Microbial Source Identification for the Watsonville Slough Locations during Dry and Wet Flow Periods

Prepared by

B. H. Olson, Ph.D.
Professor Environmental Health Science and Policy
School of Social Ecology
Environmental and Community Medicine
School of Medicine
Interim Director Urban Water Research Institute
and
Joann Le, B.S. Laboratory Technician
Master's Student
Department of Environmental Health Science and Policy
University of California
Irvine, California

April 13, 2003

Executive Summary

- **Patterns in wet and dry flow periods show that different waste sources are prevalent**
  - **In the summer, July 2003 bird waste was dominant and decreased as a percentage of the total fecal biomarkers identified at the sampling locations, although mean MPN values remained the same for summer and winter samples.**
  - **In the winter, December 2003 cow and human waste sources (MPN values/ 100 ml) increased during the winter.**
    - **Dog sources increased at two sampling points during the winter sample.**
  - **Rabbit sources increased at some locations during the winter sampling.**
- **The low occurrence of human input (0. 2% to 3%) is important because disease organisms that affect humans are most likely to be present in this waste source. Albeit at these concentrations, risk is likely to be reduced.**
  - **Occurrence of human waste sources will undergo an additional confirmation at UCI by sequencing the biomarker.**
- **Cow wastes is also important due to its potential to contain organisms that produce disease in humans**
  - **This waste source appears to be entering during storm periods from upstream areas or from re-suspended sediments.**
  - **Cow waste inputs, especially if redistributed from the sediments, could be very important because of the ability to carry O157:H7. Previously published work by our laboratory showed that approximately 1:185 *E. coli* had the O157 genotype (hemorrhagic diarrhea) in cow/cattle waste lagoons (Chern et al. 2004). Also, this genotype has been shown to survive in excess of 200 days in some soil environments (Kudva et al. 1998).**
- **Although dog fecal waste is not a high proportion of the waste in most instances, it did represent 70% of the identified biomarkers at the Har location during the winter sample. Dog waste can be controlled through extensive educational activities, as well as placing dog-waste disposal bags, in walking areas.**

Recommendations

- **If not in place, a video log detection program for sewer leaks should be implemented at yearly intervals due to seismic activity. Evaluation of septic tanks or leach fields as well as any toilet facilities and associated piping installed in county or state parks in these areas should also occur. Understanding sanitation facilities at parks and recreational camping areas as well as land use patterns of transients who would not use public facilities is important in determining human inputs.**

- **Importantly, the building code allows small separations between sewage pipes that might be a source for streams or lagoons that are recharged by subsurface waters. Locations where the geologic formations are alluvial are at high risk.**
    - **Follow up testing on occurrence of human waste sources during the winter when no irrigation or rain is occurring should be conducted, in order to better pin point human waste sources from sewers. Leaks are most easily detected during this time, because the human biomarker will be the most concentrated if sewer systems are not combined. Higher occurrence of human and cow waste biomarkers were found at most locations in the winter.**
    - **In reality, birds cannot be controlled. An investigation of the use of raptors to frighten birds from beaches is being studied at Huntington Beach, California.**
    - **Local city or county action should reduce dog input. This has been successful in a number of communities, but a very low input is likely to remain.**

Introduction:

The occurrence of fecal coliforms in natural waters from non-point source run-off is an important issue facing agencies that have the responsibility to reduce the total maximum daily load of pollutants entering their waters. A number of methods have been developed by investigators to determine the origin of these fecal wastes. Sources of fecal waste are an important issue as some wastes do not contribute any disease potential to humans using these waters or the coastal areas into which these waters flow. Other sources such as cattle or human inputs from septic tanks or the homeless could pose human health threats. **Importantly, no method to differentiate fecal waste today is more than experimental in nature. Each has demonstrated drawbacks and all are under development. Several reviews are available and two have been included as supplements to this report.**

The method used in the current study employs single or nested PCR to identify toxin genes in *E. coli* associated with certain host groups. The toxin gene method does not require a library to be established within a geographical region, because the traits used have been shown to have a world wide distribution. Also, the sensitivity of the method is very good, being able to identify one cell within millions or billions of *E. coli*, not carrying the trait of interest. The method has been shown to have geographical efficacy across the U.S. with the biomarker for cows (LTIIa, being found in 11 states including Hawaii). The biomarker used for humans, STh (STIb), has a worldwide distribution and has been found in major sewage treatment plants from several states in the U.S. besides California (Olson et al, 2004 submitted). The pig biomarker, STII is endemic in pig farms across America, while the *ral*G, biomarker for rabbits has been most extensively

investigated in Europe and Australia (Adams et al. 1997). The *tsh* gene has been found to be associated with birds of economic importance (Province et al 1994). Our research has shown that it is present in a wide variety of birds from both zoos and the wild in California. This biomarker has been reported in *E. coli* isolated from human urinary infections, probably from dogs that are the source of *E. coli* that cause urinary tract infections in humans. The last biomarker used in this study was the biomarker for the dog, *pap*G allele III (Johnson et al. 2000). This gene helps uropathogenic *E. coli* attach to urinary cells and is being widely studied in humans and other animals currently by clinical microbiologists. We have been able to identify unique sequences in isolates from dogs and humans, which we use to distinguish the two.

Another advantage of the toxin biomarker method is that all *E. coli* in a sample are screened. The summer samples clearly show that certain biomarkers require much greater numbers to be screened to produce a positive; suggesting low occurrence. All ribotyping or PCR based 16S RNA or rDNA library methods must isolate pure cultures and rely on a subset of *E. coli* isolates. These methods may test only five isolates per sample. In this study, that would have meant that the percentage of *E. coli* screened would have ranged from a high of 0.5% using the *E. coli* MPN value 930/100ml at Har to 0.1% in most of the other samples. The occurrence of human inputs may have been missed since a maximum of five isolate would have been typed. Thus, the *E. coli* toxin biomarker method screened 99.9% more *E. coli* than other commonly used methods.

