frontiers
in Microbiology

The Development of a Novel qPCR Assay-Set for Identifying Fecal Contamination Originating from Domestic Fowls and Waterfowl in Israel

Shoshanit Ohad[1]*, Shifra Ben-Dor[2], Jaime Prilusky[2], Valeria Kravitz[1], Bareket Dassa[3], Vered Chalifa-Caspi[3], Yechezkel Kashi[4] and Efrat Rorman[1]

[1] National Public Health Laboratory Tel Aviv, Ministry of Health, Tel Aviv, Israel, [2] Bioinformatics Unit, Department of Biological Services, Weizmann Institute of Science, Rehovot, Israel, [3] Bioinformatics Core Facility, National Institute of Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva, Israel, [4] Faculty of Biotechnology and Food Engineering, Technion – Israel Institute of Technology, Haifa, Israel

The emerging microbial source tracking (MST) methodologies aim to identify fecal contamination originating from domestic and wild animals, and from humans. Avian MST is especially challenging, primarily because the Aves class includes both domesticated and wild species with highly diverse habitats and dietary characteristics. The quest for specific fecal bacterial MST markers can be difficult with respect to attaining sufficient assay sensitivity and specificity. The present study utilizes high throughput sequencing (HTS) to screen bacterial 16S rRNA genes from fecal samples collected from both domestic and wild avian species. Operational taxonomic unit (OTU) analysis was then performed, from which sequences were retained for downstream quantitative polymerase chain reaction (qPCR) marker development. Identification of unique avian host DNA sequences, absent in non-avian hosts, was then carried out using a dedicated database of bacterial 16S rRNA gene taken from the Ribosomal Database Project. Six qPCR assays were developed targeting the 16S rRNA gene of *Lactobacillus*, *Gallibacterium*, *Firmicutes*, *Fusobacteriaceae,* and other bacteria. Two assays (Av4143 and Av163) identified most of the avian fecal samples and demonstrated sensitivity values of 91 and 70%, respectively. The Av43 assay only identified droppings from battery hens and poultry, whereas each of the other three assays (Av24, Av13, and Av216) identified waterfowl species with lower sensitivities values. The development of an MST assay-panel, which includes both domestic and wild avian species, expands the currently known MST analysis capabilities for decoding fecal contamination.

Keywords: birds, fecal contamination, HTS, MST, qPCR

INTRODUCTION

Microbial source tracking (MST) methodologies (Stoeckel and Harwood, 2007; Gourmelon et al., 2010; Marti et al., 2013) improve microbial monitoring resolution of surface and underground water influenced by human and animal fecal contamination. MST facilitates analysis of fecal contamination sources and their impact in a given geographical area. Data evaluation can support

environmental management decisions. For example, the efficacy of corrective actions (Converse et al., 2012) can be monitored with MST. Avian species are a known reservoir of diverse pathogens (VerCauteren et al., 2012; Sandor et al., 2014), and may contribute to local and global epidemiology.

The fecal microbiome is dynamic, influenced by age, nutrition, and health. The Aves class (Zhu et al., 2002; Scupham et al., 2008; Lu et al., 2009; Xenoulis et al., 2010), comprising both domestic and wild birds, demonstrates highly diverse habitats and dietary characteristics which challenge MST marker design and evaluation. These factors have a direct influence on assay sensitivity and specificity values, which are crucial for assay assessment, because they reflect true positive and true negative detection rates. Assay sensitivity indicates the incidence of the marker within the target host population, and assay specificity points to the exclusivity of the marker to the target host. MST markers (Shanks et al., 2008; Boehm et al., 2013; Reischer et al., 2013) can be highly sensitive and specific; serial MST markers for gulls (Sinigalliano et al., 2013), targeting the *Catellicoccus marimammalium* bacteria, demonstrated a high percent of sensitivity values ranging from 81 to 100%. In this publication, the original specificity of assays varied from 37 to 85%, and was increased to 86 to 96% based on the following calculations: standardization of data, interpretation detected not quantified (DNQ) results as true negative and pigeon results as true positive. Notably, testing pigeon droppings using the Gull quantitative polymerase chain reaction (qPCR) assays yielded amplification levels comparable to those found in gull droppings, and in selected cases, even higher. This finding emphasizes other aspects of the complexity of avian MST design as different birds share the same habitats/environmental niches and possess similar bacterial communities. Moreover, they can easily fly and change locations carrying bacteria to different geographic locations.

Molecular human MST markers target unique sequences in bacterial, viral, coliphage, and mitochondrial DNA. Human MST, based on qPCR or PCR, identifies bacterial DNA sequences affiliated with (Ahmed and Katouli, 2008) *Enterococcus faecium*, (Harwood et al., 2009) *Methanobrevibacter smithii*, (Balleste and Blanch, 2011) *Bifidobacterium dentium*. The 16S rRNA gene (Bernhard and Field, 2000; Layton et al., 2006; Haugland et al., 2010) is currently the most prominent amplification target among *Bacteroidales* assays. The 16S rRNA gene (Clarridge, 2004) is an important housekeeping gene composed of highly conserved sequences as well as variable regions. Targeting PCR to its conserved regions allow DNA amplification originating from a large range of prokaryotes, while the hyper-variable sequences can be used for discriminatory sequence analyses. The number of base substitutions in the 16S rRNA gene is relatively low in close related bacteria species presenting low discriminatory resolution at sub genus and species level. These features turned 16S rRNA gene into a prominent gene used in bacterial taxonomy and phylogeny. However, other genes such as (Yampara-Iquise et al., 2008) the *thetaiotaomicron*-specific α-1-6, mannanase, and (Shanks et al., 2009) *Bacteroidales*-like cell surface-associated genes have also been reported. Avian MST markers developed thus far, identify bacteria of various taxonomic groups, including (Shen et al., 2013) *Faecalibacterium*, (Green et al., 2012; Ryu et al.,

