APPENDIX B: *ESCHERICHIA COLI* RIBOTYPING AND WATER QUALITY MONITORING FOR THE COLCHESTER, VERMONT IWRMP, FINAL REPORT, 2ND YEAR, 2010 FIELD SEASON

Escherichia coli Ribotyping and Water Quality Monitoring for the Colchester, Vermont IWRMP

Final Report
2^{nd} Year (2010 Field Season)

July 2011

Dr. Stephen H. Jones
University of New Hampshire
Jackson Estuarine Laboratory
Durham, NH 03824

INTRODUCTION

One of the most common issues facing environmental managers concerned with surface water quality is fecal-borne microbial contamination and the threat of diseases to humans who come in contact with contaminated water or shellfish. For purposes of monitoring the sanitary quality of surface waters, fecal coliforms, enterococci and *Escherichia coli* (*E. coli*) have traditionally served as indicators of water quality for classifying waters to protect public health. However, as untreated sewage from inadequately designed wastewater treatment facilities has been eliminated or reduced in significance, the residual contamination that limits uses of surface waters is typically of unknown origin. Efforts to reduce contamination have often revolved around making a best guess of what potential sources may be significant, conducting extensive sampling programs, eliminating sources and then re-sampling surface waters to see if improvements in water quality have occurred. This process is expensive and oftentimes less fruitful than desired.

Recent adoption of biotechnological techniques for application to water quality issues has spawned a number of approaches to address identification of sources of fecal-borne contamination. These new approaches, often called "microbial source tracking" (MST), have been used successfully for well over 10 years in a number of areas in the U.S. Ribotyping of *E. coli* isolates cultured from target surface waters is one approach that can provide information on a wide range of potential sources of fecal contamination.

Various studies have reported on the use of ribotyping for tracking sources of fecal-borne microbial contaminants. The University of New Hampshire's Jackson Estuarine Laboratory (UNH/JEL) lab has conducted well over 30 studies, mostly in New Hampshire and Maine, but also in Massachusetts, Vermont and New York states. Starting with the 1st UNH study conducted in the Colchester area in 2000, these ribotyping studies have been conducted in freshwater watersheds and beaches, and marine and estuarine waters (Jones 2007, Jones 2008, Nelson et al. 2008).

Because ribotyping can provide information on the identity of source species of bacteria found in surface waters, follow-up efforts to identify and eliminate contamination sources can be directed towards those types of sources where the few species responsible for the most significant amounts of contamination can be targeted for management action. Through an iterative process of then finding possible sources of fecal contamination from significant species, ribotyping can be used again to match strains for a given species to specific sources. Thus, the overall effort to improve water quality can be targeted because the most significant sources actually found in surface waters of concern are directly identified and eliminated. Such an approach also provides significant savings of time and expense compared to traditional approaches.

Watersheds with direct discharges to inner Malletts Bay are currently listed as impaired for *E. coli* under the Clean Water Act, Section 303(d) list (VT DEC, 2008). All water quality sampling associated with this project will occur in surface waters located in the Town of Colchester, Vermont.

Exceedences of *E. coli* water quality standards in Colchester, along Lake Champlain, Malletts Bay, and some tributatary streams have resulted in the posting of some swimming beaches during summer recreation months(Table 1).

Site#	Location	2000 Ribotyping study identified sources*	2004 Geometric mean	2004 90th percentile	2005 Geometric mean	2005 90th percentile	2006 Geometric mean	2006 90th percentile	2007 Geometric mean	2007 90th percentile	2008 Geometric mean	2008 90th percentile	2009 Geometric mean	2009 90th percentile	Geomean 2009 order descending
1	Delta Park Beach		24	82	13	53	34	224	22	67	37	249	6	20	14
2	Colchester Point		10	29	22	296									
3	Mills Point		8	81	9	162									
4	Porters Point		9	125	6	99	5	30	7	31	20	111	8	23	10
5	Camp Holy Cross		4	55	5	60	3	28							
6	Spalding West Beach		10	141	5	33	8	207	6	119	9	35	13	324	5
6A	Spalding West Culvert		45	1,267											
7	Moorings Stream		46	462	46	360	31	99	84	1,733	23	201	9	143	9
8	Smith Hollow Beach	cat,duck,coyote,human,raccoon	8	103	20	1,366	17	178	11	47	25	122	13	168	5
8A	Smith Hollow Creek		224	710	242	1,186							307	509	1
9	60 West Lakeshore Drive	cat,gull,human,raccoon	7	43	13	96	17	101	10	39	15	29	12	104	8
10	4 West Lakeshore Drive		9	59	7	33									
11	Crooked Creek Beach		26	153	25	201	31	305	18	145	71	488	33	73	4
11A	Crooked Creek	raccoon,gull,cow	215	729	733	2,420							208	356	2
12	Malletts Creek												37	50	3
B-1	Bayside Beach West	deer,raccoon,cat	11	69	14	152	8	36	28	204	18	35	8	16	10
B-2	Bayside Beach Center		25	180	22	191	7	29	24	216	14	68	13	53	5
B-3	Bayside Beach East		13	127	12	68	5	24	18	201	12	58	5	7	15
R-1	Rossetti Beach West								4	35	11	58	8	225	10
R-2	Rossetti Beach East								6	25	9	48	8	257	10

Table 1. Geometric mean *E. coli* concentrations expressed as colony forming units/100 milliliters (CFU/100ml) at beach and watershed sites from 2000-09.

An earlier study (Jones 2002a) of Malletts Bay and the lower Winooski River included four sites used in the present study, where birds, pets, wild animals and human sources were identified:

1. Smith Hollow Beach (site #8) – cat, duck, coyote, human, racoon
2. 60 West Lakeshore Drive (site #9) – cat, gull, human, racoon
3. Crooked Creek (site #11A) – racoon, gull, cow
4. Bayside Beach West (site #B-1) – deer, racoon, cat

Identification of the source(s) of the contamination from the water samples collected in the 2002 study, and in more extensive sampling conducted over the past two years, will help direct management activities for eliminating significant sources of microbial pollution that limit recreational uses of the lake.

