Scenario F – An E. coli Outbreak

Background:

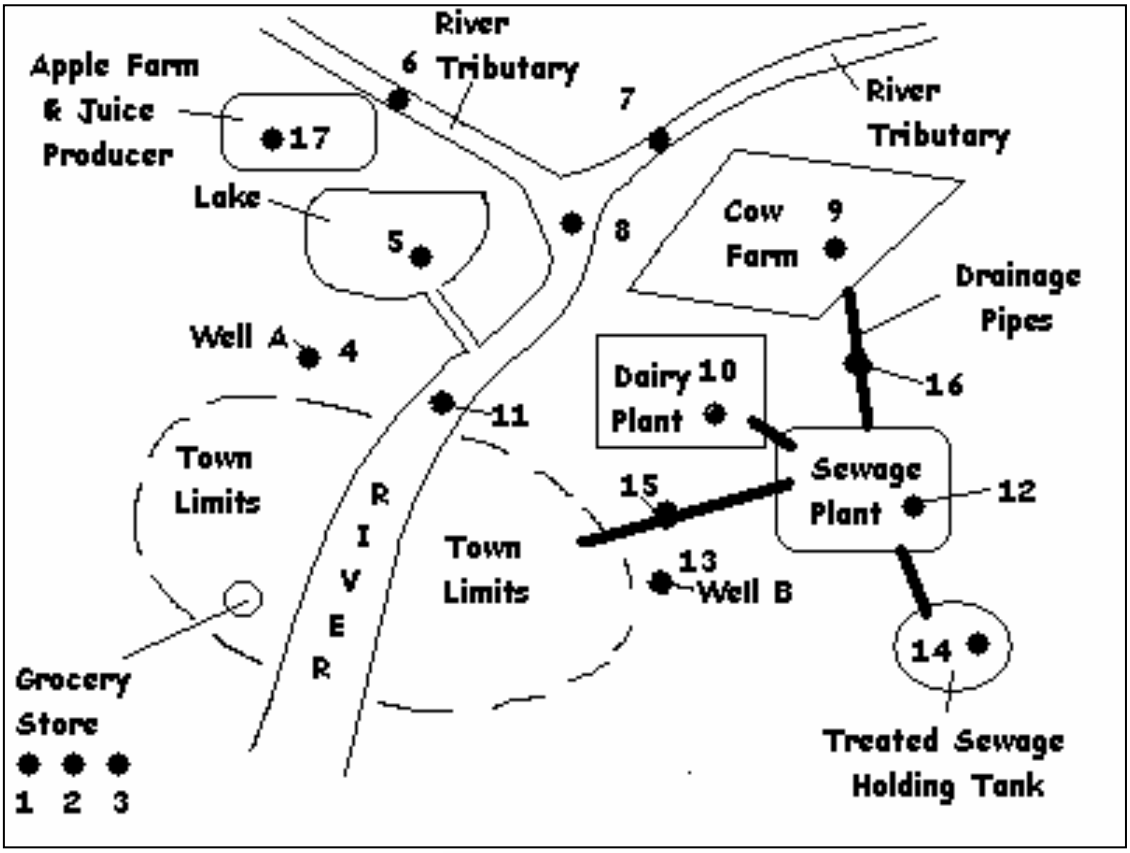
In a small rural town, a group of people – mostly young adults and children – have come down with a case of E. coli poisoning. Their hypothesis is that the culprit could be the dangerous and sometimes deadly E. coli strain, O157:H7 which has contaminated the food or water. No one is really sure, but preliminary tests do show it to be some type of E. coli that is making people sick. No one has died yet, but over twenty people are very ill. Your group of scientists has been assigned to go to the town and test all possible sources of the outbreak and answer these questions. What is the strain of E. coli that is causing the sickness? Where is it coming from? How is it getting there? And how are the people being infected?

You have decided to use GeneChip microarray technology and quickly get in communication with the manufacturing company. This company has the ability to custom order the array for you. You order a resequencing array that has the ability to resequence DNA in a sample that comes from specific E. coli strains. There are hundreds of strains of E. coli besides O157:H7, but you decide to resequence only the 6 most common strains that cause food and water borne illnesses, including the 0157:H7 strain.

When you arrive in town, you are given a map of the area which identifies possible areas of contamination as well as possible sources of contamination. You decide to take a total of 17 samples from various areas around the town. Here are some facts about the town that could be helpful:

- there is a nearby river that flows right through the town
- there are two wells that the town draws most of its public water from
- there was a huge storm and flooding a few months ago & the areas north of the town (near the farms and lakes) were particularly affected
- there is a huge farm northeast of the town that has a large herd of cows
- there is a milk processing plant also northeast of the town
- there is an apple farm that produces its' own juice northwest of the town
- there is a lake that people commonly go to swim during the summer months just north of the town
- there is a sewage treatment plant that takes sewage from the town and cow farm and treats the raw sewage before moving it to an outdoor percolation pond just east of the town
- people commonly swim in both the lake and the river during the summer months (it now coming to the end of summer)
- the pipes in and outside of the town are getting very old
- the cow farm and milk plant provide most of the meat and milk consumed by the people of the town
- the town has one main grocery store where the people get most of their food (and all of the meat, milk, and apple juice come from the local producers)

Here is a map that has all the above areas and the spots of the seventeen samples you decide to test (marked with a spot and number).



