



Article

Moving Beyond Oxford Nanopore Standard Procedures: New Insights from Water and Multiple Fish Microbiomes

Ricardo Domingo-Bretón ¹, Federico Moroni ¹, Socorro Toxqui-Rodríguez ^{1,2}, Álvaro Belenguer ¹, M. Carla Piazzon ²,
Jaime Pérez-Sánchez ^{1,*} and Fernando Naya-Català ^{1,*}

¹ Nutrigenomics and Fish Growth Endocrinology Group, Instituto de Acuicultura Torre de la Sal (IATS, CSIC), 12595 Ribera de Cabanes, Castellón, Spain; ricardo.domingo@csic.es (R.D.-B.); federico.moroni@csic.es (F.M.); socorro.toxqui@csic.es (S.T.-R.); a.belenguer@csic.es (Á.B.)

² Fish Pathology Group, Instituto de Acuicultura Torre de la Sal (IATS, CSIC), 12595 Ribera de Cabanes, Castellón, Spain; carla.piazzon@csic.es

* Correspondence: jaime.perez.sanchez@csic.es (J.P.-S.); fernando.naya@iats.csic.es (F.N.-C.)

Abstract: Oxford Nanopore Technology (ONT) allows for the rapid profiling of aquaculture microbiomes. However, not all the experimental and downstream methodological possibilities have been benchmarked. Here, we aimed to offer novel insights into the use of different library preparation methods (standard-RAP and native barcoding-LIG), primers (V3–V4, V1–V3, and V1–V9), and base-calling models (fast-FAST, high-HAC, and super-accuracy-SUP) implemented in ONT to elucidate the microbiota associated with the aquatic environment and farmed fish, including faeces, skin, and intestinal mucus. Microbial DNA from water and faeces samples could be amplified regardless of the library–primer strategy, but only with LIG and V1–V3/V1–V9 primers in the case of skin and intestine mucus. Low taxonomic assignment levels were favoured by the use of full-length V1–V9 primers, though in silico hybridisation revealed a lower number of potential matching sequences in the SILVA database, especially evident with the increase in Actinobacteriota in real datasets. SUP execution allowed for a higher median Phred quality (24) than FAST (11) and HAC (17), but its execution time (6–8 h) was higher in comparison to the other models (0.6–7 h). Altogether, we optimised the use of ONT for water- and fish-related microbial analyses, validating, for the first time, the use of the LIG strategy. We consider that LIG–V1–V9–HAC is the optimal time/cost-effective option to amplify the microbial DNA from environmental samples. However, the use of V1–V3 could help to maximise the dataset microbiome diversity, representing an alternative when long amplicon sequences become compromised by microbial DNA quality and/or high host DNA loads interfere with the PCR amplification/sequencing procedures, especially in the case of gut mucus.

Keywords: aquaculture; microbiota; 16S rRNA; third-generation sequencing; Oxford nanopore; MinION; native barcoding kit



Citation: Domingo-Bretón, R.; Moroni, F.; Toxqui-Rodríguez, S.; Belenguer, Á.; Piazzon, M.C.; Pérez-Sánchez, J.; Naya-Català, F. Moving Beyond Oxford Nanopore Standard Procedures: New Insights from Water and Multiple Fish Microbiomes. *Int. J. Mol. Sci.* **2024**, *25*, 12603. <https://doi.org/10.3390/ijms252312603>

Academic Editor: Emina Pustijanac

Received: 30 October 2024

Revised: 20 November 2024

Accepted: 21 November 2024

Published: 23 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Microorganisms form complex communities that played key roles in the preservation of the health and welfare of aquaculture species [1,2]. They are found in mucosal tissues such as the gut, skin, and gills, as well as in the faeces and the water surrounding the farmed fish. Real-time and precise understanding of the dynamics of these populations is key for unveiling the host–microbe network of interactions within the holobiont that directly impact aquaculture productivity and sustainability [3–7]. In this scenario, advances in DNA sequencing technologies, particularly Oxford Nanopore Technologies (ONTs), have emerged as a door-opener for rapid microbiome analysis through full-length 16S rRNA gene amplification [8]. Some efforts within the aquaculture field have already been conducted for designing standardised protocols for this platform. These involved the utilisation of the 16S amplification protocol recommended by ONT for the MinION™

device, equipped with R9.4.1 flow cells [9–12]. With this chemistry, the amount of input DNA material and the PCR conditions (annealing temperature and number of cycles) of the ONT recommended protocol were optimised, maximising the yield of this platform using gilthead sea bream gut mucosal samples and making the technology suitable for further studies [13,14]. However, as an in-development technology, ONT is continuously experiencing equipment updates and improvements in terms of read quality and data output, which need to be taken into account while the work with this technology advances [15–17]. Among these innovations, a keystone was the appearance of R10.4.1 flow cells, accompanied by an evolved chemistry (V14) and updated basecalling algorithms that improve the conversion process transforming raw electrical current data into the final nucleotide sequence [18]. This combination increased ONT read accuracy up to Q20+ quality indexes by reducing sequence anomaly detections. However, to fully exploit these novelties, users must have access to a computer equipped with a sufficiently powerful dedicated graphical processing unit (GPU) or even high-performance computing or cloud environments to manage the data produced by the highest-output ONT sequencers [19–24]. In any case, ONT innovations directly promote not only microbial analyses, but also other novel ONT applications, such as structural variation detection and methylation marks calling, enhancing a more specific taxonomic assignment and the non-redundant genome mapping of long-reads [25]. Therefore, the traditional assumption that ONT is a low-quality platform is being reconsidered, though this quality jump and the extension of the ONT catalogue of resources have not been properly tested using aquaculture-related samples yet.

Besides flow cells, the flexibility of ONT continues with alternative library preparation strategies that admit different primer choices [16,26]. The following two library preparation strategies are possible with ONT for metabarcoding studies: amplification using the 16S Barcoding Kit 24 V14 (RAP) or the ligation-based Native Barcoding Kit V14 after custom-amplicon generation (LIG). Nowadays, RAP is the standard method used for 16S microbiome studies within the aquaculture field [9,12,14]. This kit offers several clear advantages, such as a shorter protocol duration, reduced laboratory resources, and fewer pipetting or sample-handling steps. These features make it an ideal choice for *in situ* microbiome analysis or for use in minimally equipped laboratories. However, in well-equipped molecular biology laboratories, the LIG method offers distinct advantages that should not be overlooked [27]. LIG supports a much higher level of multiplexing, with up to 96 barcodes compared to the 24 barcodes offered by RAP. This makes the sequencing of a larger number of samples more cost-effective and time-efficient. Another major benefit of such an approach is the flexibility that it provides by separating the amplification of the targeted region from the barcoding and library preparation. This allows users to customise the targeted region length to suit their specific research needs and experiment with different primer pair modifications to achieve the best representation of the studied population [28].

The choice of primers is another well-described critical step in 16S rRNA amplification, as it directly influences the accuracy, efficiency, and reliability of microbial identification and diversity studies [29–31]. Certainly, unlike short-read sequencing, longer amplicons ensure low-level taxonomic assignments, usually down to the species rank [32]. However, these amplicons can be difficult to obtain when analysing DNA/RNA samples of animal and human origin, such as skin and gut mucus, among others [33]. The first reason for this is due to the host DNA, which is usually present and very abundant in these types of samples. In fact, host DNA depletion is a technique commonly used in human samples for microbiome enrichment [34], which can be also required in fish studies, as metatranscriptomic datasets of gilthead sea bream highlighted a high host/microbial RNA ratio (9:1) [35]. Secondly, DNA degradation substantially limits the fragment length that a PCR can successfully amplify. This process could be linked to sampling procedures and conditions, or even biological processes, such as the different mucosal epithelial turnover rhythms among seasons or feeding regimes [36]. When long-amplicon sequencing is not possible, the LIG procedure allows for the use of short universal primers covering certain hypervariable regions (i.e., V3–V4 or V1–V3) or ad hoc designed primers [34,37]. Traditionally, ONT

has not been considered as a short-read sequencing platform, but several authors have claimed the high cost-effectiveness of attempting this approach for DNA metabarcoding studies [28,38,39]. However, due to the current lack of research using the LIG method, we are unable to know what the level of comparability is in fish-related 16S amplicon datasets using distinct primer combinations, or if short-amplicon primers also maintain the level of quality that the new R10.4.1 flow cells and basecalling algorithms offer.

Taking into account all the above findings, the higher-quality ONT equipment and LIG protocols need to be benchmarked and compared to standard procedures to choose the optimal strategy that fits the user sequencing needs. Differences in taxonomic assignment, taxa identification bias, execution time, and resource consumption remain elusive when using aquaculture-related samples. Accordingly, in order to understand whether these gaps can be filled, we aimed to optimise the 16S rRNA amplification procedures implemented in ONT when sequencing fish-associated samples of gut and skin mucosa, faeces, and water from several trials, evaluating, at the same time, the feasibility and output differences of different library preparation strategies, primers, and basecalling models.

2. Results

2.1. Assessment Strategy of ONT Performance and Sequencing Results

A total of 36 samples from a mock community (1), *E. coli* pure cultures (1), gut (4) and skin mucus (4), faeces (12), and water (15) were used to test the performance of ONT MinION™ with R10.4.1 flow cells and V14 chemistry (Table 1). A mock community (ZymoBIOMICS™ Microbial Community Standard II) and a pure isolate of *E. coli* DH5α (CTRL+) were used as bacterial standards and positive controls. Gut and skin mucus samples were taken from European sea bass (*Dicentrarchus labrax*) specimens. The following two sets of water samples were taken: one from the surrounding milieu of the same European sea bass specimens and one from tanks with gilthead sea bream (*Sparus aurata*) in both the summer and winter seasons. Faeces samples were taken from gilthead sea bream (more details on sampling can be found in the Methods section). The mock, gut, skin, and water samples were processed with two different library preparation strategies (RAP and LIG), while faeces were only processed with the LIG strategy. The use of RAP only allowed for the use of full-length V1–V9 primers (27F-1492R), whereas the LIG libraries could be amplified with V1–V9 (same primer sequences as RAP), V3–V4 (341F-805R), and V1–V3 (27F-533R) primers. Available low- (FAST), high- (HAC), and super-accuracy (SUP) basecalling algorithms were applied to all the sequenced samples. After all these combinations, a total of 207 datasets were successfully obtained, with an average of 174,450 high-quality raw reads per sample, with >93% of these reads assigned to the genus level (Supplementary Table S1).

Table 1. Tissues and experimental conditions of the ONT data used in this study. Conditions in red failed at some experimental point, hindering the sequencing of the corresponding samples.

| Tissue/Standard | Sp. | Samples | Libr. Strategy | Primers | Basecalling |
|-----------------|---------------------|---------|----------------|------------------------|-------------|
| Mock | MOCK | 1 | RAP; LIG | V3–V4; V1–V3; V1–V9 | |
| CTRL+ | <i>E. coli</i> DH5α | 1 | RAP; LIG | V1–V9 | FAST; HAC; |
| Water | - | 12 | RAP; LIG | V3–V4; V1–V3; V1–V9 | SUP |
| | - | 3 | LIG | V3–V4; V1–V3; V1–V9 | |
| Skin | ESB | 4 | RAP; LIG | V3–V4; V1–V3; V1–V9 | |
| Intestine | ESB | 4 | RAP; LIG | V1–V9 | |
| Faeces | GSB | 12 | LIG | V1–V9 | |

RAP: 16S Barcoding Kit 24 V14; LIG: Native Barcoding Kit V14 after custom-amplicon generation; VX-X: Hypervariable regions overlapped with primer set; FAST: Fast-accuracy basecalling model; HAC: High-accuracy basecalling model; SUP: Super-accuracy basecalling model; MOCK: ZymoBIOMICS™ Microbial Community Standard II; CTRL+: pure isolate of *E. coli* DH5α; ESB: European sea bass; GSB: gilthead sea bream.