2012) *C. marimammalium*, and (Weidhaas and Lipscomb, 2013; Ryu et al., 2014) *Brevibacterium* sp. 16S rRNA gene clone libraries (Lu et al., 2009) from geese droppings were used to explore microbial communities from these sources to design host-specific assays. Recently (Shen et al., 2013), *in silico* data mining focusing on *Faecalibacterium* 16S rRNA gene from various animal species recognized an intervening sequence unique to poultry. The assay also identified feces of poultry and turkey, but not goose or seagull.

Our goal was to screen fecal samples of different avian species, both domestic and wild (waterfowl), and to identify specific DNA sequences within each group. Avian species may have overlapping bacterial profiles. We attempted to reveal subtle variations using the same experimental setting simultaneously. This approach required massive data collection which was accomplished using high throughput sequencing (HTS).

MATERIALS AND METHODS

Fecal Sample Collection and Sample Processing

A total of 213 fecal samples from humans (25 wastewater), domestic and wildlife animals (52 battery hen, 22 poultry, 12 turkey, 23 waterfowl, 40 bovine, 20 swine, 9 equine, 10 gazelle) were collected at diverse farms and at various geographic locations in Israel from April 2010 to August 2014. Samples of gazelle and waterfowl droppings were taken at a single location, and samples from other hosts were collected at several separate locations (the amount of which appears in brackets) as follows: urban wastewater treatment plants (5); battery hens (4 enclosures), poultry (2 enclosures), turkey (2 enclosures), bovine (5 cow sheds), swine: (2 pig sties), and equine (2 stables). Feces and dropping samples collected were as fresh as possible, and grab samplings of wastewater primary effluents were carried out into sterile polypropylene tubes and containers, respectively. Samples were shipped to the laboratory on ice within up to 3 h and kept at −80°C until the DNA extraction step.

DNA extraction from all environmental samples was conducted using the PowerSoil DNA isolation kit (MoBio, Carlsbad, CA, USA), according to the protocol of the manufacturer. The quality and quantity of avian source DNA preparations were analyzed by NanoDrop ND 1000 UV spectrophotometer (Thermo Fisher Scientific, Inc., Vienna, Austria) and samples were kept at −20°C until use.

High Throughput Sequencing

The HTS initial sample preparation step comprised two sequential PCR amplifications. First, a portion of 16S rRNA gene spanning V1–V3 regions was amplified by the TaKaRa Ex Taq polymerase (TAKARA Bio, Inc., Otsu, Japan), using the primers 27F, 5′-AGAGTTTGATCCTGGCTCAG-3′, and 518R, 5′-GTATTACCGCGGCTGCTGG-3′. PCR was conducted in a Biometra thermocycler (Biometra, Goettingen, Germany), under the following cycling conditions: an initial PCR step for 7 min at 95°C, followed by 35 cycles: 30 s at 95°C, 30 s at 55°C and

60 s at 72°C, and a final extension step at 72°C for 10 min. PCR products were column-purified by Wizard SV Gel and PCR Clean-up System (Promega, Mannheim, Germany). A second PCR amplification was carried out using primers of 27F and 518R fused to Titanium Primers: A, 5′-CGTATCGCCTCCCT CGCGCCATCAG-MID-27F -3′, and B, 5′-CTATGCGCCTTGC CAGCCCGCTCAG-MID-518R-3′, under the same cycling condition. The sequences of the six Multiplex Identifiers (MID) were as follows: turkey (MID1, ACGAGTGCGT); waterfowl 1 (MID2, ACGCTCGACA); waterfowl 2 (MID3, AGACGCACTC); stork (MID4, AGCACTGTAG): battery hen (MID5, ATCAGACACG); poultry (MID6: ATATCGCGAG). Amplification products of 650 bp length were extracted from 0.8% agarose gel using Wizard SV Gel and PCR Clean-up System (Promega, Mannheim, Germany). Quality assessment of PCR amplification products, 454 pyrosequencing library construction and sequencing were performed at DYN Labs Ltd. (Dyn labs, Caesarea, Israel). The samples were tested using HTS and named poultry, turkey, battery hen, waterfowl 1, waterfowl 2, and stork. The reads have been deposited in the SRA database (NCBI) accession number SRP065761 under the BioProject accession number PRJNA300726.

High Throughput Sequencing Data Analysis

High-quality multiplex reads were processed using QIIME Version 1.5.0 (Caporaso et al., 2010), selecting minimum quality score >25, minimal length >200 bp, containing no ambiguous bases or mismatches in the primers. The workflow for *de novo* operational taxonomic units (OTUs) picking included clustering of all the sequences into OTUs, based on their sequence similarity using the uclust clustering algorithm (Edgar, 2010). A representative set of sequences for each OTU was selected for subsequent analyses, including a taxonomy assignment by the RDP-classifier Version 2.2, using Naive Bayes classification (Wang et al., 2007). Training reference sequence set used for the classifier was the Greengenes sequence database release 12_10 (McDonald et al., 2012). Sequence identities of 97, 95, and 80% were used to approximate the species, genus and family taxonomical levels respectively, as previously described (Kuczynski et al., 2011).

