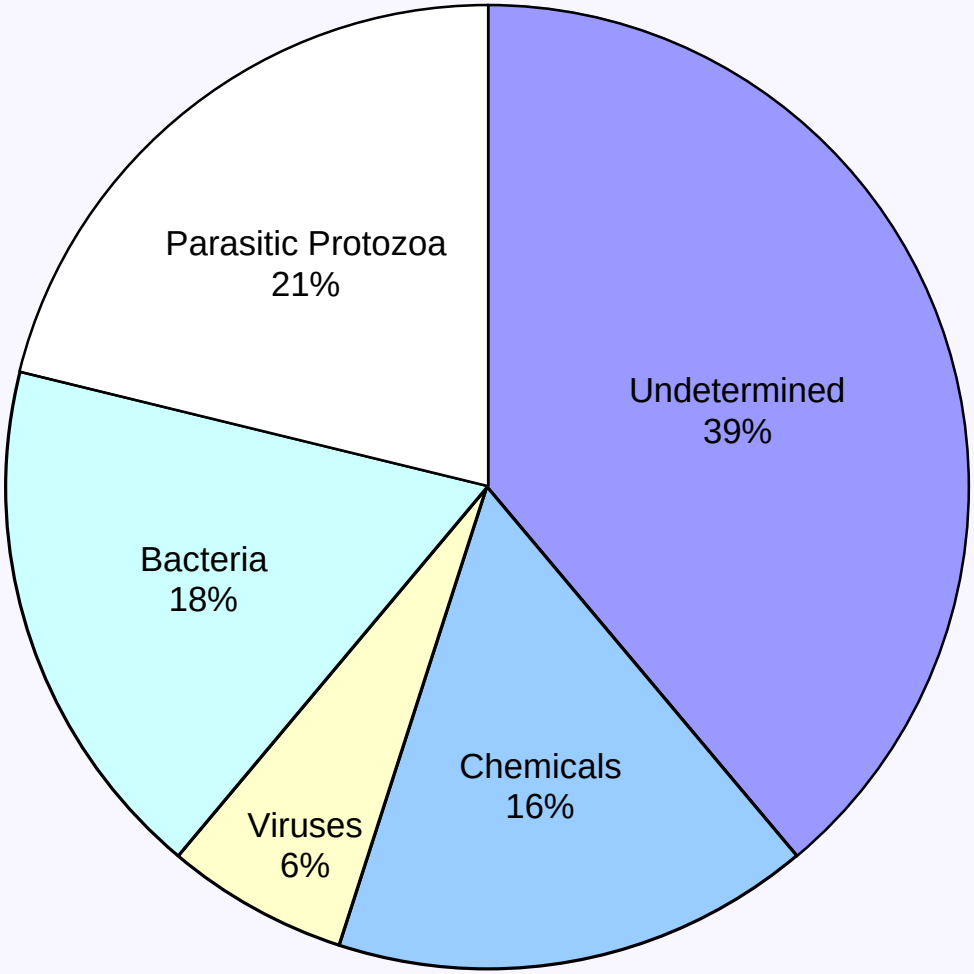
“Status of Microbial Source Tracking and its Implications for Drinking Water Protection”

Yucheng Feng

Department of Agronomy & Soils
Auburn University

Causes of drinking water disease outbreaks in the USA, 1991-2000



Source: *Drinking Water Chlorination: A Review of Disinfection Practices and Issues*

Background

- Preventing drinking water contamination should begin at the source.
- Source water is untreated water from streams, rivers, lakes, or underground aquifers which is used to supply private wells and public drinking water.
- Fecal contamination in surface waters and groundwater
  - 34,005 impaired surface waterbodies in the US
  - Pathogen indicators were the second most common cause of impairment.
  - Contaminated well water
- Potential microbial contaminants in source water: viruses, bacteria, and protozoa.
- Fecal indicator organisms: fecal coliforms, *E. coli,* and entercocci

What is *E. coli*?

- *Escherichia coli* is a Gram-negative, facultatively aerobic, nonspore-forming, rod shaped bacterium.
- One type of fecal coliforms
- Normal inhabitants of the intestinal tract of warm-blooded animals
- Many different types (strain or subtype)
- Most are benign; a few associated with human diseases
- An indicator of fecal contamination in soil and water
- Potential sources: agriculture, wildlife, leaky septic tanks, sewer overflow, etc.

What is Microbial Source Tracking (MST)?

- MST (BST) is new technology being developed to determine the origins of enteric microorganisms from environmental samples.

- Accurate source identification is essential to develop effective pollution control strategies.
- Various approaches have been used to identify fecal sources in waters.
- Categories:
  - Qualitative vs. quantitative
  - Library dependent vs. library independent
  - Phenotypic vs. genotypic

Overview of MST

- Phenotypic vs. genotypic methods
  - Phenotypic: physiological and biochemical characteristics expressed by organisms
    - Antibiotic resistance analysis (ARA)
    - Carbon utilization profile
  - Genotypic: genetic characteristics
    - Repetitive sequence-based polymerase chain reaction (rep-PCR)
    - Ribotyping
    - Pulsed field gel electrophoresis (PFGE)
    - Host specific genetic markers
- Library dependent vs. independent methods
  - Library dependent methods: ARA, rep-PCR, ribotyping, PFGE
  - Library independent methods: viruses, host specific genetic markers

Logistics of common library-based MST methods

Method	Tested target Organism	Cultivation	Major Costs	Time Required
Antibiotic resistance	*E. coli* Fecal streptococci *Enterococcus* sp.	Individual isolates	Antibiotics 96-well microplates	4-5 days
rep-PCR	*E. coli*	Individual isolates	PCR reagents PCR disposable Gel electrophoresis reagents	1 day
PFGE	*E. coli* *Enterococcus* sp.	Individual isolates	Plug prep reagents Restriction enzymes Gel electrophoresis reagents	2-4 days
Ribotyping	*E. coli* Fecal streptococci *Enterococcus* sp.	Individual isolates	DNA purification reagents Gel electrophoresis reagents Restriction enzymes Hybridization/detection solutions Labeled gene probe	1-3 days

Source: Microbial source tracking guide document (US EPA, 2005)

Which method is the best?

