Analysis of *E. coli* DNA from Selected Source(s) using Pulsed-Field Gel Electrophoresis

Thesis submitted to
The Graduate College of
Marshall University

In partial fulfillment of the
Requirements for the Degree of
Master of Science
Forensic Science

by

© Sheema Doshi

Marshall University

Huntington, West Virginia

2001

This thesis was accepted on April 30 2001
Month Day Year

as meeting the research requirements for the master's degree.

Advisor Terry W. Fenger

Department of FORENSIC SCIENCE

Leonard J. Deutsch
Dean of the Graduate College

ABSTRACT

Doshi, Sheema. 2001. Analysis of *E. coli* DNA from Selected Source(s) using Pulsed-Field Gel Electrophoresis. Unpublished master's thesis, Marshall University, Huntington, West Virginia.

The presence of fecal coliforms in surface waters is a public health issue throughout the United States and Canada. The presence of *Escherichia coli* (*E. coli*) indicates recent fecal contamination from either animal or human sources. To solve this growing problem one must identify the source(s) of such contamination by creating a database of DNA profiles of *E. coli* isolates. Hence, tracking the source(s) of *E. coli* can help link water contamination to animal source(s) of contamination. In this study, pulsed field gel electrophoresis (PFGE) was used to analyze DNA from *E. coli* isolates from known animals. For PFGE, the bacterial DNA was embedded in agarose plugs (to prevent shearing of DNA) and then cut with a restriction endonuclease (*Xba*I). The digested DNA was resolved on an agarose matrix to yield PFGE profiles. Unique DNA isolates that had been previously cut and resolved with *Xba*I were then cut with a second restriction endonuclease, *Not*I. Different DNA profiles of chromosomal fragments were generated for different animal sources. The following aspects, which are fundamentals of this study, were investigated to identify possible sources of contamination. Similar DNA profiles of *E. coli* were statistically analyzed to determine good identifier regions for rapid analysis. Four of the six established regions of a DNA profile were considered to be good identifier regions. It was determined that of the two restriction enzymes used, *Xba*I was a better enzyme to use to establish a database, since it allowed for better differentiation of strains within animals. Thus, it was possible to differentiate between animal types and individuals of the same animal in this database.

TABLE OF CONTENTS

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Special thanks to Dr. Terry Fenger, Dr. Charles Somerville, and Dr. Hong-Wei Yu who have provided an incredible amount of support and advice as thesis committee members.

Appreciation is given to the Department of Agriculture of West Virginia and the Department of Forensic Sciences of Marshall University which funded this research and aided greatly in obtaining samples.

Special thanks is given to Nizar Noureddine whose help has been invaluable and without whom this research would never have started. He and his wife showed me so many kindnesses during the time I spent at Marshall.

Thanks to all who work with the Potomac River Project. Special thanks to Krista Haught who helped me over the past year by creating a friendly and professional working area.

Thanks to Lenna Chambers and Joseph Oakley for making the summer bearable.

Thanks to Nathan Head, a friend, who helped me with certain difficulties that I had with some computer work.

Appreciation to all who work in Dr. Simmons Jr.'s Laboratory at Virginia Tech. Special thanks to Dr. Simmons Jr., Sue Herbein and Sharon Myers for taking the time to answer all of my incessant questions.

Special thanks to Dr. Monica Valentovic, who was always willing to share her time with the statistical analysis of my data.

Karma Outhouse, is my roommate and good friend. She has helped to ease my frustrations and supported me throughout my research.

Appreciation is given to my family for their support and help during the time that I spent at Marshall.

Sheema Doshi

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

CHAPTER I.

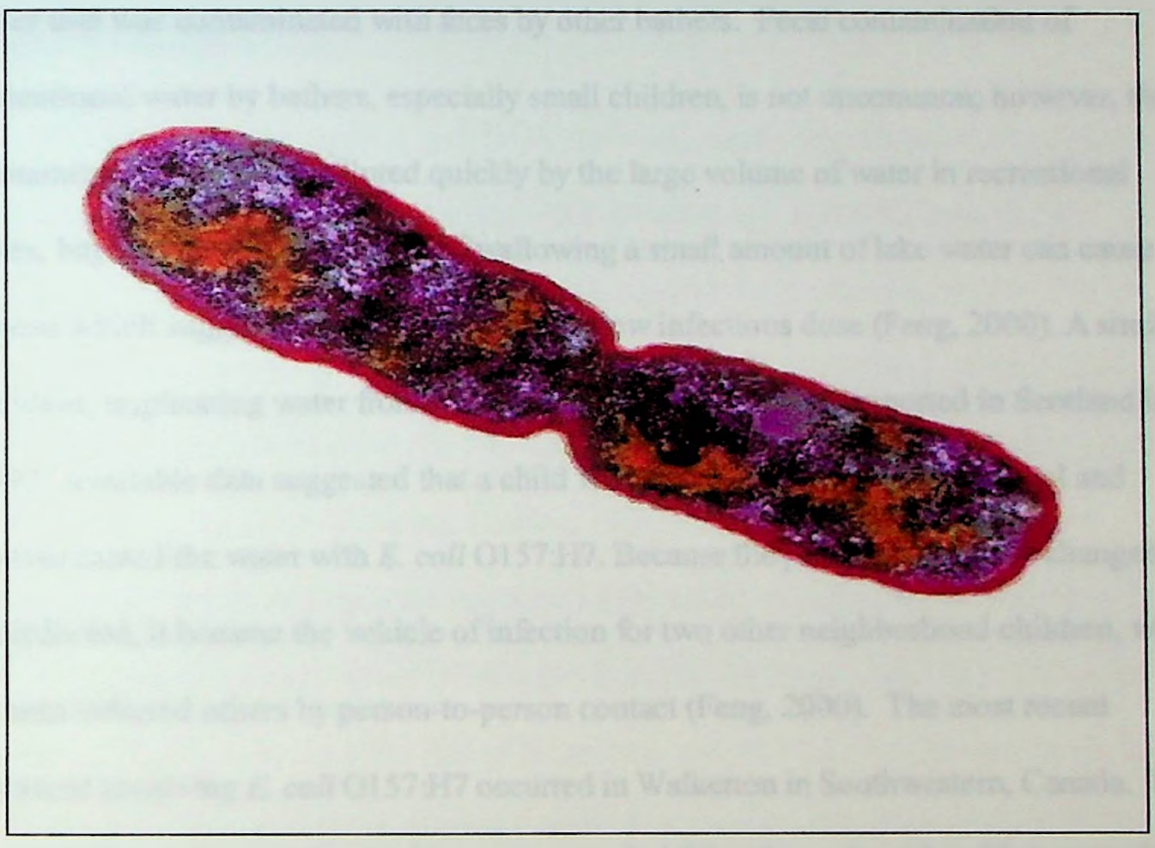
INTRODUCTION

BACKGROUND

E. coli was chosen for the purpose of generating a DNA database because it is normally found in the intestine and feces of warm-blooded animals. *E. coli* is a part of the Enterobacteriacae family and is a commonly used fecal coliform indicator (Figure 1). Coliform is defined as aerobic, facultative anaerobic, gram negative, non-spore forming, rod-shaped bacteria that ferments lactose with acid and gas formation (Ottaway, 2000). Demonstrating the presence of *E. coli* represents confirmation of fecal contamination (Micro Consultants, 1998). Since *E. coli* is a good indicator of fecal contamination, it can be used to track sources of feces that contribute to a variety of diseases in drinking water (hepatitis, cholera, typhoid fever, traveller's diarrhea) thereby, causing the water to be unsafe (Micro Consultants, 1998). In addition, *E. coli* can be tested and analyzed easily. It has been known to generate characteristic mixed acids and gases as end-products of metabolism which allows the bacteria to be selected for using a variety of presumptive tests (Micro Consultants, 1998). Also, this bacteria can be analyzed with PFGE, which has been shown to be one of the most sensitive and discriminatory methods of analysis of *E. coli* (Maslow et al., 1993).

Over the past 10 years, *E. coli* has been reported in the media as a pathogen: *E. coli* O157:H7. This pathogenic form of *E. coli* has been found to contaminate sources of drinking water, recreational lakes and rivers. For instance, the first and largest waterborne outbreak associated with this pathogen occurred in Missouri in 1989 (Feng, 2000). Two hundred and forty people were infected, thirty two hospitalized, and four died. The source of the outbreak was not identified, but backflow during a water main

Fig. 1. This figure represents a colorized transmission electron micrograph (TEM) of an *E. coli* that is dividing. TEM magnification is 92, 750x. (Kunkel, 1998-1999).



break might have contaminated the drinking water supply (Feng, 2000). In 1991, an outbreak involving *E. coli* O157:H7 occurred in recreational lake water close to Portland, Oregon. Of the 59 people affected, 21 (all children) expressed disease symptoms. Transmission of the bacteria probably occurred when the swimmers swallowed lake water that was contaminated with feces by other bathers. Fecal contamination of recreational water by bathers, especially small children, is not uncommon; however, the contaminants are usually diluted quickly by the large volume of water in recreational lakes, bays, or rivers (Feng, 2000). Swallowing a small amount of lake water can cause illness which suggest that the pathogen has a low infectious dose (Feng, 2000). A similar incident, implicating water from a children's paddling pool, was reported in Scotland in 1992. Available data suggested that a child with diarrhea had played in the pool and contaminated the water with *E. coli* O157:H7. Because the pool water was not changed or disinfected, it became the vehicle of infection for two other neighborhood children, who in turn infected others by person-to-person contact (Feng, 2000). The most recent incident involving *E. coli* O157:H7 occurred in Walkerton in Southwestern, Canada. The bacteria had infiltrated the drinking water supply killing 9 people and making more than 700 people ill (Kondro, 2000). The source of the bacterial contamination in Walkerton has not yet been identified. *E. coli* O157:H7 leads to hemolytic uremic sydrome (HUS), a leading cause of acute kidney failure in young children, the elderly, and immunosuppressed individuals.

A method of identifying sources of *E. coli* is by using its DNA profiles. Besides PFGE, there are a number of techniques and approaches that could have been used to

analyze *E. coli* DNA. Barret et al, used phage typing to identify *E. coli* O157:H7 (1994). Wiggins et al., used antibiotic resistance to identify nonpoint sources of fecal pollution (1999) and Kariuki et al., analyzed *E. coli* strains phenotypically (1999). Tynnkkynen et al used ribotyping, and randomly applied polymorphic DNA, to type selected bacterial strains (1999). All of these methods could have been used for DNA analysis. PFGE was chosen for several reasons.

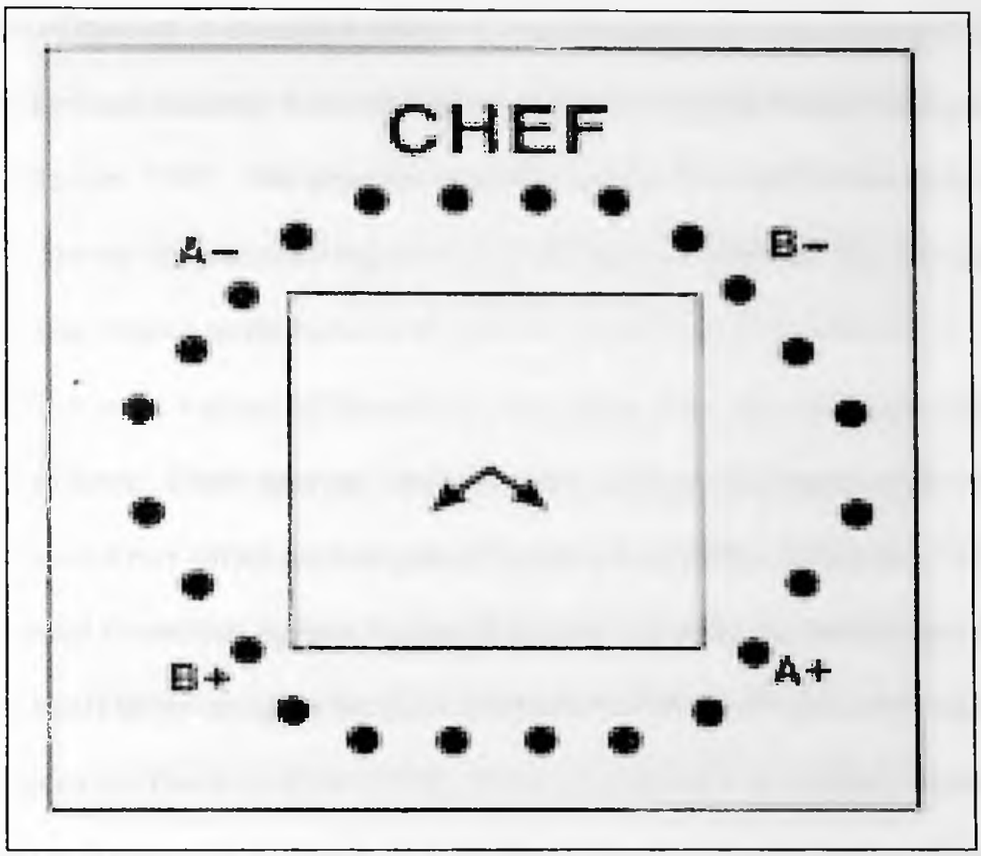
With PFGE, chromosomal *E. coli* DNA can be digested with restriction enzymes that have few restriction sites. In this study, *Xba*I (Roche) and *Not*I (Promega Corp.) restriction endonucleases were used. *Xba*I recognizes the restriction site T^CTAGA generating approximately 20 chromosomal fragments (Roche Catalog, 1999, p.150). and *Not*I recognizes the restriction site GC^GGCGCCGC generating approximately 12-15 chromosomal fragments (Promega Catalog, 1999, p.9.60). PFGE typically shows distinct, well-resolved fragments representing the entire bacterial chromosome in a single gel and allows for resolution of DNA larger than 100, 000 base pairs (bp). This is an advantage for *E. coli*, since the size of its genome is 4, 639, 221 bp (Blattner et al., 2000).

Dr. George Simmons Jr.'s Laboratory of the Department of Biology at Virginia Tech had conducted a similar analysis of *E. coli* DNA using PFGE over the past 10 years. Their samples originated from wildlife sources of selected study sites along the Eastern Shore of the Chesapeake Bay (the tailwaters of the Potomac River).

Another reason PFGE was decided upon was its use by the Centers for Disease Control (C.D.C.) to establish the PulseNet (FoodNet) DNA database. It was anticipated that DNA profiles of *E. coli* generated in this study could be compared with DNA profiles from the C.D.C. database to determine any commonalities or differences.

E. coli is a prokaryote whose chromosome is a single circular molecule of DNA, which resides in the cytoplasm (C.Kennedy and L.S. Pierson III, 1999). The supercoiled (twisted) DNA is packed inside the *E. coli* as a semi-orderly aggregate called a nucleoid. DNA is negatively charged due to its negatively charged phosphate groups. During the process of agarose gel electrophoresis, DNA is constantly pulled towards the positive charged electrode through an agarose matrix (Goering, 1993). Agarose gel electrophoresis is a unidirectional process, but PFGE is multidirectional. In PFGE, the location of the positive charged electrode continuously changes, thereby, causing the DNA molecules to continuously re-orient their direction of migration through the agarose gel (Goering, 1993). In addition, the length of the time of the electrical pulses changes the size range of separation (Joppa et al., 2000). For example, at 5.4 V/cm, the 1.6 Mb and 2.2 Mb chromosomes from S. cerevisiae appear as a single band with 90 second pulse length. Increasing the pulse length to 120 seconds resolves these into two bands (Gemmill, 1991). PFGE allows for separation (resolution) of DNA larger than 1000 kb by varying direction and duration of the electrical pulses. In the study reported here, a contour clamped homogeneous electric field (CHEF) pulsed field unit was used. CHEF uses a complex electrophoresis chamber with multiple electrodes to produce a highly uniform electrophoretic field (Figure 2). To separate extremely large DNA (greater than 100, 000 bp), a reorientation angle between 96° and 105° is almost a requirement to get a good separation in the shortest possible time (Joppa et al., 2000). Usually, the CHEF unit reorients DNA molecules over a 120° angle (Goering). Thus, it was decided that PFGE would be used to type DNA profiles of *E. coli* from nonpoint source(s).

Fig. 2. This picture exhibits the electrode configuration of the contour-clamped homogeneous electric field (CHEF) mapper unit. The gel is depicted as a square in the figure. The negative electrode (-) is at the top of the gel and at the bottom of the gel is the positive electrode (+). The negatively charged DNA is being pulled through the agarose gel toward the positive electrode. The electric field surrounds the entire gel allowing for better resolution of large sized DNA (greater than 100, 000 bp).


CHEF
A−
B−
B+
A+

Potential nonpoint sources are defined as "poor management and handling of animal waste from feedlots and poor management and handling of animal waste from poultry litter, failing or ill-sited septic systems, poor management of pasture lands, excess application of manure or municipal sludge in cropland and other agricultural areas, and natural background loadings from wildlife" (U.S. Environmental Protection Agency Region III Report, 1998). The potential nonpoint sources involved in this study were waste from known domesticated (eg. cows and chickens) and wildlife (eg. deer) animals. This study was created as the basis for identifying possible animal source(s) of contamination in the regions of Moorefield, Petersburg, Rig, Old Fields, Maysville, Kessel, and Fisher. These sites are found along the northern headwaters of the south fork of the Potomac River which run into potential watersheds of West Virginia. The U.S. Environmental Protection Agency Region III Report stated that the poultry and cattle (dairy and beef) farms along the northern headwaters of the south fork were suspected to be polluters of the Potomac River (1998). Other sources, such as, wildlife, were not considered as a means of pollution (U.S. Environmental Protection Agency Region III Report, 1998). In constructing a database using PFGE, one would be able to determine the source of the pollution and bioremediate to help minimize the pollution of the rivers of West Virginia.

OBJECTIVES

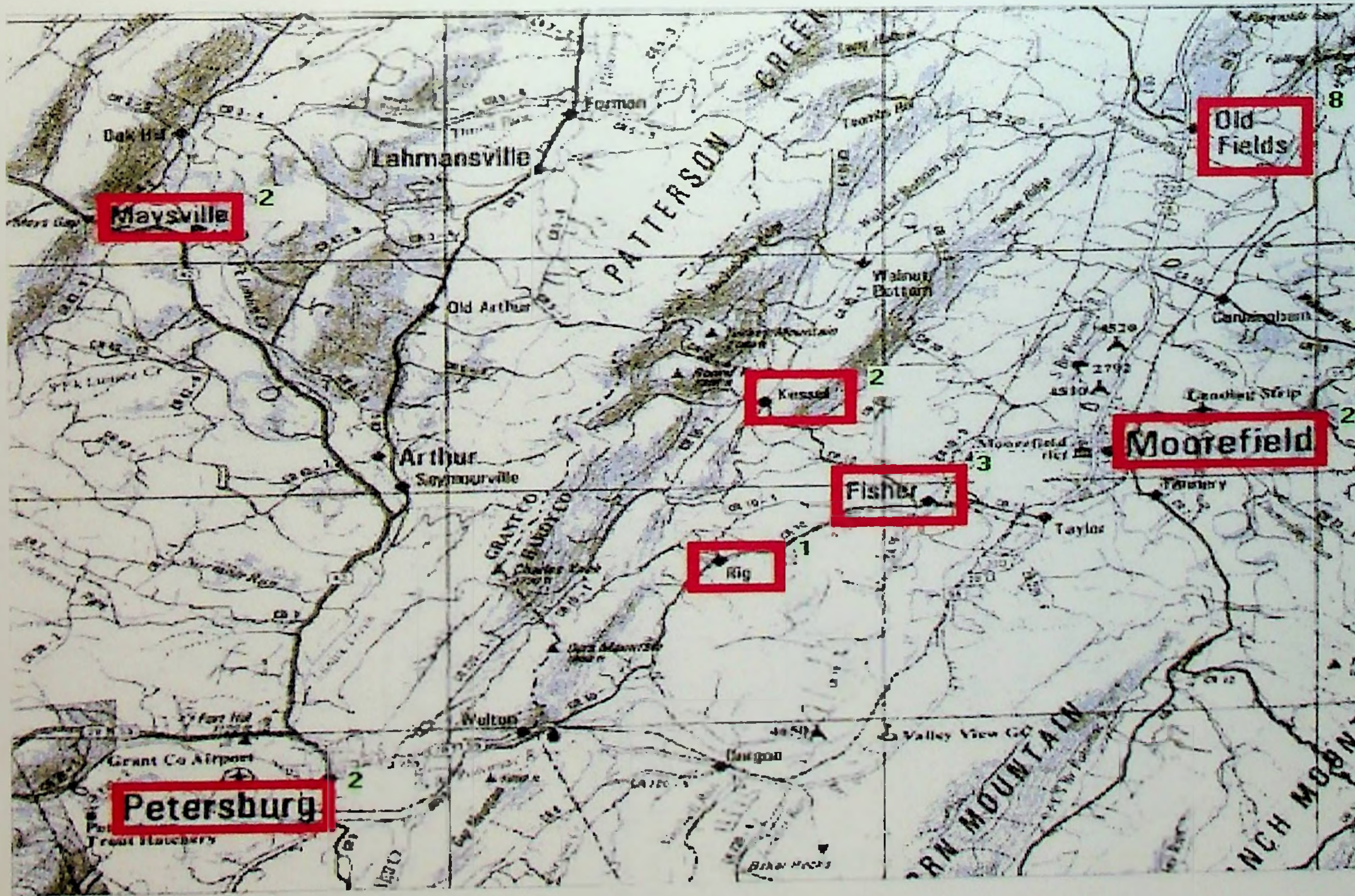
The main objective of this study was to answer the following question: Is it possible to differentiate between animal types and individuals of the same animal type using PFGE analysis of their DNA? An animal type was defined as one of the following:

Chicken, Deer, Angus Cow, Herford Cow, Charlais Cow, and Pensgar Cow. Animal fecal samples were obtained by the Moorefield, West Virginia laboratory from a small geographical region that is located in the northeastern part of West Virginia (Figure 3). Samples within this select region were from Petersburg, Maysville, Rig, Old Fields, Kessel, and Fisher. These locations border 6 tributaries that flow into the headwaters of the Potomac River. In total 14 cow samples were obtained: 2 from Petersburg, 2 from Maysville, 1 from Rig, 5 from Old Fields, 1 from Moorefield, and 3 from Fisher. Each particular type of cow was represented by 3 to 4 of the collected samples: 4 Angus Cow samples, 4 Herford Cow samples, 3 Charlais Cow samples, and 3 Pensgar Cow samples. In addition, 3 chicken samples from Old Fields had been obtained as well as 3 deer samples (1 from Moorefield and 2 from Kessel).

After the samples were collected, the following process was administered (Figure 4). Each fecal sample generated 10 *E. coli* isolates, which, in turn, generated 10 DNA profiles using PFGE. The resulting DNA profiles were analyzed for similarities between animal types and within an individual animal using visual observations (from Gel Doc 2000, Bio-Rad Laboratories, 1998) and statistics (using both Scanalytics software and Jandel Scientific software). Also, the pulsed-field generated DNA profiles were used to determine which restriction enzyme gave better discrimination between DNA profiles: *Not*I or *Xba*I. *Not*I, being an 8 base pair restriction endonuclease and *Xba*I, being a 6 base pair restriction endonuclease.