MATERIALS AND METHODS

The general study area is the beach area of Malletts Bay in Colchester VT and the surrounding watersheds. Previous years of water quality monitoring by the Town of Colchester provided data on *E. coli* concentrations that helped to focus the present study efforts (Table 1).

Samples from 12 Town beach and related sites, including routine monitoring sites and sites previously shown to have high *E. coli* concentrations, were collected from July 7 to September 8, 2010 by the Town of Colchester and processed for enumerating putative *E. coli* concentrations by membrane filtration at Endyne (Table 2). Up to 20 isolates from samples with *E. coli* concentrations ranging from 20 to >2000 cfu/100 ml were shipped to UNH/JEL for further analysis.

Site/Date	7/7/10	7/12/10	7/19/10	8/2/10	8/9/10	8/18/10	8/23/10	8/30/10	9/1/10	9/8/10	Geometric mean*		
Rainfall 48h (in)	0.00	1.28	0.02	0.00	0.16	0.04	0.87	0.00	0.00	0.14			
	cfu/100 ml										cfu/100 ml		
Weather		***Wet***		***Dry***	***Dry***	***Dry***	***Wet***		***Dry***	***Dry***	***All***	***Wet***	***Dry***
M4-PP	10	3	49								11	3	22
M6-SW	2	3	< 1	< 1	***82***	***21***	3	3	***>1000***	***22***	8	3	16
M7-MS	18	28	24	7	***> 400***	10	***1300***	***115***	***18***	***32***	45	191	31
M8-SH	***47***	***127***	***72***	***95***	***15***	***18***	***120***	***105***	***45***	***135***	62	123	52
M8A-SH	***240***	***243***	***220***	***290***	***475***	***170***	***>2000***	***>2000***	***28***	***280***	320	731	261
M9-CT	12	***63***	3	***51***	26	3	***68***	5	6	***6***	13	65	8
M11-CC	***50***	***112***	***40***	***41***	***110***	***200***	***580***	***32***	***17***	***86***	75	255	55
M11A-CC	***700***	***990***	***870***	***530***	***> 2000***	***440***	***2000***	***400***	***490***	***>1000***	818	1407	714
M12-MC	27	***62***	31	34	33	***22***	***120***	***46***	***39***	***56***	42	86	35
MB2-Bayside	37	40	23	43	30	***75***	***34***	6	***11***	***>400***	36	37	35
MR1-Rossetti	25	3	6								8	3	12
MR2-Rossetti	5	8	7	< 1	11	***39***	2	6	4	***20***	6	4	7

*Geometric means included data based on estimates for low ("<#"; 10% less than detection limit) and high (">#"; 10% more than detection limit) concentrations.

Table 2. *E. coli* concentrations at beach and related sites during the study period. Numbers in italics are samples sent to UNH/JEL for speciation confirmation; highlighted numbers are samples chosen for ribotyping

Samples were also collected for determination of *E. coli* concentrations and possible ribotyping as part of two other related studies in 2010. Samples for a phosphorus study were collected from 12 sites in the watersheds surrounding the beach area on three dates from May 24 to August 19, 2010 (Table 3). Samples were also collected as part of a targeted watershed study conducted by Stone Environmental. Samples were collected from 32 sites on three dates from August 18 to October 26, 2010 (Table 4). All samples from these two studies were also processed for enumerating putative *E. coli* concentrations by membrane filtration at Endyne. Up to 20 isolates from samples with *E. coli* concentrations ranging from 40 to >1000 cfu/100 ml for the phosphorus study and from 52 to 500 cfu/100 ml for the targeted watershed study were shipped to UNH/JEL for further analysis.

Phosphorus Study Sampling: 2010

Site/Date Weather	5/24/10 Dry	8/3/10 Rain	8/19/10 Dry	Geometric mean*
	cfu/100 ml			cfu/100 ml
BH	*77*	*680*	40	128
VI	*700*	*88*	*830*	371
MC	42	*> 1000*	*40*	123
EH	*> 1000*	*> 1000*	*590*	894
PB	*460*	*390*	*350*	397
SC	45	*620*	*150*	161
MS	*63*	*850*	13	89
PP	< 2	*> 1000*	*470*	98
S4	27		*100*	52
SB	28	*1000*	*87*	135
SBT	47	*410*	*87*	119
CD	10	*138*	*60*	44
SC - Duplicate	*70*			
CD - Duplicate	27			
MC - Duplicate		> 1000		
MS - Duplicate		1120		
PB - Duplicate			250	

* Geometric means included data based on estimates for low ("<#"; 10% less than detection limit) and high (">#"; 10% more than detection limit) concentrations.

Table 3. *E. coli* concentrations at sites during a phosphorus study. Numbers in italics are samples sent to UNH/JEL for speciation confirmation; highlighted numbers are samples chosen for ribotyping.

Stone Environmental Study: 2010

Site*	Sample date	*E. coli* cfu/100 ml	Sampling location
WQ01	**8/18/10**	2	
WQ02	**8/18/10**	2	
WQ03	**8/18/10**	2	
WQ04	**8/18/10**	13	
WQ05	**8/18/10**	***210***	
WQ06	**8/18/10**	***104***	Up the Crooked Creek
WQ07	**8/18/10**	***148***	Watershed
WQ08	**8/18/10**	***52***	
WQ09	**8/18/10**	28	
WQ10	**8/18/10**	28	
WQ20	**9/14/10**	***36***	SH Crk - impounded water, no flow
WQ21	**9/14/10**	***40***	SH Crk @ Bay
WQ22	**9/14/10**	***84***	SH Crk - impounded water, no flow
WQ23	**9/14/10**	<5	SH Crk
WQ24	**9/14/10**	20	SH Crk
WQ25	**9/14/10**	***360***	SH Crk @ Edgewood Dr subdvn
WQ26	**9/14/10**	<8	
WQ27	**9/14/10**	4	
WQ28	**9/14/10**	***58***	Outer Bay, west side of Marble I.
WQ29	**9/14/10**	<4	
WQ30	**9/14/10**	125	
WQ31	**9/14/10**	***500***	Malletts Crk @ hwy 2
WQ32	**9/14/10**	***320***	Indian Brk just past trailer park & stripmall
WQ40	**10/26/10**	***29***	SH Crk
WQ41	**10/26/10**	***27***	SH Crk
WQ42	**10/26/10**	***29***	SH Crk
WQ43	**10/26/10**	***26***	Pond Brk
WQ44	**10/26/10**	***91***	Indian Brk
WQ45	**10/26/10**	***18***	Crooked Crk
WQ46	**10/26/10**	***26***	Moorings Strm watershed
WQ47	**10/26/10**	***238***	Moorings Strm watershed
WQ48	**10/26/10**	***12***	Moorings Strm watershed

*Geometric means included data based on estimates for low ("<#"; 10% less than detection limit) and high (">#"; 10% more than detection limit) concentrations.