- Method comparison studies
  - SCCWRP (Southern California Coastal Water Research Project) study (2003)
  - USGS study (2004)

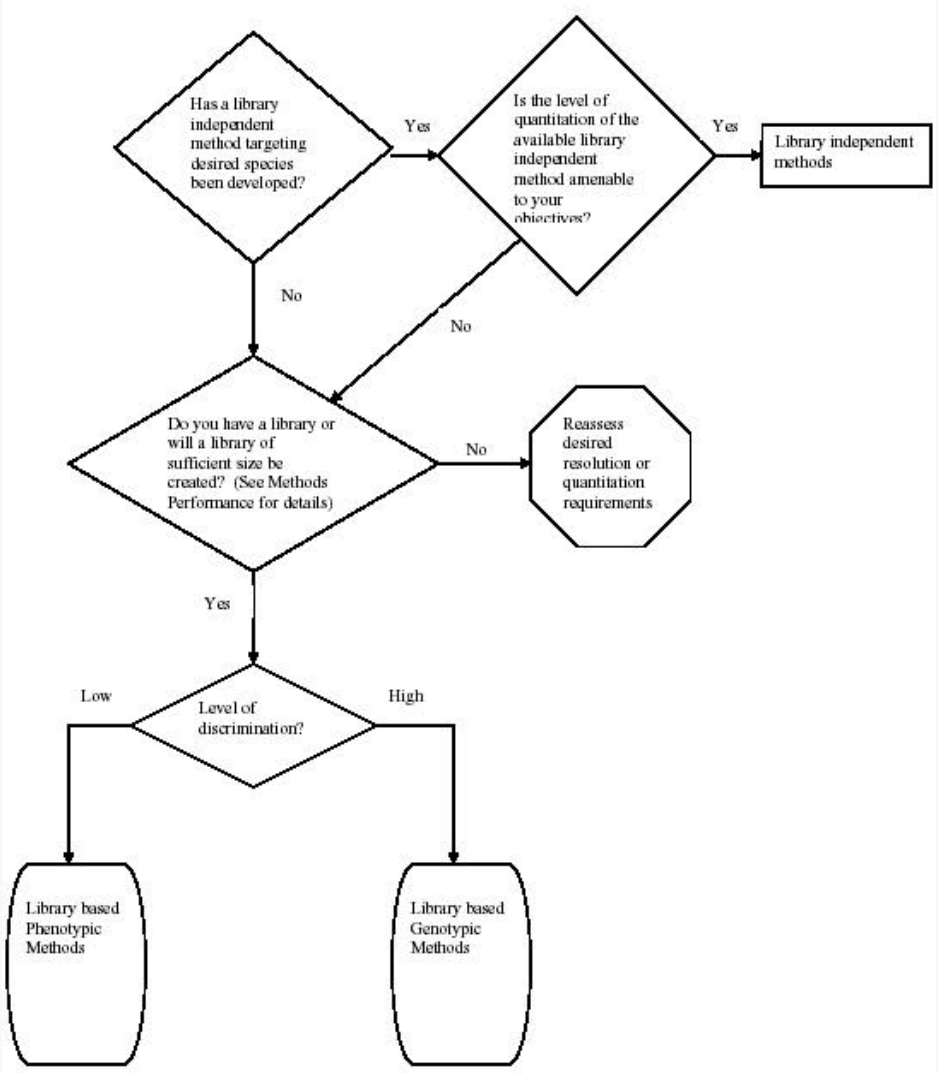
SCCWRP (12 methods, 5 fecal sources)	**USGS** (7 methods, 8 fecal sources)
ARA	ARA
MAR	Carbon utilization profiles
Kirby-Bauer antibiotic susceptibility	Ribotyping with *Hin*dIII
Carbon utilization profiles	Ribotyping with *Eco*RI
PFGE	PFGE
rep-PCR with BOX primer	rep-PCR with BOX primer
Ribotyping	rep-PCR with REP primer
Host specific PCR (Bacteroidetes)	
E. coli toxin genes	
tRFLP	
Human viruses	
F+ specific coliphages	

Which method to use?

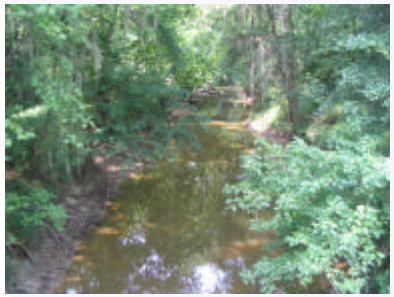
Library dependent + Phenotypic	Library dependent + Genotypic
Library independent + Phenotypic	Library independent + Genotypic

MST Decision Tree (US EPA, 2005)