In order to identify OTUs exclusive to each sample, we selected OTUs with a minimum of eight reads which were assigned only to a single group out of the six samples. A representative sequence of each 97% sequence distance defined OTU was compared for similarities with sequence reads derived from the various samples, using BLAST algorithm (Altschul et al., 1990) version 2.2.28+, in order to verify that it uniquely matches only one sample (BLASTn, with a minimum e-value of 0.001). An additional BLAST search was performed with each selected read against the non- redundant database at NCBI to ensure species specificity.

Database Construction

The database of bacterial 16S rRNA gene was taken from the RDP database (Cole et al., 2014) unaligned sequences in GenBank format, Release 11.1. The sequences were filtered for those whose source indicated feces (feature table tag isolation_source, search terms: feces, fecal, faeces, faecal, stool), were not from a viral or vertebrate source organism, and had a clearly indicated host species). On several entries the host species were automatically corrected to translate from common names to those listed in NCBI's Taxonomy, e.g., the host swine replaced by *Sus scrofa* and the host dog replaced by *Canis lupus*. The final database is composed of 127,218 sequences, and is available upon request.

Search for Unique Sequences

The sequences were input into the open access MST search server: http://mst.weizmann.ac.il. The input sequence was compared to the database fecal 16S rRNA gene (described above) with FASTA Version 36.3.6 (Pearson and Lipman, 1988), with the E-value set to 1e-10. Clustering on the hits of the similarity search was performed to reduce redundancy using CD-HIT Version 4.6 (Fu et al., 2012), with the following parameters changed from the defaults: -M 0 –T 2. The longest representative sequence was taken for each cluster. Sequences left after removing redundancy were aligned with Muscle Version 3.8.425 (Edgar, 2004), with the following parameters: -maxhours 1, -maxiters 1, -diags. The alignment was trimmed to the region that overlapped with the input sequence. In order to find unique sequences that could be taken as probes, a sliding window of 18 was used, and windows with a difference of 90% were marked as potential unique sequences using an in-house script. In addition, these regions were checked manually (Cole et al., 2014) with Probe Match and Seqmatch at the RDP website and BLAST at NCBI (Johnson et al., 2008).

Primers and probes for qPCR assays were designed from identified unique sequences in using Primer Express software (Applied Biosystems, Foster City, CA, USA) and consulting Agentek Ltd. (Agentek, Tel Aviv, Israel).

Quantitative Polymerase Chain Reaction

Quantitative polymerase chain reaction assays were performed using the StepOnePlus platform (Applied Biosystems, Foster City, CA, USA), using Universal ABI Mix (Applied Biosystems, Foster City, CA, USA) under the following conditions: 15 min at 95°C, followed by 45 cycles: 15 s at 95°C and 60 s at 60°C. DNA preparations were diluted 1:5 in Ultra Pure Water, PCR grade (Fisher Biotec, Australia) to be further tested in qPCR assays. Duplicate qPCR reactions were carried out in 20 µl final volume in which primer and probe (**Table 1**) concentrations were 500 and 250 nM, respectively.

The insert-vector constructs resulting by the ligation of amplicons into pGEM vectors (Promega, Mannheim, Germany) were used for generating the standard curves. Plasmids were quantified spectrophotometrically, from which the gene copies were calculated. Serial dilutions of cloned pGEM plasmid were carried out independently in triplicates. Plasmid -constructs with the following qPCR amplicons Av13, Av24, Av163, Av216, and Av4143 were sequenced using a commercial sequencing service (HyLabs, Rehovot, Israel).

TABLE 1 | Primers and probes of MST qPCR assays used in the study.

Assay	Primers and probe sequence (5′–3′)
Av4143	Av4143F: TGCAAGTCGAACGAGGATTTCT
	Av4143R: TCACCTTGGTAGGCCGTTACC
	Av4143P: [FAM]-AGGTGGTTTTGCTATCGCTTT-[BHQplus]
Av163	Av163F: TCCGGACTACGATGCACTTTC
	Av163R: GCATACAGAGGGAGGCGAAG
	Av163P: [FAM]- AGTTTCGCTCCGTATCGC- [BHQplus]
Av43	Av43F: GCAAGTTGAGCGGAGATATGG
	Av43R: ATCGGCCTATCCCCCAATATA
	Av43P: [FAM]-CTCTTTATATTTTAGCAGCGAACG-[BHQplus]
Av216	Av216F: ATAAGCGAGGGATAACTATTGGAAAC
	Av216R: AACTAGCTAATGCACCGCAGAT
	Av216P: [FAM]-AAGCAACTGTTTCACTTATGGAT-[BHQplus]
Av24	Av24F: GGAAACGACAGCTAATACCGGATA
	Av24R: CTCTTGGCGCATATAGCTTTCA
	Av24P: [FAM]- ATGAGACTTTCGCATGAGAGAC-[BHQplus]
Av13	Av13F: AGTTTGATCCTGGCTCAGGATG
	Av13R: GAGGCAAGTTCCTTACGCGTT
	Av13P: [FAM]- AAGTTACCTTCGGGTAATGAGGAT-[BHQplus]

Sensitivity values of the qPCR assays are presented in percentages and are calculated as the fraction of actual positive [true positive (TP)] host samples divided by all expected positive hosts, including both false negative (FN) and TP, as follows: Sensitivity = TP/(FN + TP).