2.2. Amplicon Qualities

Before sequencing, the PCR products obtained with the RAP and LIG methods were checked and their concentrations were measured (Figure 1). The values in the RAP method ranged from 35.9 ng/ μ L in the mock community and decreased in water and faeces down to 19.2 ng/ μ L and 9.36 ng/ μ L, respectively (Figure 1A). Regardless, this difference in yield was not particularly concerning, as the PCR product was sufficient for the downstream sequencing of these microbiomes. Contrarily, in the skin and intestinal mucus samples, the yields from the RAP method were notably low (less than 1 ng/ μ L), and the expected ~1.5 kb band was not observable in agarose gel for most of the samples, hindering the posterior sequencing. Within the LIG strategy, the yield ranged from 52.8 ng/ μ L in the mock community to 50.1 ng/ μ L in water and 38.9 ng/ μ L in the faeces samples when V1–V9 primers were used (Figure 1B). Interestingly, the LIG method produced amplicons with average concentrations of 10.3 ng/ μ L and 11.7 ng/ μ L in skin and intestinal mucus, with this strategy being the only one that allowed for obtaining a full-length amplicon when working with these host-associated mucus samples.

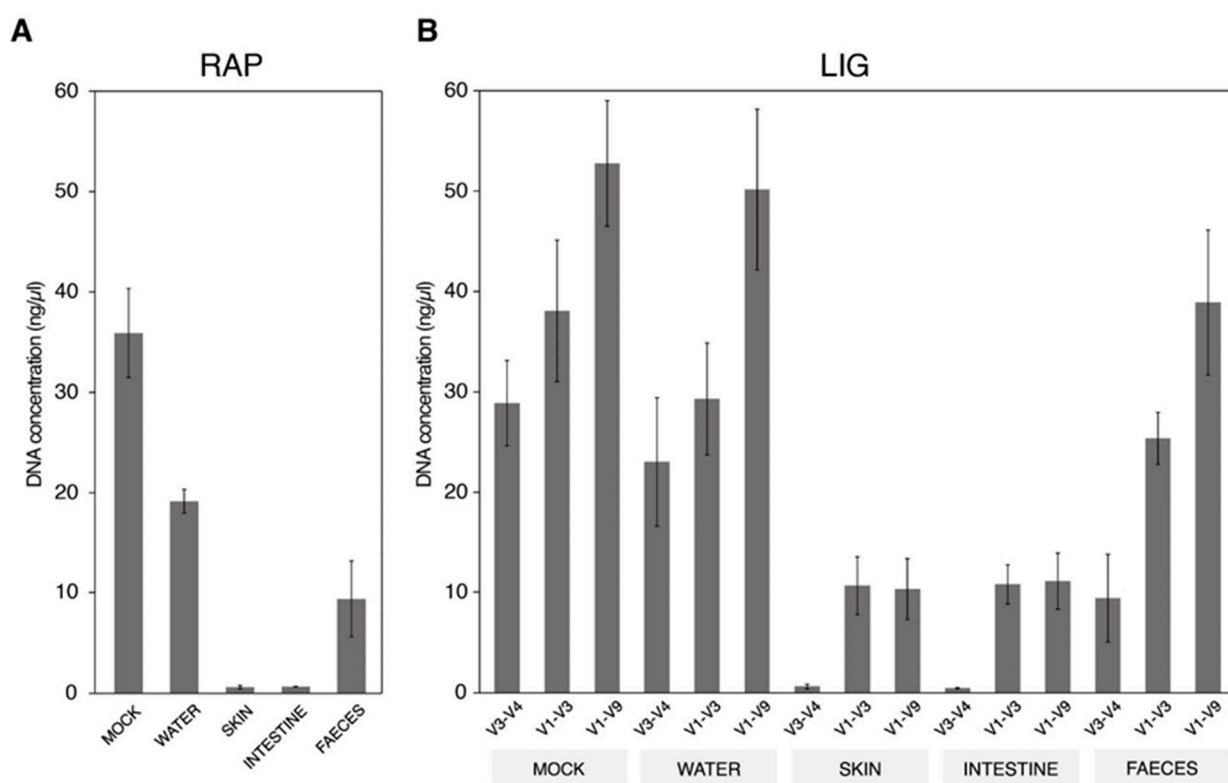


Figure 1. Bar plots representing the DNA concentration of PCR products in each one of the microbiomes included in this study for (A) RAP and (B) LIG methodology with different primer pairs.

Microbiome-specific differences were observed when comparing the PCR product yields from the three primer pairs tested under the LIG method (Figure 1B). In the water and mock communities, all primer sets showed a good performance, producing amplicons with concentrations ranging from 23.0 to 50.1 ng/ μ L. The PCR product concentration in the faeces samples followed a similar pattern, despite their yields being lower compared with the mock and water samples (9.4–38.9 ng/ μ L). In contrast, the V1–V3 primers performed similarly to V1–V9 in both the skin (10.6 compared to 10.3 ng/ μ L) and intestine (10.8 compared to 11.1 ng/ μ L) samples in terms of amplicon concentration. In fact, the V1–V3 region showed a higher estimated number of copies than the V1–V9 region (30.9 compared with 10.4 fmol in skin and 31.3 compared with 11.2 fmol in intestine mucus samples). Additionally, some replicates that failed to amplify with the V1–V9 region were successfully amplified when targeting the V1–V3 region. Despite the good performance of V1–V3 and

V1–V9, the amplification of the V3–V4 region was not achieved in the skin and intestine mucus samples, producing very low PCR yields and lacking the expected ~450 nt band in agarose gels in most cases, making posterior sequencing unfeasible.

2.3. Effects of Primers on Taxonomy Assignment

To test the performance of the RAP and LIG methods in assigning taxa, the correlation between the theoretical composition of the mock community and the values obtained after its sequencing was tested. Regardless of the method and primer choice, a high accuracy ($r = 0.93\text{--}0.99$; $p < 0.0001$) and minimal dispersion between combinations were found (Figure 2). However, the primer choice had a strong influence on the final dataset assignment in terms of the taxonomic level resolution and taxa group representation (Figure 3). In our study, the accuracy of genus-level read assignments to the theoretical taxa in the mock communities and *E. coli* pure cultures was very high (97.5–99.9%) across all primers, library preparation strategies, and basecalling models, with few unmatched or incorrectly classified reads, even for shorter regions and less accurate basecalling models (Figure 3A). However, at the species level, the assignment accuracy dropped notably, particularly for the shorter V1–V3 (47.8–68.28%) and V3–V4 regions (37.5–51.7%). Only the full-length 16S region yielded a high percentage (57.3–91.2%) of correctly classified reads at the species level. This increase was supported by the use of more accurate models (FAST < HAC < SUP). By using V1–V9, similar matching percentages were obtained at the species level using the HAC and SUP models (85.4% and 91.2%, respectively) compared to the less accurate FAST model (57.3%), with the improvement from HAC to SUP being less pronounced than the substantial gain from FAST to HAC (Figure 3A).

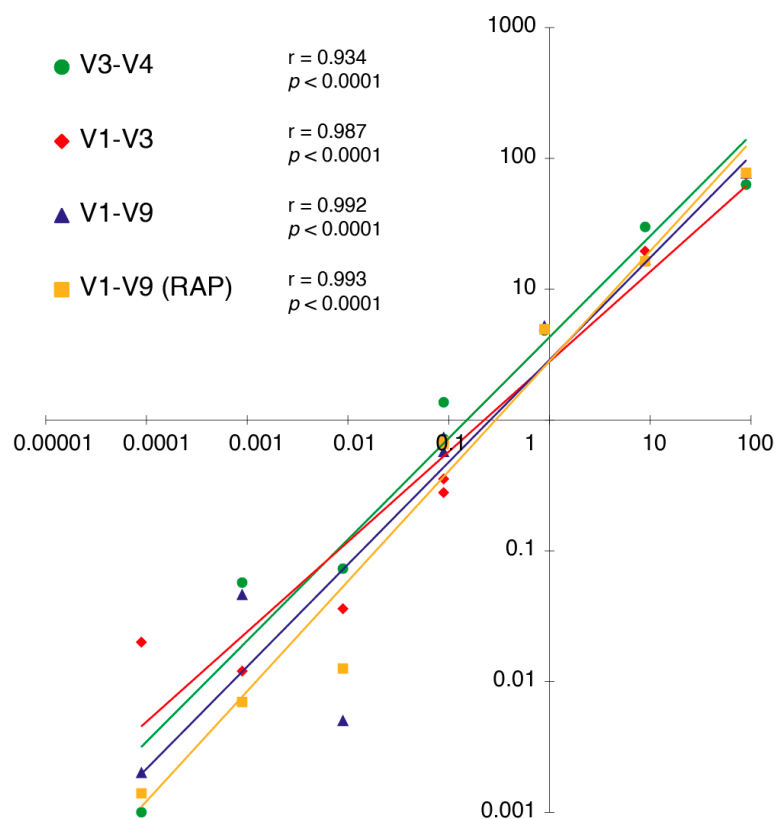


Figure 2. Correlation plots between the expected standard distribution of the mock community (X axis) and the relative abundances detected in our sequencing output (Y axis). A logarithmic scale was used to represent the data and a Pearson correlation coefficient (r) was calculated for each one of the library–primer combinations used in the study.