● = sample point

Here what each sample was taken from:

#1 – locally processed Apple Juice from store
#2 – locally processed Milk from store
#3 – locally processed meat from store
#4 – water from well A
#5 – water from lake (common swimming and boating area)
#6 – water from upper river tributary
#7 – water from upper river tributary
#8 – water from river after tributaries meet
#9 – cow feces from cow farm
#10 – cow feces from dairy plant
#11 – water from river after lake tributary (common swim area for town people)
#12 – raw sewage from treatment plant
#13 – water from well B
#14 - water from sewage holding tank (post treatment)
#15 – water from drainage pipe taking sewage water from town to treatment plant
#16 – water from drainage pipe taking sewage from cow farm to treatment plant
#17 – samples from soil taken at the farm

The GeneChip microarray you are using sequences any E. coli DNA found in the samples. It is a resequencing array that outputs a unique section of each strain. Remember, only if DNA from E. coli strains that you are sequencing for is found in the samples taken will you view results. The probes are specific to sequences in these strains. A blank result means there is no E. coli from the 6 strains you chose in the sample at all.

The E. coli genome is 1,725,750 bases in length. You have decided to look for a 20 base pair segment that is highly variable among E. coli strains and unique for each of the six strains at bases 1,725,600 to 1,725,620. Here are the 20 base pair sequences for each strain along with a bit of information on each.

JM 101

- harmless to humans
- commonly found contaminating dairy products
- unique sequence: TCCTGCAAACCTATTACAAC

RR1

- harmless to humans
- commonly found in water from lakes and rivers and other recreational areas
- unique sequence: GACTGCCCACAGATTACAAC

O157

- dangerous and potentially deadly to humans
- causes intestinal strong problems and leads to very high fevers
- normally found contaminating raw meat, unprocessed dairy and apple juice
- unique sequence: TCCTGCAGGCAGATTACAAC

0111

- non-dangerous but does cause some sickness in humans
- causes intestinal problems and slight fevers
- commonly found in intestinal tract of cows
- unique sequence: TCCTGCAGGCAGATTAGTTC

Q358

- harmless to humans
- found in the gut of humans, but aides in normal digestion of our food
- unique sequence: TCCTGCACCCAGAGGACAAC

Y1088

- non-dangerous to humans but does lead to some severe illness (intestinal cramps, fever)
- commonly found in moist soil, but known to also be water borne
- important part of natural soil ecosystem
- unique sequence: TCCTGCAGGGTCTTTACAAC

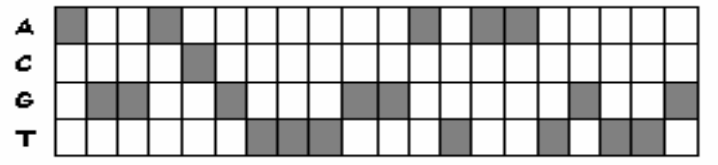
Results:

Here are the results of each of the seventeen samples:

Sample #1 - No E. coli found (blank output)

Sample #2 – Some E. coli found

Output:

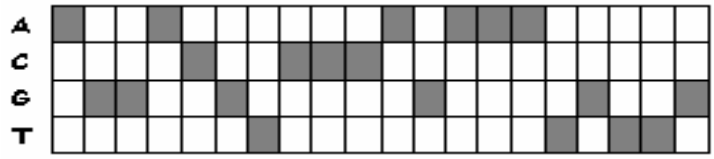


Sample #3 – No E. coli found (blank output)

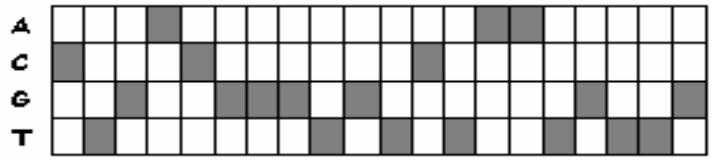
Sample #4 – No E. Coli found (blank output)

Sample #5 – Some E. coli found

Output #1:



Output #2:



