

Showcasing research from Dr. Faizan Saleem, Dr. Herb Schellhorn and Dr. Tom Edge of the MacWater Research Initiative at McMaster University, Hamilton, Ontario, Canada ([www.macwater.org](http://www.macwater.org)).

eDNA metabarcoding-based source attribution of fecal indicator bacteria exceedances in urban freshwater beaches, sand and rivers

eDNA metabarcoding and microbial source tracking digital PCR methods were applied to identify fecal contamination sources at four urban Lake Ontario beaches. Combined use of MST and eDNA methods provided a more comprehensive characterization of potential fecal contamination sources, including diverse wildlife species, that can guide targeted beach-specific water monitoring and risk management strategies.

Image reproduced by permission of Faizan Saleem, Herb Schellhorn and Thomas Edge, *Environ. Sci.: Adv.*, 2025, **4**, 456.

As featured in:



See Thomas A. Edge *et al.*, *Environ. Sci.: Adv.*, 2025, **4**, 456.



Cite this: *Environ. Sci.: Adv.*, 2025, 4, 456

## eDNA metabarcoding-based source attribution of fecal indicator bacteria exceedances in urban freshwater beaches, sand and rivers†

Faizan Saleem, Jennifer L. Jiang, Enze Li, Kevin L. Tran, Herb E. Schellhorn and Thomas A. Edge \*

Freshwater beach quality is routinely tested by measuring fecal indicator bacteria, which can assess water quality but cannot identify sources of fecal contamination. We compared eDNA metabarcoding and microbial source tracking (MST) digital PCR methods to identify fecal contamination sources in water and sand at four urban Lake Ontario beaches and two nearby river mouth locations. eDNA sequences matched mammal, bird, and fish taxa known in the study area. Human eDNA sequences were prominent in all water and sand samples such that they had less value for discriminating between sewage occurrence at sites. Mallard duck, muskrat, beaver, raccoon, gull, robin, chicken, red fox, and cow eDNA sequences were common across all locations. Dog, Canada goose, and swan eDNA sequences were more common in Toronto beach waters, suggesting localized sources. MST results were generally consistent with eDNA, such as finding the Gull4 DNA marker and the human mitochondrial DNA marker in most water and sand samples. Chicken, cow, and dog eDNA sequences and the human bacterial MST DNA marker often showed a higher frequency of occurrence on Beach Action Value (BAV) exceedance days. The surprisingly widespread detection of chicken and cow eDNA sequences was likely from incompletely digested human food, raising caution for interpreting eDNA results related to food animals in sewage-contaminated urban settings. Combined use of MST and eDNA methods provided a more comprehensive characterization of potential fecal contamination sources, including diverse wildlife species at the human–animal One Health interface, that can guide targeted beach-specific water monitoring and risk management strategies.

Received 25th June 2024  
Accepted 17th January 2025

DOI: 10.1039/d4va00221k

rsc.li/esadvances

### Environmental significance

Freshwater beaches are tested by measuring fecal indicator bacteria which assess water quality, but cannot identify sources of fecal contamination. Fecal source tracking is needed to better assess human health risks and correctly target remedial actions. It usually involves performing multiple PCR assays for host-specific gut microbes from humans and domestic animals (*e.g.* cow, dog). Our study applied a broader mammal and bird eDNA metabarcoding approach to complement microbial PCR assays. eDNA results detected numerous wildlife species likely contributing to fecal pollution in our urban settings. We also identified incompletely digested food (*e.g.* chicken, cow) likely carried over into human sewage requiring further investigation. Use of both microbial and eDNA metabarcoding source tracking techniques provided more comprehensive fecal pollution profiling.

## Introduction

Fecal contamination is one of the main causes of the deterioration of water quality and is a concern for the long-term sustainability of recreational water ecosystems.<sup>1</sup> Traditional routine beach monitoring strategies rely on fecal indicator bacteria levels for water quality assessment. However, controlling the impact of fecal contamination on the sustainability of

recreational waters requires comprehensive identification of fecal sources and such information cannot be obtained from fecal indicator bacteria levels alone.<sup>2–4</sup> Fecal source tracking to date has been mainly based on a microbial source tracking approach targeting microorganisms specific to a host organisms' gut, including humans, birds, and mammals.<sup>5–7</sup> Several microbial fecal source tracking markers, like the human *Bacteroides* HF183 marker, have been extensively tested.<sup>8</sup> However, there are still relatively few well-validated host-specific microbial DNA markers for animals, particularly for non-domestic animals. Mitochondrial sequences from human, mammal, and avian cells in environmental DNA (eDNA) offer another approach to expand the toolbox for detecting fecal

Department of Biology, McMaster University, 1280 Main St W., Hamilton, Ontario, L8S 4L8, Canada. E-mail: edget2@mcmaster.ca; Tel: +1-4168941129

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4va00221k>



contamination sources.<sup>9</sup> These eukaryotic host cells are continuously sloughed off in large numbers from the lining of human, mammal, and bird gastrointestinal tracts into fecal matter that is then shed into the environment.<sup>10</sup>

DNA derived from mammalian or avian cells in water samples can provide information on the taxonomic diversity of potential fecal contamination sources in aquatic ecosystems.<sup>9</sup> Common methods for detecting human or animal host sequences in eDNA have employed species-specific primers and quantitative PCR<sup>10</sup> or digital PCR.<sup>11</sup> However, fecal contamination in environmental waters can be from diverse animal sources, and such PCR-based approaches have typically targeted one marker at a time and are unable to provide a comprehensive diversity profile of potential fecal sources. In contrast, next generation sequencing-based metabarcoding facilitates the comprehensive characterization of universal marker genes (such as mitochondrial 16S rRNA gene) in environmental DNA. This combination of mammal and avian eDNA metabarcoding, along with growing numbers of host-specific microbial source tracking DNA markers, offers the potential to comprehensively identify potential fecal contamination sources for recreational waters.<sup>12,13</sup>

In this study, we compared eDNA metabarcoding and microbial source tracking DNA markers for more comprehensive profiling of potential fecal contamination sources in water and sand at four urban Lake Ontario beaches and two nearby river mouth locations. The questions we addressed were: (1) What are the potential fecal contamination sources for our four urban freshwater beaches? (2) How different are fecal contamination sources as determined by eDNA metabarcoding and microbial source tracking approaches? (3) What are the different fecal contamination sources between urban beach waters, beach sands, and rivers? And (4) Which fecal

contamination sources are most associated with Beach Action Value (BAV) exceedances at beaches?

## Materials and methods

### Study plan, sample collection and fecal indicator bacteria data

We collected water samples over the bathing season from four freshwater beaches (two urban beaches in the City of Toronto and two urban beaches in the City of St Catherines on Lake Ontario), as well as from beach sands at the St Catherines beaches and two rivers adjacent to the two Toronto beaches. For Toronto sampling sites, water samples were collected on each sampling day from Marie Curtis Park East Beach, Sunnyside Beach, Humber River, and Etobicoke Creek. For St Catherines sampling sites, water and sand pore water samples were collected on each sampling day from Sunset Beach and Lakeside Beach. Fig. 1 indicates the geographical locations of each sampling location from beaches and rivers. There are municipal wastewater treatment plants within one kilometer of each study area that discharge treated effluent in the general vicinity of all four beach sites. These wastewater treatment plants (WWTPs) provide for at least secondary treatment, and include the Humber WWTP (near Sunnyside Beach), Lakeview WWTP (near Marie Curtis Beach), Port Dalhousie WWTP (near Lakeside Beach) and Port Weller WWTP (near Sunset Beach).

