fimbrial genes in 18 and 21 faecal samples, respectively (Ct ≤ 35) and F18 fimbrial genes in 18 and 21 faecal samples, respectively (Ct ≤ 35) Table 1). As part of LVK regular diagnostics, *Brachyspira pilosicoli* and *Lawsonia intracellularis* were also investigated and detected in 14 and 13 samples, respectively.

Bacterial isolate characterisation

Genomic characteristics of the *Escherichia coli* and *Clostridium perfringens*

The quality of the *E. coli* and *C. perfringens* assemblies (number of contigs, contig length, coverage) are in Table S6. Flye generated assemblies with 25.9 contigs for the 26 *E. coli* isolates (on average, SD = 13.1), and 26 contigs for the 9 *C. perfringens* isolates (average, SD = 30.5). The average *E. coli* genome size was 5.57 Mbp (SD = 253,737 bp), and 3.46 Mbp (SD = 84,034 bp) for *C. perfringens*. The coverage was higher for the *C. perfringens* compared to the *E. coli* (546× and 253× respectively) due to its smaller genome. Species assignments by KmerFinder confirmed all isolates as *E. coli* or *C. perfringens* (Table S2). More detailed genomic characterisation of the *E. coli* and *C. perfringens* isolates is presented in Table S3.

Phylogenetic relationship among the *Escherichia coli* and *Clostridium perfringens* isolates

To study the phylogenetic relationships among *E. coli* and *C. perfringens* isolates, maximum-likelihood (ML) trees were constructed based on core-genome alignments and single nucleotide polymorphisms (SNPs) (Figure S2 A). The tree topologies derived from core-genome and SNPs

were highly similar in *E. coli*, while for *C. perfringens* the SNP-based phylogeny was not consistent with MLST assignment and grouped the isolates with lower confidence (bootstrap support values < 70).

Core-genome and SNP-based phylogenies revealed several well-supported *E. coli* clusters (bootstrap > 83). Isolates from different pig microbiomes occasionally showed close relatedness and shared same MLST (Figure S2A), e.g., DTU 52/DTU 56 (ST224, 462 SNPs) and LVK 94/LVK 97 (ST100, 711 SNPs). Within-microbiome isolates such as LVK 60/DTU 60 (ST641) and LVK 83/DTU 83 (ST48) showed variable SNP differences (845 and 120, respectively), indicating different strain-level diversities from the same microbiome (Figure S2; Table S4).

C. perfringens isolates showed more phylogenetic divergence with weaker support (bootstrap < 70), reflecting greater genomic variability (Figure S1; Table S4), e.g., CP 61 and CP 83 showed the highest divergence and were the most distant from the rest of the isolates. The majority of the remaining isolates were more closely related (CP 94, CP 60, CP 43 and CP 11) and belonged to ST850. The remaining isolates were either of undetermined ST or more distant (Figure S1; Tables S3 and S4).

Escherichia coli **pathotyping and virulence factors detection**

Across all *E. coli* isolates, three pathotypes were detected and 466 virulence genes were identified (Figure S2B; Table S5), primarily associated with adhesion and secretion systems. Fourteen isolates were identified as ETEC, three as EPEC, three as STEC/EHEC, and the remaining were undetermined. Mixed pathotypes were also detected in the same sample (e.g., LVK 48), referring to pathovar differences in the same faeces. Pathotype distribution did not follow phylogeny (SNPs) (Figure S2; Table S16), consistent with the plasmid-borne nature of the virulence determinants used to determine the pathovars.

Clostridium perfringens **toxino-typing and virulence factors detection**

C. perfringens virulence is associated with toxin production (7), and all our nine *C. perfringens* were toxinotype A carrying the *cpa* (or *plc*) alpha-toxin producing gene and no other toxin genes as described in the toxinotyping scheme by Rood et al. (2018) (24) (Figure S2 and Table S5). Other diarrhoea-associated genes included *cpb*2 gene, encoding the *C. perfringens* beta 2 toxin (CPB2), found in one isolate (CP 11) and *pfo* gene encoding Perfringolysin O or theta toxin (PFO) carried by all isolates. In addition, all strains harboured *ccp* (encodes possible clostridial myonecrosis-associated alpha-clostripain toxin) and the two-component regulatory system of *pfoA* and *cpa*/*plc*, *virS* and *virR*. None of the *C. perfringens* strains harboured *netB* gene, implicated in necrotic enteritis, or *cpe* gene for *C. perfringens* enterotoxin production. However, one isolate (CP 12) carried the heat-stable enterotoxin 1-encoding *east1* gene, commonly associated with ETEC.

Plasmid replicon genes screening

Between three and eight plasmid replicon genes were identified per *E. coli* isolate, with IncFIC(FII) and IncFIB(AP001918) being the most common, predominantly among ETEC. These plasmids carried key virulence genes including *eltAB, estIa* (encoding enterotoxins in ETEC) and *faeG, and fedA* encoding F4 and F18 subunits (Tables S5 and S6).

***Perfringens* plasmid content was not further examined as all isolates were toxinotype a and *C. perfringens* plasmid replicons are not included in the plasmidfinder database.**

Metagenomic data analysis

Sequencing output and microbiome community analyses PromethION ONT sequencing of 26 diarrhoeal and six healthy samples produced 68, 041 Gbp (mean 2, 126 Gbp per sample, Q > 9), whereas GridION 4 -plex and 23-plex runs yielded 414 and 415 Gbp in total (means 130 and 18 Gbp, respectively) (Figure S3; Table S7). PromethION thus generated 20× and 200× higher output per sample compared to GridION (4- and 23-plex, respectively).

PromethION research-based sequencing revealed marked differences in microbial composition between diarrhoeal and healthy pigs (Fig. 2; Figure S4). Diarrhoeal samples showed significantly higher bacterial and ARG richness (Chao1, $p < 0.05$) (Figure S4; Table S8). Whereas evenness and diversity (Simpson) were higher in healthy pigs compared to diarrhoeal ones in bacterial taxa, and no apparent ARG diversity differences were observed between the two groups (Figure S4; Table S8), consistent with Rydal et al. [38]. Beta-diversity using the Bray-Curtis dissimilarity index clearly separated diarrhoeal and healthy microbiomes by bacterial composition, ARGs and virulence genes (Fig. 2). Virulence gene and ARG profiles showed, however, a more diluted clustering between the two groups compared to the bacterial taxa (Fig. 2). Differential abundance analysis carried out with ALDEx2 v1.18.0 in R [37] identified several bacterial families that differentiated diarrhoeal from healthy pigs (Figure S6B; Table S9A).

Although similar microbiome shifts between diarrhoeal and healthy pigs have been reported previously (e.g., Rydal et al. [38]), our healthy samples originated from a separate study of same age pigs. Differences in diet, or management may therefore contribute to the observed variation, and these comparisons should be interpreted cautiously.

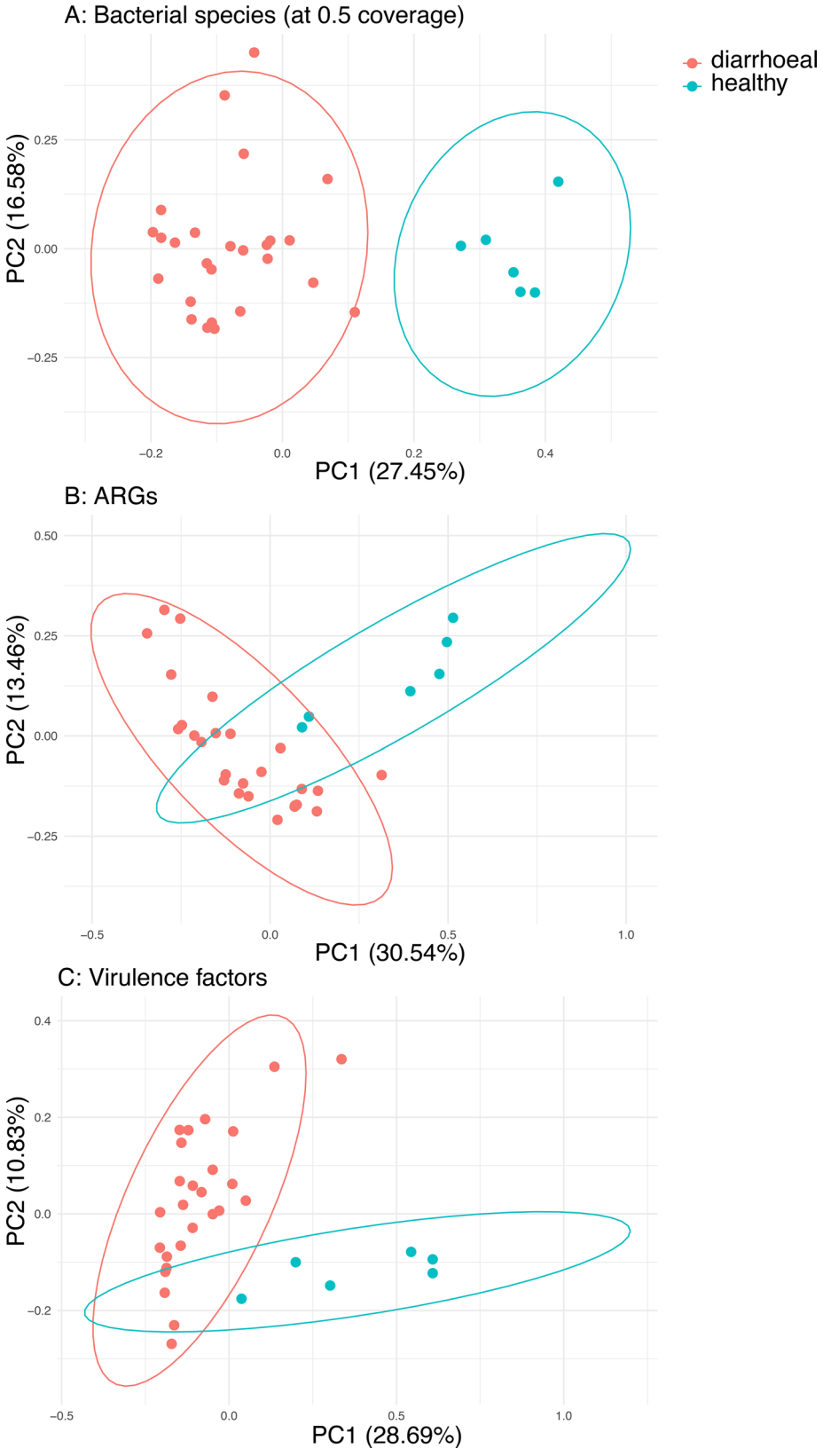

Fig. 2 Principal coordinate analysis (PCoA) of CLR-transformed normalised depth values from diarrhoeal and healthy pig microbiomes sequenced on PromethION (research-based approach). (**A**) PCoA of the bacterial species, detected at a coverage threshold of 0.5. (PCoA of bacterial species detected at coverage thresholds of 0.1 and 0.9 and at family level at 0.5 coverage are shown in Figure S5); (**B**) virulence factors; (**C**) ARGs. Normalised depths for virulence factors and ARGs were calculated at coverage threshold of 0.9

Detection of bacteria and intestinal parasites

PromethION research-based metagenomes consistently detected a significantly ($p < 0.05$) greater number of bacterial species and yielded higher sequencing depths compared to both GridION approaches (rapid or research based) (Figure S6A; Table S9B; Table S10). Several *Lactobacillus* species (including *L. amylovorus, L. johnsonii, L. crispatus*) and *Limosilactobacillus reuteri*, were among the top 50 most abundant bacterial species in the pigs microbiomes that were detected across all platforms and approaches. *Clostridium* sp. was also among the most abundant bacterial species, although identified by the PromethION research-based approach in the majority of the cases. GridION 4- and 23-plex also detected additional species, including *L. johnsonii, Streptococcus alactolyticus, L. crispatus*, and *Dorea longicatena* (known pig microbiome healthy bacteria [39](39)), although at lower depths. Interestingly, several species detected at very low depth by PromethION, such as *Kurthia huakuii, Phocea massiliensis,* and *Vagococcus fluvialis*, were also captured by GridION 4- or 23-plex (both approaches) (Figure S6A).