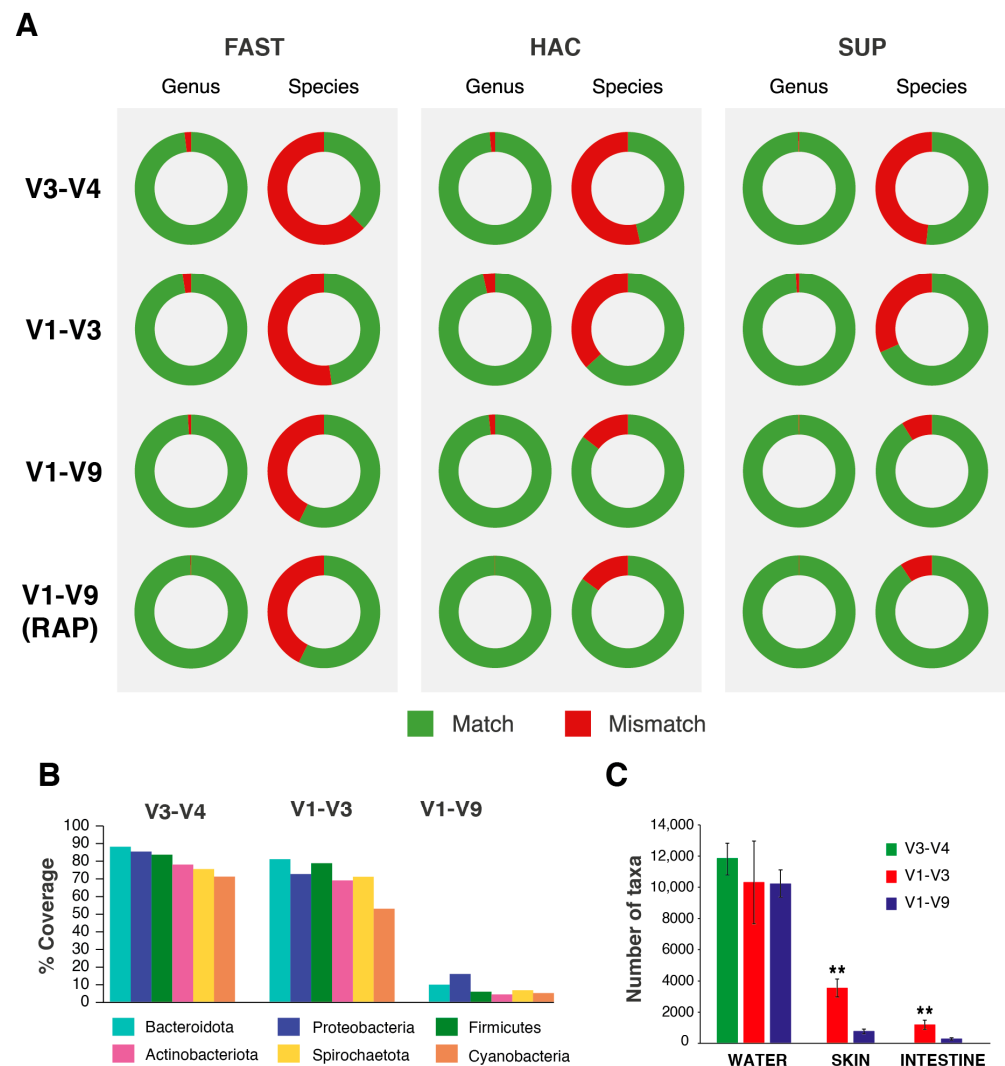


Figure 3. In silico and in vivo primer performance. (A) Pie plots representing the percentage of reads in known taxonomy samples (mock community and positive controls) assigned correctly to genus and species level in each one of the library–primer–basecalling combinations used in the study. (B) Bar plots showing the percentage of taxonomies of the 6 dominant phyla in aquaculture-related samples matched in SILVA after in silico hybridisation of each set of primers used in this study. (C) Effect of primer choice on the number of taxa found in water, skin, and intestine samples of this study (One-way ANOVA + Tukey’s post-test, ** $p < 0.001$).

To measure the in silico impact of primer choice on the database and, thus, on taxonomic assignment, we calculated the hybridisation coverages of the six most abundant phyla in the fish and water samples (Proteobacteria, Actinobacteriota, Firmicutes, Bacteroidota, Spirochaetota, and Cyanobacteria). The percentages ranged between 70 and 90% using the V3–V4 primers, and 50 and 84% using V1–V3, with the lowest value being found using the V1–V9 combination (5–20%) (Figure 3B). With this information, we investigated if the taxa number was also reduced in the real datasets, with differences depending on the microbiome. The water microbiome was not highly influenced by the primer choice, with the total number of taxonomies assigned decreasing from ~12,000 using the V3–V4 primers to ~10,000 using the V1–V3 and V1–V9 primers ($p > 0.05$). Otherwise, the skin and gut microbiomes were strongly influenced by the primer choice ($p < 0.05$), with the taxonomies retrieved decreasing from ~4000 and ~2000 with V1–V3 down to ~1000 and ~800 using the V1–V9 combination (Figure 3C).

2.4. Basecalling Effects on Sequencing Performance

For the water, skin, intestine, and faeces microbiomes, we ran the three available basecalling algorithms implemented in the Dorado v0.7 tool with a GPU-based (NVIDIA RTX 4090) system, which showed significant differences in the execution speed. For two datasets of 5.1 and 7.4 Gb, the FAST model ran in 7–9 min, with this time increasing to 37–41 min with HAC and up to 8 h with the SUP model (Figure 4A). After basecalling and read filtering by the length and quality thresholds, the mean quality scores obtained for all the samples sequenced in this study were 11.3 for the FAST model (92.5% accuracy), 15.9 for HAC (97.5% accuracy), and 23.4 for SUP (99.5% accuracy). The q-scores were consistent across all compartments analysed. Similarly, neither the primer choice nor the library preparation methods significantly affected the sequencing accuracy, as all basecalled samples under the same model maintained similar q-score values (Figure 4B–F). The use of different basecalling models also affected the number of singletons obtained in the datasets. Singletons, considered as features with a unique count across all samples, are usually removed and as considered sequencing noise. With an increase in the sequencing accuracy, the number of global singletons was reduced, with this observed tendency being independent of the microbiome sample studied, the library preparation method, or the region targeted (Figure S1). The FAST model produced percentages of singletons ranging from 0.6% to 7.9%, showing significantly ($p < 0.05$) higher values compared with the HAC model, which ranged from 0.3% to 5.3%. The decrease in the number of singletons with the SUP model (0.2% to 4.8%) compared with the HAC model was not significant ($p > 0.05$), with HAC and SUP being very close in terms of singletons despite the big gap in sequencing quality between them.

2.5. Effects of Library Preparation Strategy, Primers, and Basecalling Models in Taxonomy Correlations

Improvements in read quality due to basecalling algorithm updates were corroborated by analysing the taxonomic assignment in each microbiome. Using data from water samples, we were able to compare the assignment effects of the two library preparation strategies (RAP and LIG), the three primer combinations (V3–V4, V1–V3, and V1–V9), and the three basecalling models (FAST, HAC, and SUP). Regardless of the analysed variable, the unique shared taxonomies between the methods showed a very significant abundance rate in the datasets (94–99.9%) in comparison to those that were found in a particular strategy, primer, or model (Figure S2). In the first water dataset, 12 samples sequenced using each method were formed by the following two groups: one collected during the summer and the other during the winter. This selection of two distinct groups aimed to determine whether the library preparation method would yield consistent results in a real experimental setting (Figure 5). Regardless of the library preparation method, the samples from the two seasonal groups were differentiated, as shown by clustering based on Pearson's correlation coefficient (Figure 5A). To further investigate the differences between the LIG and RAP datasets, LEfSe analysis was performed to identify microbial markers between the two seasonal groups based on data from both library methods run with the SUP basecalling model. The list of biomarkers from the LIG and RAP datasets exhibited a high degree of overlap. Each analysis identified 15 bacterial biomarkers (LDA score threshold = 4, $\alpha < 0.05$) at the genus level (Figure 5B), with 12 of these markers being above the significance threshold in both protocols and only 3 showing discrepancies. However, upon closer examination of the LDA scores for the divergent markers, these scores were found to be just above the threshold in one dataset and slightly below in the other, demonstrating a strong consistency in the results from both datasets (Figure 5B).

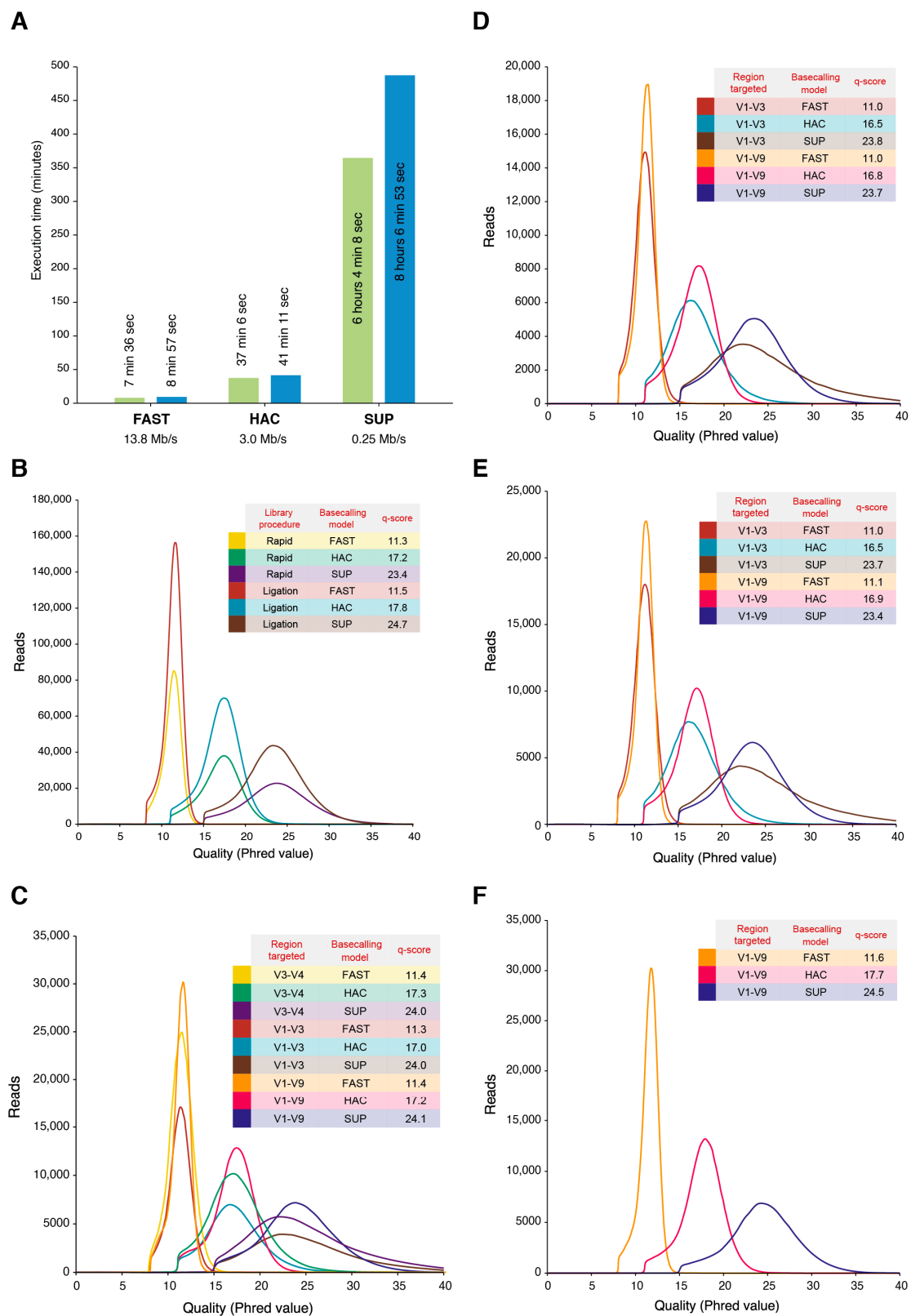


Figure 4. Basecalling time and quality performance. (A) Time of execution for each basecalling model (FAST, HAC, and SUP) used in the study for two datasets with 5.1 (green) and 7.4 (blue) Gb. Distribution of R10.4.1 reads average quality Phred score from each library preparation kit and basecalled dataset within (B,C) water, (D) skin mucus, (E) intestine mucus, and (F) faeces samples.

