Journal of Experimental Microbiology and Immunology (JEMI) Vol. 15: 130 – 135

Genotyping *Escherichia coli* Isolates from Duck, Goose, and Gull Fecal Samples with Phylogenetic Markers using Multiplex Polymerase Chain Reaction for Application in Microbial Source Tracking

Cherry Chu Yee Lee
Department of Microbiology & Immunology, University of British Columbia

Multiplex polymerase chain reaction with phylogenetic markers, *chuA*, *yjaA*, and TspE4.C2 along with virulence gene, *eae*, was performed for determination of genotypic differences in *Escherichia coli* isolated from gull, goose, and duck. Results showed that isolates from goose frequently contained TspE4.C2 and an unknown 600 bp product (87%) whereas isolates from gull and duck contained varied combinations of the three phylogenetic markers. In seven of the 15 duck samples, all three phylogenetic markers were present as in human *E. coli* isolates. Similar results were evident in two of the 15 gull samples. In addition to amplification of the unidentified 600 bp product, a 1500 bp product was observed in five of the 45 isolates tested. None of the isolates contained *eae*. Due to genotypic variations in gull and duck *E. coli* isolates, the multiplex PCR method studied would not be able to positively identify hosts for general or robust use in microbial source tracking.

Fecal contamination in drinking water is of concern to general public health and hygiene in all parts of the world. Water sources may be contaminated by manure used for agricultural purposes, sewage overflows, ineffective aseptic water decontamination systems, and false analysis of water quality (7). Identification of the source of bacterial contamination may be useful for developing plans to resolve or reduce water contamination occurrences. *Escherichia coli* are often used in microbial source tracking (MST) as an indicator for existence of other bacterial contamination in water. To determine the source of *E. coli* isolates, ribotype analysis, pulse-field gel electrophoresis, antibiotic resistance profiling, and rep-PCR DNA fingerprinting techniques have been studied for their effectiveness, efficiency, and reliability to correctly identify infected hosts (3).

The use of antibiotic resistance profiling of *E. coli* depends on the geographic location and time as the microorganisms are exposed to different environmental factors that affect their development of antibiotic resistance (3). DNA fingerprinting using repetitive elements, like primer BOX A1R, requires optimized polymerase chain reaction (PCR) and gel electrophoresis conditions to produce cleanly resolved bands for analysis and initial library generation of the DNA fingerprints representing different source *E. coli* (7, 11). Also, individual organisms of the same species can have different fingerprint patterns, and different species require varying statistical analysis (3). Often, DNA fingerprinting and other library-based methods correlate to misclassification and continuous library enlargement due to the diversity of *E. coli* strains from fecal sources (8). Although many approaches seem feasible, a completely reliable method is yet to be discovered.

Thus, an alternative approach relying on the phylogenetic grouping of *E. coli* strains based on genetic markers, *chuA*, *yjaA*, and TspE4.C2 DNA fragment, was recently studied in application for MST. Based on the study by Clermont et al. (2) and recently, Carlos et al. (1), *E. coli* can be categorized into four phylogenetic groups: A, B1, B2, or D, and then into subgroups: A_0, A_1, B1, $B2_2$, $B2_3$, D_1, and D_2. The groups were determined based on the presence or absence of *chuA*, *yjaA*, and DNA fragment TspE4.C2. *chuA* is known to be involved with heme transport while *yjaA* is involved in cellular response to hydrogen peroxide and acid stress (ecocyc.org). The function of TspE4.C2 is not yet discovered. *E. coli* containing combinations of one, two, all, or none of these genes were categorized into seven subgroups. These researchers hypothesized that certain subgroups were specifically found in certain species. In support of this hypothesis, Escobar-Paramo et al. (4) found that *E. coli* strains isolated from bird feces commonly represented groups B1 and D while human fecal *E. coli* isolates prevailed in groups A and B2. Recently, Carlos et al. (1) found that subgroup $B2_3$ (*chuA*+, *yjaA*+, and TspE4.C2+) *E. coli* was only found in humans. The above mentioned studies both demonstrated that phylogenetic grouping of *E. coli* isolates using the three

biomarkers (*chuA*, *yjaA*, and TspE4.C2) in multiplex PCR can lead to correlation between *E. coli* subgroups and their host; thus, further investigations using this method in attempt to identify source *E. coli* contamination should be considered.

