Diversity of *Escherichia coli* in the Salish Sea: Potential Implications to Marine and Human Health

Jenny L. Grunwald

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Public Health

University of Washington

2022

Committee:

Marilyn C. Roberts

Peter Rabinowitz

Daniel Enquobahrie

Program Authorized to Offer Degree:

Health Systems and Population Health

School of Public Health

University of Washington

Abstract

Diversity of *Escherichia coli* in the Salish Sea: Potential Implications to Marine and Human Health

Jenny L. Grunwald

Chair of the Supervisory Committee:
Marilyn C. Roberts, PhD
Department of Environmental and Occupational Health Services

E. coli is a species of enteric bacteria found in the intestinal tract of humans and animals that can also persist in the environment. Anthropogenic activity has led to opportunistic pathogenic *E. coli* from humans and animals contaminating environments through the discharge of fecal wastes in sewage and agricultural runoff. Studies of the diversity of *E. coli* found in environments near anthropogenic sources are needed to understand the risk to human and animal health. Gaps remain in scientific knowledge about *E. coli* diversity, survivability, and potential public health threats in marine environments. The Salish Sea is an area undergoing fast population growth with increasing nutrients and oxygen depletion due to the activity of wastewater treatment plants, and other human activity. The anthropogenic activity may impact the survivability and concentration of *E. coli* in the environment. We analyzed a dataset of 332 *E. coli* isolates from the Salish Sea ecosystem. A total of 196 multi-locus sequence types (ST) were identified from samples taken from marine water near shellfish beds, marine wildlife, river otters and a small number of freshwater samples from inlets into the Salish Sea. The isolates were compared to metadata for *E. coli* STs in EnteroBase, an international *E. coli* database. Additional information on *E. coli* STs was derived from searches of published studies in PubMed. We analyzed these data descriptively and spatially. Of the 196 ST in the data set, 80 (40.8%) did not have literature available in PubMed. For five STs: (5869, 1065, 11343, 9001, and 2164), no previous evidence of occurrence in marine water or sea mammals was identified. The greatest diversity of *E. coli* STs was found in the Strait of Juan de Fuca, an area which has wastewater treatment plants that dump partially treated wastewater into the Strait. These findings illustrate the diversity of environmental E coli in the marine environment and the need for further research the human and animal health implications of this diversity.

TABLE OF CONTENTS

List of Appendices

List of Tables

List of Figures

Abbreviations and Acronyms

EAEC: Enteroaggregative *E. coli*

E. coli: *Escherichia coli*

EIEC: Entero-invasive *E. coli*

EHEC: Enterohemorrhagic *E. coli*

EPEC: Enteropathogenic *E. coli*

ETEC: Enterotoxigenic *E. coli*

ExPEC: Extra-intestinal pathogenic *E. coli*

MLST: Multilocus Sequence Types

ST: Sequence Type

STEC: Shiga toxin-producing *E. coli*

UPEC: Uropathogenic *E. coli*

WWTP: Wastewater treatment plant

VTEC: Verotoxigenic *E. coli*

Acknowledgments

First, I would like to thank my student advisors, Christina Hughes and Vickie Ramirez. Vickie worked with me over the last 2 years to not only achieve my graduate certificate in One Health but helped link me to this project. Christina, you have been a constant cheerleader, helping me through some very difficult times, and always a source of constant encouragement. Thank you for putting up with me and always being there to answer my multitude of questions.

Next, I would like to thank Dr. Peter Rabinowitz and Dr. Marilyn Roberts. Thank you both for your patience and grace with me during these difficult past two years. Thank you to Dr. Daniel Enquobahrie for being on the thesis committee.

I would like to thank my friends and co-workers. You all have been my rock and we have gone through some crazy times. Even though we were all going through something extremely difficult, you always asked how classes were going, listened to my ongoing frustrations, and offered support and encouragement. We held each other up and I will never forget it

And finally, thank you to my husband. You have been on this journey with me since day one. Thank you for helping with the constant puppy shenanigans. Look I did it!

INTRODUCTION

While *E. coli* is part of the normal flora of warm-blooded mammals' intestines, it can be an opportunistic pathogen, causing severe disease inside and outside of the intestinal tract (CDC, 2014). *E. coli* is known to survive in freshwater, and there is increasing evidence that it can also survive in marine environments (Jang et al., 2017; Quero et al., 2015). There is tremendous genetic diversity among *E. coli*, and Multi-locus Sequencing Typing (MLST) has been used to partition *E. coli* into genetically related groups by assigning sequence types (ST) based on allelic variation into specified housekeeping genes (Ibarz Pavón and Maiden, 2009). Since its inception, MLST has been used to characterize the epidemiologic behavior of various clones of *E. coli* (Jolley et al., 2018; Vingino et al., 2021). EnteroBase is a public database that characterizes diversity within several bacterial genera, including *E. coli* [https://enterobase.warwick.ac.uk/]. The database is voluntarily populated with laboratory and metadata, including information on specific STs. Users enter metadata including sources (animal, human or environmental), source type, source details, country of origin, and lab details. The data is not centrally curated, so there is potential for data heterogeneity. Despite these limitations, EnteroBase data have been used for epidemiologic investigation of *E. coli* (Zhou et al., 2020).

There are multiple routes by which *E. coli* can enter water systems. Animal manure is often used for fertilizer in crop agriculture, thus contaminating vegetables and the surrounding soil (CDC, 2022). Once the manure is applied to fields, bacteria can make their way to groundwater or wash into nearby bodies of water from rain, flooding, or inadequate containment of animal waste (McEwen and Collington, 2018). The bacteria may ultimately end up in the marine environment. Wastewater treatment plants (WWTP) may dump partially treated effluent directly into ocean waters containing viable bacteria. Raw sewage may also be released into the environment due to flooding, line breaks, and power outages, resulting in polluting water and environmental sources (Environmental Protection Agency, 2021). The illegal discharge of ship ballast water can contribute to bacterial contamination in marine waters with enteric bacteria, including strains of *E. coli* O157:H7 (Altug et al., 2012; Ruiz et al., 2000).

The Salish Sea is an international inland sea that includes Campbell River, British Columbia, to Olympia, Washington, encompassing the Puget Sound and San Juan Islands. The coastline spans about 7,470 km and is home to approximately 37 species of animals, 172 birds, and 250 fish species (SeaDoc, 2021). The Strait of Juan de Fuca is the primary direct access to the Pacific Ocean from the Salish Sea. According to the State of Washington Department of Ecology, the Salish Sea area is home to about 4.5 million people. Population growth is expected to increase by 1.8 million people by 2050. Approximately 99 WWTP line the shores of the Salish Sea. Seventy-nine WWTP are in the United States and twenty on the Canadian side (2019). WWTPs near Vancouver and Vancouver Island, British Columbia, dump partially treated wastewater, into the Salish Sea (Johannessen et al., 2015).

This study aims to fill in some of the knowledge gaps on the diversity of *E. coli* in marine ecosystems and the potential of opportunistic pathogenic *E. coli* present in marine water and mammals in the area to impact human and animal health. The study analyzes data on previously collected isolates of *E coli* from marine water (primarily around commercial shellfish operations) and wildlife samples, while comparing the diversity of STs across four quadrants of the Salish Sea (north, south, central, and Strait of Juan de Fuca) (Vingino et al., 2021; Norman et al., 2021). We simultaneously compared the STs from our dataset to metadata about associated hosts for each ST appearing in EnteroBase (https://enterobase.warwick.ac.uk/) and the published scientific literature in PubMed.

METHODS

Collection

The *E. coli* isolates in this study were collected as part of previously published studies (Norman et al., 2021; Vingino et al., 2021). In those studies, five hundred fifty-one samples of *E. coli* were obtained from the Salish Sea ecosystem. From these, 332 *E. coli* isolates were further studied with whole genome sequencing. These included marine water samples near shellfish beds (n=238), marine water samples near local Seattle beaches (n=3), freshwater samples of estuaries entering into the Salish Sea (n=5), fecal samples from river otters (*Lontra canadensis*) (n=24), rectal swabs, and fecal samples from both dead and alive seals (*Phoca vituline*) (n=53), fecal sampling from dead porpoise (*Phocoena phocoena*) (n=7), and samples isolated from English sole (*Parophrys vetulus*) (n=2) (**Table S1**).

Specific sampling details such as parties responsible for sampling, collection methods, transportation, and laboratory analysis were described previously by Norman et al. (2021) and Vingino et al. (2021). Differences in sampling numbers are due to the previous study (Vingino et al., 2021) assessing for antibiotic resistance in the *E. coli* isolates. While 332 isolates were sent for whole-genome sequencing (WGS) and MLST, not all were able to be sent for antibiotic resistance testing. The previous study by Vingino et al. (2021) used 305 isolates, while the current study uses all 332 whole genome sequences **(Table S1).** The previous study by Vingino et al. (2021) described how the whole-genome sequencing (WGS), as well as multi-locus sequencing typing (MLST), was done in cooperation with the Washington Department of Health in conjunction with the FDA GenomeTrakr Project (ID 283914 - BioProject - NCBI) (2021). NCBI accession numbers for the isolates in that study have previously been described (Vingino et al., 2021) (**Table S2**). Additional NCBI accession numbers used for this study are as follows:

[SAMN] 13392847, 14137889, 14214489, 15182309, 15182314, 15182317, 15344669, 15483653. [SRR] 12424353, 12643355, 12618578, 12618581, 12643343, 12643347, 12643348, 12643351, 12643352, 12643360, 12643362, 12643364, 12643366-12643368, 12643370, 12663941.

