

Chapter 4 Molecular Genetic Pathogen Analysis (Microbial Source Tracking)

Research Task

The principal objectives of this task are: (i) to determine the sources of *E. coli* bacteria and their ribotype from animal populations in the study watersheds [including wildlife (birds, deer), domestic animal (horses, cattle, birds, etc.) and human (septic) sources]; (ii) to study the occurrence of pathogens (enteric bacteria, *Giardia*, and *Cryptosporidium*) in the source populations (as above) inhabiting the study area; and (iii) to use source data concerning level of pathogens associated with each of the sources of microbial pollution, in combination with *E. coli* type data from each station, to predict occurrence of pathogens in the stream.

Methods

Through consultation with NYC DEP, four integrative sampling stations were selected for our sampling effort (Figure 1.3). The four sites represent the outflow water of the (1) New Croton Reservoir [NYC DEP sampling site CRO (GH)], (2) collective Delaware reservoirs (Cannonsville, Pepacton, Neversink, Rondout collected at NYC DEP sampling site DEL (17)), (3) collective Catskill reservoirs (Ashokan, Schoharie) collected at NYC DEP sampling site CAT (ALUM), and (4) Cannonsville watershed (collected at NYC DEP sampling site WDBN - Beerston).

A total of 234 water samples were collected by NYC DEP staff during the course of their normal sampling operations at the designated stations between June and November, 2000. The water samples were filtered and plated prior to being sent to Dr. Mansour Samadpour at the Institute of Environmental Health (IEH) for ribosomal RNA typing using IEH's existing ribotyped *E.coli* database. Water samples taken by NYC DEP personnel adhered to the following protocols: NYC DEP Doc. Control Nos. 02005-d and 02130-a, and Method 04026-e.

Fresh animal fecal samples were collected in a fecal coliform free manner from the study watersheds and placed into fecal coliform free containers and shipped to the Institute of Environmental Health (Dr. Samadpour's laboratory) by overnight mail, on ice. Animal fecal samples were only collected when they were positively identified as belonging to a given animal species. No more than three samples were collected from the members of the same animal species from a given location. Only a single sample was collected from an individual animal. All containers and handling procedures were fecal coliform free. Samples were not to be stored for more than 4 days prior to shipping.

A total of 57 source samples (primarily domesticated animals and waste-water treatment plant effluent) were sent to the IEH to be analyzed for the presence of *Giardia*, and *Cryptosporidium* using enzyme linked immunosorbent assay. The effluent samples, however, were not neutralized prior to being shipped to IEH and were therefore not analyzed. The samples were also analyzed for the presence of a selected list of bacteria enteric pathogens (*Salmonella*, *Shigella*, pathogenic *E. coli*, *Yersina*, *Listeria*, *Helicobacter*, *Campylobacter*), using polymerase chain reaction based detection methods. *Giardia* and *Cryptosporidium* analysis will be performed only on fecal samples using the ELISA method. The method for *Giardia* has been published (Stibbs et al 1990). The method for *Cryptosporidium* has been published in the form of a thesis (J. Beeh, 1992, University of Washington, M.S. Thesis). The other pathogen screening is done by PCR methods. All the primers used have been published and validated with every experiment containing internal positive and negative controls to further validate the results.

Isolation and purification of *E. coli* strains from water and fecal samples involved the following methods. Water samples were received in the form of mFC plates, fecal samples arrived in specimen containers. Fecal samples were plated on MacConkey agar and incubated at 35 °C, overnight. The next day 3-5 lactose fermenting, non-mucoid colonies were picked and replated on MacConkey agar for purification. Five non-mucoid blue colonies were picked from mFC plates corresponding to each water sample, and plated on MacConkey agar for purification. A single well isolated non-mucoid colony was picked from each MacConkey plate and plated on Tryptic Soy Agar. After overnight incubation at 35°C, each culture was tested by Spot indol test using appropriate positive and negative controls. Indol positive cultures were further tested for the ability to utilize citrate using the Simmon's citrate agar. Indol positive, citrate negative colonies were identified as *E. coli* and were given isolate numbers.

