npj | science of food

Article

Published in partnership with Beijing Technology and Business University & International Union of Food Science and Technology

https://doi.org/10.1038/s41538-025-00452-5

Application of MSB selective enrichment followed by amplicon MinION sequencing for multipathogen detection in smoked salmon

Ana Costa-Ribeiro[1,2], Alexandre Lamas[3], Marta Prado[1,3] & Alejandro Garrido-Maestu[1,4]

Millions of foodborne infections are reported yearly worldwide due to a variety of pathogens. A promising approach to tackle this issue relies on Next Generation Sequencing (NGS). In this work a novel multi-foodborne pathogen detection was developed. The method combines a selective enrichment step in a novel broth, multiplex PCR with interlaboratory-validated primers, long-read Flongle MinION sequencing, and data analysis in three cloud-based pipelines to overcome complex, command line-based bioinformatic data analyses. The method, was evaluated in salmon samples spiked with fresh, heat and cold stressed, bacterial cultures of *Salmonella* spp., *E. coli* O157:H7 and *L. monocytogenes*, as well as with in-house, and commercial mock communities, doped with *Y. enterocolitica* and thermotolerant *Campylobacter* spp. No major deviations from the expected results were obtained, reaching a limit of detection <10 CFU/25 g for *Salmonella* spp., *E. coli* O157 and *L. monocytogenes* with sensitivity, specificity and accuracy values > 90% regardless the bioinformatic pipeline selected and concordance values between 0.9 to 1.0. Taken together, the proposed method allows for simple and reliable implementation of NGS as testing tool to streamline foodborne pathogen detection, and overcoming the typical limitations associated with this technology.

Detection of foodborne pathogens remains a major challenge worldwide. Even though the advent of molecular biology techniques, like qPCR, has allowed to overcome certain limitations of culture-based methods, such as results subjectivity and turnaround time, others remain. One of these might be multipathogen detection that is limited by two main factors. One of them is related to technological development as, for instance, regular laboratory equipment may differentiate up to five fluorophores thus, this is the maximum number of targets detectable by qPCR. The other major challenge relies of the fact that bacterial pathogen detection in food samples can be challenging task due to low initial pathogen concentration, presence of high amounts of interfering microorganisms, and food matrix interference[1,2]. A typical way to overcome these limitations relies on selective enrichment of foodstuffs, as this would allow to recover stressed bacteria, increase the concentration of the pathogens of interest, while limiting the growth of the regular microbiota[3,4]. A step further in this approach relies on the selection of a selective broth suitable for several pathogens so that by using one single medium, several bacteria can be detected with downstream techniques such as PCR, qPCR, LAMP, among others[5].

This approach may be combined with a non-targeted detection technique such as Next-Generation Sequencing (NGS) which would allow to detect virtually any pathogen present in the sample[6,7]. Even though this approach seems simple and straight forward, it was previously demonstrated that completely removing interfering background microbiota, and reaching sufficiently high pathogen concentration without appropriate treatment is not always an easy task[8,9]. Additionally, host DNA, due to its high concentration compared to the target microbes, has also been reported to be problematic to get the desired microbial signal[10,11]. Attempts to overcome this problem have already been reported, being 16S amplicon sequencing the most popular. Unfortunately, this approach tends to provide low taxonomic resolution[12], and when seeking for pathogens, the 16S of the

[1]International Iberian Nanotechnology Laboratory, Braga, Portugal. [2]Department of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain. [3]Food Hygiene, Inspection and Control Laboratory (LHICA), Department of Analytical Chemistry, Nutrition and Bromatology, Veterinary School, Campus Terra, University of Santiago de Compostela (USC), Lugo, Spain. [4]Laboratory of Microbiology and Technology of Marine Products (MicroTEC), Instituto de Investigaciones Marinas (IIM), CSIC, Eduardo Cabello, 6, Vigo, Spain. e-mail: alejandro.garrido@inl.int; agarrido@iim.csic.es

natural microbiota, present in higher concentration, will also be sequenced, making difficult to identify the real threats present[13]. Alternatively, more oriented approaches may be followed, such is the case of the "Quasimetagenomics" method developed by Hyeon et al.[14] which provided good results when targeting *Salmonella* spp., *L. monocytogenes* and *E. coli* O157:H7[15,16] or the "Semi-targeted" method developed by Azinheiro et al. which consisted on performing parallel selective enrichments of a number of pathogens which may be later combined for MinION, long-read sequencing analysis[17]. However, even though these approaches reported good results they also present limitations. The "Quasimetagenomics" method relies on several steps of selective enrichment, immunomagnetic separation and whole genome amplification prior to the sequencing analysis, and it is limited to one single pathogen. On the other hand, the "Semi-targeted" approach allows for the detection of several pathogens but requires the use of different culture media. One last limitation related to NGS-based methods, relies on the bioinformatic analysis of the data retrieved, as well as the cost of sequencing platforms[18].

The goal of the current study was to develop a flexible methodology suitable for multipathogen detection, and capable of overcoming most, if not all, the limitations presented above. To this end, the method included a selective enrichment step in a novel broth, to recover *Salmonella* spp., *E. coli* O157:H7/ STEC and *L. monocytogenes*, while reducing the interfering bacteria. This step was followed by a mPCR amplification of species-specific genes taking advantage of interlaboratory validated primers, to further increase the sequences associated to the pathogens, reducing DNA from other bacteria as well as from the host, and opening the door for the detection of additional genes associated to these bacteria, or other bacteria such as *Y. enterocolitica* or thermotolerant *Campylobacter* spp. as demonstrated at the end of the study with mock communities. The last steps of the method included multiplex MinION sample sequencing of the amplicons, which is a highly affordable third generation sequencing platform, taking advantage of Flongle flow cells to keep the cost per assay constrain, followed by a simple, automated, cloud-based data analysis.

Results and discussion

Application of the methodology in salmon spiked samples

In the present work a novel selective broth, MSB, was evaluated for the simultaneous detection of three main bacterial pathogens namely, *Salmonella* spp., *E. coli* O157, and other Shiga toxin-producing serotypes (STEC), and *L. monocytogenes*, by targeted MinION sequencing. Costa-Ribeiro et al. already demonstrated the suitability of the MSB for the detection of these pathogens by qPCR in ready-to-eat salad samples[19]. Even though good results were reported, multiplex qPCR-based methods can detect a limited number of targets depending on the thermocycler used. Contrary to this, by metagenomic NGS the identification of any microbial threat present in a given sample is feasible, as it does not target any particular species[20]. However, the presence of high concentration of host DNA and microbial competitors, may hinder the accurate identification of the pathogens potentially present as these tend to be in lower concentrations[21]. Including a selective enrichment step partially solves these problems as by diluting the sample, the host DNA decreases, and the growth of part of the companion microbiota is limited, favoring the pathogens of interest[1,22]. In addition to this, most metagenomic-based methods tend to perform 16S rRNA PCR, and amplicon sequencing to improve the sensitivity however this approach also enriches the DNA from non-pathogenic bacteria, and in many cases, it does not provide enough resolution for serotype, or even species, differentiation[23].