Table 4. *E. coli* concentrations at sites during a targeted sampling events conducted by Stone Environmental. Numbers in italics are samples sent to UNH/JEL for speciation confirmation; highlighted numbers are samples chosen for ribotyping.

Fecal samples from local source species were collected on five dates from February 9 to October 26, 2010 and shipped on ice to UNH/JEL. Fecal samples were decimally diluted to 10^{-8}. Aliquots (2.5 milliliters (ml)) from the dilution tubes and water sample were filtered through membrane filters (0.45 micrometers (µm) pore size) and the filters placed onto mTEC agar. The agar plates were incubated at 37°Celsius (C) for 2 hours then 44.5°C for 22 hours. Yellow colonies were counted as fecal coliforms. Following urease testing on urea substrate, the remaining yellow colonies were counted as *E. coli* and plates giving countable colonies were used for selection of *E. coli* strains for ribotyping analysis.

The *E. coli* strains were subject to a battery of biochemical tests to confirm their identity as *E. coli*. The procedures used for isolating and identifying *E. coli* strains for this study were according to standard lab protocols (Jones 2002b, Jones and Bryant 2004). After inspection of *E. coli* concentration trends at sites on different sample dates, a subset of samples were chosen for ribotyping. The study team decided to use only five isolates from each source species and water sample to be ribotyped to allow for analysis of as many samples as possible.

Generally five confirmed *E. coli* isolates from each sample were analyzed for determining ribopatterns. *E. coli* isolates were stored in cryovials at -80°C and re-cultured onto trypticase soy agar (TSA). Cultures on TSA were incubated overnight at room temperature (20°C). Some of the resulting culture was transferred to duplicate cryovials containing fresh glycerol/DMSO cryo-protectant media for long-term storage at -80°C. The culture was then ready for ribotyping.

A RiboPrinter was used to analyze *E. coli* cultures for ribotype determinations. After preparation of the samples, the automated process involved lysing cells and cutting the released deoxyribonucleic acids (DNA) into fragments via the restriction enzyme EcoR1. These fragments were separated by size through gel electrophoresis and then transferred to a membrane, where they were hybridized with a DNA probe and mixed with a chemiluminescent agent. The DNA probe targeted 5S, 16S and 23S ribosomal ribonucleic acids (RNA) genes. A digitizing camera captured the light emission as image data, from which the system extracted a RiboPrint® pattern. This pattern was initially compared to others in the RiboPrinter database for characterization and identification based on densiometry data. However, our source species identification analysis approach has conformed to other ribotyping studies in using banding patterns as the basis for comparing patterns.

Band Identification

The images were transferred from the RiboPrinter into GelComparII (Applied-Maths) analytical software. The bands in lanes containing the standard were labeled and entered into the memory for optimization of gel pattern images. The densiometry data were processed for band identification. Due to technique or supply anomalies, some banding patterns were lighter than usual and too few bands were identified based on standard GelComparII parameters. All ribopatterns were inspected and in a relatively few cases obvious significant bands were manually added. The ribopattern data for each water sample isolate were then selected for identification of source species.

Data Analysis

The analysis of the Colchester water sample isolates for identification of source species was based first on the local source species database then on a more comprehensive regional database (Table 5). The local database consisted of 79 unique ribopatterns for *E. coli* isolates collected from 13 species and septic systems, all located in the study area watersheds. The regional database consisted of 1228 unique ribopatterns from 35 species and sources collected from the Northeast United States. The term 'unique (ribo)pattern' reflects the occurrence of identical patterns for closely related strains and clonality of *E. coli* strains in individual animals or sources. Clonality refers to instances where multiple strains were isolated from a given sample and with identical ribopatterns, indicative of them being multiple individuals from the same strain, i.e., they are clones. The redundant patterns are not counted under 'unique patterns'.

	REGIONAL			COLCHESTER		
Species	**# Samples**	**# Ribotypes**	**# Unique Ribotypes**	**# Samples**	**# Ribotypes**	**# Unique Ribotypes**
Alpaca	1	3	2	-	-	-
Buffalo	2	10	8	-	-	-
Cat	7	44	21	-	-	-
Chicken	5	33	25	-	-	-
Cormorant	8	48	25	-	-	-
Cow	11	89	68	-	-	-
Coyote	13	55	43	3	14	12
Deer	48	191	120	4	21	16
Dog	25	167	88	1	4	4
Duck	8	21	14	-	-	-
Fox	19	75	53	-	-	-
Goat	2	10	8	-	-	-
Goose	23	140	93	3	10	7
Horse	14	65	54	1	5	2
Human	8	115	54	-	-	-
Landfill Trash	4	20	20			
Mouse	1	3	2	-	-	-
Muskrat	6	37	19	1	5	2
Otter	4	19	13	1	5	4
Oxen	1	10	4	-	-	-
Pig	1	16	5	-	-	-
Pigeon	2	7	4	-	-	-
Rabbit	6	35	27	1	5	3
Racoon	34	92	69	3	13	8
Robin	1	4	2	-	-	-
Seagull	33	180	111	1	5	3
Septage	8	53	36	3	15	13
Sheep	2	8	5	-	-	-
Skunk	1	6	4	-	-	-
Sparrow	1	4	3	-	-	-
Starling	1	3	1	-	-	-
Unidentified Avian	2	20	14	-	-	-
Unidentified Wildlife	6	45	31	-	-	-
Wastewater	37	193	169			
Wild Turkey	3	17	13	1	5	5
Totals	348	1838	1228	23	107	79

Table 5. Local Colchester and regional source species databases.