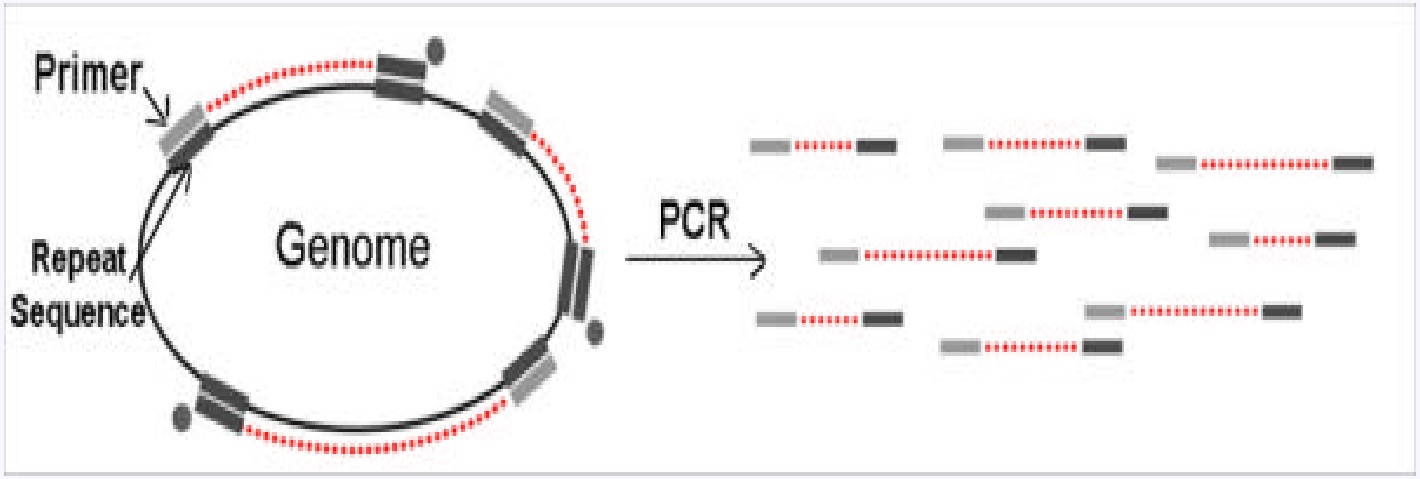
1. Is the fecal pollution problem defined?
2. Has an adequate sanitary survey been conducted?
3. How many sources were identified in the sanitary survey?
   - Proceed if multiple sources
4. Is the watershed/study area of manageable size?
5. What is the desired level of discrimination?
   - Human vs nonhuman
   - Group separation
   - Species specific
   - Specific individual host



Examples of MST

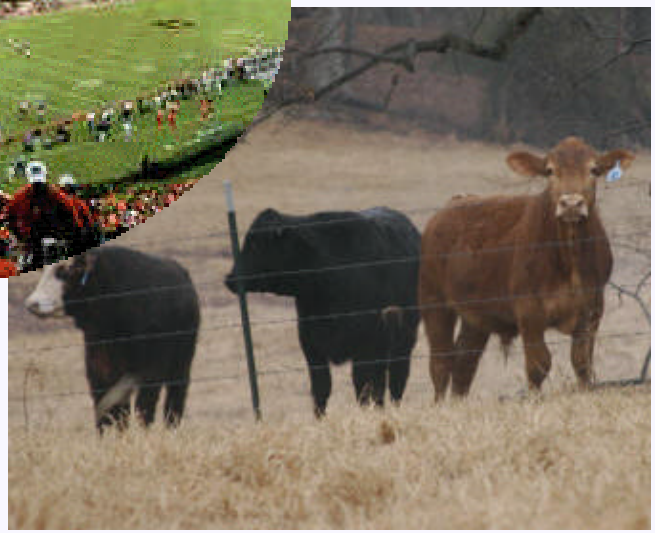


Repetitive sequence-based PCR (rep-PCR)



- *E. coli* genome contains many copies of repetitive DNA sequences. During the polymerase chain reaction, primers bind to these specific repetitive sequences and multiple DNA fragments with various lengths are generated (www.bacbarcodes.com).
- Advantages: simple, rapid, cost-effective, differentiation to the strain level

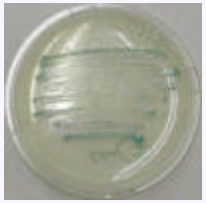
Fecal samples for the known-source library