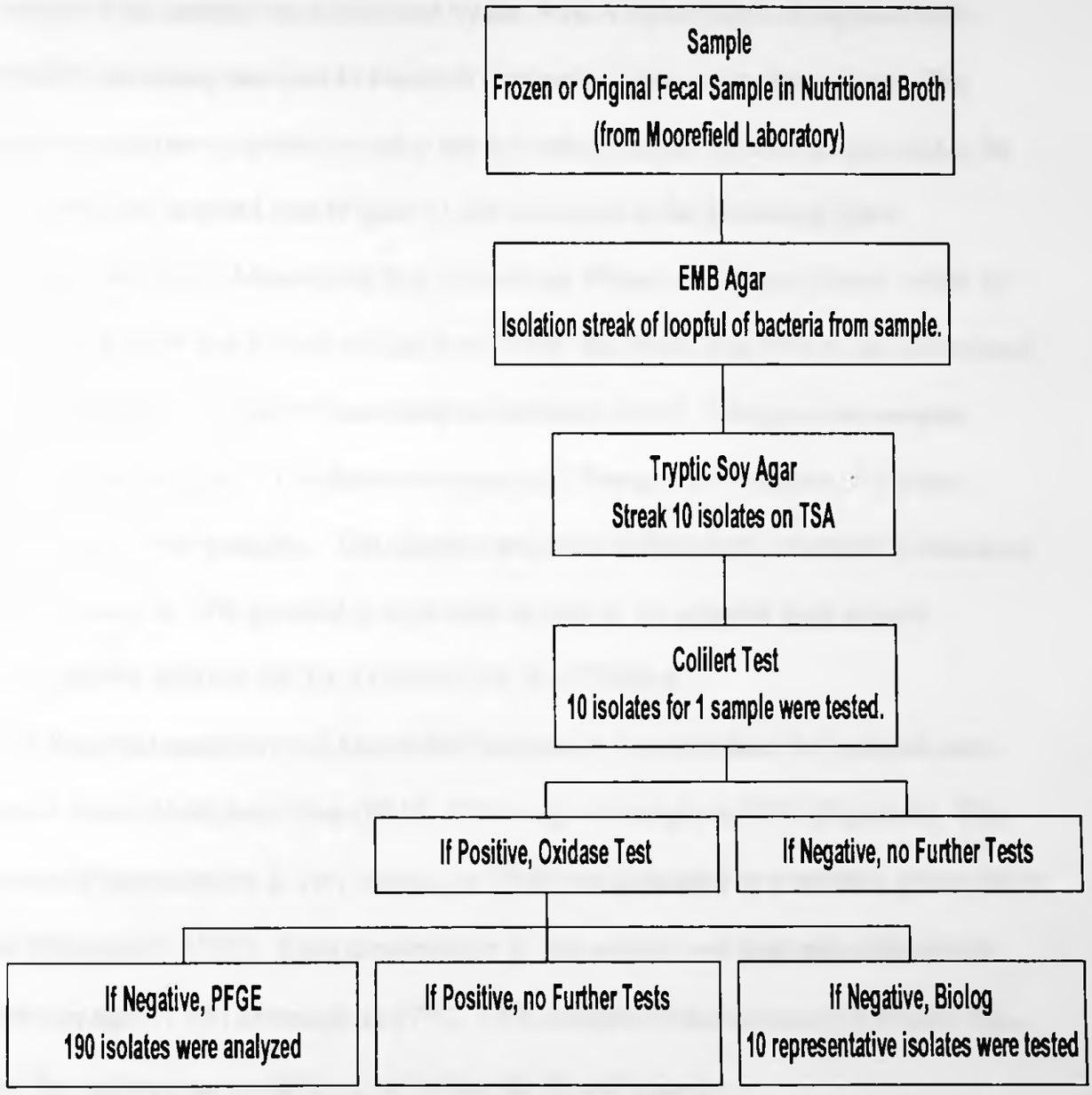
Presumptive *E. coli* tests had been administered to ensure that the DNA being analyzed was from *E. coli* (Figure 4). The tests used for this purpose were the Colilert* test, the oxidase test, using selective media for *E. coli* (Eosin Methylene Blue), and Biolog*.

Fig. 3. Map of area from which fecal samples were collected by the Moorefield laboratory. This region is found in the northeastern part of West Virginia. Samples were taken from Petersburg, Maysville, Rig, Old Fields, Moorefield, Kessel, and Fisher (indicated in red). These sites are found along the edge of 6 tributaries that flow into the headwaters of the Potomac River. The number of samples taken from each site are indicated in green. In total, 14 fecal samples from cows were collected: 2 from Petersburg, 2 from Maysville, 1 from Rig, 5 from Old Fields, 1 from Moorefield, and 3 from Fisher. In addition, 3 chicken samples were collected from Old Fields, and 3 deer samples were collected (2 from Kessel and 1 from Moorefield).



Old Fields
8
Moorefield
2
Maysville
2
Lahmansville
Forman
Oak Hill
Old Arthur
Arthur
Seymourville
Walnut Bottom
Cunningham
Kessel
2
Fisher
3
Rig
1
Taylor
Tannery
Landes Strip
Petersburg
2
Grant Co Airport
Trout Hatchery
Welton
Burgoo
Valley View GC
Baker Rocks
PATTERSON
GREEN
MOUNTAIN
GRANT CO
HARDY CO

Fig. 4. Flow chart showing the process used to obtain *E. coli* DNA profiles and tests used to determine whether or not the bacteria to be analyzed was *E. coli*.


Sample
Frozen or Original Fecal Sample in Nutritional Broth
(from Moorefield Laboratory)
EMB Agar
Isolation streak of loopful of bacteria from sample.
Tryptic Soy Agar
Streak 10 isolates on TSA
Colilert Test
10 isolates for 1 sample were tested.
If Positive, Oxidase Test
If Negative, no Further Tests
If Negative, PFGE
190 isolates were analyzed
If Positive, no Further Tests
If Negative, Biolog
10 representative isolates were tested

CHAPTER II.

MATERIALS AND METHODS

SAMPLE COLLECTION

Animal fecal samples were collected by the West Virginia Dept. of Agriculture's Moorefield laboratory and sent to Marshall University's laboratory for analysis. The samples for analysis were chosen using the following criteria: a) each sample had to be found within the selected area (Figure 3) that consisted of the following cities: Maysville, Old Fields, Moorefield, Rig, Petersburg, Fisher, and Kessel found within 50 miles of each other and b) each animal type within this select area were to be represented by 3 or 4 samples. In total 20 fecal samples had been tested: 4 Angus cow samples, 4 Herford cow samples, 3 Charlais cow samples, 3 Pensgar cow samples, 3 chicken samples, and 3 deer samples. The samples were sent to Marshall University's laboratory either as frozen in 25% glycerol in cryo-vials or sent as the original fecal sample.

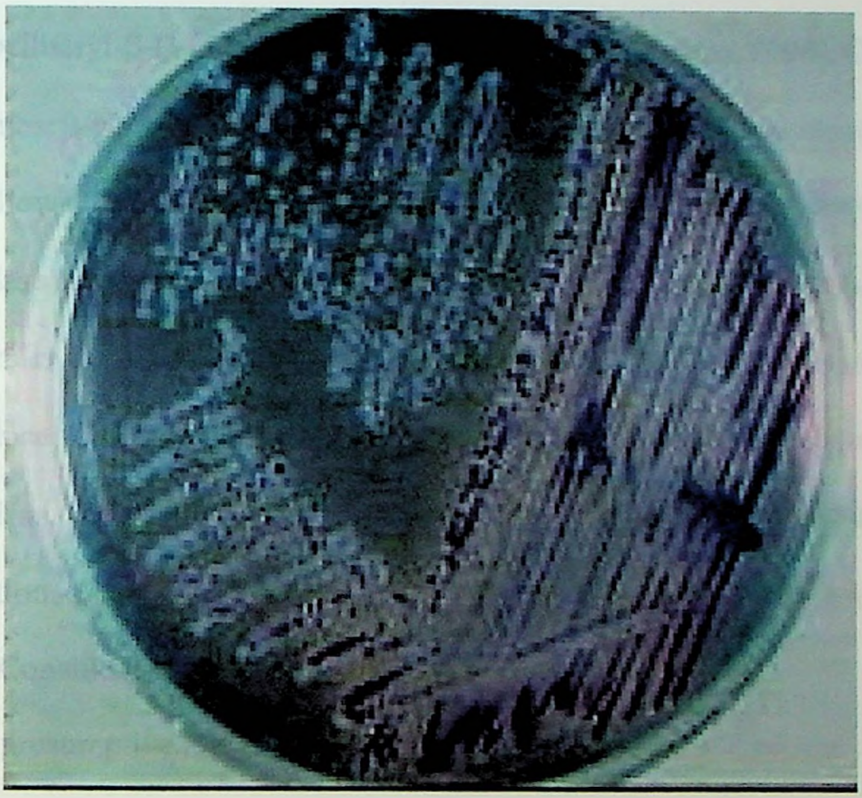
CULTURING AND SUBCULTURING OF BACTERIA

The fecal test samples from Moorefield laboratory were streaked for isolation onto Levine's Eosin Methylene Blue (EMB, Difco) agar overnight at 37°C (Figure 5). The presence of presumptive *E. coli* isolates on EMB was indicated by a metallic green sheen color (Weemhoff, 1997). Each presumptive *E. coli* colony was then sub-cultured on tryptic soy agar (TSA) overnight at 37°C. TSA consists of the necessary nutrients that, under the appropriate conditions, will allow for *E. coli* growth.

PRESUMPTIVE TESTS FOR *E. coli*

In total three presumptive tests were used to show that the particular colony tested was *E. coli*. The first presumptive test was determined by the phenotypic appearance of

Fig. 5. This picture exhibits presumptive *E. coli* bacteria that has been isolated using Levine's Eosin Methylene Blue (EMB) agarose (University of Texas-Houston Medical School, 1995). Note the isolation streaking of the bacteria on the EMB. The bacteria appears to have a metallic green sheen colour.



E. coli on EMB agar. The *E. coli* isolate to be sub-cultured should have the appearance of a round colony with a green metallic sheen on EMB agar plates.

The second presumptive test is the Colilert test. For this test, a loopful of bacterial cells, from isolated colonies on TSA plates, was placed into 5 ml of Colilert solution. The Colilert solution consists of ortho-nitrophenyl- β-D-galactopyranoside (ONPG) and 4-methylumbelliferyl-β-D-glucoronide (MUG) (Micro Consultants, 1998). Colilert tests for two enzymes: β-galactosidase and β-glucoronidase. β-galactosidase cleaves ONPG to produce a yellow compound called o-nitrophenol (Micro Consultants, 1998). β-glucoronidase reacts with MUG to produce methylumbelliferone which fluoresces blue when seen under ultra-violet (UV) light (Micro Consultants, 1998). Gram-negative bacteria produce the enzyme β-galactosidase and *E. coli* produces the enzyme, β-glucoronidase. Thus, after *E. coli* cells incubate overnight at 44.5°C for 24 hours in the Colilert solution, a visible yellow color and blue fluorescence should be seen under UV light (Micro Consultants, 1998).

The third presumptive test for *E. coli*, in this study, was the oxidase test. This test is based on the production of an enzyme by oxidase positive bacterial cells called indophenol oxidase. This enzyme oxidizes the phenylenediamine oxidase reagent to form a dark purple compound, indophenol (Becton Dickinson MicroBiology Systems, 1995). Since *E. coli* is an oxidase negative organism, it does not have the indophenol oxidase enzyme. Thus, once the reagent is added to the bacteria no reaction should occur and no color change should be observed.

BIOLOG

The final test, a determinative test for *E. coli*, was Biolog. For Biolog, the test colonies were incubated overnight on TSA plus 5% sheep's blood. *E. coli* cells were then suspended in an inoculating fluid to a transmittance of 63% ± 3% (established by gram negative, oxidase negative standard). The cell suspension was then transferred into each well of a biolog plate specific for gram-negative bacteria, since *E. coli* is gram negative. A biolog plate consists of 96 wells-95 are metabolic substrates and 1 is a negative control (MiroLog System User Guide, 1999). After a 20h and 24h period of incubation at 37°C, each plate creates a visual pattern. This visual pattern is due to the ability or inability of the bacteria to utilize the substrate, tetrazolium redox dye, in each well. Substrate metabolism causes the conversion of the redox dye from colorless to purple (MiroLog System User Guide, 1999).

To identify the bacterial colony, the visual pattern in a plate was read using a microtiter reader and the pattern was matched to the appropriate gram-negative database. Ten presumptive *E. coli* colonies were chosen and tested with biolog in triplicate. To determine if Biolog version 4.0 was precise in its conclusions, the clustering of the strains were analyzed from rectangular cladograms. For each strain, presence/absence of purple color in a well was transcribed into binary scores. As a result, these scores generated rectangular cladograms using the dollop program of the Phylip version 3.5c software (Felsenstein, 1993).

WHOLE CHROMOSOME RFLP

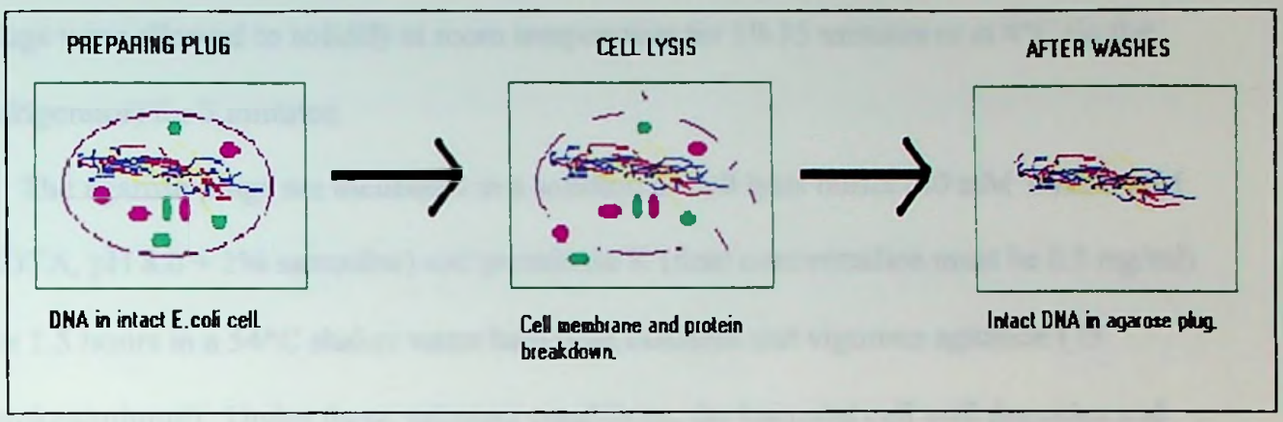
In order to electrophorese *E. coli* DNA with PFGE, DNA must be contained within an agarose plug, 1 mm thick. Plugs were prepared using a modified version of the protocol by the Centers for Disease Control as stated in the Pulse Net protocol. Ten colonies per sample (believed to be *E. coli* as indicated by the presumptive tests) were incorporated into plugs using the modified Pulse Net method. The *E. coli* cells from the TSA plates were suspended in 2 ml of cell suspension buffer (100mM Tris:100mM EDTA, pH 8.0). EDTA is a chelating agent that binds and removes calcium and/or magnesium, to help prevent degradation of the DNA via Dnases. The cells were evenly dispersed and adjusted to a concentration of 10% transmittance using the bioMerieux Vitek colorimeter (by diluting with sterile CSB or adding additional cells) as determined by previous experiments that had resulted in well resolved DNA profiles using PFGE (Noureddine, 1998). Positive and negative controls were prepared using the same method. For *Xba*I gels, the positive control was the *E. coli* O157:H7 standard, G5244 (generously donated by Dr. Toney of the Virginia Public Health Department) and the negative control is a plug that contains no bacterial cells whatsoever. For *Not*I gels, the positive control was JM107 and the negative control was the same as that for *Xba*I gels.

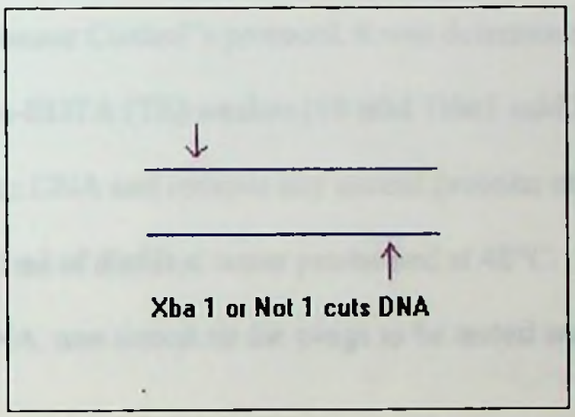
200 µl of melted 1.6% InCert containing 1% SDS at 60°C, 10 µl proteinase K (20 mg/ml) and 200 µl *E. coli* cell suspension were mixed gently with a pipettor. SDS or sodium dodecyl sulfide is a denaturing agent that denatures proteins. Proteinase K is an enzyme that cleaves protein. Immediately after mixing the plug components, 200 µl of the mixture was dispensed into the appropriate well of a disposable plug mold and the

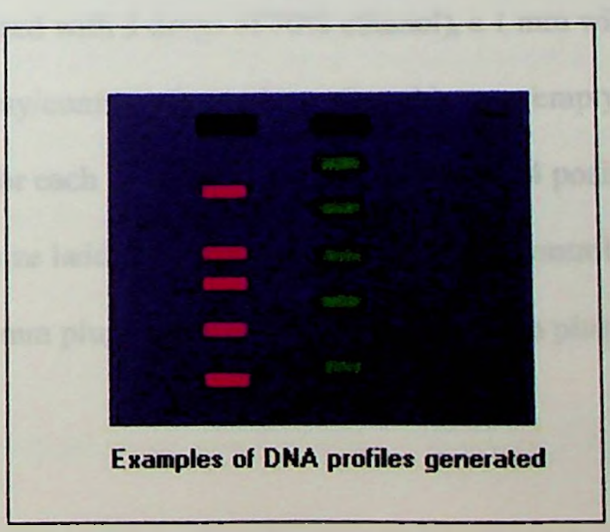
Fig. 6. Isolation of *E. coli* cell DNA in agarose plug.

Fig. 7. Action of a restriction endonuclease on chromosomal DNA of *E. coli*.

Fig. 8. Example of DNA profiles generated from a pair of selected colonies using PFGE.


PREPARING PLUG
CELL LYSIS
AFTER WASHES
DNA in intact E.coli cell
Cell membrane and protein breakdown.
Intact DNA in agarose plug.


Xba 1 or Not 1 cuts DNA


Examples of DNA profiles generated

plugs were allowed to solidify at room temperature for 10-15 minutes or at 4°C (in the refrigerator) for 5 minutes.

The agarose plugs are incubated in a solution of cell lysis buffer (50 mM Tris:50 mM EDTA, pH 8.0 + 1% sarcosine) and proteinase K (final concentration must be 0.5 mg/ml) for 1.5 hours in a 54°C shaker water bath with constant and vigorous agitation (75 strokes/minute). Under these stringent conditions, the bacterial cell wall degrades and proteins are further denatured. The proteins and cell wall wash out into the buffer solution leaving the *E. coli* DNA in the agarose plug (Figure 6).

Unlike the Centers for Disease Control's protocol, it was determined that two water washes followed by two Tris-EDTA (TE) washes (10 mM Tris:1 mM EDTA, pH 8.0) was sufficient to clean up the DNA and remove any excess proteins or loose membrane. A water wash consists of 10 ml of distilled water pre-heated at 48°C. As a final step for the isolation of bacterial DNA, one stored all the plugs to be tested and controls in TE at 4°C.

Each plug is carefully removed from the cold TE and placed onto a glass slide. Using a cover slip (sterilized with 3 drops of 70% ethanol), a 1 mm wide slice was cut from the test sample colony/control plug and transferred into an empty tube. This procedure was repeated for each of the10 colonies. In addition, 4 positive control plugs (which act as molecular size ladders in the gel) and 1 negative control plug were cut. In total there were fifteen 1 mm plugs processed for each gel. Each plug was placed in a separate tube.

For restriction of DNA (Figure 7) 10X H buffer (Roche) was diluted 1:10 with sterile reagent grade water and the *Xba*I (Roche) restriction enzyme was added to the buffer as shown in Table 1:

Table 1: Restriction Digestion of DNA in Agarose Plugs with *Xba*I

	Sterile Reagent Grade Water	H Buffer	Enzyme (10 U/µl)	Total Vol.
µl/Plug Slice	87 µl	10 µl	3 µl	100 µl
µl/10 Plug Slices	870 µl	100 µl	30 µl	1000 µl
µl/15 Plug Slices	1305 µl	150 µl	45 µl	1500 µl

100 µl aliquots of the total restriction enzyme mixture were added to each tube containing a 1 mm plug. To ensure that each plug was under the enzyme mixture, each tube was closed and mixed by tapping gently. The plug slices were then incubated at 37°C in a standing water bath for 1 hour.

Each strain that generated unique DNA profiles with *Xba*I was digested with *Not*I (Promega Corp.). A unique DNA profile is defined as a DNA profile of a test colony that was observed to be different from any other DNA profile of all strains analyzed. A positive control (JM107) and a negative control were also digested with *Not*I. First, 10x D Buffer (Promega Corp.) was diluted 1:10 with reagent grade water and BSA (Promega Corp.) using the following method as established by Krista Haught, MS Forensic Science of the Potomac River Project at Marshall University:

18.2 µl Buffer D x # of tubes containing plug slices

182 µl sterile reagent grade water x # of tubes containing plug slices

2 µl BSA x # of tubes containing plug slices

Table 2. Restriction Digestion of DNA in Agarose Plugs with *Not*I

Reagent	µl/Plug Slice	µl/10 Plug Slice	µl/15 Plug Slice
Sterile Reagent Grade Water	182 µl	1820 µl	2730 µl
D Buffer	18.2 µl	182 µl	273 µl
BSA	2 µl	20 µl	20 µl
Enzyme (10 U/µl)	2 µl	20 µl	20 µl

200 µl aliquots of the *Not*I reaction mixture were added to the 1 mm plug slices. Plugs were digested overnight at 37°C.

0.5x Tris-Borate EDTA (TBE) was needed for both the gel and electrophoresis running buffer. 2 liters of 0.5x TBE was poured into the electrophoresis chamber and warmed up to 14°C. The gels were cast according to the Centers for Disease Control protocol. 1 g of Seakem agarose was added to 100 ml of 0.5x TBE to make a solution of 1% Seakem Gold agarose (SKG). The 100 ml solution was heated and gently swirled every 30 seconds until the agarose had completely dissolved. Most of the melted SKG was poured into a gel box fitted with a 15-well comb. After 20 minutes, the gel polymerized so that it could be loaded with the pre-digested plugs.

For *Xba*I digested plugs: the isolates to be analyzed were loaded into their appropriate wells. The G5244 positive control, which, also functioned as the ladder, was loaded every fourth lane in the gel to allow for accurate measurement of the molecular weights and to compensate for smiling of bands. Thus, for *Xba*I digested plugs: the G5244 positive control/ladder would be in lanes 1, 5, 9, and 13, allowing for more precision to

determine the molecular weights of the bands of the DNA profiles to be analyzed. For *Not*I digested plugs: two 1 mm wide lambda ladders were cut with a cover slip and loaded into their appropriate wells. In addition, the positive (JM107) control, negative controls and *E. coli* isolates, that generated unique DNA profiles via *Xba*I DNA analysis, were loaded into their appropriate wells. Whether the plugs were digested with *Xba*I or *Not*I, the negative control was always loaded in the furthest left lane of the gel. The excess melted agarose (left over from when the gel was cast) was poured over the wells and allowed to harden. This kept the plugs from floating out of the wells and prevented shearing of the DNA.

For *Xba*I digested plugs: the gel was placed into the electrophoresis chamber and ran for 17.5 hours, at 200V, initial switch time 2.16 sec., and final switch time at 54.17 sec. For *Not*I digested plugs: the gel was placed into the electrophoresis chamber and allowed to run overnight for 18 hours at the same voltage and pulse times.

When the electrophoresis run was over, the gel was removed and stained. The gel was stained with diluted ethidium bromide (add 40 µl of ethidium bromide stock solution (10 mg/ml) to 400 ml of distilled water) for 30 minutes and destained for 30 minutes to 1 hour in 500 ml of distilled water. Ethidium bromide is an intercalating agent that intercalates between double stranded DNA and fluoresces under ultra-violet light. After the gel was stained and destained, the image was recorded and documented (Figure 8) using the gel documentation system Gel Doc 2000 (Bio-Rad Laboratories, 1998).