All data were analyzed with GelComparII software on a Dell computer, where the source species database was also stored. Similarity indices between the unknown isolates and the known source isolates were determined by using Dice's coincidence index. For this study, 1% band tolerance and 1.5% optimization settings were used. Both of these parameters are used to adjust the ability to differentiate between bands for the degree of accuracy desired, and also to compensate for possible misalignment of homologous bands caused by technical problems.

The source species profile with the best similarity coefficient at a given set of optimization and tolerance settings was accepted as an indication of the possible source species for the water sample isolate. For this study, the predetermined threshold similarity index that was considered to be a minimum value for identifying source species was 90%. The decision for using a 90% similarity index threshold was based on the inter-gel variability within cumulative Dice's coincidence indices determined for *E. coli* positive controls run for every ribotyping study. Thus, the identification of the source species was considered successful if the value calculated for a given water isolate was equal to or greater than the threshold value; if the calculated value was below the threshold similarity index, the water sample isolate was considered to be of unknown origin. In some cases, the similarity threshold was 89% where differences between source and sample ribotype involved only one DNA band.

Cluster analyses were performed to determine the relationships among isolates from the same source species and the same sites, as well as banding patterns that were identical for different isolates. The cluster analyses were based on the un-weighted pair group method by arithmetic averaging (UPGMA) or the neighbor joining algorithms.

The last step in data analysis was visual inspection of the band matching results. Hard copies of ribotype patterns and similarity coefficients for the unknown and most closely related source species were printed for interpretation. Interpretation and accompanying tabular representations of the data were done using MS Excel on Macintosh computers. The results of identification of source species are summarized according to both the actual and type of source species identified.

RESULTS & DISCUSSION

<u>2010 *E. coli* concentrations and screening of samples for possible ribotyping</u>

Previous year monitoring showed the highest *E. coli* concentrations at the Colchester beach sampling sites were found in Crooked Creek (M11A-CC; upstream of beach site M11-CC) and in Smith Hollow Creek (M8A-SH; upstream of beach site M8-SH) (Table 1). Two of the three highest beach site geometric mean *E. coli* concentrations in 2008 and 2009 were at the corresponding beach sites, M11-CC and M8-SH. In 2010, the highest *E. coli* concentrations were again observed at M11A-CC and M8A-SH, followed by the two corresponding beach sites, M8-SH and M11-CC (Table 2). *E. coli* concentrations during 2010 ranged from below detection (<1 cfu/100 ml) in three samples to a high of >2000 cfu/100 ml at M8A-SH and M11A-CC on several dates. These observations point to Crooked and Smith Hollow creeks as potentially significant sources of *E. coli* to Malletts Bay beaches.

The other two sets of sampling events, as part of a synoptic phosphorus study and targeted sampling in Malletts Creek watersheds, were only conducted in 2010, to provide characterization of spatial and temporal variability of *E. coli* and host species upstream of the Lake and Bay.

E. coli concentrations formed the main basis for deciding which water quality samples would be used for ribotyping to determine sources of indicator bacteria under worst-case and baseline conditions. Most of the samples with the highest *E. coli* concentrations were chosen, although some samples with high *E. coli* concentrations were not chosen (Tables 2-4). In some cases, samples from the same date across sites were chosen instead. Preceding weather conditions were also a factor; dates were chosen to reflect both wet and dry conditions (Tables 2-4). For 2010, 213 isolates were ribotyped out of a total of 588 confirmed (preserved) *E. coli* isolates from 884 isolates cultured from 45 separate water quality samples collected on 11 dates from July through October, 2010 (Table 6). Generally five isolates were ribotyped, though some samples had too few confirmed *E. coli* isolates.

Date	Site	*E. coli* cfu/100 ml	# isolates tested	# isolates preserved	# isolates ribotyped
7/12/10	**M8-SH**	**127**	**20**	**18**	**5**
	M8A-SH	**243**	**20**	**19**	**5**
	M9-CT	**63**	**20**	**18**	**5**
	M11-CC	**112**	**20**	**12**	**5**
	M11A-CC	**990**	**20**	**18**	**2**
	M12-MC	**62**	**20**	**14**	**5**
8/2/10	**M9-CT**	**51**	**20**	**19**	**4**
8/3/10	**BH**	**680**	**20**	**12**	**5**
	VI	**88**	**20**	**13**	**5**
	MC	**>1000**	**20**	**14**	**5**
	EH	**>1000**	**20**	**13**	**5**
	PB	**390**	**20**	**10**	**5**
	SC	**620**	**20**	**11**	**6**
	MS	**850**	**20**	**12**	**3**
8/9/10	**M6-SW**	**82**	**20**	**14**	**5**
8/18/10	**M8-SH**	**18**	**18**	**10**	**3**
	M8A-SH	**170**	**20**	**17**	**1**
	M11-CC	**200**	**20**	**17**	**5**
	M11A-CC	**440**	**20**	**14**	**5**
	MB2-Bayside	**75**	**20**	**19**	**5**
	MR2-Rossetti	**39**	**20**	**19**	**5**
	WQ05	**210**	**20**	**16**	**5**
	WQ06	**104**	**20**	**7**	**7**
	WQ07	**148**	**20**	**10**	**6**
8/19/10	**VI**	**830**	**20**	**15**	**5**
	MC	**40**	**6**	**5**	**5**
	EH	**590**	**20**	**17**	**5**
	PB	**350**	**20**	**14**	**5**
	SC	**150**	**20**	**16**	**5**
8/23/10	**M7-MS**	**1300**	**20**	**12**	**5**
	M8-SH	**120**	**20**	**6**	**3**
	M8A-SH	**>2000**	**20**	**11**	**4**
	M9-CT	**68**	**20**	**4**	**4**
	M11-CC	**580**	**20**	**4**	**4**
	M11A-CC	**2000**	**20**	**15**	**5**
	M12-MC	**120**	**20**	**14**	**5**
	MB2-Bayside	**34**	**20**	**8**	**5**
9/1/10	**M6-SW**	**>1000**	**20**	**12**	**5**
9/8/10	**MB2-Bayside**	**>400**	**20**	**15**	**6**
	MR2-Rossetti	**20**	**20**	**11**	**5**
9/14/10	**WQ25**	**360**	**20**	**11**	**5**
	WQ31	**500**	**20**	**13**	**5**
	WQ32	**320**	**20**	**5**	**5**
10/26/10	**WQ44**	**91**	**20**	**15**	**5**
	WQ47	**238**	**20**	**19**	**5**
		TOTALS	**884**	**588**	**213**

Table 6. *E. coli* concentrations, isolate species confirmation success (# preserved) and samples recommended for ribotyping.