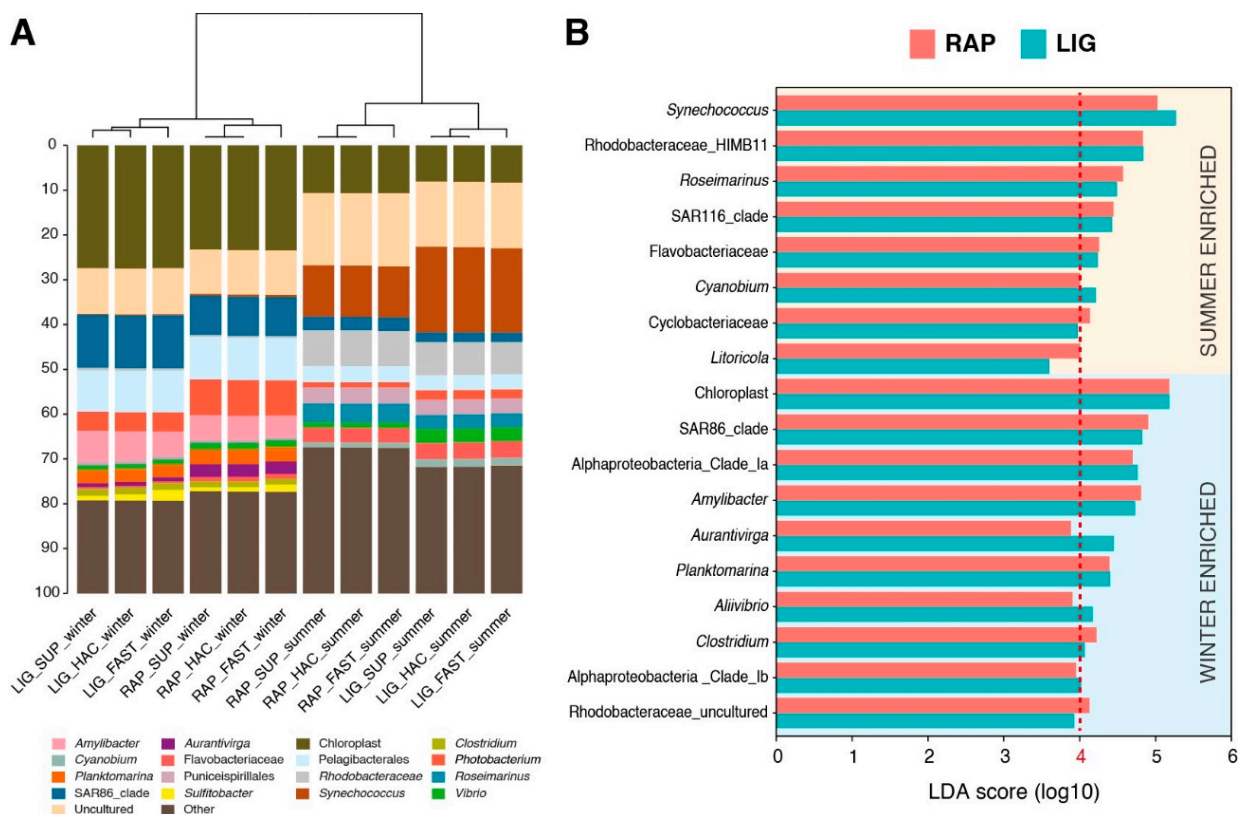


Figure 5. (A) Clustered stacked barplots showing relative abundances (sequence-based) of the most abundant genera (>1%) in summer and winter-extracted water samples. Clustering was based on Pearson correlation analysis ($p < 0.05$). Taxa with less than 1% abundance were included together in the group labelled *Other*. (B) Linear discriminant analysis effect size analysis performed at the level of genus representing the significant biomarkers for each group and their LDA score (log₁₀).

The remaining samples of water, skin, and intestinal mucus included in this study were used in the experimental trial using European sea bass. These samples were used to study the differences in the taxonomic assignment in different microbiomes due to the primer choice or the basecalling model used. All successfully amplified and sequenced samples, including their three basecalling alternatives, were put together and clustered based on their similarity represented by Pearson’s correlation coefficient. The compartment variable was, as expected, the first clustering observed, with water microbiome being separated from the skin and intestine mucus microbiomes (Figure 6A). In the water samples, where all the primers were tested, V3–V4 showed a very high correlation with the V1–V9 region ($r = 0.86$). However, a higher correlation was obtained when comparing the V1–V3 region with the full-length 16S ($r = 0.91$). In host mucosal samples, the comparison of the V1–V3 and V1–V9 regions showed a correlation coefficient of ~0.5. Despite these low values, among the most abundant taxa (>1%) represented in Figure 6A, 11 out of 16 taxa in gut mucus and 10 of 15 taxa in skin mucus varied less than 3% in the final dataset. The greatest changes were observed in the Actinobacteriota representations. Indeed, the water microbiome barely contained < 1% of this taxonomic group. Otherwise, the skin and gut mucus microbiomes accounted for 18% and 10% of the Actinobacteriota-assigned reads with the V1–V3 primers, whereas this number decreased significantly ($p < 0.05$) to 6% and 2%, respectively, with the use of the V1–V9 primers (Figure 6B). Reads assigned to *Cutibacterium* and *Microbacterium* were the most differentiated Actinobacteriota taxa between the primer combinations. In the lower level of clustering, fewer differences were observed between basecalling models compared to primer pairs. Regardless of the microbiome or the amplified region, and including faeces samples (Figure 6C), a consistent

pattern was noted. The samples basecalled with the HAC and SUP models showed a very strong correlation ($r > 0.93$ in all samples). In contrast, the samples basecalled with the FAST model were less correlated with the more accurate models ($r > 0.81$ in all samples). As a result, the SUP and HAC basecalled samples clustered together and were distinct from the FAST basecalled samples across all microbiomes and regions sequenced.

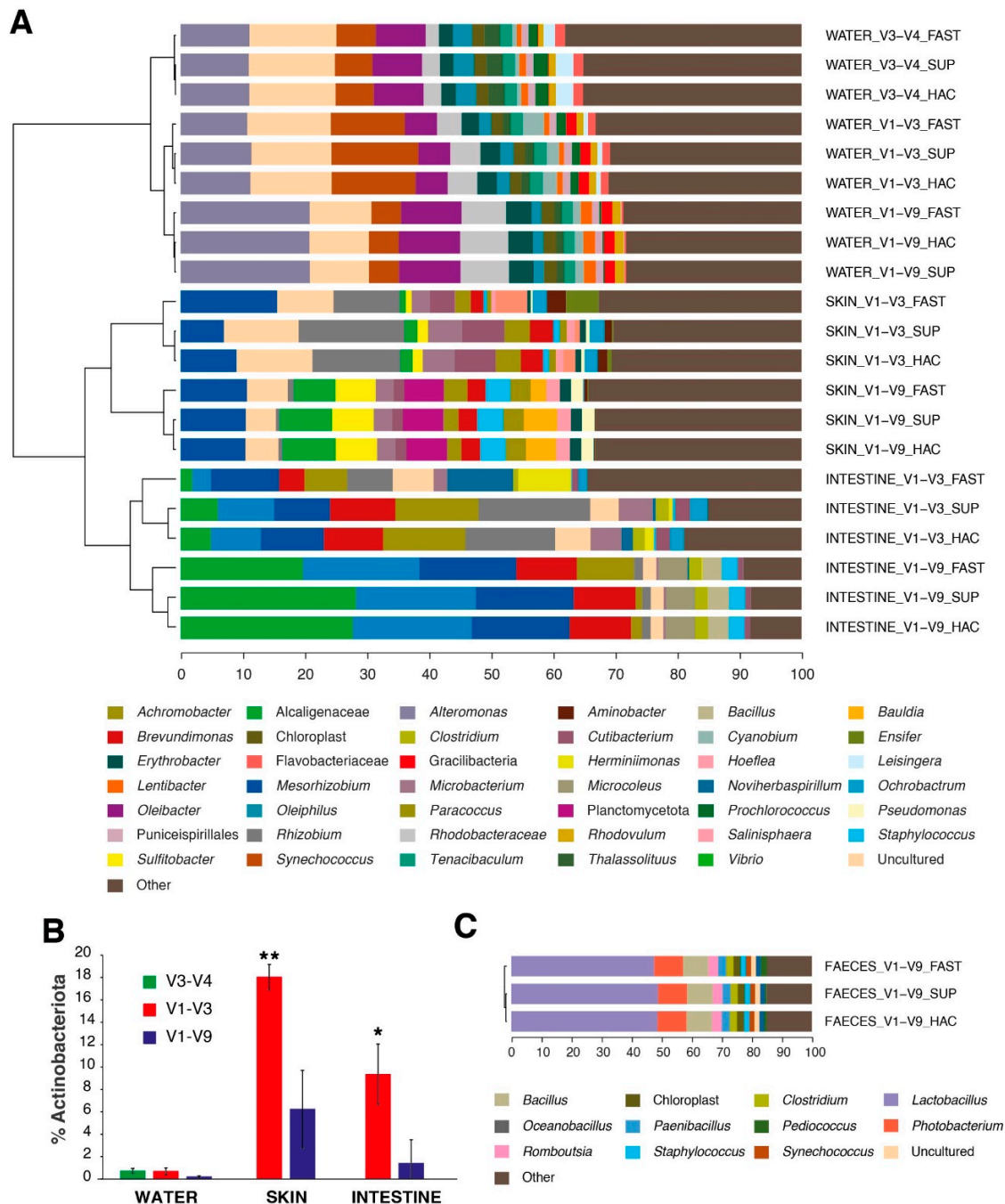


Figure 6. (A) Clustered stacked barplots showing relative abundances (sequence-based) of the most abundant genera (>1%) of the water, skin, and mucus. Clustering was based on Pearson correlation analysis ($p < 0.05$). Taxa with less than 1% abundance were grouped together in the group labelled *Other*. (B) Effect of primer choice on the number of Actinobacteriota taxa found in water, skin, and intestine samples of this study (Kruskal–Wallis + Dunn’s post-test, * $p < 0.05$; ** $p < 0.001$). (C) Clustered stacked barplots showing relative abundances (sequence-based) of the most abundant genera (>1%) of the faeces.

3. Discussion

The third-generation ONT sequencing platform allows for the rapid and in-house characterisation of microbiomes as a key method to advance the understanding of hologenic relationships between hosts and microorganisms [3,4,6,7,40]. In this study, we report the outcomes after using the ONT MinION™ device equipped with R10.4.1 flow cells for metataxonomic data through the amplification of the 16S rRNA gene. To pursue this issue in a comprehensive manner, we first scanned the catalogue of options that ONT offers in order to benchmark the different library preparation strategies, depicting differences depending on the primer and basecalling model choice in water and fish-related gut, skin, and faeces samples (Table 1). In this sense, it must be noted that the RAP method is, nowadays, the reference protocol for metabarcoding studies in aquaculture-related samples [12,14]. However, the utilisation of the LIG method has been tested successfully in environmental DNA (eDNA) metabarcoding studies using short reads as an accurate and cost-effective methodology [28,41]. Recent studies have also successfully used full-length 16S amplification procedures under the LIG methodology with human samples and bacterial cultures [42,43]. In contrast, low yields using the RAP strategy have already been noted in different fish species, including gilthead sea bream and rainbow trout [12,14]. In fact, to obtain enough amplicon amounts for microbiome sequencing, these last authors increased the input of the DNA template up to 1000 ng and modified the PCR conditions. Despite this, we obtained herein very low amplification yields (<1 ng/μL) using the RAP methodology in skin and intestinal mucus samples, greatly hindering their posterior sequencing (Figure 1A). At this point, the LIG strategy emerged as a viable alternative to the standard procedure, increasing the yields of the PCR products compared to RAP, especially in host-associated samples (Figure 1B). Furthermore, the feasibility of LIG was reinforced by taxonomy assignments that were closely correlated with those achieved with RAP when using mock (Figure 2) and real (Figure 5A) datasets, without affecting low-level taxonomy assignment (Figure 3A), data noise (Figure S2), or quality loss (Figure 4B).