The deep PromethION research-based sequencing allowed us to identify bacterial families that separate diarrheal from healthy pig microbiomes (Figure S6B). Carnobacteriaceae and Corynebacteriaceae detected in the diarrheal pig microbiomes are likely the key drivers that separate healthy from diarrheal groups. Carnobacteriaceae are lactic acid bacteria present in the porcine gut and responsive to diet transitions around weaning, consistent with well-described weaning-associated microbiome shifts [40–43], whereas Corynebacteriaceae include opportunistic pathogens of livestock, such as *C. pseudotuberculosis* and the *C. renale* group, reported in small ruminants and sometimes associated with gastrointestinal signs [44].

To assess the presence of clinically relevant taxa, the metagenomic reads of the diarrhoeal pig samples were additionally screened against a panel of pathogenic bacteria defined previously [13] and updated in this study. *E. coli* was the dominant taxon, consistently detected across nearly all samples and two approaches when present in high depth (Fig. 3; Table S11). Other bacterial pathogens, identified by the researched-based PromethION approach, including *Staphylococcus aureus, Klebsiella oxytoca*, *Klebsiella pneumonia*, *Enterococcus faecium*, *Listeria monocytogenes* and *Acinetobacter baumannii*, were also identified by all sequencing approaches but generally at very low abundances. Overall, the PromethION research-based approach detected a wide range of pathogenic taxa, while the GridION 4- and 23-plex identified only a few of these species (Fig. 3). Regarding intestinal parasites, *B. pilosicoli* and *L. intracellularis*, amongst other major contributors to PWD, were identified by qPCR in 14 and 13 samples, respectively, with threshold of $Ct \leq 35$ (Table 1). PromethION research-based approach confirmed *L. intracellularis* in eight of these and detected *Brachyspira* spp. in 15 samples (Figure S7). Additionally, *Enterocytozoon bieneusi* was the only parasite previously associated with intestinal dysbiosis in pigs that was consistently detected across all diarrhoeal and healthy pigs, primarily by PromethION (Table S12). Since parasites affecting pigs have larger genomes than prokaryotes (e.g., *Blastocystis*: 12 Mbp), the raw data were also screened at lower genome coverage (i.e., 0.1), which resulted in detection of *Blastocystis* spp. in four of the diarrhoeal samples by PromethION and *Cryptosporidium* spp. across all diarrhoeal samples and approaches

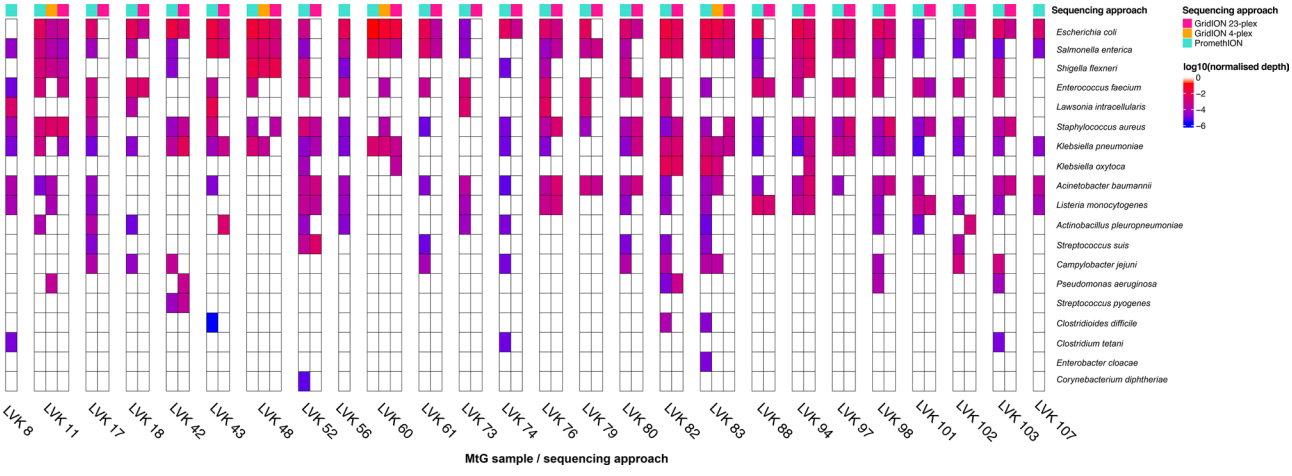

Fig. 3 Pathogenic bacterial species detected by the research-based approach (PromethION) and the rapid surveillance approach (GridION 4-plex and GridION 23-plex), presented as decreasing log10 of the depth (in blue are the most abundant and in red - least abundant species). Empty cells indicate that no bacterial pathogens were detected at the applied coverage of 0.5

(diarrhoeal and healthy) albeit with very low depth (Table S12).

Detection of the diarrhoea-causing *E. coli* in the metagenomic samples

High-quality reads from each metagenomic dataset were mapped to a database of the assembled genomes from the 26 *E. coli* isolates recovered by culturing, allowing to assess whether the cultured isolates could be detected in the corresponding faecal metagenomes. Strong alignment (71% – 96% coverage) between the *E. coli* genomes and the respective metagenomic sample was observed in five cases (e.g., MtG 11 to LVK 11; MtG 43 to LVK 43; MtG 48 to LVK 48; MtG 60 to LVK 60 and LVK 60; MtG 83 to LVK 83) (Fig. 4). In other cases (e.g., MtG LVK 48, MtG LVK 60, MtG LVK 83), the metagenomic reads aligned to multiple *E. coli* genomes at coverage ≥ 50%, with several *E. coli* isolates dominating the metagenome (Fig. 4). This indicates that the approach is hindered by the presence of closely related *E. coli* isolates in the metagenomes (e.g., MtG LVK 82 and MtG LVK 83, Figure S1), and the limits of discriminating bacteria at strain level using direct read mapping.

Detection of virulence factors in *E. coli* isolates and the metagenomes

ETEC is the *E. coli* pathovar that has been mostly associated with PWD [38]. Fourteen of the 26 *E. coli* isolates were assigned to ETEC (Figure S2A). Unique characteristic of this pathotype is carriage of genes encoding F4 or F18 fimbriae and/or two types of enterotoxins, heat-stable (STa and STb) and/or heat-liable (LT), located on plasmids with replicon types IncFIC and/or IncFIB [4].

Three ETEC isolates carried *est* and *elt* genes for both enterotoxins (LT and ST), 10 harboured either LT or STa, and none of them were detected in one ETEC isolate (Fig. 5A).

Regarding fimbriae genes, seven ETEC carried genes for F4 fimbriae, six for F18 and one isolate had no fimbriae genes (LVK 83). Overall, 12 ETEC carried genes for at least one enterotoxin and one fimbriae, while the rest of the ETEC isolates carried either enterotoxin or fimbriae genes. All ETEC isolates carried genes for production of either fimbriae or enterotoxins (Fig. 5A). F4 fimbriae, typically associated with ETEC in post-weaning pigs, was detected only in ETEC isolates in this study, a variant of F18 (F18ab, which differs from the one found in ETEC - F18ac) found in STEC causing edema disease [45] was also detected in three STEC isolates and in four *E. coli* isolates with undetermined pathotype (Fig. 5A).

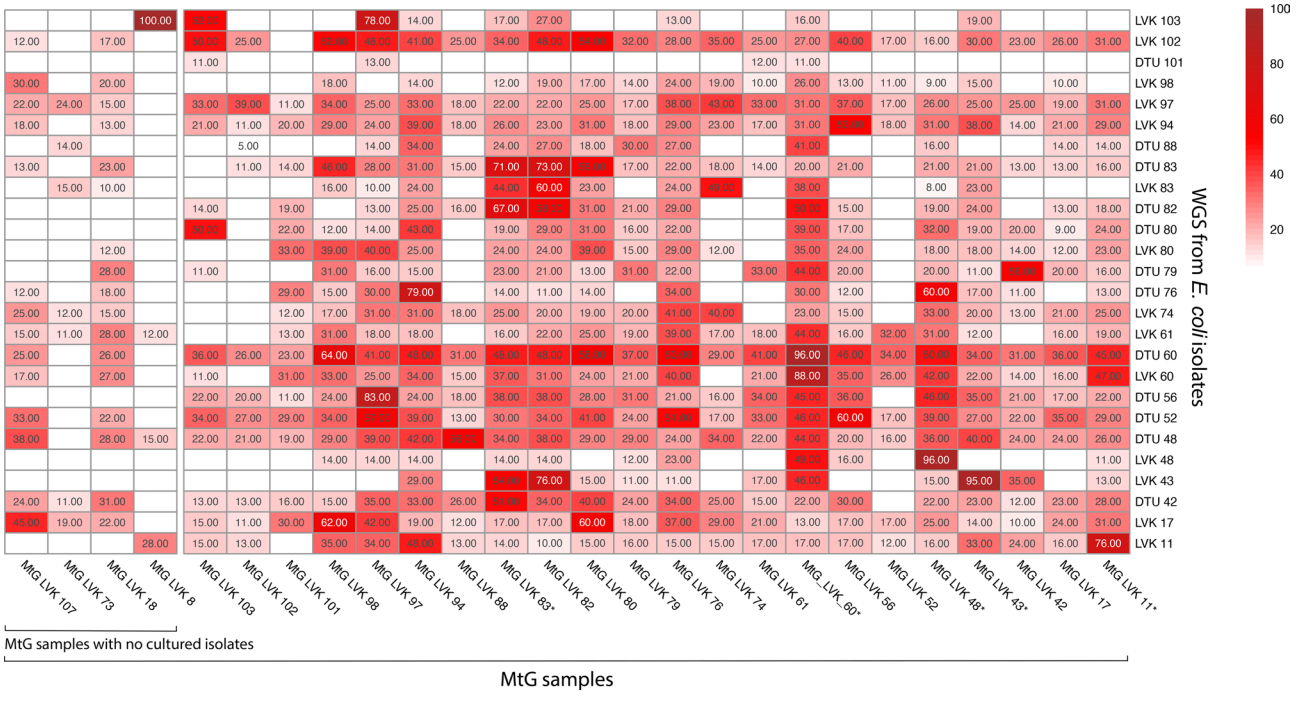

Fig. 4 Heatmap with values representing percentage alignment (coverage) of the MtG samples to a database consisting of the genomes of all 26 *E. coli* cultured isolates, recovered from the MtG samples. The alignment was obtained by mapping (using KMA) the MtG raw data (.Fastq) to the *E. coli* database. Alignments of ≥10% are shown on the heatmap. The four MtGs with no cultured *E. coli* are shown separately on the left side. MtG samples with strong alignment to an *E. coli* isolate recovered from the same sample (71% – 96%) are indicated with an asterisk on the heatmap

