Molecular Characterization of Bacteriophages for Microbial Source Tracking in Korea

Jung Eun Lee,¹ Mi Young Lim,¹ Sei Yoon Kim,¹ Sunghee Lee,¹ Heetae Lee,¹ Hyun-Myung Oh², Hor-Gil Hur³, and GwangPyo Ko¹∗

Institute of Health and Environment, Department of Environmental Health, School of Public Health, Seoul National University, 28 Yeongeon-dong, Chongro-gu, Seoul, South Korea, 110-799¹; Department of Marine Sciences, College of Natural Sciences, Inha University, Yong-Hyun-Dong, Nam-Gu, Incheon, South Korea, 402-751²; Department of Environmental Science & Engineering, Gwangju Institute of Science and Technology, Gwangju, South Korea, 500-712³

*Corresponding author: Dr. GwangPyo Ko
Mailing address: Department of Environmental Health, School of Public Health, Seoul National University, 28 Yeongeon-dong, Chongro-gu, Seoul, South Korea, 110-799
E-mail: gko@snu.ac.kr
Phone: +82 2 3668 7881.
Fax: +82 2 745 9105.
Abstract

We investigated coliphages from various fecal sources, including humans and animals, for microbial source tracking in South Korea. Both somatic and F+ specific coliphages were isolated from 43 fecal samples from farms, wild animal habitats, and human wastewater plants. Somatic coliphages were more prevalent and abundant than F+ coliphages in all of tested fecal samples. We further characterized 311 F+ coliphage isolates using RNase sensitivity assays, (RT)-PCR, and nucleic acid sequencing. Phylogenetic analyses were performed based on the partial nucleic acid sequences of 311 F+ coliphages from various sources. F+ RNA coliphages were most prevalent among geese (95%), and were least prevalent in cow (5%). Among the genogroup of F+ RNA coliphages, most F+ coliphages isolated from animal fecal sources belonged to either group I or IV, and groups II and III from human wastewater sources. Some of the group I coliphages were present in both human and animal source samples. F+ RNA coliphages isolated from various sources were divided into two main clusters. All F+ RNA coliphages isolated from human wastewater were grouped with Qbeta-like phages, while phages isolated from most animal sources were grouped with MS2-like phages. UniFrac significance statistical analyses revealed significant differences between human and animal bacteriophages. In the principal component analysis (PCA), F+ RNA coliphages isolated from human waste were distinctively separate from other animal sources. However, F+ DNA coliphages were not significantly different, or separate in the PCA analysis. These results demonstrate that proper analysis of F+ RNA coliphages can effectively distinguish fecal sources.

Keywords: Microbial source tracking, somatic coliphage, F+ coliphage, genotyping, UniFrac significance
**Introduction**

Fecal contamination of various water resources poses a serious risk to human health through consumption of microorganisms that inhabit these environments, as well as through recreational exposure (20). To correctly assess and properly manage human health risk associated with water contamination, it is necessary to have information regarding the source of fecal contamination (12, 19, 30-32). In previous studies, several microbiological methods have been applied to distinguish fecal contamination sources, particularly between human and animal sources (20, 30). For example, different types of antibiotic resistance patterns (14), molecular markers (2), identified animal viruses (26), and whole genome patterns based on Rep-PCR or ribotyping (4, 28) have been applied to microbial source tracking (MST). However, these MST methods have not yet been fully evaluated and characterized.

Coliphages are viral indicators of fecal contamination in ground water proposed by the U.S. Environmental Protection Agency (U.S. EPA) (37) and could be useful target microorganisms for MST. Based on the infectivity through the host’s sex pili, coliphages can be categorized into somatic and F+ specific coliphages. In addition, coliphages can be categorized based on the type of nucleic acids (RNA vs. DNA). At present, F+ coliphages consist of *Leviviridae* (icosahedral, single-stranded RNA phages), including *Levivirus* and *Allolevivirus*, and *Inoviridae* (filamentous, single-stranded DNA phages) (5). *Levivirus* contain both MS2-like (group I) and GA-like viruses (group II), whereas *Allolevivirus* contain both Qbeta-like (group III) and SP-like (group IV) viruses(33). These subgroups of coliphages were initially classified based on serological typing (18). Different genotypes of F+ RNA coliphages are associated with different types of fecal sources (30, 32). A number of previous studies have reported that groups II and III are mainly isolated from human feces, and that groups I and IV are mainly associated with animal feces (10, 17, 18, 27). However, the specificity of this association can vary because group I of F+ RNA coliphages were isolated from human dominated municipal wastewater (13). In addition, the applicability of this MST method in different geographical
regions, such as Asia, is still unclear. Furthermore, the application of DNA coliphages to MST has not been much investigated.

