

QT-AMP: Quanti-Tray-based amplicon sequencing for simultaneous quantification and identification of enterococci for microbial source tracking

Hidetoshi Urakawa^{a, b}, Taylor L. Hancock^{a, b}, Michael A. Kratz^a, Rick A. Armstrong^c

a - Department of Marine and Ecological Sciences, Florida Gulf Coast University, Fort Myers, Florida, USA

b - School of Geosciences, University of South Florida, Tampa, FL, 33620, USA

c – Lee County Environmental Laboratory, Fort Myers, FL, 33907, USA

Keywords

Enterococci Enterolert Microbial source tracking High-throughput amplicon sequencing

Highlights

- A selected primer set (27f-519r) can differentiate over 50 *Enterococcus* species and is suitable for Illumina amplicon sequencing.
- The median of the relative contribution of enterococci reads among total sequencing reads was 82.7% and ranged between 0% and 100%.
- Enterolert signals are most likely the mixture of enterococci and other facultative aerobes and anaerobes.
- We identified eight cosmopolitan enterococci species and three rare species.
- The median of the relative contribution of non-enterococci reads among total sequencing reads was 17.3%, respectively, and ranged between 0% and 100%.

ABSTRACT

Enterococcus is ubiquitous in human feces and has been adopted as a useful indicator of human fecal pollution in water. Although regular enterococci monitoring only examines their numbers, identification of human-specific *Enterococcus* species or genotypes could help in the discrimination of human fecal contamination from other environmental sources. We documented a new approach to characterize enterococci using an amplicon sequencing platform from Quanti Trays after following the counting of most probable numbers (MPN) of enterococci. We named this method as QT-AMP (Quanti-Tray-based amplicon sequencing). We tested surface water samples collected from three rivers in southwest Florida. We detected 11 *Enterococcus* species from 45 samples in 1.1 million sequence reads. The method detected three rare species and eight cosmopolitan species (*Enterococcus faecalis*, *E. faecium*, *E. casseliflavus*, *E. hirae*, *E. mundtii*, *E. gallinarum*, *E. avium*, and *E. durans*) which have been commonly documented in various enterococci isolation studies. It is likely that the approximate detection level of QT-AMP is four orders of magnitude higher than regular 16S rRNA gene amplicon sequencing. QT-AMP revealed that a majority of Enterolert positive signals are actually the mixture of both enterococci and other facultative aerobes and anaerobes. QT-AMP may have the potential to monitor not only enterococci but also other pathogenic bacteria commonly found in natural environments. This QT-AMP could be a powerful tool to streamline the quantification and identification of enterococci and allows us to do more accurate and efficient microbial source tracking in various water management projects and human health risk assessment.

1. INTRODUCTION

Enterococcus is a gram-positive facultative anaerobic bacteria. Because of their ubiquity in human feces and persistence in the environment, *Enterococcus* has been adopted as a useful indicator of human fecal pollution in water. Traditionally, detection of enterococci has been made by membrane filtration method (e.g., U.S. EPA method 1600 membrane filtration technique) with a combination of selective medium and incubation temperature. More recently, enterococci have also been quantified using an enzymatic reaction and the most probable number (MPN) format. Now, this semi-automated MPN method for enumeration of enterococci is standardized and commercially available as Enterolert (IDEXX, Westbrook, ME, USA), a defined substrate methodology that measures a fluorescent endpoint based on enterococci metabolizing 4-methylumbelliferone- β -D-glucoside. The method has been widely adopted worldwide (Abbott et al., 1998; Eckner, 1998; Ramoutar et al., 2020). The method greatly improves the standardization of the counting procedure, which may cause less contamination and counting errors, and greatly reduced labor and time (Budnick et al., 1996).

Although the method is standardized and accepted broadly, the use of enterococci as indicators of human fecal pollution or contamination could be problematic (Budnick et al., 1996). Identification of origins of enterococci in water bodies is confounded by multiple sources as well as by the many factors that influence the ultimate fate of microorganisms once they are released into the environment. Enterococci are not only found in human feces but also in both wild and domestic animal feces (Harwood et al., 2000; Layton et al., 2010), birds (Devriese et al., 1992), river beds and banks, soils (Byappanahalli and Fujioka, 2004), beach sand (Ran et al., 2013; Yamahara et al., 2007), and on aquatic and terrestrial plants (Byappanahalli et al., 2003; Imamura et al., 2011).

Due to the uncertainty about how long enterococci are persistent in nature, many researchers have focused on studying their dynamics and persistence in both freshwater and saltwater environments, which can be changed by multiple environmental factors controlled by

seasons, weather, hydrology, water quality, and microbial communities (Ran et al., 2013; Yamahara et al., 2007).

Identification of human-specific enterococci species or genotypes could aid in the discrimination of human fecal contamination from other environmental sources of the organisms. However, regular enterococci monitoring only addresses their numbers (Lipp et al., 2001). In the past, isolation, purification, and characterization of *Enterococcus* strains have been used to identify this bacterium at the species level (Ferguson et al., 2013; Ran et al., 2013). As an alternative approach, a cloning method of the 16S rRNA gene has been used to characterize enterococci without cultivation efforts (Sercu et al., 2011). However, these two approaches are only possible in the microbiology laboratory and not applicable in a routine manner.

To solve this issue and quantify and identify enterococcus in a streamlined fashion, we set a goal to develop a new method named QT-AMP (Quanti-Tray-based amplicon sequencing), in which first Enterolert Quanti-Tray is used for MPN counting and then pools of positive wells are used as a source for high-throughput 16S rRNA amplicon sequencing. To develop this method, we first compared three primer sets to examine which primer pair is the best to efficiently discriminate many enterococci species. Second, we tested surface water samples collected from three rivers in southwest Florida to examine how efficiently enterococci sequences were recovered and could be used for identification.

2. MATERIALS AND METHODS

2.1. Sampling sites

Water samples were collected by 12 sampling events from three rivers at routine water monitoring sampling sites of Lee County, Florida (**Fig .1**). Five sampling sites were set in the Caloosahatchee River, three in the Estero River, and four in the Imperial River. The wet and dry

season samplings were conducted in August and September 2020, and January and February 2021, respectively. The Caloosahatchee River flows west from Lake Okeechobee reaching San Carlos Bay, approximately 108 km long and 62 km² of surface area. The average depth of the river is 4.3 m. To provide flood control for settlement areas, intensive dredging operations were conducted in the late 1800s. After modification, now the river is a part of the Okeechobee Waterway, a manmade waterway system of southern Florida. The river is now formed by two components, compartmentalized by Franklin Lock and Dam: a narrow freshwater upstream segment and a wide tidal zone segment (Caloosahatchee River Estuary). The Estero River is a 13.7 km stream located within the Estero Bay watershed. The Imperial River is a 12.6 km stream surrounded by residential areas and has its mouth at the north end of Fish trap Bay, near the southern end of Estero Bay. A large portion of the river is covered by riverbank forests (**Fig. 1**) and intensively used for tourism (e.g., kayaking, fishing). Both rivers are impaired by fecal coliform according to the Florida Department of Environmental Protection's (FDEP) implementation of the Impaired Waters Rule (IWR) (Gulfbase <https://www.gulfbase.org/> & CHNEP Water Atlas <https://chnep.wateratlas.usf.edu/>). Therefore, a high abundance of fecal indicator bacteria (FIB) is problematic in these rivers.

2.2. Water sampling

Surface water samples were collected using a clean bucket from bridges or accessible shore points. Water samples were filled in acid-washed high-density polyethylene (HDPE) bottles, kept below 8°C (iced cooler), and transported to the Lee County Environmental Lab.

2.3. Water quality analyses

Water quality (temperature, pH, dissolved oxygen [DO], and conductivity) analysis was conducted by a YSI multiparameter probe. Nutrients were analyzed using a SEAL AA3 autoanalyzer (Seal Analytical, Mequon, WI, USA). Total Kjeldahl nitrogen (TKN) and total phosphate (TP) were measured after heat digestion. Total nitrogen (TN) was calculated as a total of TKN and ammonia. A portion of water samples was filtered using GF/F and chlorophyll *a*

was measured using a spectrophotometer after 90% acetone extraction (USEPA 445.0). Chlorophyll *a* was corrected for pheophytin. All laboratory analyses were conducted at a certificated laboratory (Lee County, Florida).