EnteroBase

EnteroBase version 1.1.3 is an online public and voluntary repository of scientific data regarding *E coli* (Zhou et al., 2020). The database was accessed on 8/11/2021 (https://enterobase.warwick.ac.uk/) to retrieve and download data for investigating the 196 STs identified in this study. A total of 153,428 entries existed in EnteroBase at the time of the query. Of the 153,428 entries, 47,319 were for the 196 STs found in this study. Entries for All 196 STs were found in the database. Information about the source (human, animal, environmental) for isolates was categorized based on available information found in EnteroBase using the source of the isolates: 1) found in human, animal, and environment, 2) found in animals, 3) found in humans, 4) found in the environment, 5) in animals and environment, 6) in humans and environment, and 7) in humans and animals. STs with ten or fewer entries in EnteroBase were coded as “novel” **(Table S2).** Attention was made to any reference in the data about an ST regarding a documented marine water source, marine animals, shellfish, and fish (Zhou et al., 2020)

Many *E. coli* STs are commensal and do not have pathogenic potential to cause disease (CDC, 2014). While other *E. coli* STs are associated with one or more pathogenic types, which can cause varying degrees of illness in humans and animals (Jolley et al., 2018). The pathogenic types examined included: 1) Shiga toxin-producing *E. coli* (STEC), 2) verocytoxic *E.*

coli (VTEC), and 3) enterohemorrhagic *E. coli* (EHEC), which can produce a toxin that causes various diseases with associated mortality (CDC, 2014). *E. coli* (EIEC) causes intestinal inflammation and bloody diarrhea similar to STEC (Belotserkovsky and Sansonetti, 2018). Enterotoxigenic *E. coli (ETEC)* causes diarrheal illness and is the leading cause of traveler diarrhea and illness in low-income countries, especially among children (CDC, 2014). Enteroaggregative *E. coli* (EAEC) is an emerging foodborne pathogen that may cause acute or chronic diarrhea (Kaur et al., 2010), while Enteropathogenic *E. coli* (EPEC) is a common cause of diarrheal illness in children in developing countries (Ochoa and Contreras, 2011). Extraintestinal Pathogenic *E. coli* (ExPEC) can cause *E. coli* infections outside the intestinal tract, including urinary tract and bloodstream infections. It is also the second most common cause of meningitis in neonates (Russo and Johnson, 2003). Uropathogenic *E. coli* (UPEC) is a subset of extraintestinal pathogenic *E. coli*, causing only urinary tract infections (Spurbeck and Mobley, 2013). EnteroBase was queried for mention of those pathogenic types listed above. We additionally coded ST types as whether they had been associated with STEC, EPEC, EAEC, UPEC, ExPEC, EIEC, ETEC, and EHEC strains. The number of pathogenic types for each ST was identified and added to create a "pathogen score" for a particular ST **(Table S3).**

Spatial Distribution

ST mapping across different areas of the Salish Sea was performed using QGIS version 3.22. Maps were created to demonstrate the field collection sites and anthropogenic *E coli* sources. We also used maps to give a visualization of the distribution of ST diversity. Metadata for WWTP in the Puget Sound area were obtained from the U.S. Geological Survey (USGS Water Mission Area NSDI Node). Metadata for agricultural areas were obtained from the Washington State

Department of Agriculture (Agricultural Land Use | Washington State Department of Agriculture).

PubMed

PubMed (PubMed (nih.gov) was searched between July – September of 2021 to determine the number of citations in the PubMed database for specific STs. The following search terms were utilized: ("Escherichia coli"[Mesh:NoExp] OR "E coli"[TIAB] AND ("ST (number)" OR "Sequence type (number) " OR "ST (number)"[TIAB]). The Result was categorized as follows: 0 articles found; 1-10 articles found; 11-25 articles found;26-50 articles found; 51-100 articles found; 101-200 articles found; 201-500 articles found; 501-1000 articles found, and >1000 articles found.

Statistical Analysis

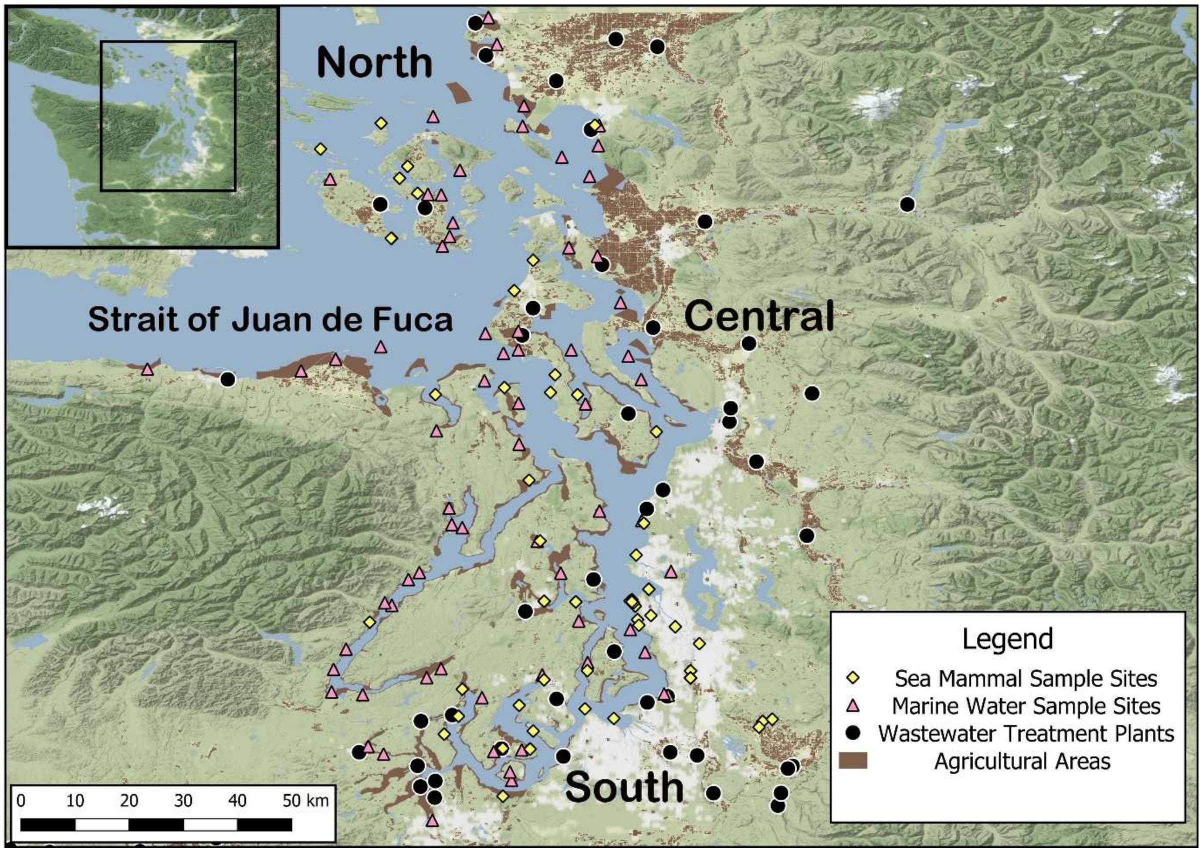
Statistical analysis of *E. coli* diversity was performed using Shannon diversity Index, Shannon equitable index (evenness), and Hutcheson t-test to characterize diversity of *E coli* in the 4 quadrants of the Puget Sound. These statistical measures were used to test the null hypotheses that 1) diversity measures of STs are similar across the four quadrants of the Puget Sound, 2) there are no geographic differences in diversity measures of isolates with source metadata derived from EnteroBase, and 3) there are no geographic differences in diversity of uncommon isolates with ten entries or less in EnteroBase. Source metadata from EnteroBase were used to characterize isolates as only those found in the environment and those in animal and the environment. Utilizing these same tests, diversity among the quadrants for STs with pathogenic potential was also assessed. Data analysis was performed using R studio version 3.6.2.

RESULTS

Sampling Source and STs

Amongst the 332 *E. coli* isolates, 196 STs were identified. In total, 238 isolates were obtained from marine water (72%) (**Figure S1**). A total of 86 isolates with 56 different STs were from animals; seal (n=17), deceased seal (n=36), porpoise (n=7), sole (n=2), and river otter (n=24)**,** while 33 different STs were found in animals only. Five isolates were found in freshwater samples (ST297, ST405, ST616, ST681, ST2557), and three were sampled from marine water by beaches (ST58, ST117, ST1308). Only two STs (ST10 and ST720) were found in all four quadrants of the marine water in the Salish Sea (**Table S3)**. The most common isolate found overall was ST10, with 12 isolates (3.6% of the total), followed by ST162, with eight isolates (2.4%). Two STs had seven isolates each (ST297, ST372). Three STs had six isolates found (ST117, ST967, and ST 1079). Most STs identified were represented by a single isolate (n=139). The remaining (n=43) *E. coli* isolates, had a total of 2-5 isolates for each ST (**Table S3)**

Figure 1. All sample locations in Salish Sea for water and marine mammals and anthropogenic sources. [a]



[a] Multiple ST may be represented by a marker

Diversity Measures and Statistical Analysis

A high level of diversity was found amongst all marine water samples in each of the four quadrants of the Salish Sea. The Strait of Juan de Fuca had the greatest richness of diversity among ST found there (n=51) with a Shannon diversity index (H=3.852), followed closely by the south quadrant (H=3.827). The north quadrant contained the least diversity (H=3.718). Using the Hutcheson t-test to compare the Shannon diversity index values of the Strait of Juan de Fuca and the north quadrant, there was no statistical difference in the diversity of the two quadrants (p=0.13). Of those STs found in animals and the environment, the central quadrant contained the greatest diversity (H=2.398), while the south quadrant was least diverse (H=0.95). The central quadrant had statistically more significant diversity than the southern quadrant (p=0.0018). Of the isolates only found in the environment, the south quadrant had the highest level of diversity (H=1.7), while the north and Strait of Juan de Fuca had the least (H=1.38). There was statistically significant higher diversity in the southern quadrant than in the northern quadrant and Strait of Juan de Fuca (p=0.02). Of the isolates coded as novel (n=62), the Strait of Juan de Fuca contained the highest amount of diversity (H=2.75), while the south quadrant had the least diversity (H=2.39). The Strait of Juan de Fuca had a statistically significant higher degree of novel diversity than the southern quadrant (p=0.009) (**Table 1**). Of those STs previously described as found in humans and the environment (n=8), humans only (n=2), humans and animals (n=4), and animal only (n=2), sample size did not allow for statistical analysis.