Contrary to what has been stated, in the present study a novel targeted sequencing method is presented. As a proof of concept, smoked salmon samples were spiked with *Salmonella* spp., *E. coli* O157, and other Shiga toxin-producing serotypes (STEC), and *L. monocytogenes*. The method included a selective enrichment in MSB to limit the growth of the natural microbiota followed by mPCR and MinION sequencing. This method addresses most limitations previously highlighted. Including an enrichment step with a standard 1/10 dilution allows to reduce the interference associated to foodstuffs, and likelihood of detecting non-viable bacteria. Likewise, by using a selective broth the presence of non-pathogenic bacteria in the final sample is also limited. Additionally, by implementing a pathogen-specific mPCR, the interference of natural food microbiota is also reduced. One last advantage worth mentioning relies on the fact that, at least, three major foodborne pathogens, *Salmonella* spp., *E. coli* O157, and *L. monocytogenes*, can be co-enriched in the broth selected, thus increasing the throughput of the method, saving hands-on work and reducing costs.

To obtain reliable results with this type of approach it is of outmost importance to perform a reliable PCR amplification, and this is greatly dependent on the genes and primers selected. To this end, we targeted the genes *invA*, *rfbE* and *prfA* for the detection of *Salmonella* spp., *E. coli* serogroup O157 and *L. monocytogenes*, while the detection of pathogenic *Y. enterocolitica* was based on the presence of the gene *ail*, and finally thermotolerant *Campylobacter* spp. were detected with a specific fragment of the 16S. The oligonucleotides chosen have all been previously validated in interlaboratory ring trials[24–28]. In addition to these, the detection of STEC was based on the presence of the *stx* gene, the selected primers allowed to detect *stx*1 and *stx*2[29], and they have been included as part of the ISO method for the detection of STEC[30].

To perform the evaluation of this novel method, a total of 26 samples were analyzed including 21 simultaneously spiked with *Salmonella* spp., *E. coli* O157, and *L. monocytogenes* in the range of 1–100 CFU/ 25 g, along with 5 negative, non-spiked samples, see Table 1. Within the 21 spiked samples, 6 were inoculated with thermally stressed bacteria, 2 heat-treated, 2 stored overnight at 4 °C, and another 2 heat-treated and stored overnight at 4 °C, to better assess the performance of the method with stressed microorganisms, and the capacity of the MSB broth to recover them, an important aspect not covered in the original study of Costa-Ribeiro et al.[19].

The first step in the evaluation of the method presented herein was to determine its LOD taking advantage of the mathematical model described by Wilrich and Wilrich[31], version 12, and recommended by international standards[32,33]. As NGS data results are dependent upon the bioinformatic tools used, three different user-friendly workflows were explored, WIMP[34], CZ_ID[35], and Galaxy[36]. It is noteworthy that after analyzing the reads with the three workflows, similar results were obtained in terms of LOD50 as in all cases the values of *Salmonella* spp. and *L. monocytogenes* were below 5 CFU/25 g, and below 10 CFU/25 g for *E. coli* O157 as shown in Tables 1 and 2, and Fig. 1A–C.

Once this critical parameter was set, we proceeded with the evaluation of the overall performance of the method taking advantage of the results from the samples spiked above the corresponding LOD50s. In all cases one ND was detected in a sample spiked above the LOD50 (15, 41 and 27 CFU/ 25 g for *Salmonella* spp., *E. coli* O157 and *L. monocytogenes* respectively) which resulted negative with all three workflows. This deviation was associated to a low total number of reads obtained after the sequencing experiment, while all the other samples, including the negative ones, obtained more than 200 reads This particular sample had only 59 reads highlighting the importance of sequencing depth[37]. For this reason, the sample was re-sequenced and in this second experiment the number of reads increased up to the range of 20,000, with 8000–14,000 reads classified depending on the workflow followed. This time, all three pathogens present were correctly detected, thus the sample was not considered as a ND. This observation, along with the rest of the results obtained, suggest that a minimum of 200 reads per barcode are needed to reach the cutoff of 50 specific reads per species, to consider a sample as positive. In the present method Flongle flow cells, along with sample multiplexing, were used to keep the cost per sample constrained, however, running single samples in a Flongle flow cell, or replacing these for regular MinION flow cells are both strategies to increase the likelihood of retrieving a higher number of reads. At present, implementing a regular flow cell would cause a price increase, for the same set of 6 samples, of more than 600 €; even if acquire the largest pack of flow cells, which would allow to decrease the price from 700 € to 535 € would have a large impact as the current price for an individual Flongle flow cell is below 70 € (809 € for a 12 flow cell pack, not sold individually). In

addition to flow cell changes, choosing an alternative DNA extraction may result in further cost reduction, however PowerSoil Pro is a highly contested kit, extensively applied for metagenomic studies[38,39].

When focusing the analysis in WIMP, all the samples spiked above the calculated LOD50 returned the expected positive results, which were further confirmed by aligning the reads against the corresponding reference genomes using Geneious. Regarding the SP, even though one negative sample reached the minimum number of reads to be considered positive for *E. coli*, it was confirmed to be negative after Geneious analysis, as no reads aligned neither in the *rfbE* nor in the *stx* genes, as would be expected for positive samples as shown in Supplementary Fig. S1B–D, thus the SP obtained was 100%, and this resulted in an AC higher than 100%, and k of 1.0 for all the pathogens tested (Table 2).

When the analyses were done with CD_ID, one ND was obtained for *L. monocytogenes* in a sample spiked with 17 CFU/25 g, and another for *E. coli* O157 spiked with 43 CFU/25 g. These deviations resulted in a slight decrease in the SE which ranged from 93.3 to 94.7%, but the SP was maintained in 100%. Taken together now the AC when using this workflow ranged from 95 to 95.8%, and the k was stablished in 0.9–1.0. *Salmonella* spp. was correctly identified in all samples, as shown in Table 2.