Although the primer combination, enzymes, and input template were replicated exactly in RAP and LIG, the primers provided with the RAP kit also contain specific barcoding sequences and attachment chemistry, which facilitates the subsequent addition of sequencing adapters without the need for ligation reactions [44]. It can be hypothesised that, at some point, these elements can reduce the PCR yield, but the actual reasons for this are not understood yet. To try to solve this problem, some authors have performed the two following separate PCR reactions: an initial PCR with conventional primers to amplify the target region, followed by a second PCR using the barcoded primers of the 16S kit to prepare the sequencing libraries [9]. However, the use of two consecutive PCR reactions with the same primer sequence can introduce artifacts and lead to PCR bias, which needs to be considered during the data analysis step [45,46]. Furthermore, the reduced yield of the RAP method was particularly concerning in mucosal tissues, where host DNA usually represents a major fraction of the total sample compared to bacterial DNA [47–50]. Then, the results of the present study reinforce the already posed idea that RAP works better with environmental samples, high-quality samples, or samples less contaminated with host DNA [14], while LIG extends the possibilities for sequencing challenging samples and precludes their loss, as occurs when RAP amplification is not possible. Altogether, this work represents the first study to optimise the use of the ONT LIG strategy in fish- and water-related samples, as well as the first direct comparison between the LIG and RAP methods for microbiome analysis. In this sense, it is noteworthy that our findings highlight a strong correlation between the two methodological approaches (Figure 5B), making it feasible that the choice of method can depend exclusively on the resources and analytical needs of each experiment.

In addition to the above findings, the use of the LIG strategy opens the door to both long- and short-read sequencing in ONT with different primer combinations that the user can define ad hoc. In this study, the V1–V3 and V1–V9 regions could be amplified in all mock, fish, and water microbiomes, but with our experimental conditions (i.e., PCR

conditions and ONT-recommended polymerase for RAP), the use of the V3–V4 primers remained limited to mock, water, and faeces samples (Figures 1B and 2). At the same time, an increase in the amplicon length (V1–V9) led to the assignment of lower taxonomic levels (genus and species) in our final datasets (Figure 3A), confirming the actual advantage of sequencing with ONT, as already confirmed by several studies [51–53]. However, this taxonomic assignment performance was accompanied by lower *in silico* hybridisation using long primers, which was evidenced herein when we analysed the skin and gut mucus microbiomes (Figure 3B,C). The hybridisation of the chosen primers in the sample DNA would directly impact amplicon generation and, then, posterior mapping to the reference database. In this study, we used the SILVA v138.1 database, one of the most referenced and updated 16S rRNA databases, with almost 1.5 M non-redundant entries [54]. However, the high sequence similarity between SILVA and the rest of the main databases (i.e., RDP and Greengenes) suggests a more than probable parallelism [55]. Thus, good hybridisation should be promoted in the first instance to avoid taxa under-representations [56,57], which was the case for our gut and skin mucus datasets, where V1–V9 clearly yielded a decrease in the abundances of the Actinobacteriota phylum (Figure 6B). In fact, this trend has already been observed by several authors, who reported an Actinobacteriota taxa under-estimation and nucleotide mismatch in primer hybridisation [50,58]. Taxa belonging to the Actinobacteriota phyla have been isolated from fish gut and related to gut and skin functionality in fish and aquaculture species, as seen in diverse studies regarding fish genetics, ecology, and nutrition [4,59–61]. Thus, solving this limitation is key for an accurate and complete representation of aquaculture microbiomes and holobionts, and will allow for developing ONT's full potential. Some steps are being taken towards this goal, such as the design of ad hoc primers and databases [50,62] or the testing of the ~4.3 kb 16S-ITS-23S region from the *rrn* operon as a better microbial marker [63]. As long-read sequencing continues to evolve, targeting alternative regions of the 16S rRNA gene could serve as a viable option to ensure the accurate characterisation of diverse microbiomes. Indeed, the use of V1–V3 has previously depicted a resolution comparable to that of full-length 16S rRNA sequences analysed over diverse human microbiomes with the long-read PacBio platform [52,64,65]. In this line, our results suggest that the V1–V3 region can serve as a reliable alternative to the V1–V9 region, offering a higher representation of the phylum Actinobacteriota (Figure 6B) and a greater consistency in amplifying all types of microbial samples. Additionally, the sequencing of a shorter region could also lead to a very cost-effective sequencing procedure on ONT platforms [28] for monitoring microbial communities when the species level is not required, as the taxonomic resolution remains the major trade-off compared with the V1–V9 region. Certainly, arriving to a species-level resolution is not widespread in 16S analysis, as most users often collapse their results to the genus level. In this sense, the use of shorter primers can help to raise the microbial diversity in final datasets. Therefore, based on our results, we consider the amplification of the V1–V3 region, using a conventional PCR with subsequent attaching barcodes and sequencing adapters through the LIG strategy, as an alternative to long-read amplicons.

Lastly, it must be noted that the ONT procedure is based on the foundation that an electric current potential is transformed into a nucleotide sequence in a basecalling process. For this reason, the major efforts of ONT have been addressed to improve basecalling algorithms and tools during the last few years [44]. This work has paid off, and the mean q-score achieved with the most advanced SUP model represented, in our work, a substantial improvement over previous chemistry iterations, surpassing the barrier of a q-score of 20 or a 99% accuracy, regardless of the type of library preparation and the primer combination (Figure 4B–F). However, this increase in accuracy comes at a significant computational cost. Running the SUP model requires considerably more hardware resources than the FAST and HAC models, as occurred with the fish samples and R9.4.1 chemistry [12]. This increased computational requirement may not be a major issue when adequate resources are available, as was the case in this study, where a GPU-based computer was sufficient to perform the real-time basecalling of the MinION™ output (Figure 4A). However, for

systems with low-spec GPUs, or without a dedicated one, the execution time of more accurate models may increase considerably, surpassing the sequencing run time by far or even making it directly impracticable [18,66]. In terms of taxonomy assignment, the jump in read quality due to the evolution of basecalling models did not have a strong effect, and was accompanied by a higher species-level assignment (Figure 3A). Interestingly, this increase was very comparable between the HAC and SUP models, the two basecalling models that performed equally in this study. Indeed, the singleton number obtained using HAC and SUP was significantly lower compared to that with FAST (Figure S1), reducing the noise in the dataset. Clustering also reinforced this parallelism, rendering HAC and SUP taxonomy assignments always with correlation coefficients of 0.99 (Figures 5A and 6A,B). Therefore, although the SUP model is recommended to maximise the quality of produced data, the HAC and SUP models remain closely related in terms of sequencing noise and taxonomic assignment in metabarcoding studies, offering HAC an increased speed in result generation and an adequate cost-effective alternative when computational resources are not available.

In conclusion, this study provides a comprehensive benchmarking of the variables involved in 16S microbiome analysis on the ONT platform (Figure 7). The LIG and RAP library preparation methodologies performed equally, and the choice between them only depends on the specific experimental requirements. For challenging samples, targeting the V1–V3 region serves as a practical alternative to V1–V9, yielding an increased diversity at the genus level. Finally, the HAC model offers a cost-effective alternative to SUP for basecalling metabarcoding sequencing data when computational resources are limited. Altogether, these guidelines can serve as a reference for more accurate, efficient, and tailored sequencing approaches for diverse aquaculture microbiome studies.

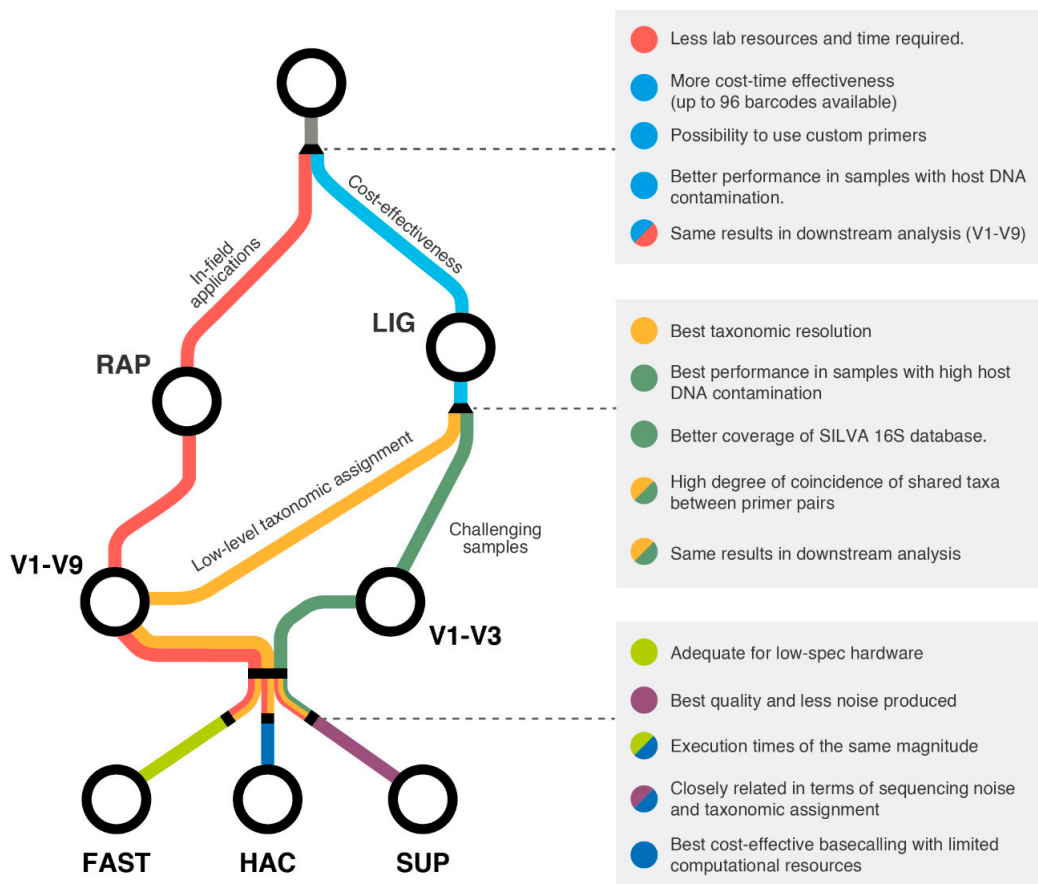


Figure 7. Proposed framework for decision making of the most cost-effective protocol when using ONT devices, including pros and cons of the different catalogue resources tested in this study.

4. Materials and Methods

4.1. Ethics Statement

The fish manipulation and tissue collection were carried out according to the Spanish (Royal Decree RD53/2013) and current EU (2010/63/EU) legislation on the handling of experimental fish. All procedures were approved by the Ethics and Animal Welfare Committee of the Institute of Aquaculture Torre de la Sal (IATS-CSIC, Castellón, Spain), CSIC Ethics Committee (permission 1135/2021), and Generalitat Valenciana (permission 2021-VSC-PEA-0192).

4.2. Mock Community, Water, and Fish Samples

For bacterial standard samples, the ZymoBIOMICS™ Microbial Community Standard II (Log Distribution) (Zymo Research, Tustin, CA, USA) was used in this study. This mock community is composed of eight bacteria strains and two yeasts (not applicable in 16S rRNA gene studies), following a logarithmic distribution of theoretical concentrations based on genomic DNA, as follows: *Listeria monocytogenes* (89.1%), *Pseudomonas aeruginosa* (8.9%), *Bacillus subtilis* (0.89%), *Escherichia coli* (0.089%), *Salmonella enterica* (0.089%), *Lactobacillus fermentum* (0.0089%), *Enterococcus faecalis* (0.00089%), and *Staphylococcus aureus* (0.000089%). Additionally, another standard consisting of bacterial DNA from a pure isolate of *E. coli* DH5 α was used as a positive control for PCR and was also sequenced as a negative control, reflecting possible contaminations during the library preparation. The DNA extraction of both standards was carried out following the same procedure, using the High Pure PCR Template Preparation Kit (Roche, Basel, Switzerland) with a previous step of enzymatic lysis with lysozyme (250 μ g/mL; 15 min; 37 °C) [67].