2.4. MPN counting using the Enterolert Quanti-Tray method

Water samples for enterococcus and *Escherichia coli* were aseptically collected and brought back to the laboratory using Thermo Scientific Security-Snap Coliform Polypropylene Water Sample Bottles with 10mg sodium thiosulfate (Thermo Fisher Scientific) and kept in ice (below 8°C). All samples were analyzed within six hours of collection following US EPA standards and the Enterolert manufacturer's instructions. Enterolert-E ampules (IDEXX Laboratories, Westbrook, ME, USA) were aseptically added into 100 ml of water samples. The sample-reagent mixture was poured into a Quanti-Tray/2000 and sealed with the Quanti-Tray Sealer PLUS (IDEXX). The trays were incubated at $41 \pm 0.5^\circ\text{C}$ for 24 h. Under UV light in a dark room, yellow fluorescing positive wells were counted, and the enterococci density was calculated from the IDEXX MPN table.

2.5. DNA extraction from Quanti-Trays

For DNA extraction, the back of the Quanti-Tray was disinfected with 70% alcohol, and media from up to five fluorescing (positive) wells were withdrawn using sterile syringes. The withdrawn volume was recorded in each operation and approximately 10 ml of culture medium from positive wells were filtered through disposable 0.22 μm polysulfone filters (Millipore Sigma). DNA was extracted using a modified phenol-chloroform method (Urakawa et al., 2010). The purity and concentration of DNA were tested using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Foster City, CA, USA) and a Qubit 2.0 Fluorometer with the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). DNA was stored at -20°C until further analysis.

2.6. High-throughput DNA sequencing

A PCR primer set (27f [AGRGTTTGATCMTGGCTCAG] and 519r [GTNTTACNGCGGCKGCTG]) was used and the 27F primer contained a DNA barcode for sample identification. Samples were amplified using the HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA, USA) with cycling conditions of 94°C for 3min, followed by 28 cycles of 94°C for 30s, 53°C for 40s, and 72°C for 1min. A final elongation step at 72°C was set for 5min. The success of PCR amplification was verified using agarose gel electrophoresis (2%). Samples were purified using Ampure XP beads and pooled together in equimolar concentrations. Sequencing was performed using the Illumina MiSeq platform with v3 300 base single read protocol at MR DNA (Shallowater, TX, USA). Sequence data were processed through the removal of barcodes and low-quality reads (low average quality (<Q25) or short length (<200 bases), denoising, and chimera removal (Edgar et al., 2011). Forward and reverse reads were merged and operational taxonomic units (OTU) were binned by clustering sequences at 3% divergence (i.e., 97% similarity). OTUs were taxonomically classified using BLASTn against a curated database derived from GreenGenes/RDP/NCBI. The raw sequence reads were deposited in the Short Reads Archive (SRA) and are publicly available with the accession number from SRR0000000 through SRR0000000.

2.7. Assessment of primer design

The best primer set was examined by in silico analysis. We collected 16S rRNA gene sequence data sets of enterococci from LPSN (List of Prokaryotic names with Standing in Nomenclature), GenBank Taxonomy, and a review manuscript (Byappanahalli et al., 2012), which were used to make the most updated species list of the genus *Enterococcus* (as of January 1, 2021). One DNA sequence was selected for one species and a total of 60 species were used for analysis (Table S1). The first FASTA data set was made by using 16S rRNA gene sequences provided by LPSN. Some DNA sequences originally selected were replaced by another of the same species due to relatively large numbers of undetermined nucleotide positions (*N*). Therefore, sequences containing fewer numbers of *N* and longer base pairs were selectively used for analysis. The FASTA data set were aligned by MUSCLE and manually inspected. Three sets of alignment data

were created to examine the validity of the three primer sets. We assessed three primer sets: (1) 27f and 519r, (2) 515f and 806r, and (3) 515f and 926r. Sequence similarity was calculated for each species pair and a p-distance matrix was prepared for each primer set. MEGA X was used to analyze these sequence data (Kumar et al., 2018).

2.8. Statistical analysis

A Pearson correlation test was performed to examine the reproducibility of samples. Descriptive statistical analyses and histogram design were made using XLSTAT (Addinsoft Inc, NY, USA). A heat map was generated using Heatmapper (Babicki et al., 2016).

3. RESULTS

3.1. Which primer pair has the strongest species discrimination power?

To determine an optimum primer pair that has the strongest species discrimination power, sequence similarity was calculated among 60 *Enterococcus* species and a distance matrix was prepared for each primer set (**Table S2**); a total of 1,830 paired-distance was examined (**Table 1**). The mean paired-distance of 27f-519r pair ($8\% \pm 3.8\%$ [SD]) was greater than the pairs of 515f-806r ($1.3\% \pm 1.5\%$) and 515f-926r ($1.3\% \pm 1.7\%$). In the case of the 27f-519r pair, a peak of relative frequency was the fourth column of the histogram (bound 0.0678-0.0904, relative frequency 0.285), while the first two columns occupied the majority of relative frequencies in the cases of 515f-806r and 515f-926r pairs (**Fig. 2**). These data suggested that the 515f-806r pair has a stronger species discrimination power than the two other primer pairs and was echoed by the DNA sequence alignment data, which contained the largest numbers of variable sites ($n = 179$) that aid the discrimination of *Enterococcus* species and were concentrated in the front part of the 16S rRNA gene (V1-V3 regions) (**Table 1**).

3.2. Assessment of species discrimination limitation

Our DNA database survey showed that the 16S rRNA gene sequences of 60 species of *Enterococcus* species are available (as of December 2020). All primer pairs showed that some species have 100% identities with other species, indicating that some *Enterococcus* species are not distinguishable by this high-throughput sequencing approach. We examined 100% paired match cases for each species and identified groups of *Enterococcus* clustered together. Four to five groups were found in the case of each primer pair and they consisted of 2 to 13 species (**Table S3**). Among three primer pairs, 27f-519r showed the best result and four groups found were (1) *E. caccae*, *E. haemoperoxidus*, *E. moraviensis*, *E. silesiacus*, and *E. ureasiticus*, (2) *E. massiliensis* and *E. songbeiensis*, (3) *E. pseudoavium* and *E. viikkiensis*, and (4) *E. porcinus* and *E. villorum*. Therefore, these listed species cannot be identified as a single species by using the 27f-519r primer pair.

3.3. Environmental factors and *Enterococcus* distribution

In southwest Florida, MPN counts of enterococci are routinely high and it exceeds USEPA's enterococci criteria. The data normally show no typical point source continuously serves as a reservoir of enterococci to the river water, suggesting that the sources of enterococcus are not limited to wastewater but multiple unidentified sources (**Fig. 1**). In total, 48 water samples were collected from three rivers by 12 sampling events (six in wet and six in dry seasons) (**Table 2**). The conductivity data suggested that all three rivers were partially influenced by the mixing of seawater. The positive detection rates of enterococci in the tested water samples were 100% (**Table 3**). The geometric mean MPN counts of enterococci were highest in the Imperial River and followed by the Estero River (**Table 3**). The lowest geometric mean MPN counts were detected in the Caloosahatchee River. Over 75% of tested water samples exceeded USEPA's statistical threshold value (STV) of 130 CFU/100 ml in the Imperial and Estero Rivers and only 5.2% in the Caloosahatchee River. Over 100% of tested water samples from each site exceeded USEPA's geometric mean criterion of 35 CFU/100 ml in the whole sampling period in the Imperial and Estero Rivers but not from the Caloosahatchee River. Only one high number was found in the Caloosahatchee River (CES06SUR, September 2020, 782 MPN/100 ml). On that

day, the water sample was characterized by low turbidity, pH, dissolved oxygen, and high nutrients, suggesting the impact of a heavy rain event (**Table 2**). No statistically significant difference was observed for enterococci between wet and dry season samples ($p > 0.05$). The abundance of *Enterococci* and *E. coli* highly correlated in all three rivers as total ($R^2 = 0.83$, $p < 0.001$, $n = 36$) and within each river (**Table 2**).