Lastly, performing the analyses in Galaxy, intermediate results to those obtained with WIMP and CZ_ID were obtained. *Salmonella* spp. and *L. monocytogenes* were correctly identified, as well as all the negative samples, however, one deviation associated to *E. coli* O157 was recorded in a sample spiked with 43 CFU/ 25 g. These results generated SE values ranging from 90.9 to 100%, SP of 100%, giving ACs between 93.8 and 100% with k values of 1.0 for *Salmonella* spp. and *L. monocytogenes* and 0.9 for *E. coli* O157. In Fig. 2 the percentage of reads obtained in each samply is sample is shown.

As it may be observed by the data presented so far, regardless the bioinformatic workflow followed, the method developed returns excellent results as demonstrated by the Cohen's k values, as values between 0.9 to 1.0 were obtained for all pathogens, and it has been reported that values higher than 0.81 for this parameter are already interpreted as "almost complete concordance"[40].

Table 1 | Inoculation pattern

Treatment*	N	*Salmonella* spp.	*E. coli* O157:H7	*L. monocytogenes*
No	2	15	41	27
No	1	20	15	25
No	2	16	15	24
No	2	11	8	17
No	2	4	13	11
No	2	3	3	5
No	2	1	1	3
No	2	1	<1	5
No	5	–	–	–
Heat	2	11	8	17
Cold	2	64	43	24
Heat-Cold	2	64	43	24

Numeric values provided are expressed in CFU/ 25 g.*"No" treatment were samples processed directly after inoculation; "heat" indicates bacterial cultures heated at 60 °C for 10 min; "cold" denotes samples which were stored at 4 °C overnight after inoculation; and "heat-cold" were samples where the bacterial cultures were heater prior to spiking, and after that, were stored refrigerated overnight before starting the analysis.

Evaluation of the methodology in cold and heat-stressed bacteria

Stressed bacteria may also be challenging to detect as they may not properly grow in selective media[41]. To address this issue, in the present study a total of six samples, in groups of two, were submitted to different types of thermal stress, which included heat treatment at 60 °C, cold treatment storage of spiked samples at 4 °C overnight, and a combination of these two by heating the pure bacterial cultures, spiking and storing those samples refrigerated overnight. Regardless the type of stress applied, all samples were correctly identified using WIMP. One sample was not correctly detected neither by CZ_ID, nor by Galaxy. This particular sample was inoculated with 43 CFU/ 25 g, followed by the cold stress treatment. Upon sequencing, it was correctly identified with WIMP (159 reads) but neither with CZ_ID, nor by Galaxy, the minimum threshold to be considered positive, 50 reads, was not reached, as only 18 and 23 reads were obtained with these two workflows. These differences may be related to differences in taxonomic assignment program, as CZ_ID implements minimap2 instead of Kraken like the other two pipelines. These results highlight the importance of stablishing an appropriate bioinformatic pipeline as already indicated by several previous authors[42–44]. A key element of any bioinformatic pipeline relies on proper selection of the database to perform the taxonomic assignment; note that WIMP uses RefSeq, while Galaxy and CZ_ID the Standard_prebuilt_pluspf_2022-06-07 and NCBI nt respectively, being this another potential reason behind this discrepancy. Importance of database selection was also highlighted by Siegwald et al.[45]. In Fig. 2 the percentage of reads obtained with stressed bacteria is shown.

Evaluation of the methodology with mock communities

The last part of the evaluation of the methodology presented included the analysis of mock communities to incorporate reference material from a different source rather than our laboratory, and to assess the potential to detect a wider number of pathogens, in this case, *Campylobacter* spp. and *Y. enterocolitica*.

The method provided satisfactory results with the commercial mock community regardless the workflow followed as *Salmonella* spp. and *L. monocytogenes* were detected, while it was negative for *E. coli* O157/ STEC,

Table 2 | Summary of the results obtained for the LOD50 and the performance parameters

Workflow	Microorganism	LOD50	PA	PD	NA	ND	SE	SP	AC	k
WIMP	*Salmonella* spp.	3.3	15	0	5	0	100	100	100	1.0
	E. coli O157	9.1	11	0	5	0	100	100	100	1.0
	L. monocytogenes	3.5	19	0	5	0	100	100	100	1.0
CZ_ID	*Salmonella* spp.	2.2	16	0	5	0	100	100	100	1.0
	E. coli O157	8.1	13	0	5	1	93.3	100	95.0	0.9
	L. monocytogenes	4.6	18	0	5	1	94.7	100	95.8	0.9
Galaxy	*Salmonella* spp.	2.9	14	0	5	0	100	100	100	1.0
	E. coli O157	9.1	9	0	5	1	90.9	100	93.8	0.9
	L. monocytogenes	3.5	18	0	5	0	100	100	100	1.0

LOD50 expressed in CFU/25 g.