In several previous studies, F+ coliphages were isolated and analyzed using culture and serological methods (7, 22, 34). Nucleic acid sequence-based analyses of F+ RNA specific genes have not been performed or are very limited (40). Stewart et al. (33) performed F+ RNA group III sequence analysis, and MST based on direct nucleic acid sequence was found to be more reliable than nucleic acid hybridization. However, a comprehensive analysis based on nucleic acid sequence has not been applied to coliphages. Thus, the objectives of this study were 1) to test the applicability of both RNA and DNA coliphage-based microbial source tracking in South Korea, and 2) to apply nucleic acid sequence-based analyses of both RNA and DNA coliphages isolated from various animal and human fecal sources to MST.

Materials and Methods

Collection of fecal and municipal wastewater samples

1) Animal feces

Thirty-two fecal samples from pigs, cows, chickens, and wild animals were collected from different animal farms or wild animal habitats. All of animal farms were located in rural areas near Seoul, Korea. And the sampling sites of wild animal habitats were located in Chonnam area, South Western part of South Korea where migrating birds from Siberia rest every winter. The fecal samples were collected using disposable polyvinyl bags, which were stored in polyethylene containers at 4°C until analysis. All animal fecal samples, except wild animal feces, were processed within 8 h of sampling and analyzed within 24 h. Since the sampling site for the wild animal feces was relatively far away, the wild animal feces was transported into the laboratory on dry ice, and then analyzed within 24 h.

2) Human wastewater
Municipal wastewater (containing human sewage and feces) samples were collected from the influents of sewage treatment plants in the metropolitan area of Seoul, South Korea. Since human feces and animal feces are treated separately at different waste treatment plants, we collected the samples from wastewater treatment plants only for human wastes. In total, 11 samples were collected from these wastewater treatment plants. Human wastewater samples were collected using autoclaved polyethylene bottles and stored at 4°C until analysis. Samples were analyzed within 24 h of collection.

**Recovery of somatic and male-specific bacteriophages**

Approximately 5 g of collected fecal samples were suspended in 20 ml phosphate buffered saline (PBS, pH=7.4) and vortexed for 5 min. The fecal suspensions were centrifuged at 5000 × g for 20 min, and the supernatant was carefully collected for further analysis. For human wastewater samples, each sample was centrifuged at 5000 × g for 20 min to remove large debris, and the supernatant was carefully collected for further analysis. The resulting supernatants were analyzed using single agar layer (SAL) methods (36). Briefly, to isolate somatic and male-specific (F+) coliphages from the supernatant, either *Escherichia coli* CN13 (ATCC #700609) or *E. coli* F<sub>amp</sub> (ATCC#700891) strains were used, respectively. From the plaques observed on the SAL plate, approximately 20 coliphage isolates were selected using a sterile wooden toothpick, and each isolate was subjected to two-step enrichment procedures (35). The sterile toothpick used in the plaque assay was suspended in 5 ml tryptic soy broth (TSB) containing 100 µl log phase of either *E. coli* CN13 or *E. coli* F<sub>amp</sub>. The tube was gently vortexed for 1 min, and the phage isolated from a single plaque was incubated at 37°C for 16-18 h at 150 rpm. After cultivation, chloroform was added, and the tube was vigorously vortexed for 5 min. The suspension was centrifuged at 5000 × g for 20 min, and the single phage isolate in the resulting supernatant was removed and stored at −70°C until further analysis. For each experiment, either MS2 (ATCC# 15597-B1) or phiX 174 (ATCC #13706-B1) was used as a positive control of somatic and male-specific coliphages, respectively. Sterile PBS was used as the negative control.

