

Prepared in cooperation with the
Vermont Department of Environmental Conservation

Using Host-Associated Genetic Markers to Investigate Sources of Fecal Contamination in Two Vermont Streams



Scientific Investigations Report 2011–5113

Using Host-Associated Genetic Markers to Investigate Sources of Fecal Contamination in Two Vermont Streams

By Laura Medalie, Leslie J. Matthews, and Erin A. Stelzer

Prepared in cooperation with the
Vermont Department of Environmental Conservation

Scientific Investigations Report 2011–5113

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2011

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit <http://www.usgs.gov> or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod>

To order this and other USGS information products, visit <http://store.usgs.gov>

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

Suggested citation:

Medalie, Laura, Matthews, L.J., and Stelzer, E.A., 2011, Using host-associated genetic markers to investigate sources of fecal contamination in two Vermont streams: U.S. Geological Survey Scientific Investigations Report 2011–5113, 30 p., at <http://pubs.usgs.gov/sir/2011/5113>.

Acknowledgments

Many individuals and organizations helped in various ways. The project would never have gotten started without guidance and enthusiasm from Donald Stoeckel, formerly a hydrologist with the U.S. Geological Survey (USGS) Ohio Water Science Center. Neil Kamman of the Vermont Department of Environmental Conservation (VTDEC) also was instrumental in getting started. Timothy Clear of VTDEC helped to select sampling stations. Gerald DiVincenzo and Daniel Needham from the VTDEC LaRosa Laboratory generously allowed use of their equipment and facilities. Aaron Worthley, Robert Low, Marli Rupe, and Ethan Swift shared their expertise and local knowledge, and many other volunteers of the Huntington and Richmond Conservation Commissions and the Poultney Mettawee Natural Resource Conservation District contributed time by collecting water-quality samples for several years. Generous staff at Audubon Vermont in Huntington and the Nature Conservancy in West Haven provided time and local expertise. Several farmers and property owners graciously acquiesced to the odd request for samples of their animal waste. Susan Morse was a knowledgeable and enthusiastic instructor in finding and identifying scat. Special thanks to Heather Pembroke and Robert Low and their neighbors in Huntington and Richmond for organizing the Huntington dog-waste collection day.

Very helpful reviews were provided by Joseph Duris and Kenneth Hyer, both of the USGS.

This page has been left blank intentionally.

Contents

Abstract.....	1
Introduction.....	1
Microbial Source Tracking.....	2
<i>Bacteroidales</i> Genetic Markers.....	2
Limitations to Using <i>Bacteroidales</i> Genetic Markers.....	2
Purpose and Scope	3
Study Areas.....	3
Huntington River Watershed.....	3
Mettawee River Watershed.....	5
Sampling Design, Collection, and Processing	8
Selection of Sampling Stations	8
Sampling Design	10
Sample Collection and Pre-Processing	10
Water Samples.....	12
Fecal Reference Samples	12
Use of Fecal Reference Samples to Adjust Marker Results in Water	14
Results of Genetic-Marker Analyses	15
Fecal Reference Samples	15
Huntington River.....	17
Mettawee River and Flower Brook.....	21
Limitations of Study and Evaluation of Hypotheses.....	21
Relation of Results to Land Use.....	25
Relation of Results to Discharge.....	25
Ratios of <i>E. coli</i> to Genetic-Marker Material in Fecal Material and in Water	25
Summary and Conclusions.....	27
References Cited.....	28

Figures

1. Map showing locations of the Huntington River and the Mettawee River watersheds in Vermont.....4
2. Map showing land use in the Huntington River watershed in Vermont and stations sampled for genetic-marker testing
3. Map showing land use in the Mettawee River watershed in Vermont and stations sampled for genetic-marker testing
4. Graph showing example of load-duration curve for *Escherichia coli* (*E. coli*) for station M3 on the Mettawee River, Vermont, showing that most instantaneous loads were higher than the target load over the range of flow conditions.10
5. Boxplots showing summaries of historical *Escherichia coli* (*E. coli*) concentrations at selected stations in the A, Huntington (Huntington Conservation Commission, 2010) and B, Mettawee (The Poultney Mettawee Natural Resources Conservation District, 2009) Rivers

6. Photographs showing several varieties of scat collected in or near the Huntington and Mettawee River watersheds in Vermont during 2009.....	13
7. Bar graph of median genetic-marker results showing with log-transformed data that there is some potential for cross reactivity among all reference groups.....	17
8. Graphs showing concentrations of <i>Escherichia coli</i> (<i>E. coli</i>), turbidity, and potential for sources to explain observed <i>E. coli</i> concentrations in samples collected from stations on the Huntington River; discharge at the Mad River streamgage 04288000 adjusted to the Huntington River watershed by drainage area; and approximate precipitation in Huntington, Vermont, on A, July 16, 2009; B, August 2 and 3, 2009; and C, October 7 and 8, 2009.....	18
9. Graphs showing concentrations of <i>Escherichia coli</i> (<i>E. coli</i>) and potential for sources to explain observed <i>E. coli</i> concentrations in samples collected from stations on the Mettawee River and Flower Brook; discharge at the Mettawee River streamgage 04280450 adjusted to the most downstream sampling station M3 by drainage area; and approximate precipitation in the Mettawee River watershed, Vermont, on A, July 29 and 30, 2009; B, August 21 and 22, 2009; and C, November 12, 2009.....	22
10. Boxplot showing comparison of ratios of log-transformed (base 10) <i>Escherichia coli</i> concentration to log-transformed (base 10) genetic-marker concentration for all fecal-reference samples used to compute theoretical thresholds and water samples in the Huntington and Mettawee Rivers.....	26

Tables

1. Basin characteristics and land use within the Huntington and Mettawee River watersheds.....	5
2. References for analytical methods of genetic-marker tests using quantitative polymerase-chain reaction for <i>Bacteroidales</i> genetic markers, Vermont, 2009.....	8
3. Sampling stations for <i>Bacteroidales</i> genetic-marker tests of the Huntington and Mettawee Rivers.....	9
4. Sources of fecal reference samples by reference group.	14
5. Concentrations of <i>Escherichia coli</i> and host-associated genetic markers for fecal reference samples collected in Vermont, 2009.....	16
6. Concentrations of <i>Escherichia coli</i> and host-associated genetic markers and field parameters for water samples collected in the Huntington River in Vermont, 2009.....	20
7. Concentrations of <i>Escherichia coli</i> and host-associated genetic markers and field parameters for water samples collected in the Mettawee River and Flower Brook in Vermont, 2009.....	24
8. Time since previous substantial runoff event for sample collection dates in the Huntington and Mettawee Rivers	27

Conversion Factors, Datum, and Abbreviations

Multiply	By	To obtain
Length		
foot (ft)	0.3048	meter (m)
inch (in.)	2.54	centimeter (cm)
mile (mi)	1.609	kilometer (km)
Area		
square mile (mi ²)	2.590	square kilometer (km ²)
Flow rate		
cubic foot per second (ft ³ /s)	0.02832	cubic meter per second (m ³ /s)
Volume		
quart (qt)	0.9464	liter (L)
Mass		
ounce, avoirdupois (oz)	28.35	gram (g)

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$^{\circ}\text{F}=(1.8\times^{\circ}\text{C})+32$$

Vertical and horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83) and to the Vermont State Plane Coordinate System.

Concentrations of chemical constituents in water are given in parts per million (ppm); volumes of liquids are given in milliliters (mL) and liters (L).

List of Acronyms

CFU	colony forming unit
DNA	deoxyribonucleic acid
<i>E. coli</i>	<i>Escherichia coli</i>
HCC	Huntington Conservation Commission
MPN	most probable number
MST	microbial source tracking
NRCD	Natural Resources Conservation District
OB	optical brighteners
OWML	Ohio Water Microbiology Laboratory
ppm	parts per million
qPCR	quantitative polymerase chain reaction
RCC	Richmond Conservation Commission
rRNA	ribosomal ribonucleic acid
USEPA	U.S. Environmental Protection Agency
USGS	U.S. Geological Survey
VTDEC	Vermont Department of Environmental Conservation

This page has been left blank intentionally.

Using Host-Associated Genetic Markers to Investigate Sources of Fecal Contamination in Two Vermont Streams

By Laura Medalie, Leslie J. Matthews¹, and Erin A. Stelzer

Abstract

The use of host-associated *Bacteroidales*-based 16S ribosomal ribonucleic acid genetic markers was investigated as a tool for providing information to managers on sources of bacterial impairment in Vermont streams. The study was conducted during 2009 in two watersheds on the U.S. Environmental Protection Agency's 303(d) List of Impaired Waters, the Huntington and the Mettawee Rivers. Streamwater samples collected during high-flow and base-flow conditions were analyzed for concentrations of *Escherichia coli* (*E. coli*) and *Bacteroidales* genetic markers (General AllBac, Human qHF183 and BacHum, Ruminant BoBac, and Canid BacCan) to identify humans, ruminants, and canids as likely or unlikely major sources of fecal contamination. Fecal reference samples from each of the potential source groups, as well as from common species of wildlife, were collected during the same season and from the same watersheds as water samples. The results were combined with data from other states to assess marker cross reaction and to relate marker results to *E. coli*, the regulated water-quality parameter, with a higher degree of statistical significance.

Results from samples from the Huntington River collected under different flow conditions on three dates indicated that humans were unlikely to be a major source of fecal contamination, except for a single positive result at one station that indicated the potential for human sources. Ruminants (deer, moose, cow, or sheep) were potential sources of fecal contamination at all six stations on the Huntington River during one high-flow event and at all but two stations during the other high-flow event. Canids were potential sources of fecal contamination at some stations during two high-flow events, with genetic-marker concentrations in samples from two of the six stations showing consistent positive results for canids for both storm dates. A base-flow sample showed no evidence of major fecal contamination in the Huntington River from humans, ruminants, or canids.

Results from samples from the Mettawee River watershed collected during high-flow conditions (12 storm

samples on 2 dates at 6 stations) indicated that there was no evidence of fecal contamination from humans in seven samples and possible evidence in five samples. Results for humans were positive for only one station during both storm events. For two of the five samples with evidence for human fecal contamination, results for two different human genetic markers agreed, but results from three samples were inconsistent. In samples from five of the six Mettawee stations, ruminants were a potential source of fecal contamination on at least one of the three sampled dates, including three positive results for the base-flow sample. Yet samples from all of the stations that showed positive results for ruminants did so for only one or two of the three sampled dates. Samples from only one of the six stations gave consistent results, which were negative for ruminants for all three dates. In the Mettawee River base-flow sample, humans were an unlikely source of major fecal contamination.

Factors that may influence results and conclusions include the timing of sample collection relative to the storm event; variability of *E. coli* and *Bacteroidales* concentrations in fecal reference samples and in water; sampling and analytical errors; the potential cross reactivity of host-associated genetic markers; and different persistence and survival rates of *E. coli* bacteria and *Bacteroidales* genetic markers on land, in water, and by season. These factors interfere with the ability to directly relate *Bacteroidales* concentrations to *E. coli* concentrations in river samples. It must be recognized that while use of *Bacteroidales* genetic markers as a source tracking tool coupled with the interpretive approach described in this report cannot be used quantitatively to pinpoint sources, it can be used to exclude potential sources as major contributors to fecal contamination.

Introduction

Waters contaminated with fecal material pose a health risk to humans from pathogenic viruses and bacteria, protozoa, or invertebrates. Fecal contamination is commonly determined by testing water for the indicator *Escherichia coli* (*E. coli*) (Dufour, 1984), a generally harmless intestinal bacterium found in all mammals and birds. Vermont has established the

¹ Vermont Department of Environmental Conservation.

water-quality standard for *E. coli* in class B water (suitable for recreational activities) to be 77 colony forming units (CFU) per 100 milliliters (mL) (State of Vermont Natural Resources Board, 2008). The U.S. Environmental Protection Agency (USEPA) has a less stringent criterion of a single sample maximum of 235 CFUs per 100 mL for a designated beach area in fresh recreational waters (U.S. Environmental Protection Agency, 1986). For 37 of the 101 Vermont streams listed as impaired—not meeting water-quality criteria established by the State—*E. coli* or fecal coliform bacteria is at least one of the contaminants of concern (Vermont Department of Environmental Conservation, 2008). The Vermont Department of Environmental Conservation (VTDEC) is investigating approaches for managing this issue; however, there remains a large gap between measuring concentrations of *E. coli* in recreational waters that sometimes exceed standards or criteria and the larger management question of how to mitigate the problem.

Microbial Source Tracking

The desire to identify human and animal sources of fecal contamination in surface waters used for recreational beaches and shellfish harvest has prompted the development of a variety of Microbial Source Tracking (MST) methods for analyzing environmental water samples. The methods are predicated upon the assumption that feces from different sources have distinguishing characteristics that enable identification in water samples. Dozens of MST approaches, characterized by differences in cost, speed, practicality, and type of information yield, have been developed and are in different stages of use. MST practitioners generally agree that no single method is appropriate for all applications (Stoeckel and Harwood, 2007; U.S. Environmental Protection Agency, 2005). Library-dependent MST methods are costly and time consuming and require the construction of a library of local fecal reference samples (U.S. Environmental Protection Agency, 2005). Some studies have suggested that library-dependent MST methods may be excessively prone to false-positive (Griffith and others, 2003) or false-negative (Stoeckel and others, 2004) source classifications. Some library-independent MST methods, such as coliphage abundance (Cole and others, 2003), are nonquantitative and are useful only to broadly distinguish between human and animal sources of fecal contamination.

A USEPA MST-guide document details several characteristics of the ideal microbial-source marker (U.S. Environmental Protection Agency, 2005). These characteristics include host specificity; ubiquitous distribution in the host; stability with regard to mutations, time, and geography; consistent rate of chemical decomposition; similar distribution in water to that found in the original fecal material; conduciveness to representative sampling in water; ability to assess relative and absolute contributions from various hosts; and relevance to regulatory tools and health risks.

Bacteroidales Genetic Markers

Bacteroidales-based 16S ribosomal ribonucleic acid (rRNA) genetic markers (hereafter called “*Bacteroidales* genetic markers”) have been shown to possess some of the ideal characteristics described by USEPA (2005). These markers, detected using quantitative polymerase chain reaction (qPCR), demonstrate host specificity—they exploit host-specific sequence variations in the 16S rRNA gene of gut-associated anaerobic bacteria in the order *Bacteroidales*. As of 2009, distinct markers were identified for feces from human, ruminant, canid, horse, and swine sources (Bernhard and Field, 2000; Simpson and others, 2004; Kildare and others, 2007; Layton and others, 2006). Using *Bacteroidales* genetic markers confers two major advantages: bacteria of the order *Bacteroidales* are found in human and animal feces and compose the largest portion of intestinal microbiota, and they are obligate anaerobes—they have limited survival and reproduction in the oxygenated surface-water environment.

Limitations to Using *Bacteroidales* Genetic Markers

There also are some shortcomings to using qPCR-detected *Bacteroidales* genetic markers for MST. Genetic markers have not been developed for any wildlife species (the ruminant marker cannot distinguish among deer, moose, cow, or sheep hosts, and the canid marker cannot distinguish among domestic dogs, wolves, coyotes, or fox hosts), host specificity may be lost for some species that live in close contact (for example, dogs may eat feces of humans or other potential hosts), and the persistence of individual genetic markers in the environment may vary (Field and Samadpour, 2007) with respect to each other and with respect to indicator bacteria such as *E. coli*. Also, the use of *Bacteroidales* genetic markers means that the species used for source tracking is different than the indicator species widely used for quantifying fecal contamination and for setting water quality standards (*E. coli*). Therefore, the results need to be translated into quantitative estimates for the regulated species, *E. coli*, to provide a meaningful result to managers.

In reality, most streams in natural environments carry fecal contamination from several or many different sources. In this report, including or excluding a source group as a potential source of fecal contamination indicates that the load coming from the particular source was or was not large enough to meet a defined numerical criterion. If results associated with a source do not meet the numerical criteria, even if detected, then that source may still be present but not in quantities that account for a large percentage of the *E. coli* bacteria. Use of the term “potential source” is not intended to convey any information about other potential sources that were not tested, which could also be large contributors to the observed *E. coli* concentrations. Information gained from the practical experience of applying the method in Vermont

will provide some context for evaluating the usefulness of this technique in helping to manage impairment due to fecal contamination in Vermont streams.