Table 1. Diversity measures for quadrants in Salish Sea

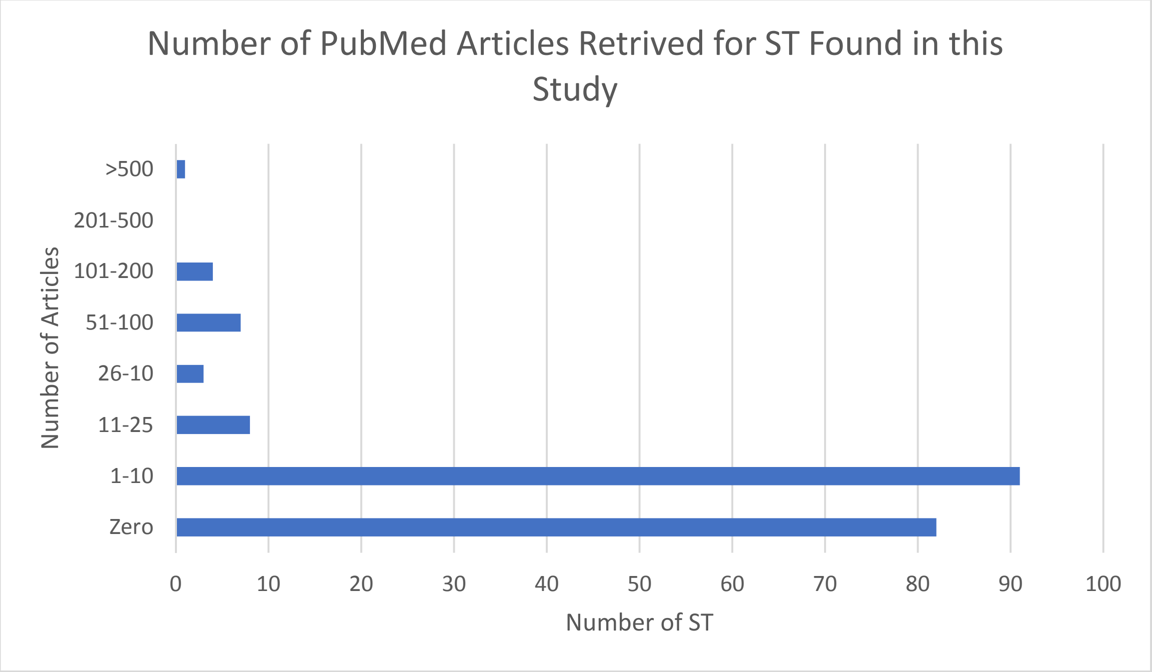
All Marine samples	**# Of unique ST (richness)**	**Shannon Diversity Index**	**Evenness**	**Hutcheson t-test[1]**
North Quadrant	45	3.718	0.977	
South Quadrant	49	3.827	0.983	
Central Quadrant	48	3.798	0.981	
Strait of Juan de Fuca	51	3.852	0.975	
				P=0.1287
Environmental and Animal				
North Quadrant	5	1.609	1	
South Quadrant	3	0.95	0.865	
Central Quadrant	11	2.398	1	
Strait of Juan de Fuca	3	1.09	1	
				P=0.0018
Environmental				
North Quadrant	4	1.386	1	
South Quadrant	6	1.7	0.816	
Central Quadrant	5	1.6	1	
Strait of Juan de Fuca	4	1.386	1	
				P=0.02
Novel				
North Quadrant	13	2.565	1	
South Quadrant	12	2.39	0.96	
Central Quadrant	14	2.615	0.991	
Strait of Juan de Fuca	16	2.75	0.993	
				P=0.009

[1] Hutcheson t-test was performed comparing the quadrants with the highest and lowest Shannon diversity index

PubMed

Of the 196 STs found in this study, 82 (41.8%) STs had no published articles retrieved on a PubMed search, suggesting that these ST types are rarely identified and/or reported. An additional 91 STs had only 1-10 articles available (46%). Only one isolate, ST131, had greater than 1,000 articles retrieved (**Figure 2**).

Figure 2. Number of articles retrieved for ST found in this study in PubMed



EnteroBase

In contrast, in EnteroBase, only 47 (24%) had ≤ five entries, and 15 (7.7%) had ≤ ten entries in EnteroBase. Compared to source data from EnteroBase, 129 (65.8%) STs from this study were found in humans, animals, and the environment. Of the remaining STs, 21(10.7%) had known sources in the environment, 30 (15.3%) were known environmental and animal sources, eight had known sources in humans and environment, four were found in humans and animals, two were known to be found in humans only, and two were known animal sources. (**Table S3**). All quadrants contained STs that were associated with animal sources. In the northern quadrant, 84% of all STs were associated with animals (n=38), in comparison to 85% in the southern quadrant (n=41), 88% in the central quadrant (n=43), and 78% in the Strait of Juan de Fuca (n=40). STs associated with human sources were also isolated in the four quadrants, 76% (n=34) in the north, in comparison to 70% (n=34) in the central, 78% (n=38) in the south, and 86% (n=44) in the Strait of Juan de Fuca. For those isolates found in humans and the environment (n=8), five isolates were found in the Strait of Juan de Fuca. The two human-only isolates were obtained in dead seals. The two animal-only isolates were found in the Strait of Juan de Fuca and the southern quadrant. The four isolates found in humans and animals were found in two deceased seals, the central, and the southern quadrants.

Five STs stood out when compared to known sources in EnteroBase (ST5869, ST1065, ST2164, ST9001, and ST11343) which revealed no previous evidence of occurrence in marine water and sea mammals. Upon further review, our study STs included four isolates found for which EnteroBase listed the source as marine water; 22 isolates were found in marine fish, 11 in marine mammals, and 40 in shellfish (**Table 2).**

Table 2. ST found in this study that had a marine source in EnteroBase

Marine Mammal [a]	Shellfish [b]	Fish (specific saltwater) [c]	Fish (non-specific)	Salt/Marine water
ST: 10, 38, 69, 75, 127, 131, 132, 155, 162, 491, 4219	ST: 10, 38, 48, 58, 69, 73, 101, 117, 127, 155, 162, 206, 215, 224, 297, 327, 345, 349, 372, 446, 540, 547, 641, 666, 681, 906, 942, 1049, 1056, 1079, 1304, 1423, 1611, 2144, 2521, 3601, 4038, 4162, 4481, 6096, 6188	ST: 132, 212, 327, 1704, 1706,	ST: 10, 12, 38, 48, 46, 58, 101, 109, 127,131, 162, 224, 362, 746, 1629, 2522, 6998	ST: 38, 40, 46, 130

[a] Seal, dolphins, sealions
[b] Oysters, mussels, bi-valves, shrimp, crab, mollusk
[c] Sea bream, tuna, sardine, salmon

Information on Pathogenic Types according to EnteroBase

The 196 STs found in this study were queried in EnteroBase to determine if there were known STEC/VTEC, EPEC, EAEC, ExPEC, EIEC, ETEC, EHEC, or UPEC isolates associated with these STs. Thirty-five of the STs found in our study were UPEC strains. ST10 was the most abundant ST found in this study (n=12) and was associated with STEC/VTEC, EPEC, EAEC, ExPEC, and ETEC strains. Six STs found in this study (ST10, ST58, ST101, ST155, ST297, ST1248) found in this study were associated with STEC strains, ten EPEC strains, three EIEC strains, 19 EAEC, 28 ExPEC strains, seven ETEC, and three EHEC (**Table S4**). In those STs with at least one pathogen type, the south quadrant contained the highest diversity (H=2.955), while the central quadrant contained the least diversity (H=2.58). All quadrants sampled for marine water, freshwater, seals (alive and deceased), porpoise, and sole had at least one pathogen associated ST isolated. To further identify potential severe disease-causing strains of *E. coli* in food sources, EnteroBase was queried for pathogenic type ST found in this study with sources found in marine mammals and shellfish, with known disease-causing strains. There were 21 STs found in shellfish and 8 STs found in marine animals with pathogenic potential (**Table S4**)

DISCUSSION

This study describes a high degree of *E. coli* ST diversity in a marine environment as evidenced by Shannon diversity index scores of greater than three in the marine water samples from the four quadrants of the Salish Sea. Five novel STs may be among the reported in marine and animal environments in published literature. Within the quadrants, there was a high degree of diversity in *E. coli* STs that were novel isolates, those with known environmental sources, and those with known environmental and animal sources. Evenness across the four quadrants was consistently close to one. This finding is not surprising given the high proportion of identified STs represented by one isolate; thus, there is a similar abundance, suggesting there is no dominant strain present in the quadrants at the sampling time. Previous studies have described *E. coli* diversity utilizing the Shannon diversity index in freshwater environments. A high degree of *E. coli* diversity was found in freshwater wells in Nigeria, which the authors speculated was due to multiple sources of contamination, including proximity to septic tanks and erosion (Odetoyin et al., 2022). Chandran and Mazumder described a high level of *E. coli* diversity, with seasonal variations, within a freshwater lake on Vancouver Island, British Columbia, Canada (2015). The authors had previously described greater *E. coli* diversity in animal and avian hosts than in human hosts (2014, 2013). They hypothesized that the high level of diversity and seasonality variation of *E. coli* in the 2015 study was due to the variations in animal populations in the area as the lake was in a forested watershed. This lake is located near the Strait of Georgia, which connects to the north quadrant of Puget Sound. Given the Puget Sound is home to 37 species of animals and 172 birds that depend on the marine ecosystem (SeaDoc, 2021), this may account for the overall high diversity and differences in diversity in the quadrants when comparing isolates with known sources of animal and environment and environmental only isolates.