3.4. Occurrence and diversity of *Enterococcus* species

A total of 45 surface water samples containing seven duplicates were used for QT-AMP. The total and mean sequencing counts were 1,121,843 and $24,929.8 \pm 1234.6$ (standard error [SE], $n = 45$), respectively. The mean and median of the relative contribution of enterococci reads among total sequencing reads was $72.3 \pm 4.6\%$ (SE, $n = 44$) and 82.7%, respectively, and ranged between 0% and 100% (**Fig. 3**). A heat map was used to show the spatiotemporal distribution of enterococci species in three river water samples (**Fig. 4**). The data show a clumped distribution pattern and strong Z scores concentrated on samples from the Imperial River, which was followed by the Estero River samples. Limited enterococci species were found in the water samples from the Caloosahatchee River. This order of species distribution harmonized with the abundance of enterococci, which was measured by MPN counting (**Table 2**). In total 11 and 10 *Enterococcus* species were found from the wet and dry season samples, respectively. Unexpectedly, 10 out of 11 species overlapped each other, and *E. malodoratus* was solely found from wet season samples as a minor species (**Fig. 5**). The two highest DNA sequence reads were from *E. faecalis* (42.0% in wet season and 36.2% in dry season) and *E. casseliflavus* (34.9% in wet season and 30.1% in dry season) (**Fig. 5**). The prevalence of these two species was also shown in the heat map and these two species were major enterococci in the Caloosahatchee River (**Figs. 1 & 4**). The occurrence of each *Enterococcus* species in the water samples ranged between 100% (19/19 samples in the wet season) (*E. casseliflavus*) to 0% (0/19 samples in the dry season) (*E. malodoratus*). Many species were similarly found in both wet and dry seasons, but a few species were more frequently found in wet (*E. casseliflavus*) or dry (*E. mundtii* and *E.*

rivorum) seasons (**Fig. 5**). Eleven *Enterococcus* species found in this study also well-overlapped with major *Enterococcus* species previously documented from various environments (**Table 4**).

3.5. Reproducibility of QT-AMP

Seven pairs of samples were sequenced as duplicated samples, which used two different Enterolert trays, DNA extraction, and amplicon sequencing reactions (**Fig. 6**). Four sample pairs, ES28GR-Aug, ES28GR-Sep, IMPRGR51-Jan, and IMPRGR41-Feb, showed high correlation (R^2) among each sample and ranged between 0.893 and 0.989 ($p < 0.0001$) (**Fig. 6**). One other sample (ES47A-28GR-Jan) showed a moderate correlation (0.628, $p = 0.029$). Although, two samples (CES01SUR-Sep and ES47A-28GR-Feb) showed poor reproducibility. The high reproducibility was not assured by the high number of *Enterococcus* species found in each sample and the relative abundance of *Enterococcus* sequences in total sequencing counts.

3.6. Non-enterococci species found in Enterolert

Enterolert has been designed and used for quantifying enterococci in natural environments. However, it has been reported that Enterolert counting may contain non-enterococci counting (Budnick et al., 1996; Ferguson et al., 2013). In this study, we found some other non-target bacteria also grew in the wells of Enterolert. In total 16 genera in the wet season and 12 genera in the dry season were detected along with *Enterococcus* species (**Fig. 7**). As predicted, a majority of these species were facultative aerobes or anaerobes. However, a few detected bacterial taxa (*Synechococcus* and *Pelagibacter*) were not likely growing in Enterolert and at the incubation condition (no light, high nutrient, and high incubation temperature) but abundant in the original sampled water. Both *Synechococcus* and *Pelagibacter* are abundant in Caloosahatchee River water (Garcia et al. 2013).

4. DISCUSSION

Understanding the interaction between FIB density and surrounding environmental conditions has been one of the major interests of previous studies because it improves our capability to more accurately interpret the relationships between FIB densities and human health risks (Byappanahalli et al., 2012). The relationship of enterococci density and types of water bodies was previously examined in southwest Florida, and the authors concluded that the larger volumes of water allow for greater dilution and buffering against high enterococci densities; the highest densities of enterococci were found in a small stream while the lowest densities were found in Tampa Bay, which is connected to the Gulf of Mexico (Badgley et al., 2010). Our results echoed the previous study and showed that high enterococci densities were found in two small-sized rivers, the Estero River and Imperial River, while the Caloosahatchee River maintains a small number of enterococci throughout the year. Strong UV irradiation negatively influences the survival of enterococci in environments, especially in southwest Florida. Estero River and Imperial River have thick riparian vegetation that partially protects the river water from direct sunlight exposure and may prolong the survival of enterococci in river water (**Fig. 1**).

4.1. *Enterococcus* as indicators of fecal contamination in recreational water

In natural water bodies, enterococci are used as a microbial indicator of environmental contamination, because they are found in high concentrations in feces, and exposure to enterococci is linked to adverse health effects in beachgoers (Lipp et al., 2001). *Enterococci* concentrations play a role in recreational water quality standards, therefore understanding of origins, survival and transport mechanisms of enterococci are necessary for effective water management. Enterolert has some advantages over the traditional counting methods and it revolutionized FIB monitoring methods (Budnick et al., 1996). In regular usage, Enterolert is designed to provide quantitative data, which is equivalent to other membrane filtration and MPN methods (Budnick et al., 1996). It does not provide qualitative data for users unless the Quanti-Tray is used to isolate and characterize enterococci after counting. Therefore, a method to allocate sources of enterococci found in surface water would be beneficial to water managers. One advantage of enterococci as FIB is that the genus contains multiple species and

their habitats are diverse. A comparison of multiple studies suggests that the population structure of FIB in the environment is highly variable and the dynamics are complex (Badgley et al., 2010). QT-AMP assay is ideal for this purpose because this assay has a high sensitivity to detect various *Enterococcus* species in samples through its high-throughput format without any cultivation and identification processes. Currently, less than hundreds to up to a few thousand isolates have been phenotypically and genotypically tested (Badgley et al., 2010; Ferguson et al., 2013; Nayak et al., 2011; Ran et al., 2013), however, QT-AMP can easily provide millions of sequence reads in a short period and may be suitable to find minor species. In the present study, 11 enterococci isolates were found from water samples (**Fig. 4**), which exceeded nine species obtained from 1,279 presumptive enterococci isolates from a wide variety of environmental samples (beach water, urban runoff, and wastewater treatment plant influent & effluent) (Ferguson et al., 2013). It also exceeded eight species obtained from 2,441 *Enterococcus* strains originated from sand, sediment, water, and soil from Lake Superior. (Ran et al., 2013), suggesting the high sensitivity of QT-AMP (**Table 4**).

4.2. Detection of fecal indicator bacteria using high-throughput sequencing methods

One of our early challenges of the use of an amplicon sequencing technique to detect fecal indicator bacteria was the development of the method of signal amplification. For instance, in an early application of 16S rRNA amplicon sequencing to detect fecal coliform bacteria, the authors only identified a very small number of target sequences and decided to analyze the sequence data at the family level, rather than the genus or species level (Schang et al., 2016). As a part of our different projects, regular 16S rRNA gene amplicon sequence analysis was applied for water samples from the Estero River and Imperial River (unpublished data). No enterococci sequences were found in surface water samples from the Imperial River (402K in total reads from 13 samples) while 35 sequence reads from two wastewater contaminated ditch samples (all sequences were identified as *E. faecalis*) were found in the mixture of surface water, ditch water, and groundwater samples from the Estero River (623K in total reads from 31 samples). In this case, the relative abundance of enterococci found in two samples were

0.007% and 0.26% of the total reads for each sample, and 0.0056% of all sequence reads from Estero River ($n = 31$). In the present study, on average 72.3% of sequences (1,121K in total reads from 45 samples) were identified as enterococci. Thus, we successfully increased the detection sensitivity of enterococci targets. With the comparison of these two enterococci detection efficiency data (72.3% of QT-AMP and 0.0056% of regular amplicon sequencing), we confirmed that QT-AMP is 12,910 times more sensitive than the regular amplicon 16S rRNA gene sequencing. Therefore, we concluded that the approximate detection level of QT-AMP is four orders of magnitude higher than the regular 16S rRNA gene amplicon sequencing.