Samples from four different microbiomes were obtained for this work, as follows: water, skin and anterior intestine mucus, and faeces. Animal-derived samples were obtained from fish grown under natural photoperiod and temperature conditions fed with standard diets at the experimental facilities of IATS (40°5' N; 0°10' E). Water samples were collected directly from the fish tanks with sterile glass bottles (1 L), and bacterial biomass was then filtered with a manifold system using mixed cellulose ester filters with a pore size of 0.22 μ m. DNA from the water samples was extracted using the DNeasy PowerSoil Pro kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The filters were transferred to sterile Petri dishes, cut into small pieces, and submitted to mechanical lysis with the ceramic bead tubes provided in the kit, using the FastPrep 24 homogenizer (MP Biomedicals, Irvine, CA, USA) at 6 m/s for 30 s. Skin mucus samples were obtained by gently scrubbing the surface of the fish skin with a sterile microscope slide in favour of the scales, and then transferred to sterile 1.5 mL tubes and stored at -80 °C until DNA extraction. A portion of the anterior intestine (~2 cm) was opened and washed with sterile Hank's balanced salt solution to discard non-autochthonous bacteria, collecting only adherent bacteria by scraping off intestinal mucus with the blunt edge of a sterile scalpel. DNA was extracted from the skin and anterior intestine mucus using the High Pure PCR Template Preparation Kit (Roche, Basel, Switzerland), as described above. Faeces were obtained extracting the whole intestine and squeezing out all the intestinal content directly into a sterile tube with the aid of sterile forceps. DNA from the faeces was extracted using the DNeasy PowerSoil Pro kit (Qiagen, Hilden, Germany) using 250 mg of faeces as a starting material. The concentration and quality of the extracted DNA were determined in all the samples with Nanodrop 2000c (Thermo Fisher, Waltham, MA, USA).

4.3. Rapid 16S Barcoding Kit Sequencing

The complete V1–V9 region rRNA was amplified and barcoded using the 16S Barcoding kit 24 V14 (RAP; SKQ-16S114.24) following the manufacturer's protocol (version 16S_9199_V114_revB_06Dec2023), using provided 27F-1492R barcoded primers (Table 2) and including modifications of input DNA and PCR conditions optimised elsewhere for this kit [14]. PCR products were visualised in agarose gel (1% *w/v* TAE buffer) to check for the presence of the specific band of ~1500 bp, and DNA concentrations were determined

by fluorescence using PicoGreen dye (Thermo Fisher, Waltham, MA, USA). Libraries were pooled, purified using Agencourt AMPure XP beads (Backman Coulter, Brea, CA, USA), and loaded in an ONT MinION™ (Oxford Nanopore Technologies, Oxford, UK) sequencing device following the manufacturer's protocol, then sequenced using a flow-cell R10.4.1 (FLO-MIN114). Sequencing data were acquired using MinKNOW v24.02.6 software.

Table 2. Sequences of the forward (F) and reverse (R) primers used for the amplification of the hypervariable regions of the 16S rRNA gene, as well as Denaturation (D), Annealing (A), and Extension (E) temperatures induced during PCR amplification with each primer pair.

| 16S Gene Region | Primer Pair | PCR Conditions |
|----------------------|--|--|
| V1–V3 (27F-533R) | F: AGA GTT TGA TCM TGG CTC AG R: TTA CCG CGG CKG CTG GCA CG | D: 95 °C 1 min A: 30 × [95 °C 20 s, 56 °C 30 s, 65° 1 min] E: 65 °C 5 min |
| V3–V4 (341F-805R) | F: CCT ACG GGN GGC WGC AG R: GAC TAC HVG GGT ATC TAA TCC | D: 95 °C 1 min A: 30 × [95 °C 20 s, 56 °C 30 s, 65° 1 min] E: 65 °C 5 min |
| V1–V9 (27F-1492R) | F: AGA GTT TGA TCM TGG CTC AG R: CGG TTA CCT TGT TAC GAC TT | D: 95 °C 1 min A: 30 × [95 °C 20 s, 52 °C 30 s, 65° 2 min] E: 65 °C 5 min |

4.4. Ligation Sequencing of Amplicons

The Native Barcoding kit 96 V14 (LIG; SQK-NBD114.96) was used to amplify three different hypervariable regions of the 16S rRNA gene (V3–V4, V1–V3, and V1–V9), using universal primers and adapted PCR conditions (Table 2). All PCRs were performed in a total volume of 25 µL; 12.5 µL of Long Amp Hot Start Taq 2x Master Mix (New England Biolabs, Ipswich, MA, USA), 1 µL of each primer (9 µM), and 10.5 µL of template DNA at the specific concentration for each type of sample made up with Ultrapure DNase/RNase-Free Distilled Water (Invitrogen, Waltham, MA, USA). Negative controls were included to check for possible contamination. PCR products were purified using Agencourt AMPure XP beads (Backman Coulter, Brea, CA, USA) using a beads/sample ratio of 0.4 for full-length amplicons and 0.6 for short-amplicon fragments. After clean-up, the amplicons were visualised in agarose gel (1% *w/v* TAE buffer) to ensure the amplification of the desired length fragment (V3–V4 ~ 460 bp; V1–V3 ~ 500 bp; and V1–V9 ~ 1500 bp), and the DNA concentration was quantified using Picogreen (Thermo Fisher, Waltham, MA, USA). Libraries from each length were separately sequenced in a R. 10.4.1 flow cell after flushing it with the Flow Cell Wah Kit (EXP-WSH004), always using unique barcodes in each flow cell to avoid cross-contamination between subsequent runs. Sequencing data were acquired using the MinKNOW v24.02.6 software.

4.5. Basecalling and Bioinformatic Analysis

Raw sequencing POD5 files from all the runs performed in this work were basecalled using Dorado v0.7 (<https://github.com/nanoporetech/dorado>; last accessed: 25 August 2024) using a computer equipped with an Nvidia RTX 4090 24 GB GPU. All the samples were basecalled using all the three available models in their latest versions, as follows: FAST (dna_r10.4.1_e8.2_400bps_fast@v5.0.0); HAC (dna_r10.4.1_e8.2_400bps_hac@v5.0.0); and SUP (dna_r10.4.1_e8.2_400bps_sup@v5.0.0). The basecalled samples were then demultiplexed and trimmed from barcodes and adapters using Dorado v0.7. Resulting BAM files were converted into FASTQ format using samtools v1.10 [68]. These raw basecalled FASTQ files were uploaded to the Sequence Read Archive (SRA) under the Bioproject accession number PRJNA1177626 (BioSample accession numbers: SAMN44444729-935). Raw reads

were filtered using Chopper v.0.8.0 [69], using different length thresholds depending on the region targeted (V3–V4 = 300–550 bp; V1–V3 = 300–600 bp; and V1–V9 = 1200–1800 bp). In the same line, samples were filtered for quality with different minimum thresholds based on the basecalling model used in each case (FAST, $q = 8$; HAC, $q = 11$; and SUP, $q = 15$). These variable quality values were calculated to reach a uniform number of reads in all the samples from one specific tissue, regardless of the basecalling model used. Quality and length metrics were obtained for each sample using NanoPlot v1.42.0 [69]. Filter reads were then taxonomically assigned with minimap2 v2.28-r1209 [70] using SILVA v138.1 as a reference database [71].

4.6. Data Analysis

In order to evaluate the potential performance of the primer pairs, an *in silico* PCR test was performed using TestPrime v1.0, which computes the coverages for each taxonomic group in all the taxonomy entries found in the SILVA v138.1 database [72]. Differences in the number of taxa were obtained from each basecalling model by one-way ANOVA (Tukey's post-test, $p < 0.05$), while the number of singletons and phylum abundance were analysed by Kruskal–Wallis test (Dunn's post-test, $p < 0.05$) using SigmaPlot v14 (Systat Software Inc., Chicago, IL, USA). The normality of the data was verified by the Shapiro–Wilk test. Analysis of the bacterial compositional data was performed in R v4.3.0 using the packages *phyloseq* v1.44.0 [73], *corrplot* v0.94 [74], and *microbiomeMarker* v1.9.0 [75]. Pearson's correlations ($p < 0.05$) were used for the comparison of the bacterial compositional data between samples obtained using different primers, library preparation strategies, and basecalling models. The samples were clustered depending on their Pearson correlation coefficients. To determine the bacteria genera that most likely explain the differences between seasons within the water samples, a linear discriminant analysis (LDA) effect size (LEfSe) was conducted with *microbiomeMarker* v1.9.0 (LDA cutoff = 4, Wilcoxon cutoff = 0.05).

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms252312603/s1>.

Author Contributions: Conceptualisation: J.P.-S.; methodology: R.D.-B., S.T.-R., Á.B., J.P.-S. and F.N.-C.; software: R.D.-B., J.P.-S. and F.N.-C.; validation: R.D.-B., J.P.-S. and F.N.-C.; formal analysis: R.D.-B., F.M., Á.B., J.P.-S. and F.N.-C.; investigation: R.D.-B., J.P.-S. and F.N.-C. data curation: R.D.-B., J.P.-S. and F.N.-C.; writing—original draft preparation: R.D.-B., J.P.-S. and F.N.-C.; writing—review and editing: R.D.-B., F.M., S.T.-R., M.C.P., J.P.-S. and F.N.-C.; visualisation: R.D.-B.; supervision, J.P.-S.; F.N.-C.; project administration, J.P.-S.; funding acquisition: J.P.-S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the EU H2020 AQUAEXCEL3.0 Project (871108). Additionally, this study forms part of the ThinkInAzul program and was supported by MCIN with funding from European Union NextGenerationEU (PRTR-C17.I1) and by Generalitat Valenciana (THINKINAZUL/2021/024). This output reflects only the author's view and the European Union cannot be held responsible for any use that may be made of the information contained herein.

Institutional Review Board Statement: Fish manipulation and tissue collection were carried out according to the Spanish (Royal Decree RD53/2013) and the current EU (2010/63/EU) legislations on the handling of experimental fish. All procedures were approved by the Ethics and Animal Welfare Committee of the Institute of Aquaculture Torre de la Sal (IATS-CSIC, Castellón, Spain), CSIC (permit number 2021/VSC/PEA/0192).

Informed Consent Statement: Not applicable.