Source species sampling and choice of isolates for ribotyping

Source species fecal matter samples for 2010 were collected beginning in February 2010 at Colchester Pond and again in on four other dates in March, August, September and October (Table 7). Of the 34 samples collected, 20 had *E. coli* concentrations high enough to warrant isolate testing. From those 20 samples, 15 were chosen for ribotyping up to five isolates. Source samples were from a variety of different locally occurring species or sources, including septage, dogs and different birds and wild animals. Overall, 408 isolates were tested, 288 preserved and 91 ribotyped for the 2010 local source species database.

	Sampling regime	Species	Sample #	*E. coli* conc/g WW	# isolates tested	# isolates preserved	# isolates ribotyped
2/9/10	Keeping Tracks Colchester Pond	rabbit	RA11	83	10	6	5
		deer A	DE11	65240	20	17	6
		deer B	DE21	2000	20	20	5
		coyote	CO11	77233	20	20	5
		grouse A	GR11	BDL	0		
		grouse B	GR21	BDL	0		
		avian	AV11	BDL	0		
		muskrat	MU11	BDL	0		
3/16/10	Brent Toth	coyote	CO11	17673	20	5	2
		muskrat	MU11	11194	20	20	5
		muskrat	MU21	75	0		
		raccoon	RC11	82659	20	20	5
		otter	OT11	BDL	0		
		mink	MI11	BDL	0		
		red fox	RF11	BDL	0		
8/18/10	Brent Toth	septage	SS01	4	0		
		septage	SS02	264	20	11	5
		septage	SS03	1600	25	11	5
		dog	DO11	>80000000	20	17	5
		dog	DO21	12000000	23	16	5
9/1/10	Keeping Tracks various places	duck	DU11	20400000	21	9	3
		unknown bird	UK11	BDL	0		
		otter	OT11	28800000	24	14	5
		raccoon	RC11	>80000000	20	11	5
		goose	GE11	36000000	29	23	5
		coyote	CO11	344000000	20	15	5
10/26/10	Brent Toth	sediment	SED11	4	0		
		sediment	SED21	BDL	0		
		sediment	SED31	BDL	0		
		goose	GE11	BDL	0		
		goose	GE21	BDL	0		
		raccoon	RC11	18800000	24	17	5
		deer	DE11	20000000	26	18	5
		deer	DE21	24000000	26	18	5
				TOTAL =	**408**	**288**	**91**

*BDL = Below Detection Limit

Table 7. Local source species sample *E. coli* concentrations, confirmed (preserved) *E. coli* isolates and ribotyping choices.

Five isolates from each of the 2010 sources/species were chosen for ribotyping as an initial screening of the diversity and usefulness of these sources and samples. Identical patterns from the same sample are not useful for identifying sources from water sample isolates, though all sources but the 3/16/10 coyote and the 9/1/10 raccoon sources had at least three unique patterns from the five chosen isolates, suggesting good diversity and the potential for getting new patterns from more isolates from these samples if needed. Some of the sources had one or more patterns that were 'shared', or identical to a pattern from another source isolate from the local source species database.

Ribotyping analysis using the local and regional databases

All chosen water sample isolates were ribotyped and then analyzed to determine source species using the combined 2009-2010 database of local sources (Table 8). Several isolates for some samples did not match well with the RiboPrinter *E. coli* database of ribopatterns and were thus not considered to be *E. coli*. Alternative isolates were ribotyped, but some samples had inadequate numbers of confirmed *E. coli* isolates to provide five for ribotyping.

Overall, 90 of the eventual 213 isolates (42%) tested were identified to sources, i.e., they matched source species patterns at ≥90% similarity. The remaining unidentified isolates either matched local source species patterns at <90% similarity (72 isolates; 34% of the total) or matched multiple, or 'mixed' source patterns at ≥90% similarity (46 isolates; 22% of the total). There were only single isolate matches to the raccoon, gull and horse source patterns, with single isolates matching to a mix of either bird or wild animal sources. The rest of the water sample patterns (85 isolates) matched to geese (28 isolates), deer (21 isolates), wild turkey (15 isolates), rabbit (10 isolates), septage (7 isolates), or otter and dog (3 isolates). Of note is the high number of isolates from geese (Canada geese) and deer, both of which had ≥10% identifications relative to all isolates analyzed. Wild turkey and rabbit were identified as sources for ≥10% of all identified isolates, suggesting these four sources are significant sources of contamination at the sample sites. The seven isolates from septage are also noteworthy, as this source is of most concern for public health reasons.