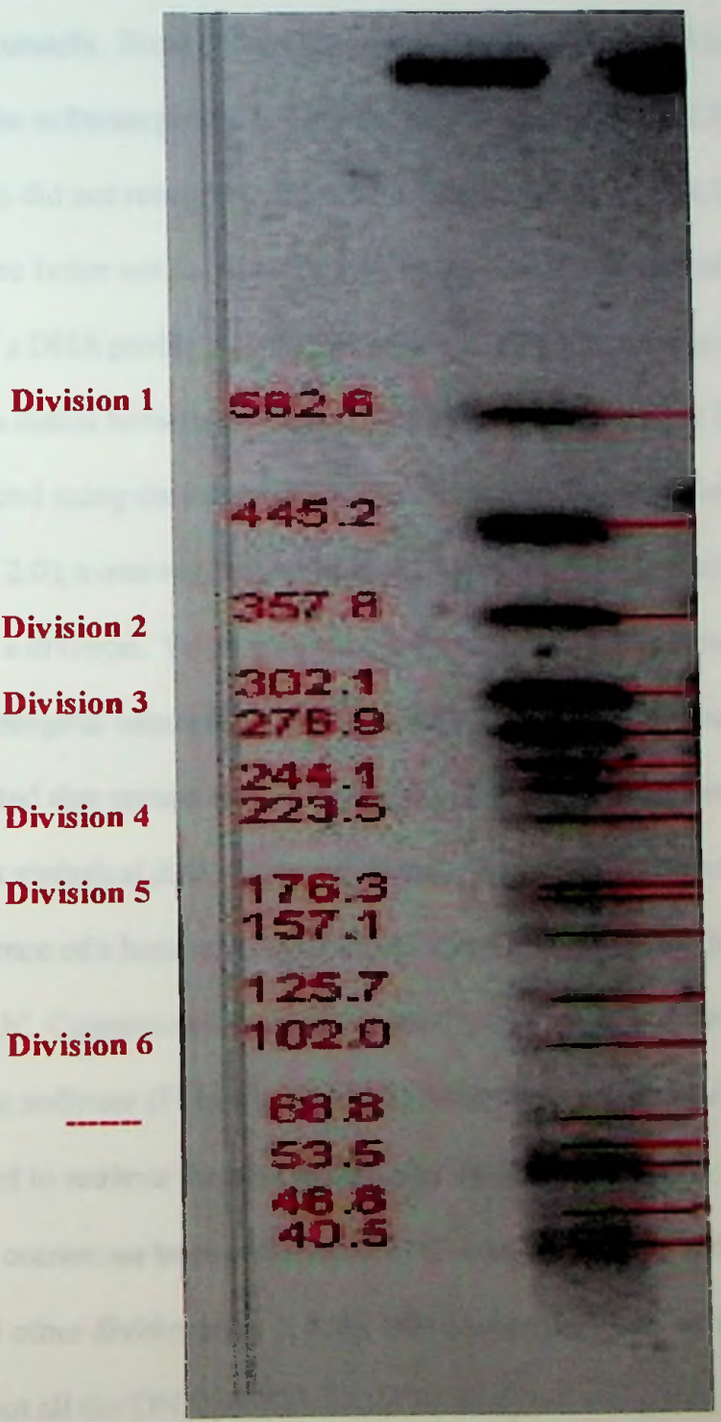
STASTICAL ANALYSIS OF DATA

Each DNA profile that had been observed to be the same was given a strain identification name. Based on visual observation the frequency of similar DNA profiles obtained by *Xba*I digestion was determined for each individual animal: Herford Cow 3179910, Herford Cow 7219903, Angus Cow 3319904 etc.

The gels were analyzed and molecular weights of bands were obtained using the Scanalytics software RFLPscan Forensic Edition version 3.0 (Scanalytics, A division of CSPI, 1994). There was not enough data to analyze the DNA profiles generated by *Not*I digestion. Only observations were noted for these DNA profiles. On the other hand, there was a large amount of molecular weight data obtained from DNA profiles that had been generated with *Xba*I. Each DNA profile generated by PFGE and digested with *Xba*I was divided into 6 divisions: 582.6 kb to 357.8 kb, 357.79 kb to 302.1 kb, 302.09 kb to 223.5 kb, 223.49 kb to185.3 kb, 185.29 kb to 102.0 kb, 101.99 kb to 68.8 kb (Figure 9). These divisions were based upon the molecular weights of the G5244 standard bands. The molecular weights of the bands, from top to bottom, of the G5244 standard are: band #1: 582.6 kb, band #2: 445.2 kb, band #3: 357.8 kb, band #4: 302.1 kb, band #5: 276.9 kb, band #6: 257.5 kb, band #7: 244.1 kb, band #8: 223.5 kb, band #9: 185.3 kb, band #10: 176.3 kb, band #11: 157.1 kb, band #12: 125.7 kb, band #13: 102.0 kb, band #14: 68.8 kb, band #15: 60.3 kb, band #16: 53.5 kb, band #17: 46.8 kb, band #18: 40.5 kb, and band #19: 36.3 kb (Figure 9).

It was decided that the DNA profiles generated by PFGE would be analyzed from the first band of the standard, which has a molecular weight of 582.6 kb to the fourteenth band, which has a molecular weight of 68.8 kb. Bands of a size above that of 582.6 kb,

Fig. 9. A representation of the molecular weights of the G5244 standard bands. Based on this criterion, the 6 divisions for statistical purposes were generated for the *Xba*I DNA database.


Division 1
582.8
445.2
Division 2
357.8
Division 3
302.1
276.9
244.1
Division 4
223.5
Division 5
176.3
157.1
125.7
Division 6
102.0
68.8
53.5
46.8
40.5

which were seen in some DNA profiles, were excluded from any statistical analysis because they were not found within the ladder region. These particular bands would be outside of the standard curve for the molecular weights and their molecular weights could not be estimated accurately. Bands of a molecular weight less than 68.8 kb were too light to be measured by the software program. The Scanalytics software (RFLPscan, Forensics Edition, version 3.0) did not recognize these faint bands indicating that, for comparison purposes, it would be better not to include these bands into the statistical analysis.

Each division of a DNA profile was further divided into subdivisions to allow for better precision of a match between groups of similar molecular weight bands. All DNA profiles were analyzed using the same procedure. Using the Jandel Scientific software (SigmaStat version 2.0), a one-way anova analysis was performed on the raw data of each subdivision within a division. To do pair-wise comparisons for faster analysis, the Tukey Test was used. Analogous bands from identical banding patterns were compared. The null hypothesis stated that groups of similar molecular weight bands were the same at $p<0.05$. Using this statistical data, presence/absence reports were generated. A one (1) indicated the presence of a band or a 'match' and a zero (0) indicated the absence of a band or a 'no match'. Consensus trees were created using the dollop program of the Phylip version 3.5c software (Felsenstein, 1993) from these presence/absence reports. Tree view was used to retrieve the best rectangular cladogram for each consensus tree. For division 1, 52 consensus trees were created. 83 consensus trees were generated for division 5. For all other divisions (2, 3, 4, 6), 100 consensus trees were generated.

Unfortunately, not all the DNA profiles could be analyzed using this method because the Phylip software was unable to analyze 190 strains in a reasonable amount of time.

Thus, the large data sets were reduced to 48 strains to allow for faster analysis of the data sets. 38 of the strains analyzed were observed to be unique DNA profiles in the animal database. The other 10 strains came from 1 animal to ensure that clustering of the strains were occurring in the appropriate manner, since it was known which strains were identical for that animal. In total there were 6 data sets for the 48 strains, one per division.

CHAPTER III.

RESULTS AND DISCUSSION

PRESUMPTIVE TESTS

All 190 colonies tested appeared to be green and shiny when grown on EMB. When incubated in Colilert solution, the bacteria appeared yellow and fluoresced blue using 365 nm UV light, indicating that they were colilert positive for *E. coli*. In addition, all colonies had been found to be oxidase negative.

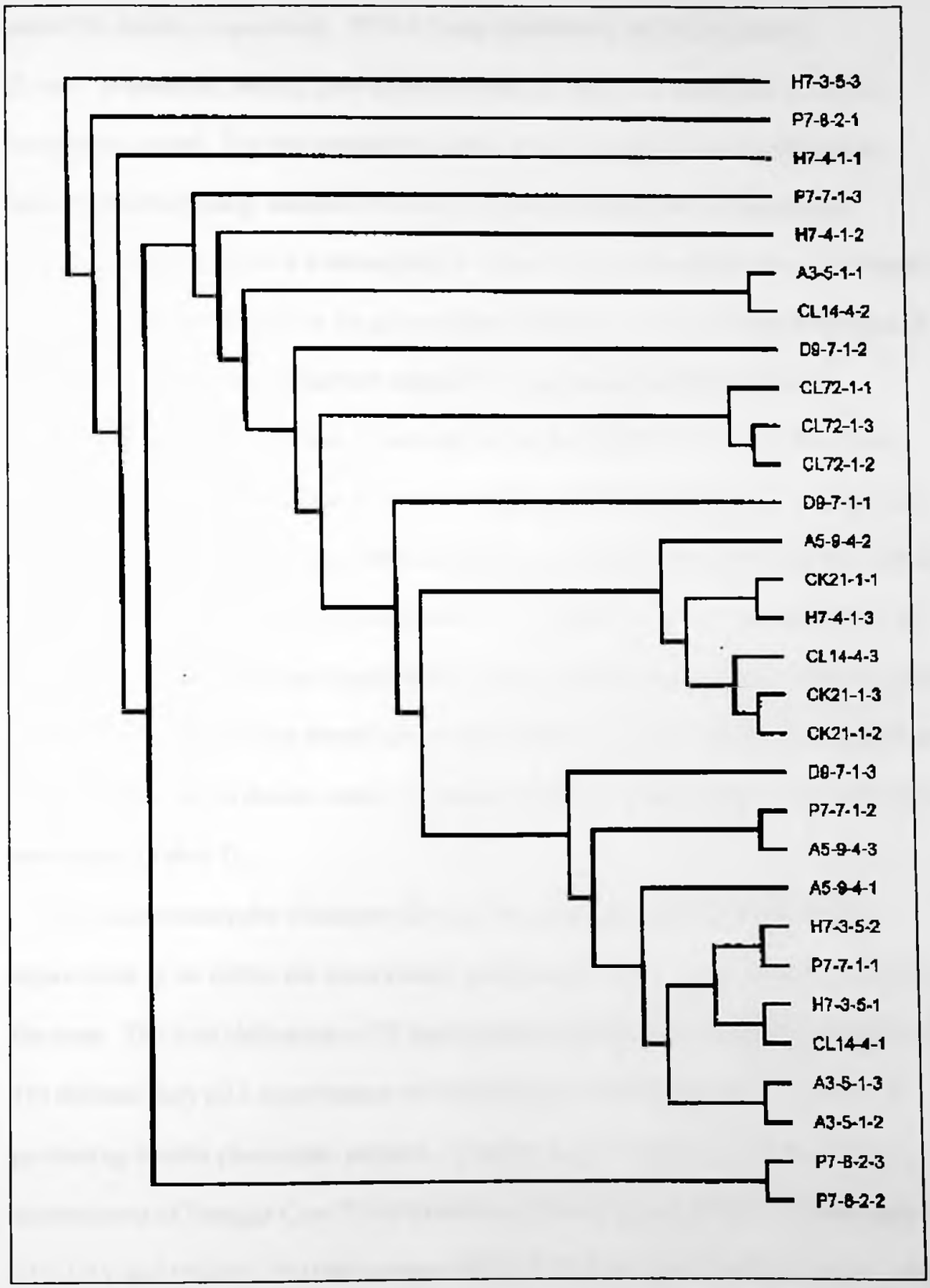
BIOLOG

The 10 bacterial isolates tested with Biolog version 4.0 were Pensgar Cow 7219908 colony 2 (P7-8-2), Pensgar Cow 7219907 colony 1 (P7-7-1), Deer 9807 colony 1 (D9-7-1), Angus Cow 5059909 colony 4 (A5-9-4), Angus Cow 3319905 colony 1(A3-5-1), Charlais Cow 10069904 colony 4 (CL1-4-4), Charlais Cow 7219902 colony 1 (CL7-2-1), Herford Cow 7219903 colony 5 (H7-3-5), Herford Cow 7219904 colony 1 (H7-4-1), and Chicken 7219901 colony 1 (CK21-1). Each isolate was tested in triplicate. Each experiment was given the designation 1, 2, and 3. For example, A5-9-1 experiment 1 was called A5-9-1-1, A5-9-1 experiment 2 was called A5-9-1-2, and A5-9-1 experiment 3 was called A5-9-1-3. The same designations were given to the rest of the bacterial isolates tested with Biolog (Table 3). Biolog could not identify CL7-2-1-1 at the 24h incubation period because of human error. This was signified in the table as NR (no reaction).

When compared to the gram-negative database, Biolog gave different strain identifications for each experiment of the same strain (Table 3). For example, after a 20-

Figure 10. Rectangular Cladogram of Biolog data at 20h incubation period, Phylip version 3.5c (Felsenstein, 1993)

BIOLOG DATA AT 20h



hour incubation at 37°C, Biolog identified P7-7-1-1 and P7-7-1-2 as *E. Hermanii*, 100% and 81% similar, respectively. P7-7-1-3 was identified to be 99% similar to *E. coli*. In addition, Biolog gave different results for the same strain after a 24-hour incubation period. For this incubation period, P7-7-1-1 and P7-7-1-2 could not be identified with Biolog, whereas, P7-7-1-3 was determined to be 92 % similar to *E. coli*. Similarity index is a determination of how similar the sample strain is compared to a known bacterial strain in the gram-negative database. For those strains mentioned that Biolog could not conclusively identify, it was determined that the closest identification for P7-7-1-1 was *E. coli* with a similarity index of 0.29 and the closest identification for P7-7-1-2 was *E. Hermanii* with a similarity index of 0.24. Thus, this indicated a variation of Biolog results within one strain not only for one incubation period but between two incubation periods as well. According to Biolog's protocol, the 24h incubation period results are supposed to be more reliable than the 20h incubation period results. Although, Biolog should show similar phenotypic patterns for all 3 experiments of each strain, for all strains tested, a variation of Biolog's identification was seen within each strain (Table 3).

The ideal rectangular cladogram for each incubation period should show each experiment to be within the same cluster unit for each strain. Unfortunately, this was not the case. The best cladogram of 23 trees generated for the 20h incubation period (Figure 10) showed only all 3 experiments of Charlais Cow 7219902 colony 1 (CL7-2-1) as producing similar phenotypic patterns. In addition, it showed that 2 of the 3 Biolog experiments of Pensgar Cow 7219908 colony 3 (P7-8-3), Angus Cow 3319905 colony 1 (A3-5-1), and Chicken 7219901 colony 1 (CK21-1) produced similar phenotypic patterns

Figure 11. Rectangular Cladogram of Biolog data at 24h incubation period, Phylip version 3.5c (Felsenstein, 1993)

BIOLOG DATA AT 24h

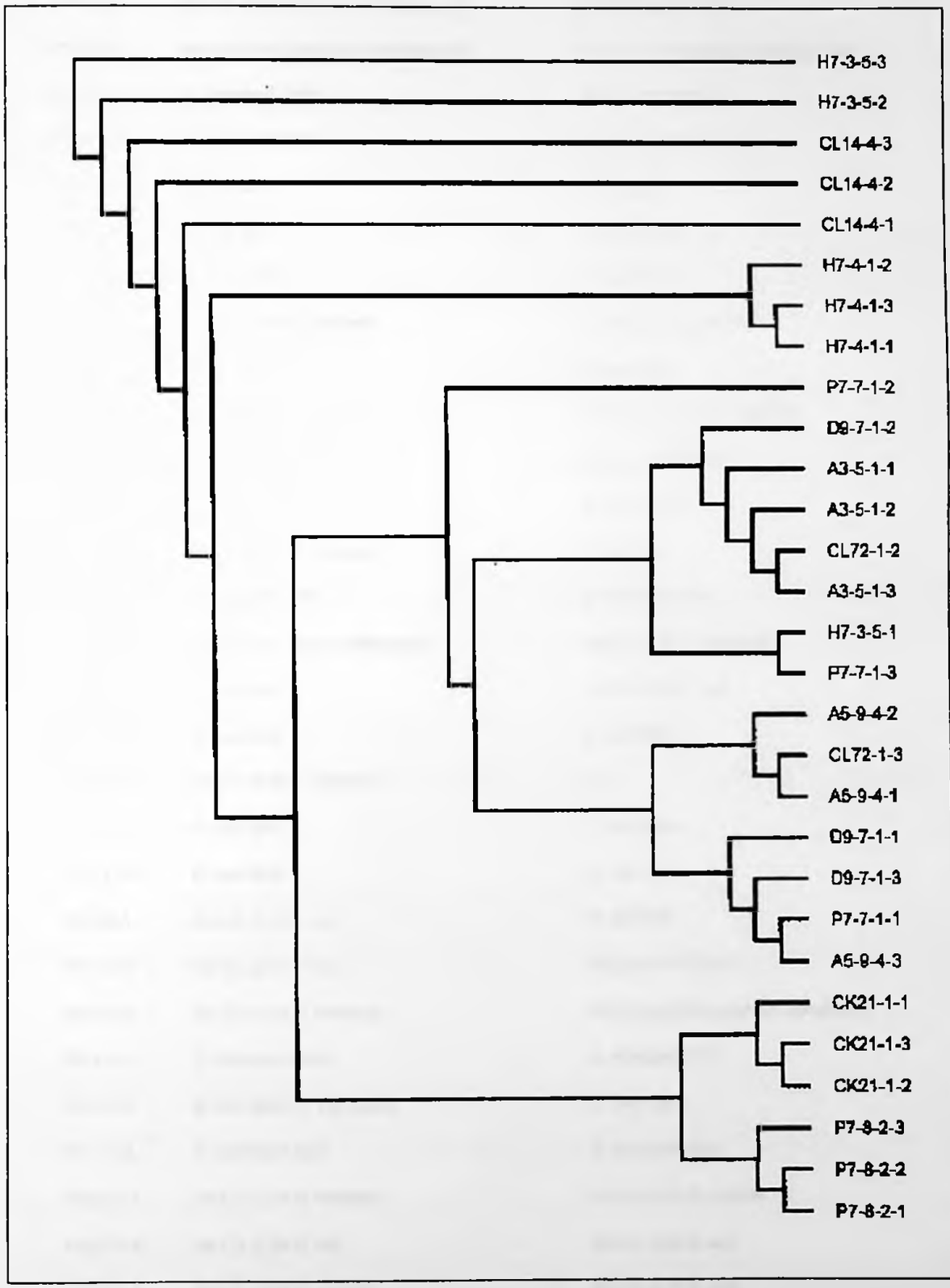


Table 3. Strain Identification by Biolog

Strain	20h incubation	24h incubation
P7-8-2-1	No I.D.-0.44 *E. Hermanii*	*E. Hermanii*-98%
P7-8-2-2	No I.D.-0.35 *Leclercia Adecarboxylata*	No I.D.-0.46 *E. vulnaris*
P7-8-2-3	No I.D.-0.16 *Leclercia Adecarboxylata*	No I.D.-0.31 *Leclercia Adecarboxylata*
P7-7-1-1	*E. Hermanii*-100%	No I.D.-0.29 *E. coli*
P7-7-1-2	*E. Hermanii*-81%	No I.D.-0.24 *E. Hermanii*
P7-7-1-3	*E. coli*-99%	*E. coli*-92%
D9-7-1-1	*E. coli*-97%	*E. coli*-100%
D9-7-1-2	*E. coli*-100%	*E. coli*-100%
D9-7-1-3	No I.D.-0.25 *E. Hermanii*	*Citrobakter Braakii*-84%
A5-9-4-1	*E. coli*-75%	*E. coli*-73%
A5-9-4-2	*E. coli*-89%	No I.D.-0.42 *E. coli* O157:H7
A5-9-4-3	No I.D.-0.48 *E. coli*	No I.D.-0.49 *E. coli*
A3-5-1-1	*E. coli*-95%	*E. coli*-100%
A3-5-1-2	No I.D.-0.29 *E. Hermanii*	*E. coli*-95%
A3-5-1-3	*E. Hermanii*-88%	*E. Hermanii*-94%
CL1-4-4-1	No I.D.-0.21 *Buttiauxiella Izardii*	No I.D.-0.38 *E. Hermanii*
CL1-4-4-2	*E. coli*-93%	No I.D.-0.46 *E. coli*
CL1-4-4-3	*E. coli*-89%	*E. coli*-94%
CL7-2-1-1	No I.D.-0.49 *E. Hermanii*	NR
CL7-2-1-2	*E. coli*-86%	*E. coli*-98%
CL7-2-1-3	*E. coli*-80%	*E. coli*-100%
H7-3-5-1	No I.D.-0.35 *E. coli*	*E. coli*-99%
H7-3-5-2	No I.D.-0.31 *E. coli*	No I.D.-0.31 *E. coli*
H7-3-5-3	No I.D.-0.14 *E. Hermanii*	No I.D.-0.30 *Enterobacter Amnigenus*
H7-4-1-1	*E. Hermanii*-99%	*E. Hermanii*-97%
H7-4-1-2	*E. coli*-99%, *E. coli*-100%	*E. coli*-100%
H7-4-1-3	*E. Hermanii*-80%	*E. Hermanii*-99%
CK21-1-1	No I.D.-0.20 *E. Hermanii*	No I.D.-0.33 *E. vulnaris*
CK21-1-2	No I.D.-0.38 *E. coli*	No I.D.-0.26 *E. coli*
CK21-1-3	No I.D.-0.40 *E. vulnaris*	No I.D.-0.17 *E. coli*

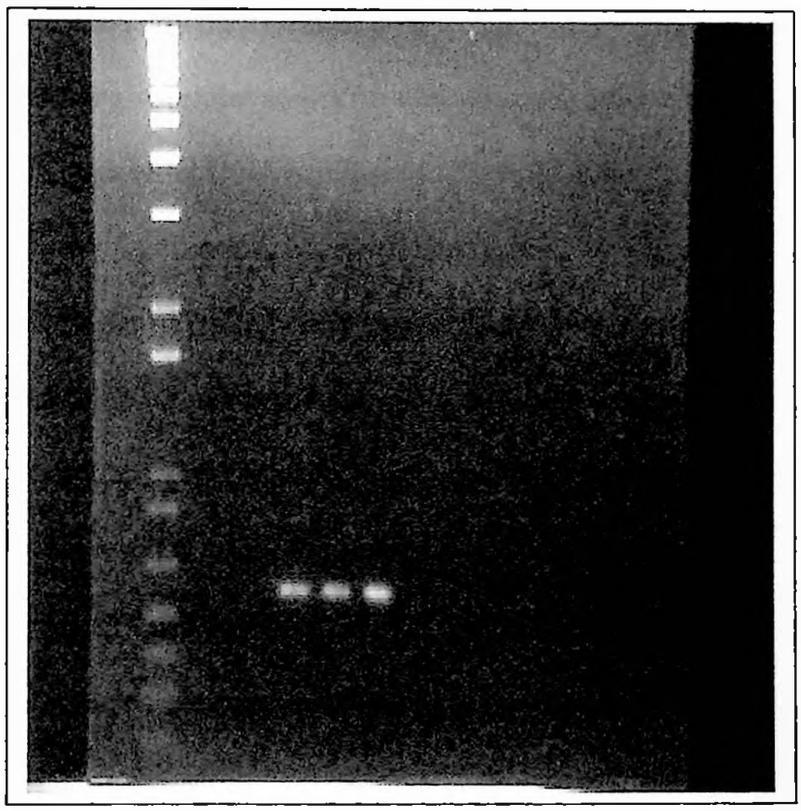
as read by Biolog. Of the two incubation periods, the 24h incubation period gave the better rectangular cladogram (Figure 11). The best cladogram of 100 trees generated by the Phylip version 3.5c (Felsenstein, 1993) showed that all 3 experiments of each of the following 3 bacterial strains gave similar patterns: Pensgar Cow 7219908 colony 1 (P7-8-1), Chicken 7219901 colony 1 (CK21-1), and Herford Cow 7219904 colony 1 (H7-4-1). Unfortunately, this indicated that the other 7 bacterial strains tested did not cluster properly into each of their similar units. Therefore, it was determined that Biolog version 4.0 was not sensitive enough to differentiate from strain to strain.