MacConkey Agar

MacConkey

ChromAgar



MRVP

methyl red

----------> Red color

37°C

Kovac's reagent

1% tryptone ----------> Red ring

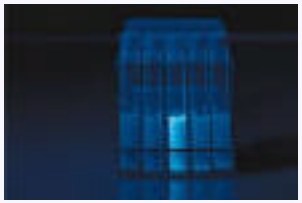
37°C

Citrate agar ----------> No color change

37°C

EC mug ----------> Gas production & fluorescent under UV

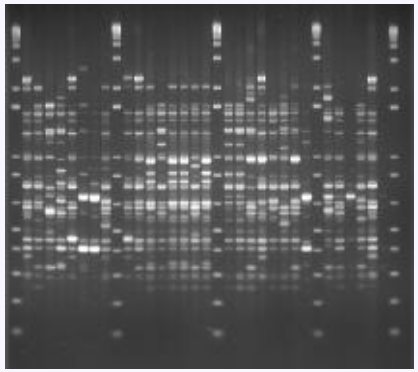
44.5°C



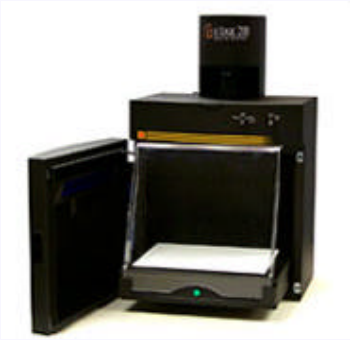
Rep-PCR DNA fingerprinting

E. coli cells mixed with PCR reagents

PCR with Box A1R primer



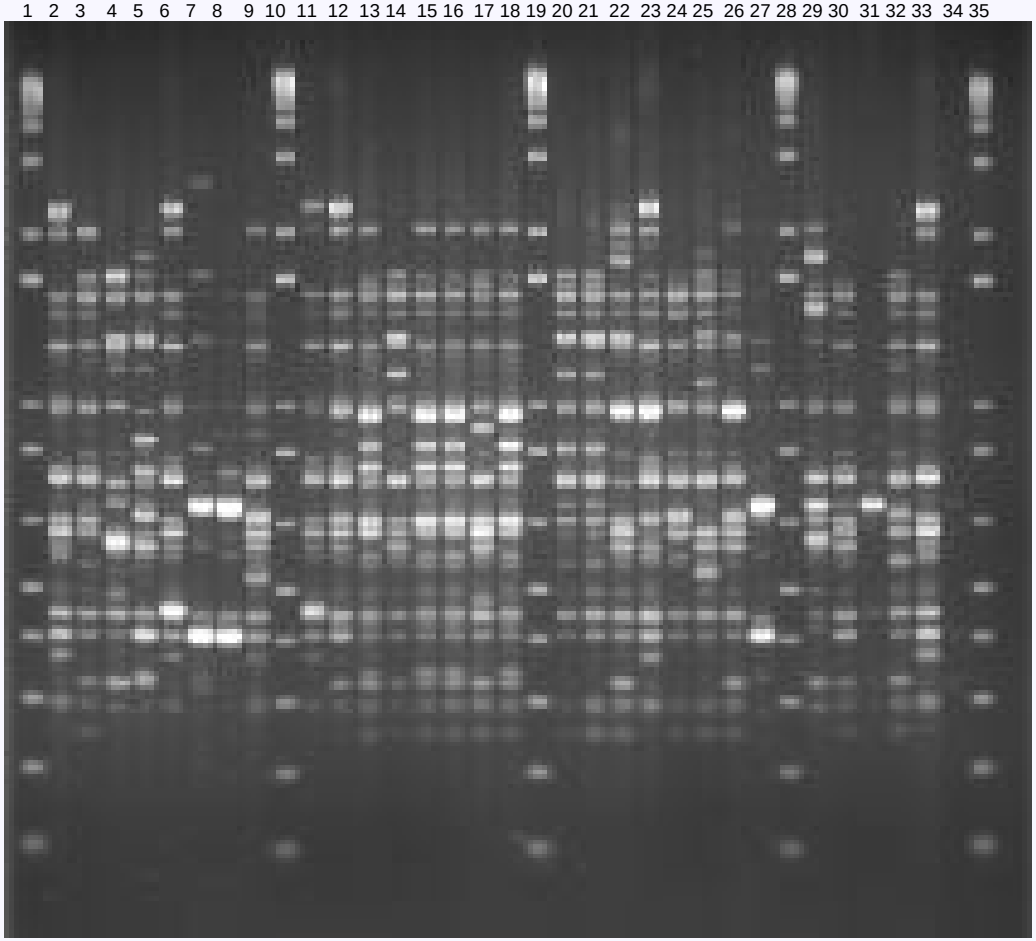
Electrophoresis



Gel imaging

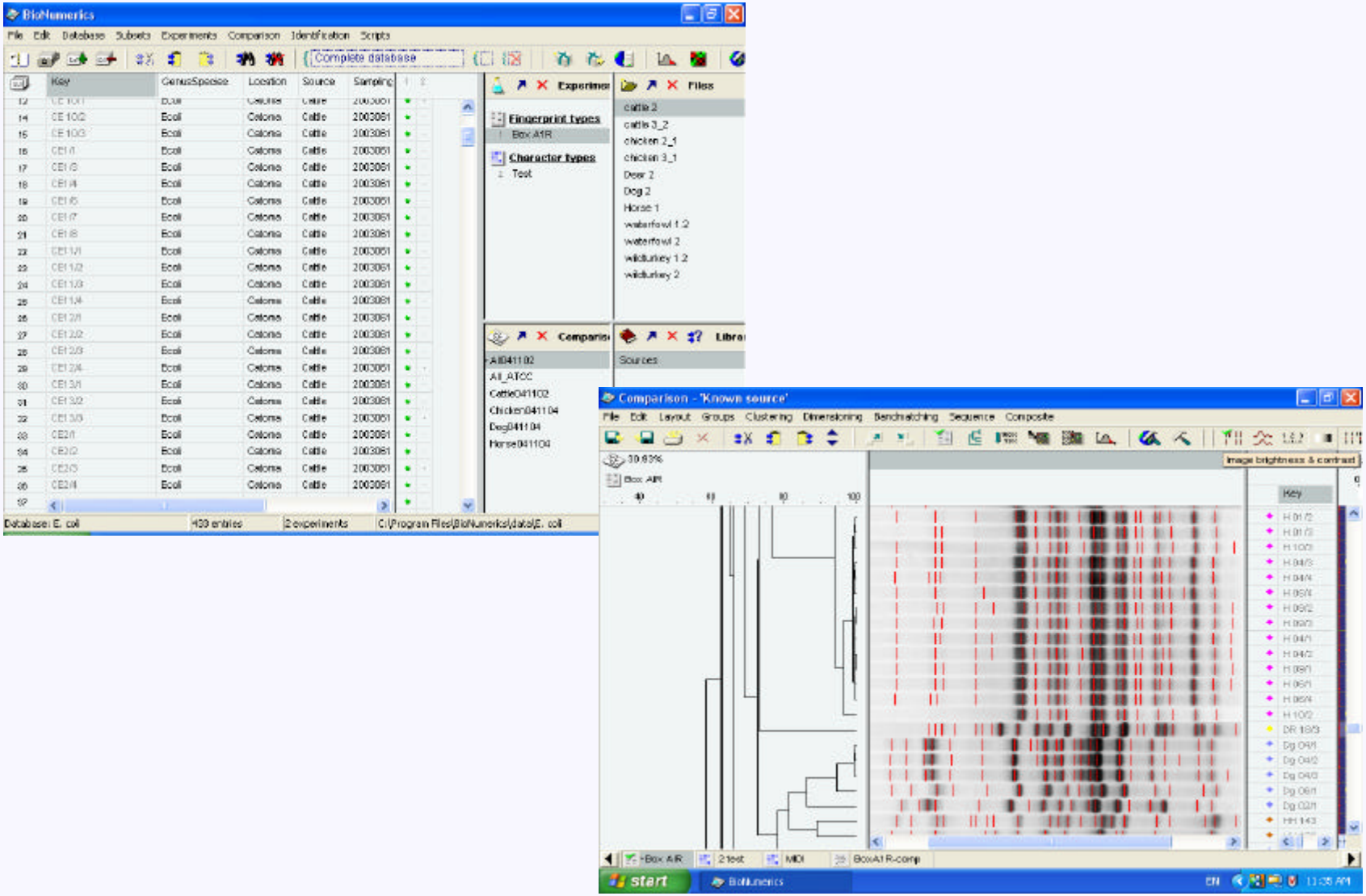
Data analysis

rep-PCR fingerprints generated using the BOX A1R primer