For water sample collection from Toronto beaches and rivers, samples were collected three times a week for the 2021 summer season between June 01 and August 26 (total = 309 samples). On each sampling day, eight samples were collected across Marie Curtis Park East Beach (3 samples: transects 30W, 30W replicate, and 32W), Sunnyside Beach (3 samples: transects 18W, 21W, and 21W replicate), Etobicoke Creek (1 sample), and Humber River (1 sample). Samples from St Catherines beaches

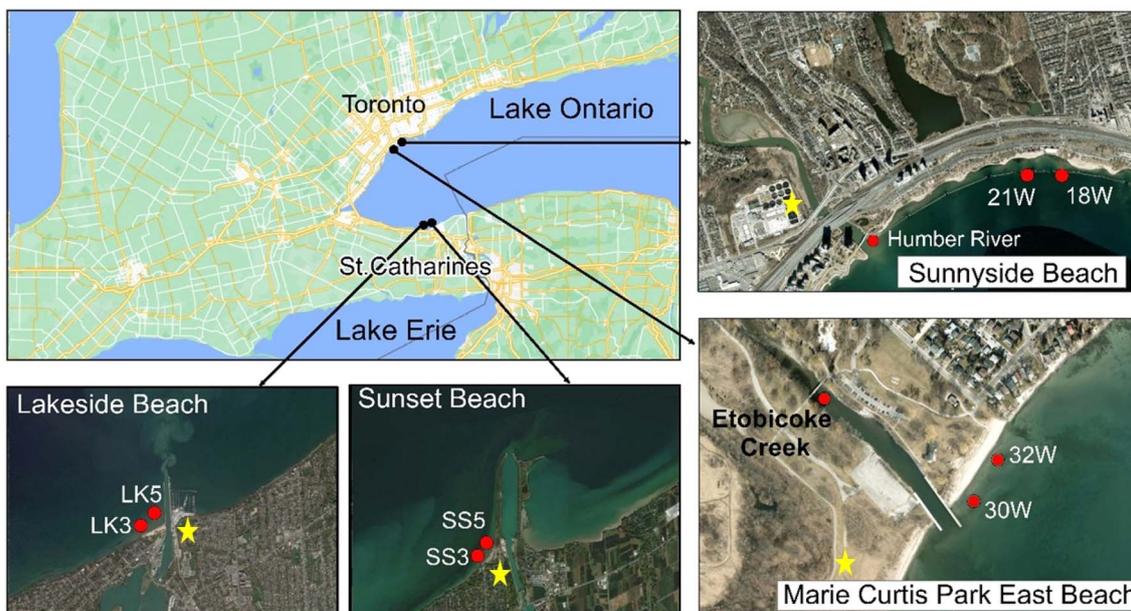


Fig. 1 Geographical locations of Toronto and St Catherines (Niagara) sampling sites. ★ symbols mark the approximate location of wastewater treatment plants.



were collected three times a week in the 2022 summer season between May 31 and September 06 (total = 336 water and sand pore water samples). Eight samples were collected for each sampling day, including 2 water samples and 2 sand pore water samples at Lakeside Beach (transects LK3 and LK5 sampling sites) and 2 water and 2 sand pore water samples at Sunset Beach (transects SS3 and SS5 sampling sites). Sand pore water samples were obtained by digging to collect groundwater in the beach swash zone within about 1 meter of the lake. Samples were collected between 5:30 and 7 am for the 2021 sampling season (Toronto) and 8:00 and 9 am for the 2022 sampling season. Grab samples were collected in 500 ml sterile polyethylene terephthalate (PET) bottles and were delivered to the lab on ice within 1 hour of sample collection. After delivery to the lab, samples were processed for *E. coli* by membrane filtration and *Enterococcus* by qPCR within a few hours. Fecal indicator bacteria data and categorizations of water samples in relation to Beach Action Values (BAV) were taken from our previously published work.<sup>14,15</sup>

### DNA extraction and sample selection/pooling for metabarcoding

For eDNA extraction, 300 ml of sample was filtered through a 0.22-micron nitrocellulose membrane filter (Millipore Corp., Bedford, MA). Filters were then processed for DNA extraction using the Norgen Soil Plus DNA Extraction kit (Norgen Biotek Corp., Canada) with minor modifications. We increased the bead beating time and changed the glass beads to zirconium beads to improve cell disruption. All the other steps were performed using the manufacturer's protocol. Filters and DNA extracts were stored at  $-80^{\circ}\text{C}$  prior to analysis. For digital PCR and DNA sequencing (eDNA metabarcoding), a subset of 48 samples were selected from across beach water, beach sand, and river sampling locations (ESI Table 1†). Sample selection first categorized samples as beach action value exceedance or non-exceedance samples based on fecal indicator levels (*E. coli* and *Enterococcus*). Then, for each beach location, 3 BAV-exceedance and 3 non-exceedance samples were selected, except for Lakeside Beach, for which we selected 2 samples each (due to only two BAV exceedance days). The 3 BAV-exceedance and 3 non-exceedance samples were selected for each river location according to the same sampling dates for their adjacent beaches. For beach sand locations at Lakeside and Sunset Beaches, the BAV-exceedance and non-exceedance samples were selected, corresponding to the same sampling dates selected for their adjacent beach water samples. DNA corresponding to sampling sites from a single location and for a single sampling day was pooled in equivalent volumes as a single sample for library preparation and DNA sequencing.

### eDNA metabarcoding sequencing

Mitochondrial 16S PCR was performed in two parts following Ragot and Villemur (2022):<sup>13</sup> (1) amplification of  $\sim 400$  bp fragment with PCR cycles limited to 10 to attempt to reduce PCR amplification bias due to dominant taxa, and (2) nested PCR using Illumina linker-attached primers and PCR product from

first PCR with 35 cycles. For the first PCR, each PCR reaction (25  $\mu\text{L}$ ) comprised of 12.5  $\mu\text{L}$  Hot Start PCR master mix (Thermo-scientific Inc., USA), 1.0  $\mu\text{L}$  of forward and reverse primers (10  $\mu\text{M}$ ), 2.0  $\mu\text{L}$  DNA, and 8.5  $\mu\text{L}$  of nuclease-free water. The PCR protocol for the first PCR included initial denaturation at  $95^{\circ}\text{C}$  for 10 min, 10 cycles of  $95^{\circ}\text{C}$  for 30 s,  $58^{\circ}\text{C}$  for 1 min, and  $72^{\circ}\text{C}$  for 40 s, followed by final extension at  $72^{\circ}\text{C}$  for 5 min. The reaction composition and PCR protocol remained the same for the nested PCR, except the PCR product from the first PCR was used as the DNA template, and the PCR protocol included touchdown PCR for annealing ( $69$  to  $59^{\circ}\text{C}$  for 10 cycles) to improve the annealing efficiency before the cycling steps. After amplicon PCR, amplicons were purified using Ampure XP magnetic beads (Beckman Coulter, California, USA). For index PCR (attachment of unique DNA barcodes to each sample's amplicons), we designed 7 P5 and 8 P7 primers, corresponding to 56 unique barcode combinations and ordered as Ultramers from IdtDNA (Coralville, USA). For the index PCR, each PCR reaction (25  $\mu\text{L}$ ) comprised of 12.5  $\mu\text{L}$  Hot Start PCR master mix (Thermo-scientific Inc, USA), 2.0  $\mu\text{L}$  of P5 and P7 indexing primers (5  $\mu\text{M}$ ), 5.0  $\mu\text{L}$  of purified PCR product, and 6.5  $\mu\text{L}$  of Nuclease-free water. The index PCR protocol included initial denaturation at  $95^{\circ}\text{C}$  for 3 min, 8 cycles of  $95^{\circ}\text{C}$  for 30 s,  $55^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 30 s, followed by final extension at  $72^{\circ}\text{C}$  for 5 min. Indexed PCR amplicons were purified using Ampure XP magnetic beads (Beckman Coulter, California, USA), followed by normalization to 4 nm for each sample and pooling. All the PCR reactions were performed on Bio-Rad CFX96 Touch Real-Time PCR (Bio-Rad Inc. USA). Paired-end sequencing ( $2 \times 250$ , V2 kit) was performed on the Illumina Mi-Seq sequencing platform at Farncombe Sequencing Institute (McMaster University).

### Digital PCR for microbial fecal source tracking markers

Assays for digital PCR (dPCR) used previously published primer and probe sets for the human HF183/BacR287 marker,<sup>16</sup> the seagull Gull4 marker,<sup>17</sup> the human mitochondrial DNA marker,<sup>10</sup> the Canada goose mitochondrial DNA marker,<sup>18</sup> and the dog Dog3 marker.<sup>19</sup> The ruminant Rum2Bac DNA marker<sup>20</sup> and the swine Pig2Bac DNA marker<sup>21</sup> were run only on samples with cattle and pig sequences detected from eDNA analysis. Additional description of the dPCR method and sequences of primers and probes are shown in ESI.† Digital PCR reactions were conducted in singleton format as a duplex assay for HF183 and Gull4 markers and for the human mitochondrial and goose mitochondrial markers, while other markers were run in a singleplex format. Comparative assay testing in singleplex and duplex formats has not indicated any interference between marker assays in each duplex. Yield and purity of DNA extracts were assessed by spectrophotometer (NanoDrop Lite, Thermo-Fisher) prior to dPCR analyses. Each duplex dPCR reaction consisted of 1  $\mu\text{L}$  nuclease-free water, 0.75  $\mu\text{L}$  of each 900 nM forward and reverse primer, 0.75  $\mu\text{L}$  of each 250 nM probe, 7.5  $\mu\text{L}$  of QuantStudio™ 3D Digital PCR Master Mix v.2 (Thermo-Fisher) and 2.0  $\mu\text{L}$  of extracted DNA template. Each singleplex dPCR reaction consisted of 1.95  $\mu\text{L}$  nuclease-free water, 1.4  $\mu\text{L}$