The Strait of Juan de Fuca had the greatest ST richness, and the highest total number of isolates found. The total number of isolates sampled in each quadrant is similar and variation from quadrant to quadrant may have been dependent on sampling methods. Five of the eight isolates found in humans and the environment were also sampled in this region. While the Strait of Juan de Fuca had the highest Shannon diversity index score, followed very closely by the southern quadrant, there was no statistical significance in the amount of diversity between the Strait and the quadrant with the lowest score (north). This suggests that the Strait and south quadrant may be slightly more diverse than the other quadrants but not at a statistically significant level. Interestingly, the Strait of Juan de Fuca also had the most diversity of novel isolates as compared to the other quadrants. The significance and explanation for this are unclear, though it is important to note the other quadrants had high diversity with index scores greater than two when assessing the biodiversity for the novel isolates.

The central quadrant had the most diversity in STs found in animals and the environment, which was statistically significant compared to diversity in the other quadrants. This quadrant also had the second-highest diversity index score for environmental isolates. This may be due to a large number of agricultural areas and animals present in the region. In those STs with known environmental sources, the highest diversity was in the southern quadrant, followed closely by the central quadrant. It is unclear why the southern quadrant had the most diversity, as the central quadrant appears to have more agricultural areas (**Figure 1**). The diversity of STs may be influenced by the density of animal populations and presence of WWTP in the southern quadrant similar to the findings of Chandran and Mazumder (2015).

The northern quadrant had the fewest total number of isolates found and did not have any statistical significance in any of the categories investigated. This is somewhat surprising due to the presence of multiple untreated sewage outfalls in Vancouver and Victoria, British Columbia (Krogh et al., 2017). We hypothesized due to the presence of untreated sewage, there would be a very high level of diversity and more human isolates in this quadrant. One possible explanation is that tidal currents and winds can carry the sewage into the Strait of Juan de Fuca, which may account for the variation found in this region (Krogh et al., 2018).

In additional to the high diversity of *E. coli* ST in this study, there were also interesting findings in regard to novel STs. Many STs in the current study have minimal available literature in PubMed, while five STs had no literature describing an occurrence in marine water and sea mammals. For example, in this study, *E. coli* ST5869, found in marine water, had entries in EnteroBase listing human and animal sources, including animals from China, Belgium (Brussels), and the United States (Tennessee), human sources from Cambodia and the Netherlands, and an avian source in Kenya. Scholarly articles have previously described this ST as being found in food sources in Germany and Belgium (Pauly et al., 2020; Garcia-Graells et al., 2015), animals treated in a veterinary clinic in Switzerland (Schmitt et al., 2021), and humans in Nigeria (Jesumirhewe et al., 2020). We could not find a previous report of this ST in a marine environment.

In the current study, *E. coli* ST1065 and ST11343 were associated only with human sources in EnteroBase, but in the current study were found in dead seals. In EnteroBase, ST1065 was entered as a human source from France and United States (Washington). In the expanded literature review, one article linked ST1065 to cattle (Isiko et al., 2015). We did not find any previous case reports of this ST in a marine animal. No scholarly articles were found regarding

ST11343, even in an expanded search from multiple databases. While this ST was isolated in human sources in EnteroBase, this may be the first reported case in animals published.

ST9001, isolated from a live seal in the current study, was referenced in human and environment in EnteroBase with sources from North America (not further specified) from humans and water. ST2164 isolated from a river otter in the current study was referenced in the environment only in EnteroBase with sources from Canada and Massachusetts and water sources from Florida and Washington. Neither ST9001 nor ST2164 were found in an expanded literature search in multiple databases. We did not find previous reports of these STs in animals.

The significance of the high degree of *E. coli* diversity in this water system is unclear. Enteric bacteria have been generally considered to have poor survivability in marine waters due to cold temperature, UV light, and salinity. Until recently the presence of *E. coli* in marine sources was once thought to only be an indicator of recent fecal contamination. However, there is growing evidence that *E. coli* can survive outside of mammalian hosts and in the environment (Jang et al., 2017). Biofilm production may increase survival in marine sediment (Lee et al., 2006), while algae may act as a reservoir for *E. coli* (Quero et al., 2015). There is some evidence that *E. coli* can enter a viable but non-culturable state, and those isolates carrying pathogenic genes will be able to retain their ability to cause infections (Pommepuy et al., 1996). While the current study does not measure fitness but rather detection at the time of sampling (although detection required the ability to be cultured), the potential health risks posed to humans and marine mammals by the presence of *E. coli* in the Puget Sound remain poorly understood.

Marine and Human Health

There have been several previous studies that describe cultured pathogens from marine animals and marine food sources. It is difficult to ascertain the significance of *E. coli* cultured from fish in food stalls and markets as they may have been cross contaminated during processing and preparing. However, Brahmi et al. isolated antibiotic-resistant *E. coli* from the gut of wild-caught fish in the Mediterranean Sea using sterile techniques (2015). Marine Mammals are known to be reservoirs for *E. coli*. In a 1991 study by Buck et al., *E. coli* was isolated from stranded cetaceans from the heart, lung, oral cavity (including trachea), and skin lesions. These animals ranged from the Northeast Coast of the United States to Florida. *E. coli*, which rarely causes endocarditis in humans, was found to be the cause of infective endocarditis in a sea lion (*Zalophus californianus*) five days after the animal came under the care of veterinarians for a fin injury (Kim et al., 2001). Previous studies have described the presence of pathogenic *E. coli* in seal pups from pristine environments in Antarctica (Hernandez et al., 2007). Other studies have described the presence of *E. coli* in deceased seal pups though the significance of disease burden is unknown. Some seals did show signs of septicemia and pneumonia (Spraker et al., 2020). *E. coli* has been isolated in the feces of Orca whales (*Orcinus orca*) with an unknown significance on the animal's health (Melendez et al., 2019). While increasing anthropogenic activities are causing an increase in the number of enteric bacteria and possibly promoting their survivability, the risk to the health of marine animals remains uncertain. However, *E. coli* found in marine animals may also serve as sentinels to human health.

E. coli's presence in marine waters may impact human health. The marine water samples positive for *E. coli* were obtained near shellfish beds for routine monitoring. Most shellfish feed on organic materials by drawing in ocean water with cilia on their gills, and they can filter two to three gallons of water within a one-hour time frame. (Salvini-Plawen, 2018). As the shellfish filter, bacteria and viruses can accumulate in their tissues as they feed. Often these pathogens do not affect the taste or smell of the shellfish. The risk of illness is most significant with raw shellfish is consumption, particularly oysters (CDC, 2022). Previous studies have described the presence of *E. coli* in oysters at a concentration that would not be safe for fresh consumption (Forcelini et al., 2009). In contrast, other studies described high *E. coli* concentrations in bivalve tissues, with the surrounding marine water having smaller concentrations or even undetectable (Bobanović-Ćolić et al., 2021). Potentially pathogenic serotypes of *E. coli* have also been isolated from shellfish (Miotto et al., 2019), including STEC and EPEC strains (Baliere et al., 2016). While widespread outbreaks of enteritis caused by *E. coli* contaminated shellfish consumption have not been reported, there are cases of human illness from STEC in shellfish and crab consumption (Matulkova et al., 2013; CDC, 2019). *E. coli* has also been isolated in other marine food sources such as fish, lobster, and shrimp (Fadaeifard et al., 2014). Even though there have been no widespread outbreaks of illness due to *E. coli* in seafood, health risks associated with increased survivability and presence in food sources should not be minimized. Most food-borne illnesses go unreported due to mild-moderate and self-limited symptoms (CDC,2020). Even in patients who seek medical attention, accurate diagnosis may not be made as most food-borne illnesses are not detected on routine laboratory testing (CDC, 2020). Therefore, it is not unreasonable to suggest that the incidence of enteritis caused by *E. coli* from a marine food source is higher than reported in the available literature.

Study Strengths and Limitations

There were several strengths of this study. Samples of marine water were taken in a coordinated manner by the Washington State Department of Health as part of their shellfish monitoring program (Vingino et al., 2021). To our knowledge, no samples were collected after significant rainfall or flooding events which may have influenced the amount of *E. coli* in the area. Additionally, because the sampling of deceased animals was opportunistic and there was no human contact with the otters and alive seals while collecting fecal samples (Vingino et al., 2021), there was no harm or stress caused to marine mammals during the collection of samples.

This study had several limitations. EnteroBase is a public database not centrally curated and, as such, relies on the users to enter accurate information. We noted instances when the source niche and the source type did not match (e.g., the source niche was listed as human, but the source type entered listed as ovine/goat). Additionally, the source was not entered for every isolate. While every attempt was made to identify and mitigate these inconsistencies, some source information for isolates may be inaccurate. It is also possible that isolates in this study have been recovered from other sources but were not listed. Only a few isolates were reported by pathogen type (e.g., STEC, ExPEC, and EHEC). Thus, we may have under-identified pathogenic potential in our isolates. In addition, STs, especially those seeming to be novel, may be underreported as the database relies on human entry and is not auto populated by lab detection software.

The literature review and counting of articles associated with each ST were performed with PubMed only. Additional articles may be found on different scholarly publication databases. This study aimed not to search multiple databases for each ST, but to give a broad overview of available literature in an extensive, primarily peer-reviewed database associated with each. Of note, an expanded literature review utilizing multiple databases was performed for ST5869, ST1065, ST11343, ST9001, and ST2164 as described above to identify whether our reporting was novel. Finally, because this study utilized a previously collected dataset, the assumption is made that the data contained in the dataset is accurate and without errors. Samples for dead seals and porpoises collected for the original dataset were opportunistic and taken postmortem.

Conclusion

There is growing evidence that *E. coli* has greater survivability in marine environments than previously thought. To our knowledge, this study is the most extensive characterization of *E. coli* diversity in a marine environment to date. Pathogenic strains of *E. coli* have previously been described in shellfish and marine animals with unknown health impacts. Because the impact of the diversity of *E. coli* in this study is still relatively unknown, future studies should continue to assess public health threat potential.

.