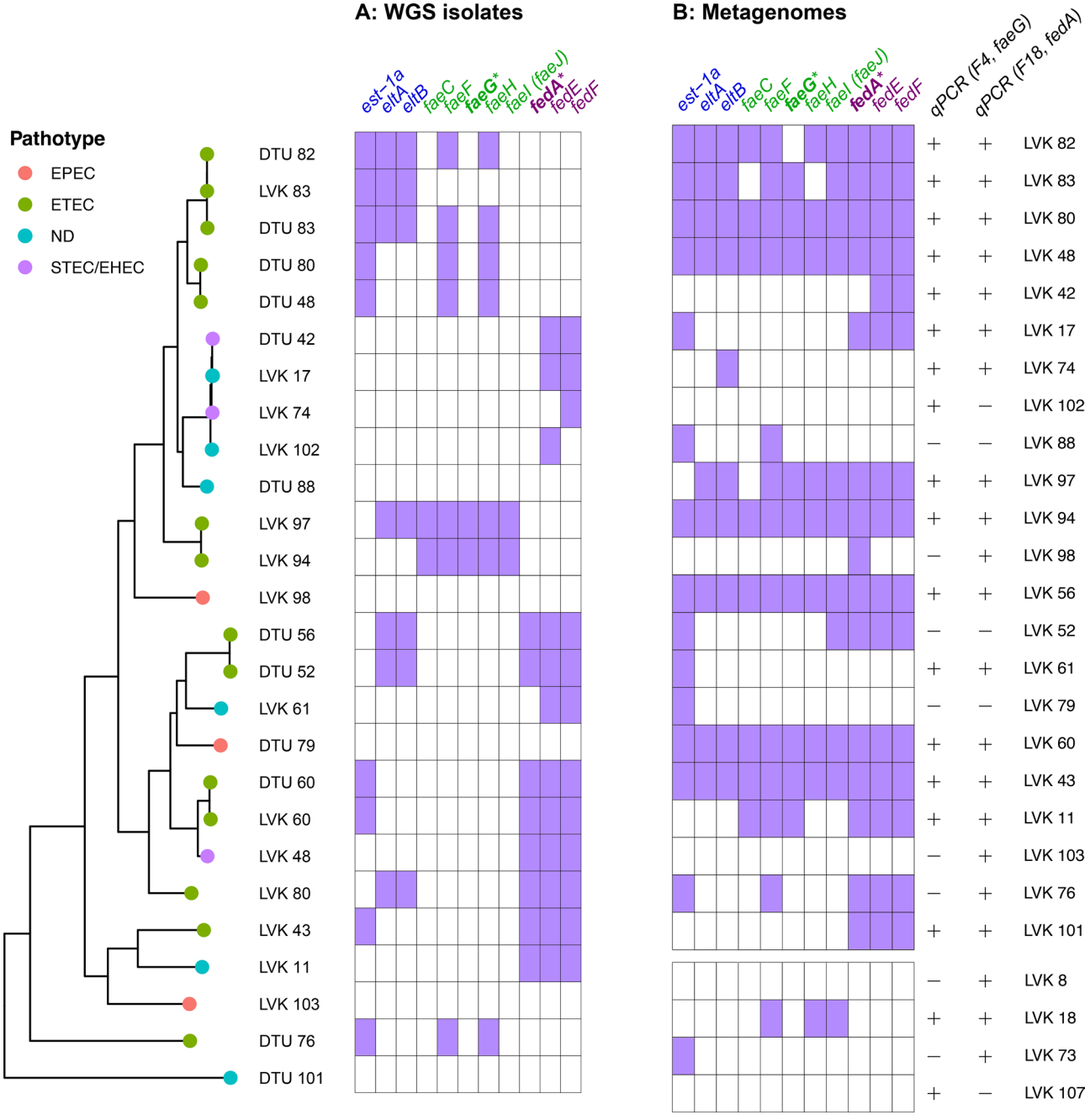

Fig. 5 Genes encoding fimbrial adhesins (F4: green labels and F18: violet labels) and enterotoxins (blue labels) in *E. coli* detected at a gene coverage ≥ 0.9 in (**A**) *E. coli* single isolates and (**B**) metagenomic samples. The tree is a SNP-based phylogeny. *est(A)* gene encoding for heat-stable toxin (STa) and *eltAB* encoding for heat-labile (LT) toxin; F4 (K88) fimbriae: F4 major sub-unit gene *faeG* and genes for F4 minor sub-units (*faeC*, *faeF*, *faeH*, *faeI*, *faeJ*); F18 (F107) fimbriae: F18 major sub-unit gene *fedA* and genes for F18 minor sub-units (*fedE* and *fedF*). ETEC: enterotoxigenic *E. coli*, EPEC: enteropathogenic *E. coli*, STEC/EHEC: shiga toxin-producing *E. coli*/Enterohemorrhagic *E. coli*, ND: not determined. *E. coli* F4 and F18 genes detection by qPCR is based on a value of $ct < 35$

Asterisks denote the major sub-units genes of F4 (*faeG*) and F18 (*fedA*) fimbriae.

Metagenomics sequencing by the research-based PromethION approach detected all virulence factors related to enterotoxin production, F4 and F18 fimbriae that were found in the *E. coli* isolates, with only few inconsistencies (Fig. 5B). Additionally, metagenomics identified more enterotoxin and fimbriae genes in the faecal samples (Fig. 5), suggesting presence of uncultured ETEC isolates or other *E. coli* pathovars carrying variants of these virulence factors.

PromethION m etagenomics and qPCR comparison also showed overall concordance with only several inconsistencies (Fig. 5B). While 18 samples were F4-positive by qPCR, nine of these were F4 negative by metagenomics, however, in two of them other F4 sub-unit genes (*faeF*, *faeH*, *faeI*) were detected (Fig. 5B). 21 samples were F18-positive by qPCR compared to 15 by metagenomics, and one with other F18 sub-unit genes (*fedE, fedF*). Most discordances were cases where qPCR was positive but metagenomics negative, however, one metagenomic sample (LVK 52) was F18-positive with metagenomics yet negative using qPCR (Fig. 5B).

Detection of antimicrobial resistance genes in the *E. coli* isolates and in the metagenomes

The cultured *E. coli* WGS harboured genes for resistance to eight antibiotic classes, while in the PromethION metagenomic samples resistance to 15 antibiotic classes was observed (Figures S8A and S8B). All ARGs identified in the *E. coli* isolates were also detected in the corresponding metagenomes (Table S13). A variety of ARG profiles were identified in the *E. coli* isolates with greatest diversity of genes conferring resistance to aminoglycosides (16 different aminoglycoside resistance genes), followed by macrolide lincosamide streptogramin B (MLS) antibiotics (Figure S8A). All ETEC isolates but three (two EPEC and one of undetermined pathotype) were multidrug resistant (MDR), carrying genes for resistance to aminoglycosides, beta-lactams, macrolides, tetracyclines, sulfonamides, trimethoprim and quinolones. An STEC isolate (LVK 48) harboured resistance genes to seven of the eight antibiotic classes detected among the isolates (Figure S8A).

Some ETEC isolates within a phylogenetic cluster exhibited nearly identical ARG repertoires (e.g., DTU 82/LVK 83/DTU 83; DTU 56/DTU 52 and DTU 60/LVK 60), while other phylogenetically close ETEC isolates carried markedly different ARG profiles (e.g., DTU 80/DTU 48; LVK 97/LVK 94 (Figure S8C) attributable to the varying plasmid content (Figure S8D). Amongst the corresponding metagenomes, the highest diversity of ARGs was observed among the MLS antibiotics, followed by aminoglycosides and beta-lactam antibiotics (Figure S8B). This pattern reflects antimicrobial use in Danish pig production, where macrolides are commonly applied for *Lawsonia* infections in slaughter pigs, while aminoglycosides are widely used to treat *E. coli*-associated PWD in weaners. Importantly, no ESBL, carbapenemase, or *mcr* genes were detected neither in the single isolates nor in the metagenomic samples. PointFinder also could not detect any known mutations in the *gyrA* or *parC* genes linked to fluoroquinolone resistance.

Since beta-lactam resistance genes are one of the most abundant ARGs, especially in Gram-negative bacteria and often occur with other ARGs on mobile genetic elements [46](47), we examined their occurrence and detection in both single isolates and diarrheal samples across sequencing platforms and approaches (Figures S8 and S9). While *bla*TEM-1B was the only beta-lactam gene detected in the *E. coli* isolates (Figure S8C), 12 *bla* gene families were identified in the metagenomic samples, *bla*TEM was the most abundant *bla* gene among them, however, *cfxA*, due to its association with *Bacteroides*, was the most abundant beta-lactam gene followed by *bla*OXA, *bla*CARB, *bla*ACI and *bla*EBR gene families (Figure S9A), often co-occurring with other ARGs on mobile genetic elements in bacterial pathogens such as *Pseudomonas*, *Proteus* spp. and *Acinetobacter* spp. (some of which were among the detected pathogens in the metagenomic samples (Fig. 3)).

When comparing the detection of beta-lactam resistance genes across sequencing approaches, in 50% of the metagenomic samples none of the *bla* gene families were detected by the GridION 23-plex rapid approach compared to the PromethION approach (Figure S9A). However, when ARGs were merged into antibiotic groups to which they confer phenotypic resistance (in accordance with the ResFinder database), the three most abundant antibiotic classes (tetracyclines, aminoglycosides and MLS) were also detected by the GridION 4-plex and 23-plex approached, although at lower depths (Figure S9B). Lastly, the aminoglycoside antibiotics were the most abundant class found in nearly all *E. coli* isolates (Figure S9C), likely explained by the use of aminoglycoside antibiotics (neomycin and apramycin) for treatment of PWD in Danish pig farming.

Comparison between *E. coli* MAGs and single isolates

High-quality (HQ; >90% completeness, <5% contamination) and medium-quality (MQ; 50–90% completeness, <10% contamination) *E. coli* MAGs were assembled from 14 metagenomic samples (Table S14). Eight MAGs had > 90% completeness and genome sizes of 4.6–5 Mbp (Table S14). *E. coli* could not be cultured from faecal sample LVK 18, but an *E. coli* MAG from LVK 18 was reconstructed. MLST assignment was only possible for MAG LVK 74 (ST10, 99.96% completeness) and MAG LVK 83 (ST48, 100% completeness) as MAGs often suffer of low completeness. In both cases, the *E. coli* single isolates and the reconstructed MAGs had the same ST, indicating that the MAGs and the isolates could be the same strain.

Coli *genome comparisons using phylogeny and average nucleotide identity (ANIb)*

To examine how the constructed *E. coli* MAGs compared with the cultured *E. coli* WGS from the corresponding

microbiome, SNP and core-genome phylogenetic inferences were constructed. Both SNP and core-genome phylogenies had the same topologies and similarly clustered the *E. coli* MAGs and single isolates. Six cultured *E. coli* isolates closely clustered with the reconstructed MAGs from the same samples (Fig. 6A). Among those, MAG LVK 48/SI LVK 48 and MAG LVK 43/SI LVK 43 had 3557 and 6719 SNP differences, respectively (Fig. 6A; Table S15), and MAG LVK 74/SI LVK 74 had only 1041 SNPs difference (Fig. 6A; Table S15). Others paired MAGs and WGS isolates had lower MAG/genome coverage (58 and 64% between MAG LVK 82 and SI DTU 82 and MAG LVK 88 and SI DTU 88, respectively (Figure S10B)), yet remained closely related in the SNP matrix (Figure S10A; Tables S14 and S15).