Site & date	TOTAL Isolates	Isolates identified	Goose	Wild turkey	Gull	Mix bird	Otter	Raccoon	Rabbit	Deer	Mix wild animal	Septic	Horse	Dog	No identification <90%	Mixed >90%
7/12/10																
M8-SH	5	2								1		1			3	0
M8A-SH	5	2	1						1						2	1
M9-CT	5	1		1											0	4
M11-CC	5	1	1												3	1
M11A-CC	2	1							1						0	1
M12-MC	5	2	1							1					1	2
8/2/10																
M9-CT	4	1							1						3	0
8/3/10																
BH	5	1		1											1	2
EH	5	0													2	3
MC	5	2	2												3	0
MS	3	1	1												2	0
PB	5	3		3											2	0
SC	5	1							1						3	1
VI	5	2		1						1					2	1
8/9/10																
M6-SW	5	3		1		1								1	1	0
8/18/10																
M8-SH	3	2	1	1											1	0
M8A-SH	1	0													0	1
M11-CC	5	2		1					1						3	0
M11A-CC	5	3		1			2								1	1
MB2-Bay	5	4	2	2											0	1
MR2-Ross	5	1	1												0	4
WQ05	5	4	2						1			1			1	0
WQ06	3	3		1					1			1			0	0
WQ07	4	2							1	1					2	0
WQ08	4	0													4	0
8/19/10																
EH	7	6	1							4		1			0	1
MC	6	4	1					1		1				1	1	0
PB	5	2	2												1	2
SC	5	3		1								2			2	0
V1	5	1							1						4	0
8/23/10																
M7-MS	5	2			1					1					2	1
M8-SH	3	2	1											1	1	0
M8A-SH	4	2	2												1	1
M9-CT	4	2	2												2	0
M11-CC	4	2							1	1					2	0
M11A-CC	5	1	1												4	0
M12-MC	5	2								1		1			1	2
MB2-Bay	5	1		1											0	4
9/1/10																
M6-SW	5	1	1												0	4
9/8/10																
MB2-Bay	6	5	1							4					0	1
MR2-Ross	5	2	2												1	2
9/14/10																
WQ25	5	2	1							1					3	0
WQ31	5	1	1												1	3
WQ32	5	0													3	2
10/26/10																
WQ44	5	1								1					4	0
WQ47	5	4								3			1		1	0
TOTAL	213	90	28	15	1	1	2	1	10	21	0	7	1	3	74	46
% of total		**42%**	13%	7%		0%	1%	0%	5%	10%	0%	3%		1%	**35%**	**22%**
% of ID'd		100%	31%	17%		1%	2%	1%	11%	23%	0%	8%		3%		

Two species/sources matched to the water sample pattern:
"Mix Bird" = goose and gull; "Mix Wild Animal" = several different combinations; "Mix Human" = septage and wastewater or landfill trash

Table 8. Ribotyping analysis using the combined 2009-2010 local Colchester source species database. Highlighted numbers are identifications that exceeded 10% of the total isolates analyzed, as designated.

The lack of any identified sources for four samples (EH on 8/3/10, M8A-SH & WQ08 on 8/18/10, WQ32 on 9/14/10) suggests the source(s) was not in the local database, or that an inadequate number of isolates from the sampled sources were included for the source species database. It is interesting that sources were identified at EH and M8A-SH on other dates, so the lack of identifications on a single sample date suggests different sources with time at these sites. It appears that the Canada goose, septage, rabbit, wild turkey, dog and otter source patterns were very useful for identifying sources while the horse, gull and raccoon patterns were less helpful. The isolates used from the latter four species apparently contain specific patterns that are not contributing to the pollution, or they may not be significant sources of *E. coli*. Selecting only five isolates from each source is a minimum number and analysis of more patterns from these samples could be useful in identifying sources for more water samples.

Use of the full regional database, including the local source isolates, enabled identification of sources for a total of 141 isolates, or 66% of the total 213 isolates analyzed (Table 9). The fraction of total isolates that were identified is above average for ribotyping studies, suggesting that *E. coli* isolate ribopatterns used for identifying sources, especially the local source, were well chosen for this watershed area. A relatively low percentage (12%) of the isolate patterns were either unique or did not match well with known source patterns, while 23% of the isolate patterns matched with a mix of species. Canada geese, wild turkey and deer remained significant sources, though better matches to regional database sources were found for several isolates identified to these and other species using the local database. The regional database included patterns that provided matches to some of the same source species in the local database, including Canada goose, gull, otter, raccoon, deer, septage, horse, and dog. This helps to verify these sources, and suggests that use of more patterns from local sources of these species could be useful for further identifying sources of pollution.

The number of different sources rose from 12 in the local database, to 17 with the regional database; new sources included fox (5 isolates), coyote (2 isolates), mixed human (2 isolates), pigeon, and sheep (1 isolate each). The most frequently identified sources using the regional/local database were Canada geese and deer, constituting 26% and 23% of identified sources, respectively. Other frequently identified sources include wild turkey (11 isolates), septage (10 isolates, including two that also matched to wastewater or landfill trash), raccoon (8 isolates), dog and gull (7 isolates), rabbit, and mixed wild animal (6 isolates), and fox (5 isolates). Sources were identified for one or two isolates at EH-8/3/10 and WQ32-9/14/10 where no sources were identified using only the local database. For two samples (M8A-SH & WQ08 on 8/18/10), no sources were identified, although there was only one isolate analyzed for the M8A-SH sample. Because all four isolates analyzed for WQ08 had patterns that matched at <90% similarity, either the pollution source(s) were not included in the regional/local database, or an inadequate number of isolates from the locally sampled sources were included in the source species database.