PCR TESTING

In addition, another test was used after Biolog to determine if the strains were *E. coli.* A representational sample, of 3 strains previously tested with Biolog were selected for polymerase chain reaction (PCR) analysis. The 3 bacterial strains tested were CL7-2-1, D9-7-1, and P7-8-2. The two oligonucleotide primers, BE1 and BE2, are specific for the rpo E gene (accession #: U37089 in the Genebank) of *E. coli.* The rpo E gene is the heat shock sigma factor for the initiation of transcription for *E. coli.* PCR was used to amplify this gene. As a result, one could determine the molecular weight of the band that each strain produced using agarose gel electrophoresis. (Figure 12). The molecular weight of the band generated for CL7-2-1 was 565.12 base pairs (bp), for D9-7-1 was 560.20 bp, and for P7-8-2 was 555.33 bp. The actual size of the rpo E gene is 576 bp, thus, producing a percent error of 1.89% for CL7-2-1, 2.74% for D9-7-1, and 3.59% for P7-8-2. Percent error was determined using the formula: [(actual-experimental)/actual] x 100. Based upon the criteria that the percent error for each strain tested was tight (less than 5%) and that the bands appeared to have migrated the same

Figure 12. E.coli (PCR Test) Lane 1: 1kb Ladder, Lane 2: Charlais Cow 7219902 colony 1, Lane 3: Deer 9807 colony 1, Lane 4: Pensgar Cow 7219908 colony 2.

1 2 3 4



distance on the gel (Figure 12), indicated that the rpo E gene was amplified in the above three strains. Thus, it would seem that based upon this PCR testing that the above strains could be *E. coli.* The strain that could easily be questioned is P7-8-2, which did not give any indication of being *E. coli* when identified by Biolog but seemed to be identified as *E. coli* based on PCR testing. Further testing needs to be done to validate this method as a possible way of identifying strains of bacteria to be *E. coli.*

PULSED FIELD GEL ELECTROPHORESIS GENERATED DNA PROFILES

There have not been any publications to date comparing PFGE generated DNA profiles of *E. coli* from nonpoint source(s) with contaminated water. Currently, there is a database of PFGE generated DNA profiles of *E. coli* from known animal source(s) being developed at Virginia Tech and at Marshall University. Dr. George Simmons' Laboratory, Department of Biology at Virginia Tech has been collecting wildlife samples and generating PFGE DNA profiles of *E. coli* at selected study sites along the Eastern Shore of the Chesapeake Bay. Marshall University's Department of Forensic Science, under the supervision of Dr. Terry Fenger, have been collecting domestic and wildlife samples to generate PFGE DNA profiles of *E. coli* at selected sites along the Potomac River. These databases are being generated for the purpose of studying environmental problems.

All samples selected, for this study, were given designations by the Moorefield Laboratory of the West Virginia Department of Agriculture. This was for record-keeping purposes only. Figures 13 to 73 show all the DNA profiles that were generated by each strain using PFGE. Each strain that generated unique DNA profiles with *Xba*I was independently digested with *Not*I. The resulting DNA profiles generated by *Not*I

Figure 13. Angus Cow 3319904 (A3-4) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on October 14, 1999. Lane 1: Standard (G5244), Lane 2: A3-4 colony 1, Lane 3: A3-4 colony 2, Lane 4: A3-4 colony 3, Lane 5: Standard (G5244), Lane 6: A3-4 colony 4, Lane 7: A3-4 colony 5, Lane 8: A3-4 colony 6, Lane 9: Standard (G5244), Lane 10: A3-4 colony 7, Lane 11: A3-4 colony 8, Lane 12: A3-4 colony 9, Lane 13: Standard (G5244), Lane 14: A3-4 colony 10, Lane 15: Negative Control.

Figure 14. Extra DNA Profiles (re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 11, 1999. Lane 16: A3-4 colony 10

Figure 15. A3-4 Unique DNA Profile (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 17: A3-4 colony 1.

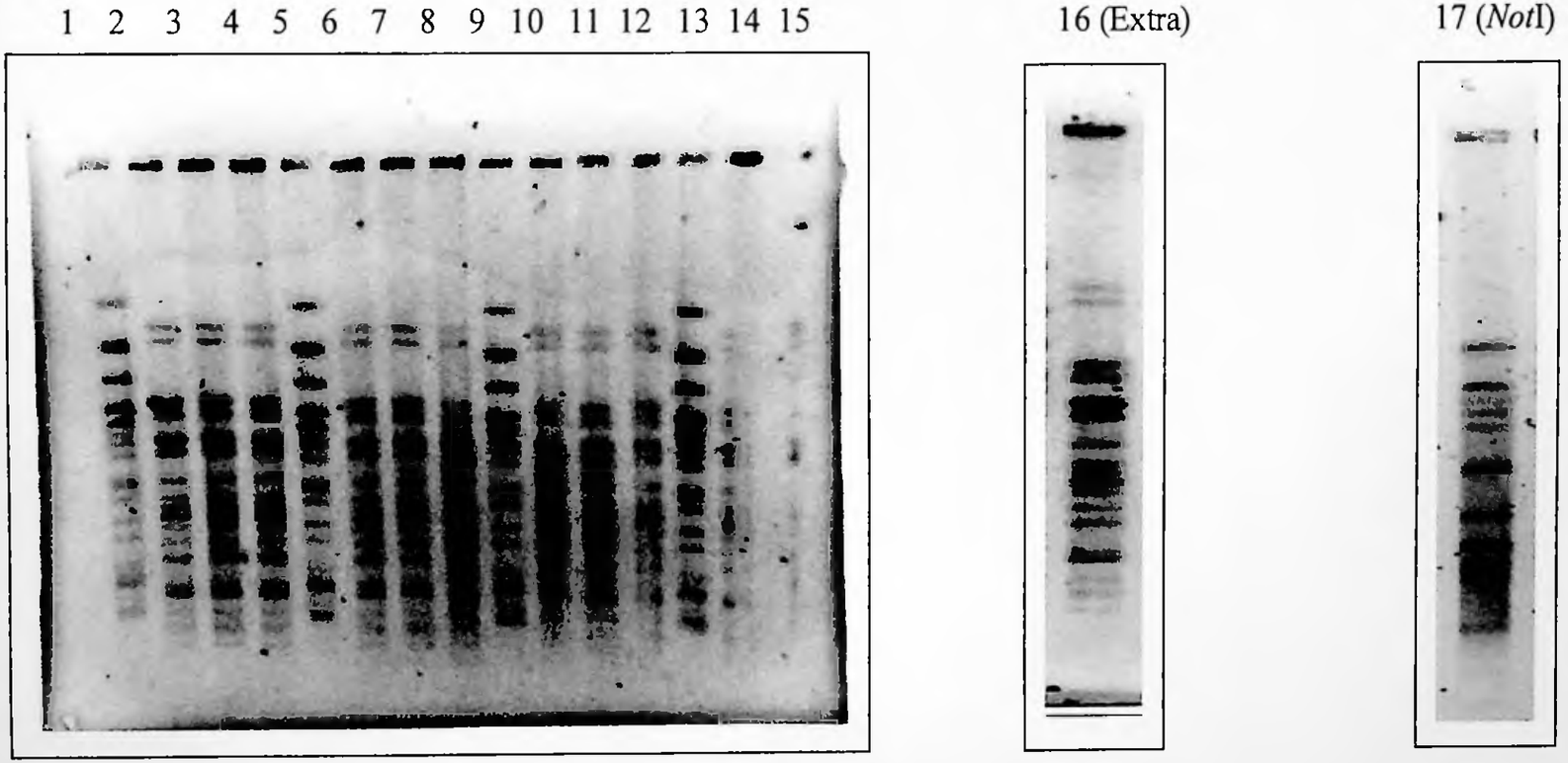

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 (Extra)
17 (*Not*I)

Figure 16. Angus Cow 3319905 (A3-5) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on October 1, 1999. Lane 1: Standard (G5244), Lane 2: A3-5 colony 1, Lane 3: A3-5 colony 2, Lane 4: A3-5 colony 3, Lane 5: Standard (G5244), Lane 6: A3-5 colony 4, Lane 7: A3-5 colony 5, Lane 8: A3-5 colony 6, Lane 9: Standard (G5244), Lane 10: A3-5 colony 7, Lane 11: A3-5 colony 8, Lane 12: A3-5 colony 9, Lane 13: Standard (G5244), Lane 14: A3-5 colony 10, Lane 15: Negative Control.

Figure 17. Extra DNA Profiles (re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 11, 1999. Lane 16: A3-5 colony 10.

Figure 18. A3-5 Unique DNA Profile (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 17: A3-5 colony 1.

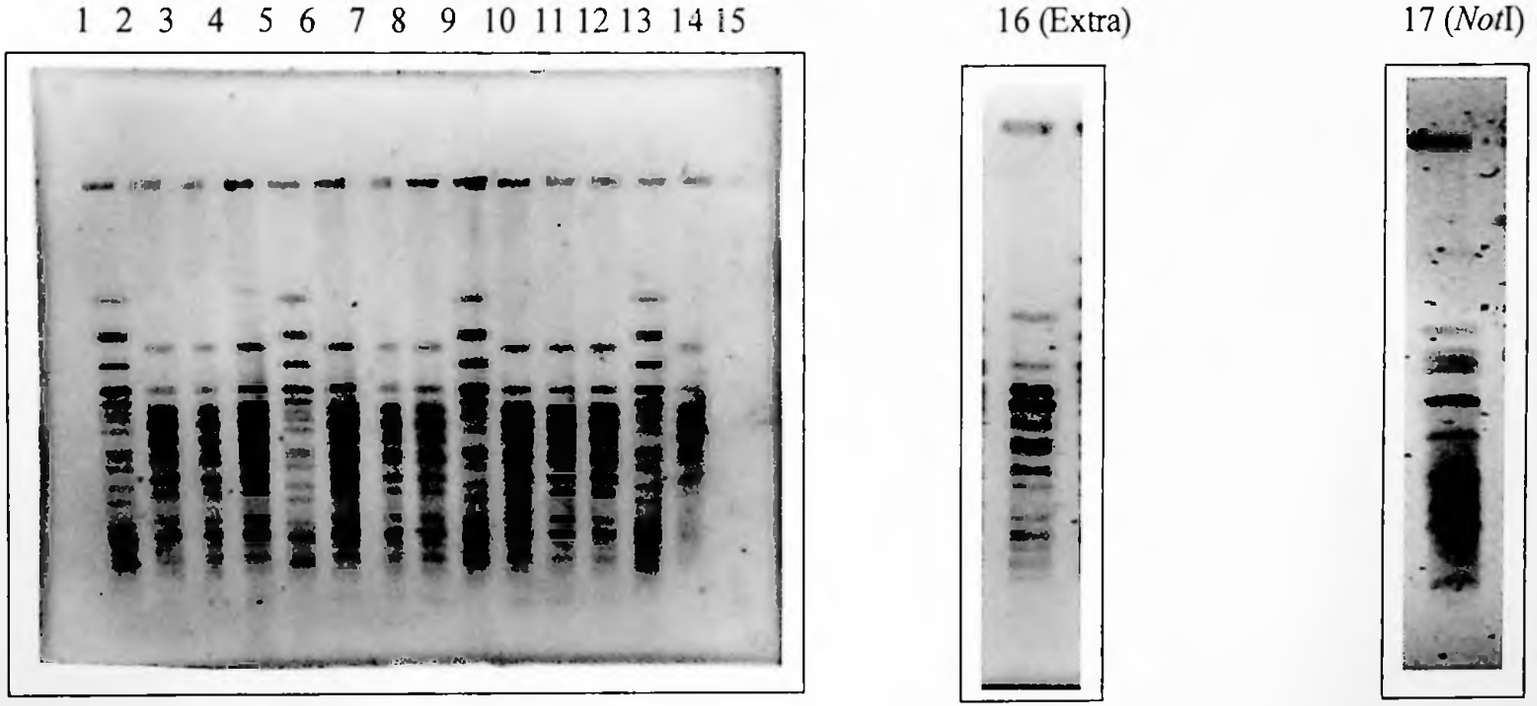
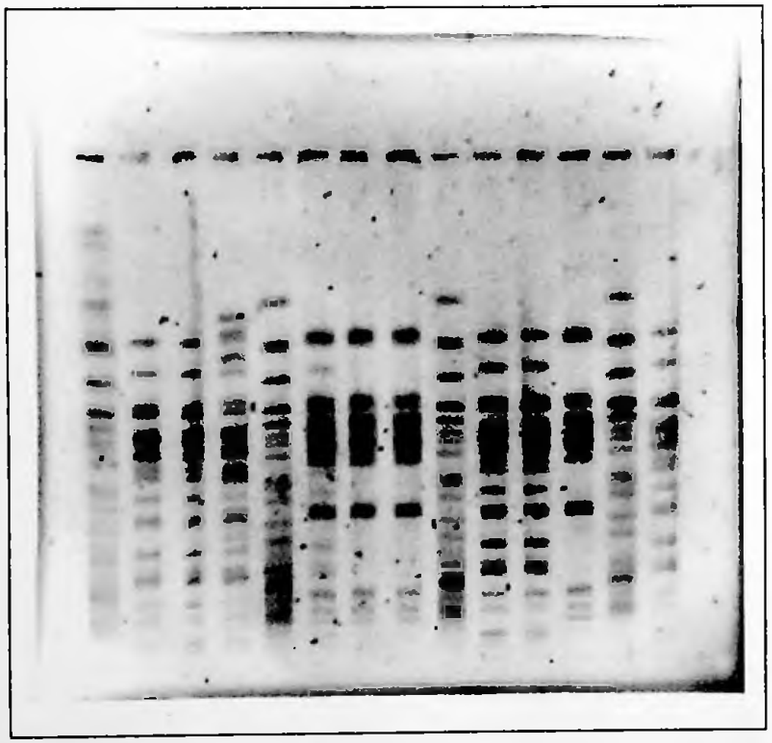

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 (Extra)
17 (*Not*I)

Figure 19. Angus Cow 3179908 (A3-8) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on October 1, 1999. Lane 1: Standard (G5244), Lane 2: A3-8 colony 1, Lane 3: A3-8 colony 2, Lane 4: A3-8 colony 3, Lane 5: Standard (G5244), Lane 6: A3-8 colony 4, Lane 7: A3-8 colony 5, Lane 8: A3-8 colony 6, Lane 9: Standard (G5244), Lane 10: A3-8 colony 7, Lane 11: A3-8 colony 8, Lane 12: A3-8 colony 9, Lane 13: Standard (G5244), Lane 14: A3-8 colony 10, Lane 15: Negative Control.

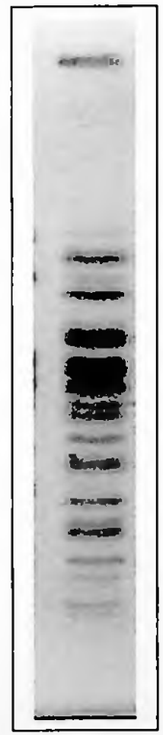
Figure 20. Extra DNA Profile(re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 11, 1999. Lane 16: A3-8 colony 10.

Figure 21. A3-8 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 17: A3-8 colony 1, Lane 18: A3-8 colony 3, Lane 19: A3-8 colony 6.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 (Extra)



17 18 19 (*Not*I)

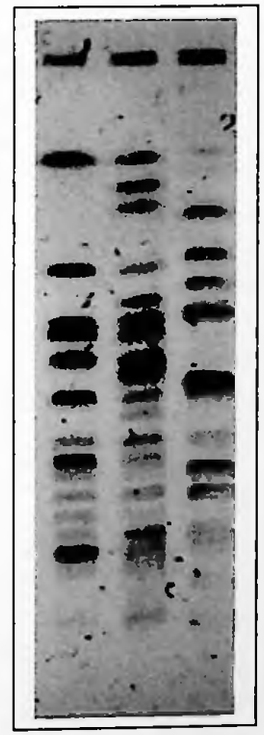


Figure 22. Angus Cow 5059909 (A5-9) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on October 14, 1999. Lane 1: Standard (G5244), Lane 2: A5-9 colony 1, Lane 3: A5-9 colony 2, Lane 4: A5-9 colony 3, Lane 5: Standard (G5244), Lane 6: A5-9 colony 4, Lane 7: A5-9 colony 5, Lane 8: A5-9 colony 6, Lane 9: Standard (G5244), Lane 10: A5-9 colony 7, Lane 11: A5-9 colony 8, Lane 12: A5-9 colony 9, Lane 13: Standard (G5244), Lane 14: A5-9 colony 10, Lane 15: Negative Control.

Figure 23. Extra DNA Profiles (re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 11, 1999. Lane 16: A5-9 colony 2, Lane 17: A5-9 colony 10.

Figure 24. A5-9 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 18: A5-9 colony 3, Lane 19: A5-9 colony 4, Lane 20: A5-9 colony 9.

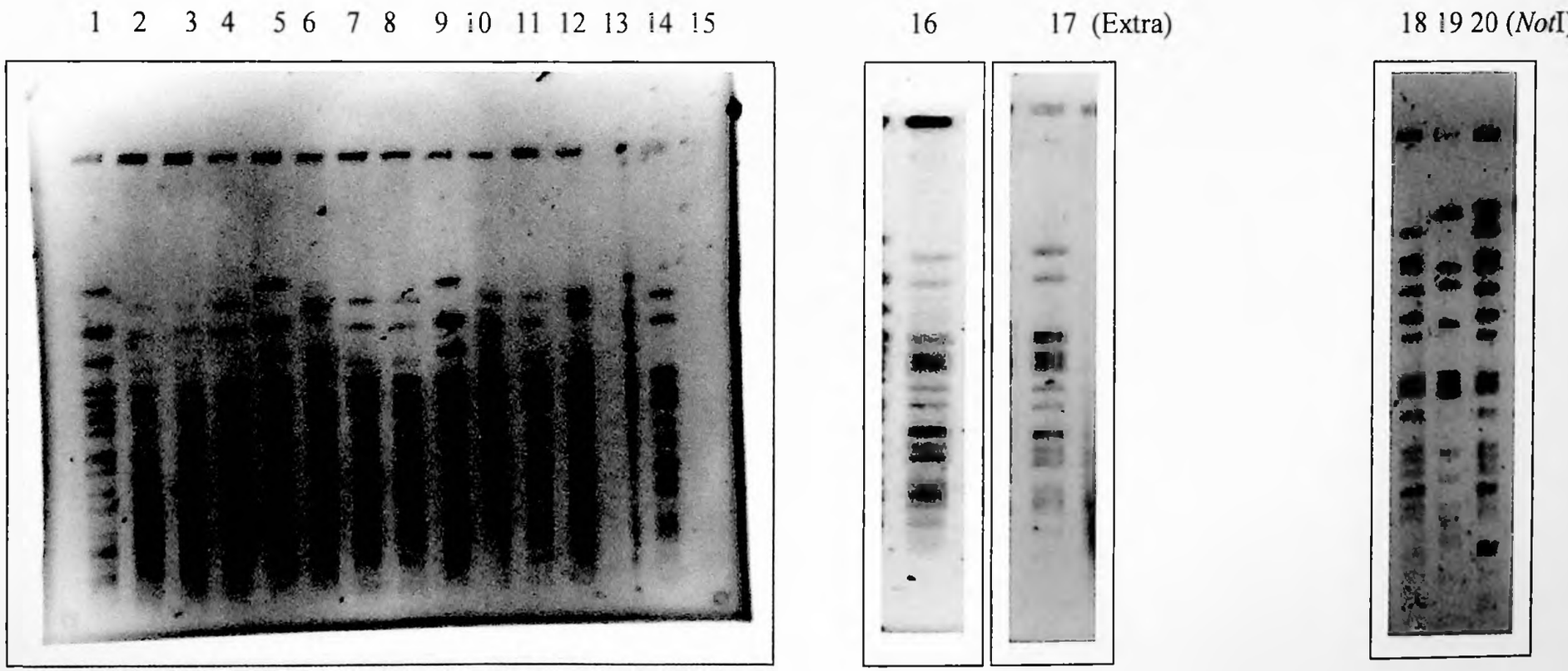

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16
17 (Extra)
18 19 20 (*Not*I)

Figure 25. Summary of Angus Cow Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 1: Lambda Ladder, Lane 2: A3-4 colony 1, Lane 3: A3-5 colony 1, Lane 4: A3-8 colony 1, Lane 5: A3-8 colony 3, Lane 6: A3-8 colony 6, Lane 7: A5-9 colony 3, Lane 8: A5-9 colony 4, Lane 9: A5-9 colony 9, Lane 10: Lambda Ladder, Lane 11: Positive Control (JM107), Lane 12: Negative Control

Figure 26. Lane 13: A3-5 colony 1 with better resolution. Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000.

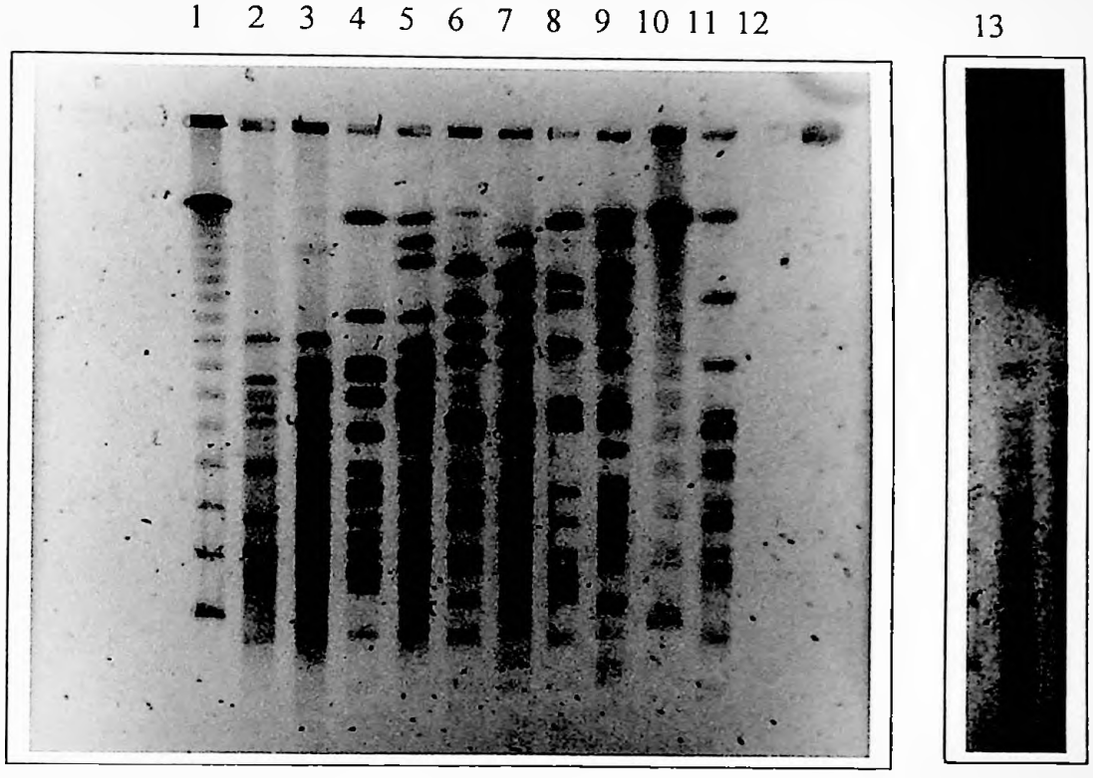
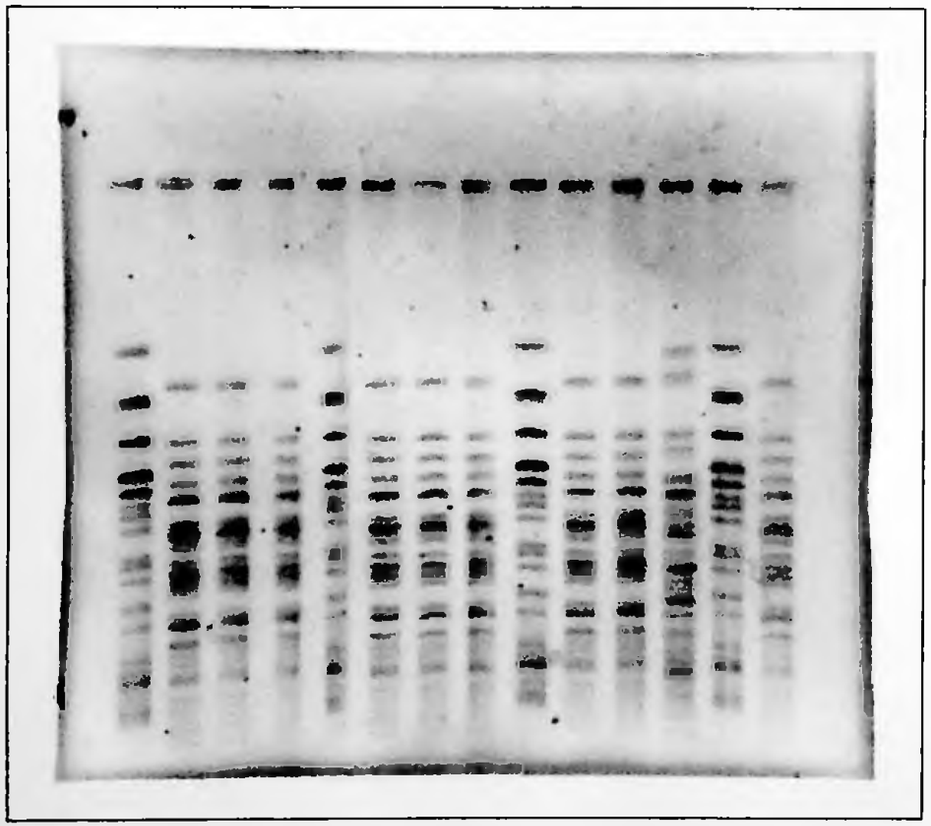

1 2 3 4 5 6 7 8 9 10 11 12
13

Figure 27. Herford Cow 3179910 (H3-0) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on September 30, 1999. Lane 1: Standard (G5244), Lane 2: H3-0 colony 1, Lane 3: H3-0 colony 2, Lane 4: H3-0 colony 3, Lane 5: Standard (G5244), Lane 6: H3-0 colony 4, Lane 7: H3-0 colony 5, Lane 8: H3-0 colony 6, Lane 9: Standard (G5244), Lane 10: H3-0 colony 7, Lane 11: H3-0 colony 8, Lane 12: H3-0 colony 9, Lane 13: Standard (G5244), Lane 14: H3-0 colony 10, Lane 15: Negative Control.