In addition to determining the source of bacterial contamination, co-identification of virulent bacteria in water samples would eliminate additional microbial testing procedures. The gene *eae* is frequently associated with bacterial virulence in *Escherichia coli*, *Citrobacter rodentium*, and *Salmonella enterica* (www.ncbi.nlm.nih.gov/gene). Specifically, *eae* is associated with Enteropathogenic *E. coli* (EPEC) and Enterohaemorrhagic *E. coli* (EHEC) (9). These pathogenic strains are good indicators for recent fecal contamination in water due to their short life span in water (approximately 4-12 weeks; 12). Both EPEC and EHEC can contaminate water via feces from animal hosts. *eae* is a chromosomal gene that encodes for intimin (outer membrane adhesive protein) and is found to be expressed in all virulent *E. coli* strains (5). Hence, by testing *chuA*, *yjaA*, TspE4.C2, and *eae* together using multiplex PCR, information concerning fecal source and presence of virulent bacteria can be determined concurrently.

Therefore, the objective of this study is to investigate whether genetic differences exist between *E. coli* isolates from duck, geese and gull by analyzing their genotype for *chuA*, *yjaA*, and DNA fragment TspE4.C2 using multiplex PCR for application in MST. In addition, the virulent gene *eae* would be included in the multiplex PCR to test for pathogenic *E. coli*.

MATERIALS AND METHODS

Reagents, media, and primers. Luria-Bertani agar was prepared by dissolving 10 g of Bacto-Tryptone (BD, cat no. 211705), 5 g Bacto-Yeast extract (BD, cat no. 212750), 10 g Sodium chloride, and 6 g of Nutrient Agar (Difco, cat no. 0001-17-0) in distilled water and pH-adjusted with Sodium hydroxide to a final volume of 1 L and pH 7.0. Four milliliter LB broth tubes were prepared similarly except without agar. All media were sterilized and stored at 4°C until use. For washing cells, 1 M NaCl solution (Fisher, cat no. S271-3) was prepared and sterilized. Primers were synthesized by Integrated DNA Technologies (San Diego, California) with standard desalting and resuspended in distilled water to 100 uM stock concentrations (Table 1). PCR optimization procedures were performed using Invitrogen *Taq* polymerase (Invitrogen, Cat no. 18038-042) and corresponding reagents (50 mM MgCl2 and 10 X buffer – MgCl2) while colony PCR of Trout Lake samples were conducted using Fermentas *Taq* polymerase (Fermentas, cat no. EP0404) and reagents (25 mM MgCl and 10 X buffer + KCl – MgCl2). Deoxynucleotide triphosphates were previously prepared at 25 mM. GeneRuler 100 bp Plus DNA ladder supplied with 6 X DNA loading dye was purchased from Fermentas (cat no. 00050508). 10 X Tris-borate EDTA (TBE) was prepared by dissolving 54 g Tris base (Fisher, cat no. BP152-1), 27.5 g Boric acid (VWR, VW1471-01), and 3.72 g EDTA disodium salt (Fisher, cat no. BP1201) in 500 mL of distilled water. Ethidium bromide staining solution (0.5 ug/mL) was prepared with 200 uL of 1 mg/mL Ethidium bromide stock solution (March 25, 2003) in 400 mL distilled water.

Preparation of *E. coli* samples. Duck, goose and gull fecal *E. coli* isolates were obtained from Trout Lake, British Columbia, during the summer of 1995 (Table 2). The frozen samples were thawed at room temperature, then streaked onto LB agar plates and incubated overnight at 37°C to obtain isolated colonies for use in PCR. *E. coli* isolates from human were also thawed and cultured similarly. EHEC O157:H7 DNA was isolated and prepared from bacterial cultures by Dr. Deng.