Specificity values are presented in percentages and are calculated as the fraction of actual negative [true negative (TN)] host samples divided by all expected negative hosts, including both false positive and true negative, as follows: Specificity = TN/(FP + TN).

RESULTS

The development of MST avian markers included HTS of bacterial 16S rRNA gene from various avian fecal sources. This was followed by two steps of bio-informatic filtering analysis: the first, identified both comparable and distinctive OTUs among the domestic and wild avian fecal sources. The second, compared the16S rRNA gene sequences of the selected OTUs to database of bacteria originating from other host fecal samples, in order to identify unique avian DNA sequences. These analyses were the basis for planning and design of qPCR assays, which were then evaluated to establish a battery of Avian MST assays.

OTU Analysis of Avian Feces

All in all, 10 phyla were detected across the tested sample sets, revealing different distribution profiles. Most OTUs were assigned to *Firmicutes, Proteobacteria*, and *Fusobacteria*, and the incidence of other phyla was lower. *Bacteroides* and *Actinobacteria* consist 9.5% of the reads and other phyla (*Verrucomicrobia, Tenericutes, Chloroflexi*, and *Cyanobacteria*, and *TM7*) representation was less than 1% each (**Figure 1**). The

difference of fecal microbial communities between various tested host species was demonstrated by dissimilar OTU affiliations, as well as dissimilar OTU frequencies. The poultry sample had a relatively homogenous taxonomical profile. Not only were 94% of sequence reads ascribed to the *Firmicutes* phylum, but 90% belong to the *Lactobacillaceae* at the family level (data not shown). 16S rRNA gene read sequences of the turkey sample illustrated a similar but not an identical pattern in which the *Firmicutes* phylum dominated, showing higher diversity at the family level. The bacterial community profiles of the two waterfowl samples overlapped at the identified phyla level but still differed in their relative quantitative representations (**Figure 1**). A high degree of variability was observed in levels of *Proteobacteria, Fusobacteria,* and *Firmicutes* phyla between avian samples. Waterfowl samples had a noticeably high incidence of the *Fusobacteria* phylum, which was almost absent in the domestic avian samples.

Avian Marker Design

Markers pipeline development included both bioinformatics analyses and laboratory assessment. A total of 59,863 sequence reads were analyzed using a *de novo* OTU generation approach according to the performed sequence pairwise comparisons. This approach resulted in 7205 OTUs, 57 of which were sample group specific. Next, *in silico* analysis for unique DNA was carried out using a dedicated server, http://mst.weizmann.ac.il. Accession numbers of sequences from which qPCR were designed are detailed in **Table 2**.

MST qPCR Assays Characterization

Six out of the nine qPCR assays met performance criteria as evaluated by triplicates of three independent calibration curves of plasmid construct dilutions (**Table 2**). qPCR assays demonstrated similar amplification efficiencies, range of quantification (ROQ) and limit of detection (LOD) defined as 90% positive detection of tested qPCR assays. Amplicon sequencing of Av13, Av24, Av43, Av163, and Av216, showed 100% identity with the corresponding HTS sequences whereas Av4143 had 99% identity.
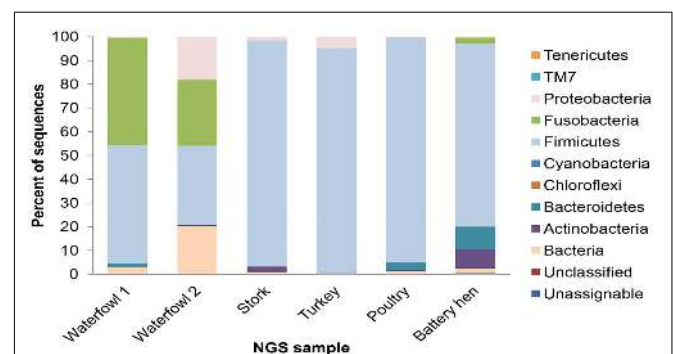

FIGURE 1 | Taxonomic distribution of associated bacteria of tested fecal avian samples: waterfowl, stork, turkey, poultry, and battery hen given in percents. Most OTUs were assigned to the phyla of *Firmicutes, Proteobacteria,* and *Fusobacteria*.

TABLE 2 | Performance features of qPCR assays.

Assay	Taxonomical affiliates	GenBank accession no.	Size (bp)	Slope	Intercept	qPCR efficiency	Range of quantification	Limit of detection (plasmid copies)[a]
Av4143	*Lactobacillus*	LN864462	244	3.75	44.05	84.7	25 – 2.5E8	6
Av163	*Gallibacterium*	LN864463	74	3.3	40.9	100	18 – 3.6E8	9
Av43	*Firmicutes*	LN864464	100	3.39	41.4	97	25 – 2.5E8	6
Av216	Other bacteria	LN864465	122	3.4	40.5	96.8	24 – 2.4E8	6
Av24	*Fusobacteriaceae*	LN864466	76	3.4	40.1	96.8	20 – 2.0E8	5
Av13	*Fusobacteriaceae*	LN864467	119	3.43	42.3	95.6	19 – 3.8E8	8

[a]*Defined as 90% positive detection of qPCR.*

Sensitivity and Specificity Assessment

Sensitivity and specificity values were calculated from qualitative analysis of qPCR results. Binary analysis, presence or absence of MST markers, disclosed a variety of marker distribution and incidence in avian hosts (**Table 3**). Av4143 was identified in all avian groups whereas Av163 was totally absent in turkey samples. qPCR profiles of poultry and battery hen samples were similar in that both were positive for Av43 and negative for the three markers Av216, Av24, and Av13. The waterfowl qPCR profile, on the other hand, was a mirror-image of the above, demonstrating no Av43 but showing Av216, Av24, and Av13. Av216, Av24, and Av13 were only detected in waterfowl, and their sensitivity values spanned from 47 to 76%.