Figure 28. H3-0 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 16: H3-0 colony 1, Lane 17: H3-0 colony 9.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 17 (*Not*I)

Figure 29. Herford Cow 7219903 (H7-3) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 12, 1999. Lane 1: Standard (G5244), Lane 2: H7-3 colony 1, Lane 3: H7-3 colony 2, Lane 4: H7-3 colony 3, Lane 5: Standard (G5244), Lane 6: H7-3 colony 4, Lane 7: H7-3 colony 5, Lane 8: H7-3 colony 6, Lane 9: Standard (G5244), Lane 10: H7-3 colony 7, Lane 11: H7-3 colony 8, Lane 12: H7-3 colony 9, Lane 13: Standard (G5244), Lane 14: H7-3 colony 10, Lane 15: Negative Control.

Figure 30. H7-3 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 16: H7-3 colony 1, Lane 17: H7-3 colony 3, Lane 18: H7-3 colony 5, Lane 19: H7-3 colony 9.

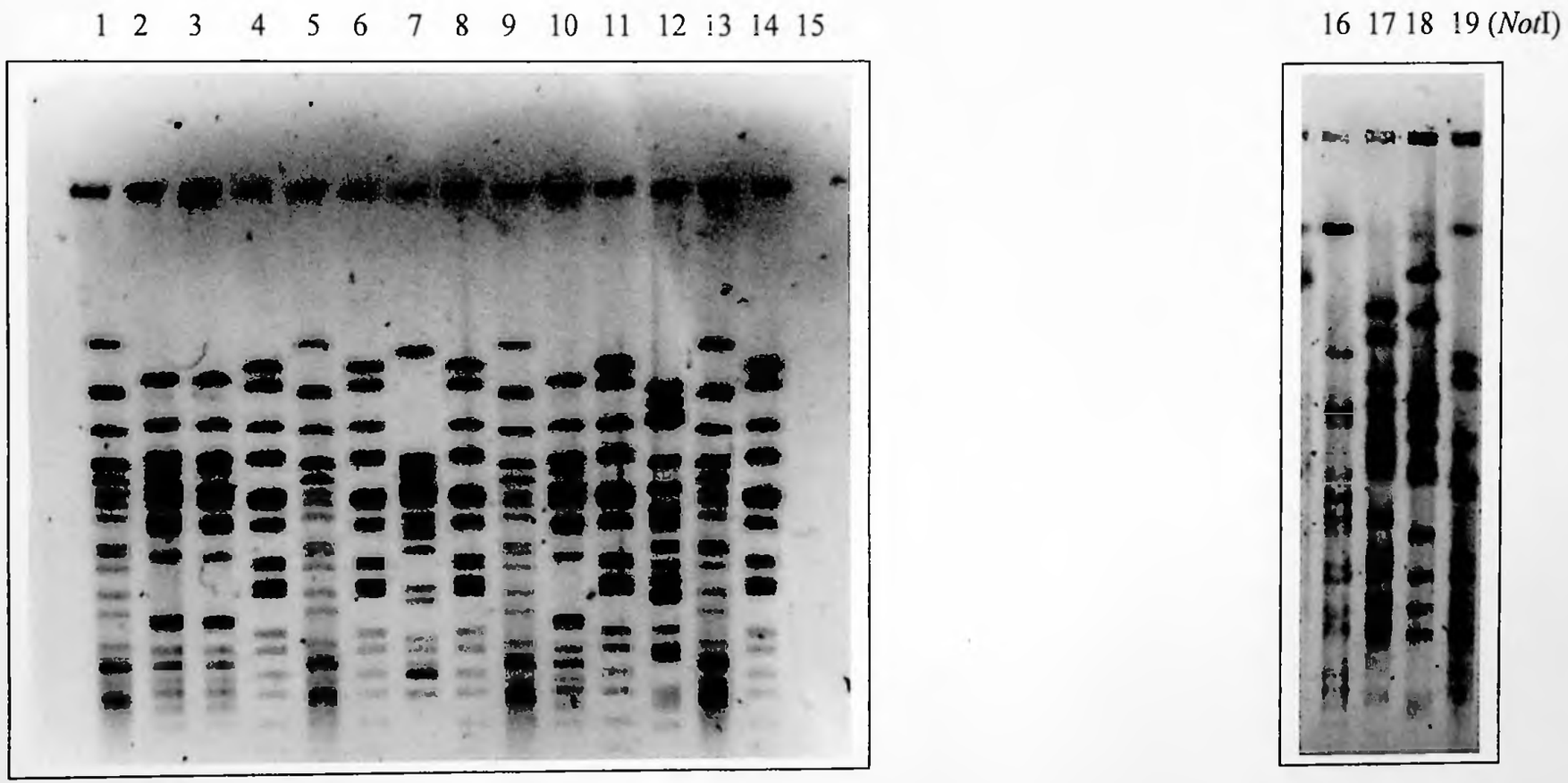

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 17 18 19 (*Not*I)

Figure 31. Herford Cow 7219904 (H7-4) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 27, 1999. Lane 1: Standard (G5244), Lane 2: H7-4 colony 1, Lane 3: H7-4 colony 2, Lane 4: H7-4 colony 3, Lane 5: Standard (G5244), Lane 6: H7-4 colony 4, Lane 7: H7-4 colony 5, Lane 8: H7-4 colony 6, Lane 9: Standard (G5244), Lane 10: H7-4 colony 7, Lane 11: H7-4 colony 8, Lane 12: H7-4 colony 9, Lane 13: Standard (G5244), Lane 14: H7-4 colony 10, Lane 15: Negative Control.

Figure 32. H7-4 Unique DNA Profiles (digested with *Not*I). Lane 16: H7-4 colony 1 Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000.

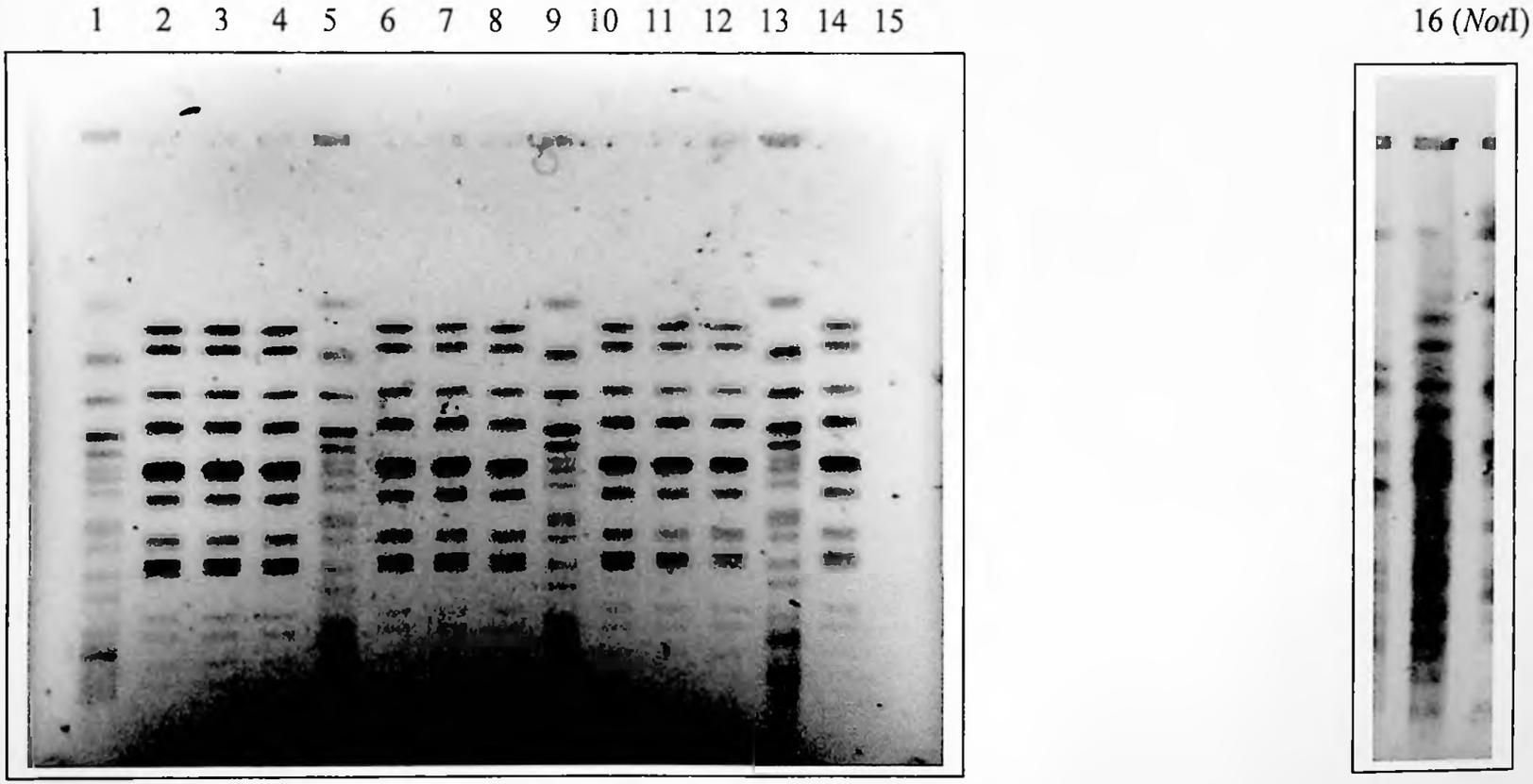
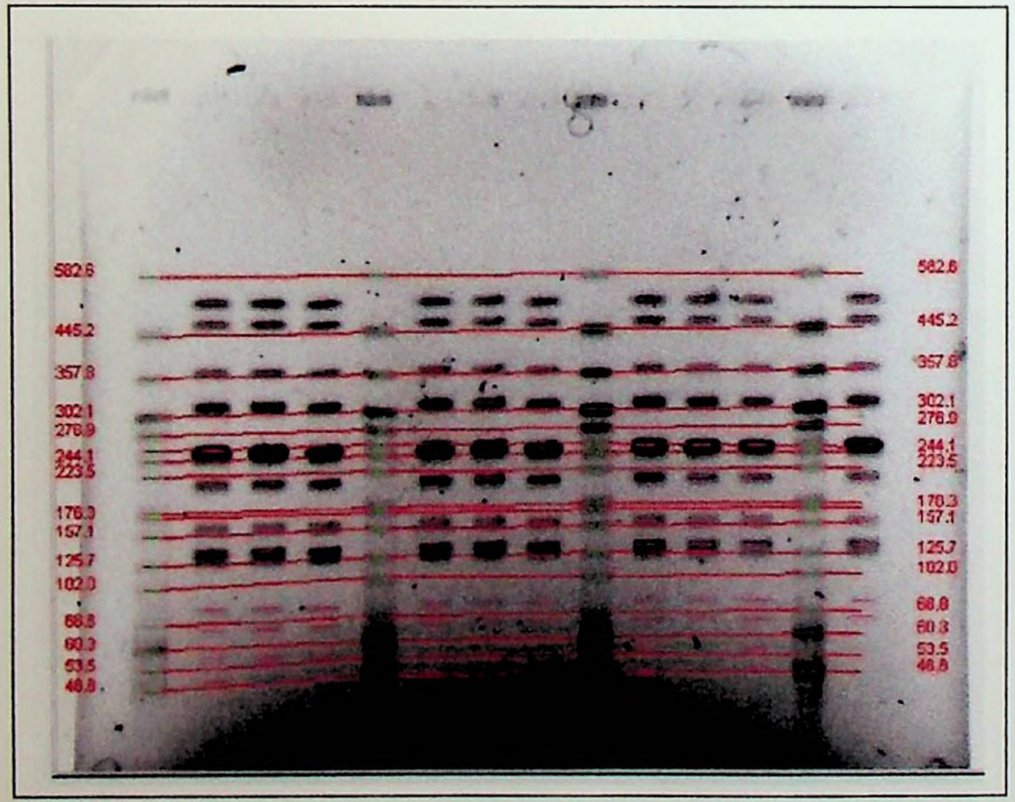

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 (*Not*I)

Figure 33. Same Gel as in Figure 31 but analyzed with Scanalytics software (RFLPscan Forensics Edition Version 3.0).

Figure 34. Same Gel as in Figure 32 but analyzed with Scanalytics software (RFLPscan Forensics Edition Version 3.0).

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 (*Not*I)

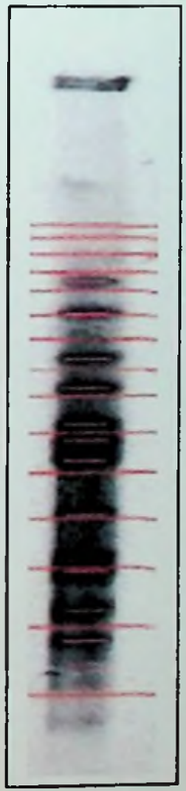
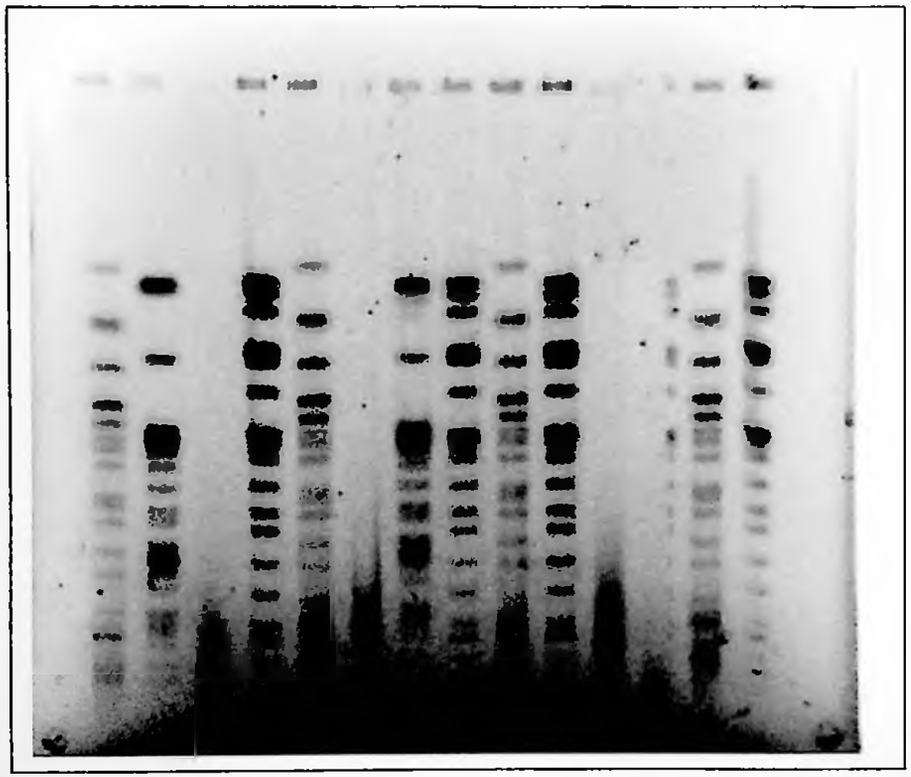


Figure 35. Herford Cow 7219909 (H7-9) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on December 5, 1999. Lane 1: Standard (G5244), Lane 2: H7-9 colony 1, Lane 3: H7-9 colony 2, Lane 4: H7-9 colony 3, Lane 5: Standard (G5244), Lane 6: H7-9 colony 4, Lane 7: H7-9 colony 5, Lane 8: H7-9 colony 6, Lane 9: Standard (G5244), Lane 10: H7-9 colony 7, Lane 11: H7-9 colony 8, Lane 12: H7-9 colony 9, Lane 13: Standard (G5244), Lane 14: H7-9 colony 10, Lane 15: Negative Control.

Figure 36. H7-9 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 16: H7-9 colony 1, Lane 17: H7-9 colony 3.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 17 (*Not*I)

Figure 37. Summary of Herford Cow Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 26, 2000. Lane 1: Lambda Ladder, Lane 2: H3-0 colony 1, Lane 3: H3-0 colony 9, Lane 4: H7-3 colony 1, Lane 5: H7-3 colony 3, Lane 6: H7-3 colony 5, Lane 7: H7-3 colony 9, Lane 8: H7-4 colony 1, Lane 9: H7-9 colony 1, Lane 10: H7-9 colony 3, Lane 11: Lambda Ladder, Lane 12: Positive Control (JM107), Lane 13: Negative Control.

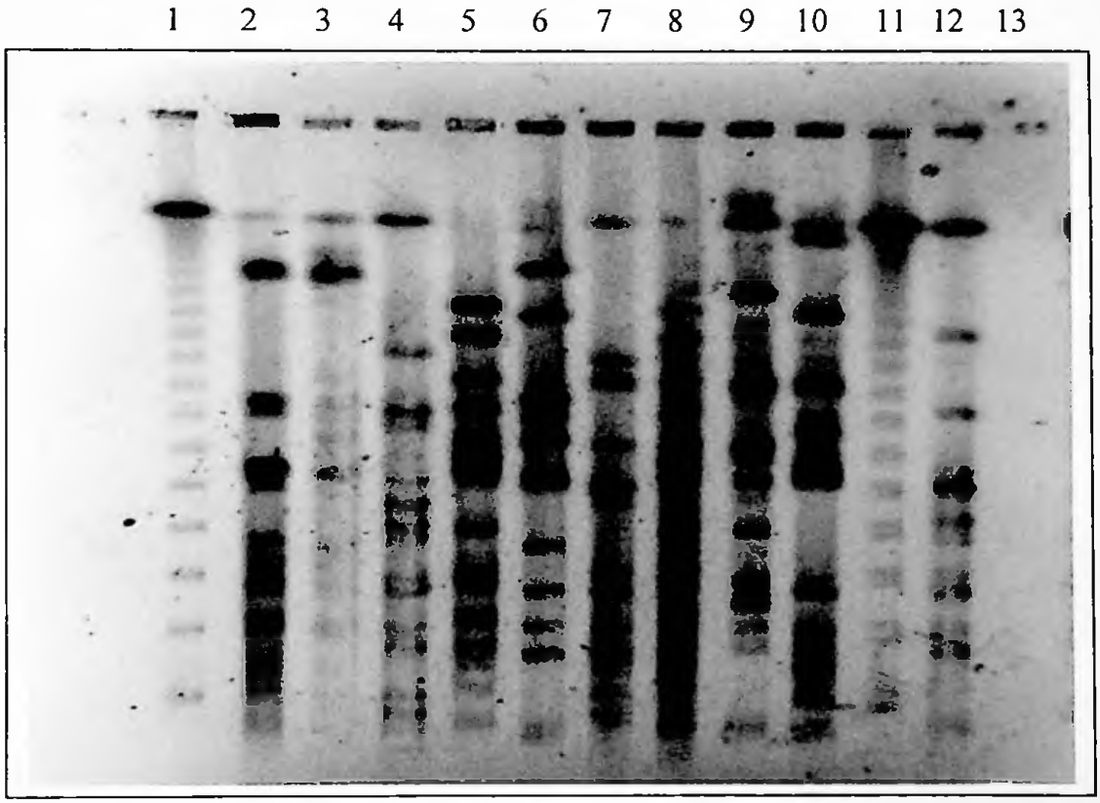
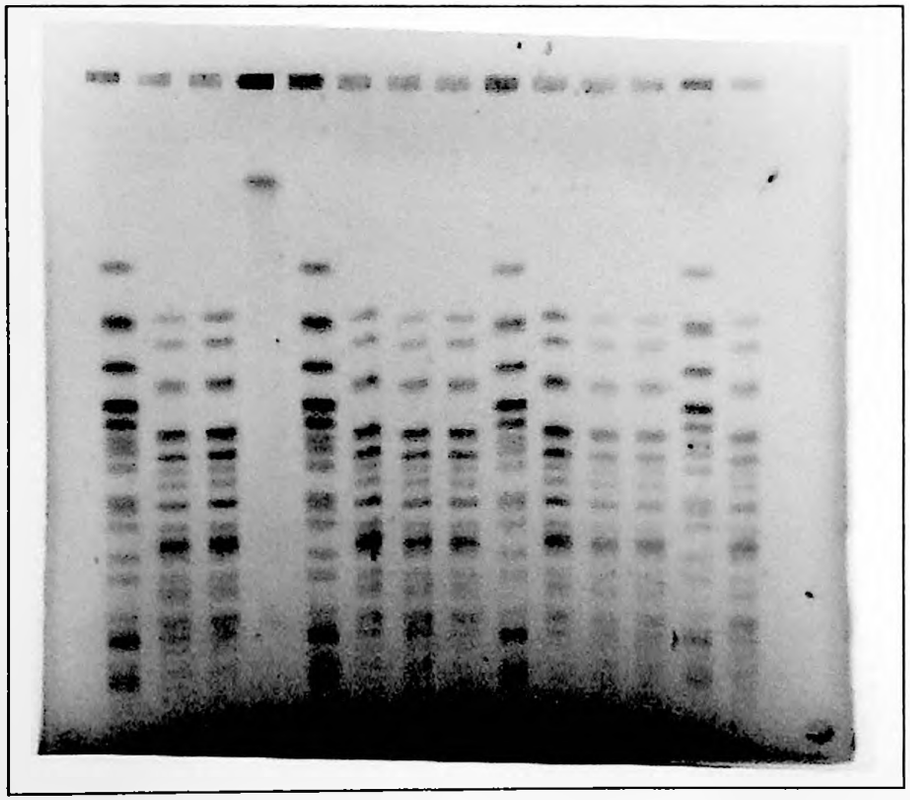

1 2 3 4 5 6 7 8 9 10 11 12 13

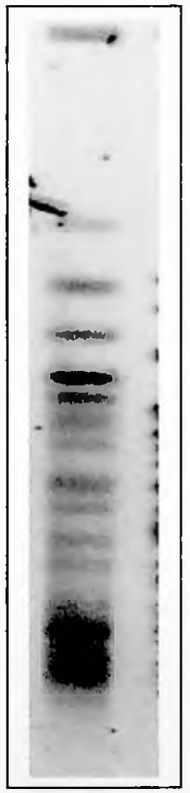
Figure 38. Charlais Cow 10069904 (CL1-4) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on December 4, 1999. Lane 1: Standard (G5244), Lane 2: CL1-4 colony 1, Lane 3: CL1-4 colony 2, Lane 4: CL1-4 colony 3, Lane 5: Standard (G5244), Lane 6: CL1-4 colony 4, Lane 7: CL1-4 colony 5, Lane 8: CL1-4 colony 6, Lane 9: Standard (G5244), Lane 10: CL1-4 colony 7, Lane 11: CL1-4 colony 8, Lane 12: CL1-4 colony 9, Lane 13: Standard (G5244), Lane 14: CL1-4 colony 10, Lane 15: Negative Control.