Purpose and Scope

The U.S. Geological Survey (USGS), in cooperation with the Vermont Department of Environmental Conservation (VTDEC), investigated the use of qPCR-detected *Bacteroidales* genetic markers to exclude or identify humans, ruminants, or canids as potential sources of fecal contamination, measured by the indicator *E. coli* bacteria, in two Vermont streams. This report describes the application in Vermont of a method based on host-associated *Bacteroidales* genetic markers to investigate some of the potential sources that contribute to fecal contamination in surface water. The method was applied to surface water collected from the Huntington River and the Mettawee River watersheds, both in the Lake Champlain Basin (fig. 1). The purpose of the study was to determine whether there was evidence of fecal contamination in the Huntington River in northwestern Vermont (from humans, dogs, or ruminants) and in the Mettawee River in southwestern Vermont (from humans or ruminants) and to assess the applicability of this method for more widespread use in other streams in Vermont. In each watershed, three sets of water samples were collected during the summer and fall of 2009 at six sampling locations. One set of samples was collected during base-flow conditions, and two sets were collected during high-flow conditions. Fecal reference samples were collected within or close to each watershed during the same time frame when the water samples were collected. An additional set of water samples collected in each watershed was analyzed for the presence of optical brighteners (OBs), a chemical added to many laundry detergents that can be used as an indicator of leaks from human sewage or septage systems, with the goal of evaluating OBs as a potential screening tool for rapid detection of human sources of fecal contamination.

Study Areas

Two watersheds representing different land uses and potentially different sources of fecal contamination were chosen for this study. Our focus was on rural watersheds because identifying the sources of fecal contamination in nonurban impaired streams has been particularly challenging for resource managers in the State of Vermont. The most important determinant for watershed selection was documented bacterial impairment by means of *E. coli* and a history of concentrations above the State standards and Federal criteria. A hypothesis that likely sources of fecal contamination include humans, ruminants, or canids (recognizing that other sources of fecal contamination also are present) was also needed to support the choice of genetic markers chosen for this study. Finally, an established group of local volunteers was needed to help,

particularly with obtaining permissions from landowners for accessing property and also collecting or providing guidance for finding fecal material to use as reference samples. The Huntington River and Mettawee River watersheds, both in the Lake Champlain Basin, met these criteria and were selected for this study (fig. 1).

Huntington River Watershed

The Huntington River watershed is a rural and generally pristine watershed in north-central Vermont (fig. 2). The Huntington River flows northward from the headwaters near Stark Mountain through Huntington Center and Huntington Village before turning northeastward in Richmond to the Winooski River near Jonesville. The ridge forming the eastern boundary of the watershed follows portions of the Long Trail—the oldest long-distance trail in the United States—through the length of Vermont and includes the Camels Hump summit, the fifth highest peak in Vermont. At the downstream sampling station for this project (H6), the Huntington River watershed is about 60 mi², and the estimated population in 2000 was 2,520 or about 42 people per square mile (people/mi²) (table 1). The sampled reach of the river was between Huntington Center and the point near where the river curves to the northeast in Richmond, Vt. (fig. 2).

In 2002, three percent of the watershed was developed land that was generally comprised of roads and residences along the valleys of the Huntington River and its tributaries (University of Vermont, 2005). Nine percent of the watershed was classified as agricultural land, also situated along the river valleys. Most of the land area, about 82 percent, was forested uplands, and the remaining 6 percent was water and wetlands.

There were no permitted point-source discharges within the Huntington River watershed—all of the residences and the few municipal or commercial properties in the watershed were served by onsite septic systems. A walking trail between two of the sampling stations was a frequently used destination for dogs and their owners. In this reach of the river as well as others, pools created from scoured bedrock were popular swimming holes. During the study period, the sampled reach of the Huntington watershed included numerous horses, one medium-sized dairy farm (200–699 mature animals), and several other small farming operations (less than 200 mature animals) with various combinations of beef cattle, sheep, pigs, chickens, and alpaca. There were large contiguous tracts of undeveloped land in and around the Huntington River watershed, including parts of the Camels Hump State Forest and Park system, several state wildlife management areas, and numerous wetlands, that provided bountiful habitat for wildlife. Large mammals in these forested and wetland areas included beaver, deer, moose, bear, coyote, fisher, raccoon, and mink, in addition to many species of smaller mammals. Many birds, both migratory and nonmigratory, also bred or passed seasonally through the forests of Vermont.

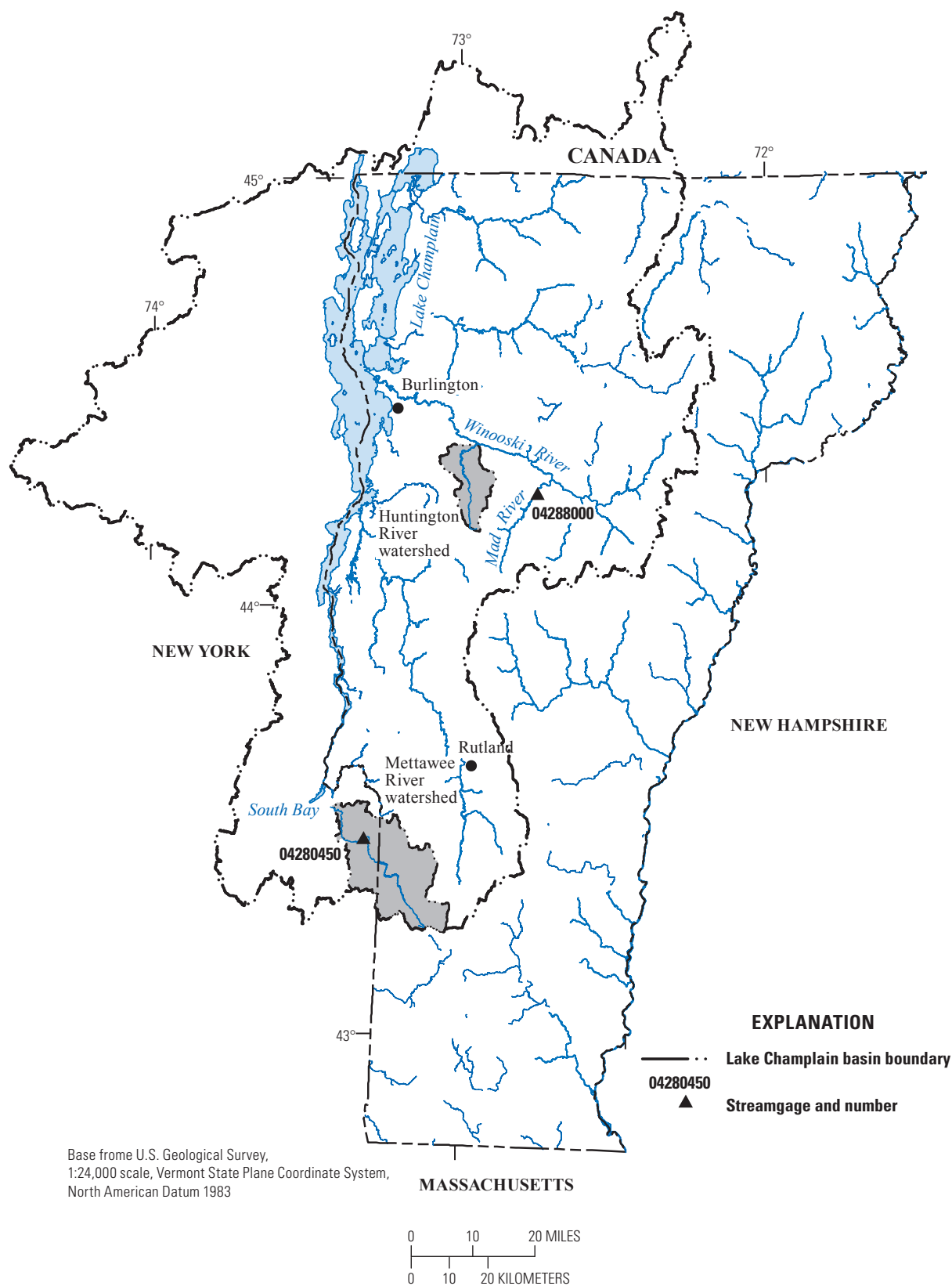


Figure 1. Locations of the Huntington River and the Mettawee River watersheds in Vermont.

Table 1. Basin characteristics and land use within the Huntington and Mettawee River watersheds.

[Characteristics are for drainage areas as delineated at the downstream sampling stations; drainage area is in square miles; population is estimated as the sum of populations of census blocks that are within the respective watersheds; population density is in persons per square mile; population data from U.S. Census Bureau (2001); land-use data from University of Vermont (2005)]

	Watershed	
	Huntington	Mettawee
Drainage area	60	69
Population in 2000	2,520	1,660
Population density	42	24
Percentage of land area		
Developed	3	9
Forest	82	65
Agricultural	9	20
Water and wetlands	6	6

The Huntington River watershed is underlain primarily by interlayered black graphitic schists of phyllite and rust-colored weathered sulfidic schist or phyllite of the Hazens Notch Formation (Thompson and Thompson, 1992). Most of the soils in the watershed are shallow to moderately deep above dense basal till and very deep to bedrock, and are moderately well to well drained; there are also some poorly drained soils in the watershed (Natural Resources Conservation Service, 2008). Most slopes are 0 to 20 percent, although about one-third of the area has slopes up to 60 percent.

The average annual temperature in Burlington is 45 degrees Fahrenheit (°F), and monthly averages vary from 18°F in January to 71°F in July. In Huntington Center, average monthly precipitation varies from 1.81 inches (in.) in February to 4.81 in. in August, with 41.1 inches being the annual average total (National Oceanic and Atmospheric Administration, 2002). About half of the average annual precipitation drains into streams and rivers (Hornbeck and others, 1997), but this amount varies throughout the year because of seasonal evapotranspiration. Data on discharge in the Huntington River are adjusted for differences in drainage area from the closest continuous USGS streamgage, 04288000 Mad River near Moretown, Vt. (fig. 1). At the Mad River streamgage, the historic record of 81 years shows that monthly summer discharges ranged from 21 to 734 ft³/s and annual discharges ranged from 134 to 430 ft³/s. Discharges are generally largest in April due to snowmelt and least in the summer months because of evapotranspiration.

Mettawee River Watershed

The Mettawee River watershed, like the Huntington, was mostly rural. The river flows from the hills of Dorset southwest for about 6 mi before turning west and then northwest (fig. 3). It enters New York State in Granville and empties to the Champlain Canal in Whitehall just south of Lake Champlain's South Bay. The drainage area of the Mettawee watershed delineated at the downstream sampling station is about 69 mi². Compared to the Huntington River watershed, the Mettawee River watershed had more than double the percentage of agricultural land (20 percent), about 3 times the percentage of developed land (9 percent), and less forested land (65 percent). Estimated population and population density in the sampled part of the watershed in 2000 were about 1,660 and 24 people/mi².

Wetlands are common near the Mettawee River and Flower Brook. Numerous small and medium-sized dairy-farm operations with corn and hay fields lined the main river corridor from the river headwaters to the New York border, although the greatest concentration of farmland was roughly from the Rupert-Dorset town line to water-sampling station M3. Horses were numerous, and pigs and beef cattle were also common in the upland part of the watershed. Leakage from manure storage facilities and manure runoff from fields at dairy- or beef-cattle farms were potential sources of ruminant fecal contamination. About 300 people in the western part of Pawlet live in sewered houses connected to the West Pawlet wastewater-treatment facility, which discharges to a tributary of the Mettawee River downstream of the sampled reach. The remaining Vermont population living within the Mettawee River watershed is served by onsite septic systems. The village of Pawlet, which occupies an area along Flower Brook just above its confluence with the Mettawee River, consists of a cluster of about 70 buildings, mostly single-family residences with a handful of commercial and municipal facilities. Failures of onsite septic systems in the village are a potential source of fecal contamination from humans. Unlike the Huntington River, there is no dog-walking destination along the Mettawee River. The Mettawee River supports populations of wild brook, brown, and rainbow trout and is a popular fly-fishing destination (Hickoff and Plumley, 2007).

The Mettawee River watershed in the vicinity of the study area is underlain primarily by purple, gray-green, and variegated slate and phyllite of the Cambrian St. Catherine Formation; dark gray to black phyllite with beds of blue-gray marble of the Cambrian Brezee Formation; and interbedded dolomite, limestone or marble, calcareous sandstone, and quartzite of the Ordovician Bascom Formation (Doll and others, 1961). Most of the upland soils of the Mettawee River watershed have slopes of 25–80 percent and are very rocky, poorly drained, and shallow above bedrock. Valley soils generally are characterized by low slopes, sandy or gravelly loams, and are more than 5 ft thick above bedrock (Natural Resources Conservation Service, 2008).

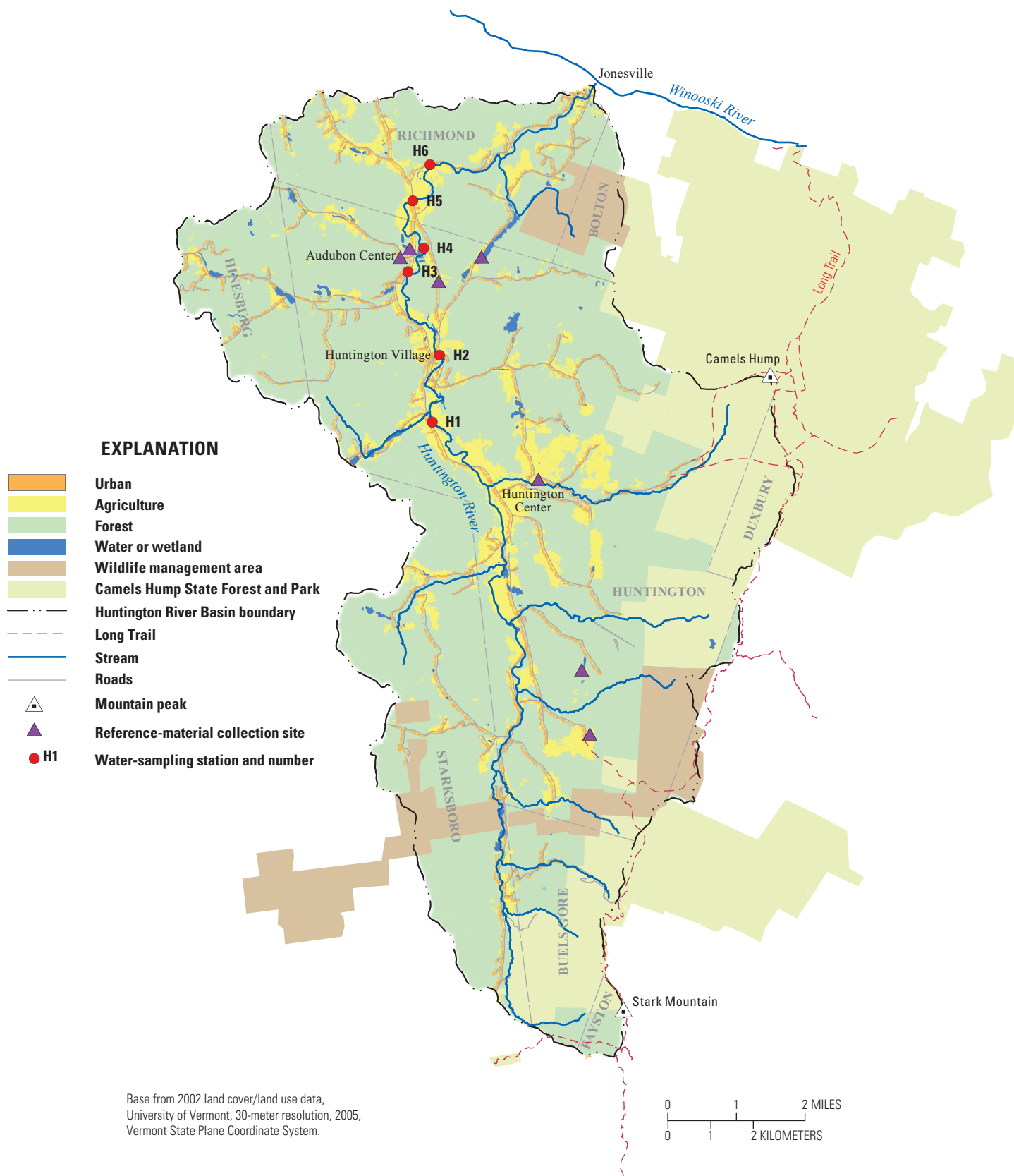


Figure 2. Land use in the Huntington River watershed in Vermont and stations sampled for genetic-marker testing.