Verification of positive controls. A human *E. coli* isolate and an EHEC DNA sample was tested for use as a primer control and a positive control for virulence, respectively. Previous studies have shown that *E. coli* from human contains all three phylogenetic markers (*chuA*, *yjaA*, and TspE4.C2) and EHEC contains *eae* and *chuA* (2). Controls were initially tested using both the cell wash method (12) and direct colony method of PCR as an additional verification that the positive controls can produce reproducible bands indicative of the robustness of the primers and PCR conditions.

Optimization of multiplex PCR annealing temperature. Direct colony PCR of human *E. coli* isolates and EHEC were used to determine the optimal annealing temperature for multiplexing *chuA*, *yjaA*, TspE4.C2, and *eae* primers. Each 10 uL PCR reaction contained 1.25 U of *Taq* polymerase, 1X Buffer-$MgCl_2$, 1.5 mM $MgCl_2$, 200 uM dNTP mix, 20 uM of each primer, distilled water, and either 1 uL of DNA (50 ng/uL) or a colony with the area of a P2 tip and height of 0.4 mm. Annealing temperatures tested include 54.4, 55.8, 58.7, and

TABLE 1. Primers used in PCR to amplify specific fragments from *E. coli*.

Identification	Primer Sequence (5' → 3')	Primer Length (bases)	Expected Product Size (bp)	Reference
ChuA.1	GACGAACCAACGGTCAGGAT	20	279	Clermont et al. (2)
ChuA.2	TGCCGCCAGTACCAAAGACA	20		Clermont et al. (2)
YjaA.1	TGAAGTGTCAGGAGACGCTG	20	211	Clermont et al. (2)
YjaA.2	ATGGAGAATGCGTTCCTCAAC	21		Clermont et al. (2)
TspE4.C2.1	GAGTAATGTCGGGGCATTCA	20	152	Clermont et al. (2)
TspE4.C2.2	CGCGCCAACAAAGTATTACG	20		Clermont et al. (2)
Eae.1	CCCGAATTCGGCACAAGCATAAGC	24	863	Zhang et al. (11)
Eae.2	CCCGGATCCGTCTCGCCAGTATTCG	25		Zhang et al. (11)

TABLE 2. *E. coli* isolates and DNA used in PCR.

Host	Source, date	Sample Identification	*E. coli* isolate (n)
Gull	Trout Lake, Summer 1995	Tr Gu 95-1 to -2 (Gull L) Tr Gu 95-3 to -12 (Gull J) Tr Gu 95-13 to -15 (Gull K)	15
Goose	Trout Lake, Summer 1995	Tr Go 95-1 to -15 (data on specific goose not available)	15
Duck	Trout Lake, Summer 1995	Tr Du 95-1 to -6 (Duck P) Tr Du 95-7 to -15 (Duck N)	15
Human	Steve Partington	Ec 20	1
EHEC O157:H7 strain EDL933	Bacterial cultures 100 ng/ul	N/A	1

61.6°C. All optimization reactions were performed using Biometra T-gradient PCR Thermocycler (Montreal Biotech Inc., Kirkland P.Q., Canada). Samples were denatured at 94°C for 4 minutes before 30 cycles of 94°C for 45 seconds, 54.4 to 61.6°C for 60 seconds, and 72°C for 90 seconds. A final extension step at 72°C for 5 minutes was also included.

Multiplex colony PCR of duck, goose, and gull *E. coli* samples. Polymerase chain reactions were prepared as above. *E. coli* isolates from human and EHEC DNA were used as positive controls while distilled water was used as a negative control. The 15 *E. coli* isolates from each of duck, goose, and gull were tested in biological replicates of five (i.e. three independent experiments containing 5 duck, 5 goose, and 5 gull samples). Cycling parameters were as follows: 94°C for 4 minutes, 25 or 30 cycles of 94°C for 30 seconds, 56°C for 45 seconds, 72°C for 30 seconds, and 72°C for 5 minutes. Both Bio-Rad GeneCycler and Biometra T-gradient Thermocycler were used.