Figure 39. CL1-4 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 16: CL1-4 colony 4, Lane 17: CL1-4 colony 8.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 (Extra)



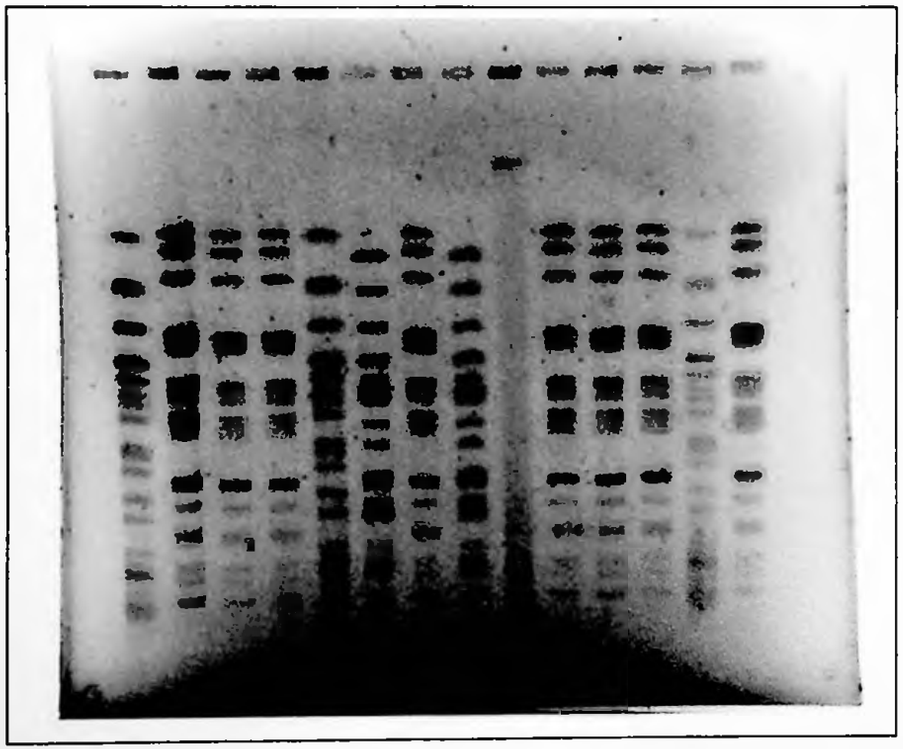
17 (*Not*I)

Figure 40. Charlais Cow 7219902 (CL7-2) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 28, 1999. Lane 1: Standard (G5244), Lane 2: CL7-2 colony 1, Lane 3: CL7-2 colony 2, Lane 4: CL7-2 colony 3, Lane 5: Standard (G5244), Lane 6: CL7-2 colony 4, Lane 7: CL7-2 colony 5, Lane 8: CL7-2 colony 6, Lane 9: Standard (G5244), Lane 10: CL7-2 colony 7, Lane 11: CL7-2 colony 8, Lane 12: CL7-2 colony 9, Lane 13: Standard (G5244), Lane 14: CL7-2 colony 10, Lane 15: Negative Control.

Figure 41. Extra DNA Profiles (re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 17, 1999. Lane 16: CL7-2 colony 3.

Figure 42. CL7-2 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 17: CL7-2 colony 1

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



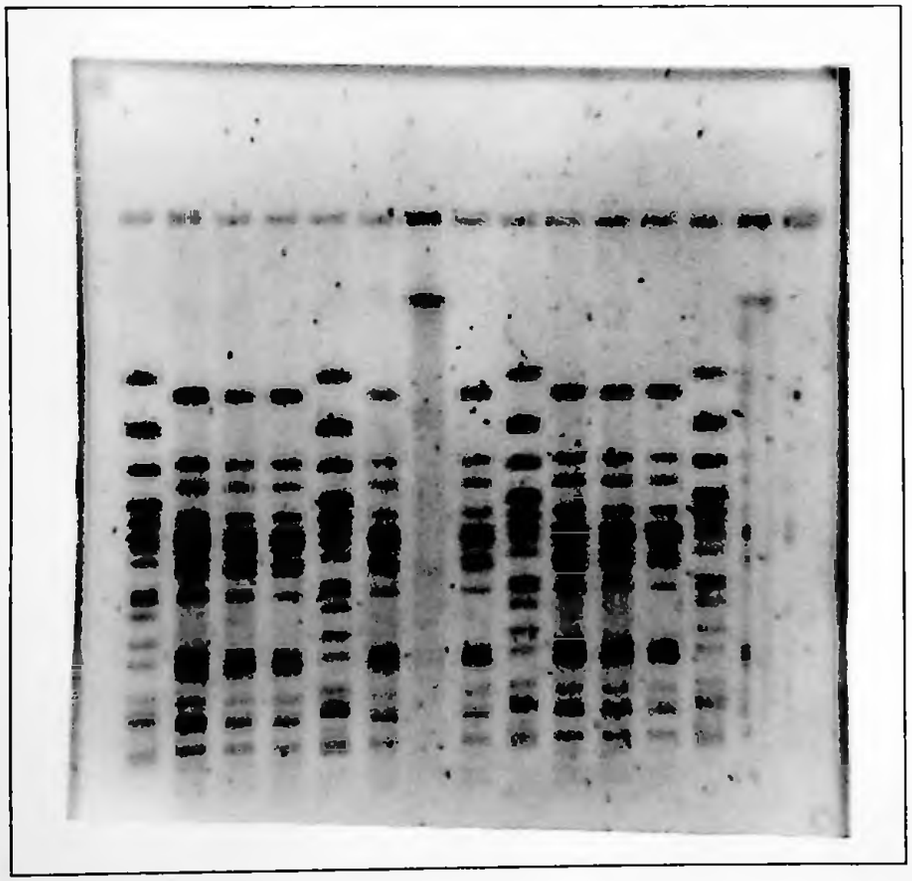
16 17 (*Not*I)

Figure 43. Charlais Cow 5059913 (CL5-3) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on September 23, 1999. Lane 1: Standard (G5244), Lane 2: CL5-3 colony 1, Lane 3: CL5-3 colony 2, Lane 4: CL5-3 colony 3, Lane 5: Standard (G5244), Lane 6: CL5-3 colony 4, Lane 7: CL5-3 colony 5, Lane 8: CL5-3 colony 6, Lane 9: Standard (G5244), Lane 10: CL5-3 colony 7, Lane 11: CL5-3 colony 8, Lane 12: CL5-3 colony 9, Lane 13: Standard (G5244), Lane 14: CL5-3 colony 10, Lane 15: Negative Control.

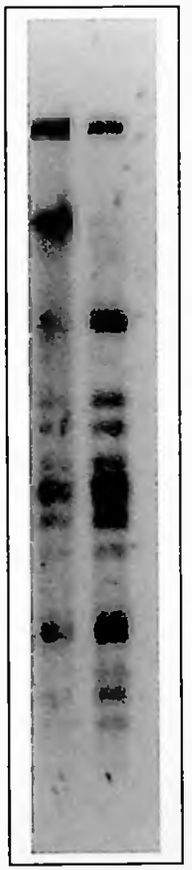
Figure 44. Extra DNA Profiles(re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 17, 1999. Lane 16: CL5-3 colony 5, Lane 17: CL5-3 colony 10.

Figure 45. CL5-3 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 18: CL5-3 colony 2.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 17 (Extra)



18 (*Not*I)

Figure 46: Summary of Charlais Cow Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 1: Lambda Ladder, Lane 2: CL5-3 colony 2, Lane 3: CL1-4 colony 4, Lane 4: CL1-4 colony 8, Lane 5: CL 7-2 colony 1, Lane 6: Lambda Ladder, Lane 7: Positive Control (JM107), Lane 8: Negative Control.

1 2 3 4 5 6 7 8

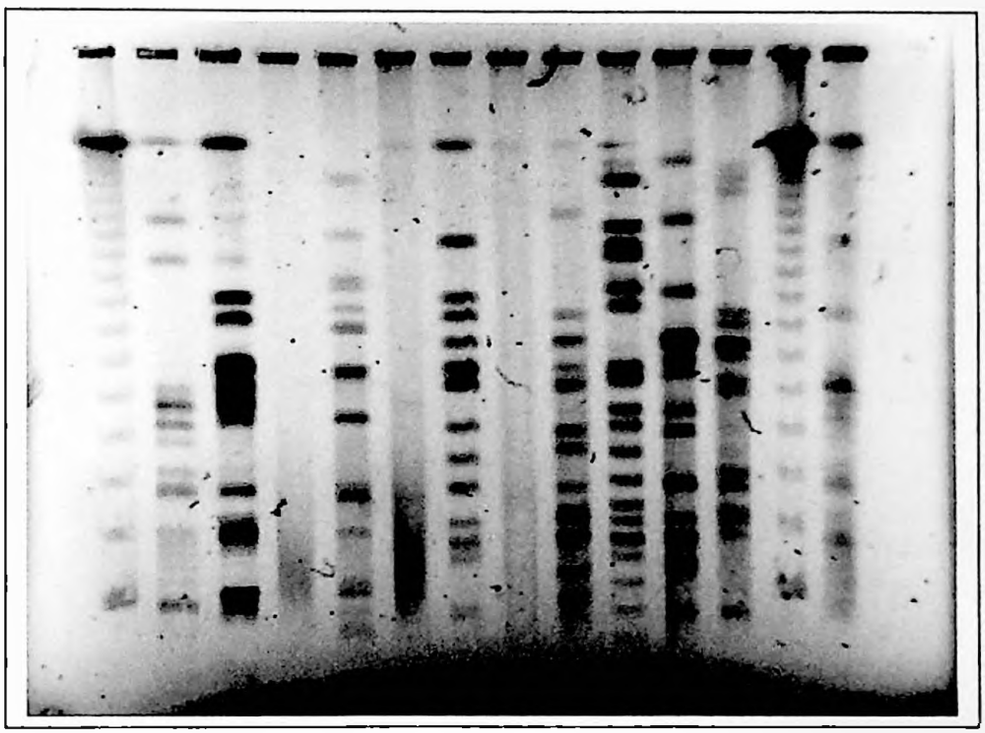
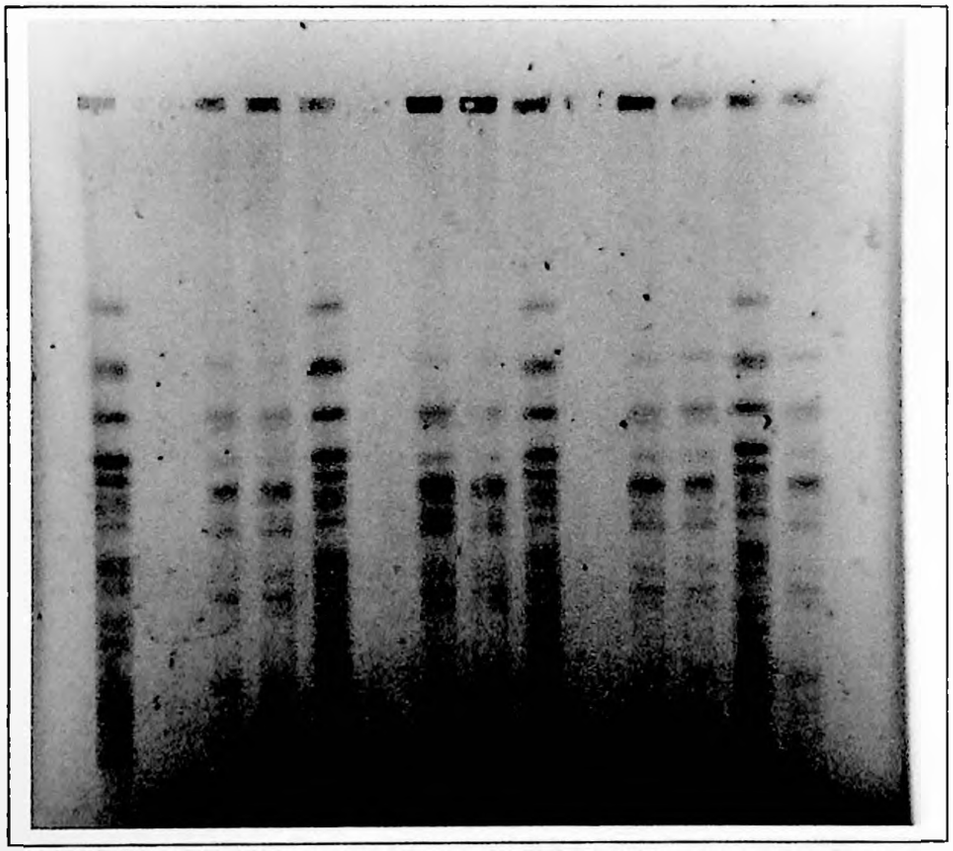


Figure 47. Pensgar Cow 7219905 (P7-5) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on December 5, 1999. Lane 1: Standard (G5244), Lane 2: P7-5 colony 1, Lane 3: P7-5 colony 2, Lane 4: P7-5 colony 3, Lane 5: Standard (G5244), Lane 6: P7-5 colony 4, Lane 7: P7-5 colony 5, Lane 8: P7-5 colony 6, Lane 9: Standard (G5244), Lane 10: P7-5 colony 7, Lane 11: P7-5 colony 8, Lane 12: P7-5 colony 9, Lane 13: Standard (G5244), Lane 14: P7-5 colony 10, Lane 15: Negative Control.

Figure 48. P7-5 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 16: P7-5 colony 2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

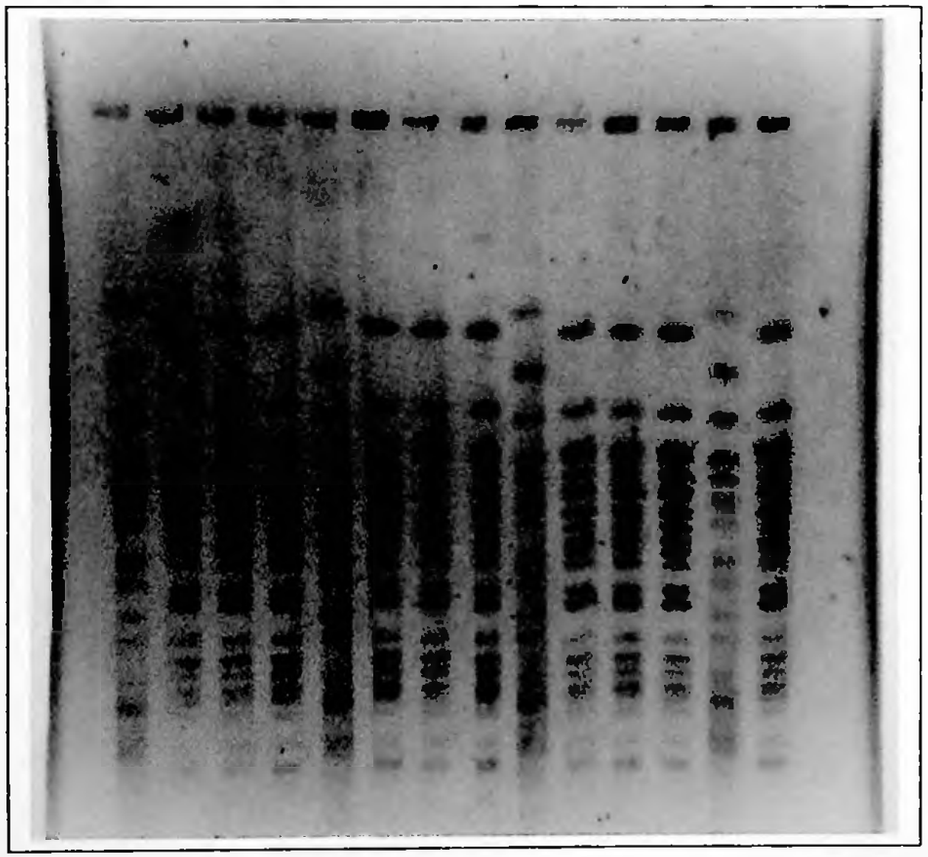


16 (*Not*I)

Figure 49. Pensgar Cow 7219907 (P7-7) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 16, 1999. Lane 1: Standard (G5244), Lane 2: P7-7 colony 1, Lane 3: P7-7 colony 2, Lane 4: P7-7 colony 3, Lane 5: Standard (G5244), Lane 6: P7-7 colony 4, Lane 7: P7-7 colony 5, Lane 8: P7-7 colony 6, Lane 9: Standard (G5244), Lane 10: P7-7 colony 7, Lane 11: P7-7 colony 8, Lane 12: P7-7 colony 9, Lane 13: Standard (G5244), Lane 14: P7-7 colony 10, Lane 15: Negative Control.

Figure 50. P7-7 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 16: P7-7 colony 1, Lane 17: P7-7 colony 4.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 17 (*Not*I)

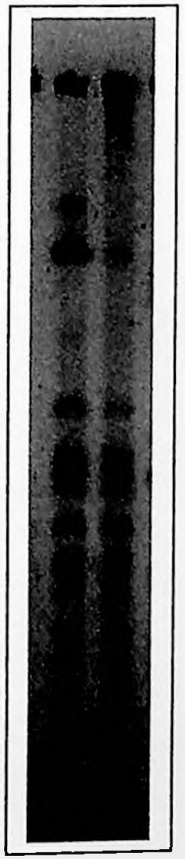


Figure 51. Pensgar Cow 7219908 (P7-8) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 16, 1999. Lane 1: Standard (G5244), Lane 2: P7-8 colony 1, Lane 3: P7-8 colony 2, Lane 4: P7-8 colony 3, Lane 5: Standard (G5244), Lane 6: P7-8 colony 4, Lane 7: P7-8 colony 5, Lane 8: P7-8 colony 6, Lane 9: Standard (G5244), Lane 10: P7-8 colony 7, Lane 11: P7-8 colony 8, Lane 12: P7-8 colony 9, Lane 13: Standard (G5244), Lane 14: P7-8 colony 10, Lane 15: Negative Control.

Figure 52. P7-8 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 16: P7-8 colony 1, Lane 17: P7-8 colony 2.

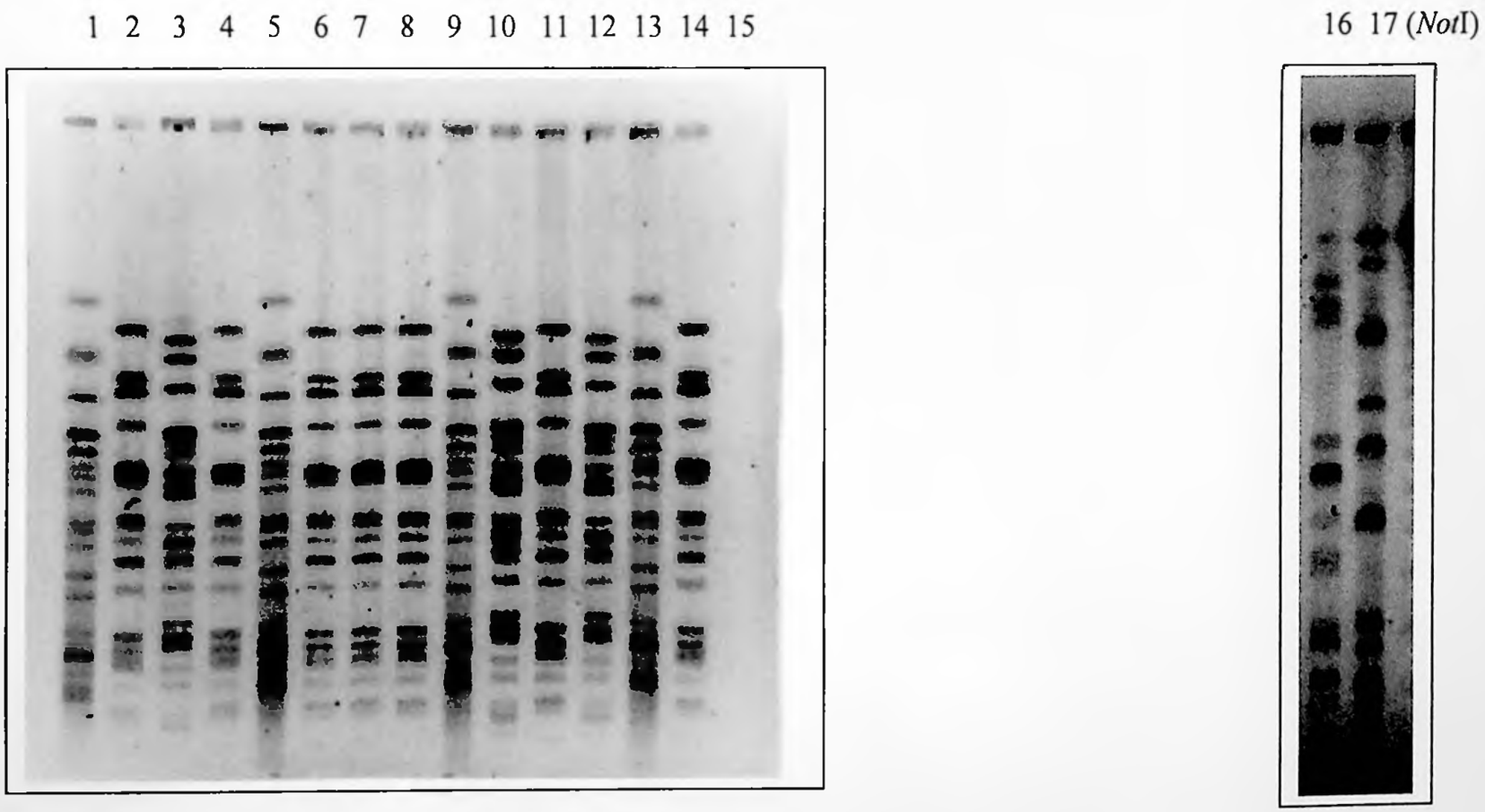

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16 17 (*Not*I)

Figure 53. Summary of Pensgar Cow Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 1: Lambda Ladder, Lane 2: P7-7 colony 1, Lane 3: P7-7 colony 4, Lane 4: P7-8 colony 1, Lane 5: P7-8 colony 2, Lane 6: Lambda Ladder, Lane 7: Positive Control (JM107), Lane 8: Negative Control.