The antibiotic resistance methods screen approximately 37 isolates per site. Thus, the biomarker method is more accurate at detecting minor inputs since subsets of isolates are not used. The toxin biomarker method is also faster since pure cultures are not necessary. A number of workers have independently tested the antibiotic methodology and validated that it can distinguish collectively wild animal population wastes from other sources.

To date, all of these markers have the potential to produce false-positives, except the biomarkers for the cow and rabbit. The incidences of these false-positive rates are very low, <1 percent in most instances. However, all data collected by all methods available today are subject to these types of issues. Ours is the only method that has made an effort to search for and identify false-positive possibilities with the traits. Two other methods, one based on antibiotic resistance and the other based on 16S RNA and restriction analysis using certain endonucleases have found that 20% of *E. coli* are cosmopolitan (inhabit the intestinal tract of several animals) and must be discarded from their analyses in order to classify waste sources. Thus, some researchers believe that these methods can only identify sources that are 20% or greater of the total.

This report summarizes results from July 2003 and December 2004.

Materials and Methods:

Bacterial strains and DNA preparations. The Enterotoxingenic *E. coli* strains containing the pTC201 plasmid with the LTIIa toxin gene, pDAS100 plasmid with the STh toxin gene, and pXE39 plasmid with the ralG toxin gene were grown in LB media (Difco, Detroit, Michigan) amended with 20 ug/mL of ampicillin. The plasmid pNS1

with the hlyE gene was grown in LB media supplemented with 50 ug/ml of kanamycin. All strains, in LB media, were incubated at 37°C with agitation at 150 rpm overnight. A one milli-liter aliquot of the each bacterial culture was centrifuged at 12,000xg for 5 minutes using a bench top Eppendorf Centrifuge (Model 5415D Netheler-Hinz, Hamburg, Germany). The pellet was collected and DNA extracted immediately or stored at 4°C until analysis.

Membrane filtration and Extraction of DNA. Samples received were serially diluted and filtered through 0.45 µm nitrocellulose membranes (Fisherbrand) then placed on mTEC agar plates. The mTEC agar plates were incubated at 35ºC for two hours then 44.5 ºC for 18 hours. Yellow *E. coli* cells grew on mTEC plates were enumerated and collected for collected for direct total DNA extraction by a freeze-thaw and phenol-chloroform method as described by Tsai and Olson (1991).

Nested PCR Amplification. Outer primers for the cow biomarker (Khatib *et al*., 2002) and human biomarker were obtained from previous research. The specificity of primers was determined in earlier studies. (Khatib *et al*., submitted 2001a,b; 20) A second set of primers for each toxin trait was developed for nested PCR. These primers were tested for cross-reactivity by screening all sequences contained in GenBank (http://www.ncbi.nlm.nih.gov/) using BLAST. Amplification was performed using 1-10 µl of DNA sample extract for each 50 µl reaction, which consisted of 5 mM Tris-HCl (pH 8.3), 25 mM KCl, 1.0 mM $MgCl_2$, 100 µM of each dNTP, 0.2 µM of each primer and 2.5 U of *AmpliTaq* DNA polymerase (Perkin-Elmer or Promega). These mixtures were heated to 95^{o}C for 1 minute followed by 30 cycles of 95^{o}C for 30s, 61^{o}C (LTIIa) for 30s or 57^{o}C (STII) for 30s, and 72^{o}C for 30s with final extension at 72^{o}C for 6 minutes. Annealing temperatures for nested PCR of the LTIIa biomarker was at 56^{o}C and 47^{o}C, respectively. Amplification of the human biomarker was done at 94^{o}C for 1 minute followed by 35 cycles of 94^{o}C for 30s, 45^{o}C (50^{o}C for nested PCR) for 30s and 72^{o}C for 30s with final extension at 72^{o}C for 6 minutes. All reactions were amplified in a Perkin Elmer (model 9600, version 1.05) DNA thermal cycler. All PCR amplicons were visualized through gel electrophoresis.

Confirmation of PCR Product. Protocols have been established to determine that the fragment produced does have similar sequences to those in the model gene using restriction enzyme digests. The restriction enzyme, BstAPI (New England Biolab), is used to produce two fragments with the sizes of 121bp and 45 bp from PCR amplicon for the STh toxin gene. Two restriction enzymes, PstI and AluI are used to confirn PCR amplicons for the LTIIa toxin genes. PstI produces two fragments with sizes of 245bp and 113bp, and AluI produces two fragments with sizes of 222bp and 136 bp. Southern Blot hybridization is all used to confirm PCR amplicons using previously designed probes specifically for each toxin.

MPN-PCR Quantification of Biomarker (Toxin gene) Occurrences. Three of five dilutions, with three replicates per dilution, from each sample were used to calculate the MPN. The occurrence of the toxin genes were calculated using the MPN-PCR analysis per filter, which was specified as negative or positive. MPN-PCR results were input into

the MPN calculator, build 23 (http//member.ync.net/mcuriale/mpn/index.html) to obtain the MPN/100 ml based on 95% confidence interval.