of each 900 nM forward and reverse primer, 0.75  $\mu\text{L}$  of 250 nM probe, 7.5  $\mu\text{L}$  of QuantStudio™ 3D Digital PCR Master Mix v.2 (ThermoFisher) and 2.0  $\mu\text{L}$  of extracted DNA template. Reactions were loaded onto a 20 000 micro-well chip (QuantStudio™ 3D Digital PCR 20 K Chip v2 with partition volume = 755 pL) using a QuantStudio™ 3D Chip Loader (ThermoFisher), and PCR assays were carried out using the ProFlex PCR System (ThermoFisher). Positive controls and a no-template negative control was run for each batch of dPCR chips analyzed. Thermocycler settings were 96 °C for 10 min, followed by 40 cycles of 60 °C for 2 min and 98 °C for 30 s, then 60 °C for 2 min. Chips were then read using a QuantStudio™ 3D Digital PCR Instrument (ThermoFisher). Results were analyzed using the QuantStudio™ AnalysisSuite™, which automatically determined threshold fluorescence values for the ROX reference dye to identify the number of qualified PCR well partitions, as well as for the FAM and VIC dye signals to identify positive reactions for the DNA markers (all thresholds were manually checked and occasionally corrected for accuracy). The AnalysisSuite™ software applies a Poisson Plus modelling technique to determine concentrations of each target within the sample, and results were reported as DNA copies per 100 ml.

All no-template PCR controls were negative, indicating a lack of contamination to compromise dPCR assays. The number of wells analyzed on dPCR chips was typically between 16 000 to 18 000 wells. The sensitivity and specificity of the dPCR assays were outlined in Edge *et al.* (2021),<sup>11</sup> and 2–3 additional samples of gull feces, dog feces, cow feces, pig feces, sewage influent, as well as DNA standards prepared by the U.S. National Institute of Standards and Technology (SRM 2917, Plasmid DNA for Fecal Indicator Detection and Identification) for HF183, Dog3, Rum2Bac, and Pig2Bac markers<sup>22</sup> were used as positive reference materials for dPCR runs. At least 4 positive wells (clearly cluster separated from negative wells) were set as a detection threshold for a DNA marker in dPCR assays. This threshold provided a clear basis for discriminating water sample dPCR results from all dPCR results for filter and DNA extraction blanks and no-template PCR control samples. The threshold was equivalent to a detection limit of about 14 DNA copies per 100 ml.

### Bioinformatics and data analysis

For each sample, ~200 000 paired-end sequences were obtained following eDNA sequencing. Data quality was checked using FastQC,<sup>23</sup> and sequences below the quality threshold 30 were removed using the Fastp FastQ preprocessing tool.<sup>24</sup> After initial quality control, the APSCALE pipeline<sup>25</sup> for metabarcoding analysis was used for downstream bioinformatics. Bioinformatics pipeline steps included: (1) length filtration (minimum length = 70, maximum length = 400 bp), illumina barcodes removal and primer sequences trimming were performed using Cutadapt,<sup>26</sup> (2) paired-end merging (Maxdiffpct = 25, maxdiffs = 199, and minovlen = 5), and Denoising (Alpha = 2, and minsize = 8) were performed using VSEARCH<sup>27</sup> to identify Exact Seq., and (3) LULU algorithm<sup>28</sup> was used for post-clustering curation to remove erroneous Exact Sequence

Variants (ESVs). Exact Sequence Variants were taxonomically annotated using NCBI BLASTn (percentage identity >95% and alignment length/query coverage % $\geq$ 80%) against the RefSeq nucleotide database. To normalize the datasets and remove sequencing artifacts, ESVs or taxonomic groups lower than 0.003% in abundance were removed from the datasets, as described previously for mt metabarcoding datasets.<sup>29</sup> Statistical comparison between groups was performed using Welch's *t*-test with a *p*-value cutoff  $\leq$ 0.05,<sup>30</sup> while Principle Component Analysis (PCA) was performed using Euclidean distances among samples on GGPlot and GGPUBR packages in R.<sup>31</sup>

## Results

### Quality control analyses

An initial trial was conducted to test for the validity of mitochondrial 16S rRNA gene fragment amplification (ESI Fig. 1†). All the tested samples from rivers and beach sampling sites amplified the gene fragment in the 300–400 bp range, similar to the positive control (Salmon sperm DNA). Although a non-specific gene fragment at ~800 bp was detected, it was removed during the gene library purification steps using magnetic bead size selection. After the Next-generation sequencing, more than ~200 000 DNA reads ( $22.3 \pm 2.2 \times 10^4$ ) were generated for each sample (ESI Table 2†). About 95–97% of the DNA reads passed the quality control parameters after primer trimming and removal of lower-quality (quality score <30) sequences. In total, 369 Exact Sequence Variants (ESVs) were identified, out of which 200 were processed for sequence annotation after removing potentially erroneous ESVs using LULU filtering, which identifies errors by searching for co-occurrence patterns and sequence similarity measures.<sup>28</sup>

### Taxonomic annotation and characterization of eDNA sequences

Across all 48 samples (200 ESVs), mitochondrial DNA metabarcoding identified 61 taxa, including fish, birds, mammals and arthropods (ESI Fig. 2†). The fish, mammal and bird taxa identified were consistent with those known from our study area, and lists of species for each sampling site are provided in ESI Tables 3 and 4.† In a few cases, we corrected taxonomic assignments to more generic identifications (European herring gull = gull; Northern house martin = martin; Eurasian tree sparrow = sparrow). ESI Fig. 2† shows the distribution of ESVs according to the mammal, bird, and fish categories in all samples. Mitochondrial DNA sequences associated with fish were relatively most abundant (55–78%), followed by mammals (20–43%) and birds (2–5%). While not the focus of this study, the most predominant fish species detected by eDNA sequences in every sample was the round goby (*Neogobius melanostomus*), consistent with a previous study on the Great Lakes.<sup>32</sup> However, alewife (*Alosa pseudoharengus*), green sunfish (*Lepomis cyanellus*), longnose dace (*Rhinichthys cataractae*), central stone-roller (*Camptostoma anomalum*), blacknosed dace (*Rhinichthys atratulus*), rock bass (*Ambloplites rupestris*), and bluntnose minnow (*Pimephales notatus*) were also widely prevalent.



Principle component analysis was performed using organism types (fishes, mammals, and birds) to assess the divergence patterns among the sampling sites. Similarity matrix analysis for fish and mammals sp. (ESI Fig. 3 and 4†) revealed two distinct but closely associated clusters. Samples from Toronto beaches and their adjacent rivers clustered together, while Niagara beach water and their adjacent beach sand samples aggregated into a separate cluster. Interestingly, compared to beach samples that grouped together, samples from river sources were more broadly distributed, indicating a higher diversity of fish and mammal species sequences in rivers. However, for bird species (ESI Fig. 5†), samples from both beach locations (Toronto and Niagara) were closely associated in a single cluster.

### eDNA sequence detection frequency of potential fecal contamination sources

Mitochondrial DNA sequences resolved and taxonomically assigned to the most common mammalian and bird fecal contamination sources are shown in ESI Fig. 6.† All 48 samples were positive for human ESVs, followed by *Castor canadensis* (Beaver: 97%), *Ondatra zibethicus* (Muskrat: 97%), *Anas platyrhynchos* (Mallard duck: 91%), *Larus* species (Gull: 87%), *Procyon lotor* (Raccoon: 87%), *Turdus migratorius* (Robin: 77%), *Gallus gallus* (Chicken: 73%), *Vulpes vulpes* (Fox: 50%), *Bos taurus* (Cow: 48%), *Canis lupis familiaris* (Dog: 33%), and *Branta canadensis* (Canada goose: 27%). Multiple ESVs were associated with most potential fecal sources, notably: human (7), squirrel (7), swan (7), robin (6), and 3 for beaver, dog, opossum, cow, red fox, Canada goose gull, and pigeon.

ESI Fig. 7† shows human eDNA sequences were predominant in all the samples (87–96%), followed by beaver and

muskrat (3–8%), mallard (2–5%), gull (2–5%), raccoon (1–5%), chicken (1–2%), and others (<1%). Fig. 2 and 3 show the percentage detection of common mammalian and avian eDNA sequences across all beach and river study locations. Among the mammalian sequences, human (100%), muskrat/beaver (92–100%) and raccoon (75–90%) eDNA sequences showed the highest detection frequency across all four sample types. Dog, vole and cat eDNA sequences were more frequently detected in Toronto beach water and river samples than in Niagara beach water and adjacent sand samples. Among avian sequences, mallard (100%), gull (82–100%), robin (68–100%), and chicken (70–82%) showed the highest detection frequency across all sample types. Goose, swan, martin, and pigeon eDNA sequences were more frequently detected for Toronto beaches and river samples than Niagara beach water and adjacent sand samples. Goose and swan eDNA sequences were more common at Toronto beaches than rivers.