4.3. Non-enterococci species found in Enterolert

It has been reported that Enterolert counting likely contains some non-enterococci genera. The detection of non-*Enterococcus* species was slightly higher using Enterolert (8.4%) than for EPA Method 1600 (5.1%) when a wide variety of environmental samples (beach water, urban runoff, and wastewater treatment plant influent, and effluent) were tested (Ferguson et al., 2013). A fecal indicator bacteria study conducted in Del Rey Lagoon, California reported that 54 species were identified from 277 isolates cultured from the Quanti-Trays and many of them were not enterococci and *E. coli* (Dorsey et al., 2013). According to the U.S. EPA, the false-positive rate of enterococci isolation using mEI agar is documented as 6% (Messer and Dufour, 1998; US_Environmental_Protection_Agency, 1997). A very low level of false-positive rate (1.6%) was documented from the subtropical waters in Florida (Nayak et al., 2011). Based on QT-AMP, only one case (CES10SUR-Feb) resulted in no enterococci (1/45, 2.2%), which was a similar ratio reported from southwest Florida (Nayak et al., 2011), and only three cases (3/45, 6.7%) were conversely purely occupied by *Enterococcus* species. All other cases were a mixture of enterococci and other heterotrophic bacteria. Therefore, our data showed that a majority of Enterolert positive signals are actually the mixture of both enterococci and other heterotrophic bacteria. QT-AMP may likely have the potential to monitor not only enterococci but also other pathogenic bacteria commonly found in natural environments (Dorsey et al., 2013).

5. CONCLUSION

We documented a new approach to characterize enterococci using an amplicon sequencing platform from Quanti Trays after being used for counting the most probable numbers (MPN) of enterococci. We named this method as QT-AMP (Quanti-Tray-based amplicon sequencing). The importance of this work was that QT-AMP will strongly enhance semi-automated Enterolert-based enterococci monitoring methods in terms of quantity and identification. We tested surface water samples collected from three rivers in southwest Florida. We detected 11 *Enterococcus* species from 45 samples in 1.1 million sequence reads. The method detected eight cosmopolitan species (*E. faecalis*, *E. faecium*, *E. casseliflavus*, *E. hirae*, *E. mundtii*, *E. gallinarum*, *E. avium*, *E. durans*), which have been commonly documented from various enterococci isolation works. Thus, this identification result is compatible with previous cultivation data. The approximate detection level of QT-AMP is likely four orders of magnitude higher than the regular 16S rRNA gene amplicon sequencing. QT-AMP revealed that a majority of Enterolert positive signals are actually the mixture of both enterococci and other facultative aerobes and anaerobes. QT-AMP may likely have the potential to monitor not only enterococci but also other pathogenic bacteria commonly found in natural environments. QT-AMP has the power to streamline the quantification and identification of *Enterococcus* and can allow for more accurate and efficient microbial source tracking in various water management projects and human health risk assessment.

Acknowledgments

We thank the Lee County Environmental laboratory for technical assistance. We also thank Drs. L. Don Duke and Serge Thomas for helpful discussion. This work was supported by the Florida Sea Grant College Program with support from the National Oceanic and Atmospheric Administration, Office of Sea Grant, U.S. Department of Commerce (NA18OAR4170085 SUB00002463).

References

- Abbott, S., Caughley, B. and Scott, G. 1998. Evaluation of Enterolert® for the enumeration of enterococci in the marine environment. *New Zealand Journal of Marine and Freshwater Research* 32(4), 505-513.
- Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A. and Wishart, D.S. 2016. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Research* 44(W1), W147-W153.
- Badgley, B.D., Nayak, B.S. and Harwood, V.J. 2010. The importance of sediment and submerged aquatic vegetation as potential habitats for persistent strains of enterococci in a subtropical watershed. *Water Research* 44(20), 5857-5866.
- Budnick, G.E., Howard, R.T. and Mayo, D.R. 1996. Evaluation of Enterolert for enumeration of enterococci in recreational waters. *Applied and Environmental Microbiology* 62(10), 3881-3884.
- Byappanahalli, M. and Fujioka, R. 2004. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. *Water Science and Technology* 50(1), 27-32.
- Byappanahalli, M.N., Nevers, M.B., Korajkic, A., Staley, Z.R. and Harwood, V.J. 2012. Enterococci in the environment. *Microbiology and Molecular Biology Reviews* 76(4), 685-706.
- Byappanahalli, M.N., Shively, D.A., Nevers, M.B., Sadowsky, M.J. and Whitman, R.L. 2003. Growth and survival of *Escherichia coli* and enterococci populations in the macro-alga *Cladophora* (Chlorophyta). *FEMS Microbiology Ecology* 46(2), 203-211.
- Devriese, L., Cruz Colque, J., Haesebrouck, F., Desmidt, M., Uyttendaele, E. and Ducatelle, R. 1992. *Enterococcus hirae* in septicaemia of psittacine birds. *Veterinary Record* 130(25), 558-559.

- Dorsey, J.H., Carmona-Galindo, V.D., Leary, C., Huh, J. and Valdez, J. 2013. An assessment of fecal indicator and other bacteria from an urbanized coastal lagoon in the City of Los Angeles, California, USA. *Environmental Monitoring and Assessment* 185(3), 2647-2669.
- Eckner, K.F. 1998. Comparison of membrane filtration and multiple-tube fermentation by the Colilert and Enterolert methods for detection of waterborne coliform bacteria, *Escherichia coli*, and enterococci used in drinking and bathing water quality monitoring in southern Sweden. *Applied and Environmental Microbiology* 64(8), 3079-3083.
- Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C. and Knight, R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27(16), 2194-2200.
- Ferguson, D.M., Griffith, J.F., McGee, C.D., Weisberg, S.B. and Hagedorn, C. 2013. Comparison of *Enterococcus* species diversity in marine water and wastewater using Enterolert and EPA Method 1600. *Journal of Environmental and Public Health* 2013, Article ID 848049, 6 pages, 2013. <https://doi.org/10.1155/2013/848049>.
- Harwood, V.J., Whitlock, J. and Withington, V. 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. *Applied and Environmental Microbiology* 66(9), 3698-3704.
- Imamura, G.J., Thompson, R.S., Boehm, A.B. and Jay, J.A. 2011. Wrack promotes the persistence of fecal indicator bacteria in marine sands and seawater. *FEMS Microbiology Ecology* 77(1), 40-49.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35(6), 1547.
- Layton, B., Walters, S., Lam, L. and Boehm, A. 2010. *Enterococcus* species distribution among human and animal hosts using multiplex PCR. *Journal of Applied Microbiology* 109(2), 539-547.

- Lipp, E.K., Farrah, S.A. and Rose, J.B. 2001. Assessment and impact of microbial fecal pollution and human enteric pathogens in a coastal community. *Marine Pollution Bulletin* 42(4), 286-293.
- Messer, J.W. and Dufour, A.P. 1998. A rapid, specific membrane filtration procedure for enumeration of enterococci in recreational water. *Applied and Environmental Microbiology* 64(2), 678-680.
- Nayak, B.S., Badgley, B. and Harwood, V.J. 2011. Comparison of genotypic and phylogenetic relationships of environmental *Enterococcus* isolates by BOX-PCR typing and 16S rRNA gene sequencing. *Applied and Environmental Microbiology* 77(14), 5050-5055.
- Ramoutar, S. 2020. The use of Colilert-18, Colilert and Enterolert for the detection of faecal coliform, *Escherichia coli* and *Enterococci* in tropical marine waters, Trinidad and Tobago. *Regional Studies in Marine Science* 40, 101490.
- Ran, Q., Badgley, B.D., Dillon, N., Dunny, G.M. and Sadowsky, M.J. 2013. Occurrence, genetic diversity, and persistence of enterococci in a Lake Superior watershed. *Applied and Environmental Microbiology* 79(9), 3067-3075.
- Schang, C., Henry, R., Kolotelo, P.A., Prosser, T., Crosbie, N., Grant, T., Cottam, D., O'Brien, P., Coutts, S. and Deletic, A. 2016. Evaluation of techniques for measuring microbial hazards in bathing waters: A comparative study. *PloS one* 11(5), e0155848.
- Sercu, B., Van De Werfhorst, L.C., Murray, J.L. and Holden, P.A. 2011. Cultivation-independent analysis of bacteria in IDEXX Quanti-Tray/2000 fecal indicator assays. *Applied and Environmental Microbiology* 77(2), 627-633.
- Urakawa, H., Martens-Habbena, W. and Stahl, D.A. 2010. High abundance of ammonia-oxidizing Archaea in coastal waters, determined using a modified DNA extraction method. *Applied and Environmental Microbiology* 76(7), 2129-2135.
- US Environmental Protection Agency 1997 Method 1600: membrane filter test method for enterococci in water, US Environmental Protection Agency Washington, DC.