A portion of each *E. coli* strain isolated from the samples was stored at -80°C, in nutrient broth plus 15% glycerol in order to grow pure cultures of *E. coli* for freezing (long term storage), and isolation of DNA. Restriction enzyme digestion and Agarose gel electrophoresis DNA samples was then conducted as follows. Genomic DNA was isolated from each *E. coli* strain using a standard protocol. All reagents and buffers were made according to standard protocol. Reagents and buffers were tested for sterility. Every batch of restriction enzyme reaction contained two reactions with a positive control strain which was included on two lanes on each gel. Agarose gel electrophoresis was conducted under standard conditions. Agarose gel concentration, and volume, buffer straight, pH, mA, V, and electrophoresis time were controlled for during each electrophoresis. Each agarose gel was assigned a number, and when more than one gel was run, the position of the first standard reference strain was changed in each gel (1st lane on the first gel, to the nth lane on the nth gel). After electrophoresis, gels were stained in ethidium bromide, Each of the two gels were stained in a single container and, of the two gels placed in the

same container, one corner of the gel with the higher number was clipped with the labels for each gel transferred to the staining container. Each gel was then photographed and a hard copy of the print was labeled with the gel sheet (containing the isolates numbers loaded on each lane, and the enzyme used to cut the DNA, plus date, gel number, voltage, mA, gel strength, buffer strength, and electrophoresis time information).

Southern blot hybridization using radio labeled cDNA probe for rRNA genes was performed according to standard protocol. After photography each gel is returned to the same staining container. Gels are denatured for Southern blotting in the same container. Each blotting apparatus was set in a separate container which was labeled with the gel number. Each membrane filter was labeled with the gel number, restriction enzyme designation, date, and technician's initials.

Upon arrival an agues extract of each fecal sample were prepared and frozen for future analysis by ELISA for Giardia an Cryptosporidium soluble antigens. Fecal samples were also fixed onto two slides (per sample) for direct immunofluorescent antibody staining analysis for the presence of Giardia cysts and Cryptosporidium oocysts. The inclusion of the antibody direct immunofluorescent (IFA) analysis is for the purpose of minimizing the risk of having false negatives by ELISA

QA/QC

The IEH laboratory participated in a laboratory QC effort involving the identification of reference isolate materials provided by an outside laboratory. Forty separate isolates (Table 4.1), representing a number of wildlife, farm and domestic animal, and human sources, were analyzed by the IEH laboratory during the course of the NY watersheds project sampling effort this past summer. The forty separate isolates were sent in triplicate for a total of 120 isolates. IEH laboratory personnel were blind to the isolate sample source information. All isolates were accurately identified, with each of the three copies also accurately identified. Therefore, both laboratory accuracy and precision (reproducibility) were confirmed through the analyses of these laboratory sample reference materials.

Results and Discussion

E. coli data (isolates)

A total of 489 isolates were separated from the 234 sample plates. Over half of these isolates (285) were for the Del(17) site, 83 were from the Cat(Alum) site, 97 from the Cro(GH) site, with the remaining 24 from the WDBN site. A majority (310) of the isolates were identified as coming from a particular animal or human source (Table 4.2). The identified isolate sources shown in Table 4.2 were re-categorized into one of five assumed groups; avian, domestic animal, farm animal, human, and wildlife. The distribution of these five re-categorized isolate source

groups, relative to the total number of isolates, suggests that bird species make up a large proportion of the *E. coli* sources at the four integrative sites, followed by farm animals and humans (Figure 4.1). Summary data of the fecal coliform counts taken from the plates sent to IEH are provided in Table 4.3. Fecal coliform count data were not collected for the WDBN site.