Specificity values of qPCR assays were determined by testing non-target fecal and wastewater samples from various origins, including domestic animal and urban wastewater influent samples. The results (**Table 4**) demonstrated relatively high specificity values for most assays, except for the Av163 assay, which was detected in 37% of the bovine samples.

Fecal Avian MST Profiles

Quantification qPCR feature was employed to calculate marker copies in fecal samples followed by normalization to DNA mass. The Av4143 marker had the highest median value among MST markers, with a decreasing order between poultry, battery hen and waterfowl fecal samples, respectively (**Figure 2**). Its distribution among the poultry samples demonstrated a

TABLE 3 | Number of positive detection avian fecal sample in each MST assay (the numbers of total tested samples are specified in parenthesis).

MST marker	Battery hen (50)	Poultry (20)	Turkey (10)	Waterfowl (17)	Sensitivity value (%)
Av4143	48	18	10	14	95[a]
Av163	48	10	0	10	70[a]
Av43	46	18	1	1	91[b]
Av216	0	0	0	13	76[c]
Av24	0	0	0	9	52[c]
Av13	0	0	0	8	47[c]

Assay sensitivity value was calculated as the percent of true positive detection rate of target tested samples.
[a]*True positive samples are of avian origin.*
[b]*f True positive samples are of chicken origin.*
[c]*True positive samples are of waterfowl origin.*

restricted 25 and 75 percentile range compared to battery hen and waterfowl samples. Waterfowl profiles included five markers, except the Av43; battery hens and poultry displayed a complex of three markers.

DISCUSSION

Microbial source tracking notably employs the qPCR method targeting specific bacterial sequences as unique markers of their respective hosts. We aimed to develop additional, useful avian MST assays. A major obstacle was the challenge posed by an abundance of both wild and domesticated avian species, some of which migrate over highly heterogenic environments within relatively short periods of time.

Identification of hosts such as human, bovine, and several avian species, requires targeting of specific sequences, often using the 16S rRNA gene (Marti et al., 2011), by respective qPCR assays. The development of such MST molecular markers traditionally starts with building a fecal microbial library from the desired host. The number of clones within a given library varies greatly between reported studies. For example, the Crane MST marker targeting bacteria related to *C. marimammalium* was designed following analysis of 1,151 16S rRNA gene clones whereas the (Marti et al., 2011) muskrat marker included 62 clones.

The use of HTS for the development of avian MST may enable a comprehensive screening of the microbiome of various avian hosts. Nevertheless, our HTS was performed using samples collected from only six hosts in a single country, albeit from various habitats. This might limit one's ability to reach conclusions regarding flora from avian populations worldwide.