### eDNA sequence differential abundance between the sample types/locations

Several species eDNA sequences were significantly more abundant for beaver ( $p = 4.7 \times 10^{-5}$ ), mallard ( $p = 1.0 \times 10^{-3}$ ), and chicken ( $p = 0.04$ ) at Toronto locations (ESI Fig. 8–10†). A comparison between Toronto beach and river samples identified that human ( $p = 0.08$ ) and beaver ( $p = 0.09$ ) eDNA sequences were relatively more abundant in river samples, but the observed difference was not statistically significant. Raccoon ( $p = 0.12$ ), gull ( $p = 0.15$ ), and chicken sequences were relatively more abundant in Toronto beach samples (ESI Fig. 11–15), though only chicken sequences were statistically significant ( $p = 0.004$ ). No significant differences were observed between Niagara beach water and Niagara beach sand samples.

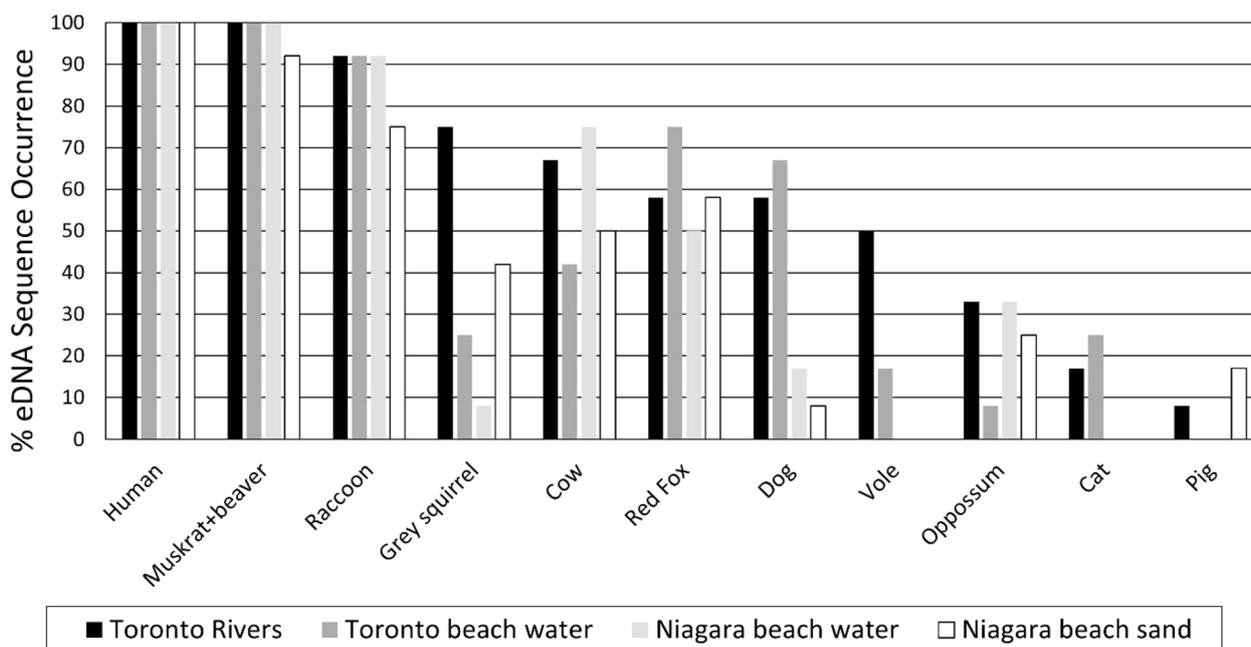


Fig. 2 Frequency of detection of most common mammalian eDNA sequences for beach, river, and sand samples.



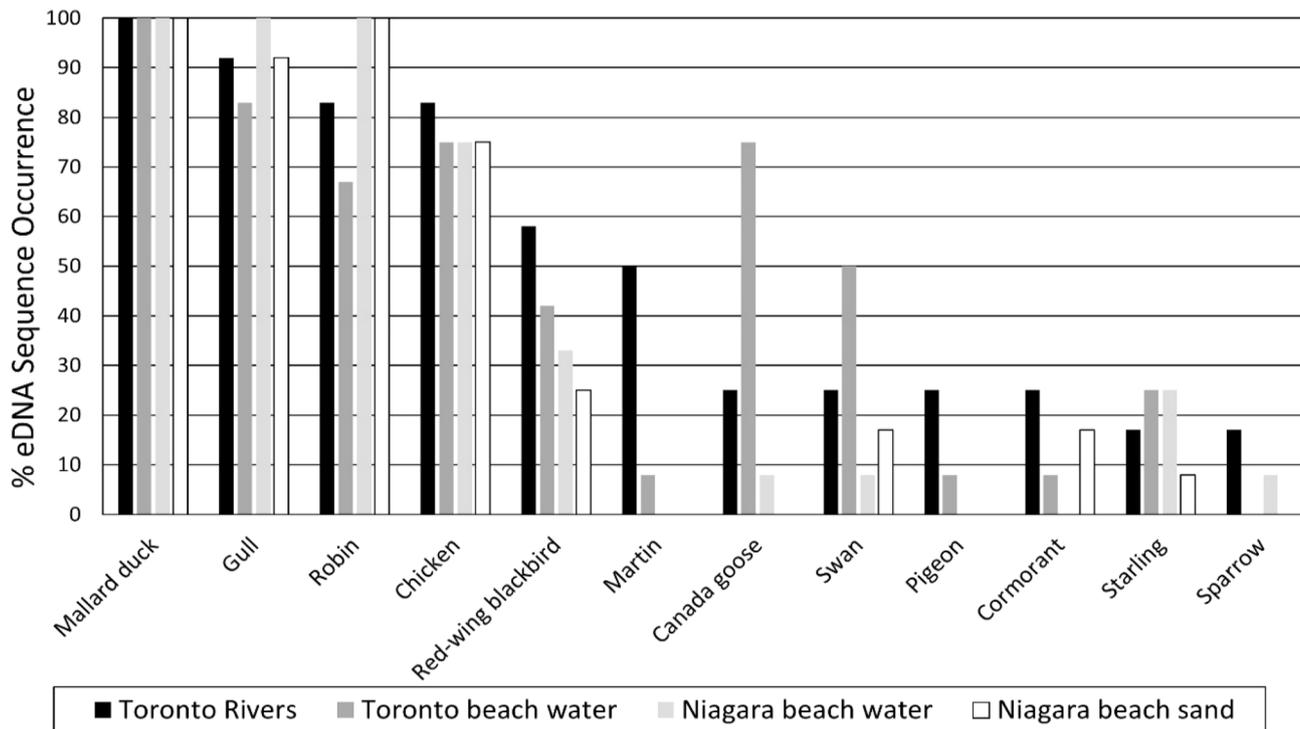


Fig. 3 Frequency of detection of most common avian eDNA sequences for beach, river, and sand samples.

#### eDNA detection frequency in association to fecal indicator beach action value exceedances

The occurrence of human, mammal, or bird eDNA sequences associated with BAV exceedance water quality conditions was determined (Table 1). Chicken and dog eDNA sequences were notably more commonly detected at Toronto beaches during BAV exceedances, while cow and chicken eDNA sequences were more common in Toronto rivers during BAV exceedances. Only fox eDNA sequences were notably more commonly detected at Niagara beaches during BAV exceedances. Comparison between beach action value exceedance and non-exceedance beach water samples identified that beaver, muskrat, mallard and chicken eDNA sequences were relatively more abundant in BAV exceedance samples (ESI Fig. 16–19†). However, only chicken ( $p = 0.02$ ) were significant. Specifically, the higher abundance of chicken eDNA sequences ( $p = 0.03$ ) in BAV exceedance samples was mainly associated with Toronto beaches (ESI Fig. 20).

#### Digital PCR detection frequency of potential fecal contamination sources and beach action value exceedances

Four relatively common microbial source tracking markers for human, mammalian and avian fecal pollution sources were tested for the same 48 samples using a digital PCR method to compare with results from eDNA metabarcoding (Fig. 4). Human mt (68–100%) and Gull4 (75–100%) markers were detected in the majority of the samples across our study locations, followed by human HF183 (9–82%) and Dog3 (22–42%). The frequency of the Gull4 marker was higher for beach water (100%) than for rivers (73%) and beach sand (82%) samples

(Table 2). The highest concentrations of the Gull4 marker were at Lakeside Beach (5300 DNA copies per 100 ml), Marie Curtis Beach (2505 DNA copies per 100 ml), and in the sand at Sunset Beach (1714 DNA copies per 100 ml).