Quality Control and Assurance Protocol

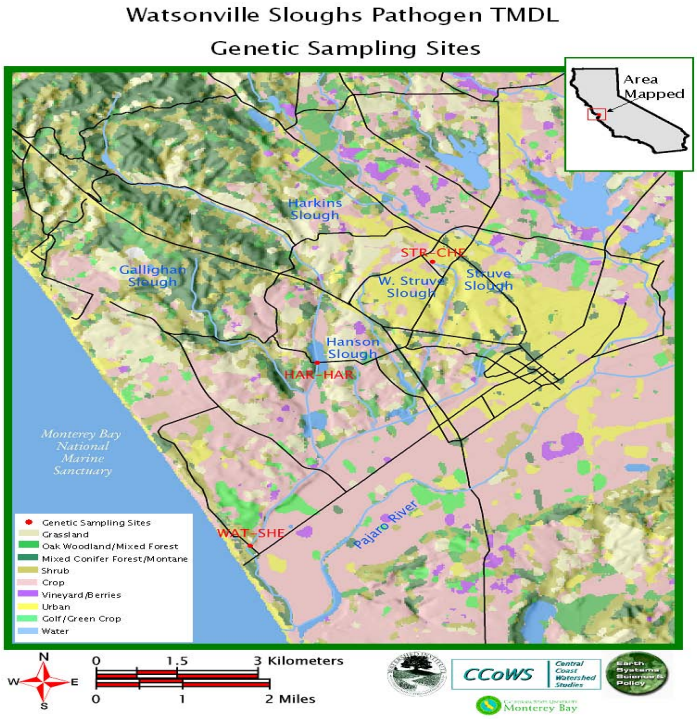
General Laboratory Procedures. For quality control purposes, laboratory equipments like refrigerators (1-4°C) and freezers (-20° to -80°) are monitored daily and temperature are recorded. Biohazard hood is monitored and tested once a month by exposing plate count agar plates to the airflow in hood to confirm the no bacterial growth. The thermal cycler, pipettes, centrifuges, spectrophotometer are calibrated prior to performing tests to ensure their accuracy.

Processing of DNA. There is no standardized protocol for quality control in regard to processing DNA. This laboratory has developed the following procedures to insure quality control. A positive and negative control is included with each set of extractions that are performed. A minimum detection limit is determined for each set of water samples.

***PCR* Analysis.** Interference with the PCR reactions is determined by a protocol reported in Oshiro et al. (1997) and Khatib et al (2002). PCR assays are performed in a designated hood with UV light. The hood is sanitized with 10% bleach follows by 75% ethanol then UV radiation for 5 minutes. Carry over in the PCR reaction is controlled for by a set of positive and negative control, which is included with each PCR run.

Additional Quality Control. If results are obtained that are unclear, the fragment is excised from the gel, sequenced and the sequence compared to the Genbank target sequence to authenticate the result.

Sampling locations: The following sampling locations were tested for the presence *E. coli* on the Watsonville Slough (Map 1).



Map 1. Watsonville slough drainage area with sampling locations: Str-che A-C (Struve Slough, replicates 1-3), Har-har A-C (Harkins Slough, replicates 1-3), Wat-she A-C (W. Struve Slough, relicates 1-3).

Results:

Introduction

Total *E. coli* per 100 ml were determined using a Most Probable Number Method or direct counts for the summer and winter periods of 2003 (Figures 1a and 1b). The data show quite clearly that the methods are slightly different, but the values in this work are relativity close. However, in the summer of 2003 another dilution should have ideally been included in the MPN testing. This can be seen by the frequency of 2400 *E. coli* per 100 ml. Thus, the direct counts more accurately reflect the values for the total direct counts. An analysis of MPN comparing a hypothetical value, if another dilution had been included, approach those reported by direct counts. This emphasizes the importance of capturing all dilutions. This should not affect the results, since all the organisms captured are screened. Because our results are based on a MPN method for the determination of

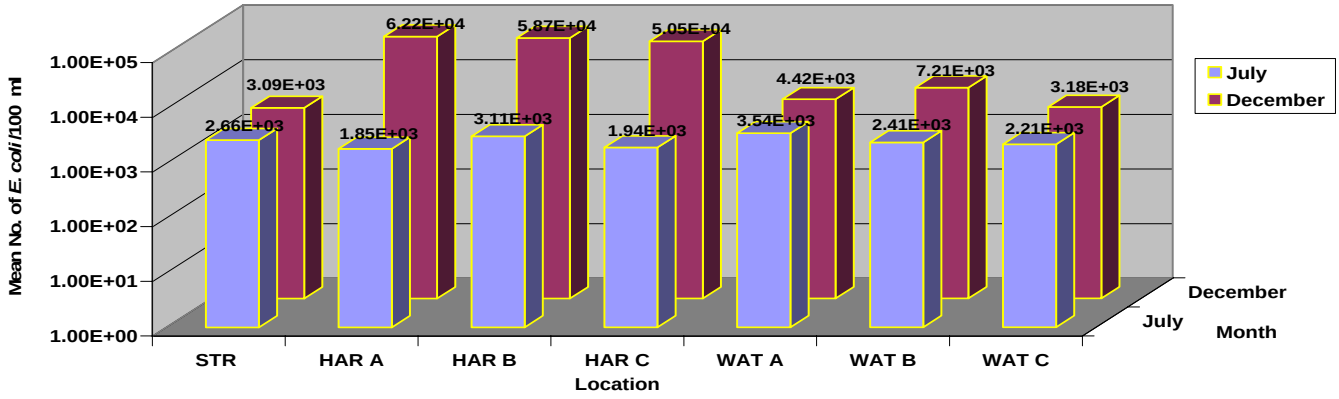


Figure 1a. Mean *E. coli* /100 ml between samples collected from Watsonville Slough during dry and wet weather flows using direct counts

fecal waste sources, we routinely compared direct counts and the MPN method. The most probable number method has been used extensively in fields where a Poisson distribution or low probability of positives exists. By capturing the entire population, the examiner has the ability to analyze the largest subsets of populations with the least effort. A Poisson distribution is a good estimation of a binomial distribution when the probabilities are small and the number of trials is large. In the most probable number method the 95% confidence intervals are broad and therefore, some precision is sacrificed to capture infrequent occurrences (see Figure 1c). As can be seen in the Har summer 2003 direct counts which seem low when compared with the MPN values, in actuality the 95% confidence intervals bracket the direct counts, reflecting some of the imprecision and indicating how the sum of individual biomarker results can be greater than the mean of MPN value for total *E. coli*