No obvious relationships were found between the percentage of impervious area or agriculture/grass area and the three human-influenced (human, domestic animals, and farm animals) isolate source groups (Figure 4.2). Grass area was lumped with agricultural land use because pasture, in addition to grass and hay fields, is included within the grass land-use category. Interestingly, the largest percentage of farm-identified isolate sources [Cat(Alum) site] corresponded to the second lowest percentage of agriculture/grass area. An apparent relationship was revealed between the percentage of water and the avian isolate source group (Figure 4.2). The Croton Reservoir watershed [Cro(GH)] had the highest percentage of avian-identified isolate sources and the highest percentage of water. Conversely, the W. Br. Delaware River watershed (WDBN) had the lowest percentages for both the avian-identified isolates and water. It should be noted that the avian category could include farm species such as chickens. Overall, these results should be viewed with caution given the large area represented by each microbial sampling site and the uneven distribution of samples collected at each site.

Pathogen data

Giardia and *Cryptosporidium* were found in samples collected from at least two of the four reservoir watersheds (Table 4.4) with most of the occurrences coming from dogs (Table 4.5). Note that a majority of the 45 non-effluent samples collected came from dogs (26). *Giardia* was found in the single wildlife (Woodchuck) sample collected from the study area. *Shigella spp.* were the only pathogen species not detected in at least one of the samples. *Listeria spp.* had the highest number of occurrences (15) of the given pathogens.

The correspondence between a pathogen source and a source of microbial pollution as identified from the *E. coli* source sampling is provided in Table 4.6. The association between pathogen source and identified *E. coli* sources was limited to the fecal matter sampling effort conducted in the four reservoir watersheds listed in Table 4.4. This limited dataset demonstrates that pathogenic bacteria are present in several different animal hosts, and that these hosts are contributing to the microbial pollution in the four reservoirs from which the data were gathered. However, a great deal of caution should be exercised in interpreting these data given the number of assumptions made in summarizing these data as outlined in Table 4.6 and the limited number of fecal matter samples that were collected in this first year of the project.

Implications of the results for the second year of the project.

The data from this first year of sampling provided some interesting insight into the potential dominate source of microbial sources of contamination, namely avian species. However, improvements to the sampling scheme can be made to better identify the sources, at least to within a major reservoir watershed, of the microbial contaminants. Furthermore, QA/QC efforts have to improve to ensure quality of the data. Therefore, the following changes will be made:

(1) Fecal matter sampling and consequently pathogen analyses, will be dropped in the second year, primarily due to financial considerations. No loss in *E. coli* source sampling (i.e. fecal matter sampling) will occur though, since Dr. Mansour Samadpour is currently involved in a number of east coast microbial source tracking studies (including NY state).

(2) Two of the four, first year sampling sites [Del(17) and Cat(Alum)] will be dropped, and replaced by five reservoir effluent sites in the Delaware/Catskill system. These new sites will be (i) CRR2 or WDT0 for the Cannonsville Res., (ii) PRR2 for the Pepacton Res., (iii) NRR2 for the Neversink Res. (iv) EAR for the Ashokan (Schoharie) Res., and (v) SRR1 or SRR2 for the Schoharie Res. Cro(GH) will continue as a sampling site in order to maintain an EOH site and along with the WDBN site, will provide continuity in the sampling effort across all three years of sampling in phase I of the project.

(3) Coordination in sampling effort between SWRC and the NYC DEP laboratories has to improve. Sampling effort was not evenly distributed across the four sites and sample tracking (e.g. chain of custody) was inadequate. Better direction from SWRC in terms of sampling effort, sample tracking, etc. will be provided to the NYC DEP laboratories involved in the project before the beginning of the year two sampling effort.

Literature Cited

Stibbs, H.H., Samadpour, M., and J. E. Ongerth. 1990. Identification of *Giardia lamblia*-specific antigens in infected human and gerbil feces by western immunoblotting. *Journal of Clinical microbiology* 28:2340-2346.