Interestingly, the human HF183 marker was only detected on BAV-exceedance beach days for Toronto beach sites. In contrast, HF183 was detected for both BAV exceedance and non-exceedance beach days for Niagara beach sites. Overall, the HF183 marker was detected more frequently for Niagara beach water (65%) than Toronto beach water (42%), however, the frequency of detection of HF183 in association with BAV exceedances was higher for Toronto beach water (83%) than Niagara beach water (50%). The highest concentrations of HF183 measured were at Sunnyside Beach (10 673 DNA copies per 100 ml) and the Humber River (1832 DNA copies per 100 ml) on a BAV exceedance day. These concentrations are lower than those previously found in other sewage-impacted Toronto waters where HF183 could exceed  $10^5$  to  $10^6$  DNA copies per 100 ml in the Don River and at CSO outfalls respectively (Edge *et al.* 2021). Interestingly, the HF183 DNA marker only occurred in 8% of Niagara beach sand samples, in contrast to 100% sand detections of the human mt DNA marker and human eDNA sequences. The highest concentrations of the human mt DNA marker were measured on BAV exceedance days in Sunset Beach (1660 DNA copies/100 ml) and at Marie Curtis Beach (753 DNA copies/100 ml).

Although the overall detection frequency of the Dog3 marker was higher for Niagara beach water (42%) than Toronto beach water (25%), the frequency of occurrence of the Dog3 marker association with BAV exceedances was higher for Sunnyside



**Table 1** Frequency of occurrence of human, mammal, and bird eDNA sequences as potential fecal contamination sources for the beaches, rivers and sand on BAV exceedance and non-exceedance days

Sampling locations	Potential fecal sources	Frequency of sequence detection (BAV exceedance)	Frequency of sequence detection (BAV non-exceedance)
Toronto beaches	Human	100%	100%
	Beaver	100%	100%
	Mallard	100%	100%
	Chicken	100%	50%
	Gull	67%	100%
	Raccoon	100%	83%
	Cow	33%	50%
	Muskrat	83%	100%
	Robin	50%	67%
	Fox	67%	50%
	Dog	83%	50%
	Goose	67%	67%
	Toronto rivers	Human	100%
Beaver		100%	100%
Mallard		100%	100%
Chicken		100%	67%
Gull		83%	100%
Raccoon		100%	83%
Cow		83%	33%
Muskrat		100%	100%
Robin		67%	100%
Fox		33%	50%
Dog		50%	50%
Goose		33%	17%
Niagara beaches		Human	100%
	Beaver	100%	100%
	Mallard	83%	67%
	Chicken	67%	50%
	Gull	83%	83%
	Raccoon	83%	100%
	Cow	50%	50%
	Muskrat	100%	100%
	Robin	83%	83%
	Fox	67%	17%
	Dog	17%	0%
	Goose	0%	17%
	Niagara beach sand	Human	100%
Beaver		100%	83%
Mallard		100%	83%
Chicken		67%	50%
Gull		83%	100%
Raccoon		67%	83%
Cow		67%	17%
Muskrat		100%	83%
Robin		67%	100%
Fox		50%	67%
Dog		17%	0%
Goose			Not detected

Beach water (100%) than other beach waters (Table 2). All the BAV exceedance samples positive for the Dog3 marker in Toronto were from Sunnyside Beach. Interestingly, the association of the Dog3 marker with BAV exceedances was higher for Toronto beach water than its adjacent rivers. While the Canada goose mt DNA marker was never detected by dPCR, Canada goose eDNA sequences were found in 67% of Toronto beach water samples, 25% of Toronto river samples, and 8% of Niagara beach water samples.

The ruminant fecal bacterial Rum2Bac DNA marker was tested on 16 of the cow eDNA +ve samples, and the pig fecal bacterial Pig2Bac DNA marker was tested on the 3 pig eDNA +ve samples. These two DNA markers for host-specific fecal bacteria were not detected in these samples. A comparison of the detection rate of food animal eDNA sequences (chicken, cow, pig) with human HF183 and mt dPCR markers and human eDNA sequences is provided in ESI Table 5.† Chicken, cow, and pig eDNA sequences were almost always detected in Toronto



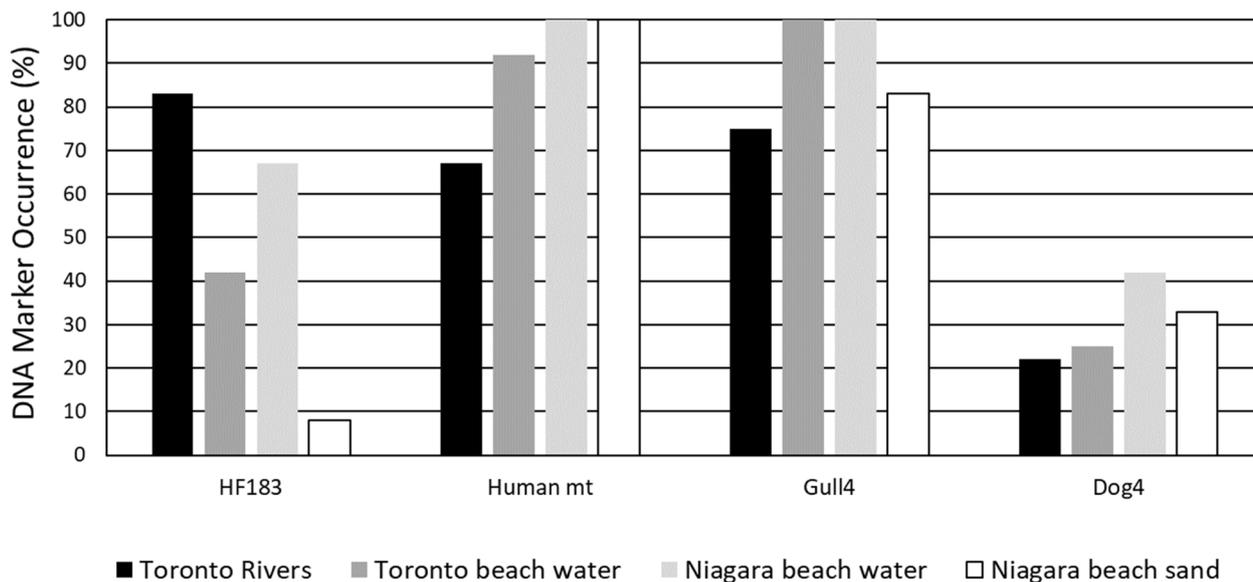


Fig. 4 Detection percentage of human, gull, and dog dPCR fecal source tracking DNA markers in beach, river and sand samples.

Table 2 Frequency of occurrence of microbial source tracking DNA markers from beach, river, and sand locations on BAV exceedance and non-exceedance days

Sampling locations	Potential fecal contamination sources	Frequency of sequence detection (BAV exceedance)	Frequency of sequence detection (BAV non-exceedance)
Toronto beaches	HF183	83%	0%
	Human Mt	83%	100%
	Gull4	100%	100%
	Dog3	50%	0%
Toronto rivers	HF183	100%	67%
	Human Mt	75%	60%
	Gull4	83%	67%
	Dog3	25%	20%
Niagara beaches	HF183	50%	83%
	Human Mt	100%	100%
	Gull4	100%	100%
	Dog3	50%	33%
Niagara beach sand	HF183	17%	0%
	Human Mt	100%	100%
	Gull4	83%	83%
	Dog3	17%	50%

river samples with the human HF183 DNA marker for sewage contamination. In contrast, these food-animal eDNA sequences were more commonly associated with human eDNA sequences and the human mt DNA marker than human-specific bacterial fecal marker (HF183) for Toronto and Niagara beach water and Niagara beach sand. In particular, chicken and cow eDNA sequences were always detected with human eDNA sequences and the human mt DNA marker in Niagara beach sand, but only rarely with the human HF183 DNA marker.

## Discussion

Fecal indicator bacteria, including *E. coli* and *Enterococcus*, commonly used for recreational water quality monitoring, cannot differentiate between fecal sources.<sup>33</sup> Most fecal source

tracking studies have applied a microbial source tracking approach focused on a limited number of targets (source-specific microbial taxa). While targets like the HF183 DNA marker have been widely tested for detecting human fecal contamination, it is increasingly recognized that no marker is 100% host-specific and there is value in having multiple targets to confirm fecal pollution sources.<sup>7</sup> In this study, we applied eDNA metabarcoding and dPCR assays for microbial DNA markers as a more comprehensive approach to assess fecal source profiles for urban freshwater beaches along with their adjacent water sources (rivers/creeks) and beach sands.