The taxonomical variability in young poultry and young turkeys was found to be lower than that isolated from battery hens or storks. Phyla identified in avian fecal samples have been previously mapped (Lu et al., 2009), demonstrating a dominance of the *Firmicutes*, *Proteobacteria*, and *Fusobacteria*. This finding was reaffirmed in the present study where *Firmicutes* had the highest occurrence. The *Bacteriodales* phylum was found in variable frequencies in the tested avian groups, ranging from 0.1% in waterfowl to 9.56% in battery hens. Members of the *Bacteriodales* phylum have not been consistently reported in avian gut and excreta; in some studies (Zhu et al., 2002) they were almost absent; in others (Scupham et al., 2008) they were identified. This variability can be attributed, at least in part, to
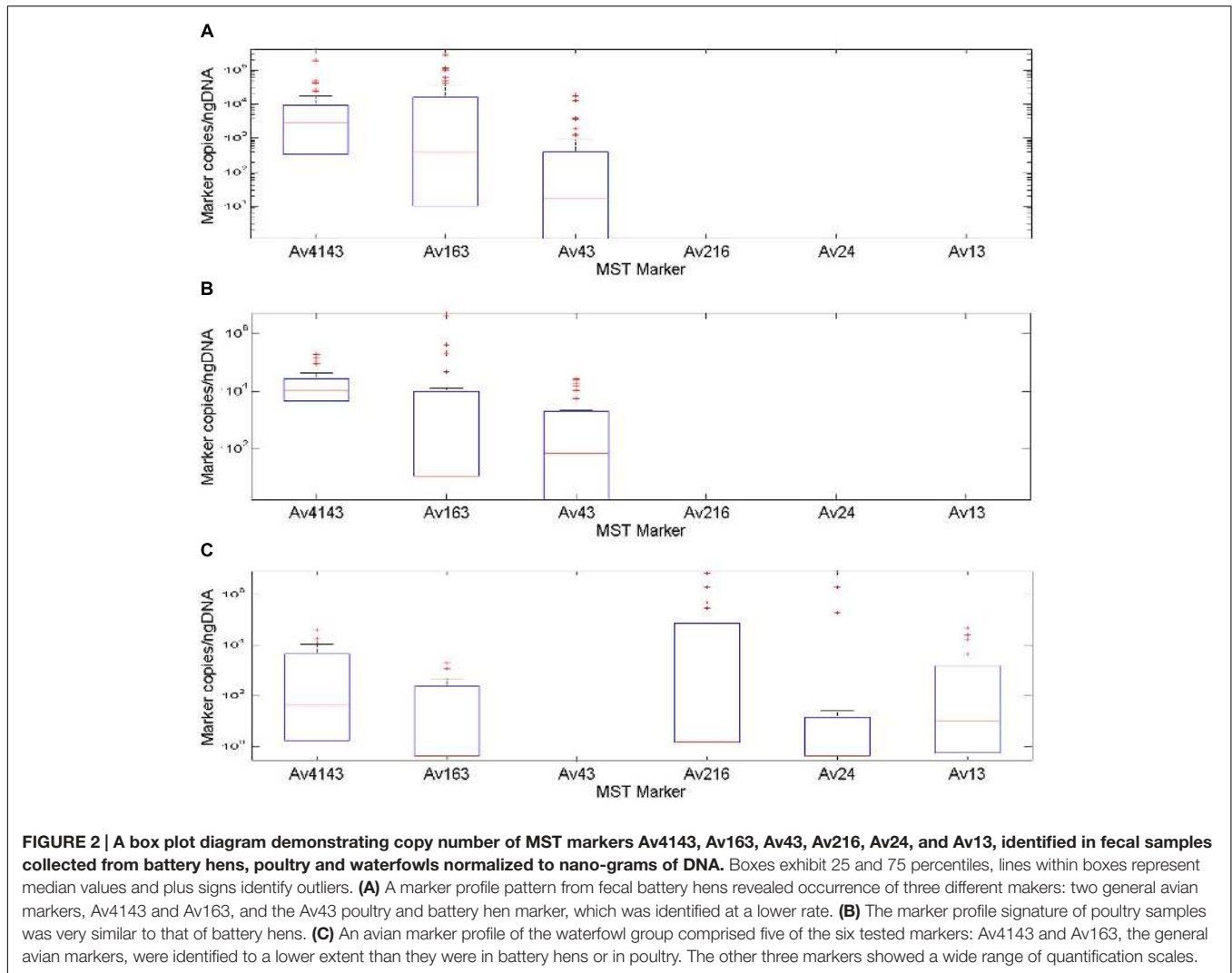
FIGURE 2 | A box plot diagram demonstrating copy number of MST markers Av4143, Av163, Av43, Av216, Av24, and Av13, identified in fecal samples collected from battery hens, poultry and waterfowls normalized to nano-grams of DNA. Boxes exhibit 25 and 75 percentiles, lines within boxes represent median values and plus signs identify outliers. **(A)** A marker profile pattern from fecal battery hens revealed occurrence of three different makers: two general avian markers, Av4143 and Av163, and the Av43 poultry and battery hen marker, which was identified at a lower rate. **(B)** The marker profile signature of poultry samples was very similar to that of battery hens. **(C)** An avian marker profile of the waterfowl group comprised five of the six tested markers: Av4143 and Av163, the general avian markers, were identified to a lower extent than they were in battery hens or in poultry. The other three markers showed a wide range of quantification scales.

disparities in study design and to the use of different PCR primer sets.

The HTS approach, which was critical in filtering specific MST sequences, included two stages of bioinformatic analysis. A high correlation was shown between the results of the bioinformatic analysis and the final qPCR confirmation results for some but not all of the assays. The three waterfowl assays targeting Av216, Av24, and Av13, were only found in fecal samples from water dwelling fowls, but not from domestic avian species; this was confirmed by qPCR testing. On the other hand, the two assays, Av4143 and Av163, were found to be qPCR positive in most tested samples; the former was only found in battery hens, and the latter was only found in poultry by QIIME software. These findings might raise concerns regarding the possibility

TABLE 4 | Number of positive qPCR detections in each avian assay, tested in non avian samples.

Assay	Human wastewater effluent (25)	Bovine (40)	Swine (20)	Horse (9)	Gazelle (10)	Specificity value (%)
Av4143	1	0	1	1	0	97
Av163	0	15	0	0	0	85
Av43	0	1	0	0	0	99
Av216	0	1	0	0	0	99
Av24	0	0	0	0	0	100
Av13	0	0	0	0	0	100

The numbers of total tested samples are specified in parenthesis. Specificity of qPCR assays calculated as percent of true negative detection rate of non-target samples.

of incomplete representation of specific taxa in the HTS experiment. HTS bias representation of bacterial communities (Zhang et al., 2011) is associated, in part, with library preparation based on 16S rRNA gene PCR amplification of heterogeneous sequences. The qPCR reaction, on the other hand, is designed to amplify specific bacterial community members, and is also characterized by a low LOD. The variations between HTS and qPCR results can be attributed to the differences between their respective abilities to detect low copy number sequences.

Simultaneous screening of multiple sequences at the HTS approach allowed identification of an array of assays for various avian groups. The sensitivity values ranged from 47 to 95%; the lower values ascribed primarily to the waterfowl assays and did not meet the 80% benchmark (Boehm et al., 2013). This demonstrates some of the limitations of studying a single geographic location, when testing only a limited sample size of waterfowl. A broader validation (Reischer et al., 2013) in different geographical locations will provide a more reliable sensitivity value, as previously reported.