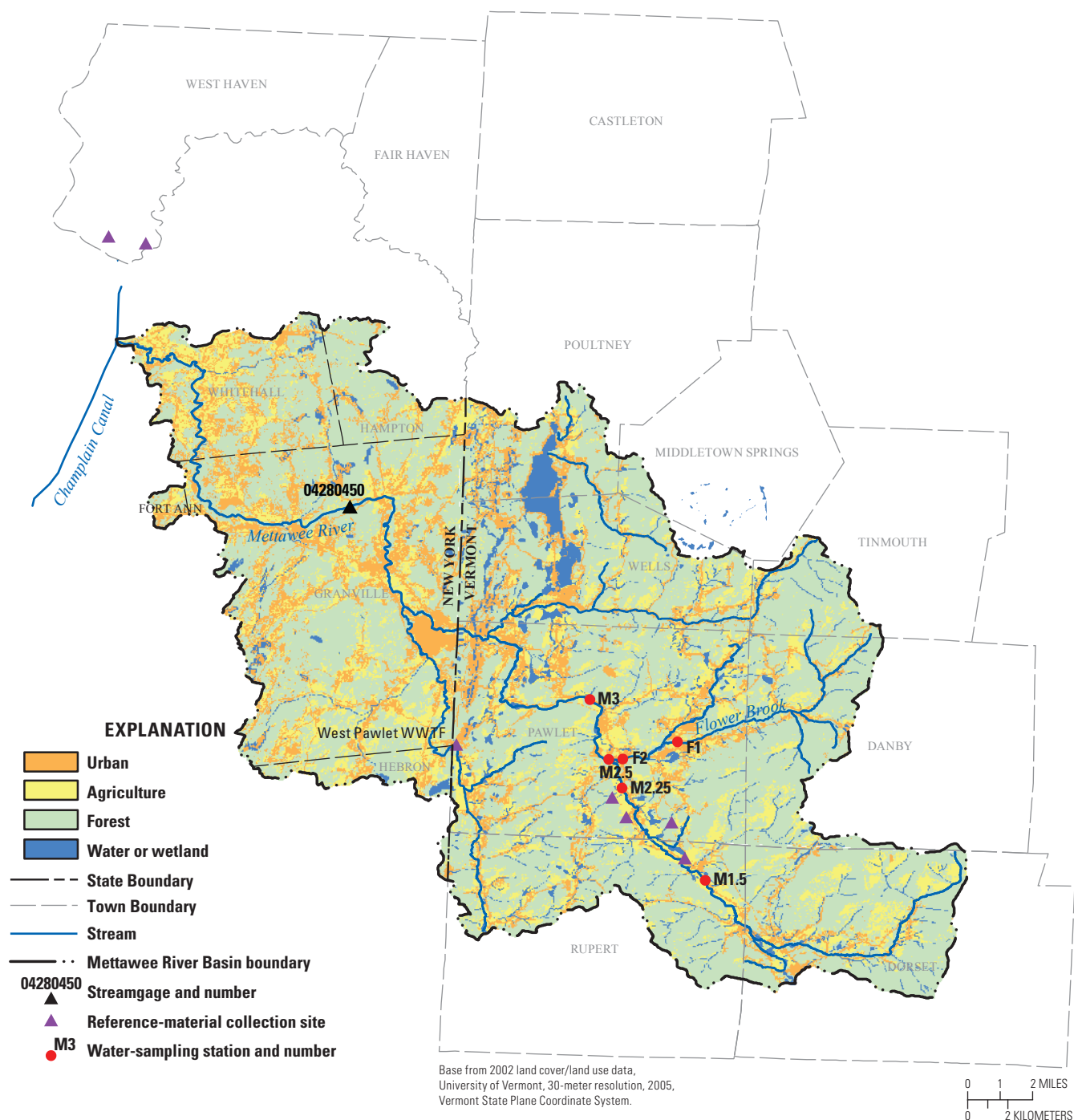


Figure 3. Land use in the Mettawee River watershed in Vermont and stations sampled for genetic-marker testing.

In Rutland, near the Mettawee River watershed (fig. 1), the average annual temperature is 46°F, and monthly averages vary from 20°F in January to 69°F in July (National Oceanic and Atmospheric Administration, 2002). Average monthly precipitation varies from a February low of 1.97 in. to a July high of 4.58 in. Annual average total precipitation is 39.1 inches. Streamflow discharges reported for the Mettawee River are adjusted for the 41 percent difference in drainage area from USGS streamgage 04280450, Mettawee River near Middle Granville, New York (fig. 1). Streamflow discharges vary greatly on a monthly and annual basis; for example, at streamgage 04280450, the minimum summer monthly discharge for the period of record since 1990 was 16.8 ft³/s in September 1995 and the maximum was 434 ft³/s in August 2004. Variations can also be great within a year, as for example in 1996, when monthly discharges progressed from 400 to 156 to 47 ft³/s from July through August.

Sampling Design, Collection, and Processing

All analyses except for genetic-marker concentrations and *E. coli* concentrations for human water and fecal samples were conducted by project personnel at the VTDEC LaRosa Laboratory (referred to as the VTDEC Laboratory in this report). After the field season, frozen samples were sent to the USGS Ohio Water Microbiology Laboratory (OWML) for the analysis of genetic-marker concentrations. References for the laboratory methods for the deoxyribonucleic acid (DNA) extraction and qPCR analysis done by the OWML are provided in table 2. The limits of quantification for the qPCR assays were based on methods described by Kildare and others (2007).

Table 2. References for analytical methods of genetic-marker tests using quantitative polymerase-chain reaction for *Bacteroidales* genetic markers, Vermont, 2009.

Genetic-marker test (primer)	Target host species	Reference
AllBac	General	Layton and others, 2006
qHF183	Human	Seurinck and others, 2005
BacHum	Human	Kildare and others, 2007
BoBac	Ruminant	Layton and others, 2006
BacCan	Canid	Kildare and others, 2007
dsRed2	Plasmid-borne dsRed insert	Stoeckel and others, 2009

Selection of Sampling Stations

Sampling stations on the Huntington and Mettawee Rivers and Flower Brook tributary were chosen from among the preexisting monitoring stations established by the volunteer monitoring programs as those from which samples were most likely to provide information on sources of bacteria through analysis of *Bacteroidales* genetic markers (table 3). No new stations were considered because we wanted to build on existing data and because stream access from private lands had already been established. Station numbers for the Mettawee River and Flower Brook are the same as those used by the Poultney Mettawee Natural Resources Conservation District (NRCD). Three methods were used to determine the most suitable subset of stations on each river for the *Bacteroidales* genetic-marker testing: physical reconnaissance of the watersheds, examination of historic data, and tests for OBs in stream water.

Physical reconnaissance of the Huntington River watershed in June and July of 2008 and of the Mettawee River watershed in May of 2009 consisted of direct observations in addition to gathering of information from local residents and conservation groups. Direct observations were made of the rivers, on the land near the rivers, and of digital orthophotographs of the watersheds, with goals of looking for specific land uses and noting any general features that could explain high levels of bacteria in the water. Information was obtained primarily from people in the local watershed groups who collected the historical *E. coli* data: the Huntington Conservation Commission (HCC), the Richmond Conservation Commission (RCC), and the Poultney Mettawee NRCD. Staff from Audubon Vermont in Huntington and The Nature Conservancy in West Haven (near the Mettawee River) also were consulted during the reconnaissance.

The HCC and RCC volunteer groups have collected water samples for *E. coli* analysis at 18 stations on the main stem of the Huntington River since 2003. Between 2005 and 2008, the Poultney Mettawee NRCD collected water samples for *E. coli* analysis at 8 stations on the Mettawee River and its Flower Brook tributary. Although water samples from no single station had consistently high relative *E. coli* concentrations, data showed that a few stations were better candidates for the genetic-marker tests on the basis of more frequent high relative *E. coli* concentrations and on the inclusion of particular land uses within the drainage areas of specific stations. Load-duration curves showing how instantaneous (measured) loads of *E. coli* compared to target loads (defined in this case by the Vermont standard for Class B waters of 77 CFU per 100 mL) by flow-duration range provided evidence that both unpermitted point and nonpoint sources of contamination had contributed to bacteria at four of the candidate stations on the Huntington and Mettawee Rivers and Flower Brook (one station is shown as an example in figure 4). The load-duration curves provide evidence for point sources of *E. coli* contamination when instantaneous amounts plot above targeted amounts under low-flow and

Table 3. Sampling stations for *Bacteroidales* genetic-marker tests of the Huntington and Mettawee Rivers.

[Hypothetical sources of fecal contamination include humans, ruminants, canids, and wildlife in the Huntington River watershed, and all of these except for canids in the Mettawee River watershed.]

Sampling station	Land use immediately above station
Huntington River (H)	
H1	Farm
H2	Village
H3	Residential, forest
H4	Forest
H5	Open land
H6	Open land, forest
Mettawee River (M) and Flower Brook (F)	
M1.5	Farm, forest
M2.25	Farm, forest
F1	Farm, forest, residential
F2	Village, forest
M2.5	Forest, village, open land
M3	Farm, forest

dry conditions and for nonpoint sources of contamination when instantaneous amounts plot above targeted amounts under moist and high-flow conditions (U.S. Environmental Protection Agency, 2007). Summary box plots of historical *E. coli* data (fig. 5) for the selected stations (table 3) show that quartiles, medians, and geometric means of *E. coli* concentrations generally were higher in samples from the Mettawee River than from the Huntington River but that there were more maximum concentrations at or near the analytical limit (2,419 most probable number per 100 mL [MPN/100 mL]) for the Huntington River than for the Mettawee River. Several of the annual data reports prepared by the HCC for the Huntington River (<http://users.gmavt.net/kevinandt/HCC3.htm> accessed on June 3, 2011) indicate the existence of a direct relation between *E. coli* concentrations and streamflow.

A screening tool using OBs was developed to detect human sources as an auxiliary element of this project. OBs have the potential to be a practical and low cost method for detecting human wastewater in urban stormwater and streams. Failed or malfunctioning septic systems have been suspected as one of the potential sources of elevated bacteria in the Huntington and Mettawee Rivers. OBs that are added

to many laundry and some dishwashing detergents fluoresce when exposed to certain wavelengths of ultraviolet light. The detection of OBs in water samples indicates probable human contamination from improperly functioning onsite septic systems or leaking municipal sewage pipes (Dickerson and others, 2007; Hartel and others, 2008). A positive result for OBs in water samples could identify a good candidate station for using genetic-marker tests to further test water samples for human sources. The first question posed was whether this method would provide a practical screening tool to help select stations for the more expensive genetic-marker analysis. A more general question was whether this method would be viable as a future screening tool for detecting human waste products in Vermont waterways.

A Turner Designs AD700 fluorometer at the VTDEC Laboratory was retrofitted for testing of OBs in water samples with an optical brightener filter kit that consisted of (1) an excitation filter (300–400 nanometers), (2) an emission filter (436 nanometers), and (3) a near-ultraviolet lamp. Water samples were collected directly as manual grabs into sterile containers, stored in a cooler on ice, and brought to the VTDEC Laboratory for immediate analysis using instrument protocols as described in the operating manual (Turner Designs, 2002). A standard curve was developed by serially diluting Tide™ liquid detergent (Hartel and others, 2008) from a 1,000 parts per million (ppm) solution (1 mL of Tide dissolved in 1 L of distilled and deionized water). The 1,000 ppm solution was used to establish the standard curve of 40, 81, 142, 398, and 556 relative fluorescence units. Two negative control samples—an OB-free laundry detergent (2.4 relative fluorescence units) and pond water from a park (5.3 relative fluorescence units)—and two positive control samples—an urban stream (30 relative fluorescence units) and effluent from a wastewater-treatment facility (80 relative fluorescence units)—confirmed the validity of the method and the standard curve.

Water from the Huntington River was tested for OBs on two dates (September 11, 2008 and September 21, 2009), and water from the Mettawee River was tested once (September 21, 2009). Since the OB method is essentially an attempt to detect a point source of contamination, low-flow conditions were targeted. Daily streamflow on September 11, 2008, at the Mad River streamgage (04288000) was 52 ft³/s, which is less than the mean of the daily values, 81 ft³/s, for that date based on 81 years of streamflow record. Daily streamflow on September 21, 2009, in the Mettawee River at USGS streamgage 04280450 was 61 ft³/s, which is close to the mean of the daily values, 63 ft³/s, for that date based on 19 years of record. None of the results yielded detections of OBs in the water for either stream. Descriptions of applying this method in the literature indicate that OB detections are much more common in water from storm drains or pipe outfalls than open water (Dickerson and others, 2007). Although the method failed to detect OBs in water samples from the Huntington or Mettawee Rivers, it did detect OBs in positive control samples of water from an urban stream and a wastewater-treatment

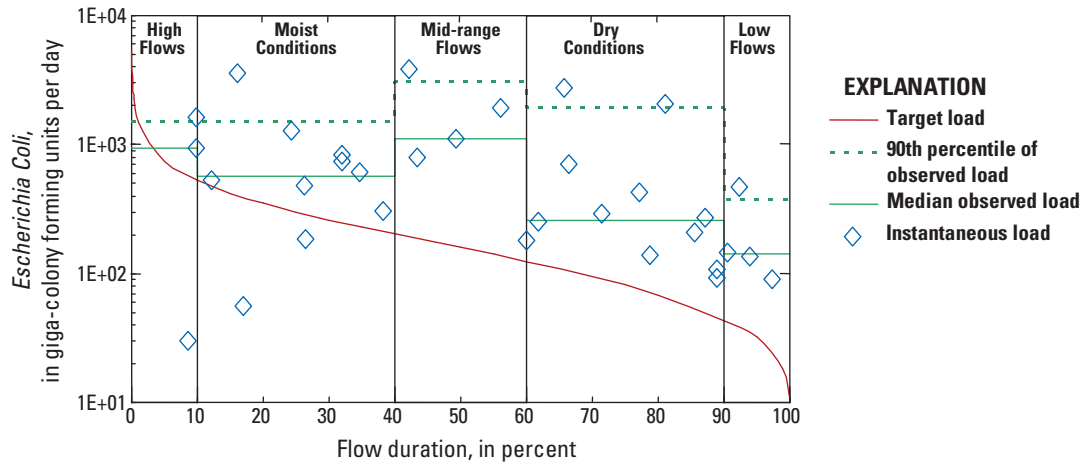


Figure 4. Example of load-duration curve for *Escherichia coli* (*E. coli*) for station M3 on the Mettawee River, Vermont, showing that most instantaneous loads were higher than the target load over the range of flow conditions. Plot created by Timothy Clear, Vermont Department of Environmental Conservation.

facility. Thus, OBs may be a practical rapid-screening tool for septic leaks or other wastewater effects on the surface waters of Vermont, particularly close to inflows, because there are minimal fixed costs and the test is quick and easy to implement.

Sampling Design

All water and fecal reference samples were collected during 2009. The general sampling scheme was to collect three to four sets of water samples from each station during the recreational swimming season of June through September and during the fall months if more samples were needed. One sample set would be collected during base-flow conditions to provide a comparison for ambient conditions, and the remaining sets would be collected during higher flows when the concentrations of *E. coli* tended to be greater. During the same time frame, fecal reference samples were collected from each of the potential source groups, as well as from common species of wildlife in each of the watersheds, to assess marker cross reaction and establish quantitative boundaries for interpreting streamwater results.

Water samples with high concentrations of *E. coli* indicate fecal contamination and are likely to yield the most informative results from the genetic-marker tests. Although *E. coli* concentrations were determined for each water sample collected, the results were not known until at least 24 hours after the samples were collected and processed. Judgment needed to be exercised to make the decision regarding whether and when to collect a storm sample. Several sources of information were considered in making the decision: Web sites that posted radar and other weather data; real-time USGS streamflow for the Mettawee River, the Mad River, and other rivers in the storm path; and telephone calls to people who lived in the study areas. When it appeared as though a

given storm had potential to generate runoff that could raise *E. coli* concentrations in the river, project personnel drove to either the Huntington or Mettawee River and measured turbidity at one of the stations. Turbidity is used as a surrogate indicator for whether *E. coli* concentrations are elevated (Rasmussen and Ziegler, 2003). If turbidity was not higher than background levels (generally less than 2 formazin nephelometric units), then samples were not collected at that time.

For the water samples from the Huntington River, the genetic markers AllBac,² BoBac (for cows and other ruminants), qHF183 (humans), and BacCan (dogs and other canids) were analyzed. For the water samples from the Mettawee River, AllBac, BoBac, qHF183, and BacHum (a second genetic marker for humans) were analyzed. Two human and no canid genetic markers were analyzed in the samples from the Mettawee River because humans were suspected contributors to *E. coli* contamination to a greater extent than dogs or other canids. Although feces from birds likely contribute to an unknown extent to fecal contamination in Vermont streams, reliable MST markers for birds are not yet available.