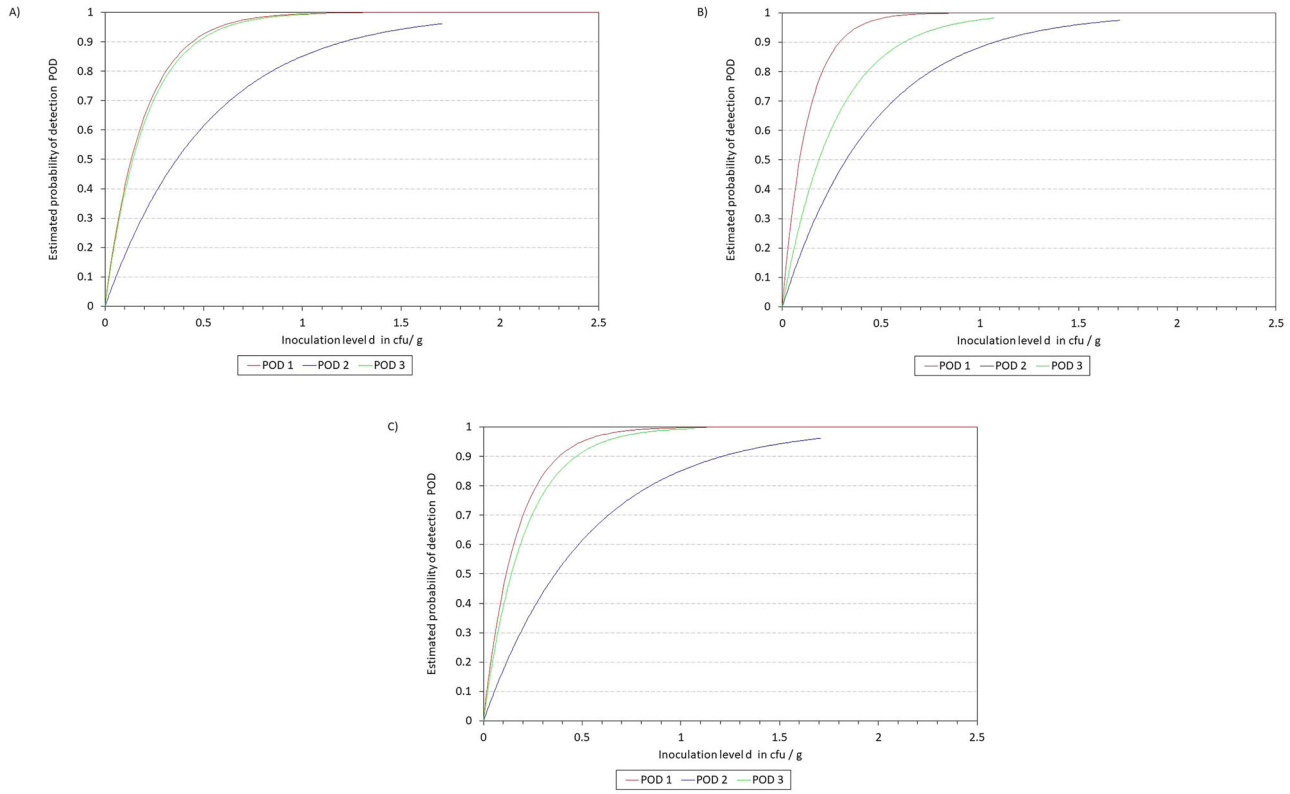

Fig. 1 | Graphical representation of the estimated POD for different pathogens. Estimated POD 1 (*Salmonella* spp.), POD2 (*E. coli* O157:H7) and POD 3 (*L. monocytogenes*) after data analysis with WIMP (**A**), CZ_ID (**B**), or Galaxy (**C**). Plot extracted from the PODLOD function of Wilrich & Wilrich.

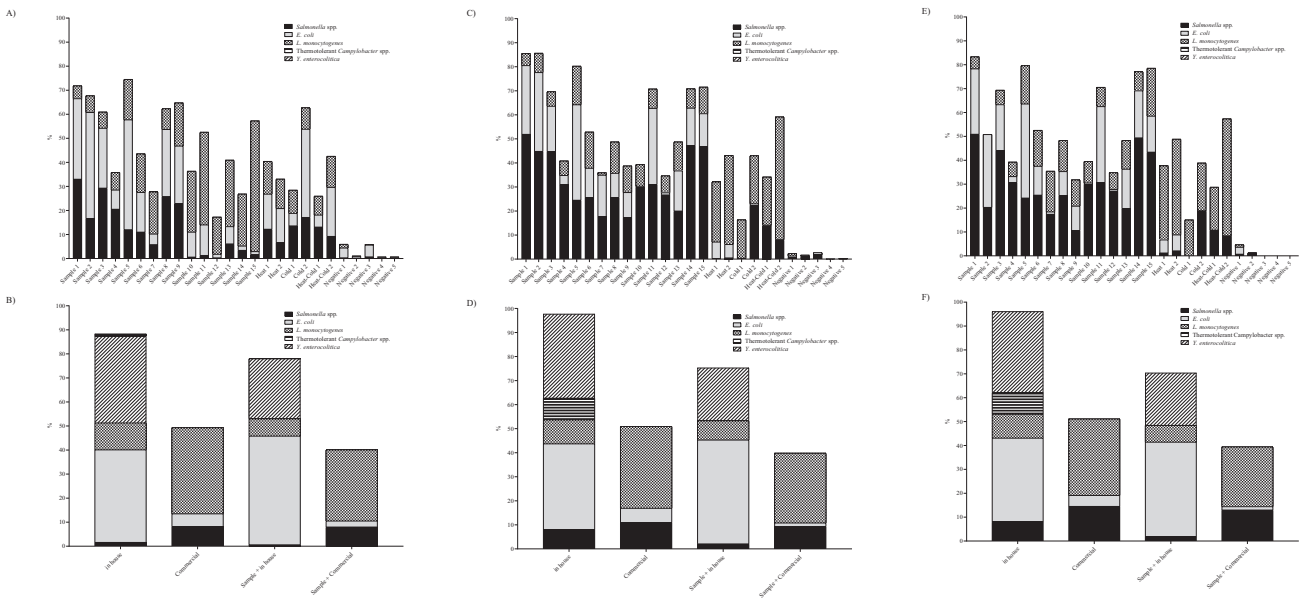

Fig. 2 | Percentage reads obtained for the different pathogens. The result from in spiked and negative samples or the in-house and commercial mock communities when analyzed with WIMP (**A**, **B**), CZ_ID (**C**, **D**) or Galaxy (**E**, **F**).

Y. enterocolitica and *Campylobacter* spp. It must be kept in mind that even though it had *E. coli*, the strain included by the manufacturer was neither O157 nor STEC. These results were consistent with native mock community, as well as after mixing it with DNA from a negative sample, thus demonstrating lack of interference from the food matrix.

When the analysis was focused on the in-house mock community, all five pathogens were successfully detected following the three workflows however, this time upon mixing it with the DNA from the negative sample resulted in lack of detection of *Salmonella* spp., *Campylobacter* spp., and the *stx* genes. After careful analysis, it was determined that the problem was related to how the samples were prepared since prior to mPCR the mixing of the different pathogen DNAs and the negative sample may have resulted in excessively low concentration hindering its amplification and subsequent detection by MinION sequencing. This result would be in agreement with the observation that about 20 times less reads were obtained. Another important point is that the commercial mock community had a significantly

higher DNA concentration from each pathogen compared to the one prepared in-house. As indicated by the manufacturer, the preparation purchased had a 12% of DNA from *Salmonella enterica* and another 12% of *L. monocytogenes*, and it was acquired at a final concentration of 100 ng/ µL, as 3 µL were used for mPCR, a total of 36 ng of each pathogen were added while for the in-house preparation the initial concentration of each pathogen was 0.4 ng/µL, which rendered a final concentration of 0.07 ng/ µL after mixing 10 µL of each of the six bacteria mixed to prepare the mock, and resulted in ~0.2 ng of DNA added to the mPCR. Thus, this is likely the reason why the pathogens present were correctly detected in the commercial mock but not in the in-house one. These issues arose given the fact that for these tests the bacteria were not cultured, thus, suitable final concentrations were not reached. We are aware that a current limitation with the results obtained relies on the identification of *Y. enterocolitica* and *Campylobacter* spp. as these were not co-enriched with the other three. However, we believe that this limitation can be easily overcome by performing a parallel enrichment for these pathogens, and mixing the enriched samples prior to DNA extraction, in a similar way as described by Azinheiro et al. in their semi-targeted detection method[17] or Ruiz-Rueda et al. and Jofré et al. in their studies using qPCR for the simultaneous detection of *Salmonella* spp. and *L. monocytogenes*[46,47].