Specificity values were all above 85% and indicated that the bioinformatic sorting process pinpoints unique determinants of specific bacteria species in avian feces.

Environmental contamination sources can be better understood using MST. The reliability and power of MST increase with the number of available assays which can be employed in evaluation of avian fecal bacteria. As a result, resolution and understanding are more attainable with simultaneous use of an array of markers.

AUTHOR CONTRIBUTIONS

Analysis using Qiime, OTU analysis, co-authorship, and critical manuscript revision: BD, VC-C; Putting together a process for seeking unique determinants and constructing a unique, 16S rRNA gene sequence database of fecal origin, followed by creation of free on-line access to this database, co-authorship and critical manuscript revision: SB-D, JP; conception and design of the work including all stages thereof, acquisition, analysis, and interpretation of all data, co-authoring the work, critically revising it, and approving the final submitted version: ER, SO, SB-D, JP, VC-C, BD, VK, YK; agreement to be accountable for all aspects of the work: BD, ER, JP, SB-D, SO, VC-C, VK, YK.

REFERENCES

Ahmed, W., and Katouli, M. (2008). Phenotypic variations of enterococci in surface waters: analysis of biochemical fingerprinting data from multi-catchments. *J. Appl. Microbiol.* 105, 452–458. doi: 10.1111/j.1365-2672.2008.03763.x

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Balleste, E., and Blanch, A. R. (2011). Bifidobacterial diversity and the development of new microbial source tracking indicators. *Appl. Environ. Microbiol.* 77, 3518–3525. doi: 10.1128/AEM.02198-10

Bernhard, A. E., and Field, K. G. (2000). A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides*-Prevotella genes encoding 16S rRNA. *Appl. Environ. Microbiol.* 66, 4571–4574. doi: 10.1128/AEM.66.10.4571-4574.2000

Boehm, A. B., Van De Werfhorst, L. C., Griffith, J. F., Holden, P. A., Jay, J. A., Shanks, O. C., et al. (2013). Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. *Water Res.* 47, 6812–6828. doi: 10.1016/j.watres.2012.12.046

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Clarridge, J. R. (2004). Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin. Microbiol. Rev.* 17, 840–862. doi: 10.1128/CMR.17.4.840-862.2004

Cole, J., Wang, Q., Fish, J., Chai, B., Mcgarrell, D., Sun, Y., et al. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, 633–642. doi: 10.1093/nar/gkt1244

Converse, R. R., Kinzelman, J. L., Sams, E. A., Hudgens, E., Dufour, A. P., Ryu, H., et al. (2012). Dramatic improvements in beach water quality following gull removal. *Environ. Sci. Technol.* 46, 10206–10213. doi: 10.1021/es302306b

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565

Gourmelon, M., Caprais, M. P., Le Mennec, C., Mieszkin, S., Ponthoreau, C., and Gendronneau, M. (2010). Application of library-independent microbial source tracking methods for identifying the sources of faecal contamination in coastal areas. *Water Sci. Technol.* 61, 1401–1409. doi: 10.2166/wst.2010.033

Green, H. C., Dick, L. K., Gilpin, B., Samadpour, M., and Field, K. G. (2012). Canada goose, duck, and chicken fecal contamination in water. *Appl. Environ. Microbiol.* 78, 503–510. doi: 10.1128/AEM.05734-11

Harwood, V. J., Brownell, M., Wang, S., Lepo, J., Ellender, R. D., Ajidahun, A., et al. (2009). Validation and field testing of library-independent microbial source tracking methods in the Gulf of Mexico. *Water Res.* 43, 4812–4819. doi: 10.1016/j.watres.2009.06.029

Haugland, R. A., Varma, M., Kelty, C. A., Sivaganesan, M., and Shanks, O. (2010). Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected Bacteroidales species and human fecal waste by real-time PCR. *Syst. Appl. Microbiol.* 33, 348–357. doi: 10.1016/j.syapm.2010.06.001

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., Mcginnis, S., and Madden, T. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Res.* 36, W5–W9. doi: 10.1093/nar/gkn201

Kuczynski, J., Stombaugh, J., Walters, W. A., Gonzales, A., Caporaso, J. G., and Knight, R. (2011). Using QIIME to Analyze 16S rRNA gene sequences from microbial communities. *Curr. Protoc. Bioinformatics* doi: 10.1002/0471250953.bi1007s36 Chap. 10, Unit 10.7.

Layton, A., Mckay, L., Williams, D., Garrett, V., Gentry, R., and Sayler, G. (2006). Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. *Appl. Environ. Microbiol.* 72, 4214–4224. doi: 10.1128/AEM.01036-05

Lu, J., Santo Domingo, J. W., Hill, S., and Edge, T. A. (2009). Microbial diversity and host-specific sequences of Canada goose feces. *Appl. Environ. Microbiol.* 75, 5919–5926. doi: 10.1128/AEM.00462-09

Marti, R., Gannon, V. P., Jokinen, C., Lanthier, M., Lapen, D. R., Neumann, N. F., et al. (2013). Quantitative multi-year elucidation of fecal sources of waterborne pathogen contamination in the South Nation River basin using bacteroidales microbial source tracking markers. *Water Res.* 47, 2315–2324. doi: 10.1016/j.watres.2013.02.009