Sample Collection and Pre-Processing

Because it was not known ahead of time which samples would be analyzed for the genetic-marker tests, all water samples were pre-processed and stored in a freezer at a temperature of -112°F pending selection of the subset of

² AllBac, a qPCR assay developed to detect conserved 16S rRNA gene sequences of *Bacteroidales* species that compose about 10 percent of the total fecal mass of all mammals (Layton and others, 2006), was used to confirm that water and fecal samples contained *Bacteroidales* bacteria in detectable quantities. AllBac marker was detected in all but one water sample, which had very low concentrations of *E. coli*.

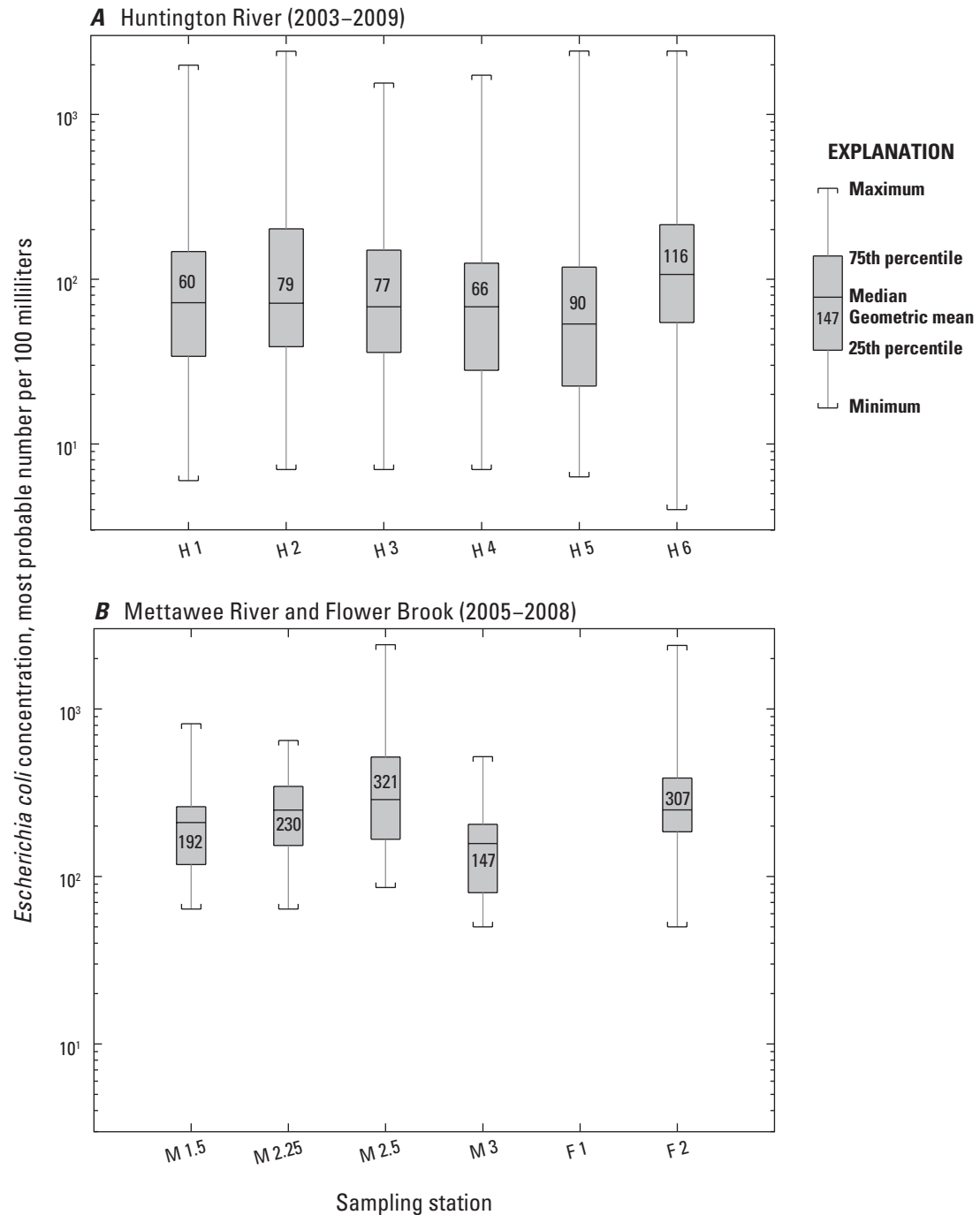


Figure 5. Summaries of historical *Escherichia coli* (*E. coli*) concentrations at selected stations in the A, Huntington (Huntington Conservation Commission, 2010) and B, Mettawee (The Poultney Mettawee Natural Resources Conservation District, 2009) Rivers. Geometric mean concentration values range from 60 to 116 most probable number (MPN) per 100 milliliters in the Huntington River and from 147 to 321 MPN per 100 milliliters in the Mettawee River and Flower Brook.

samples to be analyzed at the end of the sampling season. The initial or “pre-processing” steps of sample processing took place within hours after sample collection at the VTDEC Laboratory and concluded with freezing the samples.

Water Samples

For each water sample, three sterile prelabeled 120-mL IDEXX bottles were filled with stream water from a single point as grab samples and stored in coolers on ice for transport back to the VTDEC Laboratory. A fourth bottle was collected at one station for every set of samples collected on a given date for duplicate *E. coli* analysis. During high-flow conditions, often taking a single grab near the edge of the stream was the only safe method because few stations were at bridges. If the stream was safely wadable, the sample collector waded to a point near the center of the stream. If the stream was not safely wadable, the turbulent water was considered to be sufficiently well mixed to yield a representative sample from near the edge. Water was collected directly into the sample bottles at approximately mid-depth while the collector faced upstream.

Field parameters measured concurrently with the stream samples included specific conductance and turbidity using a YSI 600-OMS meter with a model 6136 turbidity probe and water temperature and pH using a Beckman model 255 pH meter. Both meters were checked and calibrated, if necessary, at least monthly. Once all the water samples were collected, they were driven immediately to the VTDEC Laboratory for pre-processing. If samples were collected late in the afternoon on a given day, the samples were pre-processed the following morning; however, in all cases, the samples were pre-processed within 15 hours and usually within 2–8 hours of collection.

Of the three IDEXX bottles filled with streamwater, one was for *E. coli* analysis at the VTDEC Laboratory, and the remaining two bottles were filtered in duplicate according to suggested protocols (Donald Stoeckel, U.S. Geological Survey, written commun., 2008) and stored at -112°F. Before the samples were frozen, 50 mL of an exogenous DNA spike-and-recovery control, *E. coli* carrying the plasmid-borne target gene DsRed2 (Matz and others, 1999), was added to evaluate recovery efficiency of the subsequent DNA extraction (Stoeckel and others, 2009). Samples collected at base flow were processed for *E. coli* concentrations using only the raw streamwater. For samples collected during high-flow conditions, *E. coli* concentrations were routinely analyzed using raw and diluted (1:100) streamwater, so that concentrations greater than the analytical limit of 2,419 MPN/100 mL could be quantified. Dilutions were made by pipetting 1 mL of raw river water into a clean IDEXX bottle filled with 99 mL of sterile water containing dissolved Colilert media. The VTDEC Laboratory protocol for *E. coli* analysis is the USEPA (2005) Colilert MPN method cited in the laboratory Quality Assurance plan (Vermont Department of Environmental Conservation, 2008).

Fecal Reference Samples

Fecal reference samples from four different animal reference groups were collected and subjected to the same *E. coli* and genetic-marker analyses as the water samples. Because intestinal microbe populations may vary seasonally and regionally with diet and climatic factors (Hartel and others, 2003; Stoeckel and Harwood, 2007), fecal reference samples were collected within the same time frame and space as the water samples. However, due to resource limitations, the local (Vermont) reference samples were pooled with samples from other parts of the country that had been collected and analyzed using the same techniques, after the ranges were compared to determine similarity.

The procedure for collecting fecal reference samples depended on the source of fecal material. Fresh samples, as indicated by moisture and color (or in occasional lucky cases, by observation), were targeted because they contained the most viable *E. coli* and *Bacteroidales* bacteria. Collections of cow feces at farms were scheduled ahead of time by making appointments with local farmers. Collections of dog feces also were generally arranged ahead of time by obtaining permission from dog owners to collect samples from their property. In the Huntington River watershed, dog owners were given supplies to collect their own dog feces, which were then left in a prearranged location for project personnel to retrieve later in the day. Human fecal reference samples for the Mettawee River watershed were collected from influent to the West Pawlet Wastewater-Treatment Facility³ (fig. 3) and for the Huntington River watershed by collecting a dip from the septic tank at Audubon Vermont (fig. 2). Wildlife scat (fig. 6) was collected during four field days (two days each in the Huntington and Mettawee watersheds) in or near wetlands or upland areas that were anecdotally rich in wildlife (information about wildlife management areas in Vermont, including common wildlife species, was obtained from the Vermont Fish and Wildlife Department [2009]). Although scat did not necessarily need to be identified to the species level for this project—only to separate out the ruminants (deer and moose) and canids (fox, coyote, and wolf) from other wildlife—the decision was reached to process the scat from identifiable species separately so that it might be possible to obtain more information about potential sources at a later date. Some samples could not be classified. The types and number of fecal reference samples that were collected in each watershed are shown in table 4.

The steps for collecting solid fecal reference samples were (protocol available at the USGS office in Columbus, Ohio):

³ Although it was outside of the sampled watershed, the West Pawlet Wastewater-Treatment Facility offered an accessible source of human fecal material that represented a broad diversity of composited samples from people living in the vicinity of the study area. Influent (as opposed to effluent) was used because this untreated material would most closely resemble the material that could potentially leak from septic systems.



Figure 6. Several varieties of scat collected in or near the Huntington and Mettawee River watersheds in Vermont during 2009.

1. Record the location and descriptive information,
2. Wash hands,
3. Label 50-mL centrifuge collection tube,
4. Put on new disposable gloves,
5. Place a pea-sized sample from interior of the fecal mass (or a single deer or moose pellet) into a collection tube using a sterile plastic picnic utensil,
6. Wipe outside of capped tube with ethanol,
7. Place on ice, and
8. Wash hands.

Each cooler contained a field blank sample filled with sterile buffer that was treated the same as other samples to test for cross contamination. The protocols for collecting liquid samples from wastewater influent or septic tanks were essentially the same as for solid materials, except that dip samples were collected into sterile plastic bottles.

Human fecal reference samples were sent immediately overnight on ice to the OWML. The remaining types of fecal reference samples were pre-processed at the VTDEC Laboratory. Pre-processing steps included compositing individual samples that were collected on a given day into reference groups—wild ruminants, domestic ruminants, dogs, or other types of wildlife; combining uniform subsamples of

Table 4. Sources of fecal reference samples by reference group.

[>, greater than]

Reference group	Source	Approximate number of individuals
Huntington River watershed		
Human	Septic tank at Vermont Audubon	> 6 (employees plus visitors)
Ruminant	Cow—composite from 2 farms, plus 12 beef cattle	200 ^a
Ruminant	Moose and deer	9 ^a
Ruminant	Sheep	1
Canid	Dogs	19
Wildlife	Otter	2
Wildlife	Raccoon	6
Wildlife	Composite of bear, mustelid, small rodent, and unidentified scat	5
Wildlife	Coyote	3
Mettawee River and Flower Brook watershed		
Human	Influent from West Pawlet wastewater-treatment facility	290 ^a
Ruminant	Cow—composite from 2 farms, including 10 heifers	200 ^a
Ruminant	Moose and deer	8 ^a
Canid	Dogs	7
Wildlife	Beaver	5
Wildlife	Unidentified	5
Wildlife	Bird (presumed grouse)	2
Wildlife	Coyote	2
Wildlife	Composite of raccoon, bear, muskrat	4

^a Numbers are approximate because it was unknown how many individual animals contributed to the collected samples.

the individual samples that formed each reference group in a quart-size freezer bag and kneading for at least 10 minutes to create a homogeneous reference-group sample; and dividing the reference samples into reference subgroups on the basis of specific criteria, for example, a subgroup from which information could potentially be derived at a later date by conducting targeted tests.

After each day's fecal reference samples were composited into reference groups, three pre-processing steps were done in the VTDEC Laboratory: (1) analysis of *E. coli* concentrations, (2) measurement of dry weights, and (3) determination of weights for the reference-group samples and placement into extraction tubes for genetic-marker tests. For the *E. coli* analysis, approximately 1 g of each fecal reference sample was added to 100 mL of sterile buffered water. From this aqueous sample, a 1:10 dilution was made by transferring 10 mL to 90 mL sterile buffered water. From the 1:10 dilution, three serial 1:100 dilutions were made into sterile water containing Colilert media, resulting in dilutions of 10⁻³, 10⁻⁵, and 10⁻⁷, which were analyzed for *E. coli* using the USEPA (2005) Colilert MPN test. For the second pre-processing

step, approximately 1 g of the composite was placed into a preweighed aluminum container, placed into a 140°F oven for at least 48 hours, and then weighed again to obtain the dry weight. For the third pre-processing step, approximately 0.1 g of each reference sample and 50 mL of the DsRed2 spike-and-recovery control were added to each of two plastic vials. The contents of one vial would be analyzed, and the other would be a backup duplicate. The contents of the vials were mixed using a vortex mixer and stored at -112°F until shipment to the OWML for qPCR analysis.

Use of Fecal Reference Samples to Adjust Marker Results in Water

Results for genetic markers in water are expressed as copies per 100 mL. Genetic-marker results need to be related to the fecal-indicator bacteria that are regulated by State and Federal law, which in Vermont are *E. coli* bacteria. A direct relation between concentrations of *E. coli* and genetic markers in water cannot be assumed for three reasons: *E. coli* and

genetic markers degrade at different rates in fecal material, the percent recovery of genetic markers from water and fecal material can be highly variable, and the concentrations of *E. coli* and genetic markers are highly variable in fecal material. Regarding the varying degradation rates, microcosm studies that isolated various environmental treatments on relative degradation rates have attempted to determine the relative importance of these factors; for instance, Walters and Field (2009) have shown that natural light does not affect survival rates of human markers or *E. coli* concentrations, and Dick and others (2010) have shown that decay of human markers may be similar to decay of *E. coli* under treatments of sunlight, sediment exposure, temperature, and predation. Variability caused by inconsistent recovery of DNA can be reduced by use of an exogenous DNA spike-and-recovery control, such as the DsRed2 used in this study (Stoeckel and others, 2009).

To address the variable nature of concentrations of genetic markers and *E. coli* in fecal material, three equations were used (Stoeckel and others, 2011). The first equation embodies the assumption that the ratio of the concentrations of *E. coli* to genetic marker in fresh feces approximates the ratio of the concentrations of *E. coli* to genetic marker in the water sample that is collected. Although this assumption will sometimes be untrue because *E. coli* and genetic markers decay at different rates, it is a starting point from which to address the variability issue. Terms are rearranged to solve for *E. coli* in water (equation 2). Equation 3 is used to determine whether the host associated with the genetic marker is likely or unlikely to be the source of the *E. coli* in the sample: if measured *E. coli* in water (left side of equation 3) is less than the theoretical threshold of *E. coli* in water that is calculated by the right side of equation 3, then the source associated with the genetic marker is identified as a potential contributor, perhaps one of many contributors, of fecal contamination in the stream (Stoeckel and others, 2011). However, if the theoretical threshold is less than the observed *E. coli* concentration in water, then the host associated with that genetic marker can be excluded as a major contributor to fecal contamination in the stream.

E. coli in feces/marker in feces \approx *E. coli* in water/marker in water (1)

E. coli in water = marker in water * *E. coli* in feces/marker in feces (2)

E. coli in water < marker in water * *E. coli* in feces, adjusted/marker in feces, adjusted (3)

The terms on the right side of equation 3 are adjusted to incorporate information on the natural variability of bacteria in fecal material. The adjustment, derived empirically from laboratory tests of dilute fecal suspensions of known materials (Stoeckel and others, 2011), uses information from the available pool of fecal reference samples by taking the mean *E. coli* or marker result for the particular host and adding (for the numerator) or subtracting (for the denominator) the

95-percent confidence interval. Since a larger pool of fecal reference samples better represents the range of concentrations seen in nature and increases the statistical power of the analysis, concentration data for genetic markers from Vermont fecal reference samples (table 5) were combined with concentration data from studies done in Fountain Creek, Colorado, (Stoeckel and others, 2011) and along the Ohio River, Ohio, (data for the Ohio River are available at the USGS office in Columbus, Ohio) in calculations involving equation 3.