Overall, the method presented in the current study demonstrated to provide reliable results, fulfilling the requirements of LOD and SE for alternative methods requested by NordVal[32]. By combining a selective enrichment step with mPCR, the pathogens under study were accurately detected by MinION sequencing in two working days, in a similar way to other "next-day" detection methods reported but with the added value of a high multiplexing capacity[48,49]. In addition to this, the three simple, and automated, workflows tested demonstrated to be suitable for overcoming the typical bioinformatic constrains associated to NGS-based methods.

To conclude, in the present work we have developed, described and evaluated a next-day, straightforward methodology for multiplathogen detection based on targeted MinION sequencing. To this end, interlaboratory-validated primers were chosen for higher reliability. The method included a selective enrichment step which demonstrated suitable to recover *Salmonella* spp., *E. coli* O157:H7 and *L. monocytogenes* even after thermal stress, either heat or cold treatment. Given the fact that NGS is per se a non-targeted approach, by increasing the number of fragments pre-amplified, it also opened the door for the detection of additional pathogens like *Y. enterocolitica* and thermotolerant *Campylobacter* spp. upon proper enrichment. Lastly, three simple bioinformatic workflows were compared being obtained similar results with all of them, even though slightly better results were obtained with WIMP. The percentage of reads obtained with the different mock communities, directly sequenced and after used to spike salmon samples, is depicted in Fig. 2.

Methods

Bacterial strains and reference material

Salmonella enterica serovar Typhimurium, WDCM 00031, *L. monocytogenes* WDCM 00021 and *E. coli* WDCM 00014, were selected as reference microorganisms for spiking experiments. All three were purchased from the Spanish Type Culture Collection. For all the experiments, overnight fresh cultures were prepared by adding a single colony into Nutrient Broth (NB, Biokar Diagnostics S.A., Allonne, France) and the suspension was incubated at 37 °C. After incubation, the cultures were diluted and *Salmonella* spp. and *E. coli* were plated on Tryptic Soy Agar (TSA, Biokar Diagnostics S.A., Allonne, France), while *L. monocytogenes* was plated on Tryptic Soy Yeast Extract Agar (TSYEA, Biokar Diagnostics S.A., Allonne, France). All the plates were incubated at 37 °C overnight to determine the concentration of viable bacteria present for spiking experiments.

Two different microbial mock communities were analyzed. One was commercially acquired from Zymo Research (ZymoBIOMICS® Microbial Community DNA Standard, D6306, Zymo Research, Orange, CA). The second mock community was in house prepared with the microorganisms detailed in Supplementary Table S1 of the supporting information.

Sample processing

The standard procedure consisted on weighting 25 g of smoked salmon in a BagFilter XF (filter size <20 µm, Interscience, Saint Nom, France) and 225 mL of mTA10 Selective Broth (MSB), described by Costa-Ribeiro et al., were added[19]. The matrixes were then spiked with dilutions prepared from fresh cultures of the reference bacteria; these were generated as described in 2.1. Once spiked, the matrixes were homogenized for 30 s in a stomacher and incubated at 35 °C for 24 h. After the selective enrichment 1 mL aliquots were taken for DNA as detailed below.

To better assess the performance of the sequencing-based method, three different stresses were evaluated. 1) Heating the fresh cultures (60 °C for 10 min) and followed by cooled cooling in water with ice for 2 min), 2) spiking the food sample and storing it at 4 °C overnight prior to adding the enrichment broth, and 3) heating the culture, before adding it to the food sample and followed by cold storage.

With the idea in mind of evaluating the method on a wider range of microorganisms, DNA obtained from negative samples was spiked with DNA from the different mock communities included in this study, and then this was processed by multiplex PCR (mPCR) and subsequent DNA sequencing as detailed below.

DNA extraction

Pure cultures from the bacteria detailed in Supplementary Table S1 were extracted with Monarch® HMW DNA extraction kit (New England Biolabs, Ipswich, MA, USA), following the protocol for Gram-positive bacteria recommended by the manufacturer. DNA concentration was quantified in a Qubit™ 4 Fluorometer, with the Qubit™ 1X dsDNA High Sensitivity (HS) Assay Kit (Invitrogen™, Carlsbad, CA, USA); and the quality of the extracted DNA was assessed based on the 260/280 and 260/230 absorbance ratios measured in a NanoVue Plus™ Spectrophotometer (GE Healthcare Europe GmbH, Portugal). Once quantified, the DNA concentration of all the bacteria was adjusted to 0.4 ng/ µL, and were mixed in equal amounts to form the final mock community.

The DNA extraction from spiked samples was based on the protocol followed by Costa-Ribeiro et al.[19]. Two mL of each contaminated salmon sample was removed and centrifuged at 900 × *g* for 1 min. The supernatants were recovered into another 2 mL tube and centrifuged again at 16,000 × *g* for 5 min. One mL of TE (TE 1X, Tris-HCl 10 mM, EDTA 1 mM) was used to resuspend the pellet obtained and then it was centrifuged under the same conditions as mentioned above. The supernatant was discarded, and the pellet was resuspended in 200 µL of lysozyme-achromopeptidase (20 mg/mL of lysozyme and 1 mg/mL of achromopeptidase prepared in TE 2X with 1.2% of Triton X-100 (Sigma-Aldrich, St. Louis, MO, USA)) and 25 µL of proteinase K (10 mg/ mL. Macherey-Nagel, Düren, Germany). The mixture was incubated at 37 °C for 20 min and then the DNeasy PowerSoil Pro kit (Qiagen, Barcelona, Spain) was used with some minor modifications. Four hundred µL of CD1 buffer was added to the previous solution, then transferred to a PowerBead Pro tube and 400 µL extra was added, the suspension was heated to 65 °C for 10 min and vortexed for 10 min, followed by the standard protocol. The DNA was eluted in 30 µL of elution buffer and centrifuged at 15,000 × *g* for 1 min, the filtrate was recovered and centrifuged again under the same conditions.