Other paired MAGs and WGS *E. coli* that showed weaker correspondence, e.g., MAG LVK 42 and MAG LVK 98 were of low quality (61 and 55% completeness, respectively). Other *E. coli* MAG – isolate pairs showed weaker or more complex correspondence. MAG LVK 43 and MAG LVK 48 were clearly closest to their cultured isolates, yet differed by 6719 and 3557 SNPs (Figure S10A; Table S14; Table S15), respectively, compared to their WGS isolates (SI LVK 43 and SI LVK 48). MAG LVK 82 was only 58% complete and differed from isolate DTU 82 by 4419 SNPs, but was almost equally close to DTU 83 and LVK 83, which differed by just 120 SNPs among one another, **suggesting that multiple closely related E. coli lineages co-occurred in these microbiomes**. In contrast, MAG LVK 83 was 90% complete and closely matched both isolates from the same faecal sample (278 and 503 SNPs), confirming successful recovery of that lineage. MAG LVK 88 also showed partial overlap with SI DTU 88 (6,203 SNPs at 64% coverage), while MAG LVK 74, in addition to clustering with SI LVK, also grouped with isolates from LVK 17, DTU 42, and LVK 102 (1,600–2,000 SNPs; Fig. 6A; Figure S10A), **indicating the presence of related strains across different pigs.**

The results of the SNP-based phylogeny were further supported by ANIb analysis. Hadamard values (ANIb × coverage) (Figure S10B) were highest for the same MAGs – *E. coli* genomes MAG LVK 48 to SI LVK 48 (0.82), MAG LVK 43 to SI LVK 43 (0.74), MAG LVK 74 to SI LVK 74 (0.84), and MAG LVK 83 to SI DTU 83A (0.90). MAG LVK 82 and MAG LVK 88 showed intermediate values (0.59–0.62) due to their lower completeness.

Dot plot alignments were also constructed to compare the synteny between the common regions of the cultured *E. coli* isolates and the reconstructed MAGs for the closely-related pairs in the SNP-based phylogeny (Fig. 6A) with overlapping genomic regions ranging 55–90% (Fig. 6B; Figure S11). While some of them showed strong collinearity with large genomic overlaps and only a few gaps or duplications (e.g., MAG LVK 43 and SI LVK 43, 75% and MAG LVK 74 and SI LVK 74, 82%) (Fig. 6B), other pairs exhibited larger gaps (possibly due to incomplete MAG), inversions or low genomic overlaps (Figure S11), which could be attributed to incomplete or misassembled MAGs or strain-level variations.

Sub-typing of the E. coli *MAGs*

Pathotyping was generally not possible for the *E. coli* MAGs, except for MAG LVK 42 and MAG LVK 48 (Table S16), as the majority of the genes that determine the *E. coli* pathotypes are located on plasmids (e.g., *hlyA*, *sta1*, *stb*, *eltAB*), which is consistent with the absence of plasmid replicon genes and plasmid-borne virulence genes in the MAGs. The only virulence gene that was consistently found in the MAGs was *hlyA*, and because *hlyA* is not exclusive to pathogenic *E. coli*, its presence is not indicative for a disease-associated *E. coli*. As *E. coli* share large parts of their genomes, the constructed MAGs could also belong to commensal *E. coli* strains. Interestingly, SI DTU 42C (STEC) carried *stx* and fimbrial genes, whereas its corresponding MAG (MAG LVK 42) was typed as EPEC based on the *eae* gene, suggesting that two distinct *E. coli* lineages could have been recovered in that sample, by metagenomics and culturing. Similarly, MAG LVK 48 (STEC) formed a phylogenetic cluster with SI LVK 48 (STEC), but away from SI DTU 48 (ETEC), indicating successful construction of the STEC but not the ETEC strain (Fig. 6; Table S16). Similarly, due to the incomplete MAGs only two could be sub-typed by the Achtman 7-gene MLST *E. coli* scheme (LVK 74, ST10 and LVK 83, ST48), both of which matched their cultured isolates. Screening of MAGs against VFDB failed to identify any of the fimbrial (F4, F18) or enterotoxin (*estA*, *eltAB*) genes that were present in the corresponding isolates (Fig. 5; Table S18), consistent with their plasmid location. Likewise, no plasmid replicon genes were recovered from MAGs, in contrast to the isolates where multiple replicons from various incompatibility groups were frequently detected (e.g., IncF, IncX, IncY) (Table S6). ARGs were detected in two MAGs only (MAG LVK 56 and MAG LVK 97), however their profiles differed from the corresponding WGS: MAG LVK 56 carried *erm*(B) and *ant(3")-Ia*, which were absent from SI DTU 56, while SI LVK 97 carried *dfrA* and *sul1* but MAG LVK 97 harboured *mef*(C)-*mph*(G) and *sul2*, both located on the same contig (Table S20).

Comparison between the research-based and the rapid surveillance metagenomic approaches

To evaluate whether a rapid metagenomic workflow can approximate results from a high-yield research setup, we compared a GridION-based rapid protocol (rapid DNA extraction of ~ 2 h, rapid library prep, and multiplexing of 23 samples per flow cell) with a research-based

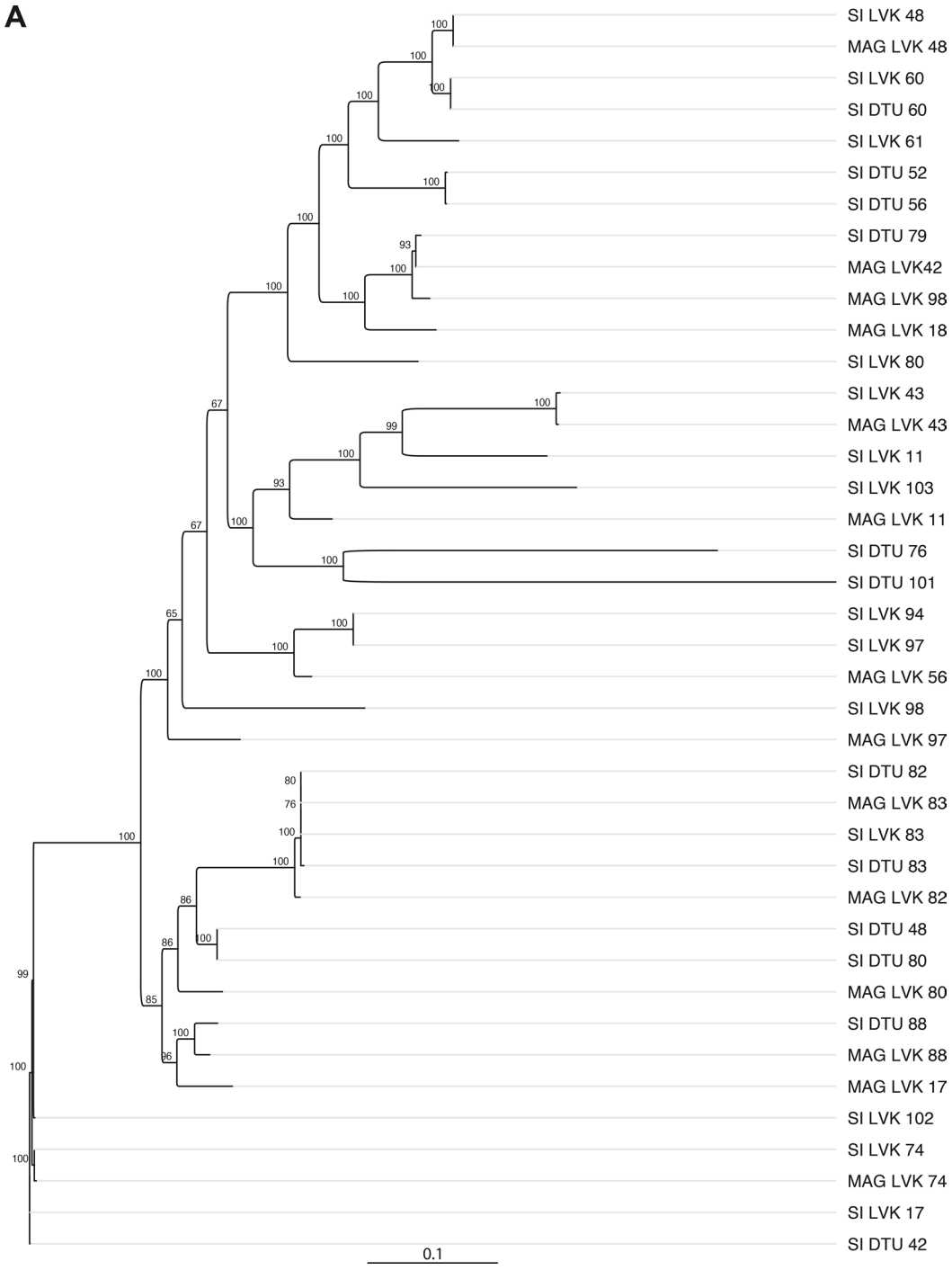

Fig. 6 (**A**) SNP-based phylogeny of *E. coli* single isolates and *E. coli* MAGs. Bootstrap support values are shown on the tree. Scale bar represents 0.1 substitutions per site. The tree was visualised in geneious Prime® 2025.2.2. (**B**) dot plots representing gene synteny between two of the *E. coli*/MAG pairs, showing strong collinearity over large genomic regions (SI LVK 43/MAG LVK 43 and SI LVK 74/MAG LVK 74), x-axis: *E. coli* SI contigs as reference, y-axis: *E. coli* MAG contigs as query. A diagonal straight line indicates synteny among the genomes, inverted horizontal lines indicate sequences that exist in both the SI and MAG but not in the same order, gaps indicate sequence that exists in one sample only and repeats - a sequence that is repeated more than once in a sample. Comparison of the SI LVK 43/MAG LVK 43 was done over 75% (with >75% nucleotide identity) and SI LVK 74/MAG LVK 74) over 82% genome alignment. The remaining *E. coli*/MAG pairs are shown in Figure S11. MAG: metagenome-assembled genome, SI: single isolate