**RNase sensitivity assay to distinguish F+ DNA and F+ RNA coliphages**
To distinguish F+ DNA and F+ RNA coliphages, each isolated F+ coliphage was subjected to an RNase sensitivity assay (18). Escherichia coli F<sub>amp</sub> was cultivated in an exponential phase and poured into a 150-mm diameter Petri dish in 0.8% TSB agar with and without RNase (100 µg/ml) (Sigma-Aldrich, St. Louis, MO, USA). Serial dilutions (5 µl; undiluted, 10<sup>-2</sup>, 10<sup>-4</sup>, and 10<sup>-6</sup>) of the isolated coliphages were spotted onto both RNase present (+) and RNase negative (−) plates. The plates were incubated for 12–18 h at 37°C. When the plaque was observed in the presence of RNase, the phage was considered to be a DNA coliphage. When a plaque was only observed on the RNase (−) plate, the bacteriophage was considered to be an RNA coliphage.

**F+ DNA coliphage PCR and F+ RNA coliphage RT-PCR**

F+ coliphage strains were amplified using primers as previously described (38). Either DNA or RNA coliphages were heat-released and diluted 1:50. The released nucleic acid was amplified using either PCR or RT-PCR assays as described in a previous study (38). Amplified (RT)-PCR products were purified using the QIAquick PCR purification Kit (Qiagen, Valencia, CA, USA) and sequenced at Cosmo Gene Tech (Seoul, Korea).

**Genotyping of F+ coliphages and phylogenetic analysis**

Chromatographs of each sample sequence were checked and edited using SeqMan software (Lagergene 6.0, DNASTAR, Inc., Madison, WI, USA). Nucleic acid sequences with unclear chromatographs or shorter fragment were excluded from further data analysis. In addition, any possible clone from the same fecal sample was excluded as well. The nucleic acid sequences of novel F+RNA coliphage strains were deposited into GenBank (assigned GenBank accession number: GQ268004 to GQ268012). Each sequence of analyzed samples was exported to GenBank and confirmed using BLAST. All sequences with both F+ RNA and F+ DNA coliphages were imported into ARB library (25). In addition, all sequences of *Leviviridae* and *Inoviridae* (M13-like) available in GenBank were downloaded and imported into the ARB software package. Sequences were aligned, and the phylogenetic analysis was performed based on the neighbor-joining method.
Principal component analysis by UniFrac

To distinguish the nucleic acids of bacteriophages from both animal and human fecal sources, a principal component analysis (PCA) was performed using UniFrac software (23). The nucleic acid sequences of bacteriophages were aligned using Clustal W and exported to the UniFrac software for PCA. PCA was performed to determine whether bacteriophages from different fecal sources are distinctive using either UniFrac significance tests or the P-test. F+ RNA and F+ DNA coliphages were analyzed separately.

Results

Incidences of coliphages in animal feces and human wastewater

Table 1 summarizes the prevalence of somatic and F+ coliphages from various fecal sources in Korea. In most fecal sources, somatic coliphages occurred with a high incidence (86-100%). However, the prevalence of somatic coliphages in cow feces was lower than in other fecal sources (58%). In the fecal samples from a pig septic tank and from chickens, the concentrations of somatic coliphages ranged from 380 to 13,700 plaque-forming units (PFU)/g (median density of 680 PFU/g) and from 3 to 10,000 PFU/g (median density of 1,450 PFU/g), respectively. On the other hand, the concentrations of somatic coliphages in cow feces ranged from 1 to 25 PFU/g (median density of 11 PFU/g).

The prevalence of F+ coliphages was lower than that of somatic coliphages (Table 1). F+ coliphages were abundantly found (73-100%) in human wastewater, the pig septic tank, and in goose feces. On the other hand, F+ coliphages had a lower prevalence in cow feces (25%). F+ coliphages in human wastewater ranged from < 1 to 285 PFU/mL (median 5 PFU/mL). The median number of coliphages found in pig fecal samples was slightly higher than in other fecal samples (Table 1).