Yamahara, K.M., Layton, B.A., Santoro, A.E. and Boehm, A.B. 2007. Beach sands along the California coast are diffuse sources of fecal bacteria to coastal waters. *Environmental Science & Technology* 41(13), 4515-4521.

Table 1. Resolution of three regions of 16S rRNA gene amplified by three primer sets

Primer pair	Length (bp)	Variable site (n)	Covering variable region	100% identical pairs (n)	Distinguishable species (n)	Pairwise distance				
						Observation	Min	Max	Mean	SD
27f & 519r	506	179	V1-V3	13	54	1830	0.000	0.216	0.080	0.038
515f & 806r	253	37	V4	225	25	1830	0.000	0.111	0.013	0.015
515f & 926r	374	62	V4-V5	149	35	1830	0.000	0.118	0.013	0.017

Table 2. Physical and chemical parameters collected from three southwest Florida rivers

RIVER	SAMPLE ID	DATE	TEMP	PH	DO	COND	Chl a	NH3	NOX	O-PO4	TN	TP	ENTERO ECOLI	
			°C	units	mg/L	µmhos/cm	ug/L	mg/L as N	mg/L as N	mg/L as P	mg/L as N	mg/L as P	0 ml	MPN/100
Caloosahatchee	CES01SUR_Aug	8/11/2020	31.8	7.5	5.1	587	21.13	0.01	0.29	0.082	1.60	0.140	10	10
	CES01SUR_Sep	9/8/2020	30.7	7.2	2.8	443	3.33	0.05	0.34	0.120	1.40	0.170	10	20
	CES01SUR_Jan	1/14/2021	19.2	7.6	7.3	460	0.61	0.01	0.23	0.045	1.00	0.083	10	10
	CES01SUR_Feb	2/10/2021	22.5	7.8	9.5	446	22.49	0.01	0.10	0.031	1.20	0.093	nd	75
	CES03SUR_Aug	8/11/2020	31.1	7.6	3.9	810	3.91	0.08	0.28	0.093	1.40	0.120	10	10
	CES03SUR_Sep	9/8/2020	30.4	7.3	3.2	458	0.50	0.08	0.29	0.106	1.40	0.160	10	20
	CES03SUR_Jan	1/14/2021	18.9	7.4	7.2	448	2.81	0.01	0.21	0.045	1.00	0.080	10	31
	CES03SUR_Feb	2/10/2021	22.8	7.7	9.1	493	13.86	0.01	0.12	0.036	1.10	0.092	20	nd
	CES04SUR_Aug	8/11/2020	31.5	7.9	5.9	4138	3.46	0.05	0.18	0.099	1.20	0.130	10	10
	CES04SUR_Sep	9/8/2020	30.6	7.6	4.9	473	1.41	0.05	0.30	0.116	1.30	0.150	10	10
	CES04SUR_Jan	1/14/2021	20.9	7.7	7.9	1517	0.75	0.03	0.18	0.044	0.99	0.110	10	53
	CES04SUR_Feb	2/10/2021	24.6	7.7	7.6	987	3.14	0.02	0.16	0.048	1.00	0.082	10	nd
	CES06SUR_Aug	8/11/2020	30.4	7.8	6.3	15470	2.86	0.03	0.02	0.080	0.97	0.110	10	20
	CES06SUR_Sep	9/8/2020	29.4	7.6	5.6	9479	3.40	0.15	0.32	0.124	1.40	0.170	782	364
	CES06SUR_Jan	1/14/2021	17.3	8.0	9.1	12195	0.84	0.01	0.09	0.029	0.85	0.051	10	nd
	CES06SUR_Feb	2/10/2021	23.1	8.0	8.5	13575	1.84	0.01	0.01	0.019	0.65	0.046	10	nd
	CES10SUR_Aug	8/11/2020	31.0	7.7	4.8	37776	3.29	0.03	0.02	0.038	0.78	0.055	10	20
	CES10SUR_Sep	9/8/2020	29.3	7.6	5.4	21898	1.05	0.13	0.24	0.087	1.20	0.120	20	42
	CES10SUR_Jan	1/14/2021	17.2	7.8	8.1	29259	0.50	0.01	0.03	0.020	0.89	0.038	10	nd
	CES10SUR_Feb	2/10/2021	22.1	7.8	7.4	30890	1.07	0.01	0.01	0.015	0.79	0.039	10	nd
Estero	ES47A-28GR_Aug	8/12/2020	28.4	7.1	1.8	667	1.56	0.06	0.16	0.004	0.83	0.008	110	113
	ES47A-28GR_Sep	9/22/2020	25.8	7.2	3.3	413	1.47	0.06	0.05	0.004	0.76	0.024	37	36
	ES47A-28GR_Jan	1/12/2021	19.9	7.3	5.2	712	2.16	0.21	0.23	0.004	1.10	0.015	56	47
	ES47A-28GR_Feb	2/24/2021	23.0	7.6	5.3	751	3.15	0.18	0.23	0.004	0.91	0.023	238	5
	ES47A-15GR_Aug	8/12/2020	28.5	7.2	3.5	797	0.92	0.08	0.31	0.011	0.94	0.028	2420	1046
	ES47A-15GR_Sep	9/22/2020	26.8	7.3	4.0	581	0.84	0.10	0.16	0.005	0.76	0.028	313	179
	ES47A-15GR_Jan	1/12/2021	20.7	7.1	5.1	1032	3.18	0.08	0.39	0.006	0.99	0.024	866	162
	ES47A-15GR_Feb	2/24/2021	22.8	7.3	4.4	2286	11.43	0.05	0.37	0.008	0.87	0.039	866	nd
	ES47A-4GR_Aug	8/12/2020	28.9	7.2	3.4	5671	5.15	0.06	0.24	0.018	0.87	0.041	219	435
	ES47A-4GR_Sep	9/22/2020	27.0	7.2	3.8	675	1.17	0.12	0.22	0.014	0.97	0.053	649	166
	ES47A-4GR_Jan	1/12/2021	19.5	7.2	5.7	5954	13.98	0.10	0.26	0.004	0.93	0.037	249	nd
	ES47A-4GR_Feb	2/24/2021	22.9	7.3	5.0	11888	6.94	0.05	0.18	0.004	0.76	0.046	299	nd
Imperial	IMPRGR80_Aug	8/25/2020	29.2	7.0	2.4	284	0.75	0.09	0.03	0.005	1.00	0.028	45	45
	IMPRGR80_Sep	9/21/2020	29.0	7.0	3.0	193	1.03	0.03	0.01	0.004	0.81	0.032	178	30
	IMPRGR80_Jan	1/20/2021	20.4	6.9	3.4	512	0.7	0.22	0.08	0.014	0.89	0.026	75	60
	IMPRGR80_Feb	2/16/2021	24.5	7.0	1.3	627	2.09	0.41	0.06	0.027	1.20	0.040	46	86
	IMPRGR51_Aug	8/25/2020	28.7	7.3	3.2	460	1.52	0.05	0.17	0.012	1.10	0.032	1414	517
	IMPRGR51_Sep	9/21/2020	28.0	7.2	3.3	361	1.07	0.05	0.15	0.035	0.82	0.081	1300	488
	IMPRGR51_Jan	1/20/2021	18.4	7.3	5.5	596	1.42	0.06	0.20	0.012	0.67	0.027	2420	1986
	IMPRGR51_Feb	2/16/2021	24.7	7.3	2.5	588	6.56	0.04	0.12	0.026	0.76	0.063	1986	1120
	IMPRGR41_Aug	8/25/2020	28.9	7.1	2.2	399	1.59	0.05	0.10	0.033	1.00	0.033	579	548
	IMPRGR41_Sep	9/21/2020	28.1	7.2	2.5	601	3.31	0.07	0.17	0.035	0.75	0.130	727	579
	IMPRGR41_Jan	1/20/2021	18.2	7.2	4.4	694	3.12	0.07	0.18	0.019	0.70	0.034	1120	770
	IMPRGR41_Feb	2/16/2021	25.2	7.3	2.3	3660	7.09	0.07	0.14	0.029	0.80	0.064	1300	nd
	IMPRGR30_Aug	8/25/2020	29.5	7.1	3.0	990	2.97	0.05	0.08	0.013	1.00	0.026	104	131
	IMPRGR30_Sep	9/21/2020	29.5	7.1	3.4	247	1.38	0.05	0.03	0.012	0.80	0.059	135	133
	IMPRGR30_Jan	1/20/2021	19.4	7.1	5.2	2081	9.12	0.04	0.16	0.004	0.81	0.027	299	365
	IMPRGR30_Feb	2/16/2021	25.8	7.3	3.9	11197	14.8	0.04	0.10	0.004	0.80	0.031	2420	nd

nd: No data are available.