Site & date	TOTAL Isolates	Isolates identified	% ID'd	Goose	Wild turkey	Gull	Pigeon	Mix bird	Mix Human	Otter	Raccoon	Rabbit	Coyote	Fox	Deer	Mix wild animal	Septic	Horse	Sheep	Dog	No identification <90%	Mixed >90%
7/12/10																						
M8-SH	5	3	60%	1											1		1				1	1
M8A-SH	5	3	60%	1							1	1									2	0
M9-CT	5	2	40%								1				1						0	3
M11-CC	5	3	60%	2											1						1	1
M11A-CC	2	1	50%												1						0	1
M12-MC	5	3	60%	1											2						1	1
8/2/10																						
M9-CT	4	4	100%	2		1						1									0	0
8/3/10																						
BH	5	3	60%												1	1				1	1	1
EH	5	1	20%	1																	0	4
MC	5	2	40%	1							1										2	1
MS	3	2	67%	1		1															1	0
PB	5	5	100%		3									1	1						0	0
SC	5	4	80%								1	1				1	1				1	1
VI	5	2	40%		1										1						2	1
8/9/10																						
M6-SW	5	5	100%		1	1		1						1						1	0	0
8/18/10																						
M8-SH	3	2	67%	1	1																1	0
M8A-SH	1	0	0%																		0	1
M11-CC	5	5	100%	1				1			1	1				1					0	0
M11A-CC	5	5	100%		1		1			2								1			0	0
MB2-Bay	5	5	100%	1	1				1					1	1						0	0
MR2-Ross	5	1	20%	1																	0	4
WQ05	5	5	100%	3											1		1				0	0
WQ06	3	3	100%		1							1					1				0	0
WQ07	4	4	100%	1											1		1			1	0	0
WQ08	4	0	0%																		4	0
8/19/10																						
EH	7	6	86%	1											5						0	1
MC	6	6	100%	2		2		1			1										0	0
PB	5	2	40%	2																	1	2
SC	5	5	100%	1	1												2			1	0	0
V1	5	3	60%	1								1				1					0	2
8/23/10																						
M7-MS	5	5	100%	1		2		1							1						0	0
M8-SH	3	3	100%	1														1		1	0	0
M8A-SH	4	3	75%	2																1	0	1
M9-CT	4	2	50%	1					1												0	2
M11-CC	4	3	75%												2	1					0	1
M11A-CC	5	2	40%	1									1								3	0
M12-MC	5	2	40%												1		1				0	3
MB2-Bay	5	3	60%		1										2						0	2
9/1/10																						
M6-SW	5	3	60%	1						1					1						0	2
9/8/10																						
MB2-Bay	6	6	100%	1										1	3					1	0	0
MR2-Ross	5	2	40%	1									1								0	3
9/14/10																						
WQ25	5	3	60%	1											1				1		1	1
WQ31	5	1	20%	1																	0	4
WQ32	5	2	40%								1			1							0	3
10/26/10																						
WQ44	5	3	60%								1				1	1					2	0
WQ47	5	3	60%												3						1	1
TOTAL	213	141		36	11	7	1	4	2	3	8	6	2	5	32	6	8	2	1	7	25	48
% of total		**66%**		17%	5%	3%	0%	2%	1%	1%	4%	3%	1%	2%	15%	3%	4%	1%	0%	3%	**12%**	**23%**
% of ID'd		100%		26%	8%	5%	1%	3%	1%	2%	6%	4%	1%	4%	23%	4%	6%	1%	1%	5%		

Two species/sources matched to the water sample pattern:
"Mix Bird" = goose and gull "Mix Wild Animal" = several different combinations "Mix Human" = septage and wastewater or landfill trash

Table 9. Ribotyping analysis using the combined regional and Colchester source species databases. Highlighted numbers are identifications that exceeded 10% of the total isolates analyzed, as designated.

Samples with elevated *E. coli* concentrations can often be caused by one or a few dominant sources. Six samples, from M6, M7, M8A, M11A, EH and MC, had *E. coli* concentrations >1000 cfu/100 ml (Table 6). There were 29 isolates analyzed from these samples and 16 (55%) were identified (Table 9). Geese were identified for 7 isolates from each sample date/site. Deer and gull were the sources for two isolates apiece, and raccoon, dog, coyote, otter and mixed bird were identified for one isolate each. There was no evidence of any dominant source for samples with elevated *E. coli* concentrations; identified sources included a range of sources and only two samples had more than one isolate identified to a given source. Each of these sources can have high *E. coli* concentrations in feces found in the environment (Table 7).

There was also no consistent evidence for different sources to be present under different weather (dry or wet) conditions (Tables 2 & 9), or for different seasons (summer or autumn). *E. coli* concentrations were higher under wet weather at all sites except those at the western side of the study area (M4, M6, MR1&2; Table 2). Only a few sites (M9, M11, M11A, MB2) had isolates ribotyped on both dry and wet weather days. Comparisons of the few source species under the different weather conditions showed completely different sources at 11A compared to similar sources at M11 and MB2, while M9 had only wild animal sources under dry conditions and mostly birds under wet conditions.

The inclusion of upstream and downstream samples at Crooked ("M11") and Smith Hollow ("M8") creeks on 7/12/10, 8/18/10 and 8/23/10 showed geese commonly identified at both sites in Smith Hollow for two dates (no isolates were identified on 8/18/10 for M8A-SH) and dogs at both sites on one date, while there was only deer as a commonly identified source at the two Crooked Creek sites on only one of the three dates. This suggests different relationships between downstream (beach) and upstream sources of contamination in these two beach areas. As suspected based on observation of congregating birds, the sources for the MB2-BAY sample included Canada geese.

The identified sources can be separated into five types of sources, including human, pet, bird, wild animal and domestic animal sources. These source types correspond to different management strategies for elimination or reductions of sources, and are thus useful for gauging what type of management strategy would be best for reducing the most significant sources.

For all samples, wild animals and birds were the dominant identified source types, including 29 and 28%, respectively, of the total isolates analyzed by the combined source species database (Table 10). Other studies in the region have shown wild animals and/or birds to be the most significant sources at freshwater beaches (Jones 2008). The public health significance of either of these fecal pollution sources is not well known, but bird-borne fecal pollution may have disease implications even for non-human species (Nelson et al. 2008).

2010	Vermont DB analysis		Regional DB analysis	
Source type	# of isolates	% of total	# of isolates	% of total
Human	7	3%	10	5%
Pet	3	1%	7	3%
Livestock	1	0%	3	1%
Bird	45	21%	59	28%
Wild animal	34	16%	62	29%
Identified	90	42%	141	66%
Unidentified	123	58%	72	34%
TOTAL	213		213	

Table 10. Types of source species for 2010 isolates identified by analysis with both source species databases.

The combined results for all samples at each of the 26 study sites shows wild animals and birds to be the most widespread source types, being identified in 23 and 21 of the sites, respectively. Pets were identified sources at 7 sites, human sources at 6 sites and domestic animals/livestock at 3 sites. Site M8-SH was the only site that had all five types of pollution sources.