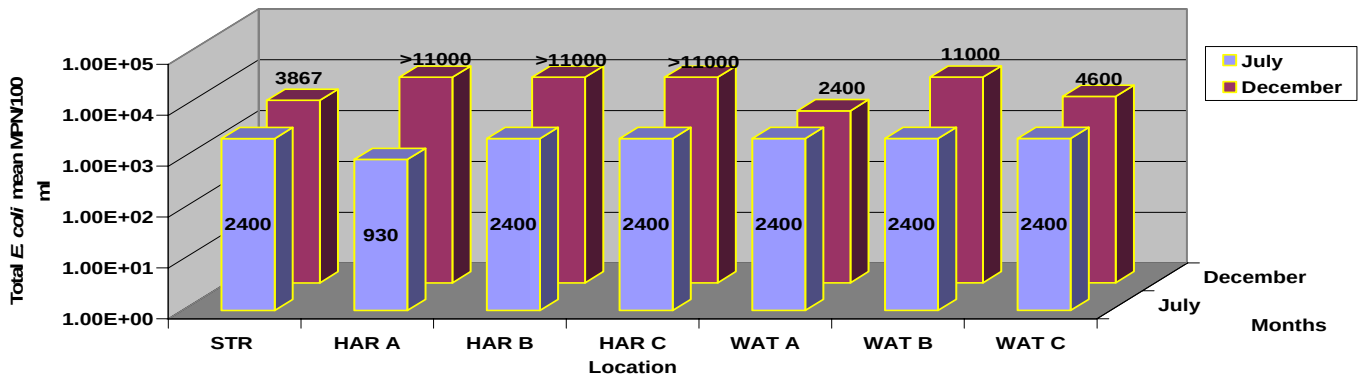


Figure 1b. *E. coli* counts in MPN/100 ml between samples collected from Watsonville Sloughs in dry and wet weather flows using MPN. **Note: The notation, >, by the MPN numbers of samples HARA-C denotes that the E. coli counts, in MPN numbers, for the sample replicates are greater than 11,000.

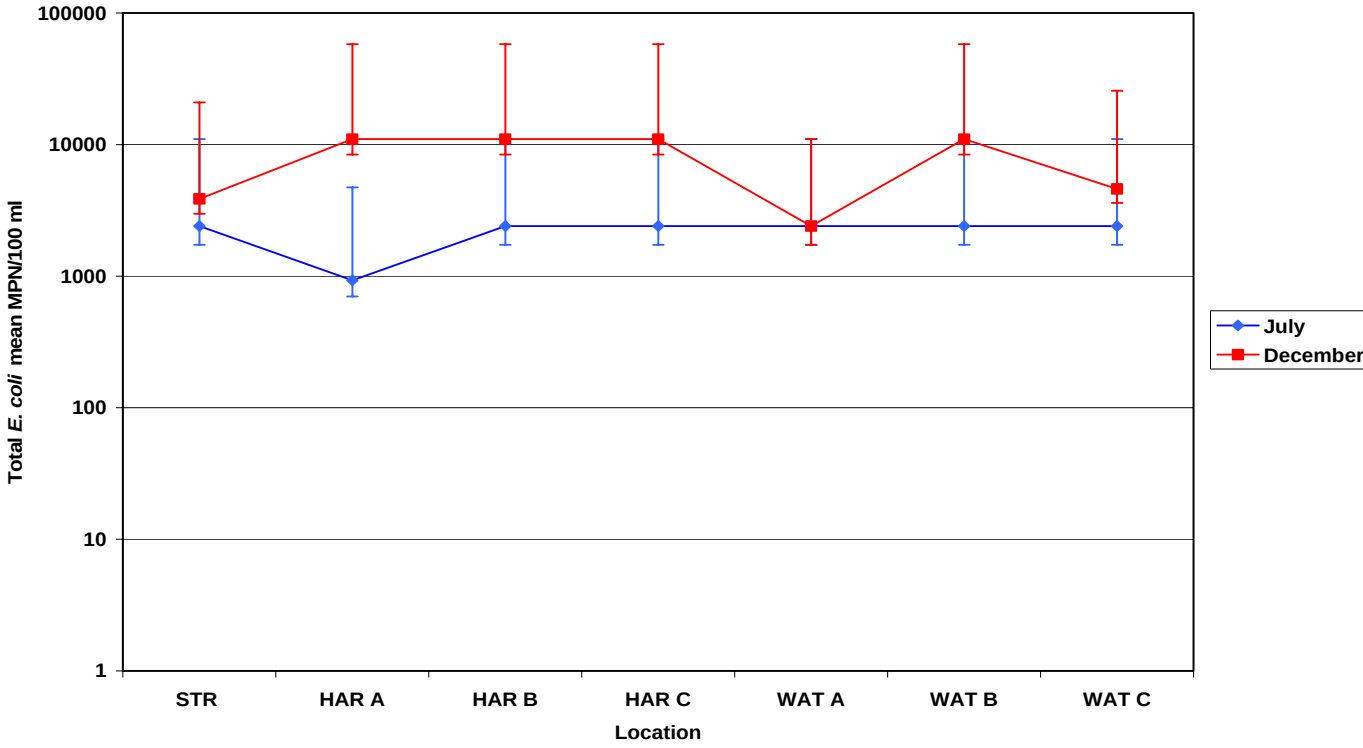


Figure 1c. Upper and lower 95% interval limits of total *E. coli* mean MPN/100 ml samples collected in July and December samples from Watsonville Sloughs

Comparison of Replicates at Sample Locations

The data comparing individual sample replicates within a sampling location showed that although variation occurred, the standard deviations overlapped. Given that bacterial counts can easily vary 0.5 logs due to clumping during normal plating procedures, these values are well within the expected range.