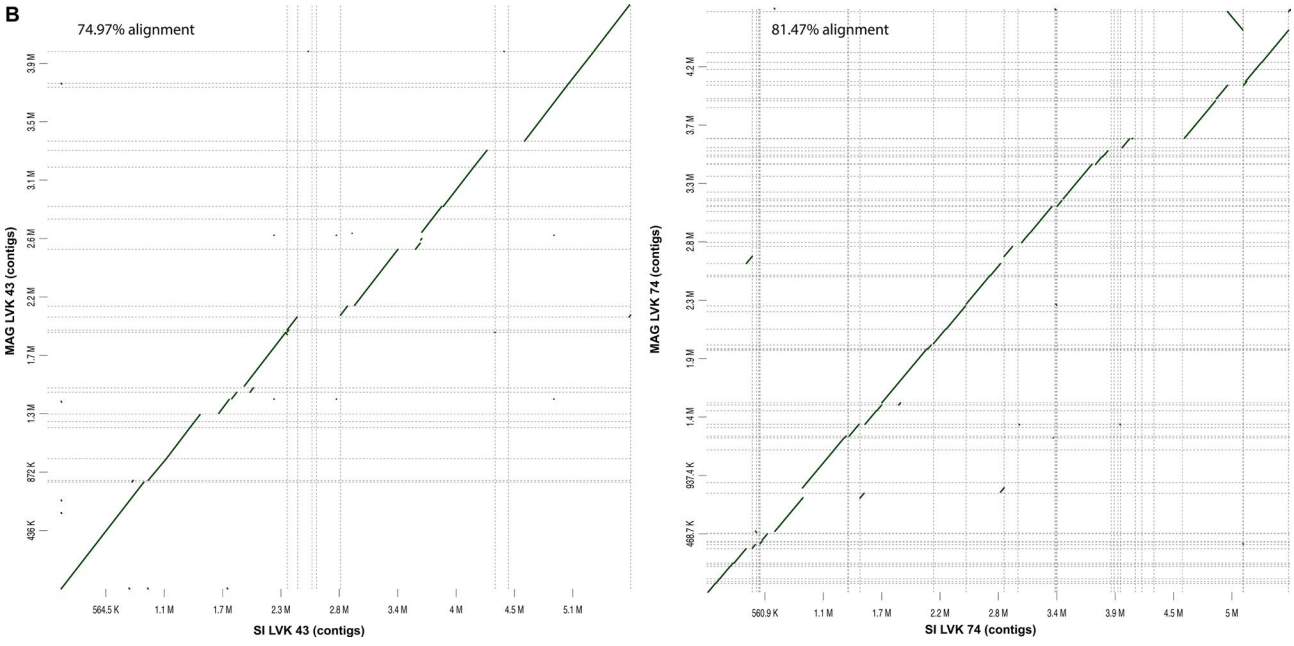

Fig. 6 (continued)

PromethION high-yield and GridION 4-plex sequencing. At bacterial species level, PCA based on Bray–Curtis dissimilarity (beta-diversity) clearly separated PromethION from both GridION approaches, i.e., PromethION samples clustered distinctly apart from both GridION runs, whereas GridION 4-plex and GridION 23-plex overlapped tightly (Figure S12A). At higher taxonomic level (bacterial family), GridION 4-plex dataset overlapped with both PromethION and GridION 23-plex datasets, while PromethION and GridION 23-plex approaches remained separated (Figure S12A). Similarly, for ARG profiles, PromethION and GridION 23-plex formed a separate cluster, and GridION 4-plex showed partial overlap to both PromethION and GridION 23-plex datasets (Figure S12B). These findings suggest that while GridION sequencing (both sequencing approaches) consistently captures the dominant bacterial and ARG profiles, it does not resolve the same community structure as the deeper PromethION sequencing and that GridION 4-plex dataset genomic and ARGs showed more similarity to the deeper PromethION sequencing due to the increased sequencing depth and similar methodology compared to GridION 23-plex rapid approach.

Output in numbers and operational time

In PromethION and GridION 4-plex workflows (research-based) for detailed bacterial characterisation, virulence and ARGs detection, high molecular weight DNA extraction (~2.5 h), barcoded ligation library preparation (~2.5 h) and ~4 h of compute time for QC, mapping, collectively ~10 h, plus 72 h sequencing.

In contrast, the rapid surveillance GridION 23-plex workflow used shorter DNA extraction (~60 min) and rapid barcoding prep (~50 min) and ~1 h compute time, collectively ~3–4 h, plus 72 h sequencing (72 h is optional can be reduced to 24 h if needed).

For culturing and bacterial identification, ~ 28 h

For qPCR for and targeted marker genes ~8 h.

Keeping in mind the overall diagnostic yield, ONT sequencing (both GridION and PromethION) detected 1891 of bacterial species, up to 257 and 15 AMR genes and classes respectively, and broad virulence gene repertoires from a single assay, while allowing multiplexing of 4–24 samples per flowcell depending on platform and depth requirements. In contrast, the qPCR used in this study targets only two virulence genes (F4 and F18 fimbrial adhesins) and does not provide taxonomic resolution, strain-level information, AMR profiles, or additional virulence factors.

Discussion

Here we demonstrate that Oxford Nanopore Technologies (ONT) metagenomic sequencing can identify the potential causative agents of post-weaning diarrhoea (PWD) in pigs directly from faecal samples, matching and extending conventional diagnostic methods. Both research-based (PromethION GridION 4-plex) and rapid (GridION 23-plex) workflows successfully detected key pathogens such as haemolytic *Escherichia coli* and *Clostridium perfringens*, as well as associated antimicrobial resistance genes (ARGs) and virulence factors (VFs). We also highlight the feasibility of integrating ONT sequencing into routine veterinary diagnostics, while revealing

important trade-offs between sequencing depth, turnaround time and analytical resolution.

Diagnostic performance of ONT compared to conventional methods

The concordance between ONT metagenomics, qPCR, and culture-based results supports its diagnostic potential. *E. coli* fimbrial adhesins (F4 and F18) and enterotoxin genes were consistently detected across methods, confirming that ONT sequencing can recover clinically relevant virulence markers directly from faeces. The few discrepancies, where qPCR detected F4/F18 but metagenomics did not, likely reflect limitations in DNA extraction, uneven sequencing depth or lower abundance of target genes within complex microbiomes. This mirrors findings from human and veterinary metagenomic studies showing that pathogens below ~1% relative abundance may evade detection at moderate sequencing depths [12, 47]. It is also documented that qPCR has a higher sensitivity at detecting targeted genes (in this case *faeG* and *fedA*, for F4 and F18 fimbriae, respectively) compared to metagenomics, while metagenomics captures a larger pool of pathogen-related genes [8, 45]. Metagenomics, however, identified additional virulence factors and ARGs not captured by targeted assays, indicating higher genomic resolution and ability to detect emerging pathotypes or co-infections. The detection of the intestinal eukaryotes *Blastocystis* spp. and *Enterocytozoon bieneusi* by ONT suggests that long-read sequencing can broaden pathogen detection beyond bacterial agents in a single step. However, low abundance and incomplete reference databases still limit sensitivity for such eukaryotic pathogens.

Comparative performance of MAGs and cultured isolates

Comparison between metagenome-assembled genomes (MAGs) and cultured *E. coli* genomes showed the genomic resolution achieved by ONT metagenomics. Fourteen high- and medium-quality *E. coli* MAGs (≥50% completeness, <10% contamination) were reconstructed from diarrhoeal samples, consistent with previous studies using alternative assembly and binning methods [46–49], and in seven of the cases clustered closely with their corresponding *E. coli* isolates in the SNP-based phylogeny, sharing high coverage (up to 90%) and strong genomic collinearity. This demonstrates that, when sequencing depth is sufficient as in PromethION, ONT reads can recover individual pathogen genomes directly from complex faecal communities without prior culturing. The reconstruction of an *E. coli* MAG from a culture-negative sample further supports that metagenomics can capture non-culturable or low-abundance lineages.

In contrast, the substantial SNP variation between closely-related WGS isolates and MAGs, incomplete MAGs, and presence of multiple *E. coli* lineages within a sample often hindered sub-typing (MLST, serotyping, pathotyping) and accurate detection of mobile genetic elements, ARGs and VFs. These challenges are consistent with previous studies, showing that MAGs reconstructed from rich and diverse bacterial communities using short- or long-read metagenomics data can lack up to 25% of core and 50% of accessory genes compared to reference isolates [46, 47, 49, 50]. Such losses primarily affect the detection of plasmid-encoded ARGs and virulence factors, as those tend to be located on short contigs with atypical tetranucleotide frequency (TNF) and genomes with unusual TNF profiles may be fragmented or underrepresented and missed by binners [48, 51].

Differentiating between closely-related *E. coli* lineages in the same sample is also particularly challenging. For instance, for metagenomic short-read data, closely-related genomes represent genome-sized approximate repeats that lead to mis-assemblies [46]. This prevents distinguishing between pathogenic strains, and between pathogenic strains and healthy commensal gut strains [47, 52]. In our study, using long-read sequencing data, the MAG and the cultured isolate in sample LVK 42 belonged to two different pathotypes (EPEC and STEC), which strengthen the idea of co-existence of multiple *E. coli* pathotypes within an individual pig microbiome, which could be identified by the reconstructed MAGs from long-read sequencing data.

Overall, our findings indicate that although sufficient sequencing depth and assembly quality can produce *E. coli* MAGs highly similar to cultured isolates, several limitations persist: (1) inability to re-construct complete genomes and associate plasmids with their hosts, which is critical since many clinically relevant virulence and resistance genes are plasmid-encoded, and (2) the presence of highly similar *E. coli* lineages in the pigs microbiomes, including substantial genomic overlap between commensal and pathogenic strains. The latter remains a confounding factor that requires optimisation and benchmarking of MAG construction pipelines. Furthermore, our study was not exhaustive in the comparison between MAGs and isolates, as only one *E. coli* isolate per sample was sequenced except for four of the samples, and one MAG per sample was re-constructed. Thus, our whole-genome comparison of the *E. coli* MAGs and cultured isolates highlights both the potential and current limitations of metagenomic assemblies for pathogen detection and characterisation.

Detection limits for *Clostridium perfringens*

Clostridium MAGs were reconstructed from 18 samples. However, t he MAGs were annotated at genus

level, i.e., *Clostridium* spp. (*Clostridium* sp000435835, GCA_000435835.1), and showed only 68–78% ANIb similarity to the *C. perfringens* cultured isolates, consistent with other studies also reporting *Clostridium* MAGs at genus level [53]. Additionally, a phylogeny based on the core-genome between the *Clostridium* spp. MAGs and the *C. perfringens* isolates could not be inferred, confirming they are genomically highly diverse. The direct mapping of the MtG samples to the *C. perfringens* isolate genomes shows low coverage (<18%) and for some isolates (CP_11D and CP_94C) hits from multiple MtG samples. This could be explained by several factors. *C. perfringens* may have been present at low relative abundance, reducing sequencing depth and assembly coverage. Additionally, *C. perfringens* genome is highly repetitive with low G+C content (28.6%) [54], plasmid-rich (isolates from these species carry up to 10 plasmids [55] with diverse prophage-like regions [56]), all of which complicating assembly and binning.

Moreover, among the detected bacterial species in all 26 diarrhoeal samples, 46 belonged to *Clostridium* spp., some of which are reported as commensal regulating pig gut health (e.g., *Clostridium porci* [39](39); *Clostridium sensu* stricto 1 [57], and could potentially outcompete *C. perfringens* in the metagenomic reconstructions. As *Clostridium* genus is polyphyletic, ANIb values between species can be as low as 70%, while others are closely related and share high ANIb values (90–95%) [58]. Functionally, this means that metagenomics alone would have not detected *C. perfringens* in these diarrhoeal samples even if present, whereas culture provided direct evidence of its presence.

The contrast between *E. coli* and *Clostridium* results in reconstructing MAGs highlights that metagenomic approaches, while comprehensive, may fail to capture clinically important low-abundant or assembly-challenging taxa. For such pathogens, culture remains essential for accurate detection and characterisation, underscoring the value of combined diagnostic workflows.