Results of Genetic-Marker Analyses

The results of genetic-marker analyses are presented numerically as concentrations. Results for individual sampling dates are presented in this section in both text and figures in a presence/absence format along with graphical depictions of *E. coli* concentrations. This difference in presentation reflects the semiquantitative nature of the interpretive approach, whereby genetic-marker results in water are related to the *E. coli* results in water by equation 3. The point of this MST study, beyond determination of the raw genetic-marker concentrations, was to determine whether the concentrations associated with particular reference groups could adequately explain or account for fecal contamination as indicated by *E. coli* concentrations.

Fecal Reference Samples

The fecal reference samples from Vermont were similar when compared to those from Colorado and Ohio with which they were pooled. The *E. coli* concentration of the Mettawee sewage sample was 0.3 log units higher than the range of the Colorado and Ohio fecal reference samples but within the range seen in typical domestic wastewater (table 3.1, Tchobanoglous and Schroeder, 1987). The concentration of the human genetic marker BacHum also was higher by 0.3 log units in the Mettawee sewage sample than the highest concentration in the samples from Colorado and Ohio, but the concentration of the qHF183 human genetic marker in the Mettawee sewage sample was within the concentration range of the samples from the other states. It is not surprising that the variability in bacteria concentrations of sewage samples is large because dilution from industrial or commercial customers or from leaky pipes can reduce concentrations to varying extents. The result from the Huntington septic sample was not usable in this analysis because a dry weight was never obtained.

Because some of the concentrations of fecal reference samples from Vermont were outside the concentration ranges of the Colorado and Ohio samples (as described in the preceding paragraph), the ratios calculated by equation 3 using just Vermont fecal reference samples were compared to the ratios calculated by pooling the Vermont fecal reference

Table 5. Concentrations of *Escherichia coli* and host-associated genetic markers for fecal reference samples collected in Vermont, 2009.[*E. coli*, *Escherichia coli*; MPN, most probable number; BDL, below detection limit; DNQ, detected not quantified]

Sampling date	Contributing source	<i>E. coli</i> , in MPN per gram dry weight	Genetic markers				
			AllBac	BoBac	qHF183	BacHum	BacCan
Huntington River fecal reference samples							
6/13/2009	Wildlife raccoon	2.42E+08	2.87E+09	2.47E+08	BDL	BDL	BDL
6/13/2009 and 10/19/2009	Ruminant wildlife	1.55E+08	6.74E+11	6.45E+09	2.78E+05	6.37E+06	5.01E+07
6/13/2009 and 10/19/2009	Other wildlife	2.07E+07	4.24E+09	BDL	DNQ	5.77E+06	1.48E+08
9/10/2009	Ruminant cow	1.30E+06	1.18E+12	2.65E+10	2.93E+06	2.85E+07	4.71E+07
9/10/2009	Canid dog	2.42E+08	1.23E+12	1.67E+09	1.01E+07	2.29E+07	9.32E+10
11/2/2009	Human–Septic ¹	1.92E+05	4.27E+10	BDL	8.08E+08	4.43E+09	DNQ
11/2/2009	Human–Septic ¹ replicate	1.92E+05	1.80E+10	BDL	4.73E+08	2.15E+09	DNQ
Mettawee River fecal reference samples							
8/12/2009	Ruminant cow	1.73E+06	1.09E+12	1.23E+12	BDL	BDL	2.53E+07
8/12/2009	Ruminant cow replicate	1.73E+06	2.60E+12	4.61E+10	DNQ	6.62E+06	6.01E+07
8/12/2009 and 10/26/2009	Canid dog	6.61E+08	2.39E+11	6.67E+07	3.15E+07	7.20E+07	3.38E+09
8/13/2009	Human–sewage ¹	1.30E+08	2.30E+10	5.56E+06	1.08E+08	9.61E+08	1.74E+08
9/29/2009	Wildlife beaver	1.55E+05	2.18E+11	7.64E+07	4.05E+05	BDL	2.16E+05
9/29/2009 and 10/26/2009	Other wildlife	4.29E+08	3.70E+09	BDL	DNQ	2.94E+05	BDL
10/26/2009	Ruminant wildlife	2.42E+06	6.92E+11	4.61E+09	BDL	5.86E+06	1.88E+06

¹Units for *E. coli* for human fecal-reference samples are MPN per 100 milliliters, and units for all other species are copies per 100 milliliters.

samples with out-of-state samples for humans, ruminants, and canids. There was no difference for any of the human or canid results. The only difference was that six additional water samples from this study would have had a positive result for ruminants if the Vermont ruminant fecal reference samples had not been pooled with those from Colorado and Ohio. Thus, for ruminants, results based on the pooled fecal reference samples provided more conservative conclusions; pooling excluded ruminants as potential major sources of fecal contamination in more water samples than if just the Vermont fecal reference samples were used to draw conclusions about potential sources.

The second use of fecal reference samples was to explore the potential for cross reactivity of genetic-marker assays among reference groups. Several pathways can lead

to detection of host-associated genetic markers in fecal contamination from alternate hosts. Dogs on farms may eat cow manure and thereby incorporate cow-associated bacteria into their intestinal flora. Runoff from different land uses transports farm manure, wildlife scat, and pet droppings into storm sewers, where it mixes with human fecal matter in combined sewer systems. Thus, sewage influent samples used to represent a human fecal sample might actually be contaminated with feces from other sources. In addition, some *Bacteroidales* species or strains with identical 16S rRNA genetic-marker sequences could have adapted to multiple hosts.

For this study, median concentrations for the three tested potential genetic markers (BoBac for ruminant sources, qHF183 and BacHum for human sources, and BacCan for

canid sources) indicated that there was some cross reactivity among all groups (fig. 7). However, even for the human group, which showed the largest potential for cross reactivity, concentrations for the BoBac marker were approximately 22 times less than concentrations for the qHF183 marker and 13 times less than concentrations for the BacHum marker. Based on these results, host specificity, although imperfect, was considered sufficient for quantitative comparisons. The next highest potential for cross reactivity was in the canid group, for which concentrations for the BacCan marker were 56 times greater than concentrations for the BoBac marker. All other combinations of potential sources and genetic markers indicated even stronger host specificity.

Huntington River

Water samples from the Huntington River were collected on three dates (fig. 8). The first sample, collected on July 16, 2009, was essentially a base-flow sample because discharge increased from 112 to only 124 ft³/s despite approximately 0.35 in. of rainfall. The approximate rainfall total of 1.1 in. on August 2nd included almost 0.4 in. that fell in less than one hour just prior to sample collection as seen in the rising limb of the hydrograph. The October 7th rainfall total of

approximately 1 in. was similar in magnitude to that of the August storm; samples were taken concurrently with the peak of discharge. The hourly observed precipitation shown in figure 8 is from a derived and preliminary product of the Rivers Forecast Center (National Weather Service Advanced Hydrologic Prediction Service, 2010). Precipitation bars in the figures are placed at the center points of the value ranges presented online, and the amounts are approximate; however, these hourly data apply to the specific area of interest, and the rainfall totals are generally close to the unpublished total daily rainfall reported by a resident observer in the watershed (Robert Low, Richmond Conservation Commission, written commun., 2010).

E. coli concentrations for all of the high-flow samples were above Vermont standards and USEPA water-quality criteria (table 6). For the six base-flow samples, *E. coli* concentrations were all below USEPA criteria, and two were below Vermont standards. *E. coli* concentrations in samples collected on the two storm dates showed opposite longitudinal patterns (figs. 8B–C; table 6). On August 2nd, concentrations of *E. coli* decreased between upstream and downstream stations from 5,170 to 308 MPN/100 mL, and on October 7th, they increased from 435 to a maximum of

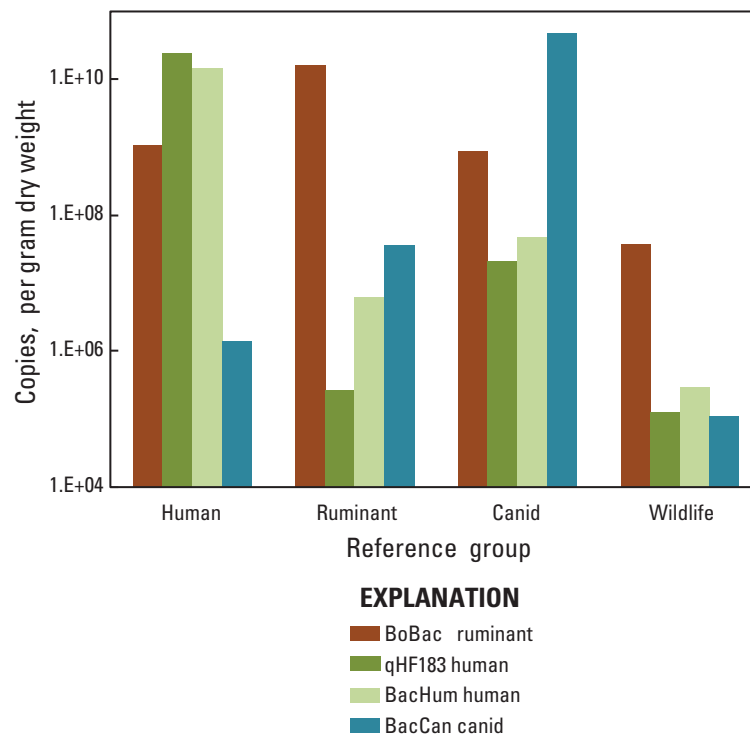


Figure 7. Median genetic-marker results showing with log-transformed data that there is some potential for cross reactivity among all reference groups.

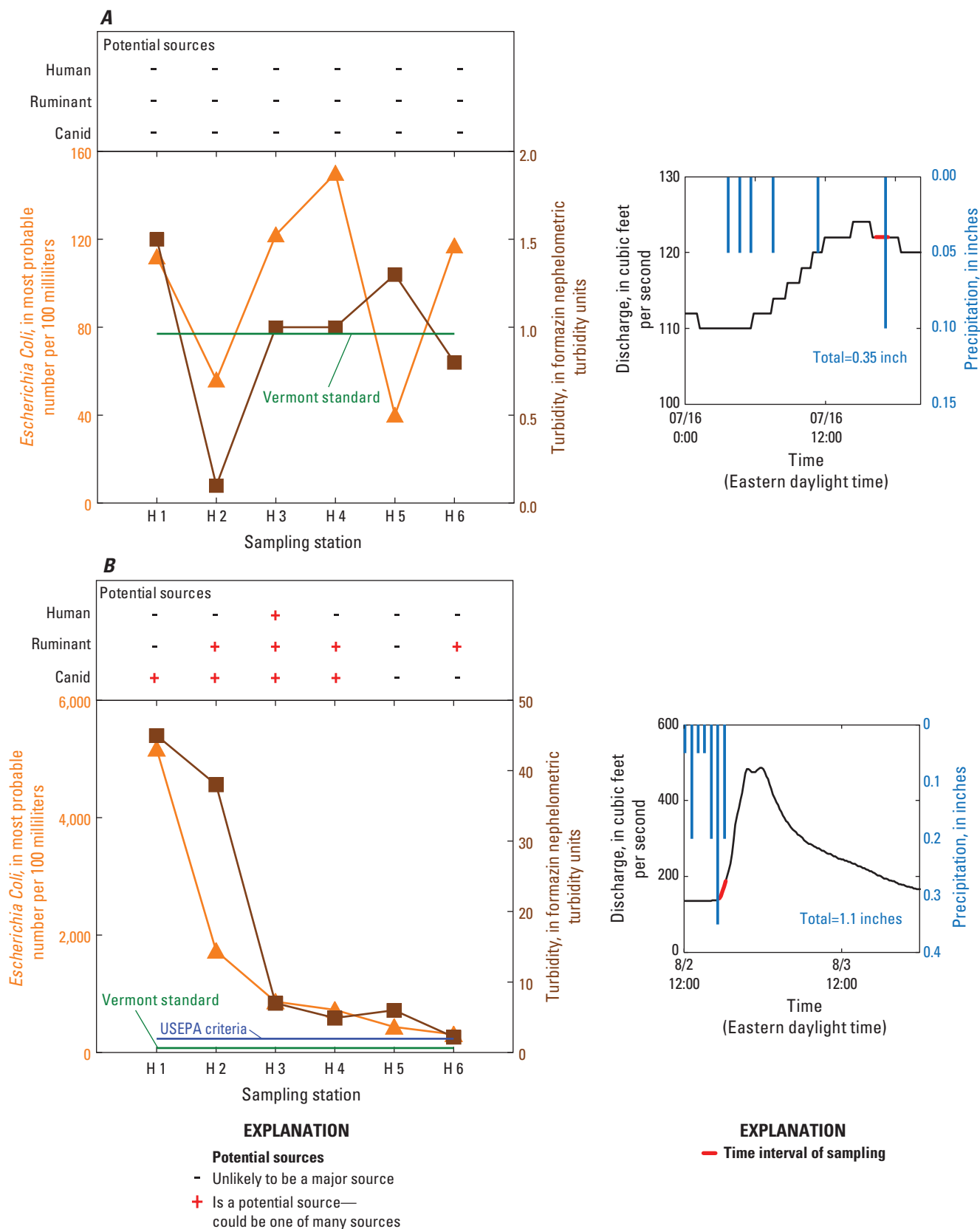


Figure 8. Concentrations of *Escherichia coli* (*E. coli*), turbidity, and potential for sources to explain observed *E. coli* concentrations in samples collected from stations on the Huntington River; discharge at the Mad River streamgage 04288000 adjusted to the Huntington River watershed by drainage area; and approximate precipitation in Huntington, Vermont, on A, July 16, 2009; B, August 2 and 3, 2009; and C, October 7 and 8, 2009.

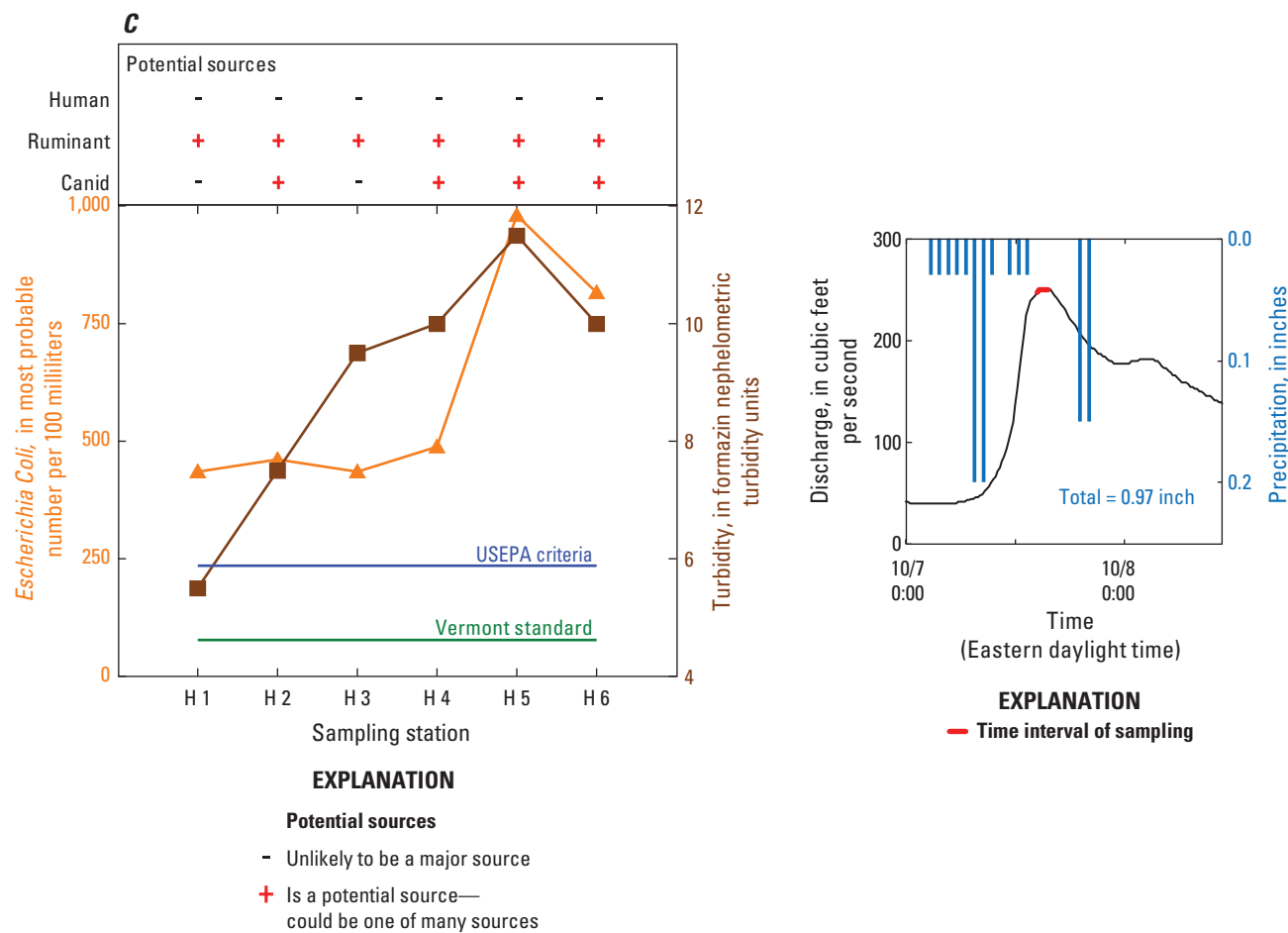


Figure 8. Concentrations of *Escherichia coli* (*E. coli*), turbidity, and potential for sources to explain observed *E. coli* concentrations in samples collected from stations on the Huntington River; discharge at the Mad River streamgage 04288000 adjusted to the Huntington River watershed by drainage area; and approximate precipitation in Huntington, Vermont, on *A*, July 16, 2009; *B*, August 2 and 3, 2009; and *C*, October 7 and 8, 2009.