Identification of F+ DNA and F+ RNA coliphages by both RNase sensitivity and (RT)-PCR assays
To identify the type of nucleic acids (either DNA or RNA) in isolated coliphages, the sensitivity of 311 F+ coliphage isolates to RNase was tested using a spot-plate assay (Table 2). Among the 311 bacteriophage samples that were tested, the numbers of F+ DNA and F+ RNA coliphages were 157 and 154, respectively. F+ DNA coliphages were more prevalent among cow (95%) and chicken feces (79%). In both pig feces and the pig septic tank, F+ RNA coliphages were more prevalent (64% and 68%), and most bacteriophages from goose fecal samples were F+ RNA coliphages (95%). In the human wastewater, the prevalence of F+ DNA and F+ RNA coliphages was 59% and 41%, respectively.

**Typing of isolated F+ coliphage sequences**

In total, 311 nucleic acid sequences from isolated F+ coliphages were confirmed and classified using both BLAST searches and phylogenetic analyses. Based on comparisons to the standard reference strains, each of the 133 RNA coliphage isolates from various fecal sources was classified into one of four groups (groups I – IV). Over half (52%) of the RNA coliphages were categorized into group I, 10% in group II, 12% in group III, and 19% in group IV (Table 3). Most coliphages isolated from animals were categorized into groups I and IV, except one isolate from a goose fecal sample. Most RNA coliphages isolated from human wastewater samples were categorized into groups II (32%) and III (42%). Ten RNA coliphage isolates (26%) were categorized into group I. Approximately 9 isolates (7%) did not share any similarity with previously identified RNA coliphages and were classified into a novel group. Among the 131 F+ DNA coliphages, 109 isolates (83%) were identified as f1-like phages. Eighteen (13%) and four (3%) isolates were identified as fd-like and M13-like phages, respectively. Of the 64 F+ DNA coliphages isolated from human wastewater samples, 47 (73%) were identified as f1-like, 14 (22%) as fd-like, and 3 (5%) as M13-like bacteriophages. Most of the F+ DNA coliphages were grouped as either f1-like or fd-like bacteriophages.

**Phylogenetic analysis of F+ coliphages using ARB software**
To investigate the genetic diversity of isolated F+ coliphage sequences, phylogenetic analyses were performed using ARB software.

(i) F+ RNA coliphages

Partial regions of replicase genes [266 bp (Levivirus) and 229 bp (Allolevivirus)] were (RT) PCR-amplified and subsequently sequenced. Phylogenetic analyses were performed with all Leviriridae nucleic acid sequences from GenBank. In total, 133 isolated F+ RNA coliphages were grouped into two major clusters (Qβ-like and MS2-like). Among the isolated F+ RNA coliphages, the majority of the phages from human wastewater sources belonged to either GA-like (group II) or Qβ-like (group III) phages (Fig. 1A). The RNA coliphages isolated from animal fecal sources belonged to either MS2 (group I)- or FI (group IV)-like phages. Bacteriophages isolated from pig septic tank fecal sources were identified as a novel separate group.

(ii) F+ DNA coliphages

To investigate the genetic variability of F+ DNA coliphages, the nucleic acids of the gene IV region were sequenced (38). All F+ DNA coliphages from various sources and reference strains were categorized into two different major clusters (fd-like and M13-like). The majority of phages from human wastewater sources were identified as fd-like phages. In addition, one phage from chicken and 16 phages from swine septic tank sources were included as well in the fd-like phages. The phages from the other animal fecal sources, including chicken, cow, and pig, belonged to the M13 and (f1)-like phage groups (Fig. 1B).

PCA analysis of F+ coliphages by UniFrac

To determine differences among F+ coliphages isolated from different fecal sources, we performed significance tests and PCA based on the phylogenetic analysis. Genetic characteristics of F+ RNA bacteriophages isolated from human fecal samples were significantly different from the other samples based on the P-test (P-value < 0.05), except the cow samples (Table 4A). The significant P-values
indicated that F+ RNA coliphages isolated from various sources have unique genetic characteristics. Among the fecal samples from various sources, only the F+ RNA coliphages from humans were significantly different (P ≤ 0.001) (Table 4 B). These results indicated that the sequences from this sample are associated with more unique branch lengths in the tree, suggesting that F+ RNA coliphages from human sources were distinctive from F+ RNA coliphages isolated from other fecal sources. However, when F+ DNA coliphages were analyzed using the same significance analysis, the P-test result was not significant among the various source samples (data not shown). F+ RNA coliphage significance test observations might best be illustrated using PCA (Fig. 2 A). When a scatter plot of the first two principal components (PCs) was analyzed, both PC1 and PC2 explained the major variation in the data (50.9 and 24.2%, respectively). PC2 clearly separated all fecal sources from human source isolates (Fig. 2 A). However, all fecal samples from animal or human sources were not clearly separated using PCA analysis of F+ DNA coliphages (Fig. 2 B).