Direct mapping of the metagenomic reads to *E. coli* genomes

The direct mapping of metagenomic reads to isolate genomes (without MAG construction) confirmed that cultured haemolytic *E. coli* strains were also captured in their corresponding metagenomes, and were the most abundant ones in roughly one-third of cases, validating the ability of ONT sequencing to detect cultured pathogens. However, metagenomes frequently aligned across multiple isolates, demonstrating co-existence of highly similar *E. coli* lineages in different pigs and co-occurrence of multiple *E. coli* strains (and pathovars) within the same animal (e.g., LVK 48 is an STEC, while DTU 48 is an ETEC; and the two LVK 80 isolates were phylogenetically distant ETEC strains). Another possibility for metagenomes alignments to multiple isolates is the extensive core-genome conservation between commensal and pathogenic *E. coli* genomes (~2,200 shared genes) [59], which suggests that alignment of 40–50% corresponds to the core genome. Unlike MAGs–cultured isolates comparison, plasmid contigs were also included in the alignments for the direct mapping, thereby increasing the complexity of this comparative approach. The largest plasmids in *E. coli* that carry ARGs and VFs have been reported to exceed 200 kb [60] and ETEC specifically can harbour large virulence plasmids, often > 150 kb [61]. Although the plasmid content in the different *E. coli* pathovars varies and can be extremely diverse, they have evolved from a single common backbone between 30 to 40 kb [62], suggesting that plasmid DNA could account for a significant proportion of the alignments to non-targeted *E. coli* isolates.

These observations indicate that while ONT metagenomics can recover near-complete pathogen genomes and detect lineages missed by culturing, strain-level discrimination remains limited in rich microbial communities. A combined workflow, culturing and WGS of multiple isolates complemented with metagenomic sequencing, would provide the most complete picture of strain diversity and plasmid-borne pathogenicity.

Plasmid recovery and limitations of metagenomic assemblies

Although ONT sequencing provided high taxonomic and functional resolution, recovery of plasmid-encoded genes in MAGs remained incomplete or uncertain, a limitation commonly observed in other metagenomic studies [46, 48–50, 63]. The reconstructed MAGs in our study lacked plasmid replicons and key virulence or resistance determinants identified in the corresponding isolates, even when chromosomal similarity was high. This limitation stems from the under-representation or mis-binning of circular plasmids and repetitive regions due to their atypical sequence composition and variable copy number [48, 63]. Additionally, plasmidome sequencing has previously been carried out separately and not as part of a regular metagenomic pipeline [64]. However, the limited plasmid-MAG recovery does not imply the absence of plasmids and plasmid-associated ARGs and virulence factors in our metagenomic data. Here, multiple plasmid-borne virulence and ARGs were detected in the metagenomes at read level, despite the inability of plasmid reconstruction in the MAGs.

Pathogen diversity and within-sample strain variation and beyond *E. coli* pathogenicity

The ETEC pathovar, carrying fimbriae F4 or F18 and/or enterotoxin genes, have mostly been associated with post-weaning diarrhoea (PWD) in piglets [65]. Phylogenetic analyses of the *E. coli* pathovars in our study revealed substantial within-sample diversity, often strains differing by hundreds of SNPs or belonging to distinct sequence types (STs). This supports previous evidence that PWD is frequently caused by several agents, involving multiple *E. coli* lineages with variable virulence and resistance repertoires [66]. Such diversity complicates diagnostics based on single isolates and underscores the value of metagenomics for capturing strain-level variation directly from faecal material. Moreover, the co-occurrence of ETEC, STEC, and EPEC pathotypes within individual samples indicates ongoing horizontal gene transfer mediated by virulence plasmids carrying fimbrial and toxin genes [67].

Recent studies suggested that the aetiology of PWD is more complex and the role of a specific etiological agent in PWD is difficult to conclude [38, 68](38, 68, 68). While enterotoxin-producing ETEC F4 has been recognised as a stable pathogen circulating in Danish pig herds [69], mixed infections involving *C. perfringens* type A, *L. intracellularis*, *B. pilosicoli* and enteric viruses, such as circoviruses and rotaviruses, have also been documented, suggesting that PWD reflects a multifactorial dysbiosis rather than a single-pathogen condition. Our ONT metagenomic analysis revealed a broader and more heterogeneous pathogen landscape than indicated by culture or qPCR alone. Beyond *E. coli* and *C. perfringens*, we detected opportunistic species including *Staphylococcus aureus*, *Klebsiella oxytoca*, *Enterococcus faecium*, *Listeria monocytogenes* and *Acinetobacter* spp. Although these organisms may not have been directly associated with PWD, their presence at low abundance suggests that diarrhoeal pigs harbour a more complex microbial imbalance, potentially involving opportunistic or emerging pathogens rather than infection with a single pathogen. The presence of these pathogens could disrupt the gut microbiota, potentially facilitating the proliferation of pathogenic species associated with PWD. The deep PromethION m etagenomic sequencing also detected the intestinal eukaryotes *Blastocystis* spp. and *E. bieneusi* previously associated with severe diarrhoea and other intestinal diseases in livestock, including pigs, alongside bacterial pathogens [70]. The consistent detection of *E. bieneusi* in both diarrhoeal and healthy pigs suggests it could also represent a common commensal rather than a specific pathogen, supported by the fact that some *E. bieneusi* genotypes are known to be zoonotic, while others are host-adapted [71, 72]. In contrast, the presence of *Blastocystis* in diarrhoeal samples only may indicate gut dysbiosis in affected animals [73].

Relative abundance and genome coverage metrics provide useful but non-definitive guidance for diagnostic interpretation. In our data, *E. coli* lineages and other pathogenic taxa detected at higher sequencing depth and broader genome coverage, particularly when supported by detection of relevant virulence factors (e.g., ETEC-associated fimbrial adhesins and enterotoxins, increased number of effector delivery system genes due to the presence of the locus of enterocyte effacement (LEE) pathogenicity island characteristic of EPEC) were prioritised as more likely contributors to disease than taxa detected at low abundance or low coverage. However, high abundance alone does not equate to causality, especially in samples containing multiple closely related *E. coli* strains that share a large, conserved core genome and differ primarily in plasmid-encoded virulence traits.

However, while both culture-based diagnostics and metagenomic profiling can robustly identify the presence of pathogens and virulence determinants, assigning causality in polymicrobial infections remains challenging. But rather than identifying a single causative agent, metagenomic sequencing provides a framework to characterise the overall pathogenic potential of the microbiome, capturing strain diversity, virulence burden, and antimicrobial resistance context. For PWD, this broader perspective may be more informative for understanding disease dynamics and guiding interpretation than a strict single-pathogen model.

Implications for antimicrobial resistance surveillance

Both the *E. coli* isolates and the metagenomes displayed the highest prevalence of ARGs against aminoglycoside, β-lactam and macrolide antibiotics, consistent with antibiotic use patterns in Danish pig production. Specifically, the aminoglycosides neomycin and apramycin, the β-lactam amoxicillin and the macrolide tylosin.

Resistance to three and more antibiotics in the majority of the cultured ETEC isolates (MDR-ETEC) indicate considerable strain and plasmid diversity similar to ETEC described in previous studies [74, 75]. All MDR-ETEC isolates in this study also carried one of the genes encoding F4 or F18 fimbriae and/or genes for enterotoxin production. Co-location of both VFs and ARGs onto conjugative plasmids or pathogenicity islands has been described previously [75] and can have marked implications in pig production as those have enhanced disease-causing abilities and are more challenging to treat. For instance, tetracycline resistance genes were found at the highest depth among all detected ARGs in the metagenomics samples and in the majority of the ETEC isolates, and co-location of *tetB* genes with enterotoxin VF genes

has commonly been reported in pig ETEC pathovars [76, 77]. The predominance of the narrow-spectrum blaTEM-1B gene and the absence of genes encoding extended-spectrum β-lactamases (ESBLs) and AmpC, colistin and fluoroquinolone resistance genes indicate that resistance profiles remain within expected numbers for enteric *E. coli* in Denmark (https://www.danmap.org/).

While isolate-based WGS provides resistance information for individual cultured pathogens, metagenomic sequencing captures the broader ARG landscape within the infected gut microbiome. This community-level information may inform diagnostic interpretation and antimicrobial decision-making in several ways. First, detection of ARGs in non-cultured or co-occurring taxa may help explain treatment failure when resistance is not observed in the cultured isolate. Second, metagenomic resistome profiles can support empiric antimicrobial selection by indicating whether resistance to specific antibiotic classes is widespread within the microbiome, even when phenotypic testing is unavailable or delayed. Finally, the absence of high-risk resistance determinants (e.g. ESBLs, carbapenemases, mcr genes) across the metagenome provides reassurance for antimicrobial stewardship and surveillance purposes. However, metagenomic resistance profiling remains complementary to, rather than an entire replacement for, phenotypic susceptibility testing.

Beyond confirming all pathogens and virulence determinants identified by culture and qPCR, metagenomic sequencing, and specifically the deep PromethION sequencing, provided additional information in all included samples. Specifically, metagenomics detected additional potentially relevant pathogens (e.g., *Salmonella*, *Staphylococcus aureus*, *Acinetobacter baumannii*) not targeted by the qPCR assay in all samples (Fig. 3). Increased number of VFs (additional fimbrial and enterotoxin genes) were detected in 86% of the samples (19 out of the 22), including in two of the samples where no haemolytic *E. coli* could be isolated, in comparison to the corresponding *E. coli* single isolates (Fig. 5). Moreover, in 9 out of the 22 samples (41%), metagenomics detected all fimbrial and enterotoxin genes (Fig. 5). Regarding the ARG repertoire, metagenomics (PromethION) detected antimicrobial resistance genes in 15 antibiotic classes, and none of which was captured by qPCR conventional diagnostics (Figure S8A-B).

Depth and accuracy trade-offs between ONT workflows

Sequencing output strongly influenced detection sensitivity. PromethION generated 20–150 × more sequence data per sample than both GridION approaches, yielding more complete community profiles and greater detection of low-abundance taxa, ARGs and VFs. The GridION workflow, while less sensitive, still achieved accurate pathogen detection in most diarrhoeal samples, as shown previously in healthy pigs [13]. This indicates it could serve as a rapid diagnostic tool when turnaround time and cost outweigh the need for deep genomic characterisation. This depth–resolution trade-off has also been reported in clinical metagenomics, where rapid nanopore sequencing provides actionable results within hours, albeit at lower coverage [78, 79].

Integration of ONT sequencing into veterinary diagnostics

Overall, ONT metagenomics bridges the gap between traditional targeted diagnostics and full genomic resolution, enabling simultaneous detection of pathogens, virulence factors, and ARGs in a single workflow. While standardisation remains necessary for routine use, our findings show that ONT can serve as a practical and scalable platform for rapid pathogen surveillance in veterinary medicine, particularly when time is crucial and culturing is a bottleneck. ONT can also be a great advantage in difficult cases where culturing is not achievable or expected antibiotics are not effective.

The integration of ONT metagenomic sequencing into diagnostics represents a major step toward real-time, genome-wide surveillance of infectious diseases at the animal–human–environment interface. By enabling simultaneous detection of pathogens, virulence factors, and resistance genes directly from faecal material, this approach can accelerate outbreak response, guide antimicrobial stewardship, and strengthen One Health surveillance frameworks, especially with several new initiatives for online pipelines for rapid analyses (without bioinformatic skills requirement). As sequencing costs continue to decline and bioinformatic pipelines become increasingly automated and coupled with AI, ONT platforms could evolve from research tools into routine frontline diagnostics for monitoring emerging bacterial and viral threats in livestock populations.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42523-026-00577-2.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5
Supplementary Material 6
Supplementary Material 7
Supplementary Material 8
Supplementary Material 9

Acknowledgements
We thank Thomas Nordahl Petersen for the help with the analytical part including read mapping, script and data generation.