Gel electrophoresis. Two percent agarose gels were prepared in 1 X TBE. To each 10 uL PCR reaction, 2 uL of 6X DNA loading dye was added and mixed. Samples were briefly centrifuged before they were loaded into the gel. A lane was also dedicated for the loading of 0.5 ug of the 100 bp DNA ladder in every gel. Samples ran at 100 V for approximately one hour. Gels were post-stained in 0.5 ug/mL Ethidium bromide for 20 to 60 minutes.

Data analysis. The presence of each phylogenetic marker and unknown amplified product of a distinct size in each species was tallied and expressed as a frequency plot. The genotype profile of each *E. coli* isolate was organized in a venn diagram denoting the three phylogenetic markers, *chuA*, *yjaA*, and TspE4.C2.

RESULTS

Multiplex colony PCR and gel electrophoresis of gull, goose, and duck *E. coli* isolates. Results from three independent experiments, representing 15 *E. coli* isolates from each of gull, goose, and duck demonstrated that all *E. coli* isolates contain at least one of the three phylogenetic markers, *chuA*, *yjaA*, and TspE4.C2 (Fig.1). Also, the virulence gene, *eae*, was not detected in any of the fecal samples using multiplex colony PCR with the specified primers (Table 1). The human *E. coli* isolate and EHEC DNA controls showed amplification of *chuA*, *yjaA*, with TspE4.C2, and *eae* with *chuA*, respectively, as expected. In addition to amplification of phylogenetic markers and *eae*, some products at 600 bp and 1500 bp were observed in samples from gull, goose, and duck as shown in Figure 1; however, this was not evident in the positive or negative controls. The 600 bp and 1500 bp bands on the gel had similar resolution and intensity compared to *chuA*, *yjaA*, TspE4.C2, and *eae* bands. In general, isolates from gull contained either one of the three phylogenetic genes and the 600 bp product. Specifically, 13 of the 15 gull samples contained the combination of TspE4.C2 and the 600 bp product. The presence of only TspE4.C2 and the 600 bp unknown product in combination was also seen in goose and duck samples. Other than the 1500 bp, 600 bp, and expected products, unpredicted bands or smears at approximately 40 bp and 65 bp were observed. Nonetheless, these bands were either at least half the intensity or had very low resolution compared to the expected products. Although, unexpected bands representing *chuA* and *yjaA* were present in the no template control, and similarly, *yjaA* was amplified in the EHEC control in one of the three independent experiments, the bands were of relatively low intensity (Fig. 1). This phenomenon would be discussed later.

Frequency of products amplified in gull, goose, and duck *E. coli* isolates. As shown in Figure 2, of the 15 *E. coli* isolates from duck, all contained at least one of *chuA*, *yjaA*, TspE4.C2, 600 bp band, and/or 1500 bp band. All gull samples also contained at least one of these products except for two samples. None of the goose *E. coli* isolates contained *yjaA*. Briefly, by comparing the frequency of amplified products in gull, goose, and duck, the presence of *chuA* seem to correlate to duck whereas the 600 bp product to goose; yet, a relationship between amplified products and gull samples cannot be clearly seen (Fig. 2). In addition, absence of *yjaA* implied goose *E. coli* isolates. In general, TspE4.C2 was most frequently detected in *E. coli* isolates of all three species, whereas the 1500 bp product was least prevalent. Within each species, TspE4.C2 was also observed at the highest frequency; however, the frequency of products from next highest to lowest differed between gull, goose, and duck. In gulls, following TspE4.C2 were *yjaA* and the 600 bp product, *chuA*, and the 1500 bp product, in decreasing occurrences. In decreasing order, TspE4.C2 and the 600 bp product, followed by *chuA* and the 1500 bp product were detected in geese. In duck, from most to least