Table 3. Quantification of enterococci using Enterolert

River	No of sample	Sampling point	Geometric mean (MPN/100 ml)	Median (MPN/100 ml)	Range (MPN/100 ml)	Statistical threshold value criterion (STV) (frequency and percent) ^a	Geometric mean criterion ^b	Correlation with <i>E. coli</i> MPN count (R ²)	Average width of river (m) ^c	Size of the watershed (Km ²)
Caloosahatchee	19	5	13.5	10	10-782	1/19 (5.2)	0/5 (0)	0.98	1373	1100
Estero	12	3	288.0	274	37-2420	9/12 (75)	3/3 (100)	0.79	17	931
Imperial	16	4	432.9	653	45-2420	12/16 (75)	4/4 (100)	0.85	16	931

^aTested water samples exceeded USEPA's statistical threshold value (STV) of 130 CFU/100 ml.

^bTested samples from each sampling location exceeded USEPA's geometric mean criterion of 35 CFU/100 ml in the whole sampling period.

^c Average river width is determined based on the location of the sampling sites.

^d The Estero River and the Imperial River share the watershed.

Table 4. List of major Enterococcus species found in each study

	This study	Ran et al. 2013	Ferguson et al. 2013	Badgley et al. 2010	Nayak et al. 2011
Number of isolates		2441	1279	50	61
	<i>E. casseliflavus</i>	<i>E. casseliflavus</i>	<i>E. casseliflavus</i>	<i>E. casseliflavus</i>	<i>E. casseliflavus</i>
	<i>E. faecalis</i>	<i>E. faecalis</i>	<i>E. faecalis</i>	<i>E. faecalis</i>	
	<i>E. faecium</i>	<i>E. faecium</i>	<i>E. faecium</i>	<i>E. faecium</i>	<i>E. faecium</i>
	<i>E. hirae</i>	<i>E. hirae</i>	<i>E. hirae</i>	<i>E. hirae</i>	<i>E. hirae</i>
	<i>E. mundtii</i>	<i>E. mundtii</i>	<i>E. mundtii</i>	<i>E. mundtii</i>	<i>E. mundtii</i>
	<i>E. gallinarum</i>	<i>E. gallinarum</i>	<i>E. gallinarum</i>		
	<i>E. avium</i>	<i>E. avium</i>	<i>E. avium</i>		
	<i>E. durans</i>	<i>E. durans</i>	<i>E. durans</i>		
	<i>E. lactis</i>				
	<i>E. rivorum</i>				
	<i>E. malodoratus</i>				
			<i>E. saccharolyticus</i>		
					<i>E. gilvus</i>
					<i>E. silesiacus/moraviensis/caccae</i>
Number of species found	11	8	9	6	6

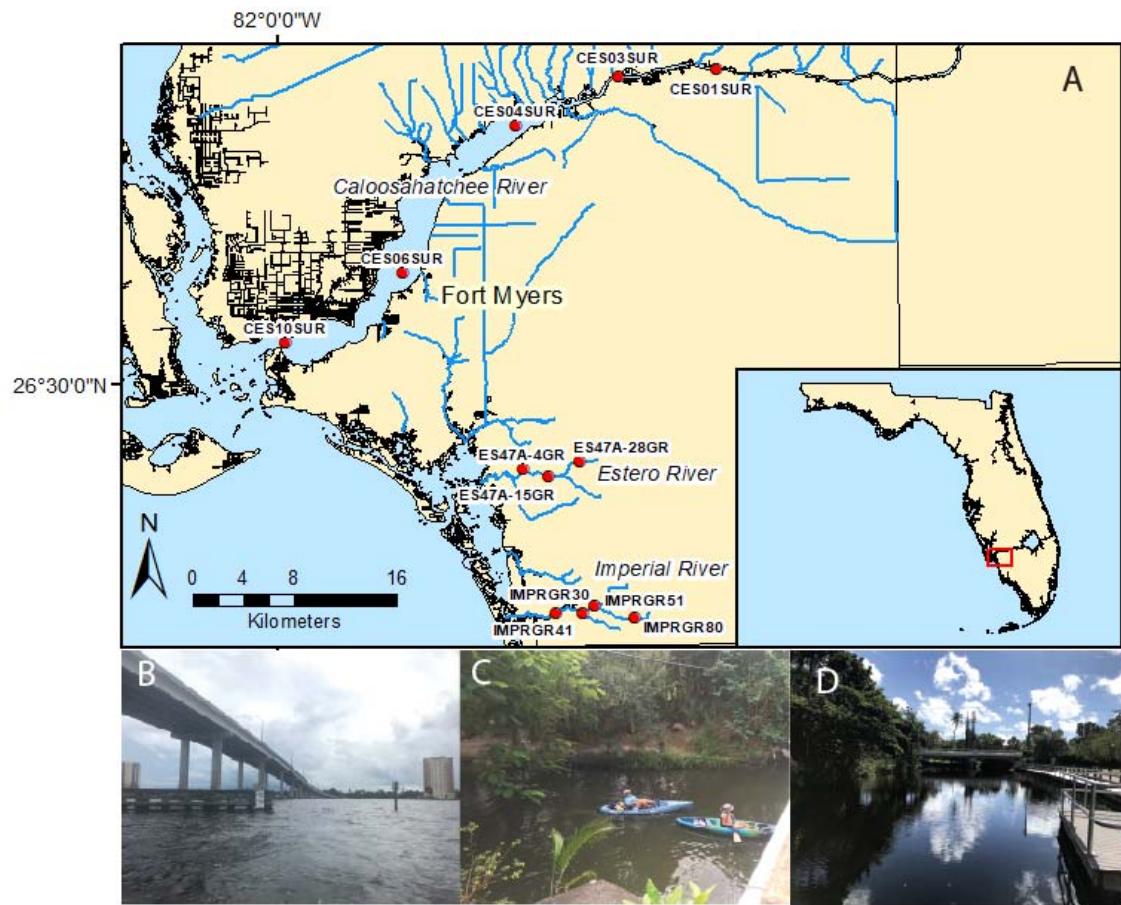


Fig. 1. Map showing the sampling sites located in three southwest Florida rivers. Sampling map (A), the Caloosahatchee River near CES06 SUR (B), the Estero River near ES47-15GR (C), and the Imperial River near IMPRGR80 (D).