REFERENCES

Altug, Gurun, S., Cardak, M., Ciftci, P. S., & Kalkan, S. (2012). The occurrence of pathogenic bacteria in some ships' ballast water incoming from various marine regions to the Sea of Marmara, Turkey. *Marine Environmental Research*, *81*, 35–42. https://doi.org/10.1016/j.marenvres.2012.08.005

Balière, Rincé, A., Delannoy, S., Fach, P., & Gourmelon, M. (2016). Molecular Profiling of Shiga Toxin-Producing Escherichia coli and Enteropathogenic E. coli Strains Isolated from French Coastal Environments. *Applied and Environmental Microbiology*, *82*(13), 3913–3927. https://doi.org/10.1128/AEM.00271-16

Belotserkovsky, & Sansonetti, P. J. (2018). Shigella and Enteroinvasive Escherichia Coli. *Current Topics in Microbiology and Immunology*, *416*, 1–26. https://doi.org/10.1007/82_2018_104

Bobanović-Ćolić, Bolotin, J., Glavić, N., & Hrustić, E. (2021). Microbiological Quality of Seawater and Mussels (Mytilus galloprovincialis, Lamarck 1819) in the Dubrovnik Coastal Area (Southeastern Adriatic). *NAŠE MORE*, *68*(3), 137–149. https://doi.org/10.17818/NM/2021/3.1

Brahmi, S., et al. (2015). CTX-M-15-producing *Escherichia coli* and the pandemic clone O25b-ST131 isolated from wild fish in Mediterranean Sea. *Clinical Microbiology and Infection*, 21(3), e18–e20.

Buck, J.D., et al. (1991). Bacteria associated with stranded cetaceans from the northeast USA and southwest Florida gulf coasts. *Diseases of Aquatic Organisms*, 10(2), 147–152

Centers for Disease Control and Prevention. (2022). Antibiotic/Antimicrobial Resistance (AR/AMR). Where Resistance Spreads: Water, Soil, Environment | CDC [Accessed April 2, 2022].

Centers for Disease Control and Prevention. (2014). E. coli. Questions and Answers. Questions and Answers | E. coli | CDC [Accessed March 26, 2022].

Centers for Disease Control and Prevention. (2020). Food Safety. Information for Healthcare Professionals <B>Error processing SSI file</B>
 (cdc.gov) [Accessed April 29, 20220]

Centers for Disease Control and Prevention. (2022). Foods That Can Cause Food Poisoning. Foods That Can Cause Food Poisoning | CDC [Accessed April 28, 2022].

Centers for Disease Control and Prevention. (2019). Multistate Outbreak of gastrointestinal illnesses linked to oysters imported from Mexico. Multistate Outbreak of Gastrointestinal Illnesses Linked to Oysters Imported from Mexico | Multistate Outbreak of Gastrointestinal Illnesses Linked to Oysters Imported from Mexico | Vibrio Illness (Vibriosis) | CDC [Accessed April 30, 2022].

Chandran, A & Mazumder, A. (2014). Occurrence of Diarrheagenic Virulence Genes and Genetic Diversity in Escherichia coli Isolates from Fecal Material of Various Avian Hosts in British Columbia, Canada. *Applied and Environmental Microbiology*, *80*(6), 1933–1940. https://doi.org/10.1128/AEM.03949-13

Chandran, & Mazumder, A. (2015). Pathogenic potential, genetic diversity, and population structure of Escherichia coli strains isolated from a forest-dominated watershed (Comox Lake) in British Columbia, Canada. *Applied and Environmental Microbiology*, *81*(5), 1788–1798. https://doi.org/10.1128/AEM.03738-14

Chandran, A & Mazumdar, A. (2013). Prevalence of Diarrhea-Associated Virulence Genes and Genetic Diversity in Escherichia coli Isolates from Fecal Material of Various Animal Hosts. *Applied and Environmental Microbiology*, *79*(23), 7371–7380. https://doi.org/10.1128/AEM.02653-13

Environmental Protection Agency. (2021). LEARN: What Affects Beach Health. LEARN: What Affects Beach Health | US EPA [Accessed April 2, 2022]

Fadaeifard, Rahimi, E., & Tajbakhsh, E. (2014). Occurrence of Escherichia coli O157 in Some Marine Fishes and Shellfish Sold in Isfahan markets, Iran. *Biosciences, Biotechnology Research Asia*, *11*(3), 1353–1358. https://doi.org/10.13005/bbra/1526

Forcelini, H., Kolm, H., & Absher, T. (2009). Total Coliforms and Escherichia coli in Oysters Marketed at the Municipal Market of Guaratuba, Parana - Brazil. *Boletim Do Instituto De Pesca, 35*(2), 275-283.

Garcia-Graells, Berbers, B., Verhaegen, B., Vanneste, K., Marchal, K., Roosens, N. H.., Botteldoorn, N., & De Keersmaecker, S. C. (2020). First detection of a plasmid located carbapenem resistant blaVIM-1 gene in E. coli isolated from meat products at retail in Belgium in 2015. *International Journal of Food Microbiology*, *324*, 108624–108624. https://doi.org/10.1016/j.ijfoodmicro.2020.108624

Hernandez, J., Prado, V., Torres, D., Waldenström, J., Haemig, P., & Olsen, B. (2007). Enteropathogenic Escherichia coli (EPEC) in Antarctic fur seals Arctocephalus gazella. *Polar Biology*, *30*(10), 1227–1229. https://doi.org/10.1007/s00300-007-0282-2

Ibarz Pavón, & Maiden, M. C. J. (2009). Multilocus Sequence Typing. *Methods in Molecular Biology (Clifton, N.J.)*, *551*, 129–140. https://doi.org/10.1007/978-1-60327-999-4_11

Isiko, Khaitsa, M., & Bergholz, T. M. (2015). Novel sequence types of non□O157 Shiga toxin□ producing Escherichia coli isolated from cattle. *Letters in Applied Microbiology*, *60*(6), 552–557. https://doi.org/10.1111/lam.12404

Jang, Hur, H. □G., Sadowsky, M. J., Byappanahalli, M. N., Yan, T., & Ishii, S. (2017). Environmental Escherichia coli: ecology and public health implications—a review. *Journal of Applied Microbiology*, *123*(3), 570–581. https://doi.org/10.1111/jam.13468.

Jesumirhewe, Springer, B., Allerberger, F., & Ruppitsch, W. (2020). Whole genome sequencing of extended-spectrum β-lactamase genes in Enterobacteriaceae isolates from Nigeria. *PloS One*, *15*(4), e0231146–e0231146. https://doi.org/10.1371/journal.pone.0231146

Johannessen, Macdonald, R. W., Burd, B., van Roodselaar, A., & Bertold, S. (2015). Local environmental conditions determine the footprint of municipal effluent in coastal waters: a case study in the Strait of Georgia, British Columbia. *The Science of the Total Environment*, *508*, 228–239. https://doi.org/10.1016/j.scitotenv.2014.11.096

Jolley, K.A., Bray, J.E. & Maiden, M.C.J. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Research*, 3, 124

Kaur, P., Chakraborti, A., & Asea, A. (2010). Enteroaggregative Escherichia coli: An Emerging Enteric Food Borne Pathogen. *Interdisciplinary perspectives on infectious diseases*, *2010*, 254159. https://doi.org/10.1155/2010/254159

Kim, J.H., et al. (2002). Endocarditis associated with *Escherichia coli* in a sea lion (*Zalophus californianus*). *Journal of Veterinary Diagnostic Investigation*, 14(3), 260–262.

Krogh, Ianson, D., Hamme, R. C., & Lowe, C. J. (2018). Risks of hypoxia and acidification in the high energy coastal environment near Victoria, Canada's untreated municipal sewage outfalls. *Marine Pollution Bulletin*, *133*, 517–531. https://doi.org/10.1016/j.marpolbul.2018.05.018

Krogh, J., Lyons, S. & Lowe, C.J. (2017). Pharmaceuticals and Personal Care Products in Municipal Wastewater and the Marine Receiving Environment Near Victoria Canada. *Frontiers in Marine Science*, 4, 2-18.

Lee, Lin, T. Y., Lin, C.-C., Kohbodi, G. A., Bhatt, A., Lee, R., & Jay, J. A. (2006). Persistence of fecal indicator bacteria in Santa Monica Bay beach sediments. *Water Research (Oxford)*, *40*(14), 2593–2602. https://doi.org/10.1016/j.watres.2006.04.032

Melendez, Roberts, M. C., Greninger, A. L., Weissman, S., No, D., Rabinowitz, P., & Wasser, S. (2019). Whole-genome analysis of extraintestinal pathogenic Escherichia coli (ExPEC) MDR ST73 and ST127 isolated from endangered southern resident killer whales (Orcinus orca). *Journal of Antimicrobial Chemotherapy*, *74*(8), 2176–2180. https://doi.org/10.1093/jac/dkz159

Miotto, Ossai, S. A., Meredith, J. E., Barretta, C., Kist, A., Prudencio, E. S., R. W. Vieira, C., & Parveen, S. (2019). Genotypic and phenotypic characterization of Escherichia coli isolated from mollusks in Brazil and the United States. *Microbiology Open (Weinheim)*, *8*(5), e00738–n/a. https://doi.org/10.1002/mbo3.738

McEwen, S., & Collignon, P. (2018). Antimicrobial Resistance: A One Health Perspective. *Microbiology Spectrum, 6*(2), Microbiology spectrum, 2018-03, Vol.6 (2).