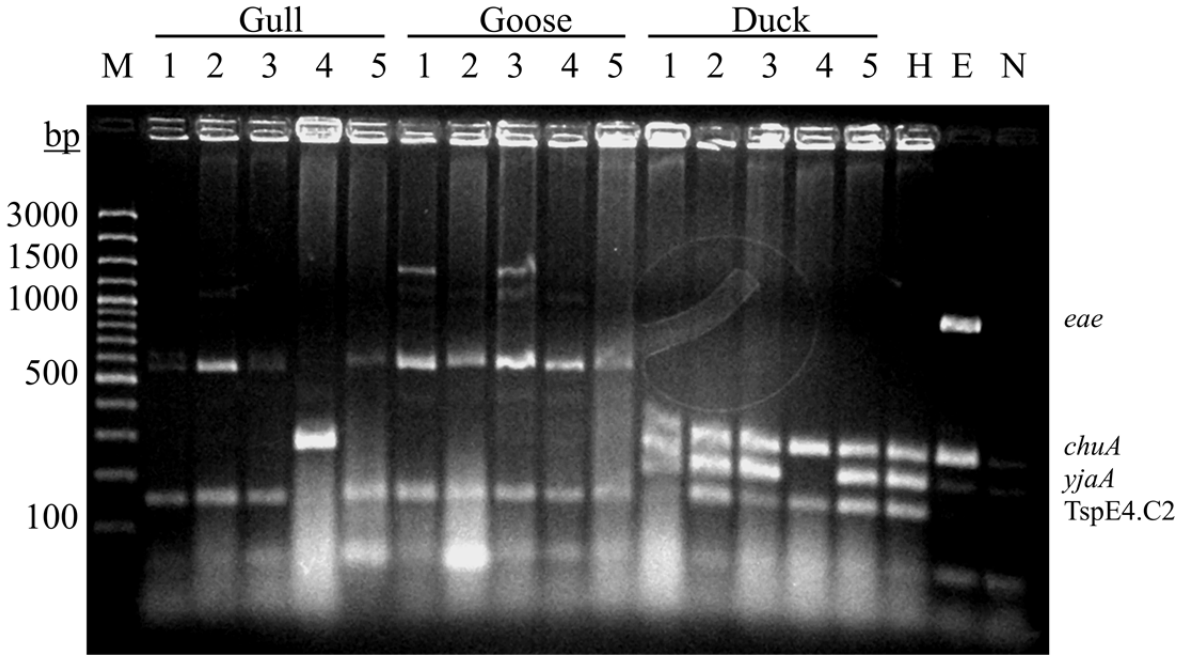

FIG. 1. Phylogenetic genes in *E. coli* isolates from gull, goose, and duck fecal samples. The presence of *chuA* (279 bp), *yjaA* (211 bp), TspE4.C2 (152 bp) and *eae* (863 bp) in *E. coli* isolates from gull, goose, and duck fecal samples was determined using multiplex colony PCR. Five isolates from each species is shown. Samples were resolved in a 2% agarose gel in 1 X TBE at 100 V for an hour. The gel was stained in 0.5 ug/mL Ethidium bromide for 30 minutes and UV-exposed for 2.6 seconds. The gel image (above) was captured using AlphaImager. A human *E. coli* isolate (H) and an EHEC DNA sample (E) served as positive controls whereas a reaction with no template (N) served as a negative control. Samples labeled 1 to 5 represent Tr 95-1 to Tr 95-5 in each of the three species.

frequently found were TspE4.C2, *chuA*, *yjaA*, the 600 bp product and the 1500 bp product.