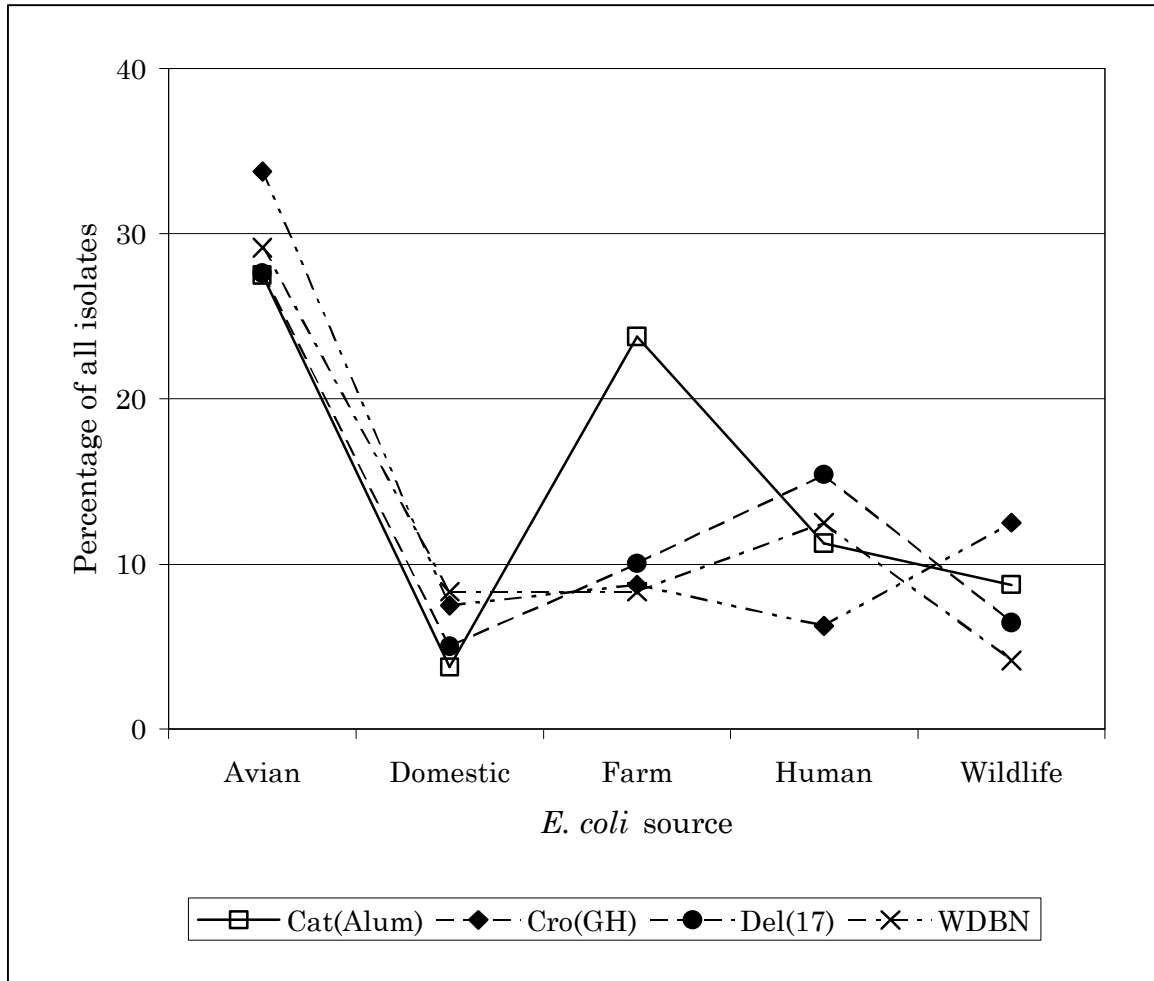


Figure 4.1 Course grouping of *E. coli* source types as a percentage of all isolate samples (not just the ribotyped isolates) for the four microbial integrative sampling sites.