Lanes 1, 10, 19, 18, 28, and 35 contain 1-kb Plus DNA ladder.
Lanes 2, and 33 contain ATCC 25922, and lane 34 is the negative control.
Lanes 3-5 contain dog isolates, lanes 6-9, 11 and 12 chicken, lanes 13-18 horse, lanes 20-25 cattle, and lanes 26, 27, and 29 -32 wild turkey.

Known source library



Source groups and DNA fingerprint patterns in the library

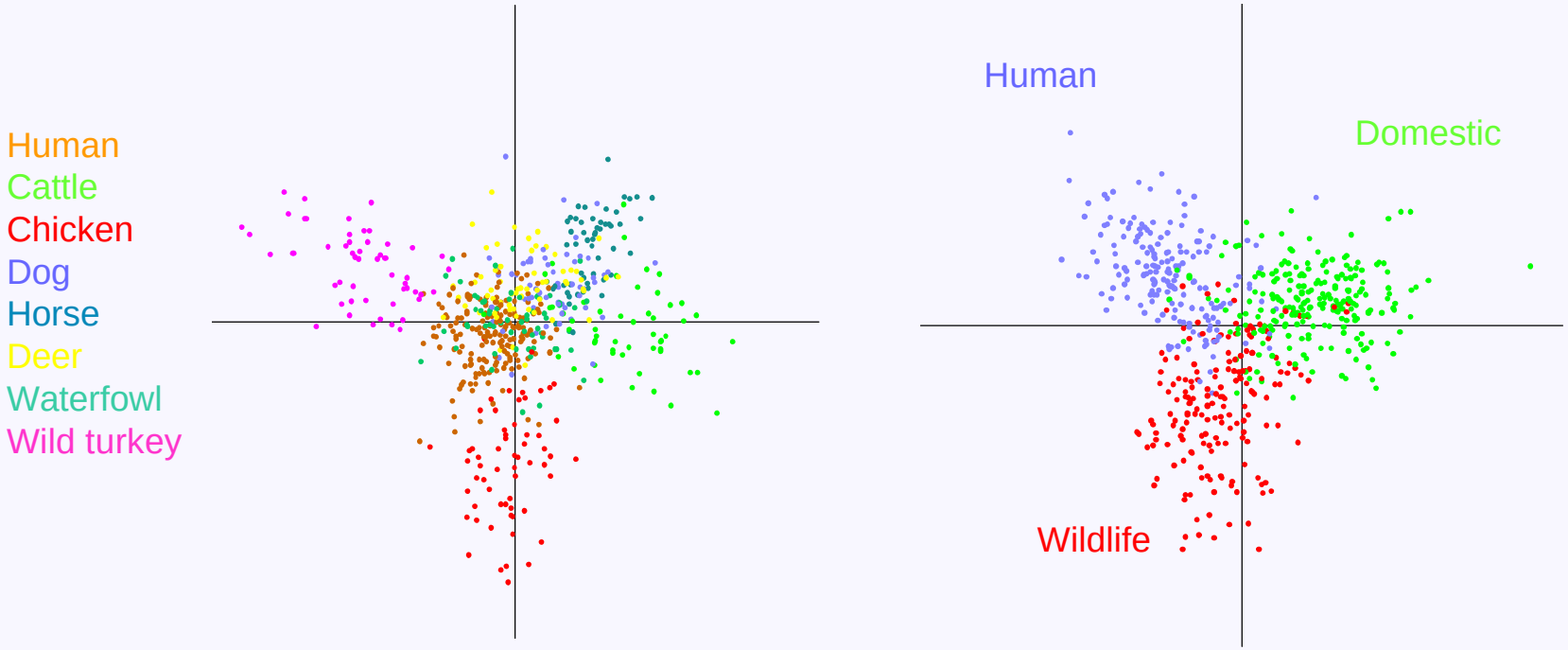
Source group	No. of fecal samples	No. of *E. coli* isolates	No. of total fingerprints	No. of unique fingerprints
Cattle	20	60	53	33
Dog	18	60	60	42
Chicken	20	60	60	28
Horse	20	54	54	33
Deer	28	61	61	46
Wild turkey	12	54	53	25
Waterfowl	21	62	62	40
Human	165	180	179	167
Total	**310**	**591**	**582**	**414**