Our study found eDNA from fishes, humans, beavers, muskrats, mallard ducks, and gulls was predominant across all our study locations. A study in two rural Quebec watersheds<sup>34</sup> found generally similar results, although that study detected



black bear and moose eDNA, and had fewer dog, red fox, raccoon and gull eDNA detections than our urban area. Our results probably reflected the continuous eDNA shedding by specific animals living in close association with urban aquatic settings and the large human populations in our study area. Our study locations are large urban centers (Toronto and St Catharines), and they are all in close proximity to municipal wastewater treatment plants. The relatively high numbers of human-associated sequences in all our samples were probably associated with treated effluent discharges from nearby municipal wastewater treatment plants, combined sewer overflows, and sewage cross-connected stormwater outfalls. The regular occurrence of sewage contamination at our Toronto study locations has been well documented,<sup>35–37</sup> and the results are in agreement with previous studies finding higher levels of human mitochondrial DNA associated with sewage contamination.<sup>11,38–41</sup>

A challenge with mitochondrial DNA is that it may not be exclusively derived from fecal sources. For example, in recreational water settings, the potential for bather shedding of human skin cells, saliva and hair, in addition to fecal matter, is well-established.<sup>12</sup> Human mitochondrial DNA has even been proposed as an indicator for detecting greywater from residential sources like showers or sinks that do not involve toilet flush water.<sup>42</sup> Human eDNA can be readily detected as a genetic by-catch in water, sand, and air environmental samples, even in areas with relatively low human habitation densities,<sup>43</sup> which raises caution for interpreting human eDNA results from water samples as automatically inferring sewage contamination in some areas. In our urban sewage-impacted study areas, the high prevalence of human eDNA sequences in all water and sand samples precluded our ability to discriminate between sites that were differentially impacted by human fecal contamination. The microbial source tracking HF183 marker was a better indicator of human fecal contamination hotspots and contributions to BAV exceedances than the human mt DNA marker or human eDNA sequences.

An aspect requiring further examination is whether microbial source tracking DNA markers from anaerobic bacteria, and eukaryotic eDNA from different tissue compartments (*e.g.* feces, skin, hair, feathers, rotting carcass) can decay at significantly different rates and have different persistence and transport characteristics. Human mt DNA from sewage has been found to have decay rates similar to those of the microbial HF183 marker in freshwater mesocosms.<sup>44,45</sup> However, it has been suggested that if the mitochondrial (mt) membrane remains intact, this could allow mt DNA to persist longer in the environment.<sup>46</sup> Some studies examining decay rates of eDNA in the environment have identified a slower decay rate of mt DNA than nuclear DNA,<sup>47–49</sup> suggesting that mitochondrial eDNA markers may persist longer in the environment compared to bacterial or eukaryotic nuclear DNA markers. It is possible that a longer persistence and accumulation in beach sand may have contributed to human eDNA sequences and the human mitochondrial DNA marker occurring in 100% of our beach sand samples compared to only 8% for the microbial HF183 marker.

Similar to studies in Ontario and Quebec watersheds,<sup>12,34</sup> our eDNA metabarcoding results detected a diverse range of animals likely impacting water quality. This was particularly the case for beavers and muskrats which may not be typically associated with urban recreational water settings. Detection of beavers and muskrats as a potential fecal contamination source for recreational waters could be very important for beach monitoring strategies as these animals are known to harbour protist pathogens like *Giardia* and *Cryptosporidium*, whose occurrence may be poorly correlated to conventional fecal indicator bacteria like *E. coli*.<sup>50,51</sup> An Ontario agricultural watershed study found muskrats to commonly occur at some sites, and they were also associated with an increased likelihood of occurrence of *Campylobacter* species in water samples.<sup>52</sup> While microbial source tracking *Bacteroides* DNA markers have been developed for beavers<sup>53</sup> and muskrats,<sup>54</sup> they have not been well tested, particularly in urban settings.

Other urban mammal eDNA sequences commonly detected in our samples were for raccoons, red foxes, and dogs. Raccoons occur widely in our study area, and their feces can significantly contribute to fecal pollution in stormwater systems<sup>55</sup> and present *Salmonella* and other pathogen health risks.<sup>56</sup> Similar to our results, dogs have also been identified as locally important sources of fecal pollution at other beaches,<sup>57</sup> which may guide the need for beach-specific remedial actions. Many of these urban wildlife species occur at the animal-human One Health interface and have the potential to contribute to the transmission of zoonotic pathogens or antimicrobial resistance to people in recreational water settings.

eDNA metabarcoding results also detected diverse potential avian fecal contamination sources, particularly Mallard ducks, gulls, and robins. Both eDNA sequencing and the Gull4 microbial source tracking DNA marker indicated the prominence of gull fecal contamination associated with beaches. The significance of gull fecal contamination at our sites is in agreement with previous studies conducted at Sunnyside Beach<sup>35,58</sup> and in riverine and urban coastal areas of Southern Ontario.<sup>59</sup> Another study on Lake Michigan beaches identified gulls as the predominant fecal contamination source, and intervention strategies to reduce gull occurrence significantly improved beach water quality.<sup>60</sup>

Our findings also commonly detected Canada goose eDNA sequences at Toronto beaches, indicating their local importance, even though dPCR did not detect the Canada goose mitochondrial DNA marker. This was consistent with our Toronto river results, where human eDNA sequences were always prevalent, but dPCR less frequently detected the human mitochondrial DNA marker. More common detection of human and Canada goose eDNA sequences may reflect the ability of eDNA sequence data to detect multiple DNA sequence variants (ESVs) for many fecal contamination sources.<sup>61</sup> The ability to detect a greater variability of DNA sequence variants may provide an advantage of eDNA metabarcoding over the specificity of PCR methods for detecting DNA markers.

An unexpected result from our eDNA metabarcoding was the common detection of chicken and cow (and some pig) sequences in all our urban study areas. These sequences could



have been transported from farms, food processing facilities, or residential food waste further away from our study sites. However, it is unlikely such sources would have been so widespread to commonly impact all our study sites at two rivers and four different beaches in two different cities, and for chicken eDNA sequences to be significantly associated with BAV exceedances. Chicken, cow, and pig eDNA sequences were always detected with human eDNA sequences, often with the human HF183 microbial source tracking marker, and in proximity to wastewater treatment plants and other known sewage sources in our study area. Potential concerns about human food DNA sequences carrying over into human feces have been previously identified.<sup>10,12,34,45,62</sup> A previous study on one of our Toronto watersheds<sup>42</sup> used a broader CO1-based eDNA metabarcoding approach and raised a similar concern based on detecting eDNA sequences for non-native food fish species like Tilapia (*Oreochromis* sp.) and Sea bass (*Serranidae* sp.) in water samples. The passage of food DNA through animal digestive tracts is increasingly investigated by metabarcoding diet analyses of wildlife scats, and a study<sup>63</sup> recently used eDNA metabarcoding to demonstrate the high occurrence of domestic dog DNA in red fox scats that was attributed to significant coprophagia. In some situations, such as sewage-impacted aquatic ecosystems, it will be important to apply microbial source tracking DNA markers for common human food sources, including cattle, pig and chicken, alongside eDNA metabarcoding, to test for the possibility of eDNA sequences from undigested food waste in human feces.

eDNA metabarcoding identified a broader range of animal species likely impacting water quality in our study area than the more limited range of microbial source-tracking DNA markers available. However, we also identified limitations of eDNA metabarcoding that require further investigation. eDNA metabarcoding studies of fecal pollution should use microbial source tracking DNA markers in some situations as additional lines of evidence for fecal source attribution. Human eDNA PCR blocking techniques may prove useful for sewage-contaminated urban settings to better associate the occurrence of animal eDNA sequences with fecal contamination. Using fecal source tracking information from both eDNA metabarcoding and microbial source tracking DNA markers, combined with conventional fecal indicator bacteria, can help design targeted beach sustainability programs and risk management actions to control fecal contamination.

## Conclusions

(1) eDNA metabarcoding of mammal and avian 16S rRNA gene sequences found human eDNA sequences to predominate in all urban beach water, sand, and river samples. PCR blocking for human eDNA sequences should be considered to avoid masking the detection of fecal contamination from less common animals in sewage-impacted areas.

(2) Beaver, muskrat, gull, and mallard eDNA sequences were the most prominent animals detected across our study sites, indicating the potential for these wildlife species to contribute

to fecal contamination and health risks at urban freshwater beaches.

(3) Dogs and Canada goose eDNA sequences were most common at Sunnyside Beach, suggesting that these fecal contamination sources can be beach-specific and may require targeted remediation strategies.

(4) Digital PCR assays for the bacterial Gull4 DNA marker and a human mitochondrial DNA marker were consistent with eDNA metabarcoding in detecting the widespread occurrence of these potential fecal sources.