Multiplex PCR (mPCR)

DNA extracted in section M 2.3.2, obtained from spiked samples, was used as template for mPCR, prior to MinION sequencing. Specific primers for the amplification of *invA*, *rfbE*, *stx*, *prfA*, *ail* and 16S, were selected for the detection of *Salmonella* spp., *E. coli* O157, Shiga toxin-producing *E. coli*, *Listeria monocytogenes*, pathogenic *Y. enterocolitica* and thermotolerant *Campylobacter* spp. The reaction was performed in a final volume of 20 µL consisting of 10 µL Platinum™ Multiplex PCR Master Mix (Applied Biosystems™, Foster City, CA, USA), 100 nM of each primer, 3 µL of the template DNA, and the remaining volume was filled with nuclease-free water (New England BioLabs, Inc., Ipswich, MA, USA). The amplification was performed using a Veriti Thermal Cycler (Applied Biosystems™, Foster City,

CA, USA) with the following PCR thermal profile, hot start activation 95 °C for 2 min, followed by 40 cycles of 95 °C for 30 s, 60 °C for 90 s and 72 °C for 1 min, with a final extension at 72 °C for 10 min. After amplification, the amplicons obtained were used for sequencing as detailed below. Additional information regarding the primers may be found in Supplementary Table S2 of the supporting information.

Pathogen detection by multiplex qPCR

The confirmation of the results of spiked samples was performed by two different multiplex qPCR assays. The reactions were run in a QuantStudio™ 5 System (Applied Biosystems™, Foster City, CA, USA) configured in "FAST" format. The unified thermal profile selected included a UDG step at 50 °C for 2 min, followed by hot start activation step at 95 °C for 2 min, and then it was followed by 50 cycles of dissociation at 95 °C for 1 s and combined annealing – extension at 63 °C for 20 s. Samples with Cq values lower than 38 were considered positive. All the reactions had a final volume of 20 µL with 10 µL of TaqMan® Multiplex Master Mix (Applied Biosystems™, Foster City, CA, USA) 3 µL of template DNA, the corresponding primers and probes, see supporting information, Supplementary Table S3 for specific concentrations, and the remaining volume was filled with nuclease-free water (New England BioLabs, Inc., Ipswich, MA, USA). One assay included the simultaneous detection of *Salmonella* spp., *E. coli* O157 and *L. monocytogenes* targeting the genes *ttr*, *rfbE* and *hly* respectively, along with an Internal Amplification Control, while the second assay detected pathogenic *Y. enterocolitica* and thermotolerant *Campylobacter* targeting the genes *ail* and a specific region of the 16S (additional details provided in Supplementary Table S3).

MinION sequencing

Libraries were prepared with the Rapid Barcoding Sequencing Kit 24 V14 (SQK-RBK114.24, Oxford Nanopore Technologies, Oxford, UK) following the protocol provided by the manufacturer for the Flongle device using the mPCR amplicon product as input sample. Once prepared, the libraries were loaded in Flongle flow cells (FLO-FLG001). No more than four samples were multiplexed per sequencing experiment.

Bioinformatic analyses

Three different, easy to use, cloud-based bioinformatic workflows were compared, and regardless the workflow followed, positive samples were considered those with more than 50 reads for a given pathogen after confirmation with Geneious as detailed below.

The first workflow was "What's In My Pot" (WIMP). This workflow was described by Juul et al.[34]. It is readily available through EPI2ME™ (https://epi2me.nanoporetech.com) and it can be configured to perform the base calling in real time so that the results can be observed as they are being generated. It uses "centrifuge" *kmer*-based read identification[10]. A minimum abundance cutoff of 0.5% was set for the species identification.

The second workflow was CZ_ID, which is a cloud-based genomic analysis platform, it may be accessed through https://czid.org/. It has recently expanded its workflow to long read metagenomic NGS data derived from Oxford Nanopore[35]. The workflow performs a quality control, and if a reference genome is provided, the reads are aligned using minimap2[50], otherwise de novo assembly is performed with metaFlye[51]. For direct comparison with WIMP, the FASTQ files were uploaded and directly analyzed without reference genome.

The third workflow was performed in Galaxy. Fastq files for each sample were merged in a single Fastq file and uploaded to European server of Galaxy (https://usegalaxy.eu). Taxonomic sequence classifier Kraken2 (Galaxy Version 2.1.1+galaxy1)[52] was used to assign taxonomic labels to the reads. For that purpose, the database Standard_prebuilt_pluspf_2022-06-07 with default conditions was used. Kraken2 output file was used to re-estimate abundance at a taxonomic level with Bracken (Galaxy Version 2.9+galaxy0)[53] using standard_prebuilt_pluspf_2022-06-07 and default conditions.

In order to confirm the results, reference genomes for each of the pathogens under study were retrieved from RefSeq (NC_022569, NZ_CP008957, NC_003210, NC_002163 and NZ_LR134161 for *Salmonella*, *E. coli* O157:H7 and STEC, *L. monocytogenes*, *Campylobacter* spp. and *Y. enterocolitica* respectively) and imported into Geneious Prime® version 2023.0.1. (Biomatters Ltd., Auckland, New Zealand). After the sequencing experiment, the Fastq files were also imported into Geneious and aligned against the different reference genomes using Minimap2[50], to confirm that the species results obtained by each bioinformatic workflow were due to the presence of specific amplicons, in terms of gene and fragment size, for each pathogen. In the particular case of *E. coli*, the presence of amplicons for the *rfbE* and *stx*1 and/ or *stx*2 genes, was screened, see Supplementary Fig. S1A–G.

Determination of the Limit of Detection (LOD)

The LOD was calculated with the version 12 of the mathematical model described by Wilrich and Wilrich[31]. Briefly, smoked salmon samples were simultaneously spiked with decreasing concentrations of the three pathogens under study, seeking for a concentration which will render positive and negative results. The data of bacterial concentration and number positive replicates were input in the model to determine the LOD50 as required by the ISO and NordVal standards[32,33].

Fitness-for-purpose

After stablishing the LOD50 for each one of the pathogens the samples above the corresponding LOD50 were used for the evaluation of the Relative Sensitivity, Specificity and Accuracy (SE, SP and AC) and the Cohen's kappa (k) as described by reported by Anderson et al., and Tomás et al.[54,55]. To this end, the samples were classified as being in Positive or Negative Agreement (PA/ NA) if they matched the expected results, or Deviation (PD/ ND) if they did not.

In addition to the spiked samples, six non-spiked samples were also processed to evaluate the SP of the method. Lastly, as it was not possible to obtain samples naturally contaminated with all the pathogens which could be potentially detected with the current method, an in-house mock community was also analyzed, directly and after mixing it with the DNA of a negative sample to mimic potential matrix effect. In a similar way, a commercial mock community was acquired and analyzed directly and after its mixture with a negative sample.

Data representation

Graphical representation of the data was performed with GraphPad Prism version 8.0.0 for Windows (GraphPad Software, San Diego, California USA, www.graphpad.com).