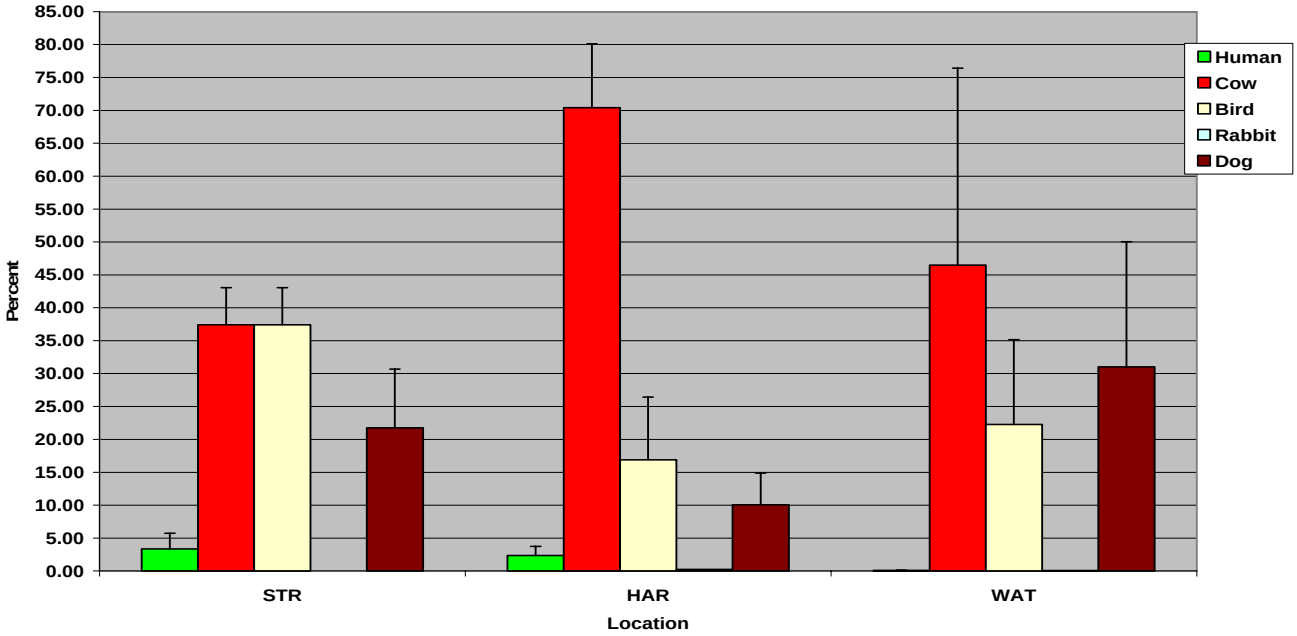


Figure 2. The means and standard deviations of percent distribution of identified biomarkers in the samples collected in December of 2003 from Watsonville Slough locations.

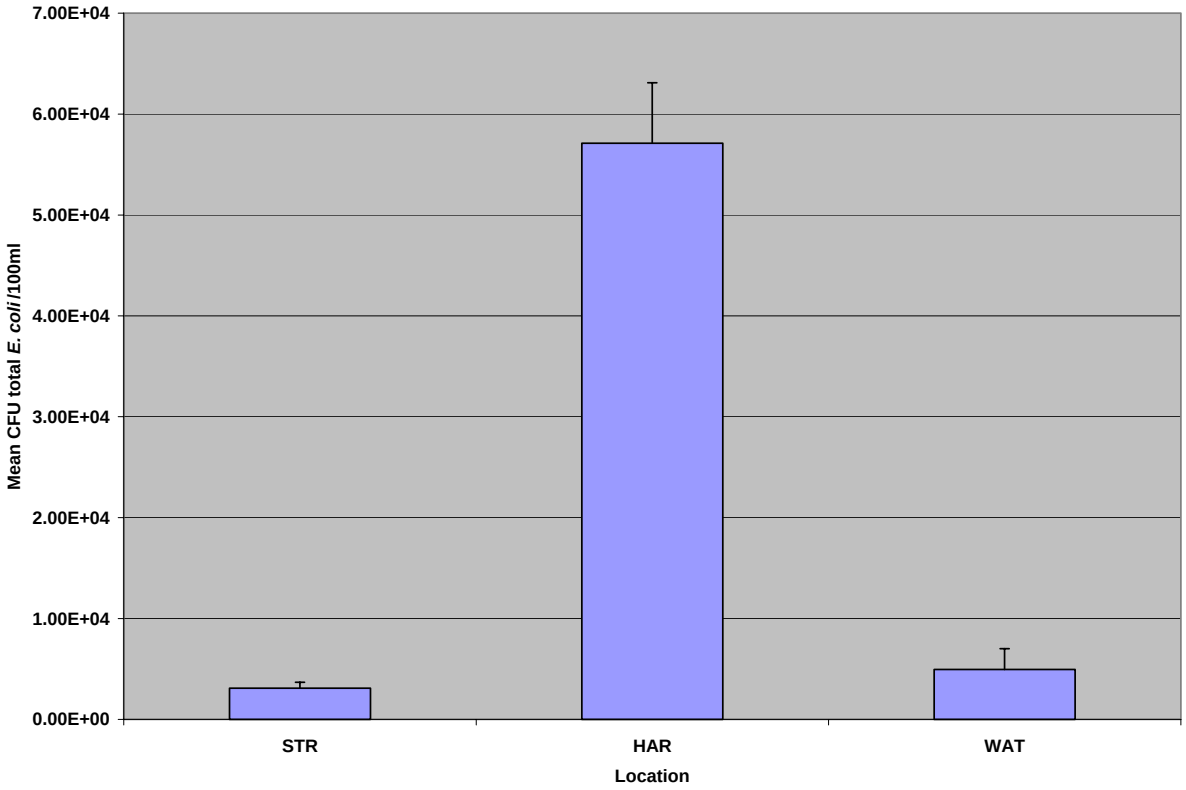


Figure 3. The means and standard deviations of *E. coli* direct count in CFU/100 ml of sample replicates collected in December of 2003.