Matulkova, Gobin, M., Taylor, J., Oshin, F., O'Connor, K., & Oliver, I. (2013). Crab meat: a novel vehicle for E. coli O157 identified in an outbreak in South West England, August 2011. *Epidemiology and Infection*, *141*(10), 2043–2050. https://doi.org/10.1017/S0950268812002816

Norman, S. et al, (2021). Antibiotic Resistance of Bacteria in Two Marine Mammal Species, Harbor Seals and Harbor Porpoises, Living in an Urban Marine Ecosystem, the Salish Sea, Washington State, USA. *Oceans (Basel, Switzerland), 2*(6), 86-104

Ochoa, T. J., & Contreras, C. A. (2011). Enteropathogenic escherichia coli infection in children. *Current opinion in infectious diseases*, *24*(5), 478–483. https://doi.org/10.1097/QCO.0b013e32834a8b8b

Odetoyin, Ogundipe, O., & Onanuga, A. (2022). Prevalence, diversity of diarrhoeagenic Escherichia coli and associated risk factors in well water in Ile-Ife, Southwestern Nigeria. *One Health Outlook*, *4*(1), 3–3. https://doi.org/10.1186/s42522-021-00057-4

Pauly, Hammerl, J. A. J. A., Grobbel, M. M., Käsbohrer, A. A., Tenhagen, B.-A. B.-A., Malorny, B. B., Schwarz, S. S., Meemken, D. D., & Irrgang, A. A. (2021). Identification of a blaVIM-1-Carrying IncA/C2 Multiresistance Plasmid in an Escherichia coli Isolate Recovered from the German Food Chain. *Microorganisms (Basel)*, *9*(1), 29. https://doi.org/10.3390/microorganisms9010029

Pommepuy, Butin, M., Derrien, A., Gourmelon, M., Colwell, R. R., & Cormier, M. (1996). Retention of enteropathogenicity by viable but nonculturable Escherichia coli exposed to seawater and sunlight. *Applied and Environmental Microbiology*, *62*(12), 4621–4626. https://doi.org/10.1128/AEM.62.12.4621-4626.1996

QGIS.org, 2022. QGIS Geographic Information System. QGIS Association. http://www.qgis.org

Quero, Fasolato, L., Vignaroli, C., & Luna, G. M. (2015). Understanding the association of Escherichia coli with diverse macroalgae in the lagoon of Venice. *Scientific Reports*, *5*(1), 10969–10969. https://doi.org/10.1038/srep10969

Ruiz, Rawlings, T. K., Dobbs, F. C., Drake, L. A., Mullady, T., Huq, A., & Colwell, R. R. (2000). Global spread of microorganisms by ships. *Nature (London)*, *408*(6808), 49–50. https://doi.org/10.1038/35040695

Russo, & Johnson, J. R. (2003). Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem. *Microbes and Infection*, *5*(5), 449–456. https://doi.org/10.1016/S1286-4579(03)00049-2

Salvini-Plawen, L. (2018). Mollusk. In *Encyclopedia Britannica*. https://www.britannica.com/animal/mollusk [Accessed April 29, 2022]

Schmitt, Kuster, S. P., Zurfluh, K., Jud, R. S., Sykes, J. E., Stephan, R., & Willi, B. (2021). Transmission Chains of Extended-Spectrum Beta-Lactamase-Producing Enterobacteriaceae at the Companion Animal Veterinary Clinic-Household Interface. *Antibiotics (Basel)*, *10*(2), 171. https://doi.org/10.3390/antibiotics10020171

Spraker, Kuzmina, T. A., & DeLong, R. L. (2020). Causes of mortality in northern elephant seal pups on San Miguel Island, California. *Journal of Veterinary Diagnostic Investigation*, *32*(2), 312–316. https://doi.org/10.1177/1040638720907100

Spurbeck, R.R., Mobley, H.L.T. (2013). Uropathogenic *Escherichia coli*, in: *Escherichia coli*: Pathotypes and Principles of Pathogenesis: Second Edition.

The SeaDoc Society. (2022). About the Salish Sea. About the Salish Sea — SeaDoc Society [accessed 5/17/2022].

Vingino, A., Roberts, M., Wainstein, M., West, J., Norman, S., Lambourn, D., . . . Rabinowitz, P. (2021). Surveillance for Antibiotic-Resistant E. coli in the Salish Sea Ecosystem. *Antibiotics (Basel), 10*(10), 1201

Washington State Department of Ecology. (2019). Puget Sound Nutrient Source Reduction Project. Puget Sound Nutrient Source Reduction Project. Volume 1: Model Updates and Bounding Scenarios (wa.gov) [Retrieved April 20, 2022].

Zhou, Z., Alikhan, N., Mohamed, K., Fan, Y., & Achtman, M. (2020). The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. *Genome Research, 30*(1), 138-152.

APPENDIX

Table S1. Comparison of ST utilized for Vingino et al. study and current study

Source	Number of *E. coli* isolated (Vingino et al, 2021)	Number of *E. coli* isolated this study
North Puget Sound	49	57
Central Puget Sound	55	58
South Puget Sound	56	59
Strait of Juan de Fuca	52	64
Seal (Alive)	17	17
Seal (Dead)	35	36
Porpoise	7	7
Sole	2	2
Otter	24	24
Fresh water	5	5
Marine Water by beaches	3	3
Total	305	332

Table S2. NCBI numbers from Vingino et al. and current study

[SAMN]13337618, 13348248, 13352752,13352855- 13352864, 13392846, 13392848-13392863, 13392951-13392953, 13392847, 13418005, 13429240, 22 13429289, 13482430, 13502693, 13502695, 13502889-13502891, 13513927-13513929, 13513928-13513930, 13513935-13513938, 13513942, 13513948, 13518346, 13518347, 13898866-13898880, 13911824, 13911825, 14057293, 14057294, 14080880-14080885, 14083856-14083863, 14083865, 14083866, 14083868-14083870, 14083873, 14084247, 14113834, 14113836-14113844, 14113847, 14113850, 14113860-14113863, 14137883-14137892, 14137896-14137905, 14137905, 14137978, 14137979, 14138286, 14140185-14140189, 14140195-14140217, 14214489, 14214490-14214498, 14270850- 14270852, 14271025, 14271030-14271033, 14291765, 14316584-14316586, 14316588- 14316590, 14316618, 14316619, 14316621, 14316622, 14316624, 14316625, 14316627, 14316629, 14316633, 14316684-14316687, 14593716-14593722, 14749987, 14749988, 14749995, 14750012, 14750852, 14750854-14750856, 15182299-15182304, 15182308-15182316, 15182319, 15182320, 15182323, 15344667, 15344669, 15344671, 15344672, 15344674, 15483653-15483656, 15777149, 15777151, 15777153-15777155, 15777158, 15777160, 15777162, 15777164, 15777165, 15777167, 16054328-16054337, 16054339- 16054347, 16054538, 16056701-16056705, 16056743-16056748, 16136466, 16136468, 16136469, 16136474-16136479, 16136481, 16136482, 16136485, 16136487, 16136489, 16136490, 16202553-16202558, 16257942-16257946, 16377217-16377219, and 1643928

Table S3. Table of all ST found in current study, number found in EnteroBase, where the isolates were found in current study and source in EnteroBase.

	Number found in EnteroBase	Number of Articles Found in PubMed	Sample Source (quadrant, animal)	EnteroBase Source
12 Isolates found ST types identified n=1				
ST10	9565	101-200	North Quadrant (n=4) Central Quadrant (n=4) River Otter (n=2) South Quadrant (n=1) SJDF (n=1)	Humans, Animals, and Environment
8 Isolates Found ST types identified n= 1				
ST162	750	11-25	Central Quadrant (n=1) Seal, Dead (n=3) Seal, Alive (n=1) Porpoise (n=1) River otter (n=2)	Humans, Animals, and Environment
7 Isolates Found ST types identified n= 2				
ST297	519	1-10	South Quadrant (n=1) SJDF (n=5) Fresh water (n+1)	Humans, Animals, and Environment
ST372	428	11-25	North Quadrant (n=1) Central Quadrant (n=2) SJDF (=1) Seal, Dead (n=3)	Humans, Animals, and Environment
6 Isolates Found ST types identified n= 3				
ST117	1492	51-100	Central Quadrant (n=1) SJDF (n=2) Seal, Dead (n=2) Marine Beach (n=1)	Humans, Animals, and Environment
ST967	31	1-10	SJDF (n=1) South Quadrant (n=1) Seal, Dead (n=2) Seal, Alive (n=2)	Humans, Animals, and Environment

ST1079	223	1-10	North Quadrant (n=3) SJDF (n=1) River Otter (n=2)	Humans, Animals, and Environment
5 Isolates found **ST types identified n= 7**				
ST40	186	11-25	South Quadrant (n=2) SJDF (n=2) Seal, Alive (n=1)	Humans, Animals, and Environment
ST73	2224	51-100	South Quadrant (n=1) Seal, Dead (n=4)	Humans, Animals, and Environment
ST130	109	1-10	South Quadrant (n=2) Central Quadrant (n=2) Seal, Alive (n=1)	Humans, Animals, and Environment
ST919	30	1-10	North Quadrant (n=2) Central Quadrant (n=2) South Quadrant (n=1)	Humans, Animals, and Environment
ST1718	14	0	North Quadrant (n=2) SJDF (n=2) Seal, Dead (n=1)	Humans, Animals, and Environment
ST2144	57	0	Central Quadrant (n=1) SJDF (n=3) River Otter (n=1)	Humans, Animals, and Environment
ST10752	5	0	North Quadrant (n=1) Central Quadrant (n=1) Seal, Alive (n=1) Dead Seal (n=2)	Humans and Environment
4 Isolates found **ST types identified n= 5**				
ST58	1580	51-100	North Quadrant (n=1) Central Quadrant (n=1) Seal, Dead (n=1) Marine by beaches (n=1)	Humans, Animals, and Environment
ST69	2228	101-200	North Quadrant (n=1) SJDF (n=1) River otter (n=2)	Humans, Animals, and Environment
ST196	110	1-10	SJDF (n=2) South Quadrant (n=1) Seal, Alive (n=1)	Humans, Animals, and Environment
ST362	215	1-10	River Otter (n=4)	Humans, Animals, and Environment
ST720	88	1-10	North Quadrant (n=1) Central Quadrant (n=1) SJDF (n=1) South Quadrant (n=1)	Humans, Animals, and Environment