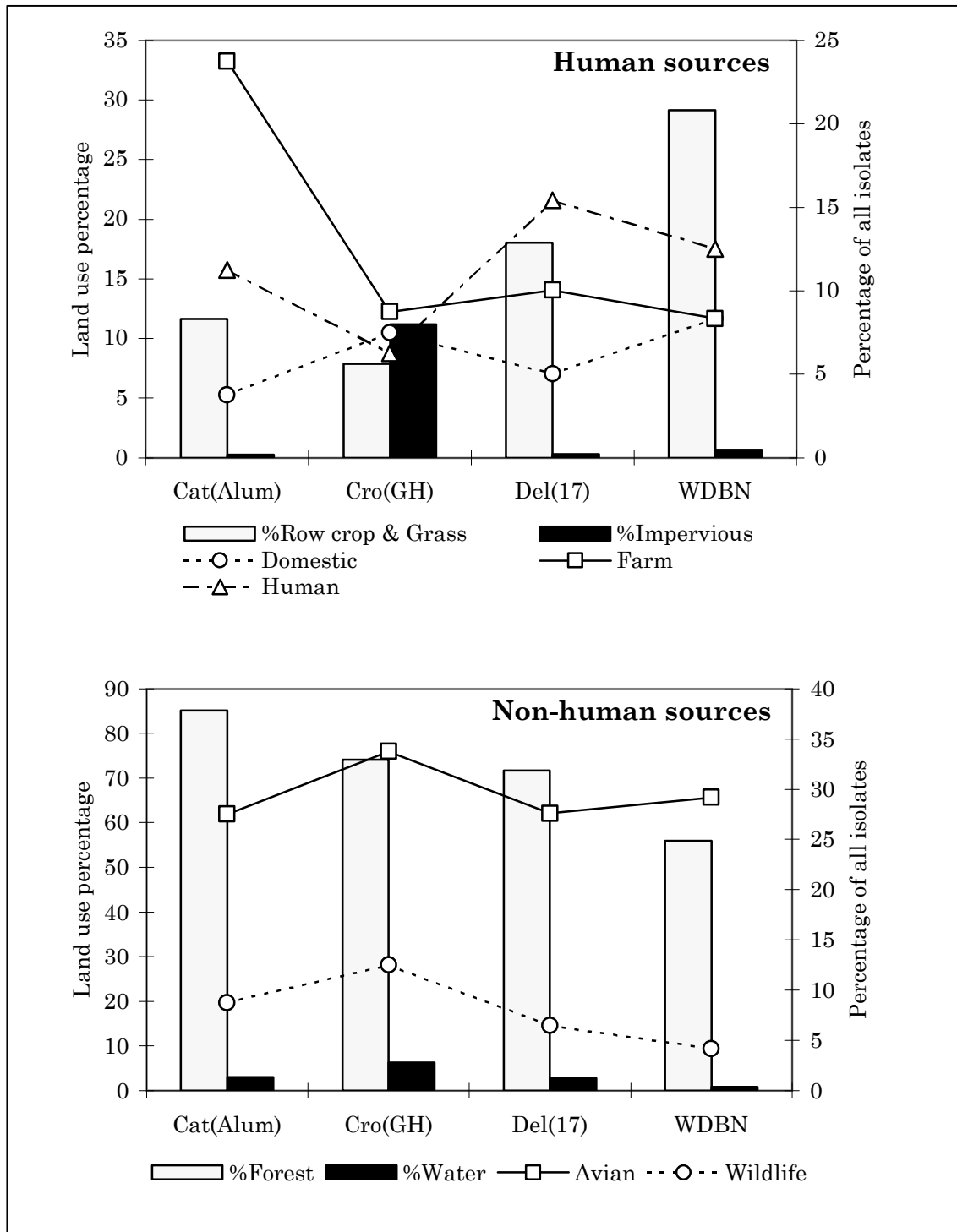


Figure 4.2. Percentages of isolate sources (assumed groupings – corresponding to groupings presented in Figure 4.1) and land use percentages separated into human isolate sources (top) and non-human isolate sources (bottom) for the four microbial integrative sampling sites.

Table 4.1. *E. coli* source types included as laboratory reference samples. Source types having the same microbial clonal type number indicate genetically similar *E. coli* ribotypes. A blank entry under the database match column indicates that the given ribotype was not part of an existing IEH *E. coli* ribotype database.