Author contributions
MI: Data curation, Validation, Formal analysis, Conceptualization, Visualization, Methodology, Investigation, Writing – original draft and review & editing. EEBJ: Data analysis, MAGs, Methodology, Writing – review & editing. BS: Conventional diagnostics and qPCR. FA: Project administration, Resources, Methodology, Supervision, Writing – review & editing, Funding acquisition, Conceptualization. HV: Resources, Methodology, Writing – review & editing, Funding acquisition, Conceptualization. SO: Project administration, Methodology, Supervision, Data curation, Investigation, Software, Conceptualization, Resources, Formal analysis, Validation, Visualization, Project administration, Writing – original draft and review & editing, Funding acquisition.

Funding
The author(s) declare financial support was received for the research and/or publication of this article. This work was supported by EUPAHW (EUPAHW_SOA8) Grant no. 101136346.

Data availability
The raw sequence data of the metagenomic and single isolates sequenced in this study are deposited in the European Nucleotide Archive under the project number PRJEB88111.

Declarations

Ethical
Not applicable. All faecal samples used in this study were obtained as leftover material from routine veterinary diagnostics. No animals were handled, treated, or subjected to any experimental procedures for the purpose of this work. According to Danish legislation on animal experimentation (Animal Experimentation Act, LBK nr 474 of 15/05/2014 and subsequent amendments), the use of discarded animal faeces from routine diagnostics does not require ethical approval. Therefore, no institutional or licensing committee approval was required for this study.

Generative AI statement
The author(s) declare that no AI was used in the creation of this manuscript, either in the writing or figure generation or editing.

Competing interests
The authors declare no competing interests.

Received: 5 December 2025 / Accepted: 17 April 2026
Published online: 28 April 2026

References

1. Bose B, Siva Kumar S. Economic burden of zoonotic and infectious diseases on livestock farmers: a narrative review. J Health Popul Nutr. 2025;44(1):158. https://doi.org/10.1186/s41043-025-00913-3.
2. Eriksen EØ, Kudirkiene E, Christensen AE, Agerlin MV, Weber NR, Nødtvedt A, et al. Post-weaning diarrhea in pigs weaned without medicinal zinc: risk factors, pathogen dynamics, and association to growth rate. Porc Health Manag. 2021;7(1):54. https://doi.org/10.1186/s40813-021-00232-z.
3. Guitart-Matas J, Ballester M, Fraile L, Darwich L, Giler-Baquerizo N, Tarres J, et al. Gut microbiome and resistome characterization of pigs treated with commonly used post-weaning diarrhea treatments. Anim Microbiome. 2024;6(1):24. https://doi.org/10.1186/s42523-024-00307-6.
4. García V, Gambino M, Pedersen K, Haugegaard S, Olsen JE, Herrero-Fresno A. F4- and F18-positive enterotoxigenic *Escherichia coli* isolates from diarrhea of postweaning pigs: genomic characterization. Appl Environ Microbiol. 2020;86(23):e01913–20. https://doi.org/10.1128/AEM.01913-20.
5. Nagy B, Fekete PZ. Enterotoxigenic *Escherichia coli* in veterinary medicine. Int J Med Microbiol. 2005;295(6–7):443–54. https://doi.org/10.1016/j.ijmm.2005.07.003.
6. Bergeland ME, Henry SC. Infectious diarrheas of young pigs. Vet Clin N Am Large Anim Pract. 1982;4(2):389–99. https://doi.org/10.1016/S0196-9846(17)30113-1.
7. Posthaus H, Kittl S, Tarek B, Bruggisser J. *Clostridium perfringens* type C necrotic enteritis in pigs: diagnosis, pathogenesis, and prevention. J Vet Diagn Invest. 2020;32(2):203–12. https://doi.org/10.1177/1040638719900180.
8. Ferreira C, Otani S, Aarestrup FM, Manaia CM. Quantitative PCR versus metagenomics for monitoring antibiotic resistance genes: balancing high sensitivity and broad coverage. FEMS Microbes. 2023;4:xtad008. https://doi.org/10.1093/femsmc/xtad008.
9. Medina JE, Castañeda S, Camargo M, Garcia-Corredor DJ, Muñoz M, Ramírez JD. Exploring viral diversity and metagenomics in livestock: insights into disease emergence and spillover risks in cattle. Vet Res Commun. 2024;48(4):2029–49. https://doi.org/10.1007/s11259-024-10403-2.
10. Hugenholtz P, Tyson TGWM. Metagenomics. Nature. 2008;455(7212):481–83. https://doi.org/10.1038/455481a.
11. Ji B, Qin J, Ma Y, Liu X, Wang T, Liu G, et al. Metagenomic analysis reveals patterns and hosts of antibiotic resistance in different pig farms. Environ Sci Pollut Res. 2023;30(18):52087–106. https://doi.org/10.1007/s11356-023-25962-1.
12. Bloemen B, Gand M, Vanneste K, Marchal K, Roosens NHC, De Keersmaecker SCJ. Development of a portable on-site applicable metagenomic data generation workflow for enhanced pathogen and antimicrobial resistance surveillance. Sci Rep. 2023;13(1):19656. https://doi.org/10.1038/s41598-023-46771-z.
13. Ivanova M, Aarestrup FM, Otani S. Impact of sample multiplexing on detection of bacteria and antimicrobial resistance genes in pig microbiomes using long-read sequencing. Front Microbiol. 2025;16:1597804. https://doi.org/10.3389/fmicb.2025.1597804.
14. Andersen VD, Aarestrup FM, Munk P, Jensen MS, De Knegt LV, Bortolaia V, et al. Predicting effects of changed antimicrobial usage on the abundance of antimicrobial resistance genes in finisher' gut microbiomes. Preventative Vet Med. 2020;174:104853. https://doi.org/10.1016/j.prevetmed.2019.104853.
15. Ståhl M, Kokotovic B, Hjulsager CK, Breum SØ, Angen Ø. The use of quantitative PCR for identification and quantification of *Brachyspira pilosicoli*, *Lawsonia intracellularis* and *Escherichia coli* fimbrial types F4 and F18 in pig feces. Vet Microbiol. 2011;151(3–4):307–14. https://doi.org/10.1016/j.vetmic.2011.03.013.
16. Clausen PTLC, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. BMC Bioinf. 2018;19(1):307. https://doi.org/10.1186/s12859-018-2336-6.
17. De Coster W, Rademakers R. NanoPack2: population-scale evaluation of long-read sequencing data. Bioinformatics. 2023;39(5):btad311. https://doi.org/10.1093/bioinformatics/btad311.
18. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37(5):540–46. https://doi.org/10.1038/s41587-019-0072-8.
19. Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 2014, 30;30(14):2068–69. https://doi.org/10.1093/bioinformatics/btu153.
20. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691–93. https://doi.org/10.1093/bioinformatics/btv421.
21. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. PLoS ONE. 2014;9(8):9. https://doi.org/10.1371/journal.pone.0104984.
22. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37(5):1530–34. https://doi.org/10.1093/molbev/msaa015.
23. Carattoli A, Zankari E, Garciá-Fernández A, Larsen MV, Lund O, Villa L, et al. In silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014, 58;58(7):3895–903. https://doi.org/10.1128/AAC.02412-14.
24. Bessonov K, Laing C, Robertson J, Yong I, Ziebell K, Gannon VPJ, et al. Ectyper: in silico *Escherichia coli* serotype and species prediction from raw and assembled whole-genome sequence data. Microb. Genomics. 2021;7(12):7. https://doi.org/10.1099/mgen.0.000728.
25. Rood JI, Adams V, Lacey J, Lyras D, McClane BA, Melville SB, et al. Expansion of the *Clostridium perfringens* toxin-based typing scheme. Anaerobe. 2018;53:5–10. https://doi.org/10.1016/j.anaerobe.2018.04.011.
26. Pheatmap KR, Heatmaps P. 2010;1.0.13. https://doi.org/10.32614/CRAN.package.pheatmap.