Rates of correct classification within specific sources after decloning

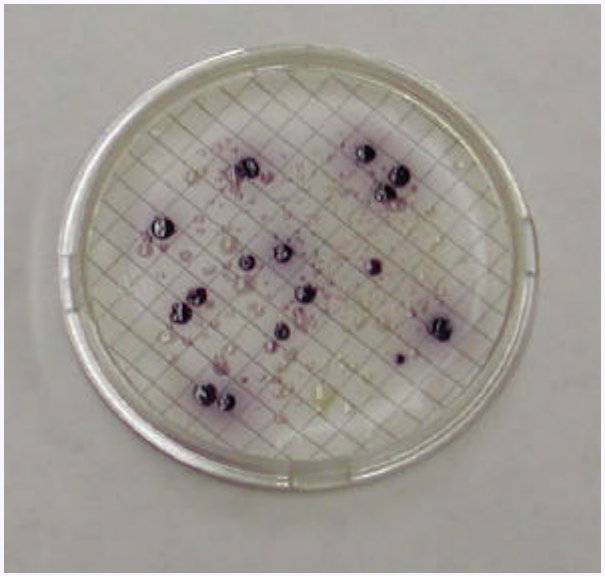
	Cattle	Chicken	Dog	Deer	Horse	Human	WF	WT
Cattle	**61**	3.6	2.4	2.2	0	0	7.7	0
Chicken	3.4	**86**	4.8	2.2	0	3.0	0	0
Dog	3.4	3.6	**50**	4.3	3.0	1.2	2.6	0
Deer	0.4	0	7.1	**70**	0	0	7.7	12
Horse	0.4	0	4.8	2.2	**91**	1.2	0	0
Human	12	3.6	26	8.7	6.1	**92**	10	8
Waterfowl	18	3.6	2. 4	8.7	0	1.2	**72**	12
Wild turkey	0.4	0	2.4	2.2	0	1.8	0	**68**

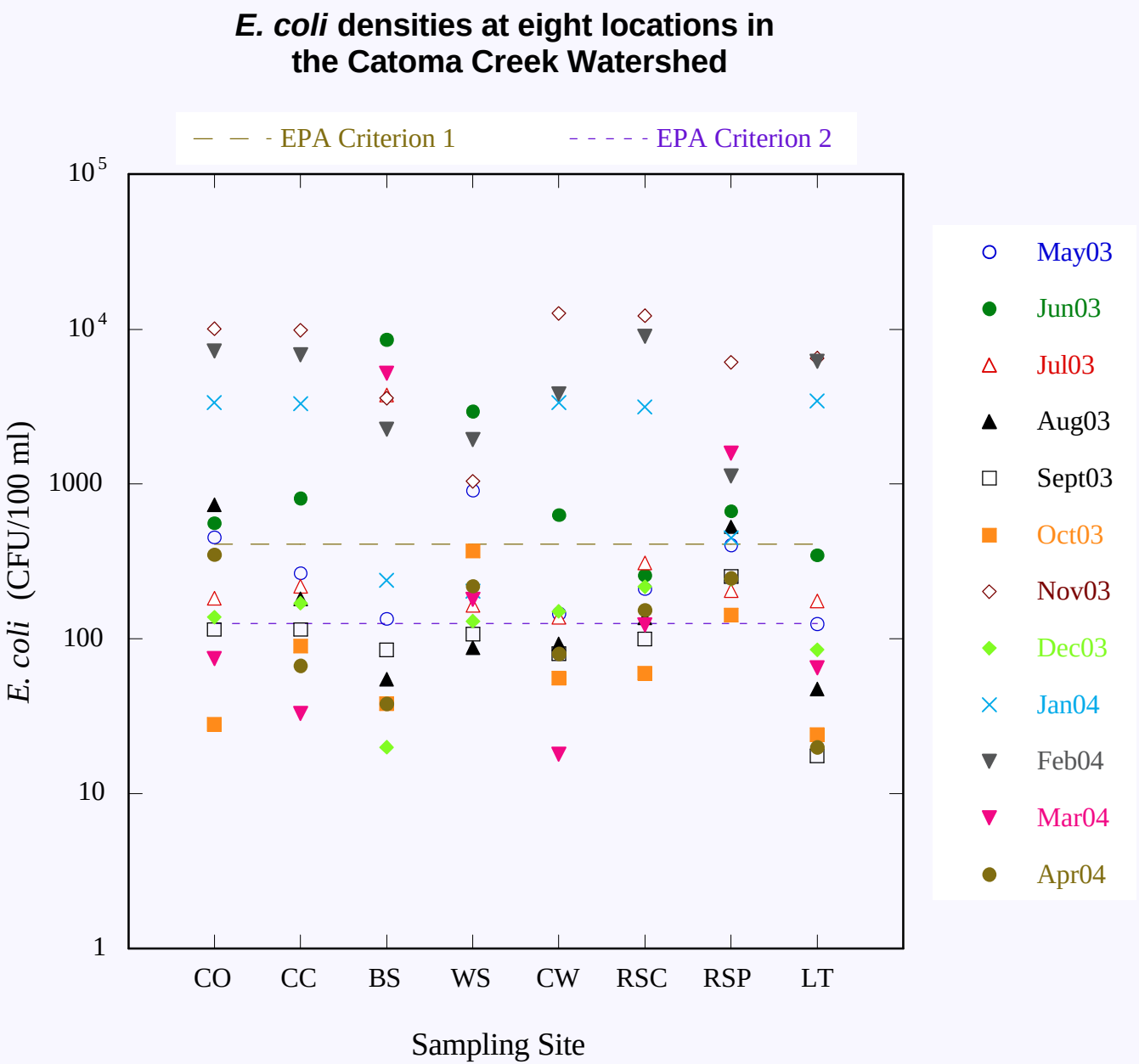
ARCC = 74%

MANOVA plot of *E. coli* rep-PCR DNA fingerprints from different source groups: eight vs. three source groups