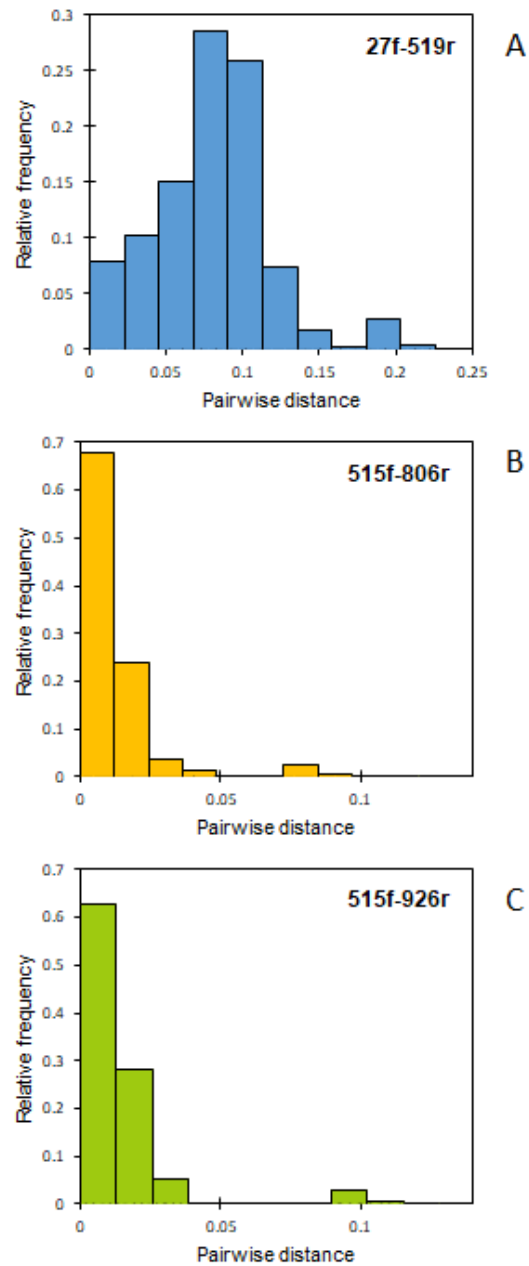


Fig. 2. Histograms showing the relative frequency of the distribution of pairwise distance calculated from the partial 16S rRNA gene sequences of 61 *Enterococcus* species. Three primer pairs are used for each histogram. The pairwise distance of two identical sequences is zero and 5% dissimilarity is shown as 0.05.

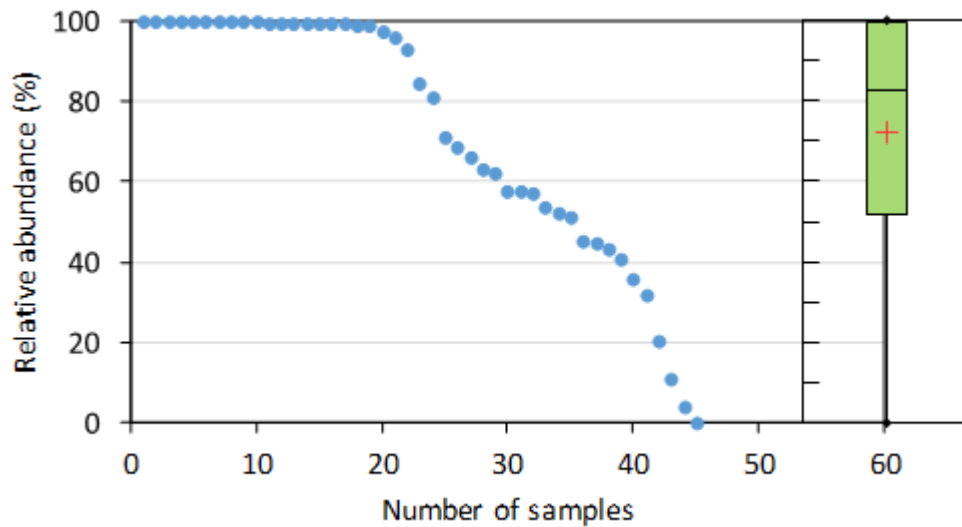


Fig. 3, Relative abundance of enterococci in the total amplicon sequence reads. An inset box plot shows median (centerline, 82.7), first and third quartiles, mean (red cross, 72.3 ± 4.6 [SE], $n = 44$), and range (0 to 100%).

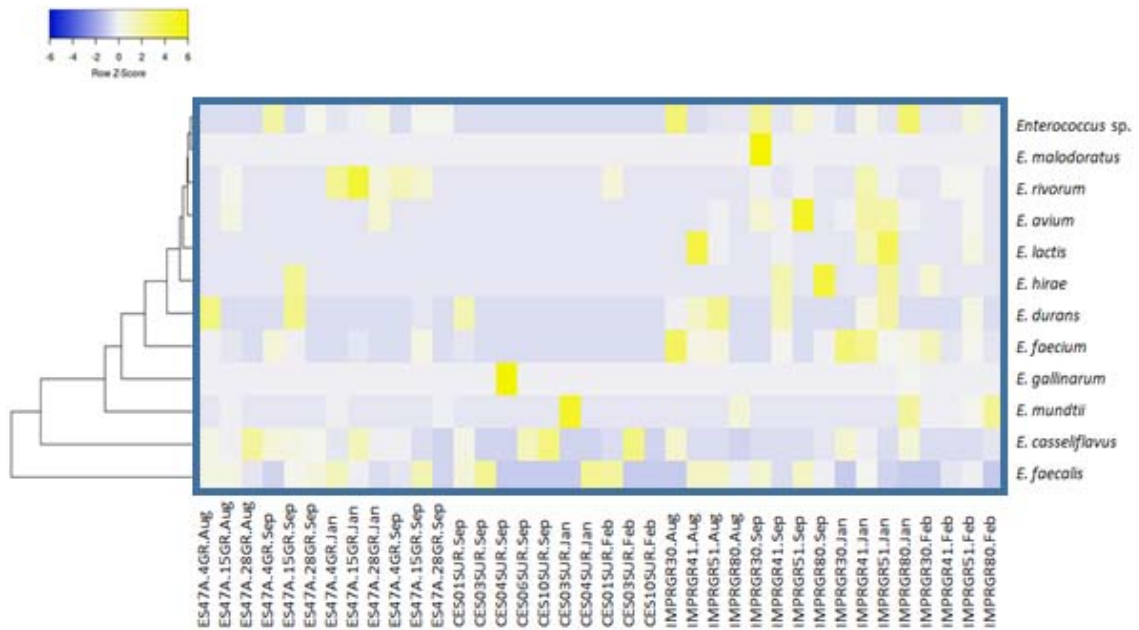


Fig. 4. Heat map of the relative abundance of sequencing reads showing enterococci composition at the species level. Intense yellow colors indicate high standardized relative abundance values (row Z-scores), while blue colors indicate low standardized relative abundance values. Taxa are clustered using the Euclidean distance method and hierarchical clustering with the average linkage method.