Site	# isolates	# ID'd	% of total isolates identified	TYPE of SOURCE % of identified isolates Birds	Wild animals	Human	Domestic animals	Pets	% isolates unknown source
M6-MS	10	8	80%	50%	38%			13%	20%
M7-MS	5	5	100%	80%	20%				0%
M8-SH	11	8	73%	50%	13%	13%	13%	13%	27%
M8A-SH	10	6	60%	50%	33%			17%	40%
M9	13	8	62%	50%	50%				38%
M11-CC	14	11	79%	36%	64%				21%
M11A-CC	12	8	67%	38%	50%		12%		33%
M12-MC	10	5	50%	20%	60%	20%			50%
MB2-Bay	16	14	88%	29%	57%	7%		7%	13%
MR2-Ross	10	3	33%	67%	33%				60%
BH	5	3	60%		67%			33%	40%
EH	12	7	58%	29%	71%				42%
MC	11	8	73%	75%	25%				18%
MS	3	2	67%	100%					33%
PB	10	7	70%	71%	29%				33%
SC	10	9	90%	22%	33%	33%		11%	18%
VI	10	5	50%	40%	60%				50%
WQ05	5	5	100%	60%	20%	20%			0%
WQ06	3	3	100%	33%	33%	33%			0%
WQ07	4	4	100%	25%	25%	25%		25%	0%
WQ08	4	0	0%						100%
WQ25	5	3	60%	33%	33%		33%		40%
WQ31	5	1	20%	100%					80%
WQ32	5	2	40%		100%				60%
WQ44	5	3	60%		100%				40%
WQ47	5	3	60%		100%				40%

Table 11. Types of source species identified at each sampling site sampled during 2010.

Most of identified human (8 of 10 isolates) and all (3) of livestock sources were identified in either the Smith Creek or the Crooked Creek watershed. Isolates from dogs were identified at seven sites in five different watersheds, and not found in only the Moorings Stream watershed. In the Bay beach (MB2), Ross beach (MR2), Moorings stream (MS) and the Mallett's

Creek (MC) watersheds, birds and wild animals made up 75 of 79, or 95% of the identified sources.

The *E. coli* isolates collected in 2009 were reanalyzed to identify source species using combined source species databases that included ribopatterns for local and regional strains collected in 2010. We were especially interested in whether the 2010 local source species would be useful in identifying sources for 2009 sample isolates. The combined 2009-10 local database included isolates from 23 source species and 107 isolates (Table 5), an increase of 17 samples and 77 isolates from the 2009 database. The inclusion of wild animals in the local source species database in 2010 gave substantially more identifications of isolates. Those identified as wild animal sources increased from 0 to 14 isolates, while the number identified as coming from birds decreased by 4 isolates (Tables 12 A&B). Inclusion of more samples and isolates from local septage sources also provided two new identifications of human sources. Overall, sources for 47% of the isolates were identified using the local database that included 2010 isolates, and improvement of 16% from the 2009 analysis.

A.	Vermont DB analysis		Regional DB analysis	
Source type	# of isolates	% of total	# of isolates	% of total
Human	1	1%	2	3%
Pet	0	0%	1	1%
Livestock	1	1%	2	3%
Bird	21	28%	25	34%
Wild animal	0	0%	9	12%
Identified	23	31%	39	53%
Unidentified	51	69%	35	47%
TOTAL	74		74	

B.	Vermont DB analysis		Regional DB analysis	
Source type	# of isolates	% of total	# of isolates	% of total
Human	3	4%	6	8%
Pet	1	1%	3	4%
Livestock	0	0%	0	0%
Bird	17	23%	18	24%
Wild animal	14	19%	19	34%
Identified	35	47%	46	62%
Unidentified	39	53%	28	38%
TOTAL	74		74	

Table 12. Source species types for samples collected in 2009 identified by analysis with A.) 2009 source species databases, or B.) 2009 and 2010 source species databases.

The impact on source identifications using the combined regional database increased the overall identifications by 7 isolates, with increases in identifications for pets (+2), humans (+4) and wild animals (+10), and decreases for livestock (-2) and birds (-7). Again, the biggest gain was for isolates identified from wild animal and human sources. All apparent decreases in identification are from shifts in identification due to the replacement of these matches with statistically higher pattern matches.

The results from analysis of both 2009 and 2010 isolates are of interest to see if source identification between years for sites and overall sources was consistent, and if not, whether changes were expected or otherwise explainable. Nearly three times as many isolates were analyzed in 2010, and the combined results for both years were highly similar in terms of % source types and overall isolate sources identified (Table 13) compared to only 2009 (Table 12B) and only 2010 (Table 10) results.

2010	Vermont DB analysis		Regional DB analysis	
Source type	# of isolates	% of total	# of isolates	% of total
Human	10	3%	16	6%
Pet	4	1%	10	3%
Livestock	1	0%	3	1%
Bird	62	22%	77	27%
Wild animal	48	17%	81	28%
Identified	125	44%	187	65%
Unidentified	162	56%	100	35%
TOTAL	287		287	

Table 13. Source species types for all samples collected during 2009 and 2010 identified by analysis with databases containing source isolates from 2009 and 2010.

REFERENCES

Jones, S.H. 2008. Environmental Sources of Microbial Contaminants in Shellfish and Their Public Health Significance. J. Foodservice 19: 238-244.

Jones, S.H. 2007. Optimized Use of *Escherichia coli* Ribotyping for Identifying Pollution Sources in New Hampshire's Coastal Waters. Final Report. New Hampshire Department of Environmental Services, Concord, NH.

Jones, S.H. 2002a. Microbial source tracking in Vermont using *Escherichia coli* ribotyping. Final report to the US EPA Section 104(b)3. US EPA Region 1, Boston MA.

Jones, S.H. 2002b. QA Plan for the Jackson Estuarine Laboratory Microbiology Lab. USEPA approved: 2002.

Jones, S.H. and T. Bryant. 2004. Standard procedure for detection of total coliforms, fecal coliforms, *Escherichia coli* and enterococci from environmental samples. Jackson Estuarine Laboratory, University of New Hampshire, Durham, NH.

Nelson, M., S.H. Jones, C. Edwards and J.C. Ellis. 2008. Characterization of *Escherichia coli* populations from gulls, landfill trash, and wastewater using ribotyping. Dis. Aquat. Org. 81: 53-63. doi: 10.3354/dao01937

VT DEC (2008). Direct smaller drainages to Inner Mallets Bay, Appendix 7, *In*. Vermont Statewide Total Maximum Daily Load (TMDL) for Bacteria Impaired Waters. Vermont Department of Environmental Conservation.