Data Availability Statement: All the basecalled data (FASTQ files) used in this work were loaded in the Sequence Read Archive (SRA) under the Bioproject accession number PRJNA1177626 (BioSample accession numbers: SAMN44444729-935). Raw sequencing data prior to basecalling (POD5 files) can also be obtained upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Brugman, S.; Ikeda-Ohtsubo, W.; Braber, S.; Folkerts, G.; Pieterse, C.M.J.; Bakker, P.A.H.M. A comparative review on microbiota manipulation: Lessons from fish, plants, livestock, and human research. *Front. Nutr.* **2018**, *5*, 80. [[CrossRef](#)] [[PubMed](#)]
2. Lorgen-Ritchie, M.; Uren Webster, T.; McMurtrie, J.; Bass, D.; Tyler, C.R.; Rowley, A.; Martin, S.A.M. Microbiomes in the context of developing sustainable intensified aquaculture. *Front. Microbiol.* **2023**, *14*, 1200997. [[CrossRef](#)] [[PubMed](#)]
3. Limborg, M.T.; Alberdi, A.; Kodama, M.; Roggenbuck, M.; Kristiansen, K.; Gilbert, M.T.P. Applied hologenomics: Feasibility and potential in aquaculture. *Trends Biotechnol.* **2018**, *36*, 252–264. [[CrossRef](#)] [[PubMed](#)]
4. Naya-Català, F.; Piazzon, M.C.; Torrecillas, S.; Toxqui-Rodríguez, S.; Caldach-Giner, J.À.; Fontanillas, R.; Sitjà-Bobadilla, A.; Montero, D.; Pérez-Sánchez, J. Genetics and nutrition drive the gut microbiota succession and host-transcriptome interactions through the gilthead sea bream (*Sparus aurata*) production cycle. *Biology* **2022**, *11*, 1744. [[CrossRef](#)] [[PubMed](#)]
5. Palladino, G.; Rampelli, S.; Scicchitano, D.; Musella, M.; Quero, G.M.; Prada, F.; Mancuso, A.; Seyfarth, A.M.; Turrone, S.; Candela, M.; et al. Impact of marine aquaculture on the microbiome associated with nearby holobionts: The case of patella caerulea living in proximity of sea bream aquaculture cages. *Microorganisms* **2021**, *9*, 455. [[CrossRef](#)]
6. Piazzon, M.C.; Naya-Català, F.; Perera, E.; Palenzuela, O.; Sitjà-Bobadilla, A.; Pérez-Sánchez, J. Genetic selection for growth drives differences in intestinal microbiota composition and parasite disease resistance in gilthead sea bream. *Microbiome* **2020**, *8*, 168. [[CrossRef](#)]
7. Soriano, B.; Hafez, A.I.; Naya-Català, F.; Moroni, F.; Moldovan, R.A.; Toxqui-Rodríguez, S.; Piazzon, M.C.; Arnau, V.; Llorens, C.; Pérez-Sánchez, J. SAMBA: Structure-learning of aquaculture microbiomes using a bayesian approach. *Genes* **2023**, *14*, 1650. [[CrossRef](#)]
8. Kai, S.; Matsuo, Y.; Nakagawa, S.; Kryukov, K.; Matsukawa, S.; Tanaka, H.; Iwai, T.; Imanishi, T.; Hirota, K. Rapid bacterial identification by direct pcr amplification of 16s rRNA genes using the Minion™ Nanopore sequencer. *FEBS Open Bio* **2019**, *9*, 548–557. [[CrossRef](#)]
9. Gonçalves, A.T.; Collipal-Matamal, R.; Valenzuela-Muñoz, V.; Nuñez-Acuña, G.; Valenzuela-Miranda, D.; Gallardo-Escárate, C. Nanopore sequencing of microbial communities reveals the potential role of sea lice as a reservoir for fish pathogens. *Sci. Rep.* **2020**, *10*, 2895. [[CrossRef](#)]
10. Morales-Rivera, M.F.; Valenzuela-Miranda, D.; Valenzuela-Muñoz, V.; Nuñez-Acuña, G.; Avendaño-Herrera, R.; Gallardo-Escárate, C. Nanopore sequencing evidenced the presence of fish bacterial pathogens in the sea louse (*Caligus rogercresseyi*) microbiota collected from distant salmon farms in Chile. *Aquaculture* **2022**, *552*, 738026. [[CrossRef](#)]
11. Stoeckl, T.; Katzenmeier, S.N.; Breiner, H.W.; Rubel, V. Nanopore duplex sequencing as an alternative to Illumina Miseq sequencing for eDNA-based biomonitoring of coastal aquaculture impacts. *Metabarcoding Metagenomics* **2024**, *8*, 45–68. [[CrossRef](#)]
12. Zarantonello, G.; Cuenca, A. Nanopore-enabled microbiome analysis: Investigating environmental and host-associated samples in rainbow trout aquaculture. *Curr. Protoc.* **2024**, *4*, e1069. [[CrossRef](#)] [[PubMed](#)]
13. Toxqui-Rodríguez, S.; Holhorea, P.G.; Naya-Català, F.; Caldach-Giner, J.; Sitjà-Bobadilla, A.; Piazzon, C.; Pérez-Sánchez, J. Differential reshaping of skin and intestinal microbiota by stocking density and oxygen availability in farmed gilthead sea bream (*Sparus aurata*): A behavioral and network-based integrative approach. *Microorganisms* **2024**, *12*, 1360. [[CrossRef](#)] [[PubMed](#)]
14. Toxqui-Rodríguez, S.; Naya-Català, F.; Sitjà-Bobadilla, A.; Piazzon, M.C.; Pérez-Sánchez, J. Fish microbiomics: Strengths and limitations of Minion sequencing of gilthead sea bream (*Sparus aurata*) intestinal microbiota. *Aquaculture* **2023**, *569*, 739388. [[CrossRef](#)]
15. Chen, P.; Sun, Z.; Wang, J.; Liu, X.; Bai, Y.; Chen, J.; Liu, A.; Qiao, F.; Chen, Y.; Yuan, C.; et al. Portable Nanopore-sequencing technology: Trends in development and applications. *Front. Microbiol.* **2023**, *14*, 1043967. [[CrossRef](#)]
16. Wang, Y.; Zhao, Y.; Bollas, A.; Wang, Y.; Au, K.F. Nanopore sequencing technology, bioinformatics and applications. *Nat. Biotechnol.* **2021**, *39*, 1348–1365. [[CrossRef](#)]
17. Zeng, J.; Cai, H.; Peng, H.; Wang, H.; Zhang, Y.; Akutsu, T. Causalcall: Nanopore basecalling using a temporal convolutional network. *Front. Genet.* **2020**, *10*, 1332. [[CrossRef](#)]
18. Pagès-Gallego, M.; de Ridder, J. Comprehensive benchmark and architectural analysis of deep learning models for nanopore sequencing basecalling. *Genome Biol.* **2023**, *24*, 71. [[CrossRef](#)]
19. Ferguson, S.; McLay, T.; Andrew, R.L.; Bruhl, J.J.; Schwessinger, B.; Borevitz, J.; Jones, A. Species-specific basecallers improve actual accuracy of nanopore sequencing in plants. *Plant Methods* **2022**, *18*, 137. [[CrossRef](#)]
20. Petrone, J.R.; Rios Glusberger, P.; George, C.D.; Milletich, P.L.; Ahrens, A.P.; Roesch, L.F.W.; Triplett, E.W. RESCUE: A validated nanopore pipeline to classify bacteria through long-read, 16s-its-23s rRNA sequencing. *Front. Microbiol.* **2023**, *14*, 1201064. [[CrossRef](#)]
21. Wang, Z.; Fang, Y.; Liu, Z.; Hao, N.; Zhang, H.H.; Sun, X.; Que, J.; Ding, H. Adapting Nanopore sequencing basecalling models for modification detection via incremental learning and anomaly detection. *Nat. Commun.* **2024**, *15*, 7148. [[CrossRef](#)] [[PubMed](#)]
22. Zalusky, M.P.G.; Gustafson, J.A.; Bohaczuk, S.C.; Mallory, B.; Reed, P.; Wenger, T.; Beckman, E.; Chang, I.J.; Paschal, C.R.; Buchan, J.G.; et al. 3-hour genome sequencing and targeted analysis to rapidly assess genetic risk. *Genet. Med. Open* **2024**, *2*, 101833. [[CrossRef](#)] [[PubMed](#)]
23. Kolmogorov, M.; Billingsley, K.J.; Mastoras, M.; Meredith, M.; Monlong, J.; Lorig-Roach, R.; Asri, M.; Alvarez Jerez, P.; Malik, L.; Dewan, R.; et al. Scalable Nanopore sequencing of human genomes provides a comprehensive view of haplotype-resolved variation and methylation. *Nat. Methods* **2023**, *20*, 1483–1492. [[CrossRef](#)] [[PubMed](#)]

24. Goenka, S.D.; Gorzynski, J.E.; Shafin, K.; Fisk, D.G.; Pesout, T.; Jensen, T.D.; Monlong, J.; Chang, P.C.; Baid, G.; Bernstein, J.A.; et al. Accelerated identification of disease-causing variants with ultra-rapid nanopore genome sequencing. *Nat. Biotechnol.* **2022**, *40*, 1035–1041. [[CrossRef](#)]
25. Ni, Y.; Liu, X.; Simeneh, Z.M.; Yang, M.; Li, R. Benchmarking of Nanopore R10.4 and R9.4.1 flow cells in single-cell whole-genome amplification and whole-genome shotgun sequencing. *Comput. Struct. Biotechnol. J.* **2023**, *21*, 2352–2364. [[CrossRef](#)]
26. MacKenzie, M.; Argyropoulos, C. An introduction to nanopore sequencing: Past, present, and future considerations. *Micromachines* **2023**, *14*, 459. [[CrossRef](#)]
27. Sauvage, T.; Cormier, A.; Delphine, P. A Comparison of Oxford Nanopore library strategies for bacterial genomics. *BMC Genom.* **2023**, *24*, 627. [[CrossRef](#)]
28. van der Reis, A.L.; Beckley, L.E.; Olivar, M.P.; Jeffs, A.G. Nanopore short-read sequencing: A quick, cost-effective and accurate method for DNA metabarcoding. *Environ. DNA* **2023**, *5*, 282–296. [[CrossRef](#)]
29. Armougom, F.; Henry, M.; Vialettes, B.; Raccach, D.; Raoult, D. Monitoring bacterial community of human gut microbiota reveals an increase in *Lactobacillus* in obese patients and methanogens in anorexic patients. *PLoS ONE* **2009**, *4*, e7125. [[CrossRef](#)]
30. Hamady, M.; Knight, R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res.* **2009**, *19*, 1141–1152. [[CrossRef](#)]
31. Schloss, P.D.; Gevers, D.; Westcott, S.L. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE* **2011**, *6*, e27310. [[CrossRef](#)] [[PubMed](#)]
32. Matsuo, Y.; Komiya, S.; Yasumizu, Y.; Yasuoka, Y.; Mizushima, K.; Takagi, T.; Kryukov, K.; Fukuda, A.; Morimoto, Y.; Naito, Y.; et al. Full-Length 16S rRNA gene amplicon analysis of human gut microbiota using Minion™ Nanopore sequencing confers species-level resolution. *BMC Microbiol.* **2021**, *21*, 35. [[CrossRef](#)] [[PubMed](#)]
33. Feehery, G.R.; Yigit, E.; Oyola, S.O.; Langhorst, B.W.; Schmidt, V.T.; Stewart, F.J.; Dimalanta, E.T.; Amaral-Zettler, L.A.; Davis, T.; Quail, M.A.; et al. A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. *PLoS ONE* **2013**, *8*, e76096. [[CrossRef](#)] [[PubMed](#)]
34. Marquet, M.; Zöllkau, J.; Pastuschek, J.; Viehweger, A.; Schleußner, E.; Makarewicz, O.; Pletz, M.W.; Ehrlich, R.; Brandt, C. Evaluation of microbiome enrichment and host DNA depletion in human vaginal samples using Oxford Nanopore’s adaptive sequencing. *Sci. Rep.* **2022**, *12*, 4000. [[CrossRef](#)]
35. Naya-Català, F.; Piazzon, M.C.; Calduch-Giner, J.A.; Sitjà-Bobadilla, A.; Pérez-Sánchez, J. Diet and host genetics drive the bacterial and fungal intestinal metatranscriptome of gilthead sea bream. *Front. Microbiol.* **2022**, *13*, 883738. [[CrossRef](#)]
36. Okada, T.; Fukuda, S.; Hase, K.; Nishiumi, S.; Izumi, Y.; Yoshida, M.; Hagiwara, T.; Kawashima, R.; Yamazaki, M.; Oshio, T.; et al. Microbiota-derived lactate accelerates colon epithelial cell turnover in starvation-refed mice. *Nat. Commun.* **2013**, *4*, 1654. [[CrossRef](#)]
37. Akaçin, İ.; Ersoy, Ş.; Doluca, O.; Güngörmüşler, M. Using custom-built primers and Nanopore sequencing to evaluate CO-utilizer bacterial and archaeal populations linked to bio-H₂ production. *Sci. Rep.* **2023**, *13*, 17025. [[CrossRef](#)]
38. Chang, J.J.M.; Ip, Y.C.A.; Bauman, A.G.; Huang, D. MinION-in-ARMS: Nanopore Sequencing to expedite barcoding of specimen-rich macrofaunal samples from autonomous reef monitoring structures. *Front. Mar. Sci.* **2020**, *7*, 448. [[CrossRef](#)]
39. Chang, J.J.M.; Ip, Y.C.A.; Ng, C.S.L.; Huang, D. Takeaways from mobile DNA barcoding with Bentolab and Minion. *Genes* **2020**, *11*, 1121. [[CrossRef](#)]
40. Carradec, Q.; Poulain, J.; Boissin, E.; Hume, B.C.C.; Voolstra, C.R.; Ziegler, M.; Engelen, S.; Cruaud, C.; Planes, S.; Wincker, P. A framework for in situ molecular characterization of coral holobionts using Nanopore sequencing. *Sci. Rep.* **2020**, *10*, 15893. [[CrossRef](#)]
41. Munian, K.; Ramli, F.F.; Othman, N.; Mahyudin, N.A.A.; Sariyati, N.H.; Abdullah-Fauzi, N.A.F.; Haris, H.; Ilham-Norhakim, M.L.; Abdul-Latiff, M.A.B. Environmental DNA metabarcoding of freshwater fish in Malaysian tropical rivers using short-read nanopore sequencing as a potential biomonitoring tool. *Mol. Ecol. Resour.* **2024**, *24*, e13936. [[CrossRef](#)] [[PubMed](#)]
42. Dommann, J.; Kerbl-Knapp, J.; Albertos Torres, D.; Egli, A.; Keiser, J.; Schneeberger, P.H.H. A novel barcoded nanopore sequencing workflow of high-quality, full-length bacterial 16S amplicons for taxonomic annotation of bacterial isolates and complex microbial communities. *mSystems* **2024**, *9*, e0085924. [[CrossRef](#)] [[PubMed](#)]
43. Vanhee, M.; Floré, K.; Vanthourenhout, S.; Hellemans, J.; Muyldermans, A.; Reynders, M. Implementation of full-length 16S Nanopore sequencing for bacterial identification in a clinical diagnostic setting. *Diagn. Microbiol. Infect. Dis.* **2024**, *108*, 116156. [[CrossRef](#)]
44. ONT. Chemistry Technical Document. 2023. Available online: <https://nanoporetech.com/document/chemistry-technical-document> (accessed on 15 January 2024).
45. Drengenes, C.; Eagan, T.M.L.; Haaland, I.; Wiker, H.G.; Nielsen, R. Exploring protocol bias in airway microbiome studies: One versus two PCR steps and 16S rRNA gene region V3 V4 versus V4. *BMC Genom.* **2021**, *22*, 3. [[CrossRef](#)]
46. Silverman, J.D.; Bloom, R.J.; Jiang, S.; Durand, H.K.; Dallow, E.; Mukherjee, S.; David, L.A. Measuring and mitigating PCR bias in microbiota datasets. *PLoS Comput. Biol.* **2021**, *17*, 3. [[CrossRef](#)]
47. Deiner, K.; Bik, H.M.; Mächler, E.; Seymour, M.; Lacoursière-Roussel, A.; Altermatt, F.; Creer, S.; Bista, I.; Lodge, D.M.; de Vere, N.; et al. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Mol. Ecol.* **2017**, *26*, 5872–5895. [[CrossRef](#)]