Figure 54. Lane 9: P7-5 colony 2 with better resolution (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 9: P7-5 colony 2

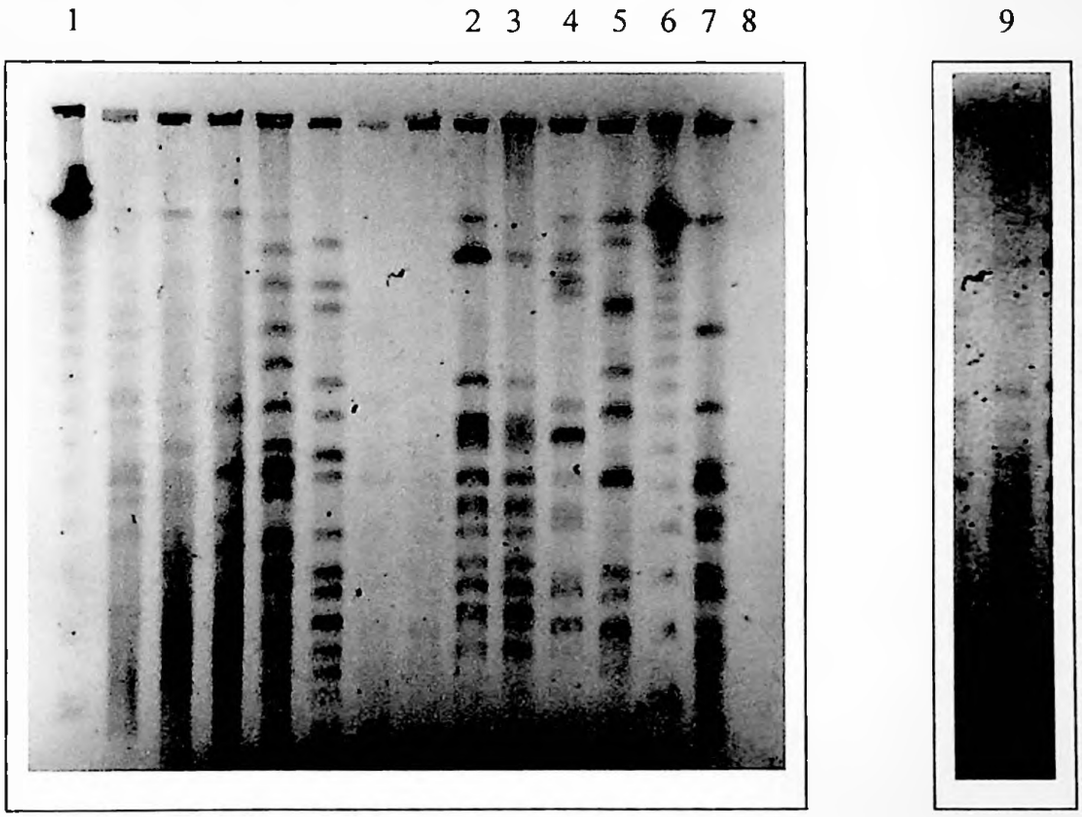
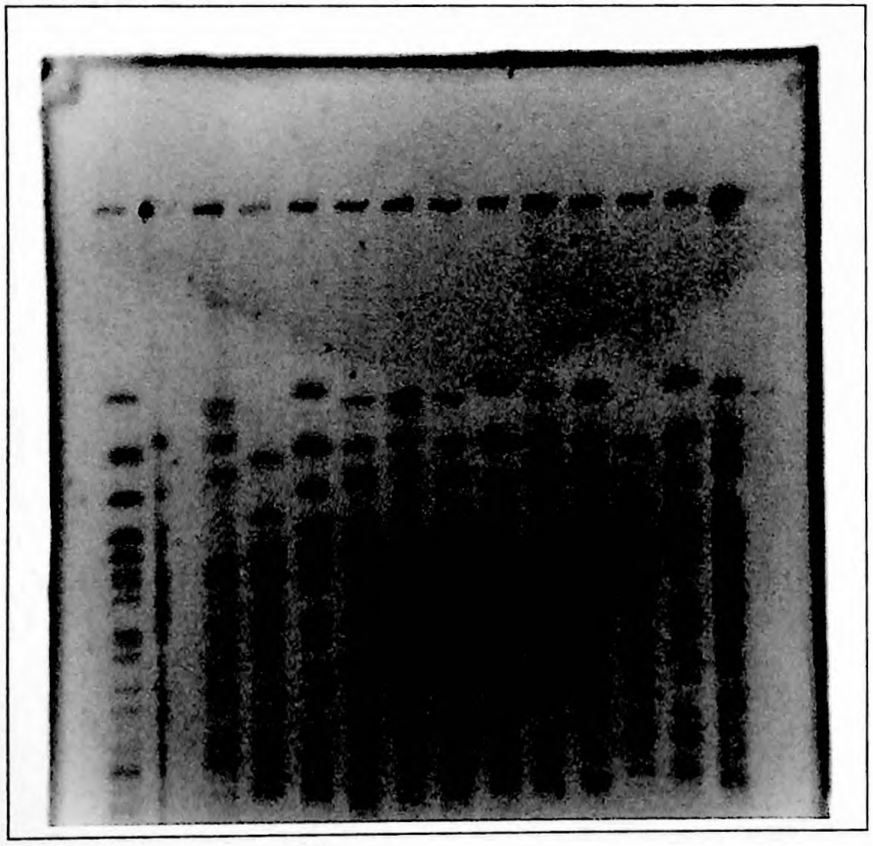

1
2 3 4 5 6 7 8
9

Figure 55. Deer 9806 (D9-6) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on September 24, 1999. Lane 1: Standard (G5244), Lane 2: D9-6 colony 1, Lane 3: D9-6 colony 2, Lane 4: D9-6 colony 3, Lane 5: Standard (G5244), Lane 6: D9-6 colony 4, Lane 7: D9-6 colony 5, Lane 8: D9-6 colony 6, Lane 9: Standard (G5244), Lane 10: D9-6 colony 7, Lane 11: D9-6 colony 8, Lane 12: D9-6 colony 9, Lane 13: Standard (G5244), Lane 14: D9-6 colony 10, Lane 15: Negative Control.

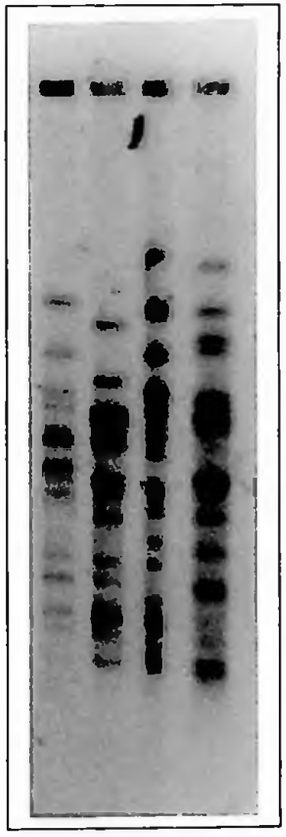
Figure 56. Extra DNA Profiles(re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 11, 1999. Lane 16: D9-6 colony 1, Lane 17: D9-6 colony 9, Lane 18: D9-6 colony 10.

Figure 57. D9-6 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 19: D9-6 colony 3, Lane 20: D9-6 colony 4, Lane 21: D9-6 colony 5.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 17 18 (Extra)

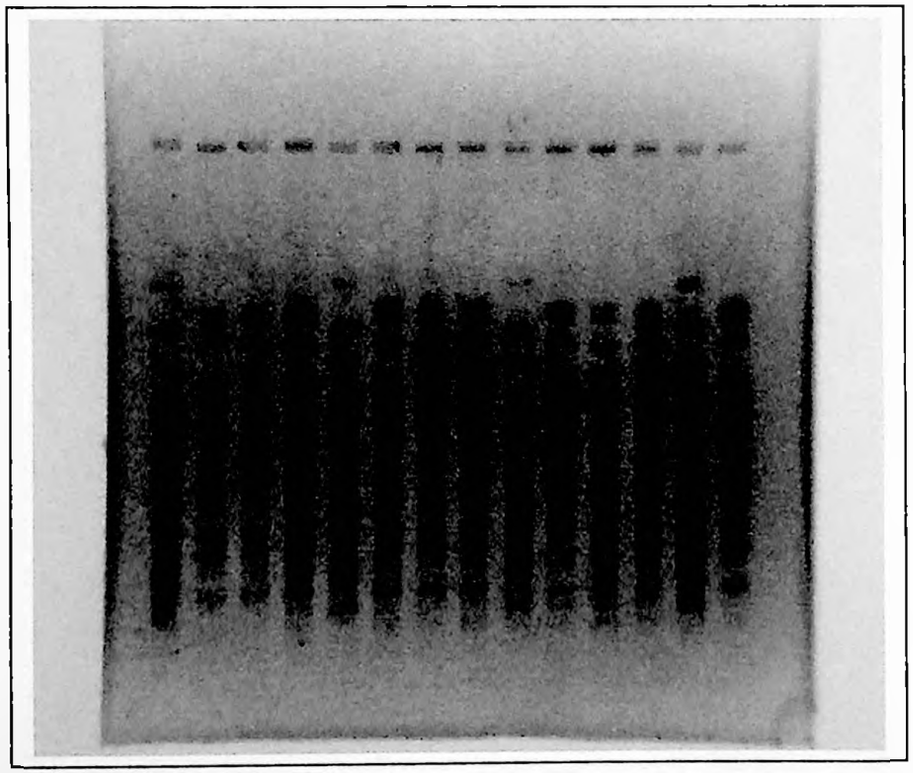


19 20 21 (*Not*I)

Figure 58. Deer 9807 (D9-7) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on October 16, 1999. Lane 1: Standard (G5244), Lane 2: D9-7 colony 1, Lane 3: D9-7 colony 2, Lane 4: D9-7 colony 3, Lane 5: Standard (G5244), Lane 6: D9-7 colony 4, Lane 7: D9-7 colony 5, Lane 8: D9-7 colony 6, Lane 9: Standard (G5244), Lane 10: D9-7 colony 7, Lane 11: D9-7 colony 8, Lane 12: D9-7 colony 9, Lane 13: Standard (G5244), Lane 14: D9-7 colony 10, Lane 15: Negative Control.

Figure 59. D9-7 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 16: D9-7 colony 1.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 (*Not*I)

Figure 60. Deer 7199901 (D7-1) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on December 4, 1999. Lane 1: Standard (G5244), Lane 2: D7-1 colony 1, Lane 3: D7-1 colony 2, Lane 4: D7-1 colony 3, Lane 5: Standard (G5244), Lane 6: D7-1 colony 4, Lane 7: D7-1 colony 5, Lane 8: D7-1 colony 6, Lane 9: Standard (G5244), Lane 10: D7-1 colony 7, Lane 11: D7-1 colony 8, Lane 12: D7-1 colony 9, Lane 13: Standard (G5244), Lane 14: D7-1 colony 10, Lane 15: Negative Control.

Figure 61. D7-1 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 16: D7-1 colony 1, Lane 17: D7-1 colony 2.

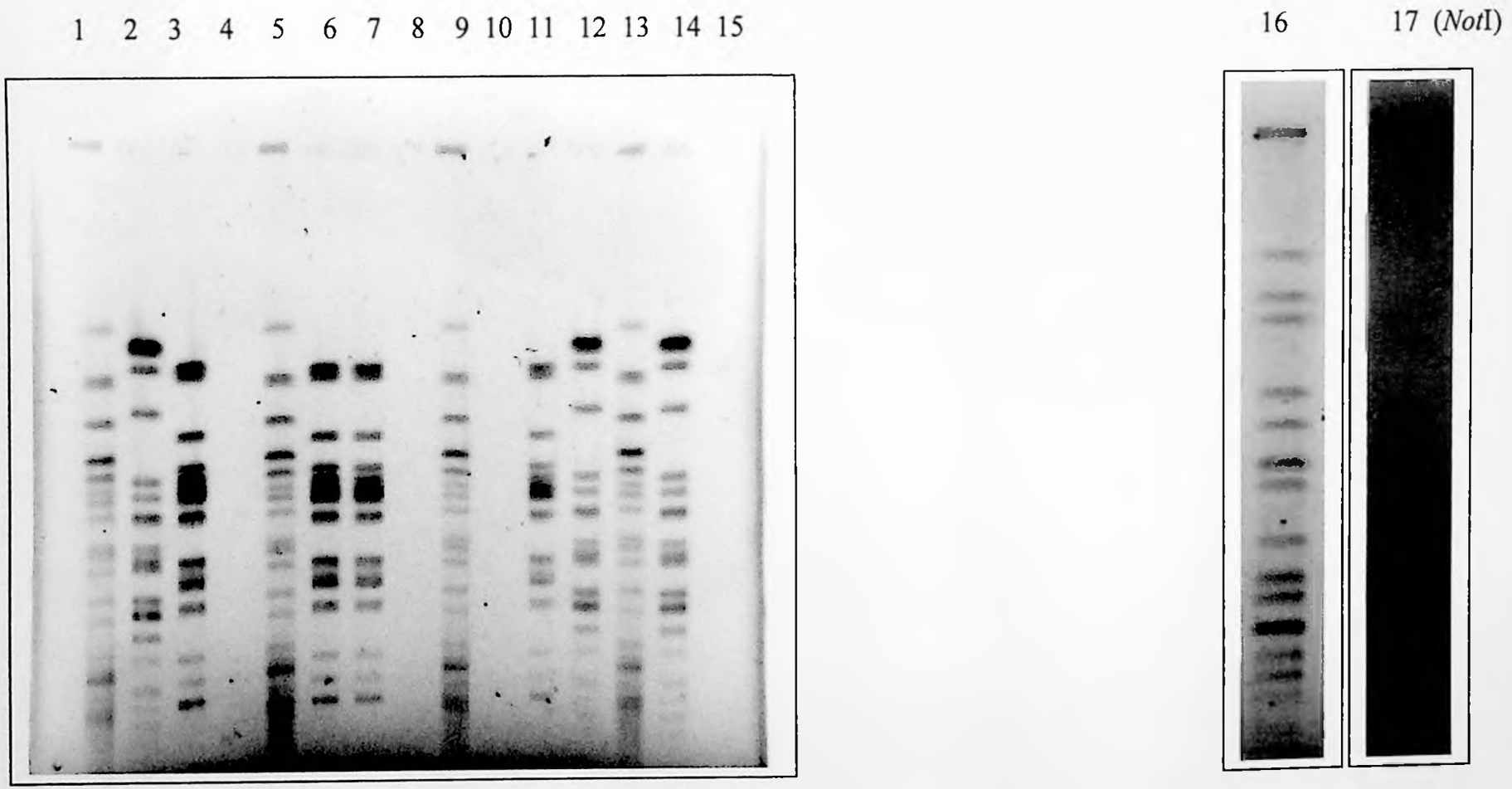

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
16
17 (NotI)

Figure 62. Summary of Deer Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 1: Lambda Ladder, Lane 2: D9-6 colony 3, Lane 3: D9-6 colony 4, Lane 4: D9-6 colony 5, Lane 5: D9-7 colony 1, Lane 6: D7-1 colony 1, Lane 7: Lambda Ladder, Lane 8: Positive Control (JM107), Lane 9: Negative Control.

Figure 63. Summary of Deer Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 10: D7-1 colony 2

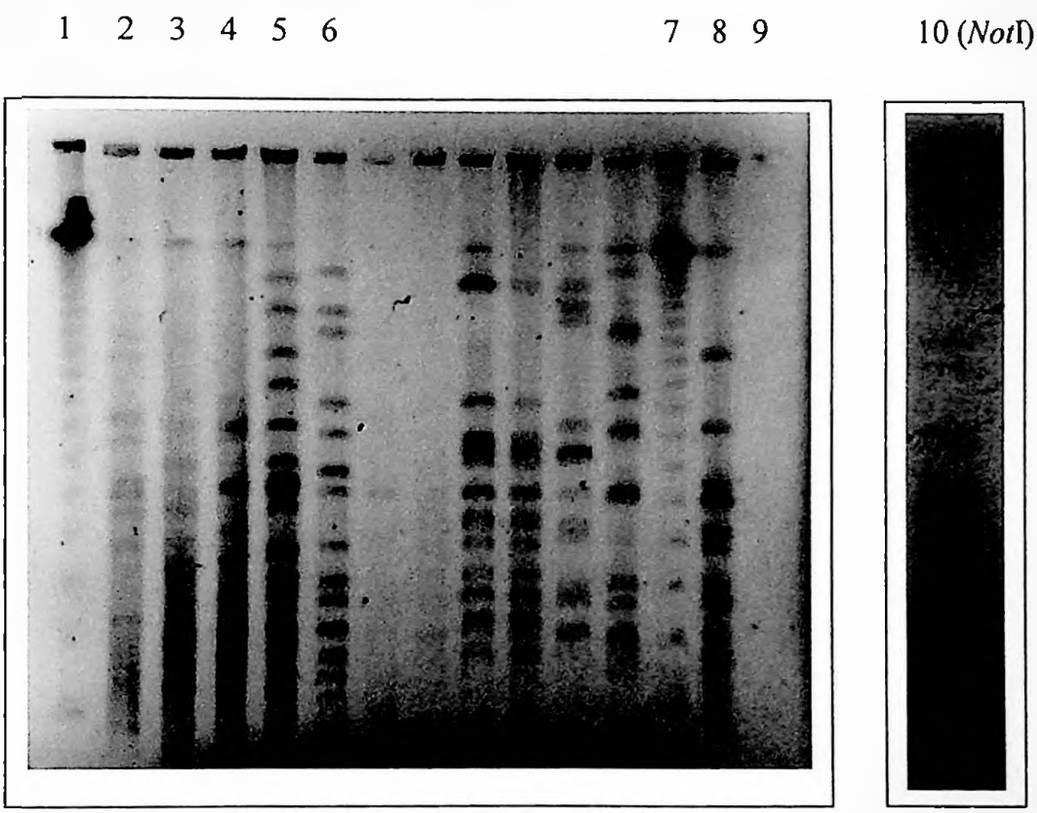
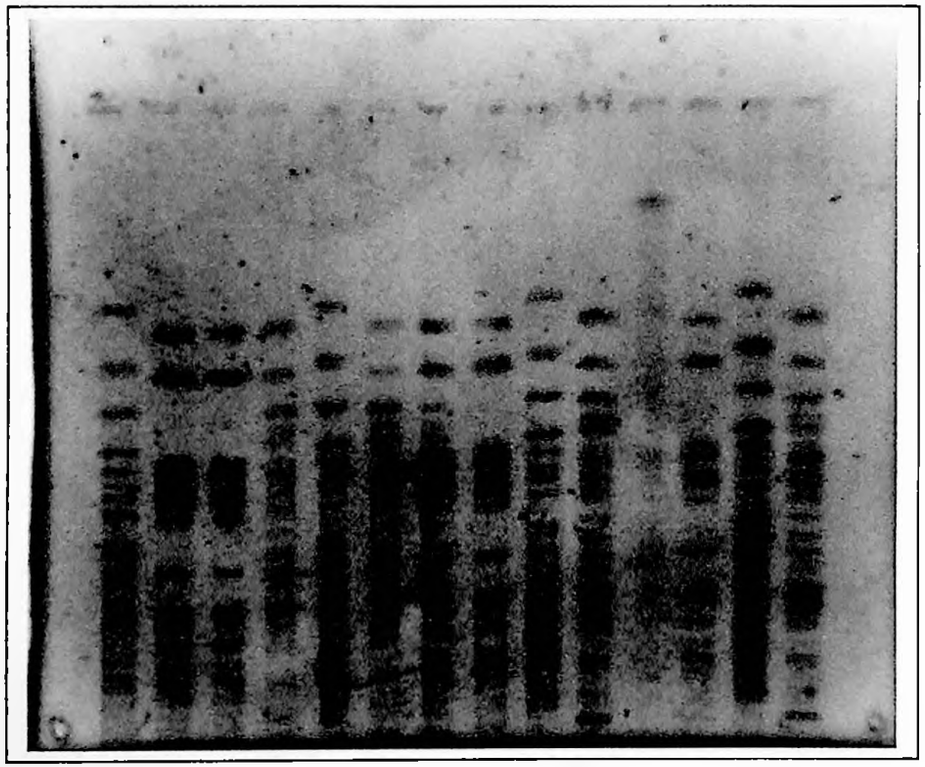

1 2 3 4 5 6
7 8 9
10 (*Not*I)

Figure 64. Chicken 22099902 (CK 2-2) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on October 5, 1999. Lane 1: Standard (G5244), Lane 2: CK2-2 colony 1, Lane 3: CK2-2 colony 2, Lane 4: CK2-2 colony 3, Lane 5: Standard (G5244), Lane 6: CK2-2 colony 4, Lane 7: CK2-2 colony 5, Lane 8: CK2-2 colony 6, Lane 9: Standard (G5244), Lane 10: CK2-2 colony 7, Lane 11: CK2-2 colony 8, Lane 12: CK2-2 colony 9, Lane 13: Standard (G5244), Lane 14: CK2-2 colony 10, Lane 15: Negative Control.

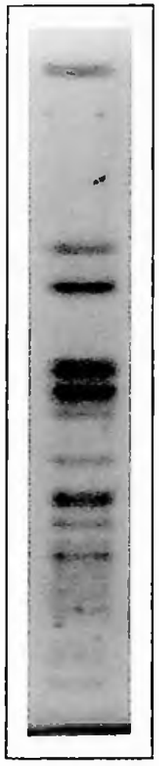
Figure 65. Extra DNA Profiles (re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 11, 1999. Lane 16: CK 2-2 colony 8.

Figure 66. CK2-2 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 17: CK2-2 colony 1, Lane 18: CK2-2 colony 3

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 (Extra)



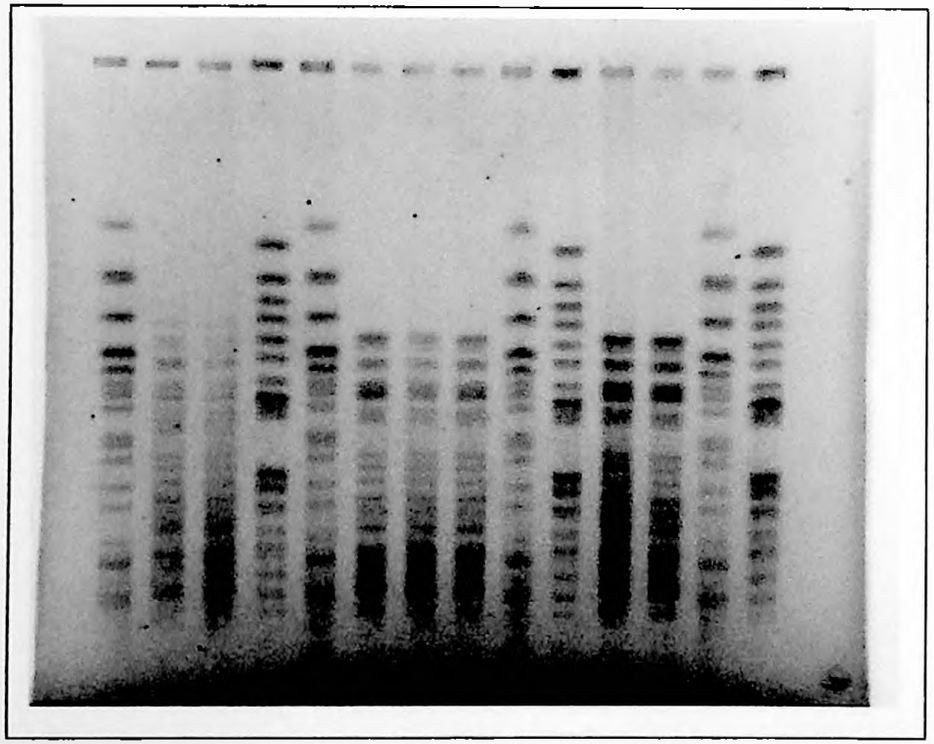
17 18 (*Not*I)

Figure 67. Chicken 7209901 (CK20) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 26, 1999. Lane 1: Standard (G5244), Lane 2: CK20 colony 1, Lane 3: CK20 colony 2, Lane 4: CK20 colony 3, Lane 5: Standard (G5244), Lane 6: CK20 colony 4, Lane 7: CK20 colony 5, Lane 8: CK20 colony 6, Lane 9: Standard (G5244), Lane 10: CK20 colony 7, Lane 11: CK20 colony 8, Lane 12: CK20 colony 9, Lane 13: Standard (G5244), Lane 14: CK20 colony 10, Lane 15: Negative Control.