Table 6. Concentrations of *Escherichia coli* and host-associated genetic markers and field parameters for water samples collected in the Huntington River in Vermont, 2009.

[*E. coli*, *Escherichia coli*; MPN, most probable number; $\mu\text{S}/\text{cm}$, microsiemens per centimeter; samples collected on 7/16/2009 were during base-flow conditions and samples collected on 8/2/2009 and 10/7/2009 were during high-flow conditions; BDL, below detection limit; DNQ, detected, not quantified; N/A, not analyzed; --, no data; bolded values indicate results that exceed the theoretical threshold for determination of sources as potential contributors to fecal contamination in the stream—it is possible for a non-bolded value to be higher than a bolded value for a given genetic marker on the same sampling date because *E. coli* concentrations, which are factors in the theoretical threshold calculation, are different for each sample.]

Sampling date	Station	<i>E. coli</i> in MPN per 100 mil- liliters	Genetic markers					Turbidity, in formazin nephelometric units	Specific conductance, in $\mu\text{S}/\text{cm}$ at 25 degrees Celsius	pH	Temperature, in degrees Celsius	
			Copies per 100 milliliters									
			AIIBac	BoBac	qHF183	BacHum	BacCan					
7/16/2009	H1	112	2.96E+05	BDL	BDL	N/A	BDL	BDL	1.5	82	6.35	18.3
7/16/2009	H2	56	4.93E+05	BDL	BDL	N/A	BDL	BDL	0.1	92	--	18.6
7/16/2009	H3	122	1.26E+06	BDL	BDL	N/A	BDL	BDL	1	95	--	18.6
7/16/2009	H4	150	1.25E+06	BDL	BDL	N/A	BDL	BDL	1	99	--	18.6
7/16/2009	H5	40	1.16E+06	BDL	BDL	N/A	BDL	BDL	1.3	100	--	19.1
7/16/2009	H6	117	1.21E+06	BDL	BDL	N/A	BDL	BDL	0.8	103	7.2	19.4
7/16/2009	H6 replicate	115	N/A	N/A	N/A	N/A	N/A	N/A	--	--	--	--
8/2/2009	H1	5,170	1.36E+07	2.15E+05	BDL	N/A	1.16E+04	45	77	--	--	--
8/2/2009	H2	1,730	5.98E+06	1.06E+05	5.99E+03	N/A	1.13E+04	38	92	--	--	--
8/2/2009	H3	866	3.97E+06	4.67E+04	1.13E+04	N/A	2.74E+04	7	97	--	--	--
8/2/2009	H4	727	2.35E+06	3.15E+04	DNQ	N/A	1.13E+04	4.9	99	--	--	--
8/2/2009	H5	435	1.85E+06	BDL	BDL	N/A	DNQ	6	101	--	--	--
8/2/2009	H5 replicate	260	2.39E+06	BDL	BDL	N/A	DNQ	--	--	--	--	--
8/2/2009	H6	308	4.12E+06	2.83E+04	BDL	N/A	DNQ	2.2	103	--	--	--
10/7/2009	H1	435	3.64E+06	2.46E+04	BDL	N/A	BDL	5.5	58	7.4	10.9	10.9
10/7/2009	H2	461	6.77E+06	1.39E+04	BDL	N/A	1.08E+04	7.5	64	7.36	11	11
10/7/2009	H2 replicate	411	N/A	N/A	N/A	N/A	N/A	--	--	--	--	--
10/7/2009	H3	435	5.96E+06	1.40E+04	BDL	N/A	DNQ	9.5	69	7.28	11.2	11.2
10/7/2009	H4	488	1.00E+07	2.42E+04	BDL	N/A	1.15E+04	10	70	6.92	11.3	11.3
10/7/2009	H5	980	1.55E+07	8.13E+04	BDL	N/A	2.09E+04	11.5	85	7.21	11.5	11.5
10/7/2009	H5 replicate	1,120	N/A	N/A	N/A	N/A	N/A	--	--	--	--	--
10/7/2009	H6	816	1.13E+07	4.11E+04	DNQ	N/A	2.15E+04	10	80	7.56	11.5	11.5

980 MPN/100 mL. Turbidity generally followed the same patterns as *E. coli* concentrations.

Genetic-marker results (table 6) were all below detection for the July 16 base-flow sample, when *E. coli* concentrations were also low. Results for the August 2 storm show that ruminants could have potentially contributed to fecal contamination at all stations but two (H1 and H5), canids could have potentially contributed to fecal contamination at the four upstream stations (H1, H2, H3, and H4), and humans could have potentially contributed to fecal contamination at one station (H3). During the storm of October 7th, there was no evidence of human sources of fecal contamination, ruminants could have contributed to fecal contamination at all sampling stations, and canids could have potentially contributed to fecal contamination at all but two sampling stations (H1 and H3), where the canid marker was below detection or quantification.

To summarize, of the 18 samples collected under different flow conditions at 6 stations, there was only one sample in which the genetic marker for humans exceeded the theoretical threshold; thus, humans probably can be excluded most of the time as a major source of fecal contamination in the Huntington River. During the two storms sampled under high-flow conditions, ruminants may have been one of perhaps several sources of the elevated *E. coli* concentrations that were measured. The evidence for canids was mixed—during the sampled high-flow events, they may have contributed to fecal contamination at some but not all stations.

Mettawee River and Flower Brook

Water samples were collected on the Mettawee River and its Flower Brook tributary during two high-flow events and once during base-flow conditions (fig. 9). Beginning on July 29, 2009, and continuing through the 30th, a total of approximately 2.85 in. of rain fell, although most of that rain fell after sample collection had ended (fig. 9A). The August 21 storm was very intense. Most (1.2 in.) of the approximately 1.65 in. of rain fell in just 1 hour, causing the Mettawee River to rise with extreme rapidity during the time interval when the samples were collected (fig. 9B). The third set of samples was collected during base-flow conditions on November 12 (fig. 9C). The hourly observed precipitation shown in figure 9 is from the same online source as described previously for the Huntington study area.

E. coli concentrations for all of the high-flow samples (table 7) were above Vermont water-quality standards and USEPA water-quality criteria. For base-flow samples, *E. coli* concentrations were below USEPA criteria at all stations and below the more restrictive Vermont standards at three stations. For all three of the sampling dates, *E. coli* concentrations were lowest at stations F1 and M1.5. Concentrations of *E. coli* were extremely high, over 20,000 MPN/100 mL, at two stations during the storm on August 21st. Concentrations of *E. coli* higher than the analytical limit of 2,419 MPN/100 mL are

not typically measured in Vermont streams because samples are not routinely diluted. Therefore, the representativeness of these high concentrations for storm conditions is unknown. For the August 21 storm, the turbidity pattern between stations was similar to that for *E. coli* concentrations. No turbidity data were available for the July 29–30 storm.

Four stations on the Mettawee River or Flower Brook (F1, M1.5, M2.25, and M3) showed evidence from genetic-marker results that humans could have potentially contributed to fecal contamination (table 7) during the sampled storms. Evidence for potential ruminant sources of fecal contamination was found in samples from one to three stations on each of the three dates. However, the stations that showed a positive genetic-marker result for ruminants differed among the dates—only one station (F2) had the same result, negative, for all three dates.

The redundancy of using two human markers in the Mettawee watershed was intended to provide additional evidence for whether humans were likely or unlikely to be a major source of fecal contamination in the Mettawee River. Results for BacHum were uniformly higher than for qHF183, when both were above detection or quantification limits (table 7). Consistent with this study, Kildare and others (2007) showed in a head-to-head comparison of the two human markers that, whereas they had similar sensitivity and specificity, the BacHum assay resulted in significantly higher numbers of copy than the qHF183 assay. Results from this study showed that for the two samples (M2.25 on July 29th and M1.5 on August 21st) in which the qHF183 marker identified humans as a potential source of fecal contamination, BacHum corroborated that conclusion. However, for 3 of the 16 cases when qHF183 results indicated that humans were unlikely to be major sources of fecal contamination, BacHum results pointed to the opposite conclusion and identified humans as a potential source.

Limitations of Study and Evaluation of Hypotheses

Reiterating what is meant by positive and negative genetic-marker results might help in the interpretation. If the theoretical *E. coli* calculated by equation (3) is less than the measured *E. coli*, the result is considered negative; this means that the genetic-marker host is unlikely to be a major source or contributor of fresh fecal contamination. The source is not necessarily ruled out, however, as potentially minor relative to other sources. If the theoretical *E. coli* calculated by equation (3) is more than the measured *E. coli*, the result is considered positive; this provides evidence, but not proof, that the source contributed to the fecal contamination. A positive result means that the particular host cannot be ruled out as a potential source, but generally alternate sources might also have contributed to the fecal contamination as measured by the indicator *E. coli* bacteria.

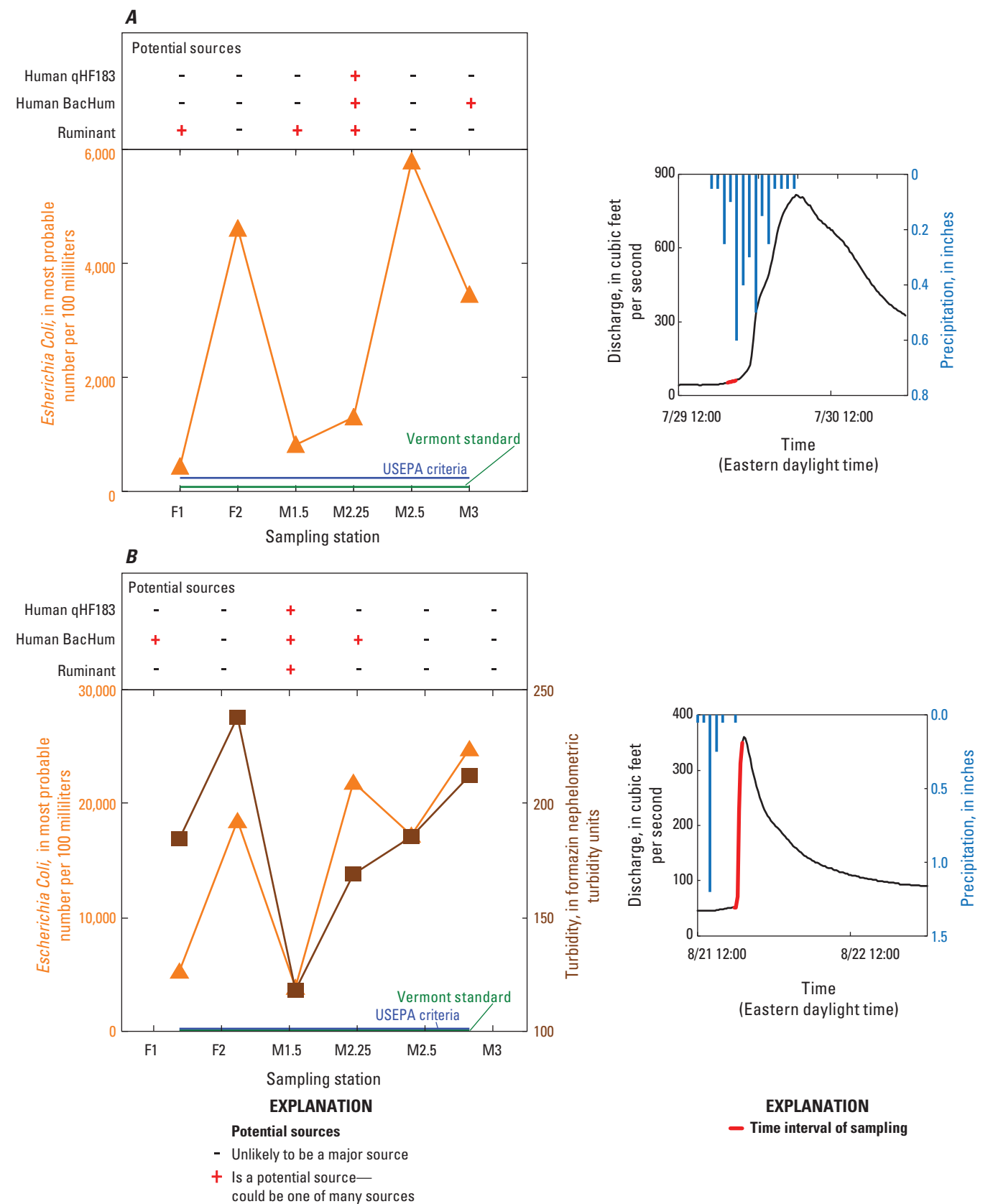


Figure 9. Concentrations of *Escherichia coli* (*E. coli*) and potential for sources to explain observed *E. coli* concentrations in samples collected from stations on the Mettawee River and Flower Brook; discharge at the Mettawee River streamgage 04280450 adjusted to the most downstream sampling station M3 by drainage area; and approximate precipitation in the Mettawee River watershed, Vermont, on *A*, July 29 and 30, 2009; *B*, August 21 and 22, 2009; and *C*, November 12, 2009.

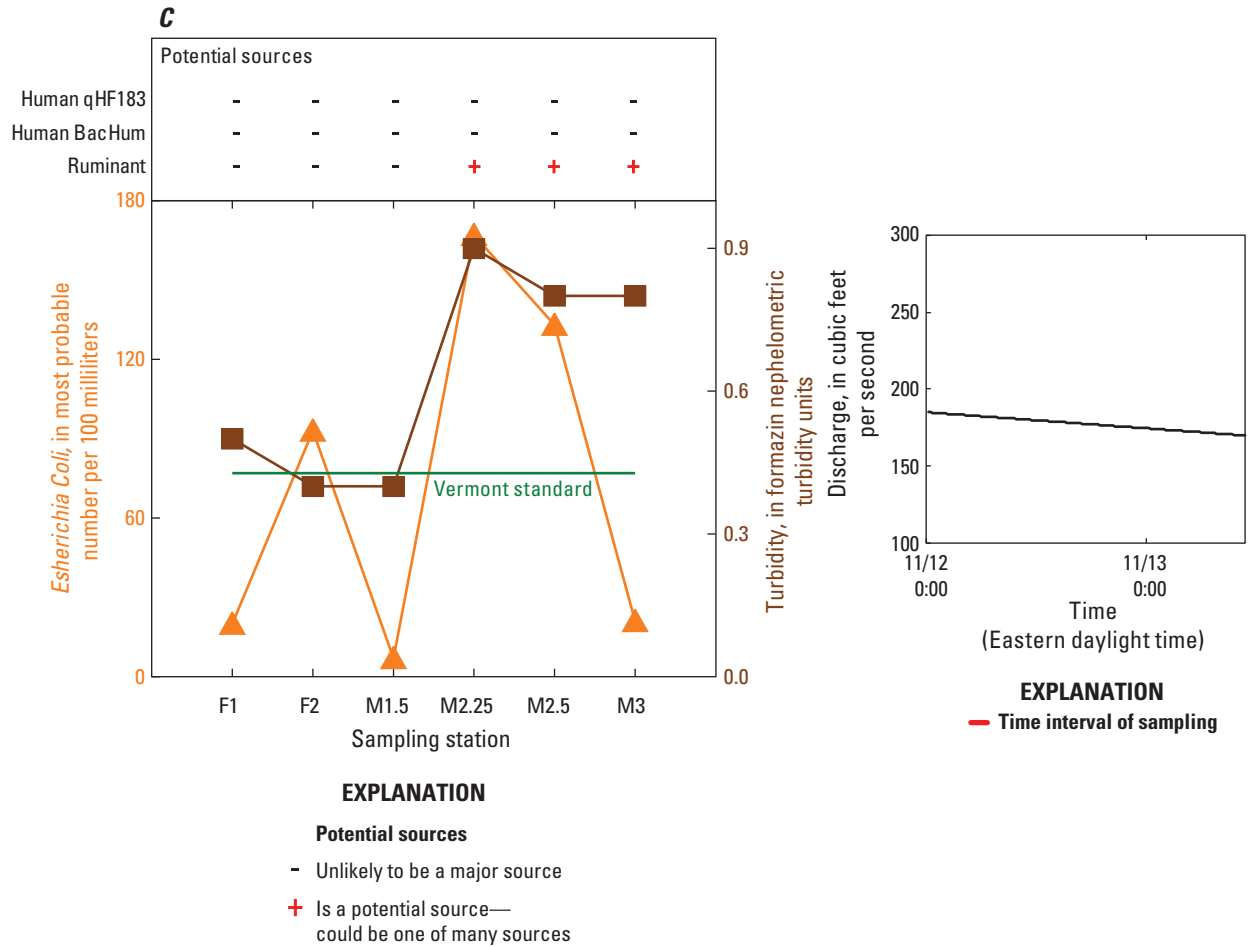


Figure 9. Concentrations of *Escherichia coli* (*E. coli*) and potential for sources to explain observed *E. coli* concentrations in samples collected from stations on the Mettawee River and Flower Brook; discharge at the Mettawee River streamgage 04280450 adjusted to the most downstream sampling station M3 by drainage area; and approximate precipitation in the Mettawee River watershed, Vermont, on A, July 29 and 30, 2009; B, August 21 and 22, 2009; and C, November 12, 2009.