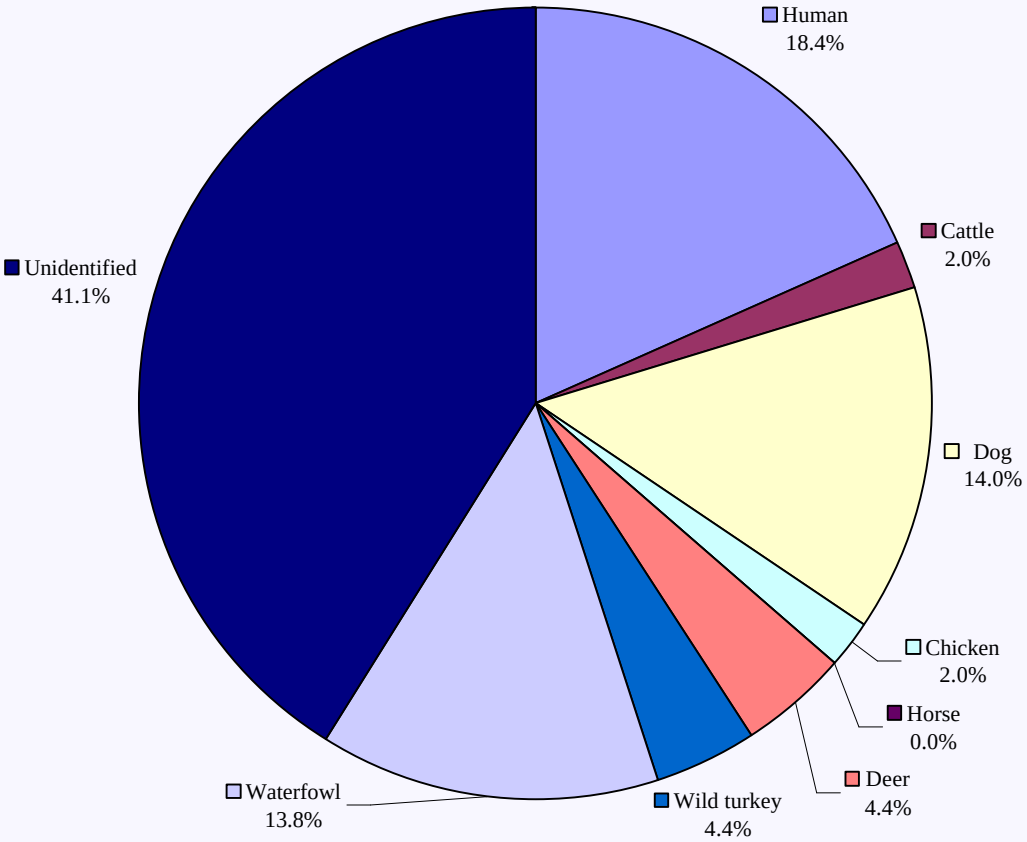


Water samples

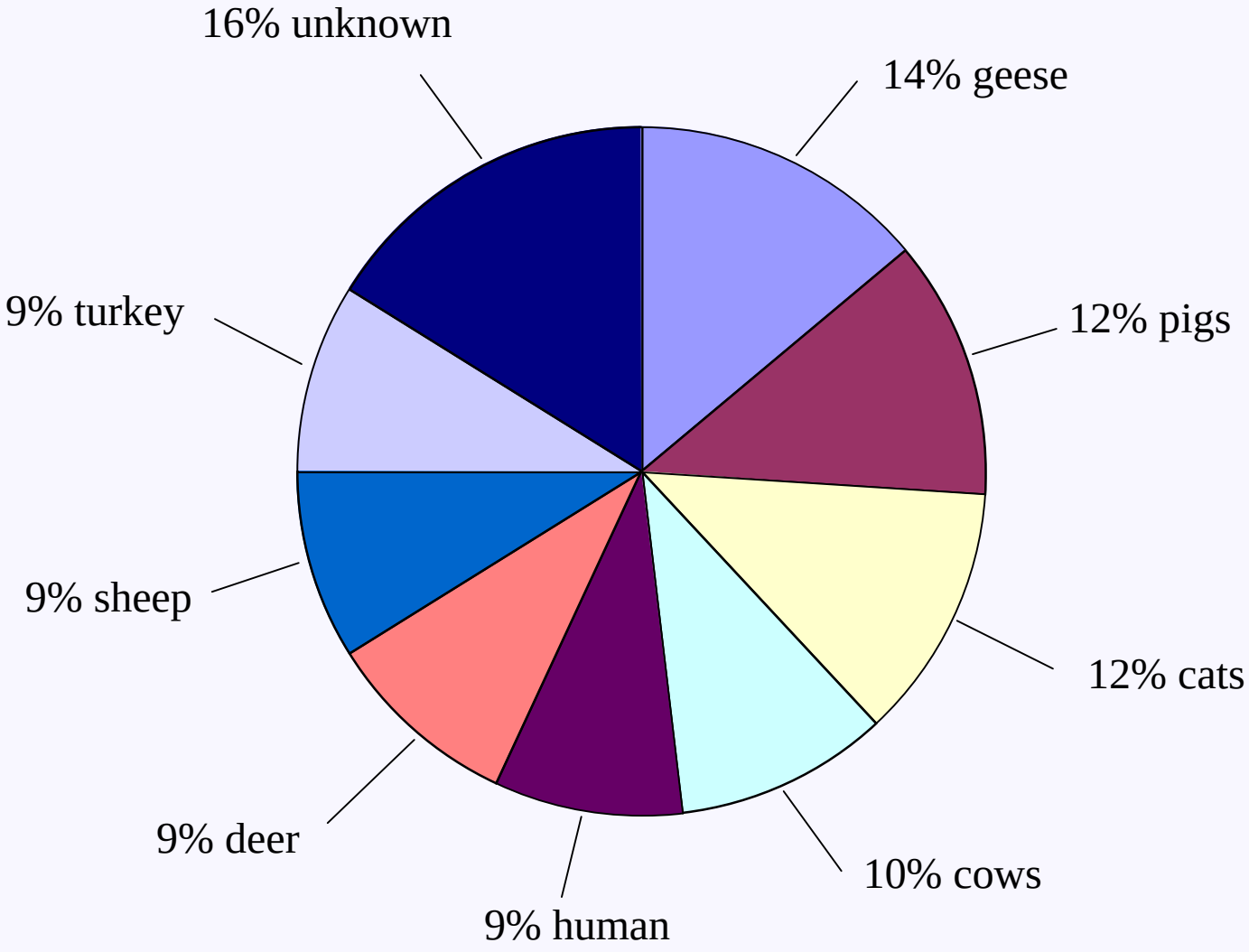



E. coli densities at eight locations in the Catoma Creek Watershed
EPA Criterion 1
EPA Criterion 2
10^5
10^4
1000
100
10
1
E. coli (CFU/100 ml)
CO
CC
BS
WS
CW
RSC
RSP
LT
Sampling Site
May03
Jun03
Jul03
Aug03
Sept03
Oct03
Nov03
Dec03
Jan04
Feb04
Mar04
Apr04