Screening for Presence of Biomarkers

We have established a procedure that allows the investigator and the agency to obtain an idea of the frequency of the biomarker amongst *E. coli* populations. The graphed data in Figure 4 indicates that as the height of the bar increases, more *E. coli* had to be screened to obtain a positive result for each biomarker in the study. It can be seen that the biomarker for rabbit fecal material was rare and a thousand *E. coli* had to be screened to obtain the positives. It also clearly shows the differences in frequencies for cows, birds, and humans. Dogs do not seem to have as much of a seasonal trend (winter vs. summer) at the HAR sampling location as at the STR and WAT sampling locations. Generally there is more fecal input of waste from the land/sediments during storm periods. Importantly, including additional biomarkers such as the pig in our analysis would change the percentage of each biomarker, but the number screened to obtain a positive result and MPN values/100 ml (figure 5) of each would not change. Thus, number of *E.* coli screened for the presence of each biomarker is an independent assessment of accuracy for this methodology.

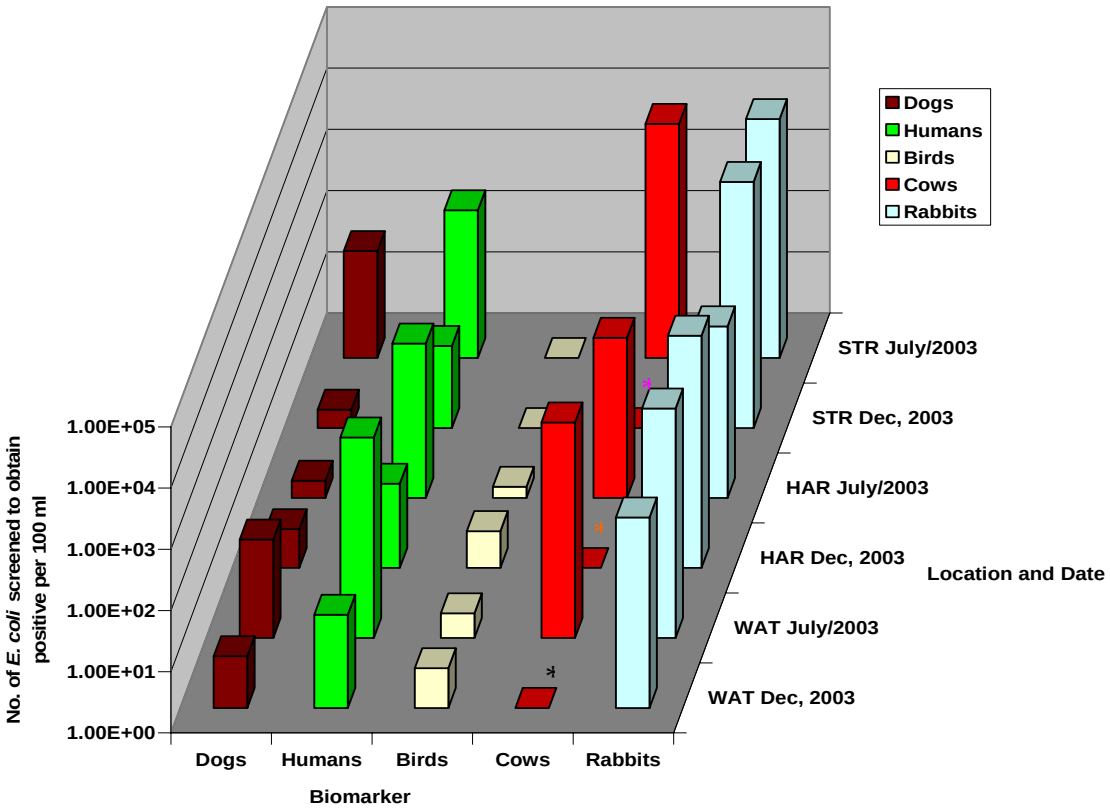


Figure 4. Comparison of the number of *E. coli* screened to obtain positive biomarker result between samples collected from Watsonville Slough during dry and wet weather flows. December samples for *STR, *HAR, *WAT indicate that $\geq 1 \leq 2$, $\geq 1 \leq 8$ and $\geq 1 \leq 3$ *E. coli* must screen to obtain positive result for the cow biomarker, respectively.

Occurrence of Fecal Waste Sources in Sampling Areas

In Figure 5, there are clear differences at most locations in the occurrence or frequency of most of the five fecal biomarker sources tested between summer and winter. Since the total numbers *E. coli* increase between summer and winter, the percent representation of a source in the population can change (see Figure 6), but the MPN value can stay the same as is the case for birds (Figure 5). The relative sampling ease of obtaining a positive is greater for birds than for dogs, regardless of season (see Figure 4). In the case of cow, human, and to some degree rabbit waste, trait frequencies increase in the winter. Again, as we suggested, either materials are being brought to sampling areas during high flows, or sediments are being re-suspended that contain these wastes. Although the time intervals between the wet and dry seasons in California seem extremely long, studies on *E. coli* 0157 in soils showed viable bacteria could be isolated 200 days after inoculation (Kudva et al 1998). The survival of *E. coli* that carry the cow biomarker in estuarine water, is >100 days, if the bacteria are protected from predation (Khatib et al. 2002). Although the frequency of *E. coli* carrying the biomarker for human waste seems low (1

to approximately 100), a study of an urban area (Laguna Nigel Report, 2001) found no *E. coli* carrying the human biomarker (up to <1 human biomarker: 10,000 *E. coli*). However, in areas where septic tanks or leach fields are more common, ratios in this range might be expected.