Table 7. Concentrations of *Escherichia coli* and host-associated genetic markers and field parameters for water samples collected in the Mettawee River and Flower Brook in Vermont, 2009.

[*E. coli*, *Escherichia coli*; MPN, most probable number; $\mu\text{S}/\text{cm}$, microsiemens per centimeter; samples collected on 11/12/2009 were during base-flow conditions and samples collected on 7/29/2009 and 8/21/2009 were during high-flow conditions; BDL, below detection limit; DNQ, detected, not quantified; N/A, not analyzed; --, no data; **bolded** values indicate results that exceed the theoretical threshold for determination of sources as potential contributors to fecal contamination in the stream—it is possible for a non-bolded value to be higher than a bolded value for a given genetic marker on the same sampling date because *E. coli* concentrations, which are factors in the theoretical threshold calculation, are different for each sample.]

Sampling date	Station	<i>E. coli</i> , in MPN per 100 milliliters	Genetic markers					Turbidity, in Formazin nephelometric units	Specific conductance, in $\mu\text{S}/\text{cm}$ at 25 degrees Celsius	pH	Temperature, in degrees Celsius
			AIIBac	BoBac	qHF183	BacHum	BacCan				
			Copies per 100 milliliters								
7/29/2009	F1	435	4.31E+05	2.28E+04	BDL	BDL	N/A	--	260	8.1	20.5
7/29/2009	F1 replicate	N/A	4.50E+05	BDL	BDL	BDL	N/A	--	--	--	--
7/29/2009	F2	4,610	1.63E+06	4.63E+04	BDL	BDL	N/A	--	216	7.5	19.1
7/29/2009	M1.5	816	1.33E+06	1.21E+05	BDL	DNQ	N/A	--	255	7.5	17.1
7/29/2009	M2.25	1,200	2.42E+06	3.35E+05	8.40E+03	1.28E+05	N/A	--	249	7.3	17.6
7/29/2009	M2.5	5,790	2.25E+06	1.96E+05	BDL	2.95E+04	N/A	--	250	7.5	18.9
7/29/2009	M3	3,450	7.33E+06	1.14E+05	7.20E+03	1.51E+05	N/A	--	217	7.8	19.7
8/21/2009	F1	5,290	2.04E+07	1.63E+05	9.03E+03	1.09E+05	N/A	260	141	7.78	19.5
8/21/2009	F2	18,500	4.13E+07	4.13E+05	DNQ	DNQ	N/A	198	110	7.56	18.9
8/21/2009	M1.5	3,880	7.23E+07	2.15E+05	3.31E+04	3.99E+05	N/A	121	167	7.68	18.3
8/21/2009	M2.25	21,870	3.11E+08	1.68E+06	8.68E+04	6.79E+05	N/A	180	126	7.39	19.4
8/21/2009	M2.5	17,230	8.72E+07	6.63E+05	DNQ	5.27E+04	N/A	199	127	7.47	19.5
8/21/2009	M3	24,810	4.20E+07	4.33E+05	BDL	DNQ	N/A	230	131	7.7	20.6
11/12/2009	F1	20	BDL	BDL	BDL	BDL	N/A	0.4	218	8.24	5.1
11/12/2009	F2	93	1.97E+06	BDL	BDL	BDL	N/A	0.5	168	8.01	4.6
11/12/2009	M1.5	7	8.85E+05	BDL	BDL	DNQ	N/A	0.4	227	7.99	6.1
11/12/2009	M2.25	167	4.95E+06	7.87E+04	BDL	BDL	N/A	0.9	206	7.82	5.6
11/12/2009	M2.5	133	5.20E+06	8.57E+04	BDL	BDL	N/A	0.8	210	7.93	6
11/12/2009	M3	21	1.69E+06	1.17E+04	BDL	DNQ	N/A	0.8	209	8.32	6.3

Relation of Results to Land Use

Because samples were collected on only three dates in each watershed, and results for individual stations varied from date to date, results for individual stations can be related only in a general way to nearby land use. On both storm dates in the Huntington watershed, ruminants could not be ruled out as potential sources of fecal contamination at most stations. Deer and moose are ubiquitous throughout the watershed, given the proximity of large tracts of State forest and wildlife-management lands. The location of a 400-cattle dairy and beef farm, the largest farm in the area and shown as one of the reference-material collection sites in figure 2, is close to a tributary of the Huntington River and is about 2.4 river miles above the upstream sampling station H1 at Huntington Center. Results for the August 2 storm show that humans were potential sources at a station just downstream of a small residential development. Dogs or other canids were potential sources of fecal contamination in samples collected on one or both storm dates at all stations. These results do not contradict hypothesized sources related to land use. Because all genetic-marker results were negative for the base-flow sample and many were positive for the high-flow samples, it appears that nonpoint sources of potential fecal contamination from the tested host groups predominate over point or continuous sources.

In the Mettawee watershed, ruminant marker results were positive at only one station (M1.5) on both high-flow sampling dates. There are many farm fields along Flower Brook above site F1 as well as along the entire Mettawee River corridor from above station M1.5 to below station M3, and also many wild ruminants in the same vicinity. An unexpected result for the Mettawee River was that samples collected at station F2 in Pawlet village were the only ones with no positive genetic-marker results for human or ruminants for any of the sample dates.

Relation of Results to Discharge

The timing of sample collection relative to the storm hydrograph is probably related to the magnitude of *E. coli* and genetic-marker concentrations in water. Concentrations of bacteria and of suspended sediment in streams during storms are closely related because bacteria are transported through physical attachment to sediment (Rasmussen and Ziegler, 2003). Figures 8 and 9B–C demonstrate the association between turbidity, a surrogate for suspended sediment (Christensen and others, 2000), and *E. coli* concentrations. The major controlling factors for concentrations of sediment (and *E. coli*) measured at a particular station during a storm are sediment supply, storm intensity, stream velocity, water temperature, and the timing of sample collection relative to peak flow (Colby, 1956). The supply of sediment from direct surface runoff usually gets depleted as a storm progresses,

resulting in greater sediment discharge during the stream rise compared to the recession (Miller and others, 2007).

High *E. coli* concentrations observed near the beginning of the stream responses to storms in this study (figs. 8A–B) were likely caused by the large initial flush of sediment being washed into the streams, augmented to some extent by finite reserves of *E. coli* from resuspended sediment (Jamieson and others, 2005). The range of *E. coli* concentrations is similar for August 2nd in the Huntington watershed (fig. 8B) and July 30 in the Mettawee Watershed (fig. 9A)—both sample sets were collected at the beginning of the rising limb. The highest *E. coli* concentrations in the study were measured in the Mettawee watershed during the intense August 21 storm, when samples were collected during the hour when discharge increased 9-fold from 40 to 360 ft³/s (fig. 9B). When samples were collected from the Huntington River during peak discharge on October 7th, concentrations of *E. coli* were only moderately high (fig. 8C) compared to samples collected during storms while the streams were rising.

Ratios of *E. coli* to Genetic-Marker Material in Fecal Material and in Water

How important—and how realistic—is the assumption that the ratio of *E. coli* to genetic marker is essentially the same in freshly deposited fecal material as it is in fecal material transported into a stream via runoff, as shown in equation 1? The validity of the assumption in this study can be addressed by comparing the ratios of log-transformed *E. coli* concentrations to log-transformed genetic-marker concentrations for fecal reference and water samples directly and seeing that, within each potential reference group, the two median values are within an order of magnitude of one another (fig. 10). Because that ratio is variable, direct quantification of fecal material in the water samples based on genetic-marker tests is precluded to some extent (Field and Samadpour, 2007). That is to say, although analytical results for genetic markers are numerical, they only can be used to deduce source allocations of fecal contamination indicated by *E. coli* concentrations if the assumption of equal ratios is true (equation 1).

On land, *Bacteroidales* and *E. coli* survival rates may differ because *Bacteroidales* are obligate anaerobes unable to survive in the presence of oxygen, whereas *E. coli* bacteria are facultative anaerobes with the ability to survive and reproduce in soils (Ishii and others, 2006). Another factor that may influence the validity of the assumption of equal ratios is that *Bacteroidales* genetic-marker assays detect undegraded DNA from both living and dead bacteria (qPCR measures DNA persistence), whereas the method for measuring *E. coli* concentration relies on cell culture and quantifies only living and reproductively competent bacteria (bacteria survival). At the time of this study (2009), no studies of relative decay rates of *Bacteroidales* genetic markers and *E. coli* bacteria in land-deposited feces had been found in the literature.

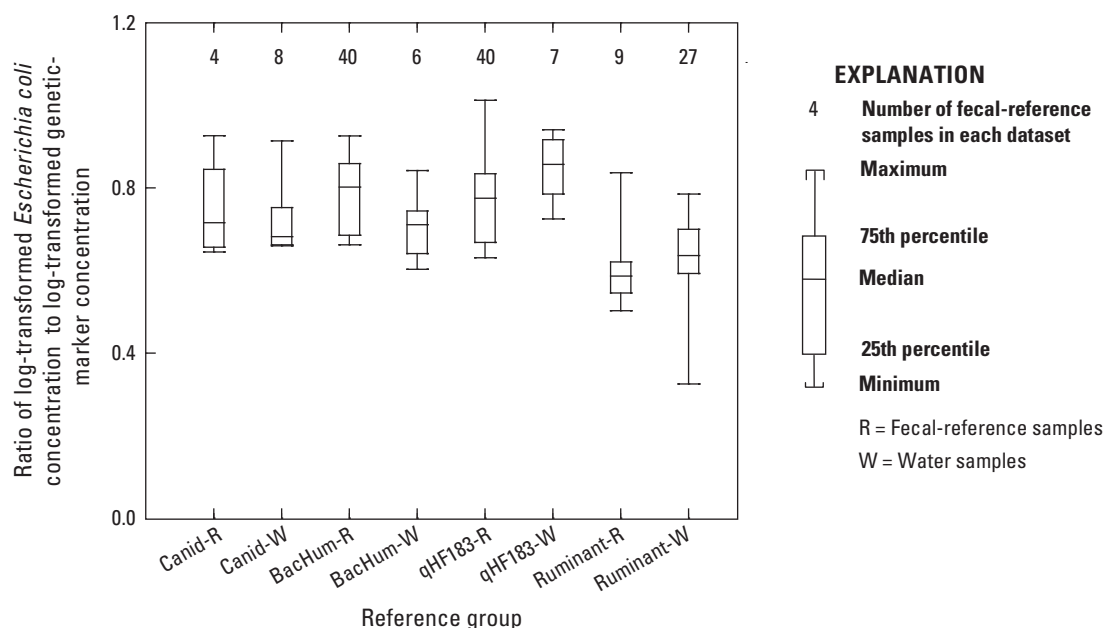


Figure 10. Ratios of log-transformed (base 10) *Escherichia coli* concentration to log-transformed (base 10) genetic-marker concentration for all fecal-reference samples used to compute theoretical thresholds and water samples in the Huntington and Mettawee Rivers. Data include fecal-reference samples from Vermont, Ohio, and Colorado.

Once fecal material reaches surface water, additional factors cause differences in persistence and survival among different genetic markers and between certain genetic markers and *E. coli* bacteria. Dick and others (2010) showed that the human marker qHF183 decays more rapidly than *E. coli* bacteria under conditions of exposure to light, higher sediment, and reduced autochthonous predation. Walters and Field (2009) found that ruminant markers persisted and survived longer than human markers, that decay rates of markers for both humans and ruminants differed from the decay rates of *E. coli*, and that light versus dark conditions differentially affected the survival of human and ruminant markers and *E. coli* bacteria. In another study, temporal patterns in rainfall, physical measures of water quality, and animal migration may explain seasonal differences in relative quantities of *E. coli* and ruminant-specific *Bacteroidales* genetic markers in bays and streamwater (Shanks and others, 2006). Anaerobic ruminant bacteria have been found to survive for extended periods of time in sediments associated with the water column (Lee and others, 2008).

Despite the uncertainty of relative decay rates of *Bacteroidales* genetic markers and *E. coli* bacteria in feces on land and in water, it can be assumed that the more recently fecal material has been deposited in a watershed, the more likely its ratio of concentrations of genetic marker to *E. coli* will remain stable when the fecal material reaches surface water. In this study, we attempted to collect fecal reference

samples that were fresh. However, source material in the water samples collected during runoff events almost certainly had decayed to an unknown extent on land prior to reaching the water.

Elapsed time since the previous sizeable runoff event could serve as a surrogate for freshness of fecal material. While more time elapsed probably means that more of the fecal material in runoff is not fresh, more time also generally translates into greater accumulation of fecal material in the watershed. Rainfalls large enough to double the magnitude of discharge at the Mettawee River streamgage station were judged to be sizeable runoff events; the time elapsed between the doubled discharge and collection of the sample gave an approximation of how long fecal material had accumulated in the watershed (table 8). Because there was no streamgage on the Huntington River, the times elapsed between the sample-collection dates and rainfalls of at least 0.25 and at least 0.5 in. were used as proxies for sizeable runoff events and for accumulation of fecal material in the watershed. The data in table 8 indicate that samples collected from the Huntington River might contain more fresh fecal material than samples from the Mettawee River; in particular, greater proportions of the *E. coli* and *Bacteroidales* bacteria in the August 2 water samples were likely from fresher sources than in samples collected on the other dates. Perhaps the finding of a large number of positive genetic-marker results in the August 2 samples is a direct manifestation of this relative

Table 8. Time since previous substantial runoff event for sample collection dates in the Huntington and Mettawee Rivers.

[Rainfall data are from observed precipitation graphics for the study area from the National Weather Service Advanced Hydrologic Prediction Service, 2010; Data from U.S. Geological Survey streamgage 04280450; --, no data]

Collection date	Days since rainfall of at least		Days since discharge doubled
	0.25 inch	0.5 inch	
Huntington River watershed			
July 16, 2009	4	4	--
August 2, 2009	2	2	--
October 7, 2009	1	4	--
Mettawee River watershed			
July 29, 2009	7	7	7
August 21, 2009	2	18	18
November 12, 2009	11	18	11

freshness. In theory, with large gaps between runoff events, *E. coli* in fecal material on land could have become enriched relative to *Bacteroidales* genetic markers. This enrichment would increase the value of the left side of equation 3, *E. coli* concentration in water, relative to the right side of the equation. In this way, streamwater samples with a large percentage of aged fecal material from runoff could lead to an interpretation of fewer positive genetic-marker results than if the same original fecal material had been fresher upon delivery into streamwater.