Discussion

Currently, little information is available for describing the genetic characteristics of both F+ and somatic coliphages for MST. This study is the first to apply nucleic acid sequence-based genotyping methods for both F+ DNA and RNA coliphages to MST. Bacteriophages have been considered useful targets in MST. However, MST is region specific and has not much been studied in Asia (11). Our study confirmed that F+ RNA coliphages isolated in Korea have similar genetic characteristics to those in other regions (30, 32). Genogroup I and IV were mostly found in animal fecal samples, and Genogroup II and III were mostly found in human dominated wastewater samples. These results coincide with other previous studies (10, 17, 27). However, 26% of F+ RNA coliphages isolated from human wastewater sample were identified Genogroup I. Thus, the identification of fecal sources solely on the genogroup I of F+ RNA coliphages would be limited (32).

F+ coliphages were commonly detected in human and animal fecal samples in this study (Table 1). These results are consistent with other studies. Previous studies reported that the prevalence of F+
coliphages was 79%, 60%, and 100% in human waste, goose, and swine fecal wastewater, respectively (3, 5). Concentrations of F+ coliphages in pig fecal samples were slightly higher than other animal sources such as chicken, cow, and goose. In previous studies, F+ RNA coliphages were rarely detected in fecal samples from adult chickens and cattle, but were more often detected in pig fecal samples (17). In this study, a much lower prevalence (25%) was only observed in cow fecal samples. Another study reported that F+ coliphages were either not detected or less commonly detected in cow sources (16). Different microbial communities may be present in the cow gut because of different digestion mechanisms and ingested food (1). Further study should be focused on the different characteristics of the microbial ecology in the cow gut and other animals with respect to diet and other factors.

The ecological mechanism by which F+ coliphages became prevalent in human wastewater and septic tanks, but not in individual feces, is not fully understood. In F+ coliphages, prolonged replication appears improbable because it would require a minimum host density to replicate (39). In both wastewater and septic tanks, host density could increase because feces from many different individuals were combined. The frequency of individual fecal samples containing F+ coliphages was not typically high (32). However, if any one individual fecal sample contains F+ coliphages within a wastewater sample, the F+ coliphages may be able to effectively replicate in the wastewater owing to the high density of the host population (32, 39).

The RNase test was applied to distinguish F+ RNA and F+ DNA coliphages. To confirm the results of RNase testing, either PCR or RT-PCR analysis for each of the bacteriophages isolates was also performed. These results indicated that both the RNase test and (RT)-PCR assays were generally consistent with each other. However, the results were not consistent for bacteriophage isolated from some human wastewater samples and animal fecal samples. In human and animal wastewater samples, several F+ DNA coliphages that were identified by spot assay were then identified as F+ RNA coliphages using RT-PCR analysis. These results suggest that the RNase sensitivity assay could be
inaccurate in some cases resulting from the inability to completely inactivate F+ RNA bacteriophages. Havelaar (15) found that F+ RNA coliphages could become resistant to RNase, which could lead to a false negative result. Hence, resistance to the RNase spot assay does not necessarily indicate the presence of F+ DNA coliphages (40). On the other hand, F+ RNA coliphages were identified as F+ DNA coliphages by PCR analysis in animal fecal samples. In this case, F+ DNA coliphages may be inactivated by RNase in some environments.