Data availability

Data will be made available upon request to the corresponding author.

Received: 5 July 2024; Accepted: 9 May 2025;
Published online: 28 October 2025

References

1. Brehm-Stecher, B., Young, C., Jaykus, L.-A. & Mary Lou, T. Sample preparation: the forgotten beginning. *J. Food Prot.* **72**, 1774–1789 (2009).
2. Nero, L. A., de Mattos, M. R., Barros Mde, A., Beloti, V. & Franco, B. D. Interference of raw milk autochthonous microbiota on the performance of conventional methodologies for *Listeriamonocytogenes* and *Salmonella* spp. detection. *Microbiol Res*. **164**, 529–535 (2009).
3. Boukharouba, A., González, A., García-Ferrús, M., Ferrús, M. A. & Botella, S. Simultaneous Detection of Four Main Foodborne Pathogens in Ready-to-Eat Food by Using a Simple and Rapid Multiplex PCR (mPCR) Assay. *Int J Environ Res Public Health* **19**, 1031 (2022).
4. Kim, H. C. & Bhunia, A. K. SEL, a selective enrichment broth for simultaneous growth of *Salmonella enterica*, *Escherichia coli* O157:

H7, and *Listeria monocytogenes*. *Appl. Environ. Microbiol* **74**, 4853–4866 (2008).
5. Qu, Y. et al. SSEL, a selective enrichment broth for simultaneous growth of *Salmonella enterica, Staphylococcus aureus, Escherichia coli* O157: H7, and *Listeria monocytogenes. J. Food Saf.* **40**, e12837 (2020).
6. Gu, W., Miller, S. & Chiu, C. Y. Clinical metagenomic next-generation sequencing for pathogen detection. https://doi.org/10.1146/annurev-pathmechdis (2018).
7. Schlaberg, R., Chiu, C. Y., Miller, S., Procop, G. W. & Weinstock, G. Validation of metagenomic next-generation sequencing tests for universal pathogen detection. *Arch. Pathol. Lab. Med.* **141**, 776–786 (2017).
8. Frey, K. G. & Bishop-Lilly, K. A. Next-generation sequencing for pathogen detection and identification. In: *Methods in Microbiology* vol. 42 525–554 (Academic Press Inc., 2015).
9. Wensel, C. R., Pluznick, J. L., Salzberg, S. L. & Sears, C. L. Next-generation sequencing: insights to advance clinical investigations of the microbiome. *J. Clin. Invest.* **132**, https://doi.org/10.1172/JCI154944 (2022).
10. Charalampous, T. et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. *Nat. Biotechnol.* **37**, 783–792 (2019).
11. Heravi, F. S., Zakrzewski, M., Vickery, K. & Hu, H. Host DNA depletion efficiency of microbiome DNA enrichment methods in infected tissue samples. *J. Microbiol. Methods* **170**, 105856 (2020).
12. Rubiola, S. et al. Comparison between full-length 16S rRNA metabarcoding and whole metagenome sequencing suggests the use of either is suitable for large-scale microbiome studies. *Foodborne Pathog. Dis.* **19**, 495–504 (2022).
13. Billington, C., Kingsbury, J. M. & Rivas, L. Metagenomics approaches for improving food safety: a review. *J. Food Protection* **85**, 448–464 (2022).
14. Hyeon, J.-Y. et al. Quasimetagenomics-based and real-time-sequencing-aided detection and subtyping of *Salmonella enterica* from food samples. *Appl. Environ. Microbiol.* **84**, 1–15 (2018).
15. Ottesen, A. et al. Quasimetagenomic source tracking of *Listeria monocytogenes* from naturally contaminated ice cream. *BMC Infect. Dis.* **20**, 1–6 (2020).
16. Townsend, A., Li, S., Mann, D. A. & Deng, X. A quasimetagenomics method for concerted detection and subtyping of *Salmonella enterica* and *E. coli* O157:H7 from romaine lettuce. *Food Microbiol.* **92**, 103575 (2020).
17. Azinheiro, S., Roumani, F., Carvalho, J., Prado, M. & Garrido-Maestu, A. Suitability of the MinION long read sequencer for semi-targeted detection of foodborne pathogens. *Anal. Chim. Acta* **1184**, 339051 (2021).
18. Cardinali, G., Corte, L. & Robert, V. Next Generation Sequencing: problems and opportunities for next generation studies of microbial communities in food and food industry. *Curr. Opin. Food Sci.* **17**, 62–67 (2017).
19. Costa-Ribeiro, A., Lamas, A., Prado, M. & Garrido-Maestu, A. Evaluation of the novel mTA10 selective broth, MSB, for the co-enrichment and detection of *Salmonella* spp., *Escherichia coli* O157 and *Listeria monocytogenes* in ready-to-eat salad samples. *Foods* **13**, 63 (2024).
20. Lewis, E., Hudson, J. A., Cook, N., Barnes, J. D. & Haynes, E. Next-generation sequencing as a screening tool for foodborne pathogens in fresh produce. *J. Microbiol. Methods* **171**, 105840 (2020).
21. Bloomfield, S. J. et al. Determination and quantification of microbial communities and antimicrobial resistance on food through host DNA-depleted metagenomics. *Food Microbiol.* **110**, 104162 (2023).
22. Kim, J. H. & Oh, S. W. Pretreatment methods for nucleic acid-based rapid detection of pathogens in food: a review. *Food Control* **121**, https://doi.org/10.1016/j.foodcont.2020.107575 (2021).
23. Hilton, S. K. et al. Metataxonomic and metagenomic approaches vs. culture-based techniques for clinical pathology. *Front. Microbiol.* **7**, 484 (2016).
24. Abdulmawjood, A. et al. Toward an international standard for PCR-based detection of *Escherichia coli* O157 - Part 1. Assay development and multi-center validation. *J. Microbiol. Methods* **55**, 775–786 (2003).
25. D'Agostino, M. et al. A validated PCR-based method to detect *Listeria monocytogenes* using raw milk as a food model - Towards an international standard. *J. Food Prot.* **67**, 1646–1655 (2004).
26. Lubeck, P. S., Cook, N., Wagner, M., Fach, P. & Hoorfar, J. Toward an international standard for PCR-based detection of food-borne thermotolerant Campylobacters: validation in a multicenter collaborative trial. *Appl. Environ. Microbiol.* **69**, 5670–5672 (2003).
27. Malorny, B. Interlaboratory diagnostic accuracy of a *Salmonella* specific PCR-based method. *Int. J. Food Microbiol.* **89**, 241–249 (2003).
28. Thisted Lambertz, S. Development of a PCR-based method for detection of pathogenic *Yersinia enterocolitica* in Pork (Dept. of Biomedical Sciences and Veterinary Public health, Swedish University of Agricultural Sciences (2005).
29. Perelle, S., Dilasser, F., Grout, J. L. & Fach, P. Detection by 5 '-nuclease PCR of Shiga-toxin producing *Escherichia coli* O26, O55, O91, O103, O111, O113, O145 and O157: H7, associated with the world's most frequent clinical cases. *Mol. Cell Probes* **18**, 185–192 (2004).
30. ISO/TS. Microbiology of food and animal feed – Real-time polymerase chain reaction (PCR)-based method for the detection of food-borne pathogens – Horizontal method for the detection of Shiga toxin-producing *Escherichia coli* (STEC) and the determination of O157, vol. 13136 (ISO, 2012).
31. Wilrich, C. & Wilrich, P. T. Estimation of the pod function and the LOD of a qualitative microbiological measurement method. *J. AOAC Int.* **92**, 1763–1772 (2009).
32. NordVal. NordVal international protocol for the validation of microbiological alternative (proprietary) methods against a reference method (NordVal, 2017).
33. ISO. ISO 16140-2:2016 Microbiology of the food chain — Method validation — Part 2: Protocol for the validation of alternative (proprietary) methods against a reference method, 16140 (ISO, 2016).
34. Juul, S. et al. What's in my pot? Real-time species identification on the MinION. *bioRxiv* 30742 https://doi.org/10.1101/030742 (2015).
35. Simmonds, S. E. et al. CZ ID: a cloud-based, no-code platform enabling advanced long read metagenomic analysis 2 3. *bioRxiv* https://doi.org/10.1101/2024.02.29.579666 (2024).
36. Gangiredla, J. et al. GalaxyTrakr: a distributed analysis tool for public health whole genome sequence data accessible to non-bioinformaticians. *BMC Genomics* **22**, 114 (2021).
37. Pereira-Marques, J. et al. Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis. *Front. Microbiol.* **10**, 1277 (2019).
38. Josefsen, M. H., Andersen, S. C., Christensen, J. & Hoorfar, J. Microbial food safety: potential of DNA extraction methods for use in diagnostic metagenomics. *J. Microbiol. Methods* **114**, 30–34 (2015).
39. Shaffer, J. P. et al. A comparison of six DNA extraction protocols for 16S, ITS and shotgun metagenomic sequencing of microbial communities. *Biotechniques* **73**, 34–46 (2022).
40. Altman, D. G. Practical statistics for medical research. https://www.taylorfrancis.com/books/mono/10.1201/9780429258589/practical-statistics-medical-research-douglas-altman (Chapman and Hall/CRC, 1990).
41. Bissonnette,, G. K., Jezeski,, J. J., Mcfeters, G. A. & Stuart, D. G. Influence of environmental stress on enumeration of indicator bacteria from natural waters. *Appl. Microbiol.* **29**, 186–194 (1975).
42. Barretto, C., Rincón, C., Portmann, A. C. & Ngom-Bru, C. Whole genome sequencing applied to pathogen source tracking in food