3 Isolates Found **ST types identified n= 10**				
ST101	994	51-100	North Quadrant (n=1) South Quadrant (n=1) Seal, Alive (n=1)	Humans, Animals, and Environment
ST131	9659	>1000	North Quadrant (n=1) Seal, Dead (n=1) River otter (n=1)	Humans, Animals, and Environment
ST428	233	1-10	Central Quadrant (n=2) SJDF (n=1)	Humans, Animals, and Environment
ST589	41	0	Central Quadrant (n=1) SDJF (n=1) Porpoise (n=1)	Humans, Animals, and Environment
ST681	101	1-10	North Quadrant (n=1) South Quadrant (n=1) Fresh water (n=1)	Humans, Animals, and Environment
ST744	469	26-50	North Quadrant (n=1) Seal, Alive (n=2)	Humans, Animals, and Environment
ST937	81	1-10	North Quadrant (n=2) South Quadrant (n=1)	Humans, Animals, and Environment
ST1170	28	1-10	Central Quadrant (n=1) Seal, Dead (n=2)	Humans, Animals, and Environment
ST6096	19	1-10	South Quadrant (n=3)	Animals and Environment
ST6998	8	1-10	South Quadrant (n=3)	Humans, Animals, and Environment
2 isolates found **ST types identified n= 28**				
ST12	871	26-50	Seal, Dead (n=2)	Humans, Animals, and Environment
ST38	2071	101-200	South Quadrant (n=2)	Humans, Animals, and Environment
ST54	14	1-10	Central Quadrant (n=2)	Humans, Animals, and Environment
ST108	32	1-10	Central Quadrant (n=2)	Humans, Animals, and Environment
ST127	773	26-50	Seal, Dead (n=1) Porpoise (n=1)	Humans, Animals, and Environment
ST136	43	1-10	South Quadrant (n=1) Porpoise (n=1)	Humans, Animals, and Environment
ST141	436	11-25	South Quadrant (n=1) Seal, Dead (n=1)	Humans, Animals, and Environment
ST155	1262	51-100	Central Quadrant (n=1) SJDF (n=1)	Humans, Animals, and Environment

ST224	393	11-25	North Quadrant (n=2)	Humans, Animals, and Environment
ST349	369	1-10	North Quadrant (n=1) South Quadrant (n=1)	Humans, Animals, and Environment
ST361	210	11-26	North Quadrant (n=2)	Humans, Animals, and Environment
ST446	131	1-10	North Quadrant (n=1) SJDF (n=1)	Humans, Animals, and Environment
ST906	237	1-10	North Quadrant (n=2)	Humans, Animals, and Environment
ST1148	47	0	South Quadrant (n=1) SJDF (n=1)	Humans, Animals, and Environment
ST1248	28	0	North Quadrant (n=1) South Quadrant (n=1)	Humans, Animals, and Environment
ST1276	23	1-10	North Quadrant (n=1) SJDF (n=1)	Humans, Animals, and Environment
ST1611	130	1-10	North Quadrant (n=2)	Humans, Animals, and Environment
ST1629	18	1-10	SJDF (n=2)	Humans, Animals, and Environment
ST1850	19	1-10	Central Quadrant (n=1) South Quadrant (n=1)	Humans, Animals, and Environment
ST2248	26	0	SJDF (n=2)	Humans, Animals, and Environment
ST2521	85	1-10	Central Quadrant (n=1) River otter (n=1)	Humans, Animals, and Environment
ST2607	11	0	River otter (n=2)	Animal and Environment
ST4162	11	0	Seal, Alive (n=2)	Animal and Environment
ST4219	5	0	SJDF (n=2)	Human and Environment
ST5442	5	1-10	South Quadrant (n=2)	Humans, Animals, and Environment
ST7250	7	0	River otter (n=2)	Animal and Environment
ST9237	4	0	North Quadrant (n=1) Central Quadrant (n=1)	Animal and Environment
ST10817	2	0	South Quadrant (n=1) SJDF (n=1)	Environment only

1 Isolate Found
ST types identified n= 139

ST2	31	51-100	South Quadrant	Humans, Animals, and Environment
ST13	108	1-10	South Quadrant	Humans, Animals, and Environment
ST24	2	1-10	SJDF	Human and Environment
ST26	32	1-10	SJDF	Humans, Animals, and Environment
ST28	147	1-10	South Quadrant	Humans, Animals, and Environment
ST46	260	11-25	Central Quadrant	Humans, Animals, and Environment
ST48	901	51-100	South Quadrant	Humans, Animals, and Environment
ST70	85	1-10	North Quadrant	Humans, Animals, and Environment
ST75	81	1-10	SJDF	Humans, Animals, and Environment
ST80	178	1-10	Central Quadrant	Humans, Animals, and Environment
ST82	1	1-10	SJDF	Humans, Animals, and Environment
ST83	51	1-10	Seal, Dead	Animal only
ST91	39	1-10	Central Quadrant	Humans, Animals, and Environment
ST109	111	0	SJDF	Humans, Animals, and Environment
ST132	15	1-10	Seal, Dead	Human and Animal
ST135	126	1-10	Central	Humans, Animals, and Environment
ST144	149	1-10	Seal, Dead	Humans, Animals, and Environment
ST154	266	0	North Quadrant	Humans, Animals, and Environment
ST160	6	1-10	River otter	Humans, Animals, and Environment
ST174	33	1-10	Seal, Dead	Humans, Animals, and Environment
ST206	378	1-10	SJDF	Humans, Animals, and Environment

ST212	139	1-10	Central Quadrant	Humans, Animals, and Environment
ST215	35	1-10	SJDF	Humans, Animals, and Environment
ST227	71	0	Central Quadrant	Humans, Animals, and Environment
ST327	112	1-10	Seal, Dead	Humans, Animals, and Environment
ST345	211	1-10	Seal, Dead	Humans, Animals, and Environment
ST355	181	1-10	SJDF	Humans, Animals, and Environment
ST388	109	1-10	South Quadrant	Humans, Animals, and Environment
ST405	975	101-200	Fresh water	Humans, Animals, and Environment
ST423	59	0	South Quadrant	Humans, Animals, and Environment
ST441	50	0	Central Quadrant	Humans, Animals, and Environment
ST491	103	0	South Quadrant	Humans, Animals, and Environment
ST526	7	0	SJDF	Human and Environment
ST538	83	0	River otter	Humans, Animals, and Environment
ST540	312	1-10	South Quadrant	Humans, Animals, and Environment
ST542	213	1-10	SJDF	Humans, Animals, and Environment
ST547	11	0	Seal, Dead	Human and Animal
ST549	13	1-10	SJDF	Humans, Animals, and Environment
ST569	88	1-10	Porpoise	Humans, Animals, and Environment
ST603	48	1-10	Central Quadrant	Humans, Animals, and Environment
ST616	92	0	Fresh water	Humans, Animals, and Environment
ST641	424	11-25	South Quadrant	Humans, Animals, and Environment
ST666	8	1-10	SJDF	Humans, Animals, and Environment

ST679	72	1-10	South Quadrant	Humans, Animals, and Environment
ST711	35	1-10	North Quadrant	Animal and Environment
ST714	19	1-10	North Quadrant	Humans, Animals, and Environment
ST737	27	1-10	South Quadrant	Humans, Animals, and Environment
ST746	244	1-10	North Quadrant	Humans, Animals, and Environment
ST772	16	1-10	Seal, Alive	Humans, Animals, and Environment
ST847	102	1-10	Sole	Humans, Animals, and Environment
ST998	210	1-10	North Quadrant	Humans, Animals, and Environment
ST1020	3	1-10	SJDF	Humans and environment
ST1049	81	0	Central Quadrant	Humans, Animals, and Environment
ST1056	44	1-10	North Quadrant	Humans, Animals, and Environment
ST1065	2	0	Seal, Dead	Human only
ST1083	4	0	Central Quadrant	Humans, Animals, and Environment
ST1112	36	1-10	South Quadrant	Animal and Environment
ST1159	58	1-10	SJDF	Humans, Animals, and Environment
ST1162	6	1-10	Central Quadrant	Animal and Environment
ST1236	39	1-10	SJDF	Humans, Animals, and Environment
ST1246	25	1-10	River otter	Animal and Environment
ST1257	9	1-10	SJDF	Humans, Animals, and Environment
ST1304	75	1-10	SJDF	Humans, Animals, and Environment
ST1308	75	1-10	Marine water by beaches	Humans, Animals, and Environment
ST1421	63	1-10	South Quadrant	Humans, Animals, and Environment

ST1423	48	0	North Quadrant	Humans, Animals, and Environment
ST1656	14	0	Central Quadrant	Humans, Animals, and Environment
ST1704	14	1-10	South Quadrant	Animal and Environment
ST1706	26	1-10	Central Quadrant	Humans, Animals, and Environment
ST1711	3	0	South Quadrant	Animal Only
ST1723	37	1-10	Porpoise	Humans, Animals, and Environment
ST1880	6	0	River otter	Animal and Environment
ST1946	23	0	Seal, Alive	Humans, Animals, and Environment
ST1955	9	0	North Quadrant	Humans, Animals, and Environment
ST2025	31	0	Seal, Dead	Humans, Animals, and Environment
ST2063	5	1-10	Central Quadrant	Animal and Environment
ST2079	33	0	Central Quadrant	Animal and Environment
ST2164	6	0	River otter	Environment only
ST2165	34	1-10	Sole	Humans, Animals, and Environment
ST2178	73	1-10	Central Quadrant	Humans, Animals, and Environment
ST2354	23	0	Central Quadrant	Animal and Environment
ST2485	30	1-10	Central Quadrant	Humans, Animals, and Environment
ST2522	83	1-10	North Quadrant	Humans, Animals, and Environment
ST2524	6	1-10	North Quadrant	Animal and Environment
ST2541	38	1-10	South Quadrant	Humans, Animals, and Environment
ST2557	7	0	Fresh water	Animal and Environment
ST2562	12	1-10	Seal, Alive	Humans, Animals, and Environment
ST2602	22	0	North Quadrant	Humans, Animals, and Environment