48. Knight, R.; Vrbanac, A.; Taylor, B.C.; Aksenov, A.; Callewaert, C.; Debelius, J.; Gonzalez, A.; Kosciolek, T.; McCall, L.I.; McDonald, D.; et al. Best practices for analysing microbiomes. *Nat. Rev. Microbiol.* **2018**, *16*, 410–422. [[CrossRef](#)]
49. Santiago-Rodriguez, T.M.; Le François, B.; Macklaim, J.M.; Doukhanine, E.; Hollister, E.B. The skin microbiome: Current techniques, challenges, and future directions. *Microorganisms* **2023**, *11*, 1222. [[CrossRef](#)]
50. Walker, S.P.; Barrett, M.; Hogan, G.; Flores Bueso, Y.; Claesson, M.J.; Tangney, M. Non-specific amplification of human dna is a major challenge for 16S rRNA gene sequence analysis. *Sci. Rep.* **2020**, *10*, 16356. [[CrossRef](#)]
51. Szoboszlai, M.; Schramm, L.; Pinzauti, D.; Scerri, J.; Sandionigi, A.; Biazzo, M. Nanopore is preferable over Illumina for 16S amplicon sequencing of the gut microbiota when species-level taxonomic classification, accurate estimation of richness, or focus on rare taxa is required. *Microorganisms* **2023**, *11*, 804. [[CrossRef](#)]
52. Zhang, H.; Wang, X.; Chen, A.; Li, S.; Tao, R.; Chen, K.; Huang, P.; Li, L.; Huang, J.; Li, C.; et al. Comparison of the full-length sequence and sub-regions of 16S rRNA gene for skin microbiome profiling. *MSystems* **2024**, *9*, e0039924. [[CrossRef](#)] [[PubMed](#)]
53. Zheng, W.; Tsompana, M.; Ruscitto, A.; Sharma, A.; Genco, R.; Sun, Y.; Buck, M.J. An accurate and efficient experimental approach for characterization of the complex oral microbiota. *Microbiome* **2015**, *3*, 48. [[CrossRef](#)] [[PubMed](#)]
54. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, 590–596. [[CrossRef](#)] [[PubMed](#)]
55. Robeson, M.S.; O'Rourke, D.R.; Kaehler, B.D.; Ziemski, M.; Dillon, M.R.; Foster, J.T.; Bokulich, N.A. RESCRIPt: Reproducible sequence taxonomy reference database management. *PLoS Comp. Biol.* **2021**, *17*, e1009581. [[CrossRef](#)]
56. Sun, Y.; Du, X.; Li, Y.; Han, X.; Fang, S.; Geisen, S.; Li, Q. Database and primer selections affect nematode community composition under different vegetations of changbai mountain. *Soil Ecol. Lett.* **2023**, *5*, 142–150. [[CrossRef](#)]
57. Baker, G.C.; Smith, J.J.; Cowan, D.A. Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Methods* **2003**, *55*, 541–555. [[CrossRef](#)]
58. Heikema, A.P.; Horst-Kreft, D.; Boers, S.A.; Jansen, R.; Hiltemann, S.D.; de Koning, W.; Kraaij, R.; de Ridder, M.A.J.; van Houten, C.B.; Bont, L.J.; et al. Comparison of Illumina versus Nanopore 16S rRNA gene sequencing of the human nasal microbiota. *Genes* **2020**, *11*, 1105. [[CrossRef](#)]
59. Thejaswini, S.; Joji, S.; Vijayan, A.; Martin Paul, A. Isolation of gut Actinobacteria from fishes. In *Methods in Actinobacteriology*; Springer Protocols Handbooks; Dharumadurai, D., Ed.; Humana: New York, NY, USA, 2022. [[CrossRef](#)]
60. Kormas, K.A.; Meziti, A.; Mente, E.; Frentzos, A. Dietary Differences Are Reflected on the Gut Prokaryotic Community Structure of Wild and Commercially Reared Sea Bream (*Sparus aurata*). *Microbiologyopen* **2014**, *3*, 718–728. [[CrossRef](#)]
61. Meziti, A.; Nikouli, E.; Papaharisis, L.; Kormas, K.A.; Mente, E. The response of gut and fecal bacterial communities of the European sea bass (*Dicentrarchus labrax*) fed a low fish-plant meal and yeast protein supplementation diet. *Sustain. Microbiol.* **2024**, *1*, qvae005. [[CrossRef](#)]
62. Molano, L.G.; Vega-abellana, S.; Manichanh, C.; Molano, L.G.; Vega-abellana, S.; Manichanh, C. Database for 16S rRNA amplicon analysis. *MSystems* **2024**, *9*, e0095023. [[CrossRef](#)]
63. Cuscó, A.; Catozzi, C.; Viñes, J.; Sanchez, A.; Francino, O. Microbiota profiling with long amplicons using Nanopore sequencing: Full-length 16S rRNA gene and the 16S-ITS-23S of the *rrn* operon. *F1000Research* **2019**, *7*, 1755. [[CrossRef](#)] [[PubMed](#)]
64. Buetas, E.; Jordán-López, M.; López-Roldán, A.; D'Auria, G.; Martínez-Priego, L.; De Marco, G.; Carda-Diéguez, M.; Mira, A. Full-Length 16S rRNA gene sequencing by pacbio improves taxonomic resolution in human microbiome samples. *BMC Genomics* **2024**, *25*, 310. [[CrossRef](#)] [[PubMed](#)]
65. Pootakham, W.; Mhuantong, W.; Yoocha, T.; Putchim, L.; Sonthirod, C.; Naktang, C.; Thongtham, N.; Tangphatsornruang, S. High resolution profiling of coral-associated bacterial communities using full-length 16S rRNA sequence data from PacBio SMRT sequencing system. *Sci. Rep.* **2017**, *7*, 2774. [[CrossRef](#)] [[PubMed](#)]
66. Wick, R.R.; Judd, L.M.; Holt, K.E. Performance of neural network basecalling tools for Oxford Nanopore Sequencing. *Genome Biol.* **2019**, *20*, 129. [[CrossRef](#)] [[PubMed](#)]
67. Piazzon, M.C.; Naya-Català, F.; Simó-Mirabet, P.; Picard-Sánchez, A.; Roig, F.J.; Caldach-Giner, J.A.; Sitjà-Bobadilla, A.; Pérez-Sánchez, J. Sex, age, and bacteria: How the intestinal microbiota is modulated in a protandrous hermaphrodite fish. *Front. Microbiol.* **2019**, *10*, 2512. [[CrossRef](#)]
68. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M. Twelve years of SAMtools and BCftools. *Gigascience* **2021**, *10*, giab008. [[CrossRef](#)]
69. De Coster, W.; Rademakers, R. NanoPack2: Population-scale evaluation of long-read sequencing data. *Bioinformatics* **2023**, *39*, btad311. [[CrossRef](#)]
70. Li, H. New strategies to improve Minimap2 alignment accuracy. *Bioinformatics* **2021**, *37*, 4572–4574. [[CrossRef](#)]
71. Yilmaz, P.; Parfrey, L.W.; Yarza, P.; Gerken, J.; Pruesse, E.; Quast, C.; Schweer, T.; Peplies, J.; Ludwig, W.; Glöckner, F.O. The SILVA and “all-species living tree project (LTP)” taxonomic frameworks. *Nucleic Acids Res.* **2014**, *42*, D643–D648. [[CrossRef](#)]
72. Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glöckner, F.O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, e1. [[CrossRef](#)]
73. McMurdie, P.J.; Holmes, S. Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **2013**, *8*, e61217. [[CrossRef](#)]

74. Wei, T.; Simko, V. R Package “Corrplot”: Visualization of a Correlation Matrix, version 0.92. 2021. Available online: <https://github.com/taiyun/corrplot> (accessed on 15 February 2024).
75. Cao, Y.; Dong, Q.; Wang, D.; Zhang, P.; Liu, Y.; Niu, C. MicrobiomeMarker: An R/Bioconductor package for microbiome marker identification and visualization. *Bioinformatics* **2022**, *38*, 4027–4029. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.