The relation between fecal freshness and genetic-marker results in streamwater might have implications for health risk. Ultimately, the goal of applying the MST tool to recreational surface waters is to assess the amount and source of contamination from human pathogens that pose a health risk for recreational use. Since direct measurements of human pathogens are not generally practical on a widespread basis, indicator genetic markers and bacteria are used as surrogates for human pathogens. The applicability of the results depends on whether important human pathogens associated with feces decay at rates similar to *Bacteroidales* genetic markers or *E. coli* bacteria. In a freshwater microcosm experiment, Walters and Field (2009) found that decay rates of human and ruminant-specific *Bacteroidales* genetic markers were comparable to that of the salmonella-causing human pathogen *S. typhimurium* and that decay rates of human-specific genetic markers were comparable to the decay rates of various infectious human enteric viruses.

Summary and Conclusions

The U.S. Geological Survey (USGS), in cooperation with the Vermont Department of Environmental Conservation (VTDEC), collected data in 2009 on concentrations of *Escherichia coli* (*E. coli*) and *Bacteroidales*-based genetic markers in the Huntington and Mettawee River watersheds. The study used *Bacteroidales* genetic markers to investigate potential sources of fecal contamination in two impaired Vermont streams and to assess the applicability of this method for use in other streams in Vermont. The selected watersheds differed with respect to land use and hypothesized sources of historically high concentrations of *E. coli* bacteria. Concentrations of *E. coli* and of genetic markers for humans (qHF183 and BacHum), wild and domestic ruminants (BoBac), and canids (dogs and their wild relatives) were the analytes investigated in this study.

Three sets of water samples were collected in each watershed, one during base-flow conditions and two during high-flow events. An auxiliary screening test using optical brighteners did not reveal any potential conduits of leaks from septic tanks to either the Huntington or Mettawee Rivers. Fecal reference samples from Vermont were collected during the same season and in or near the same watersheds as the water samples to test for cross reactivity of genetic markers and also, when results were pooled with results from fecal reference samples from Ohio and Colorado in order to increase statistical power, to determine whether particular source groups were potential contributors of fecal contamination. The greatest potential for an error in interpretation due to cross reactivity would be to falsely identify a human source as a ruminant source—the chances of false identification are 1 in 13 for the BacHum marker and 1 in 22 for the qHF183 marker. Concentrations of *E. coli* and genetic markers were probably related to the timing of sample collection on the storm hydrograph.

On the Huntington River, concentrations of *E. coli* bacteria during two sampled storms in August and October 2009 exceeded Vermont water-quality standards and Federal criteria at all stations. During base flow in July 2009, concentrations of *E. coli* were below Federal criteria at all stations. Except for one positive result at one station, humans were unlikely to be a major source of fecal contamination in the Huntington River at all stations on the dates that were sampled. Ruminants, humans, and canids were unlikely to be major sources during the sampled base flow. The two sampled storms provided evidence that ruminants could be a common source of fecal contamination in the Huntington River and that canids could be an intermittent source of fecal contamination. Other types of animals that were not tested also could be contributing to the fecal contamination.

In the Mettawee River watershed, as in the Huntington watershed, concentrations of *E. coli* bacteria exceeded Federal water-quality criteria at all stations during high-flow events and were below Federal criteria at all stations during sampled base flow. Results for humans and ruminants were mixed.

Some results provided evidence that, during the two storms, either source may have contributed to fecal contamination, whereas other results for storms showed evidence that humans and ruminants were unlikely to be major sources of fecal contamination. Comparing the two sampled storms, three of the six stations showed consistent results for humans, and four stations had consistent results for ruminants. As in the Huntington watershed, other types of animals that were not tested could be contributing to the fecal contamination. Results from the base-flow sample showed that humans were unlikely to be a major source of fecal contamination and that ruminants were a potential source at three stations on the main stem of the Mettawee River.

An important factor to consider when interpreting results of tests using genetic markers is the freshness of the fecal material. The fresher the fecal material in runoff, the more stable the ratio of *E. coli* to genetic marker is likely to be between the original feces and the fecal material measured in the water sample. A comparison of the ratios of concentration of *E. coli* to genetic markers for fecal reference and water samples shows that within each potential source group, the two median values are within an order of magnitude of one another. Differences between the reproductive abilities of *Bacteroidales* and *E. coli* bacteria result in differences in their respective rates of survival on land. The nature of the differences between the laboratory tests for the two species also may confound the interpretation of results pertaining to their relative concentrations: the test for *Bacteroidales* measures DNA in live and dead cellular material, whereas the test for *E. coli* bacteria measures only live cells. Once fecal material reaches surface water in runoff, these and additional factors influence the rates of decay of different species of bacteria. Surrogates for freshness of the fecal material in runoff suggest that for the dates sampled in this study, samples from the Huntington River contained fresher fecal material than samples from the Mettawee River.

The potential challenges to a straightforward interpretation of host-associated genetic-marker tests are many: variability of results related to the timing of sample collection on the storm hydrograph; variability of *E. coli* and *Bacteroidales* concentrations in fecal reference samples and in water; sampling and analytical errors; potential cross reactivity of host-associated genetic markers; and differential persistence and survival rates of *E. coli* and human, ruminant, and canid *Bacteroidales* genetic markers on land, in water, and by season. These factors have been identified in previous studies and are typically managed through caution against overinterpreting results.

Caution also must be used in the interpretation of these study results when considering the risk of human exposure to contamination from contact recreation (swimming and boating activities). Results show high fecal counts (measured as *E. coli* bacteria) associated with high-flow events (turbid water from storm runoff). When evaluating recreational risk of exposure, the applicability of the results for markers depends on whether important human pathogens associated with feces

decay at rates similar to *Bacteroidales* genetic markers or *E. coli* bacteria. We do not know if decay rates for pathogens and markers are similar. If they are, then the finding of fewer positive marker results when there are long periods between runoff events could mean the recreational risks may be less compared to risks that are calculated under current policies based on *E. coli* concentrations.

This study was intended to be exploratory—to provide information to the VTDEC about the application of a contemporary genetic-marker approach for investigating potential sources of fecal contamination to impaired streams so that the State can determine whether the method would be tenable in Vermont on a larger scale. The careful wording of results and conclusions from this study reflects the realities of numerous confounding factors that interfere with the ability to directly relate *Bacteroidales* concentrations to *E. coli* concentrations in water samples. While use of *Bacteroidales* genetic markers as a source identifying tool coupled with the interpretive approach described in this report does not provide quantitative sources or loads, an important strength and potential benefit of the method is to exclude potential sources as major contributors to fecal contamination.

References Cited

- Bernhard, A.E., and Field, K.G., 2000, A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA: *Applied and Environmental Microbiology*, v. 66, no. 10, p. 4571–4574.
- Christensen, V.G., Jian, Xiaodong, and Ziegler, A.C., 2000, Regression analysis and real-time water-quality monitoring to estimate constituent concentrations, loads, and yields in the Little Arkansas River, south-central Kansas, 1995–99: U.S. Geological Survey Water-Resources Investigations Report 00–4126, 36 p.
- Colby, B.R., 1956, Relationship of sediment discharge to streamflow: U.S. Geological Survey Open-File Report 56–27, 170 p.
- Cole, D., Long, S.C., and Sobsey, M.D., 2003, Evaluation of F+ RNA and DNA coliphages as source-specific indicators of fecal contamination of surface waters: *Applied and Environmental Microbiology*, v. 69, no. 11, p. 6507–6514.
- Dick, L.K., Stelzer, E.A., Bertke, E.E., Fong, D.L., and Stoeckel, D.M., 2010, Relative decay of *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater microcosms: *Applied and Environmental Microbiology*, v. 76, no. 10, p. 3255–3262.

- Dickerson, J.W., Jr., Hagedorn, C., and Hassall, A., 2007, Detection and remediation of human-origin pollution at two public beaches in Virginia using multiple source tracking methods: *Water Research*, v. 41, p. 3758–3770.
- Doll, C.G., Cady, W.M., Thompson, J.B., Jr., and Billings, M.P., eds., 1961, Centennial geologic map of Vermont: Montpelier, Vt., Vermont Geological Survey, scale 1:250,000.
- Dufour, A.P., 1984, Health effects criteria for fresh recreational waters: Washington, D.C., U.S. Environmental Protection Agency, Office of Research and Development, EPA-600/1-84-004, 33 p.
- Field, K.G., and Samadpour, Mansour, 2007, Fecal source tracking, the indicator paradigm, and managing water quality: *Water Research*, v. 41, no. 16, p. 3517–3538.
- Griffith, J.F., Weisberg, S.B., and McGee, C.D., 2003, Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples: *Journal of Water and Health*, v. 01, p. 141–151.
- Hartel, P.G., Summer, J.D., and Segars, W.I., 2003, Deer diet affects ribotype diversity of *Escherichia coli* for bacterial source tracking: *Water Research*, v. 37, p. 3263–3268.
- Hartel, P.G., Rodgers, Karen, Moody, G.L., Hemmings, S.N.J., Fisher, J.A., and McDonald, J.L., 2008, Combining targeted sampling and fluorometry to identify human fecal contamination in a freshwater creek: *Journal of Water and Health*, v. 06.1, p. 105–116.
- Hickoff, Steve, and Plumley, Rhey, 2007, Flyfisher's guide to northern New England—Vermont, New Hampshire, and Maine (2d ed.): Belgrade, Mont., Wilderness Adventures Press, Inc., 384 p.
- Hornbeck, J.W., Bailey, S.W., Buso, D.C., and Shanley, J.B., 1997, Streamwater chemistry and nutrient budgets for forested watersheds in New England: Variability and management implications: *Forest Ecology and Management*, v. 93, p. 73–89.
- Ishii, S., Ksoll, W.B., Hicks, R.E., and Sadowsky, M.J., 2006, Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds: *Applied and Environmental Microbiology*, v. 72, no. 1, p. 612–621.
- Jamieson, R.C., Joy, D.M., Lee, H., Kostaschuk, R., and Gordon, R.J., 2005, Resuspension of sediment-associated *Escherichia coli* in a natural stream: *Journal of Environmental Quality*, v. 34, p. 581–589.
- Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., and Wuertz, Stefan, 2007, 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach: *Water Research*, v. 41, p. 3701–3715.
- Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., and Sayler, G., 2006, Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water: *Applied and Environmental Microbiology*, v. 72, no. 6, p. 4214–4224.
- Lee, Y., Molina, Marirosa, Santo Domingo, J.W., Willis, J.D., Cyterski, Michael, Endale, D.M., and Shanks, O.C., 2008, Temporal assessment of the impact of exposure to cow feces in two watersheds by multiple host-specific PCR assays: *Applied and Environmental Microbiology*, v. 74, no. 22, p. 6839–6847.
- Matz, M.W., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zarskiy, A.G., Markelov, M.S., and Lukyanov, S.A., 1999, Fluorescent proteins from nonbioluminescent Anthozoa species: *Nature Biotechnology*, v. 17, p. 969–973.
- Miller, C.V., Gutiérrez-Magness, A.L., Majedi, B.L.F., and Foster, G.D., 2007, Water quality in the Upper Anacostia River, Maryland: Continuous and discrete monitoring with simulations to estimate concentrations and yields, 2003–05: U.S. Geological Survey Scientific Investigations Report 2007–5142, 43 p.
- National Oceanic and Atmospheric Administration, 2002, Climatology of the United States No. 81: Monthly station normals of temperature, precipitation, and heating and cooling degree days 1971–2000: Vermont, National Climatic Data Center, N.C.
- National Weather Service Advanced Hydrologic Prediction Service, 2010, Precipitation map, accessed July 1, 2010, at http://www.srh.noaa.gov/rfcshare/precip_analysis_hourly.php.
- Natural Resources Conservation Service, 2008, GeologicSoils_SO, GIS Database Administrator, accessed November 23, 2009, at <http://www.vcgi.org/dataware/>.
- Rasmussen, P.P., and Ziegler, A.C., 2003, Comparison and continuous estimates of fecal coliform and *Escherichia coli* bacteria in selected Kansas streams, May 1999 through April 2002: U.S. Geological Survey Water-Resources Investigations Report 03–4056, 87 p.
- Seurinck, S., Defoirdt, T., Verstraete, W., and Siciliano, S.D., 2005, Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human fecal pollution in freshwater: *Environmental Microbiology*, v. 7, no. 2, p. 249–259.

- Shanks, O.C., Nietch, Christopher, Sominish, Michael, Younger, Melissa, Reynolds, Don, and Field, K.G., 2006, Basin-wide analysis of the dynamics of fecal contamination and fecal source identification in Tillamook Bay, Oregon: *Applied and Environmental Microbiology*, v. 72, no. 8, p. 5537–5546.
- Simpson, J.M., Domingo, J.W.S., and Reasoner, D.J., 2004, Assessment of equine fecal contamination: The search for alternative bacterial source-tracking targets: *FEMS Microbiology Ecology*, v. 47, no. 1, p. 65–75.
- State of Vermont Natural Resources Board, 2008, Vermont water quality standards, Vt. Code R. 12 004 052: Montpelier, Vt., 72 p.
- Stoeckel, D.M., and Harwood, V.J., 2007, Minireview: Performance, design, and analysis in microbial source tracking studies: *Applied and Environmental Microbiology*, v. 73, no. 8, p. 2405–2415.
- Stoeckel, D.M., Mathes, M.V., Hyer, K.E., Hagedorn, C., Kator, H., Lukasik, J., O'Brien, T.L., Fenger, T.W., Samadpour, M., Strickler, K.M., and Wiggins, B.A., 2004, Comparison of seven protocols to identify fecal contamination sources using *Escherichia coli*: *Environmental Science Technology*, chap. 15, v. 38, no. 22, p. 6109–6117.
- Stoeckel, D.M., Stelzer, E.A., and Dick, L.K., 2009, Evaluation of two spike-and-recovery controls for assessment of extraction efficiency in microbial source tracking studies: *Water Research*, v. 43, p. 4820–4827.
- Stoeckel, D.M., Stelzer, E.A., Stogner, R.W., and Mau, D.P., 2011, Semi-quantitative evaluation of fecal contamination potential by human and ruminant sources using multiple lines of evidence: *Water Research*, v. 45, p. 3225–3244.
- Tchobanoglous, G., and Schroeder, E., 1987, *Water quality: Characteristics, modeling, modification*: Addison-Wesley Publishing Company, 768 p.
- Thompson, P.J., and Thompson, T.B., 1992, Bedrock geology of the Camels Hump—Bolton Mountain area North Central Vermont: Vermont Geological Survey, Special Bulletin no. 12, 4 pls.
- Turner Designs, 2002, TD-700 Laboratory fluorometer operating manual version 2.0, accessed September 9, 2008, at http://www.turnerdesigns.com/t2/doc/manuals/td700_manual.pdf.
- U.S. Census Bureau, 2001, DemoCensus_BLACK2000, accessed December 12, 2008, at http://www.vcgi.org/dataware/?page=../search_tools/search_action.cfm&query=theme&theme_id=016-0040.
- U.S. Environmental Protection Agency, 1986, Ambient water quality criteria for bacteria-1986. EPA/440/5–84–002, Office of Water Regulations and Standards: U.S. Environmental Protection Agency, Washington, D.C., 18 p.
- U.S. Environmental Protection Agency, 2005, Microbial Source Tracking Guide document, EPA/600/R–05/064, 123 p.
- U.S. Environmental Protection Agency, 2007, An approach for using load duration curves in the development of TMDLs, EPA 841–B–07–006, 68 p.
- University of Vermont Spatial Analysis Lab, 2005, 2002, Land cover/land use dataset for Vermont, accessed August 3, 2006, at <http://www.vcgi.org/dataware/default.cfm?layer=LCLU2002>.
- Vermont Department of Environmental Conservation, 2008, LaRosa Laboratory quality assurance plan, VT DEC, Waterbury, Vt., accessed January 21, 2010, at <http://www.anr.state.vt.us/dec/lab/docs/Section1.pdf>.
- Vermont Fish and Wildlife Department, 2009, Wildlife Management Areas, accessed September 28, 2009, at http://www.vtfishandwildlife.com/wma_maps.cfm.
- Walters, S.P., and Field, K.G., 2009, Survival and persistence of human and ruminant-specific fecal *Bacteroidales* in freshwater microcosms: *Environmental Microbiology*, v. 11, no. 6, p. 1410–1421.

Prepared by the Pembroke Publishing Service Center.

For more information concerning this report, contact:

Director
U.S. Geological Survey
New Hampshire-Vermont Water Science Center
331 Commerce Way, Suite 2
Pembroke, NH 03275
dc_nh@usgs.gov

or visit our Web site at:
<http://nh.water.usgs.gov>