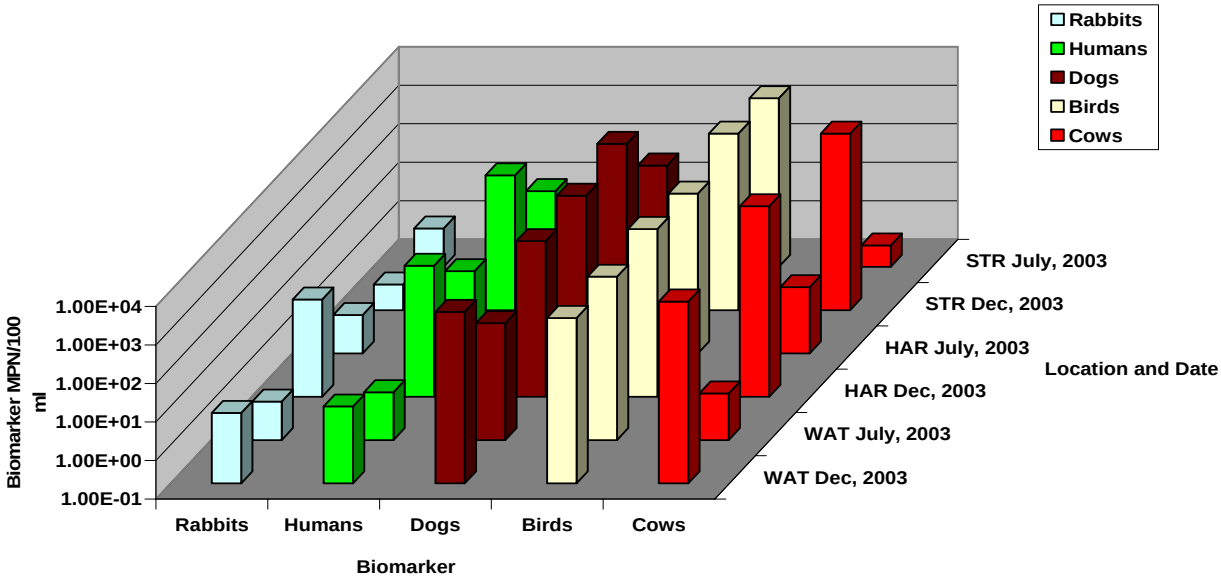


Figure 5. Comparison of occurrences of different fecal sources identified in water samples collected from Watsonville Sloughs during dry and wet weather flows

Contributions to Waste by Source

Figure 6 shows that during the summer months, birds make up the vast amount of fecal material entering the different water sampling locations. Cow waste, although persistently detected in the summer represented < 0.02% of the isolates, but increased in the winter, which might reflect grazing patterns. At the HAR location, the July sample shows a high amount of dog waste present, while the other locations show between 1 and 6% of the waste is attributable to dogs for the same period. The high concentration of dog at HAR in the summer could be an anomaly where sampling was impacted by a recent event or could reflect a use difference. As noted in the comparison samples, values changed between replicate samples, sometimes quite remarkably, although within the 95% confidence intervals. These changes have been seen in a more detailed study conducted on the Santa Ana River, where 10 liters of water were collected at 5 minute intervals for 15 minutes. The data look similar in that variations occurred and could be

substantial between samples even with a large volume collected (Shields and Olson, 2004).

Our conclusion is water is moving in a plug-flow fashion (independent packets) representing discreet segments in time and space. From a watershed viewpoint, this information is valuable, because it shows dynamic variability of the flow in the watershed. The Watsonville Slough area is not too dynamic, although approximately 33% of the samples may be different than others. Such differences raise the question of whether data should be presented as a mean value or as a median value. However, there are a too few data points in this study to determine which would be more representative of watershed conditions. At the other two locations, the dog biomarker as a percent of the population increased in the winter, which was also seen in *E. coli* carrying the cow biomarker, but to a lesser extent.

Human sources increase in the winter and to a lesser degree at WAT. This increase, although low as over all percentages, is approximately 10 fold. After obvious interventions are taken, eliminating such sources may prove difficult, for example, unauthorized campers in locations where, during the winter, materials are washed from the land into waterways. The major winter input from identified sources is cow waste and if this input is from upstream locations, then this source must be controlled at those sites (i.e. overflow from dairy farm waste containment facilities). If grasslands are the source, then better buffer zones to reduce introduction into rivulets that later flow into streams could be important.

Regional Boards have to take decisions on birds in areas and their contribution to the local ecosystem as well as the negative impact of fecal loading. Figure 5 shows that absolute amount is the same from summer to winter indicating that there may be a stationary bird population. However, in the winter the bird biomarker as a percentage of fecal sources decreases because of the increase in total fecal input. It is the authors opinion that during the winter sampling time, the increase in total *E. coli* was clearly attributable to cow, human and dog as opposed to nonbiomarker fecal sources.

It is important to consider the environmental importance of species and their usefulness in the ecosystem, apart from the negative impact of fecal loading. From the standpoint of protecting human health as opposed to meeting a water quality standard, human and cow wastes raise the greatest concerns. Human waste can contain protozoan and viral pathogens which survive longer and have a lower infectious dose to humans than most bacterial pathogens. Animal waste sources are less important from a public health standpoint, because most animal pathogens cannot infect humans. Although there are notable exceptions, such as *Cryptosporidium* and *E. coli* 0157:H7 which are present in