Marti, R., Zhang, Y., Lapen, D. R., and Topp, E. (2011). Development and validation of a microbial source tracking marker for the detection of fecal pollution by muskrats. *J. Microbiol. Methods* 87, 82–88. doi: 10.1016/j.mimet.2011.07.013

McDonald, D., Price, M., Goodrich, J., Nawrocki, E., Desantis, T., Probst, A., et al. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* 6, 610–618. doi: 10.1038/ismej.2011.139

Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* 85, 2444–2448. doi: 10.1073/pnas.85.8.2444

Reischer, G. H., Ebdon, J. E., Bauer, J. M., Schuster, N., Ahmed, W., Astrom, J., et al. (2013). Performance Characteristics of qPCR assays targeting human- and ruminant-associated bacteroidetes for microbial source tracking across sixteen countries on six continents. *Environ. Sci. Technol.* 47, 8548–8556. doi: 10.1021/es304367t

Ryu, H., Elk, M., Khan, I. U., Harwood, V. J., Molina, M., Edge, T. A., et al. (2014). Comparison of two poultry litter qPCR assays targeting the 16S rRNA gene of *Brevibacterium* sp. *Water Res.* 1, 613–621. doi: 10.1016/j.watres.2013.10.015

Ryu, H., Lu, J., Vogel, J., Elk, M., Chavez-Ramirez, F., Ashbolt, N., et al. (2012). Development and evaluation of a quantitative PCR assay targeting sandhill crane (*Grus canadensis*) fecal pollution. *Appl. Environ. Microbiol.* 78, 4338–4345. doi: 10.1128/AEM.07923-11

Sandor, A. D., Marcutan, D. I., D'amico, G., Gherman, C. M., Dumitrache, M. O., and Mihalca, A. D. (2014). Do the ticks of birds at an important migratory hotspot reflect the seasonal dynamics of *Ixodes ricinus* at the migration initiation site? A case study in the Danube Delta. *PLoS ONE* 9:e89378. doi: 10.1371/journal.pone.0089378

Scupham, A. J., Patton, T. J., Bent, E., and Bayles, D. O. (2008). Comparison of the cecal microbiota of domestic and wild turkeys. *Microbiol. Ecol.* 56, 322–331. doi: 10.1007/s00248-007-9349-4

Shanks, O. C., Atikovic, E., Blackwood, A. D., Lu, J., Noble, R. T., Santo Domingo, J., et al. (2008). Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. *Appl. Environ. Microbiol.* 74, 745–752. doi: 10.1128/AEM.01843-07

Shanks, O. C., Kelty, C. A., Sivaganesan, M., Varma, M., and Haugland, R. A. (2009). Quantitative PCR for genetic markers of human fecal pollution. *Appl. Environ. Microbiol.* 75, 5507–5513. doi: 10.1128/AEM.00305-09

Shen, Z., Duan, C., Zhang, C., Carson, A., Xu, D., and Zheng, G. (2013). Using an intervening sequence of Faecalibacterium 16S rDNA to identify poultry feces. *Water Res.* 47, 6415–6422. doi: 10.1016/j.watres.2013.08.013

Sinigalliano, C. D., Ervin, J. S., Van De Werfhorst, L. C., Badgley, B. D., Balleste, E., Bartkowiak, J., et al. (2013). Multi-laboratory evaluations of the performance of *Catellicoccus marimammalium* PCR assays developed to target gull fecal sources. *Water Res.* 47, 6833–6896. doi: 10.1016/j.watres.2013.02.059

Stoeckel, D. M., and Harwood, V. J. (2007). Performance, design, and analysis in microbial source tracking studies. *Appl. Environ. Microbiol.* 73, 2405–2415. doi: 10.1128/AEM.02473-06

VerCauteren, K., Pilon, J., Nash, P., Phillips, G., and Fischer, J. (2012). Prion remains infectious after passage through digestive system of american crows (*Corvus brachyrhynchos*). *PLoS ONE* 7:e45774. doi: 10.1371/journal.pone.0045774

Wang, Q., Garrity, G., Tiedje, J., and Cole, J. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Weidhaas, J., and Lipscomb, E. (2013). A new method for tracking poultry litter in the Potomac Basin headwaters of West Virginia. *J. Appl. Microbiol.* 115, 445–454. doi: 10.1111/jam.12231

Xenoulis, P. G., Gray, P. L., Brightsmith, D., Hoppes, S., Steiner, J. M., Tizard, I., et al. (2010). Molecular characterization of the cloacal microbiota of wild and captive parrots. *Vet. Microbiol.* 146, 320–325. doi: 10.1016/j.vetmic.2010.05.024

Yampara-Iquise, H., Zheng, G., Jones, J. E., and Carson, C. A. (2008). Use of a *Bacteroides* thetaiotaomicron-specific α-1-6, mannanase quantitative PCR to detect human faecal pollution in water. *J. Appl. Microbiol.* 105, 1686–1693. doi: 10.1111/j.1365-2672.2008.03895.x

Zhang, H., Parameswaran, P., Badalamenti, J., Rittmann, B., and Krajmalink-Brown, R. (2011). Integrating high-throughput pyrosequencing and quantitative real-time PCR to analyze complex microbial communities. *Methods Mol. Biol.* 733, 107–128. doi: 10.1007/978-1-61779-089-8_8

Zhu, X. Y., Zhong, T., Pandya, Y., and Joerger, R. D. (2002). 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. *Appl. Environ. Microbiol.* 68, 124–137. doi: 10.1128/AEM.68.1.124-137.2002