This study suggested that the ratio of F+ DNA to F+ RNA coliphages indicated the sources of the fecal samples. For example, F+ DNA coliphages were much more prevalent (95%) than F+ RNA coliphages (5%) in cow feces. On the other hand, F+ RNA coliphages were more prevalent (95%) in goose feces compared to F+ DNA coliphages (5%). These results are consistent with a previous study (5). Cole et al. (5) reported a high ratio of F+ DNA coliphages (82%) compared to F+ RNA coliphages (18%) in cow feces and a low ratio of F+ DNA (0%) to F+ RNA coliphages (100%) in goose feces. For human and pig fecal samples, a mixed ratio of F+ DNA and RNA coliphages was observed, consistent with a previous study (22). This study indicated that F+ DNA (59%) and F+ RNA coliphages (41%) were both present in human wastewater, and similarly, F+ DNA (32%) and F+ RNA coliphages (68%) were found together in pig fecal samples. A previous study also showed that F+ DNA (65%) and F+ RNA coliphages (35%) were present in human wastewater, and that F+ DNA (34%) and F+ RNA coliphages (66%) were prevalent in a hog lagoon (22).

In this study, F+ RNA coliphages were grouped based on the nucleic acid sequences of a partial replicase gene. As indicated in other studies, most F+ RNA coliphages from animal fecal samples belonged to either group I or IV, and the majority of F+ RNA coliphages isolated from human wastewater belonged to groups I, II, and III. Group I RNA+ coliphages were thus found in both human wastewater and animal fecal samples. Many previous studies have reported that group I coliphages were predominantly found in human dominated wastewater (5, 6, 32). Schaper et al. (29) found that group I coliphages were more persistent than other groups of coliphages under various
environmental conditions and stresses. This persistence would be an important factor in their presence in wastewater.

Few studies have investigated the prevalence and ecology of F+ DNA coliphages (9), and one study suggested F+ DNA coliphages as a potential indicator (22). Until now, little information has been available as to whether specific subgroups of F+ DNA coliphages are associated with specific fecal sources (38). Unlike the previous study that used the serotypes of F+ DNA coliphages (22), this study investigated F+ DNA coliphages as a potential target microorganism in MST using DNA sequence based genotypes. Typically, DNA is more stable than RNA, which suggests that F+ DNA coliphages could be better targets for MST (38). However, the results suggest that F+ RNA coliphages are more strongly associated with the ability to discriminate between human and animal sources than F+ DNA coliphages. Further studies should be conducted to investigate new types of F+ DNA bacteriophages and specific genetic markers to be used in MST.

Somatic coliphages were commonly found in both animal and human dominated wastewater samples. Concentrations of somatic coliphages varied from 30 to 3.8×10^3 PFU/ml in human dominated wastewater and from 1 to 1.4×10^4 PFU/g in animal fecal samples (Table 1). These observations are similar to those found in other studies (22, 32). A previous study reported 22 to 3.6×10^3 PFU/ml in wastewater, 3.8×10^2 to 1.4×10^4 PFU/g in hog lagoons, 1 to 2 PFU/g from cow sources, and 1×10^7 to 4.7×10^6 PFU/g from chicken litter (32). Because somatic coliphages are a more diverse group than F+ coliphages, it would be more difficult to identify target genetic markers for MST. However, if the right genetic marker were identified in somatic coliphages, they would be more attractive for MST because of their high prevalence in both individual feces and the environment.

Phylogenetic analyses on isolated bacteriophages were performed using the ARB software program (25). These analyses showed that F+ RNA coliphages isolated from various sources were divided into two main clusters (Fig. 1 A). All F+ RNA coliphages isolated from human wastewater were grouped...
with the Qbeta-like phage, while F+ RNA coliphages isolated from pig, chicken, and pig septic tank sources were grouped with the MS2-like phage. These results are encouraging because different phylogenetic characteristics are exhibited based on the types of human and animal fecal sources. The novel sequences were identified from collected animal fecal samples. However, when the phylogenetic analysis was performed for F+ DNA coliphages (Fig. 1 B), there was no distinction between human and animal fecal sources.