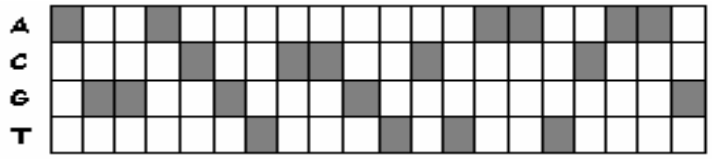
Sample #6 – No E. coli found (blank output)

Sample #7 – No E. coli found (blank output)

Sample # 8 – No E. Coli found (blank output)

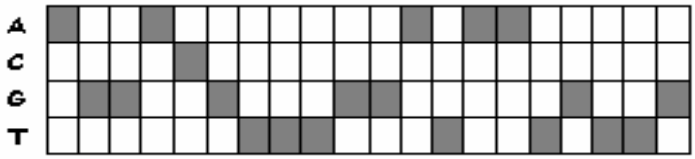
Sample # 9 – Some E. coli found

Output:



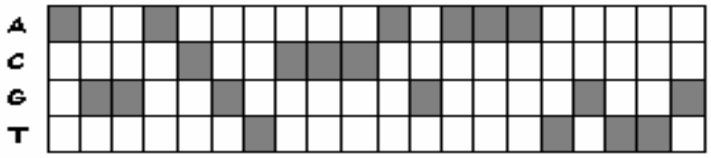
Sample #10 – Some E. coli found

Output:

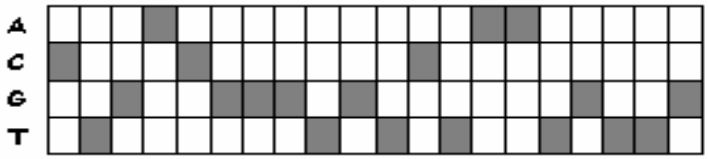


Sample #11 – Some E. coli found

Output #1:

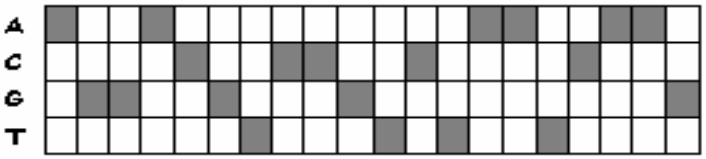


Output #2:

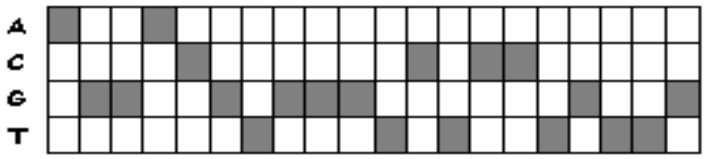


Sample #12 – Some E. coli found

Output #1:

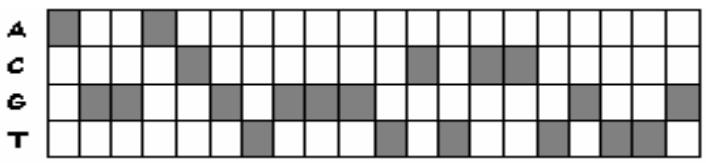


Output #2:



Sample #13 – Some E. Coli found

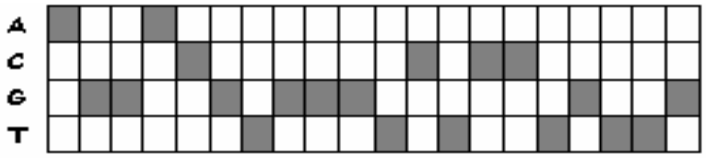
Output:



Sample #14 – No E. coli found (blank output)

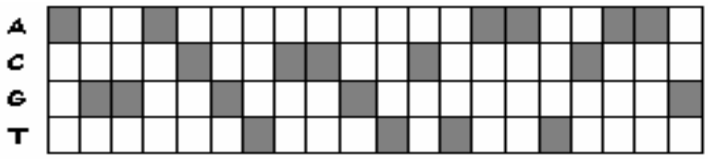
Sample #15 – Some E. coli found

Output:



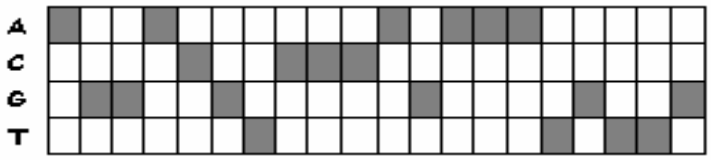
Sample #16 – Some E. Coli found

Output:



Sample #17 – Some E. Coli found

Output:



Note: If two different outputs occur, then there are two different E. coli strains in the sample tested.

Directions:

You now have all your data. Your job is to determine which E. coli strains are found where and connect them to the 20 sick people. Be sure to refer back to the facts about the town as well as the map of the area around the town. Also, review the facts about each of the 6 strains of E. coli. Good luck! Why are the people getting sick and from what? The townsfolk are depending on you!