Source Type	Microbial clonal type	Database used to match source types.
	1	IEH existing database
Horse	10	
Cat	11	
Human	12	IEH existing database
Human	13	IEH existing database
Sea Lion	14	
Sea Lion	14	
Sea Lion	14	
Sea Lion	14	
Sea Lion	14	
Sea Lion	14	
	15	
Sea Gull	16	California subset of IEH database
Human	17	IEH existing database
Sea Gull	18	IEH existing database
Sea Gull	18	IEH existing database
Bovine	19	IEH existing database
Bovine	19	IEH existing database
Sea Gull	2	IEH existing database
Human	20	IEH existing database
Human	21	California subset of IEH database
Sea Gull	22	
Bovine	23	
Bovine	24	
Sea Lion	25	
Sea Gull	26	IEH existing database
Horse	27	
Bovine	3	IEH existing database
Sea Gull	4	IEH existing database
Sea Lion	5	
Sea Lion	5	
Sea Lion	5	
Sea Lion	5	
Bovine	6	IEH existing database
Sea Gull	7	California subset of IEH database
Bovine	8	IEH existing database
Bovine	8	IEH existing database
Bovine	8	IEH existing database
Human	9	IEH existing database

Table 4.2. Isolate sample sources, broken down by identified source (including non-identified isolates), for water samples collected at the four integrative microbial sampling sites.

Isolate Source ¹	Course grouping of isolate sources (see Figure 4.1)	NYC DEP microbial sampling sites				Row Total
		Cat(Alum)	Cro(GH)	Del(17)	WDBN	
Avian	Avian	11	12	40	5	68
Avian (D-Gu)	Avian	1	2	4		7
Avian (G-G)	Avian		2	4		6
Avian (G-Gu)	Avian			1		1
Bear	Wildlife			1		1
Bovine	Farm animal	17	6	17	2	42
Canine	Domestic animal	1	1	3		5
Cat	Domestic animal			2		2
Crow	Avian		1	1		2
Deer	Wildlife	4	6	5	1	16
Dog	Domestic animal	2	2	7	2	13
Duck	Avian			1		1
Duck-Goose	Avian	6	4	9		19
Feline	Domestic animal		3	2		5
Fox	Wildlife	1	1			2
Goose	Avian	2	2	9	2	15
Horse	Farm animal	2		9		11
Human	Human	9	5	43	3	60
Opossum	Wildlife			2		2
Pig	Farm animal		1	1		2
Rabbit	Wildlife		1	3		4
Raccoon	Wildlife	1		3		4
Rodent	Wildlife	1	2	3		6
Sea Gull	Avian	2	4	8		14
Sheep	Farm animal			1		1
Squirrel	Wildlife			1		1
Not identified		23	42	105	9	179
Column Total		83	97	285	24	489

¹ Source identification by IEH.

Table 4.3. Summary of the fecal coliform counts for samples collected at three of the four integrative sampling sites (NYC DEP pathogen sampling sites).

NYC DEP sampling site	Fecal coliform counts (fc/100ml)					nos of samples
	min	25 th percentile	median	75 th percentile	max	
Cat(Alum)	<1	1	1	3	31	81
Cro(GH)	<1	1	1	2	4	74
Del(17)	<1	4	7	13	76	281
WDBN	-	-	-	-	-	-

Table 4.4. Pathogen occurrences (selected species) by WOH reservoir watershed.

Pathogenic species	Detected	Reservoir Watershed				Spp. totals
		Ashokan	Cannonsville	Pepacton	Schoharie	
<i>Campylobacter</i>	Negative	4	16	7	7	34
	Positive	2	5	1	3	11
<i>Cryptosporidium</i>	Negative	6	16	8	9	39
	Positive		5		1	6
<i>E. coli</i>	Negative	4	17	8	8	37
	Positive	2	4		2	8
<i>Giardia</i>	Negative	5	15	8	9	37
	Positive	1	6		1	8
<i>H. pylori</i>	Negative	5	20	8	10	43
	Positive	1	1			2
<i>Listeria spp.</i>	Negative	3	14	6	7	30
	Positive	3	7	2	3	15
<i>Salmonella spp.</i>	Negative	6	14	7	6	33
	Positive		7	1	4	12
<i>Shigella spp.</i>	Negative	6	21	8	10	45
	Positive					0
<i>Vibrio vulnificus</i>	Negative	6	21	7	10	44
	Positive			1		1
<i>V. parahemolyticus</i>	Negative	6	21	6	10	43
	Positive			2		2
<i>Yersinia enterocolitica</i>	Negative	5	17	8	10	40
	Positive	1	4			5
Watershed totals	Negative	56	192	81	96	425
	Positive	10	39	7	14	70

Table 4.5. Pathogen occurrences (selected species) by source (host).