Source distribution of water isolates in the Catoma Creek watershed



Source distribution of *E. coli* in a Minnesota watershed



Microbial source tracking guide document, EPA 2005

QA/QC in MST

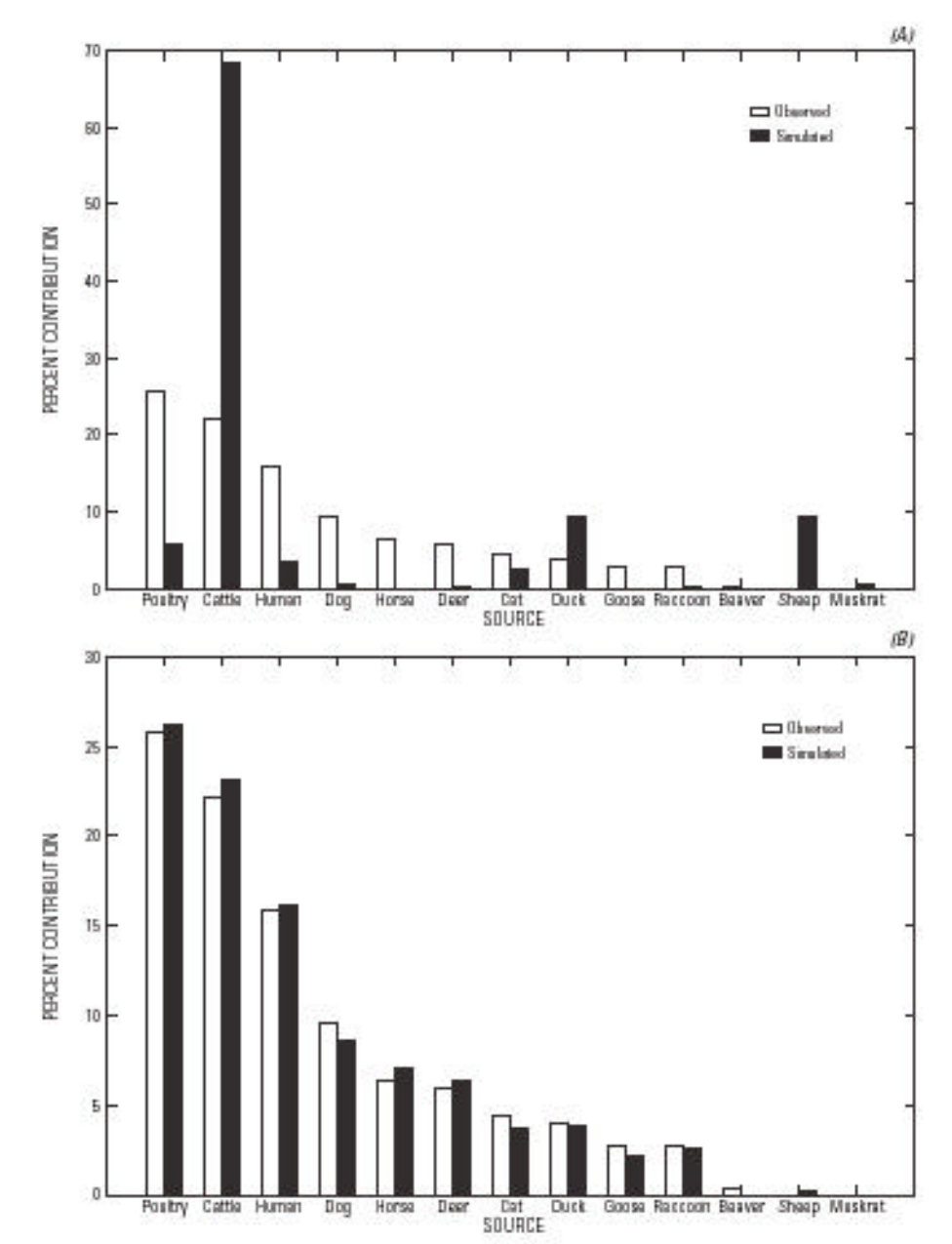
- Precision (reproducibility): replicate at least 10% of the samples
- Control samples: Positive controls, negative controls
- Performance of library: rate of correct classification
- Use more than one approach
  - Perform analyses of chemicals that are indicative of sources (caffeine, cotinine, arsenic)

TMDL Process

1. Identify quality limited waters: 303(d) Use Impairment Lists
2. Establish priority waters/watersheds
3. Develop TMDL:
   - Problem statement - cause of impairment
   - Numeric target(s) - water quality standards
   - Source analysis - relative contributions of pollutant sources
   - Loading capacity estimate
   - Allocations - wasteload allocations to point sources and nonpoint sources

TMDL = Point source loadings + Nonpoint source loadings
+ Margin of safety

TMDL at a Virginia watershed



Before

After

Source: Hyper and Moyer (2005) JAWRA

Summary

- Safe drinking water begins with source water protection.
- MST is a valuable tool for watershed management and TMDL development.
- MST techniques are still under development.
- Needs for further research:
  - Other source identifiers: *Bacteroides*
  - Library size
  - Geographic variability
  - Temporal variability
  - Search for better library-independent methods
  - Host specificity of microorganisms
  - Genetic change associated with habitat change