industry: Key considerations for robust bioinformatics data analysis and reliable results interpretation. *Genes* **12**, 1–12 (2021).
43. de Vries, J. J. C. et al. Benchmark of thirteen bioinformatic pipelines for metagenomic virus diagnostics using datasets from clinical samples. *J. Clin. Virol.* **141**, 104908 (2021).
44. D'Argenio, V., Casaburi, G., Precone, V. & Salvatore, F. Comparative metagenomic analysis of human gut microbiome composition using two different bioinformatic pipelines. *Biomed. Res. Int*. **2014**, 325340 (2014).
45. Siegwald, L. et al. Assessment of common and emerging bioinformatics pipelines for targeted metagenomics. *PLoS One* **12**, e0169563 (2017).
46. Jofre, A. et al. Simultaneous detection of Listeria monocytogenes and Salmonella by multiplex PCR in cooked ham. *Food Microbiol.* **22**, 109–115 (2005).
47. Ruiz-Rueda, O., Soler, M. & Calvo, L. Multiplex Real-time PCR for the Simultaneous Detection of *Salmonella* spp. and *Listeria monocytogenes* in Food Samples. *Food Anal. Methods* **68**, 551–556 (2010).
48. Azinheiro, S., Ghimire, D., Carvalho, J., Prado, M. & Garrido-Maestu, A. Next-day detection of viable *Listeria monocytogenes* by multiplex reverse transcriptase real-time PCR. *Food Control* **133**, 108593 (2022).
49. Rodríguez-Lázaro, D. et al. Next day *Salmonella* spp. detection method based on real-time PCR for meat, dairy and vegetable food products. *Int. J. Food Microbiol.* **184**, 113-20 (2014).
50. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
51. Kolmogorov, M. et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat. Methods* **17**, 1103–1110 (2020).
52. Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. http://ccb.jhu.edu/software/kraken/ (2014).
53. Lu, J., Breitwieser, F. P., Thielen, P. & Salzberg, S. L. Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Sci*. **2017**, e104 (2017).
54. Anderson, A. et al. Validation of a duplex real-time PCR for the detection of *Salmonella* spp. in different food products. *Food Anal. Methods* **4**, 259–267 (2011).
55. Tomás, D., Rodrigo, A., Hernández, M. & Ferrús, M. A. Validation of real-time PCR and enzyme-linked fluorescent assay-based methods for detection of *Salmonella* spp. in chicken feces samples. *Food Anal. Methods* **2**, 180–189 (2009).

Acknowledgements

A.G.-M. would like to thank the Fundação para a Ciência e a Tecnologia for the financial support through the Scientific Employment Stimulus Program (2021.02810.CEECIND). M.P. acknowledges the Ministry of Science, Innovation and Universities in Spain for her grant (BG22/00139 - Beatriz Galindo program). Last, the authors thank the financial support of the project "Agenda SMARTgNOSTICS – Global Testing & Diagnostics Solutions for antimicrobial resistances", with the reference n.º C644915155-00000024, co-funded by Component C5 – Capitalisation and Business Innovation under the Portuguese Resilience and Recovery Plan, through the NextGenerationEU Fund.

Author contributions

A.C.-R.: investigation and revision. A.L.: data curation, revision and editing. M.P.: funding acquisition, revision and editing. A.G.-M.: conceptualization, methodology, supervision, validation, writing of the original draft.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41538-025-00452-5.

Correspondence and requests for materials should be addressed to Alejandro Garrido-Maestu.