In this study, the UniFrac method was applied to distinguish between the fecal sources based on the phylogenetic analysis. UniFrac is a bioinformatics tool to test the statistical significance of phylogenetic distances among different samples. Lozupone et al. demonstrated that UniFrac is robust even for very similar samples when the sample size is large (24). Recently, several studies applied UniFrac to compare samples from multiple communities (8, 21, 23). However, this tool has never been applied to MST. The results of the P-test showed significant differences (< 0.005) between F+ RNA coliphages isolated from human and animal fecal samples. This result indicates that RNA coliphages could be applied in MST as target microorganisms. To determine whether bacteriophages isolated from various sources were different based on a unique branch length in the phylogenetic tree, the UniFrac significance test was performed (23). Among various fecal source samples, bacteriophages isolated from human wastewater were significantly different from other samples (P < 0.005; Table 4 B). This result suggests that F+ RNA coliphages isolated from human wastewater were clearly distinctive from other animal fecal samples. Furthermore, a PCA analysis confirmed these results. The PCA analysis described in Fig. 2A showed that F+ RNA coliphages isolated from human wastewater were clearly separated from other animal sources. Together, these results indicate that F+ RNA coliphages from human sources were significantly different from other fecal sources. However, when F+ DNA coliphage was analyzed, no distinctive pattern was observed (Fig. 2B). It also needs to be acknowledged that while phylogenetic analysis and PCA were able to distinguish F+ RNA coliphage isolates from known fecal sources, analysis of an environmental water sample receiving mixed fecal inputs in real world would be more complicated. In conclusion, our study demonstrated
that bacteriophages could be very useful targets in MST. Different criteria such as genogroup of bacteriophages and PCA analysis could be applied for identifying fecal sources in environmental samples.

Acknowledgments

This study was supported by the Ministry of Environment of the Republic of Korea as "The Eco-Technopia 21 Project" (900-20080010).

References


Table 1. Incidence of somatic and F+ coliphages in animal feces and human wastewater in Korea.

<table>
<thead>
<tr>
<th>Source</th>
<th>No. of samples</th>
<th>Somatic coliphages</th>
<th>F+ coliphages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive %</td>
<td>Median(^a)</td>
</tr>
<tr>
<td>Human</td>
<td>11</td>
<td>91 %</td>
<td>310</td>
</tr>
<tr>
<td>Cow</td>
<td>12</td>
<td>58 %</td>
<td>11</td>
</tr>
<tr>
<td>Chicken</td>
<td>7</td>
<td>86 %</td>
<td>1450</td>
</tr>
<tr>
<td>Goose</td>
<td>2</td>
<td>100 %</td>
<td>10</td>
</tr>
<tr>
<td>Pig</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) feces</td>
<td>8</td>
<td>86 %</td>
<td>500</td>
</tr>
<tr>
<td>2) septic tank</td>
<td>3</td>
<td>100 %</td>
<td>680</td>
</tr>
</tbody>
</table>

\(^a\) Plaque-forming unit (PFU) / g of stool or ml of wastewater

\(^b\)
Table 2. F+ coliphage characterization by RNase test and PCR or RT-PCR assay.

<table>
<thead>
<tr>
<th>Source</th>
<th>No. isolates</th>
<th>RNase test a N (%)</th>
<th>Confirmed by PCR or RT-PCR b N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>F+ DNA</td>
<td>F+ RNA</td>
</tr>
<tr>
<td>Human</td>
<td>128</td>
<td>78 (61%)</td>
<td>50 (39%)</td>
</tr>
<tr>
<td>Cow</td>
<td>21</td>
<td>17 (81%)</td>
<td>4 (19%)</td>
</tr>
<tr>
<td>Chicken</td>
<td>29</td>
<td>22 (76%)</td>
<td>7 (24%)</td>
</tr>
<tr>
<td>Goose</td>
<td>21</td>
<td>1 (5%)</td>
<td>20 (95%)</td>
</tr>
<tr>
<td>Pig</td>
<td>53</td>
<td>15 (28%)</td>
<td>38 (72%)</td>
</tr>
<tr>
<td>1) feces</td>
<td>59</td>
<td>21 (36%)</td>
<td>38 (64%)</td>
</tr>
<tr>
<td>2) septic tank</td>
<td>59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>311</td>
<td>154 (50%)</td>
<td>157 (50%)</td>
</tr>
</tbody>
</table>

a RNase test using ribonuclease A for inhibition of RNA coliphages.
b PCR for F+ DNA coliphage confirmation. Non-detected sample from PCR assay was tested by RT-PCR for F+ RNA coliphage detection.
c RT-PCR for F+ RNA coliphage confirmation. Non-detected sample from RT-PCR assay was tested by PCR for F+ DNA coliphage detection.
Table 3. Typing of F+ coliphages by BLAST search and phylogenetic analysis