(5) Chicken and cow eDNA sequences were widely detected across all study sites but are suggested to be from sewage and incompletely digested human food origins.

(6) eDNA metabarcoding can expand fecal source tracking capabilities for assessing diverse wildlife contributions to fecal pollution. However, microbial fecal source tracking markers for common food animals (*e.g.* chicken, cattle, and pigs) should be tested alongside eDNA metabarcoding analysis to address questions about the significance of coprophagy, non-fecal skin sources or occurrence of eDNA sequences from incompletely digested food in fecal pollution sources.

## Data availability

The data supporting this article have been included as part of the ESI.†

## Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors would like to thank Canada's Natural Sciences and Engineering Research Council and NSERC Alliance grants ALLRP-554507-20 and ALLRP 590321-23 for financial support, as well as support from the Ontario Ministry of Environment, Conservation and Parks.

## References

- 1 M. B. Nevers, M. N. Byappanahalli, D. Shively, P. M. Buszka, P. R. Jackson and M. S. Phanikumar, Identifying and eliminating sources of recreational water quality degradation along an urban coast, *J. Environ. Qual.*, 2018, **47**(5), 1042–1050.
- 2 K. G. Field and M. Samadpour, Fecal source tracking, the indicator paradigm, and managing water quality, *Water Res.*, 2007, **41**(16), 3517–3538.
- 3 E. Stachler, B. Akyon, N. A. de Carvalho, C. Ference and K. Bibby, Correlation of crAssphage qPCR markers with culturable and molecular indicators of human fecal pollution in an impacted urban watershed, *Environ. Sci. Technol.*, 2018, **52**(13), 7505–7512.



- 4 N. Saeidi, X. Gu, N. H. Tran, S. G. Goh, M. Kitajima, A. Kushmaro, *et al.*, Occurrence of traditional and alternative fecal indicators in tropical urban environments under different land use patterns, *Appl. Environ. Microbiol.*, 2018, **84**(14), e0028718.
- 5 A. B. Boehm, L. C. Van De Werfhorst, J. F. Griffith, P. A. Holden, J. A. Jay, O. C. Shanks, *et al.*, Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study, *Water Res.*, 2013, **47**(18), 6812–6828.
- 6 E. Jardé, L. Jeanneau, L. Harrault, E. Quenot, O. Solecki, P. Petitjean, *et al.*, Application of a microbial source tracking based on bacterial and chemical markers in headwater and coastal catchments, *Sci. Total Environ.*, 2018, **610**, 55–63.
- 7 K. Demeter, R. Linke, E. Ballesté, G. Reischer, R. E. Mayer, J. Vierheilig, *et al.*, Have genetic targets for faecal pollution diagnostics and source tracking revolutionized water quality analysis yet?, *FEMS Microbiol. Rev.*, 2023, **47**(1), 1–36, DOI: [10.1093/femsre/fuad028](https://doi.org/10.1093/femsre/fuad028).
- 8 R. E. Mayer, G. H. Reischer, S. K. Ixenmaier, J. Derx, A. P. Blaschke, J. E. Ebdon, *et al.*, Global distribution of human-associated fecal genetic markers in reference samples from six continents, *Environ. Sci. Technol.*, 2018, **52**(9), 5076–5084.
- 9 A. Martellini, P. Payment and R. Villemur, Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine, and ovine sources in fecally contaminated surface water, *Water Res.*, 2005, **39**(4), 541–548.
- 10 J. M. Caldwell, M. E. Raley and J. F. Levine, Mitochondrial multiplex real-time PCR as a source tracking method in fecal-contaminated effluents, *Environ. Sci. Technol.*, 2007, **41**(9), 3277–3283.
- 11 T. A. Edge, R. J. Boyd, P. Shum and J. L. Thomas, Microbial source tracking to identify fecal sources contaminating the Toronto Harbour and Don River watershed in wet and dry weather, *J Great Lakes Res.*, 2021, **47**(2), 366–377.
- 12 Z. R. Staley, J. D. Chuong, S. J. Hill, J. Grabuski, S. Shokralla, M. Hajibabaei, *et al.*, Fecal source tracking and eDNA profiling in an urban creek following an extreme rain event, *Sci. Rep.*, 2018, **8**(1), 14390.
- 13 R. Ragot and R. Villemur, eDNA profiling of mammals, birds, and fish of surface waters by mitochondrial metagenomics: application for source tracking of fecal contamination in surface waters, *Environ. Monit. Assess.*, 2022, **194**(9), 646.
- 14 F. Saleem, H. E. Schellhorn, A. Simhon and T. A. Edge, Same-day Enterococcus qPCR results of recreational water quality at two Toronto beaches provide added public health protection and reduced beach days lost, *Can. J. Public Health*, 2023, 1–12.
- 15 F. Saleem, T. A. Edge and H. E. Schellhorn, Application of same-day Enterococcus qPCR-based analyses for quality assessment of shorelines (water and sand) at recreational beaches, *Water*, 2023, **15**(13), 2338.
- 16 H. C. Green, R. A. Haugland, M. Varma, H. T. Millen, M. A. Borchardt, K. G. Field, *et al.*, Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples, *Appl. Environ. Microbiol.*, 2014, **80**(10), 3086–3094.
- 17 H. Ryu, J. F. Griffith, I. U. Khan, S. Hill, T. A. Edge, C. Toledo-Hernandez, *et al.*, Comparison of Gull feces-specific assays targeting the 16S rRNA genes of *Catelicoccus marimammalium* and *Streptococcus* spp, *Appl. Environ. Microbiol.*, 2012, **78**(6), 1909–1916.
- 18 J. M. Caldwell and J. F. Levine, Domestic wastewater influent profiling using mitochondrial real-time PCR for source tracking animal contamination, *J. Microbiol. Methods*, 2009, **77**(1), 17–22.
- 19 H. C. Green, K. M. White, C. A. Kelty and O. C. Shanks, Development of rapid canine fecal source identification PCR-based assays, *Environ. Sci. Technol.*, 2014, **48**(19), 11453–11461.
- 20 S. Mieszkina, J. F. Yala, R. Joubrel and M. Gourmelon, Phylogenetic analysis of Bacteroidales 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR, *J. Appl. Microbiol.*, 2010, **108**(3), 974–984.
- 21 S. Mieszkina, J. P. Furet, G. Corthier and M. Gourmelon, Estimation of pig fecal contamination in a river catchment by real-time PCR using two pig-specific Bacteroidales 16S rRNA genetic markers, *Appl. Environ. Microbiol.*, 2009, **75**(10), 3045–3051.
- 22 J. R. Willis, M. Sivaganesan, R. A. Haugland, J. Kralj, S. Servetas, M. E. Hunter, *et al.*, Performance of NIST SRM® 2917 with 13 recreational water quality monitoring qPCR assays, *Water Res.*, 2022, **212**, 118114, DOI: [10.1016/j.watres.2022.118114](https://doi.org/10.1016/j.watres.2022.118114).
- 23 J. Brown, M. Pirrung and L. A. McCue, FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool, *Bioinformatics*, 2017, **33**(19), 3137–3139.
- 24 S. Chen, Y. Zhou, Y. Chen and J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor, *Bioinformatics*, 2018, **34**(17), i884–i890.
- 25 D. Buchner, T. H. Macher and F. Leese, APSCALE: advanced pipeline for simple yet comprehensive analyses of DNA metabarcoding data, *Bioinformatics*, 2022, **38**(20), 4817–4819.
- 26 M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, *EMBnet J.*, 2011, **17**(1), 10–12.
- 27 T. Rognes, T. Flouri, B. Nichols, C. Quince and F. Mahé, VSEARCH: a versatile open source tool for metagenomics, *PeerJ*, 2016, **4**, e2584.
- 28 T. G. Frøslev, R. Kjølner, H. H. Bruun, R. Ejrnæs, A. K. Brunbjerg, C. Pietroni, *et al.*, Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates, *Nat. Commun.*, 2017, **8**(1), 1188.
- 29 L. E. Drake, J. P. Cuff, R. E. Young, A. Marchbank, E. A. Chadwick and W. O. Symondson, An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data, *Methods Ecol. Evol.*, 2022, **13**(3), 694–710.