Comparison of the phylogenetic gene profile of *E. coli* isolated from gull, goose, and duck fecal samples. The genetic relationship involving *chuA*, *yjaA*, and TspE4.C2 DNA fragment between gull, goose, and duck are outlined in Figure 3. This venn diagram shows six distinct categories of genetic profiles made up of *chuA*, *yjaA*, and TspE4.C2 alone or in combination. Generally, some isolates from duck (47%) and gull (20%) contained all three of the phylogenetic markers similar to the human *E. coli* isolate. No samples had the *chuA* and *yjaA* combination. On the other hand, a combination of *yjaA* and TspE4.C2 was observed in a duck sample while *chuA* and TspE4.C2 was evident in both duck and goose samples. Sole amplification of TspE4.C2 was detected in all three species, whereas presence of only *chuA* was seen in gull and duck *E. coli* isolates. Furthermore, gull isolates were the only ones found to contain *yjaA* only. Overall, variation in the pattern of phylogenetic markers present in each species from least to greatest was as follows: goose, gull, and duck, representing two, four, and five different categories, respectively. Note that two samples of gull *E. coli* isolates did not contain any of the markers. Moreover, the distribution of these *E. coli* isolates in different categories demonstrated that goose samples were the most confined as 13 isolates categorized in the TspE4.C2 only section. Duck samples had greater spread than goose with seven samples all containing the three phylogenetic markers and other samples largely distributed unequally in four different categories. Lastly, gull *E. coli* isolates revealed the largest distribution as only six samples dominated the TspE4.C2 only category and the rest were almost equally distributed in four other categories (including the genotype where no markers were present).

DISCUSSION

As mentioned in the results, no *eae* was detected in any of the sample isolates as expected, because *eae* is generally correlated to virulent *E. coli* (5). Since *E. coli* isolates were obtained from fecal samples of assumed healthy birds in the Trout Lake area, the isolates should be nonpathogenic. On the other hand, amplification of phylogenetic markers was expected as originally Clermont et al. then other groups (1, 2, 4) based phylogenetic grouping of *E. coli* on the presence of these marker(s). In Figure 1, amplification of these products was evident although additional products of 600 bp and 1500 bp were also unpredictably present in

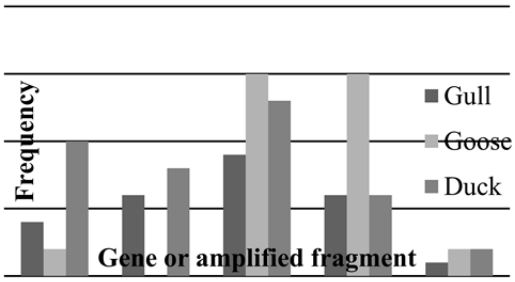

FIG. 2. Comparison of genes or fragments in gull, goose, and duck. Fifteen *E. coli* isolates from each of gull, goose, and duck were analyzed for presence of phylogenetic markers, *chuA*, *yjaA*, TspE4.C2, and other distinctive amplified products using multiplex colony PCR. Presence of each gene or fragment in each isolate was tallied and plotted. ND = none detected.

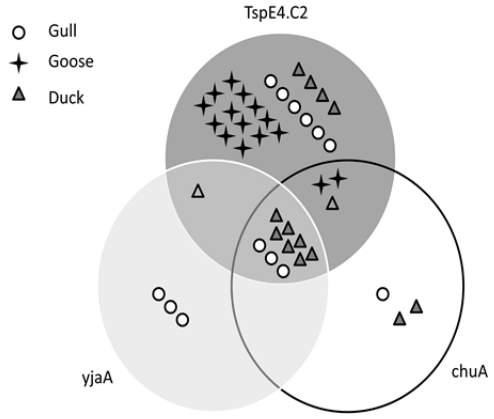

FIG. 3. Genotype comparison of *E. coli* isolates from gull, goose, and duck. Fifteen *E. coli* isolates from each of the three species were categorized according to whether they contain *chuA*, *yjaA*, and/or TspE4.C2. Note that two gull isolates did not contain any of the three genes or fragment.

some samples. These unknown products may indicate non-specific amplification due to the presence of genomic DNA in the samples containing similar sequences to any pair of the eight primers used in the multiplex PCR. However, none of the unknown products were observed in human *E. coli* isolates. Thus, the 600 bp and/or 1500 bp products deserve further investigation for their potential in differentiating human versus non-human hosts.