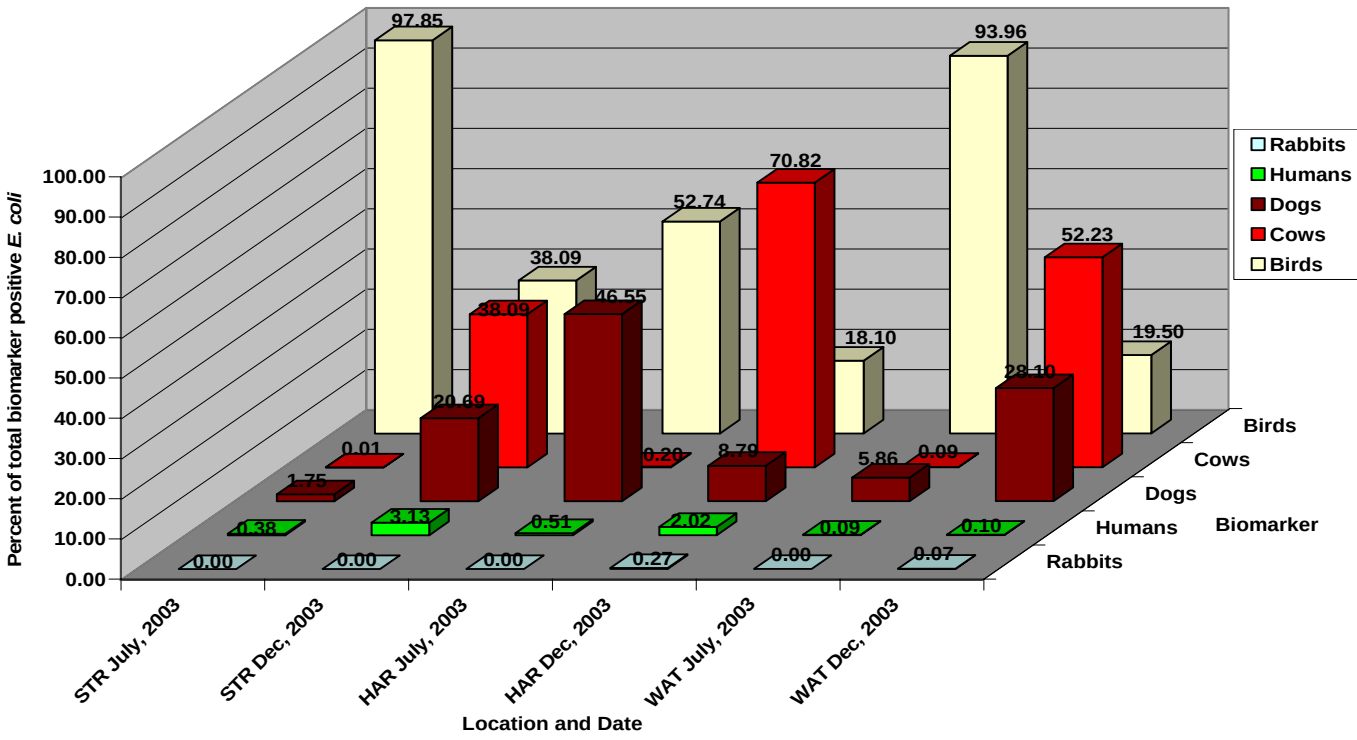


Figure 6. The distribution (%) of *E. coli* carrying the 5 biomarkers amongst biomarker positive *E. coli* in water samples from Watsonville Slough during dry and wet weather flows

cow fecal wastes. Both of these human pathogens have low infectious doses (≅10 organisms) for humans. Since our recent research report showed an average occurrence of *E. coli* 0157 to be 1:185 *E. coli* in cow waste lagoons (Chern et al. 2004), sufficient dilution may or may not have taken place before fecal wastes reach important points where primary contact may be occurring.

Future Research Recommendations:

Three research recommendations are made to aid the agency at identifying what locations are responsible for specific source inputs during wet flow periods (winter months).

- **If cow/cattle input can not be reduced, then sampling during the winter for *Cryptosporidium* and *E. coli* 0157:H7 should be conducted in areas where primary body contact occurs.**
- **Summer testing of sediments and upstream inputs to determine sources that are likely responsible for winter inputs.**
- **Sampling first flush in areas that could be suspected of contributing wet flow inputs**

If the agency can identify the incoming source, then it can likely be linked to a landowner. However, if the sediments are the cause, then historical inputs may be responsible and time for natural removal will be needed. The bird data indicate that flushing to some degree does occur.

Reference:

Adams, L.M., C.P. Simmons, L. Rezmann, R.A. Strugnell, and R.M. Robins-Browne. 1997. Identification and characterization of a K88- and CS31A-like operon of a rabbit enteropathogenic Escherichia coli strain which encodes fimbriae involved in the colonization of rabbit intestine. Infection and Immunity. 65(12), 5222-30.

Chern, E.C., Y. Tsai, B.H. Olson. 2004. Occurrence of genes associated with enterotoxigenic and enterohemorrhagic Escherichia coli in agricultural waste lagoons. Applied & Environmental Microbiology. 70(1), 356-362.

Curial, M. MPN Calculator Build 23. 2003. http//member.ync.net/mcuriale/mpn/.

Indira T. Kudva, Kathryn Blanch, and Carolyn J. Hovde. 1998. Analysis of Escherichia coli O157:H7 Survival in Ovine or Bovine Manure and Manure Slurry. Applied & Environmental Microbiology. 64: 3166-74.

Johnson, J.R., T.T. O'Bryan, D.A. Low, G. Ling, P. Delavari, Fasching, C., T.A. Russo, U. Carlino, and A.L. Stell. 2000. Evidence of commonality between canine and human extraintestinal pathogenic Escherichia coli strains that express papG allele III. Infect. Immun. 68 (6), 3327-3336.

Khatib L.A., Y.L. Tsai, and B.H. Olson. 2002. A biomarker for the identification of cattle fecal pollution in water using the LTIIa toxin gene from enterotoxigenic Escherichia coli. Appl Microbiol Biotechnol. 59(1): 97-104.

Olson, B.H. 2003. Report on water quality fecal contamination testing: City of Laguna Niguel: J02P03. County of Orange, California.

Oshiro, R.K. and B.H. Olson. 1997. Occurrence of Sth toxin gene in wastewater. In Coliforms and *E. coli*, problem or solution? *Edited by* D. Kay and C. Fricher. The Royal Society of Chemistry. Cambridge. pp 255-259.

Genbank. 2004. National Center for Biotechnology Information U.S. National Library of Medicine. http://www.ncbi.nlm.nih.gov.

Provence, D.L., and R. 3rd Curtiss. 1994. Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic Escherichia coli strain. Infect Immun. 62(4):1369-80.

Tsai, Y.L., and B.H. Olson. 1991. Rapid Method for Direct Extraction of DNA from soil and sediments. Applied & Environmental Microbiology. 57(4). 1070-1074.

Tsai Y.L., J.Y. Le, and B.H. Olson. 2003. Magnetic bead hybridization to detect enterotoxigenic Escherichia coli strains associated with cattle in environmental water sources. Canadian Journal of Microbiology. 49(6), 391-398.