27. Wickham H. Ggplot2: elegant graphics for data analysis. 2nd. Springer International Publishing; 2016 Edition, [PDF].
28. Xu S, Dai Z, Guo P, Fu X, Liu S, Zhou L, et al. ggtreeExtra: compact visualization of richly annotated phylogenetic data. Mol Biol Evol. 2021;38(9):4039–42. https://doi.org/10.1093/molbev/msab166.
29. Bogri A, Jensen EEB, Borchert AV, Brinch C, Otani S, Aarestrup FM. Transmission of antimicrobial resistance in the gut microbiome of gregarious cockroaches: the importance of interaction between antibiotic exposed and non-exposed populations. msystems. 2024;9(1):e01018–23. https://doi.org/10.1128/msystems.01018-23.
30. Jensen EEB, Sedor V, Eshun E, Njage P, Otani S, Aarestrup FM. The resistomes of rural and urban pigs and poultry in Ghana. msystems. 2023;8(5):e00629–23. https://doi.org/10.1128/msystems.00629-23.
31. Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. The J Antimicrob Chemother. 2017;72(10):2764–68. https://doi.org/10.1093/jac/dkx217.
32. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. The J Antimicrob Chemother. 2020, 75;75(12):3491–500. https://doi.org/10.1093/jac/dkaa345.
33. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal Methods. 2016;8(1):12–24. https://doi.org/10.1039/C5AY02550H.
34. Cabanettes F, Klopp C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. PeerJ. 2018;6:e4958. https://doi.org/10.7717/peerj.4958.
35. Oksanen J. Vegan: Community Ecol Package. 2024.
36. Wickham H, François R, Henry L, Müller K, Vaughan D. Dplyr: a grammar of data manipulation. 2014. 1.1.4. https://doi.org/10.32614/CRAN.package.dplyr.
37. Fernandes AD, Reid JN, Macklaim JM, McMurrough TA, Edgell DR, Gloor GB. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. Microbiome. 2014;2(1):15. https://doi.org/10.1186/2049-2618-2-15.
38. Rydal MP, Gambino M, Castro-Mejia JL, Poulsen LL, Jørgensen CB, Nielsen JP. Post-weaning diarrhea in pigs from a single Danish production herd was not associated with the pre-weaning fecal microbiota composition and diversity. Front Microbiol. 2023;14:1108197. https://doi.org/10.3389/fmicb.2023.1108197.
39. Wylensek D, Hitch TCA, Riedel T, Afrizal A, Kumar N, Wortmann E, et al. A collection of bacterial isolates from the pig intestine reveals functional and taxonomic diversity. Nat Commun. 2020;11(1):6389. https://doi.org/10.1038/s41467-020-19929-w.
40. Luo Y, Ren W, Smidt H, Wright A-DG, Yu B, Schyns G, et al. Dynamic distribution of gut microbiota in pigs at different growth stages: composition and contribution. Microbiol. Spectr. 2022;10(3):e00688–21. https://doi.org/10.1128/spectrum.00688-21.
41. Guevarra RB, Hong SH, Cho JH, Kim B-R, Shin J, Lee JH, et al. The dynamics of the piglet gut microbiome during the weaning transition in association with health and nutrition. J Anim Sci Biotechnol. 2018;9(1):54. https://doi.org/10.1186/s40104-018-0269-6.
42. Choudhury R, Middelkoop A, Boekhorst J, Gerrits WJJ, Kemp B, Bolhuis JE, et al. Early life feeding accelerates gut microbiome maturation and suppresses acute post-weaning stress in piglets. Environ Microbiol. 2021;23(11):7201–13. https://doi.org/10.1111/1462-2920.15791.
43. Garas LC, Cooper CA, Dawson MW, Wang J-L, Murray JD, Maga EA. Young pigs consuming lysozyme Transgenic Goat milk are protected from clinical symptoms of enterotoxigenic Escherichia coli infection. J Nutr. 2017;147(11):2050–59. https://doi.org/10.3945/jn.117.251322.
44. Markova J, Langova D, Babak V, Kostovova I. Ovine and caprine strains of corynebacterium pseudotuberculosis on Czech farms—a comparative study. Microorganisms. 2024;12(5):875. https://doi.org/10.3390/microorganisms12050875.
45. Lema NK, Gemeda MT, Woldesemayat AA. Recent advances in metagenomic approaches, applications, and challenges. Curr Microbiol. 2023;80(11):347. https://doi.org/10.1007/s00284-023-03451-5.
46. Sczyrba A, Hofmann P, Belmann P, Koslicki D, Janssen S, Dröge J, et al. Critical assessment of metagenome interpretation—a benchmark of metagenomics software. Nat Methods. 2017;14(11):1063–71. https://doi.org/10.1038/nmeth.4458.
47. Royer C, Patin NV, Jesser KJ, Peña-Gonzalez A, Hatt JK, Trueba G, et al. Comparison of metagenomic and traditional methods for diagnosis of *E. coli* enteric infections. mBio. 2024;15(4):e03422–23. https://doi.org/10.1128/mbio.03422-23.
48. Han H, Wang Z, Zhu S. Benchmarking metagenomic binning tools on real datasets across sequencing platforms and binning modes. Nat Commun. 2025;16(1):2865. https://doi.org/10.1038/s41467-025-57957-6.
49. Meziti A, Rodriguez-R LM, Hatt JK, Peña-Gonzalez A, Levy K, Konstantinidis KT. The reliability of metagenome-assembled genomes (MAGs) in representing natural populations: insights from comparing MAGs against isolate genomes derived from the same fecal sample. Appl Environ Microbiol. 2021;87(6):e02593–20. https://doi.org/10.1128/AEM.02593-20.
50. Chen Z, Grim CJ, Ramachandran P, Meng J. Advancing metagenome-assembled genome-based pathogen identification: unraveling the power of long-read assembly algorithms in Oxford Nanopore sequencing. Microbiol. Spectr. 2024;12(6):e00117–24. https://doi.org/10.1128/spectrum.00117-24.
51. Nelson KE, et al. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen Listeria monocytogenes reveal new insights into the core genome components of this species. Nucleic Acids Res. 2004;32(8):2386–95. https://doi.org/10.1093/nar/gkh562.
52. Huang AD, Luo C, Pena-Gonzalez A, Weigand MR, Tarr CL, Konstantinidis KT. Metagenomics of two severe foodborne outbreaks provides diagnostic signatures and signs of coinfection not attainable by traditional methods. Appl Environ Microbiol. 2017;83(3):e02577–16. https://doi.org/10.1128/AEM.02577-16.
53. Holman DB, Kommadath A, Tingley JP, Abbott DW. Novel insights into the pig gut microbiome using metagenome-assembled genomes. Microbiol. Spectr. 2022;10(4):e02380–22. https://doi.org/10.1128/spectrum.02380-22.
54. Shimizu T, Ohshima S, Ohtani K, Shimizu T, Hayashi H. Genomic map of *Clostridium perfringens* strain 13. Microbiol Immunol. 2001;45(2):179–89. https://doi.org/10.1111/j.1348-0421.2001.tb01278.x.
55. Gulliver EL, Adams V, Marcelino VR, Gould J, Rutten EL, Powell DR, et al. Extensive genome analysis identifies novel plasmid families in Clostridium perfringens. Microb. Genomics. 2023, 9;9(4). https://doi.org/10.1099/mgen.0.000995.
56. Lacey JA, Allnutt TR, Vezina B, Van TTH, Stent T, Han X, et al. Whole genome analysis reveals the diversity and evolutionary relationships between necrotic enteritis-causing strains of Clostridium perfringens. BMC Genomics. 2018;19(1):379. https://doi.org/10.1186/s12864-018-4771-1.
57. Zhu J, Sun Y, Ma L, Chen Q, Hu C, Yang H, et al. Comparative analysis of fecal microbiota between diarrhea and non-diarrhea piglets reveals biomarkers of gut microbiota associated with diarrhea. Anim nutr. 2024;19:401–10. https://doi.org/10.1016/j.aninu.2024.05.013.
58. Bhattacharjee D, Flores C, Woelfel-Monsivais C, Seekatz AM. Diversity and prevalence of clostridium innocuum in the human gut microbiota. mSphere. 2023;8(1):e00569–22. https://doi.org/10.1128/msphere.00569-22.
59. Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, et al. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. J Bacteriol. 2008;190(20):6881–93. https://doi.org/10.1128/JB.00619-08.
60. De Toro M, Garcilláon-Barcia, De La Cruz F, Garcilláon-Barcia MP. Plasmid diversity and adaptation analyzed by massive sequencing of *Escherichia coli* plasmids. Microbiol. Spectr. 2014;2(6):2.6.32. https://doi.org/10.1128/microbiolspec.PLAS-0031-2014.
61. Von Mentzer A, Blackwell GA, Pickard D, Boinett CJ, E E, Page AJ, et al. Long-read-sequenced reference genomes of the seven major lineages of enterotoxigenic Escherichia coli (ETEC) circulating in modern time. Sci Rep. 2021;11(1):9256. https://doi.org/10.1038/s41598-021-88316-2.
62. Johnson TJ, Nolan LK. Pathogenomics of the virulence plasmids of *Escherichia coli*. Microbiol Mol Biol Rev. 2009;73(4):750–74. https://doi.org/10.1128/MMBR.00015-09.
63. Nelson WC, Tully BJ, Mobberley JM. Biases in genome reconstruction from metagenomic data. PeerJ. 2020;8:e10119. https://doi.org/10.7717/peerj.10119.
64. Kirstahler P, Teudt F, Otani S, Aarestrup FM, Pamp SJ. A peek into the plasmidome of Global sewage. msystems. 2021;6(3):e00283–21. https://doi.org/10.1128/mSystems.00283-21.
65. Luppi A. Swine enteric colibacillosis: diagnosis, therapy and antimicrobial resistance. Porc Health Manag. 2017;3(1):16. https://doi.org/10.1186/s40813-017-0063-4.

66. Rhouma M, Fairbrother JM, Beaudry F, Letellier A. Post weaning diarrhea in pigs: risk factors and non-colistin-based control strategies. Acta Vet Scand. 2017;59(1):31. https://doi.org/10.1186/s13028-017-0299-7.
67. Lee W, Kim M-H, Sung S, Kim E, An ES, Kim SH, et al. Genome-based Characterization of hybrid shiga toxin-producing and enterotoxigenic Escherichia coli (STEC/ETEC) strains isolated in South Korea, 2016–2020. Microorganisms. 2023;11(5):1285. https://doi.org/10.3390/microorganisms11051285.
68. Blirup-Plum SA, Jensen HE, Nielsen SS, Pankoke K, Hansen MS, Pedersen KS, et al. Gastro-intestinal lesions are not relatable to diarrhoea or specific pathogens in post-weaning diarrhoea (PWD) in pigs. Acta Vet Scand. 2023;65(1):30. https://doi.org/10.1186/s13028-023-00693-y.
69. Eriksen EØ, Kudirkiene E, Barington K, Goecke NB, Blirup-Plum SA, Nielsen JP, et al. An observational field study of porcine post-weaning diarrhea: clinical and microbiological findings, and fecal pH-measurements as a potential diagnostic tool. Porc Health Manag. 2023;9(1):33. https://doi.org/10.1186/s40813-023-00325-x.
70. Yu X, Mu X, Yuan K, Wang S, Li Y, Xu H, et al. Occurrence and genotypic identification of Blastocystis spp. and Enterocytozoon bieneusi in Bamaxiang pigs in Bama Yao Autonomous County of Guangxi Province, China. Animals. 2024;14(22):3344. https://doi.org/10.3390/ani14223344.
71. Zhao W, Wang Y, Xin X, Liu J, Zhang X, Yan B, et al. Investigating Enterocytozoon bieneusi in pigs farmed in Zhejiang Province. In: The veterinary journal. China: Occurrence, genotype identification, evolutionary analysis, and zoonotic risk assessment. Vol. 306. 2024. p. 106191. https://doi.org/10.1016/j.tvjl.2024.106191.
72. Dashti A, Santín M, Köster PC, Bailo B, Ortega S, Imaña E, et al. Zoonotic *Enterocytozoon bieneusi* genotypes in free-ranging and farmed wild ungulates in Spain. Med Mycology. 2022;60(9):myac070. https://doi.org/10.1093/mmy/myac070.
73. Asghari A, Sadrebazzaz A, Shamsi L, Shams M. Global prevalence, subtypes distribution, zoonotic potential, and associated risk factors of Blastocystis sp. in domestic pigs (Sus domesticus) and wild boars (Sus scrofa): a systematic review and meta-analysis. Microb Pathog. 2021;160:105183. https://doi.org/10.1016/j.micpath.2021.105183.
74. Fu Y, Nawrocki EM, M'ikanatha NM, Dudley EG. Host species shapes genotype, antimicrobial resistance, and virulence profiles of enterotoxigenic *Escherichia coli* (ETEC) from livestock in the United States. Appl Environ Microbiol. 2024;90(8):e00749–24. https://doi.org/10.1128/aem.00749-24.
75. Smith MG, Jordan D, Chapman TA, Chin J-C, Barton MD, Do TN, et al. Antimicrobial resistance and virulence gene profiles in multi-drug resistant enterotoxigenic escherichia coli isolated from pigs with post-weaning diarrhoea. Vet Microbiol. 2010;145(3–4):299–307. https://doi.org/10.1016/j.vetmic.2010.04.004.
76. Olasz F, Fekete PZ, Blum-Oehler G, BoldogkÅi Z, Nagy B. Characterization of an F18+ enterotoxigenic *Escherichia coli* strain from post weaning diarrhoea of swine, and of its conjugative virulence plasmid pTC. FEMS Microbiol Lett. 2005;244(2):281–89. https://doi.org/10.1016/j.femsle.2005.01.057.
77. Goswami PS, Gyles CL, Friendship RM, Poppe C, Kozak GK, Boerlin P. Effect of plasmid pTENT2 on severity of porcine post-weaning diarrhoea induced by an O149 enterotoxigenic escherichia coli. Vet Microbiol. 2008;131(3–4):400–05. https://doi.org/10.1016/j.vetmic.2008.04.007.
78. Cheng H, Sun Y, Yang Q, Deng M, Yu Z, Zhu G, et al. A rapid bacterial pathogen and antimicrobial resistance diagnosis workflow using Oxford nanopore adaptive sequencing method. Briefings Bioinf. 2022;23(6):bbac453. https://doi.org/10.1093/bib/bbac453.
79. Freeman CN, Herman EK, Abi Younes J, Ramsay DE, Erikson N, Stothard P, et al. Evaluating the potential of third generation metagenomic sequencing for the detection of BRD pathogens and genetic determinants of antimicrobial resistance in chronically ill feedlot cattle. BMC Vet Res. 2022;18(1):211. https://doi.org/10.1186/s12917-022-03269-6.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.