The presence of *chuA* and *yjaA* was claimed to be indicators for differentiating human versus non-human sources of *E. coli* (1). Of note, Carlos et al. only studied isolates from human, chicken, cow, goat, pig, and sheep (1). In this study, *E. coli* isolated from gull, goose, and duck were investigated to determine whether a correlation truly exists between the genotype of human isolates and non-human isolates, also between species. Surprisingly, 47% and 20% of duck and gull samples, respectively, contained all three phylogenetic markers which were previously denoted as present solely in *E. coli* isolated from human (1). Previous studies indicate that *E. coli* from birds predominately have only TspE4.C2 (1, 4). As gulls, geese, and ducks are more closely related to birds than human, a higher frequency of samples having solely TspE4.C2 was expected rather than the highly variable genotypes observed in this study.

Of the three species studied, gull samples contained the highest genotypic variation whereas goose isolates contained the least. Specifically, presence of TspE4.C2 and the 600 bp product and the absence of *yjaA* can positively identify isolates from goose. Duck samples were also identified when all three phylogenetic markers were present with an accuracy of 47%. Although three gull samples (20%) also showed this genotype, they may have arose due to variation in a particular gull since the three samples were isolates from that single gull. Nevertheless, if the presence of all three markers is to be used to discriminate between human and non-human sources, duck and human would not be differentiated. Interestingly, a combination of *yjaA* and TspE4.C2 was observed in a duck isolate. This combination was novel with respect to the phylogenetic subgroups for *E. coli* specified in previous work (1, 4). In summary, results from this study indicate that determination of *E. coli* host using phylogenetic markers proposed by Clermont did not yield absolute and accurate identification of all samples. This observation was also supported by research performed to compare the multi-locus sequence typing and Clermont's PCR triplex method which showed 85-90% and 80-85% positive classification rate, respectively (6).

As noted in the results, low intensity bands representing *chuA* and *yjaA* were observed in the no template control in the first of three independent experiments. This was most probably due to cross-contamination from the human *E. coli* isolates as they were extensively used in prior pilot PCR experiments. The negative control was clean (except for presence of a 40 bp band) in the remaining two experiments after thoroughly wiping the work area and micropipettes with diluted bleach. The 40 bp smear was present in all samples, including the no template control; hence, this probably represents the surplus deoxynucleotides in the reactions. In addition, the 65 bp band in the samples may correlate to incomplete amplification of products or hybridized primer-dimer products as interaction of the eight primers in each pair combination were not studied.

Overall, isolates from goose could be identified by the presence of TspE4.C2. However, the use of solely phylogenetic markers, *chuA*, *yjaA*, and TspE4.C2, to

differentiate between *E. coli* isolates from gull and duck were not definite. Although differences in the genotype of *E. coli* from the three species existed, variations within the species were also observed, especially in gull and duck. None of the isolates contained the virulence gene, *eae*. Thus, the use of PCR with simply three phylogenetic markers and *eae* to directly identify *E. coli* host needs further investigation to ensure accurate determination and robust application in MST.

FUTURE EXPERIMENTS

Since positive identification of gull and duck *E. coli* isolates were not plausible using multiplex PCR for amplification of *chuA*, *yjaA*, TspE4.C2, and *eae*, investigation for additional markers that could be included may aid in determination of the host. The 600 bp and 1500 bp products could serve as novel markers for host identification with further analysis as they were present in some isolates and not others in combination with phylogenetic markers. Sequencing the 600 bp and 1500 bp PCR products may provide information concerning their location in the *E. coli* genome, possibly annotations, and even strain-specific details. However, as additional primers may further complicate multiplex PCR, studying all the permutations of primer pair interactions with specific PCR parameters is recommended. Primer interactions for amplification of products should be equally effective and non-competitive. If the primers behave similarly in amplifying specific targets, then the resulting products truly represent the genotype of *E. coli* isolates studied. Furthermore, to increase statistical significance of the genotypic patterns observed, increased number of isolates from each of gull, goose, and duck could be considered.