<table>
<thead>
<tr>
<th>Source</th>
<th>Group I N (%)</th>
<th>Group II N (%)</th>
<th>Group III N (%)</th>
<th>Group IV N (%)</th>
<th>Novel Sequence</th>
<th>Total</th>
<th>F+ DNA coliphages N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F+ RNA</td>
<td>F+ DNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>coliphages</td>
<td>coliphages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td>10 (26%)</td>
<td>16 (42%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>38</td>
<td>14 (22%) 3 (5%) 64</td>
</tr>
<tr>
<td>Cow</td>
<td>1 (100%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>20 (100%) - - 20</td>
</tr>
<tr>
<td>Chicken</td>
<td>-</td>
<td>-</td>
<td>5 (100%)</td>
<td>-</td>
<td>5</td>
<td>5</td>
<td>11 (92%) 1 (8%) 12</td>
</tr>
<tr>
<td>Goose</td>
<td>18 (90%)</td>
<td>1 (5%)</td>
<td>-</td>
<td>1 (5%)</td>
<td>20</td>
<td>20</td>
<td>- - - -</td>
</tr>
<tr>
<td>Pig</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) feces</td>
<td>33 (100%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>33</td>
<td>14</td>
<td>3 (18%) 1 (6%) 17</td>
</tr>
<tr>
<td>2) septic tank</td>
<td>7 (19%)</td>
<td>-</td>
<td>20 (56%)</td>
<td>9 (25%)</td>
<td>36</td>
<td>17</td>
<td>- 1 (6%) 18</td>
</tr>
<tr>
<td>Total</td>
<td>69 (52%)</td>
<td>13 (10%)</td>
<td>16 (12%)</td>
<td>26 (19%)</td>
<td>9 (7%)</td>
<td>133</td>
<td>109 (83%) 18 (13%) 4 (3%) 131</td>
</tr>
</tbody>
</table>

*E* value for nucleic acid sequence search was > 1e-30 in this study.
Table 4. P-test (A) and UniFrac significance (B) of F+ RNA coliphages.

(A) P-test

<table>
<thead>
<tr>
<th>No.</th>
<th>Sample</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Human</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Cow</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Chicken</td>
<td>≤ 0.015*</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Goose</td>
<td>≤ 0.015*</td>
<td>1.000</td>
<td>≤ 0.015*</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Pig 1) feces</td>
<td>≤ 0.015*</td>
<td>1.000</td>
<td>≤ 0.015*</td>
<td>≤ 0.015*</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2) septic tank</td>
<td>≤ 0.015*</td>
<td>1.000</td>
<td>≤ 0.015*</td>
<td>≤ 0.015*</td>
<td>≤ 0.015*</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*P-test P-values were based on comparisons of 1,000 randomized trees. Results listed are the P-values.

*P-values were < 0.05, and pairwise comparisons indicated that the F+ RNA coliphages from various sources were significantly different.
(B) UniFrac significance test

<table>
<thead>
<tr>
<th>Sample</th>
<th>P-values*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>≤ 0.001*</td>
</tr>
<tr>
<td>Cow</td>
<td>0.05</td>
</tr>
<tr>
<td>Chicken</td>
<td>0.139</td>
</tr>
<tr>
<td>Goose</td>
<td>0.728</td>
</tr>
<tr>
<td>Pig 1) feces</td>
<td>0.983</td>
</tr>
<tr>
<td>2) septic tank</td>
<td>0.936</td>
</tr>
</tbody>
</table>

*UniFrac P-values were based on comparisons of 1,000 randomized trees. Results listed are the P-values. *P-value is < 0.05 and pairwise comparisons indicated that the sequences from this sample are associated with more unique branch lengths than would be expected.
Figure 1. Neighbor-joining trees of F+ RNA (A) and F+ DNA (B) coliphage sequences using the ARB software package. The scale bar corresponds to a 10% estimated difference in nucleotide sequence positions. ‘Pig’ indicates pig feces and ‘swine’ indicates samples from a pig septic tank (see Table 1).

A. 

![Diagram of Neighbor-joining trees of F+ RNA and F+ DNA coliphage sequences using ARB software package.](image-url)
Figure 2. Principal component analysis (PCA) of F+ RNA (A) and F+ DNA (B) coliphage sequences.