Pathogenic species	Detected	Source (host)							Spp. totals
		Cat	Chicken	Dog	Rabbit	Rat	Horse	Wood chuck	
<i>Campylobacter</i>	Negative	7	2	21	2	2			34
	Positive	2	1	5	1		1	1	11
<i>Cryptosporidium</i>	Negative	6	3	24	2	2	1	1	39
	Positive	3		2	1				6
<i>E. coli</i>	Negative	8	2	21	2	2	1	1	37
	Positive	1	1	5	1				8
<i>Giardia</i>	Negative	8	3	21	3	1	1		37
	Positive	1		5		1		1	8
<i>H. pylori</i>	Negative	9	3	24	3	2	1	1	43
	Positive			2					2
<i>Listeria spp.</i>	Negative	5	1	20	1	1	1	1	30
	Positive	4	2	6	2	1			15
<i>Salmonella spp.</i>	Negative	7	2	22	2				33
	Positive	2	1	4	1	2	1	1	12
<i>Shigella spp.</i>	Negative	9	3	26	3	2	1	1	45
	Positive								0
<i>Vibrio vulnificus</i>	Negative	8	3	26	3	2	1	1	44
	Positive	1							1
<i>V. parahemolyticus</i>	Negative	9	3	24	3	2	1	1	43
	Positive			2					2
<i>Yersinia enterocolitica</i>	Negative	9	3	23	2	1	1	1	40
	Positive			3	1	1			5
Source (host) totals	Negative	85	28	252	26	17	9	8	425
	Positive	14	5	34	7	5	2	3	70

Table 4.6. Microbial source (host) groups identified as both a source of the selected pathogenic species found through the fecal matter sampling, and an *E. coli* source identified through the isolate source sampling. The microbial source groups pertain to samples collected from the Cat(Alum) [Schoharie and Ashokan Reservoirs and associated watersheds] and Del(17) [Cannonsville and Pepacton Reservoirs watersheds, only, for the pathogen source sampling] sampling sites.

Pathogenic species	Cat(Alum)			Del(17) ¹			
	AVIAN ²	DOG ³	HORSE	CAT ⁴	DOG ³	RABBIT	RODENT ⁵
<i>Campylobacter</i>	1	3	1	2	2	1	
<i>Cryptosporidium</i>				2	2	1	
<i>E. coli</i>	1	2			3	1	
<i>Giardia</i>		2		1	2		1
<i>H. pylori</i>		1			1		
<i>Listeria spp.</i>	2	2		2	3	2	1
<i>Salmonella spp.</i>	1	1	1	1	3	1	2
<i>Vibrio vulnificus</i>				1			
<i>V. parahemolyticus</i>					2		
<i>Yersinia enterocolitica</i>		1			2	1	1

¹ Fecal matter samples for pathogen source identification were only taken from within the Cannonsville and Pepacton Reservoir watersheds, however, sources of the *E. coli* isolates identified from water samples taken at the Del(17) site could have also originated from the Rondout or Neversink Reservoir watersheds.

² The AVIAN category includes those isolate sources identified strictly as “Avian” in the “Isolate Source” column in Table 4.2 and the pathogen occurrences from the “Chicken” source in Table 4.5.

³ The DOG category includes those isolate sources identified as either “Dog” or “Canine” in the “Isolate Source” column of Table 4.2.

⁴ The CAT category includes those isolate sources identified as either “Cat” or “Feline” in the “Isolate Source” column of Table 4.2.

⁵ The RODENT category includes those pathogen occurrences from the “Rat” source in Table 4.5.