ACKNOWLEDGEMENTS

This work was funded by the Department of Microbiology and Immunology, University of British Columbia. I am grateful to thank Dr. Ramey for his guidance and support. Special thanks to Dr. Wanyin Deng from Dr. B. Brett Finlay's lab at the Michael Smith Laboratories for their donation of EHEC O157:H7 strain EDL933 DNA. Also, I would like to thank Lando and Nick for providing clean glassware. Assistance from Clara, Karen, Lina, and Piriya was deeply appreciated.

REFERNCES

1. **Carlos, C., M.M. Pires, N.C. Stoppe, E.M. Hachich, M.I, Sato, T.A. Gomes, L.A. Amaral, and L.M. Ottoboni.** 2010. *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. BMC Microbiol. **10**:161-171.
2. **Clermont, O., S. Bonacorsi, and E. Bingen.** 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. Appl. Environ. Microbiol., **66**:4555-4558.
3. **Dombek, P.E., L.K. Johnson, S.T. Zimmerley and M.J. Sadowsky.** 2000. Use of repetitive DNA sequences and the PCR to differentiate Escherichia coli isolates from human and animal sources. Appl. Environ. Microbiol, **66**: 2572-2577.
4. **Escobar-Paramo, P., A. Le Menac'h, T. Le Gall, C. Amorin, S. Gouriou, B. Picard, D. Skurnik, and E. Denamur.** 2006. Identification of forces shaping the commensal Escherichia coli genetic structure by comparing animal and human isolates. Environ. Microbiol., **8**:1975-1984.
5. **Girao, D., V. Girao, K. Irino, and T. Gomes**. 2006. Classifying Escherichia coli. Emerg. Infect. Diseases. **12**:1297-1298.
6. **Gordon, D. M., O. Clermont, H. Tolley, E. Denamur.** 2008. Assigning *Escherichia coli* strains to phylogenetic groups: multi-locus sequence typing versus the PCR triplex method. Environ. Microbiol., **10**: 2484-2496.
7. **Lee, A. and E. Wong.** 2009. Optimization and the robustness of BOX A1R PCR for DNA fingerprinting using Trout Lake *E. coli* isolates. J. Exp. Microbiol. Immunol. **13**:104-113.
8. **Environment Canada.** September, 2010. Microbial source tracking in aquatic ecosystems: The state of the science and an assessment of needs. http://www.ec.gc.ca/inre-nwri/default.asp?lang=En&n=D575CDF5-1&offset=7&toc=show
9. **Yu, J. and J.B. Kaper.** 1992. Cloning and characterization of the *eae* gene of enterohaemorrhagic *Escherichia coli* O157:H7. Mol. Microbiol., **6**: 411-417.
10. **Zhang, W.L., B. Ko¨hler, E. Oswald, L. Beutin, H. Karch, S. Morabito, A. Caprioli, S. Suerbaum, and H. Schmidt.** 2002. Genetic diversity of intimin genes of attaching and effacing *Escherichia coli* strains J. Clin. Microbiol, **40**:4486–4492.
11. **Zhu, J.** 2010. Comparison of four template preparation methods and optimization of BOX A1R PCR for DNA fingerprinting of Escherichia coli isolates. J. Exp. Microbiol. Immunol. **14**:48-50.
12. **Heath Canada**. October, 2010. 4.0 Significance of *E. coli* in drinking water. http://www.hc-sc.gc.ca/ewh-semt/pubs/water-eau/escherichia_coli/significance-importance-eng.php#4