ST2766	5	1-10	SJDF	Animal and Environment
ST2772	12	1-10	Central Quadrant	Humans, Animals, and Environment
ST2792	21	0	South Quadrant	Humans, Animals, and Environment
ST3018	43	1-10	Seal, Dead	Humans, Animals, and Environment
ST3601	22	O	SJDF	Humans, Animals, and Environment
ST3672	20	0	SJDF	Humans, Animals, and Environment
ST3889	21	1-10	SJDF	Humans, Animals, and Environment
ST4038	61	1-10	South Quadrant	Humans, Animals, and Environment
ST4481	16	1-10	Porpoise	Animal and Environment
ST4657	3	0	Central Quadrant	Humans, Animals, and Environment
ST5037	4	0	North Quadrant	Environment only
ST5073	9	0	North Quadrant	Animal and Environment
ST5082	33	0	South Quadrant	Humans, Animals, and Environment
ST5869	11	1-10	Central Quadrant	Human and Animal
ST6187	2	0	Central Quadrant	Animal and Environment
ST6188	10	0	Central Quadrant	Animal and Environment
ST6476	4	0	South Quadrant	Environment only
ST6506	12	0	SJDF	Humans, Animals, and Environment
ST6521	3	0	South Quadrant	Environment only
ST6603	5	0	North Quadrant	Animal and Environment
ST7706	4	0	South Quadrant	Animal and Environment
ST7900	6	0	North Quadrant	Animal and Environment
ST8271	3	0	SJDF	Animal and Environment

ST8501	2	0	North Quadrant	Animal and Environment
ST8612	2	0	Central Quadrant	Animal and Environment
ST8820	3	0	South Quadrant	Environment only
ST8927	6	0	Central Quadrant	Animal and Environment
ST9001	2	0	Seal, Alive	Human and Environment
ST9293	4	0	SJDF	Animal and Environment
ST9422	3	0	Seal, Dead	Humans, Animals, and Environment
ST10284	2	0	SJDF	Human and Environment
ST10431	3	0	North Quadrant	Human and Environment
ST10513	1	0	South Quadrant	Environment only
ST10516	2	0	South Quadrant	Animal and Environment
ST10537	1	0	SJDF	Environment only
ST10538	1	0	North Quadrant	Environment only
ST10717	1	0	Central Quadrant	Environment only
ST10718	1	0	Central Quadrant	Environment only
ST10751	1	0	Central Quadrant	Environment only
ST10780	1	0	South Quadrant	Environment only
ST10781	1	0	Central Quadrant	Environment only
ST10782	1	0	SJDF	Environment only
ST10788	2	0	SJDF	Environment only
ST10818	2	0	SJDF	Human and Environment
ST10819	1	0	SJDF	Environment only
ST11031	1	0	South Quadrant	Environment only
ST11184	1	0	North Quadrant	Environment only
ST11222	1	0	North Quadrant	Environment only
ST11342	1	0	Seal, Dead	Human and Animal
ST11343	1	0	Seal, Dead	Human only

(SJDF = Strait of Juan de Fuca)

Table S4. ST found in this study that were associated pathogen types in EnteroBase.

	EnteroBase STEC/VTEC present	***EnteroBase EPEC present***	***EnteroBase EAEC present***	***EnteroBase ExPEC present***	***EnteroBase EIEC present***	***EnteroBase ETEC present***	***EnteroBase EHEC Present***
ST2	N	N	N	N	Y	N	N
ST10*	Y	Y	Y	Y	N	Y	N
ST12	N	N	N	Y	N	N	N
ST13	N	N	N	Y	N	N	Y
ST24	N	Y	N	N	N	N	N
ST26	N	N	N	N	N	N	Y
ST28	N	Y	N	N	N	N	Y
ST38*	N	N	Y	Y	N	N	N
ST40	N	Y	Y	N	N	N	N
ST46	N	N	Y	Y	N	Y	N
ST48	N	N	Y	Y	N	N	N
ST58	Y	N	N	Y	Y	N	N
ST69	N	N	N	Y	N	N	N
ST73*	N	N	N	Y	N	N	N
ST75*	N	N	N	Y	N	N	N
ST83	N	N	N	Y	N	N	N
ST101*	Y	Y	Y	Y	N	N	N
ST109	N	N	N	Y	N	N	N
ST117*	N	N	Y	Y	N	N	N
ST127*	N	Y	N	Y	N	N	N
ST130	N	N	Y	N	N	N	N
ST131*	N	N	Y	Y	N	N	N
ST132*	N	N	N	Y	N	N	N
ST144	N	N	Y	N	N	N	N
ST154	N	Y	N	N	N	N	N
ST155*	Y	N	Y	Y	N	Y	N
ST162*	N	N	N	Y	N	N	N
ST206*	N	Y	Y	N	N	N	N
ST224*	N	N	Y	N	N	N	N
ST297*	Y	N	N	N	N	N	N
ST327*	N	Y	N	N	N	N	N
ST349*	N	N	Y	N	N	N	N
ST355	N	N	N	Y	N	N	N
ST362	N	N	Y	N	N	N	N
ST372*	N	N	N	Y	N	N	N
ST405	N	N	N	Y	N	N	N
ST423	N	N	Y	N	N	N	N
ST526	N	Y	N	N	N	N	N

ST542	N	N	N	Y	Y	N	N
ST616	N	N	N	Y	N	N	N
ST681*	N	N	N	Y	N	N	N
ST720	N	N	Y	N	N	N	N
ST746	N	N	Y	Y	N	Y	N
ST847	N	N	N	N	N	Y	N
ST906*	N	N	N	Y	N	N	N
ST937	N	N	N	Y	N	N	N
ST998	N	N	N	Y	N	N	N
ST1049*	N	N	N	Y	N	N	N
ST1079*	N	N	N	Y	N	N	N
ST1248	Y	N	N	N	N	N	N
ST1304*	N	N	N	Y	N	N	N
ST1421	N	N	N	Y	N	N	N
ST2164	N	N	Y	N	N	N	N
ST2354	N	N	N	N	N	Y	N
ST2522	N	N	N	N	N	Y	N

Legend
Y= yes
N=no
*ST found in this study that were found in shellfish and marine animals in EnteroBase

Figure S1. Field collection sites based on source type

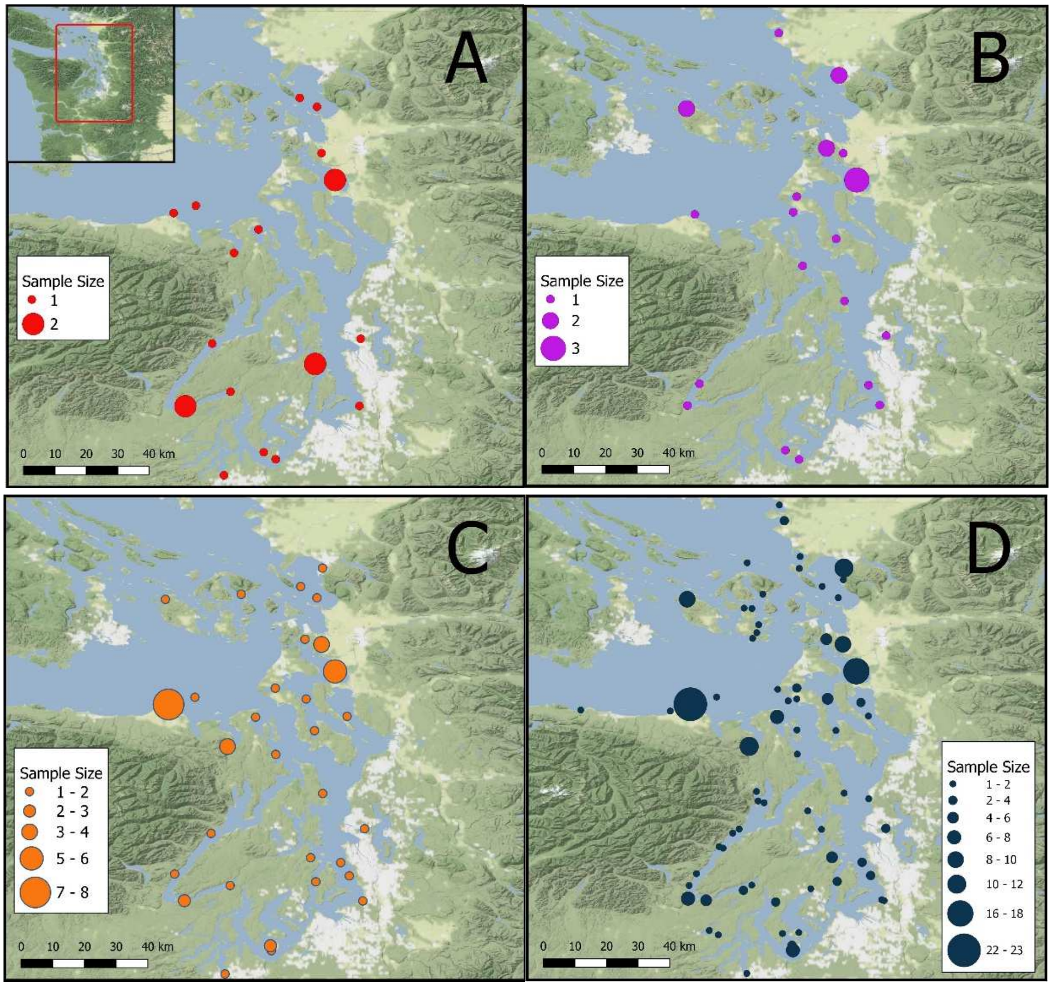


Figure S1A: All marine water ST in this study found in environment only in EnteroBase; **Figure S1B**: All marine water ST in this study found in animals and environment in EnteroBase; **Figure S1C**: All marine water ST in study designated as novel (ten or less entries in EnteroBase); **Figure S1D**.: All marine water ST in this study.