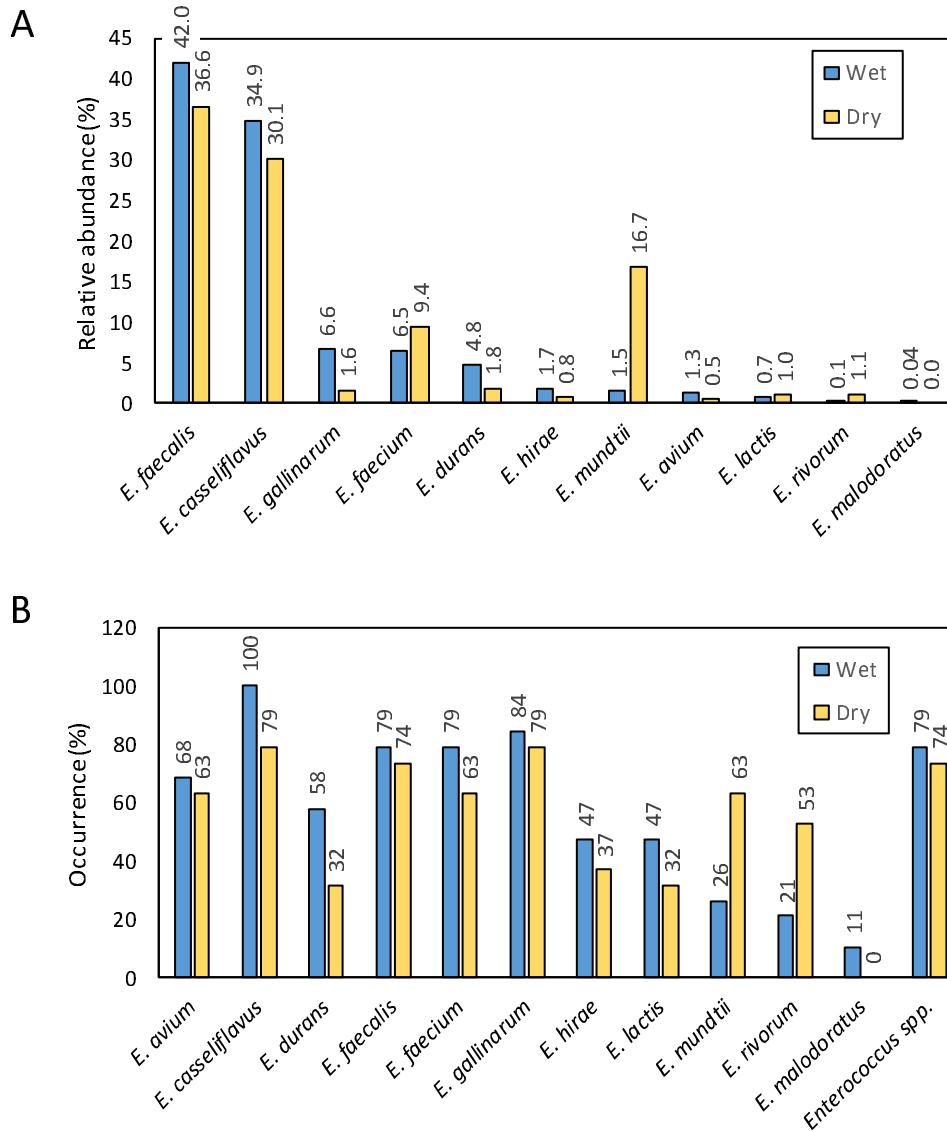


Fig. 5. Seasonal comparison of enterococcus abundance. The read frequency of each enterococci species among the total enterococci reads (A). The occurrence of each *Enterococcus* species found among water samples (B).

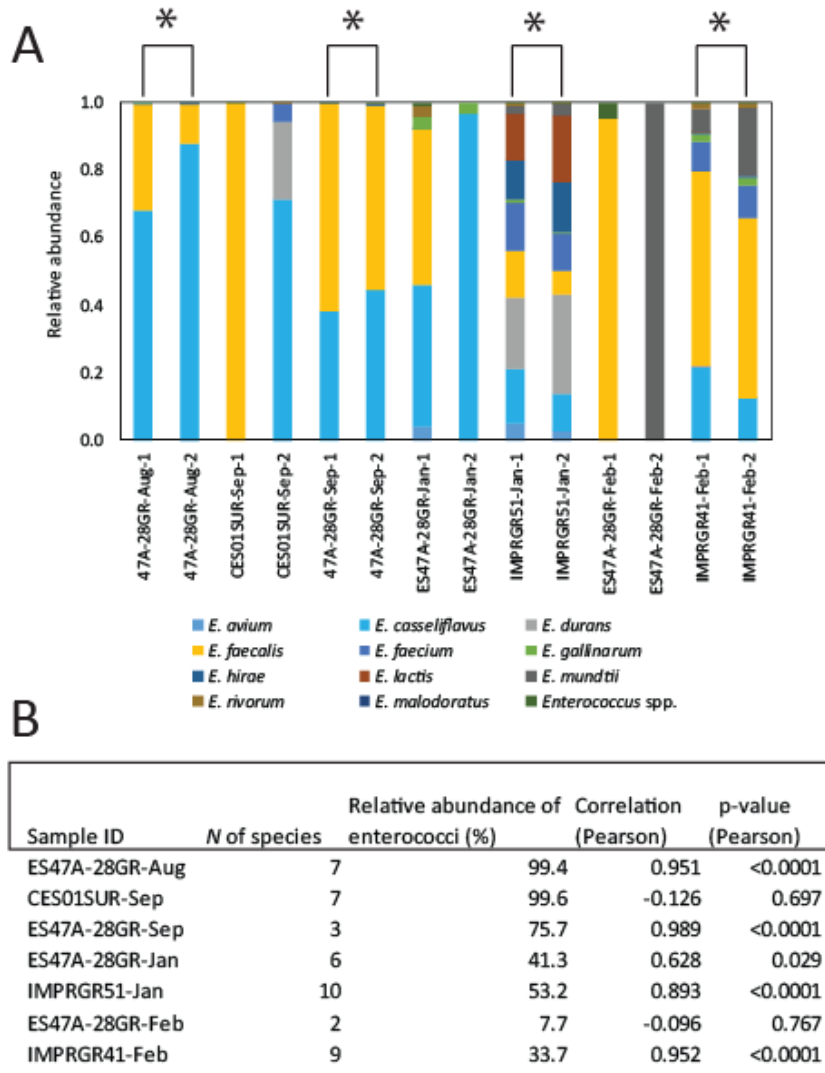


Fig. 6. Reproducibility of seven duplicated samples tested by the relative composition of enterococci species. The relative abundance of each enterococci species (A). General statistical data and the relative abundance of enterococci among the total amplicon sequencing reads (B). (A) A star is depicted if replicates are significant ($p < 0.01$). Some species are invisible because of their minor abundance. (B) The number of enterococci species found in the averaged samples is shown.

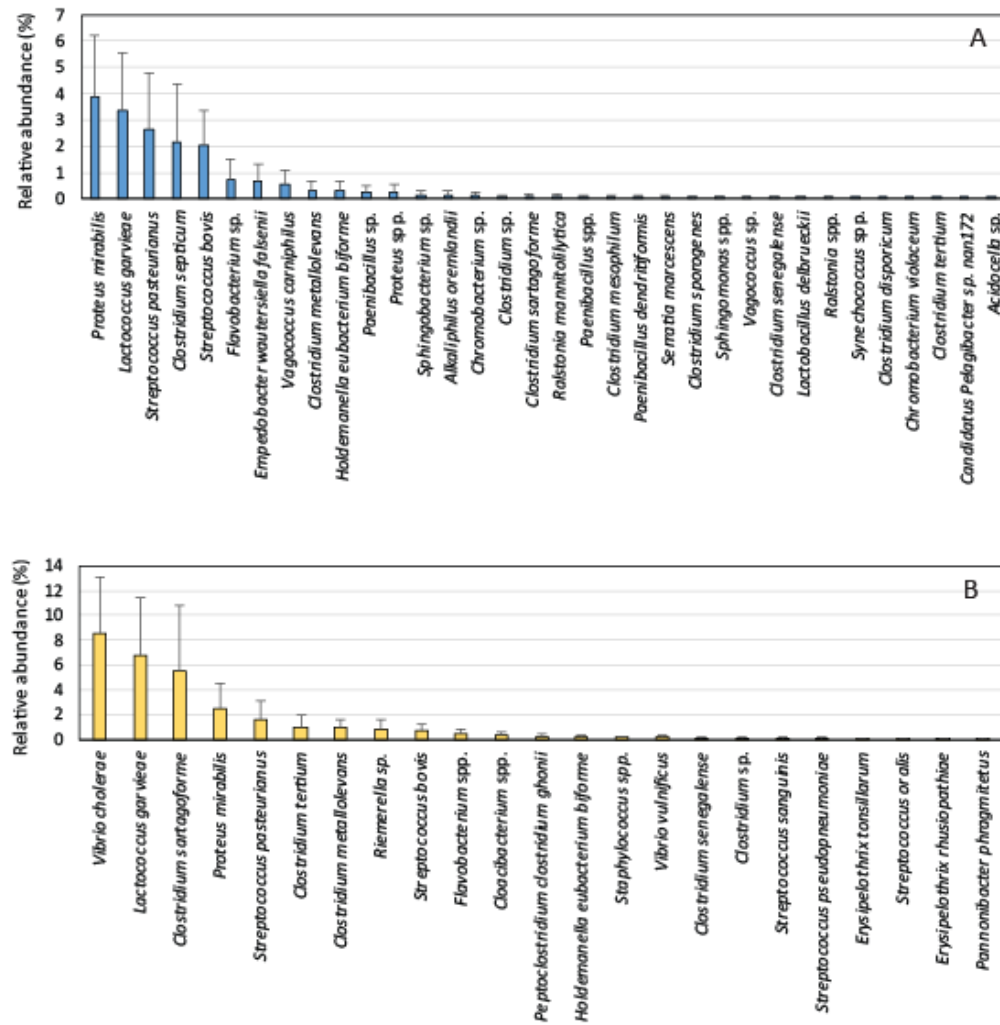


Fig. 7. Relative abundance of detected non-enterococci bacteria. Rainy season (A) and dry season (B). Data are shown as mean and standard deviation ($n =$ variable for each species).