- 30 M. Waso, S. Khan and W. Khan, Development and small-scale validation of a novel pigeon-associated mitochondrial DNA source tracking marker for the detection of fecal contamination in harvested rainwater, *Sci. Total Environ.*, 2018, **615**, 99–106.
- 31 A. Kassambara, *Ggplot2: "Ggplot2" Based Publication Ready Plots*, R package version 0.4.0, 2020.
- 32 K. J. Andresa, D. M. Lodge and J. Andrésa, Environmental DNA reveals the genetic diversity and population structure of an invasive species in the Laurentian Great Lakes, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**(37), e2307345120, DOI: [10.1073/pnas.2307345120](https://doi.org/10.1073/pnas.2307345120).
- 33 N. H. Tran, K. Y. H. Gin and H. H. Ngo, Fecal pollution source tracking toolbox for identification, evaluation and characterization of fecal contamination in receiving urban surface waters and groundwater, *Sci. Total Environ.*, 2015, **538**, 38–57.
- 34 R. Ragot, F. Lessard, A. Bélanger and R. Villemur, Assessment of multiple fecal contamination sources in surface waters using environmental mitochondrial DNA metabarcoding, *Sci. Total Environ.*, 2023, **898**, 165237.
- 35 Z. R. Staley and T. A. Edge, Comparative microbial source tracking methods for identification of fecal contamination sources at Sunnyside Beach in the Toronto region area of concern, *J. Water Health*, 2016, **14**(5), 839–850.
- 36 Z. Staley, J. Grabuski, E. Sverko and T. A. Edge, Comparison of microbial and chemical source tracking markers to identify fecal contamination sources in the Humber River and associated stormwater outfalls, *Appl. Environ. Microbiol.*, 2016, **82**(19), 6357–6366.
- 37 E. Li, F. Saleem, T. A. Edge and H. E. Schellhorn, Assessment of crAssphage as a human fecal source tracking marker in the lower Great Lakes, *Sci. Total Environ.*, 2024, **912**, 168840.
- 38 V. Kapoor, C. Smith, J. W. Santo Domingo, T. Lu and D. Wendell, Correlative assessment of fecal indicators using human mitochondrial DNA as a direct marker, *Environ. Sci. Technol.*, 2013, **47**(18), 10485–10493.
- 39 R. Villemur, M. Imbeau, M. N. Vuong, L. Masson and P. Payment, An environmental survey of surface waters using mitochondrial DNA from human, bovine and porcine origin as fecal source tracking markers, *Water Res.*, 2015, **69**, 143–153.
- 40 M. Hajj-Mohamad, M. Hachad, G. Deschamps, S. Sauvé, R. Villemur, M. A. Blais, *et al.*, Fecal contamination of storm sewers: Evaluating wastewater micropollutants, human-specific Bacteroides 16S rRNA, and mitochondrial DNA genetic markers as alternative indicators of sewer cross connections, *Sci. Total Environ.*, 2019, **659**, 548–560.
- 41 A. B. M. Tanvir Pasha, J. Hinojosa, D. Phan, A. Lopez and V. Kapoor, Detection of human fecal pollution in environmental waters using human mitochondrial DNA and correlation with general and human-associated fecal genetic markers, *J. Water Health*, 2020, **18**(1), 8–18.
- 42 B. D. Zimmerman, N. J. Ashbolt, J. L. Garland, S. Keely and D. Wendell, Human mitochondrial DNA and endogenous bacterial surrogates for risk assessment of graywater reuse, *Environ. Sci. Technol.*, 2014, **48**(14), 7993–8002.
- 43 L. Whitmore, M. McCauley, J. A. Farrell, M. R. Stammnitz, S. A. Koda, N. Mashkour, *et al.*, Inadvertent human genomic bycatch and intentional capture raise beneficial applications and ethical concerns with environmental DNA, *Nat. Ecol. Evol.*, 2023, **7**(6), 873–888.
- 44 R. Ragot and R. Villemur, Influence of temperature and water quality on the persistence of human mitochondrial DNA, human Hf183 Bacteroidales, fecal coliforms, and enterococci in surface water in human fecal source tracking context, *Sci. Total Environ.*, 2022, **838**, 156025.
- 45 W. Ren and Y. Feng, Persistence of human and cattle-associated Bacteroidales and mitochondrial DNA markers in freshwater mesocosms, *Sci. Total Environ.*, 2023, **899**, 165742.
- 46 J. B. Harrison, J. M. Sunday and S. M. Rogers, Predicting the fate of eDNA in the environment and implications for studying biodiversity, *Proc. Biol. Sci.*, 2019, **286**(1915), 20191409.
- 47 D. R. Foran, Relative degradation of nuclear and mitochondrial DNA: an experimental approach, *J. Forensic Sci.*, 2006, **51**(4), 766–770.
- 48 A. L. Emmons, J. M. DeBruyn, A. Z. Mundorff, K. L. Cobaugh and G. S. Cabana, The persistence of human DNA in soil following surface decomposition, *Sci. Justice*, 2017, **57**(5), 341–348.
- 49 J. Bylemans, E. M. Furlan, D. M. Gleeson, C. M. Hardy and R. P. Duncan, Does size matter? An experimental evaluation of the relative abundance and decay rates of aquatic environmental DNA, *Environ. Sci. Technol.*, 2018, **52**(11), 6408–6416.
- 50 C. K. M. Tsui, R. Miller, M. Uyaguari-Diaz, P. Tang, C. Chauve, W. Hsiao, *et al.*, Beaver Fever: Whole-Genome Characterization of Waterborne Outbreak and Sporadic Isolates to Study the Zoonotic Transmission of Giardiasis, *mSphere*, 2018, **3**(2), e00090.
- 51 A. Bitto and A. Aldras, Prevalence of Giardia and Cryptosporidium in muskrats in northeastern Pennsylvania and New Jersey, *J. Environ. Health*, 2009, **71**(10), 20–26.
- 52 R. Marti, V. P. Gannon, C. Jokinen, M. Lanthier, D. R. Lapen, N. F. Neumann, *et al.*, 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.*, 2013, **47**(7), 2315–2324.
- 53 R. Marti, Y. Zhang, Y. C. Tien, D. R. Lapen and E. Topp, Assessment of a new Bacteroidales marker targeting North American beaver (*Castor canadensis*) fecal pollution by real-time PCR, *J. Microbiol. Methods*, 2013, **95**(2), 201–206.
- 54 R. Marti, Y. Zhang, D. R. Lapen and E. Topp, Development and validation of a microbial source tracking marker for the detection of fecal pollution by muskrats, *J. Microbiol. Methods*, 2011, **87**(1), 82–88.
- 55 J. L. Ram, B. Thompson, C. Turner, J. M. Nechvatal, H. Sheehan and J. Bobrin, Identification of pets and raccoons as sources of bacterial contamination of urban storm sewers using a sequence-based bacterial source tracking method, *Water Res.*, 2007, **41**(16), 3605–3614.



- 56 K. J. Bondo, D. L. Pearl, N. Janecko, P. Boerlin, R. J. Reid-Smith, J. Parmley, *et al.*, Impact of season, demographic and environmental factors on Salmonella occurrence in raccoons (*Procyon lotor*) from swine farms and conservation areas in Southern Ontario, *PLoS One*, 2016, **11**(9), e0161497.
- 57 A. Shrestha, C. A. Kelty, M. Sivaganesan, O. C. Shanks and S. Dorevitch, Fecal pollution source characterization at non-point source impacted beaches under dry and wet weather conditions, *Water Res.*, 2020, **182**, 116014.
- 58 T. A. Edge, S. Hill, P. Seto and J. Marsalek, Library-dependent and library-independent microbial source tracking to identify spatial variation in faecal contamination sources along a Lake Ontario beach (Ontario, Canada), *Water Sci. Technol.*, 2010, **62**(3), 719–727.
- 59 J. Lu, H. Ryu, S. Hill, M. Schoen, N. Ashbolt, T. A. Edge, *et al.*, Distribution and potential significance of a gull fecal marker in urban coastal and riverine areas of southern Ontario, Canada, *Water Res.*, 2011, **45**(13), 3960–3968.
- 60 R. R. Converse, J. L. Kinzelman, E. A. Sams, E. Hudgens, A. P. Dufour, H. Ryu, *et al.*, Dramatic improvements in beach water quality following gull removal, *Environ. Sci. Technol.*, 2012, **46**(18), 10206–10213.
- 61 J. Skelton, A. Cauvin and M. E. Hunter, Environmental DNA metabarcoding read numbers and their variability predict species abundance, but weakly in non-dominant species, *Environ. DNA*, 2023, **5**(5), 1092–1104.
- 62 R. Kortbaoui, A. Locas, M. Imbeau, P. Payment and R. Villemur, Universal mitochondrial PCR combined with species-specific dot-blot assay as a source-tracking method of human, bovine, chicken, ovine, and porcine in fecal-contaminated surface water, *Water Res.*, 2009, **43**(7), 2002–2010.
- 63 C. N. Waggershauer, P. Taberlet, E. Coissac, K. Kortland, C. Hambly and X. Lambin, Interspecific coprophagia by wild red foxes: DNA metabarcoding reveals a potentially widespread form of commensalism among animals, *Ecol. Evol.*, 2022, **12**(7), e9029.