Figure 68. Extra DNA Profiles (re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 11, 1999. Lane 16: CK20 colony 2.

Figure 69. CK20 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 17: CK20 colony 1, Lane 18: CK20 colony 3, Lane 19: CK 20 colony 4

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15



16 (Extra)

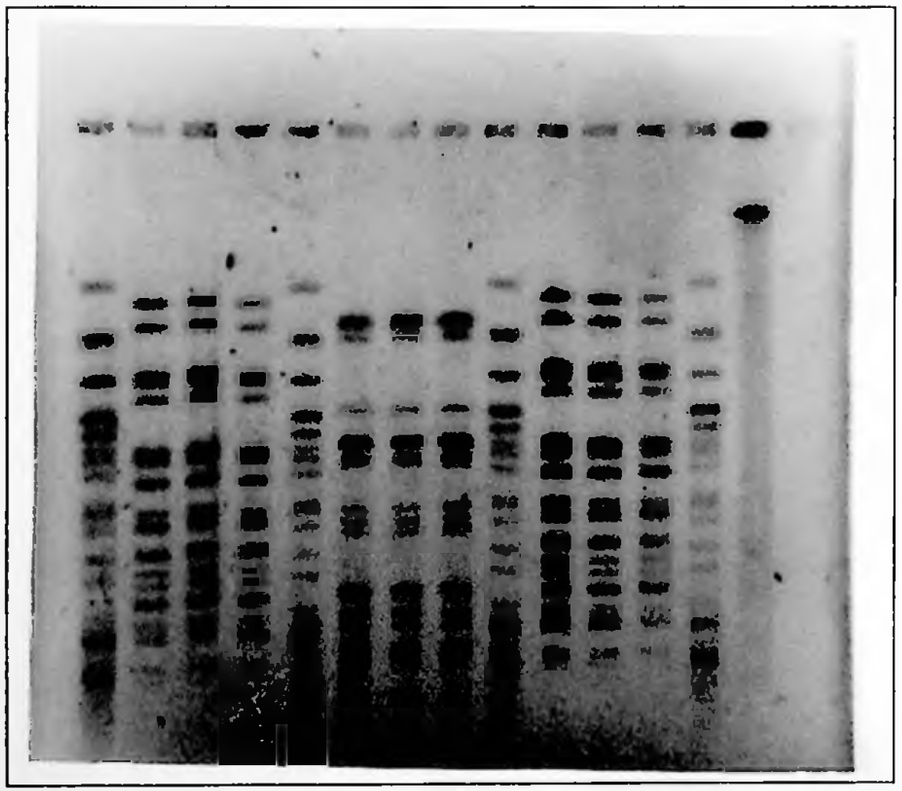
17 18 19 (*Not*I)

Figure 70. Chicken 7219901 (CK21) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 26, 1999. Lane 1: Standard (G5244), Lane 2: CK21 colony 1, Lane 3: CK21 colony 2, Lane 4: CK21 colony 3, Lane 5: Standard (G5244), Lane 6: CK21 colony 4, Lane 7: CK21 colony 5, Lane 8: CK21 colony 6, Lane 9: Standard (G5244), Lane 10: CK21 colony 7, Lane 11: CK21 colony 8, Lane 12: CK21 colony 9, Lane 13: Standard (G5244), Lane 14: CK21 colony 10, Lane 15: Negative Control.

Figure 71. Extra DNA Profiles (re-digested with *Xba*I for better resolution). Documented with Gel Doc 2000 (Quantity One Version 4.0) on December 11, 1999. Lane 16: CK21 colony 10.

Figure 72. CK21 Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 17: CK21 colony 1, Lane 18: CK21 colony 4.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

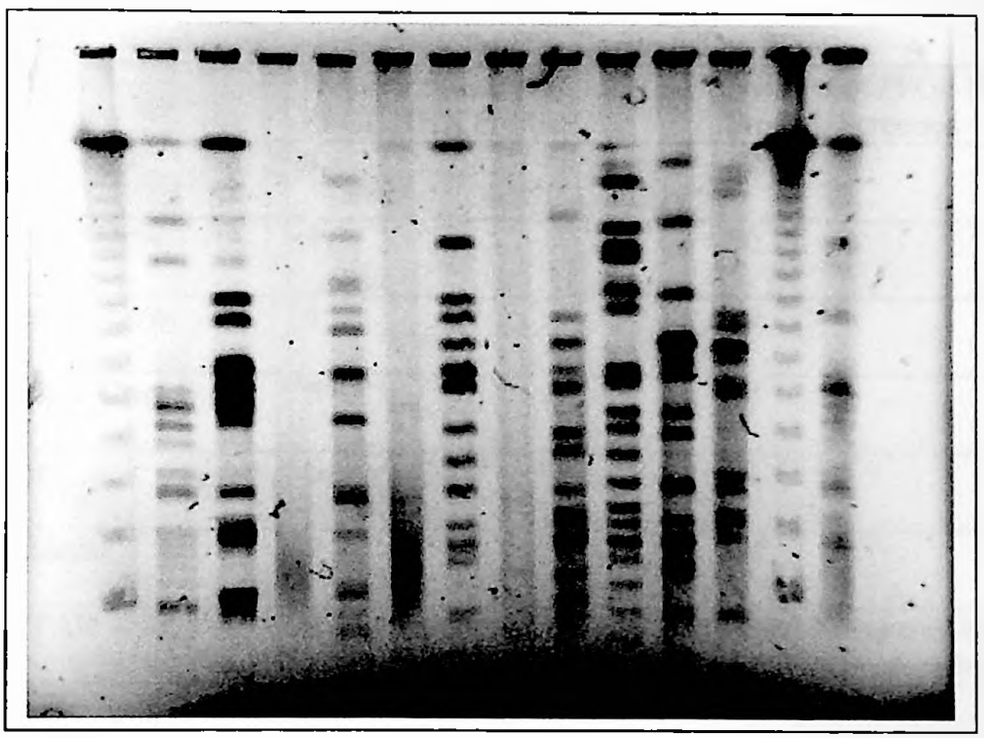


16 (Extra)

17 18 (*Not*I)

Figure 73. Summary of Chicken Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 1: Lambda Ladder, Lane 2: CK2-2 colony 1, Lane 3: CK2-2 colony 3, Lane 4: CK20 colony 1, Lane 5: CK20 colony 3, Lane 6: CK20 colony 4, Lane 7: CK21 colony 1, Lane 8: CK 21 colony 4, Lane 9: Lambda Ladder, Lane 10: Positive Control (JM107), Lane 11: Negative Control.

1 2 3 4 5 6 7 8 9 10 11



restriction can be seen in Figures 25, 26, 37, 46, 53, 54, 62, 63, and 73. The DNA profiles of each animal were documented with Gel Doc 2000 (Quantity One Version 4.0) and Scanalytics software (RFLPscan Forensic Edition Version 3.0). Figures 33 and 34 give examples of DNA profiles that have been analyzed using Scanalytics software. All documentation is shown in the figures.

Based on observation, it was determined that of 190 total DNA profiles screened, there were a total of 39 unique DNA profiles of all animals tested (Table 4).

Table 4. Number of Unique DNA Profiles Generated by PFGE

Animal	Number of Unique PFGE DNA Profiles	Number of PFGE DNA Profiles Screened
Angus Cow	8	40
Herford Cow	9	36
Charlais Cow	4	30
Pensgar Cow	5	27
Deer	6	27
Chicken	7	30
TOTAL	Total Number of Unique DNA Profiles=39	Total Number of PFGE DNA Profiles Screened=190

The data in Table 4 is based only on this study. The animals with the largest number of unique DNA profiles were Herford Cows (9). Charlais Cows had the least amount of unique DNA profiles (4). It is expected that the number of unique DNA profiles would increase as more DNA profiles from known animal sources are screened.

Each sample's DNA profiles were visually analyzed to determine the frequency of similar DNA profiles for each individual animal (Table 5). For example, Herford Cow 7219903 had 3 DNA profiles that generated the same profile pattern as strain H7-3-1, 5

Table 5. Table of Frequencies of Unique DNA Profiles

Sample Name	Unique Colony	Strain I.D.	No. of Similar Strains in Individual Animal	Freq. Of Similar Strain in Individual Animal
Angus Cow 3319904	colony 1	A3-4-1	10	(10/10) 100.00%
Angus Cow 3319905	colony 1	A3-5-1	10	(10/10) 100.00%
Angus Cow 3179908	colony 1	A3-8-1	5	(5/10) 50.00%
Angus Cow 3179908	colony 3	A3-8-3	1	(1/10) 10.00%
Angus Cow 3179908	colony 6	A3-8-6	4	(4/10) 40.00%
Angus Cow 5059909	colony 3	A5-9-3	8	(8/10) 80.00%
Angus Cow 5059909	colony 4	A5-9-4	1	(1/10) 10.00%
Angus Cow 5059909	colony 9	A5-9-9	1	(1/10) 10.00%
Herford Cow 3179910	colony 1	H3-0-1	9	(9/10) 90.00%
Herford Cow 3179910	colony 9	H3-0-9	1	(1/10) 10.00%
Herford Cow 7219903	colony 1	H7-3-1	3	(3/10) 30.00%
Herford Cow 7219903	colony 3	H7-3-3	5	(5/10) 50.00%
Herford Cow 7219903	colony 5	H7-3-5	1	(1/10) 10.00%
Herford Cow 7219903	colony 9	H7-3-9	1	(1/10) 10.00%
Herford Cow 7219904	colony 1	H7-4-1	10	(10/10) 100.00%
Herford Cow 7219909	colony 1	H7-9-1	2	(2/6) 33.33%
Herford Cow 7219909	colony 9	H7-9-9	4	(4/6) 66.67%
Charlais Cow 10069904	colony 4	CL1-4-4	2	(2/10) 20.00%
Charlais Cow 10069904	colony 8	CL1-4-8	8	(8/10) 80.00%
Charlais Cow 7219902	colony 1	CL7-2-1	10	(10/10) 100.00%
Charlais Cow 5059913	colony 2	CL5-3-2	10	(10/10) 100.00%
Pensgar Cow 7219905	colony 2	P7-5-2	7	(7/7) 100.00%
Pensgar Cow 7219907	colony 1	P7-7-1	1	(1/10) 10.00%
Pensgar Cow 7219907	colony 4	P7-7-4	9	(9/10) 90.00%
Pensgar Cow 7219908	colony 1	P7-8-1	7	(7/10) 70.00%
Pensgar Cow 7219908	colony 2	P7-8-2	3	(3/10) 30.00%
Deer 9806	colony 3	D9-6-3	2	(2/10) 20.00%
Deer 9806	colony 4	D9-6-4	6	(6/10) 60.00%
Deer 9806	colony 5	D9-6-5	2	(2/10) 20.00%
Deer 9807	colony 1	D9-7-1	10	(10/10) 100.00%
Deer 7199901	colony 1	D7-1-1	3	(3/7) 42.86%
Deer 7199901	colony 2	D7-1-2	4	(4/7) 57.14%
Chicken 2209902	colony 1	CK2-2-1	5	(5/10) 50.00%
Chicken 2209902	colony 3	CK2-2-3	5	(5/10) 50.00%
Chicken 7209901	colony 1	CK20-1	2	(2/10) 20.00%
Chicken 7209901	colony 3	CK20-3	3	(3/10) 30.00%
Chicken 7209901	colony 4	CK20-4	5	(5/10) 50.00%
Chicken 7219901	colony 1	CK21-1	7	(7/10) 70.00%
Chicken 7219901	colony 4	CK21-4	3	(3/10) 30.00%

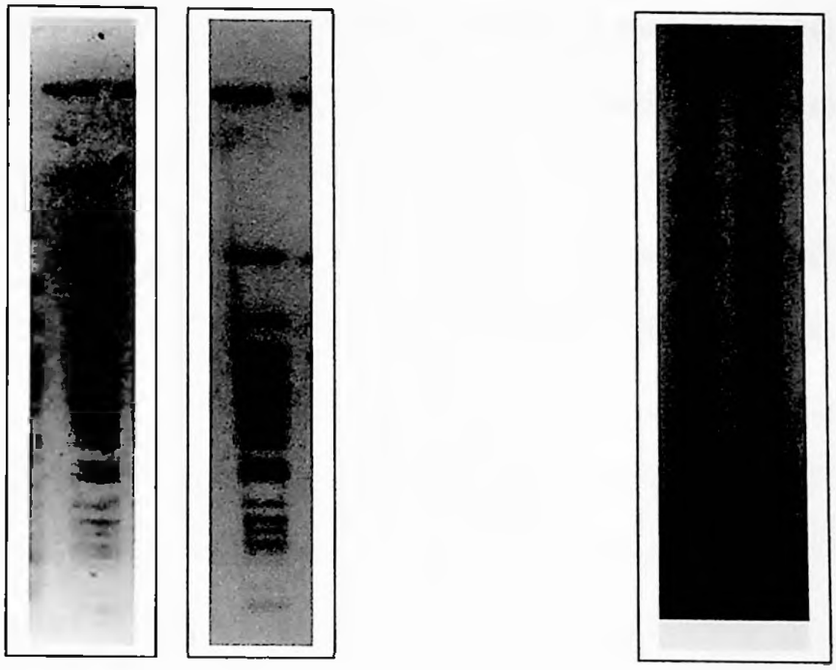
DNA profiles that generated the same profile patterns as H7-3-3, and 2 DNA profiles (H7-3-5 and H7-3-9) that were different from each other and the rest of the strains generated by this sample. Thus, the frequencies of strains similar to H7-3-1, H7-3-3, H7-3-5, and H7-3-9 in this animal were (3/10) 30%, (5/10) 50%, (1/10) 10%, and (1/10) 10%, respectively. For 13 animals, it was observed that different DNA profiles were seen within each individual animal. For example, Herford Cow 7219903 had 4 unique DNA profiles (H7-3-1, H7-3-3, H7-3-5, and H7-3-9). This accounted for a large amount of variability within this database. Based on this study, the following animals can be identified by one particular DNA profile: Angus Cow 3319904, Angus Cow 3319905, Herford Cow 7219904, Charlais Cow 7219902, Charlais Cow 5059913, Pensgar Cow 7219905, and Deer 9807. In order to determine if these animals can be identified by one DNA profile, a minimum of 100 different strains from each animal must be tested using PFGE and must generate the same profile, so that this conclusion can be statistically relevant and valid.

In addition, it was observed that one DNA profile from one sample was repeated in another sample. The DNA profile generated by colony 3 of Herford Cow 7219903 (H7-3-3) was not only identical to other strains within this animal, but it was also identical to the DNA profile generated by colony 1 of Herford Cow 7219904 (H7-4-1). The repeated identification of this DNA profile was due to these cows being from the same farm. No reproducibility of DNA profiles was seen between breeds of animals. Using this data, it was easy to distinguish one animal type from another (i.e. DNA profiles of Pensgar Cows are different from DNA profiles of Herford Cows).

Figure 74. Pensgar Cow 7219907 digested with *Xba*I. Lane 1: P7-7 colony 1 and Lane 2: P7-7 colony 4.

Figure 75. Pensgar Cow 7219907 digested with *Not*1. Lane 3: P7-7 colony 1 and Lane 4: P7-7 colony 4.

1 2 (*Xba*I) 3 4(*Not*I)



*Not*I RESTRICTION ENZYME Vs. *Xba*I RESTRICTION ENZYME

For the Pensgar Cow *Not*I DNA profiles compared to XbaI DNA profiles (Figures 74-75), a discrepancy was observed. There was a 2 band difference between the 2 *Xba*I unique DNA profiles: P7-7-1 and P7-7-4. Whereas, upon restriction with *Not*I , it was difficult to distinguish between the two unique DNA profiles. These *Not*I DNA profiles appeared to be the same. Although, *Not*I generates simpler patterns for identification purposes(less bands are generated with *Not*I restriction and the DNA bands are evenly distributed throughout the gel), it appears that, in this case, certain unique patterns could potentially be overlooked if using the *Not*I restriction enzyme alone. *Xba*I allows for better differentiation of strains within individual animals and between animals. Thus, for databasing purposes it would be better to use *Xba*I rather than *Not*I.

Furthermore, when observing the DNA profiles of chickens that had been digested with *Not*I (Figure 70), another discrepancy had been observed with this restriction enzyme. When first tested with PFGE, the three strains that had been digested with *Not*I, Chicken 7209901 colony 1 (CK20-1), Chicken 7209901 colony 4 (CK20-4), and Chicken 7219901 colony 1 (CK21-1), did not generate DNA profiles. At this point, it was thought that there was an insufficient amount of DNA available for digestion to occur. The experiment was repeated 3 times to determine the problem. For each experiment no DNA profiles were generated using PFGE. Finally, the DNA profiles that Marshall University's Potomac River Project had generated for these samples were analyzed. It was determined that the isolates used in this experiment and Marshall's experiment were the same. Thus, their documented gels were compared to this study's documented gels (Figures 76-80).

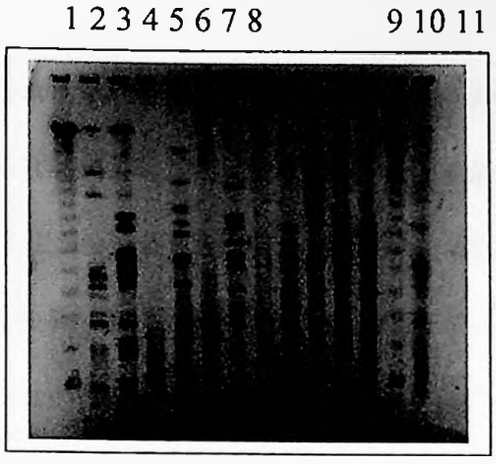
Figure 76. Summary of Chicken Unique DNA Profiles (digested with *Not*I). Documented with Gel Doc 2000 (Quantity One Version 4.0) on February 27, 2000. Lane 1: Lambda Ladder, Lane 2: CK2-2 colony 1, Lane 3: CK2-2 colony 3, Lane 4: CK20 colony 1, Lane 5: CK20 colony 3, Lane 6: CK20 colony 4, Lane 7: CK21 colony 1, Lane 8: CK 21 colony 4, Lane 9: Lambda Ladder, Lane 10: Positive Control (JM107), Lane 11: Negative Control.

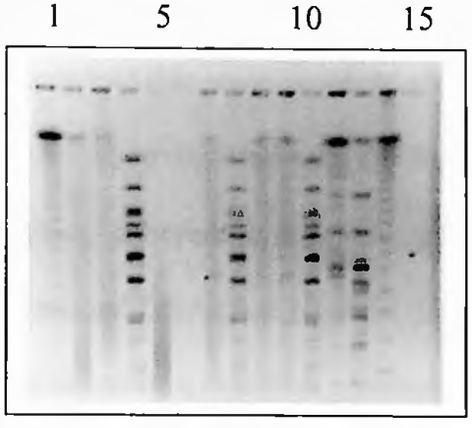
Figure 77. Chicken 7209901 (CK20) DNA Profiles (digested with *Not*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 16, 1999. Krista Haught, M.S. Forensic Science, Potomac River Project, Marshall University. Lane 1: Lambda Ladder, Lane 2: CK20 colony 1, Lane 3: CK20 colony 2, Lane 4: CK20 colony 3, Lane 5: CK20 colony 4, Lane 6: CK20 colony 5, Lane 7: CK20 colony 6, Lane 8: CK20 colony 7, Lane 9: CK20 colony 8, Lane 10: CK20 colony 9, Lane 11: CK20 colony 10, Lane 12: Standard (JM107) Lane 13: Standard (JM107) Lane 14: Lambda Ladder, Lane 15: Negative Control.

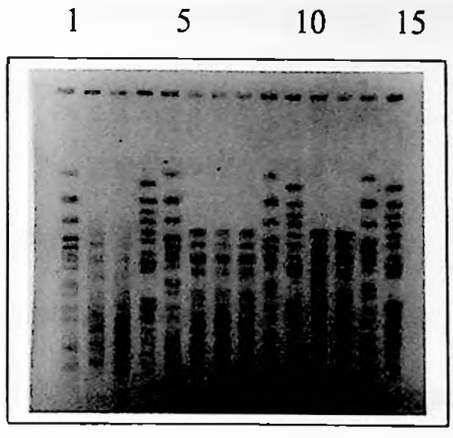
Figure 78. Chicken 7209901 (CK20) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 26, 1999. Lane 1: Standard (G5244), Lane 2: CK20 colony 1, Lane 3: CK20 colony 2, Lane 4: CK20 colony 3, Lane 5: Standard (G5244), Lane 6: CK20 colony 4, Lane 7: CK20 colony 5, Lane 8: CK20 colony 6, Lane 9: Standard (G5244), Lane 10: CK20 colony 7, Lane 11: CK20 colony 8, Lane 12: CK20 colony 9, Lane 13: Standard (G5244), Lane 14: CK20 colony 10, Lane 15: Negative Control.

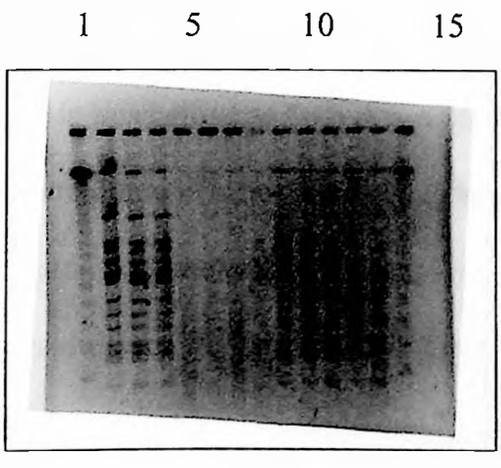
Figure 79. Chicken 7219901 (CK21) DNA Profiles (digested with *Not*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 19, 1999. Krista Haught, M.S. Forensic Science, Potomac River Project, Marshall University. Lane 1: Lambda Ladder, Lane 2: CK21 colony 1, Lane 3: CK21 colony 2, Lane 4: CK21 colony 3, Lane 5: CK21 colony 4, Lane 6: CK21 colony 5, Lane 7: CK21 colony 6, Lane 8: CK21 colony 7, Lane 9: CK21 colony 8, Lane 10: CK21 colony 9, Lane 11: CK21 colony 10, Lane 12: Standard (JM107), Lane 13: Standard (JM107) Lane 14: Lambda Ladder, Lane 15: Negative Control.

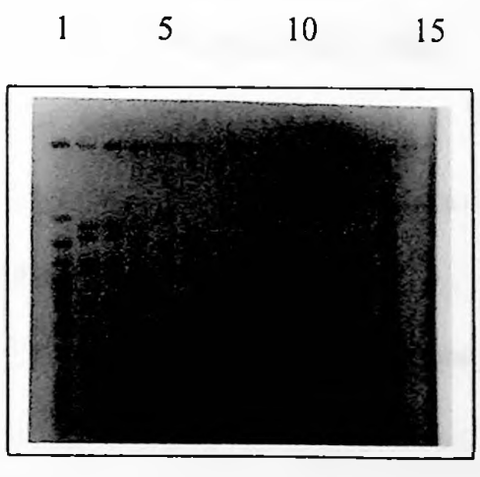
Figure 80. Chicken 7219901 (CK21) DNA Profiles (digested with *Xba*I), documented with Gel Doc 2000 (Quantity One Version 4.0) on November 26, 1999. Lane 1: Standard (G5244), Lane 2: CK21 colony 1, Lane 3: CK21 colony 2, Lane 4: CK21 colony 3, Lane 5: Standard (G5244), Lane 6: CK21 colony 4, Lane 7: CK21 colony 5, Lane 8: CK21 colony 6, Lane 9: Standard (G5244), Lane 10: CK21 colony 7, Lane 11: CK21 colony 8, Lane 12: CK21 colony 9, Lane 13: Standard (G5244), Lane 14: CK21 colony 10, Lane 15: Negative Control.


1 2 3 4 5 6 7 8
9 10 11


1
5
10
15


1
5
10
15


1
5
10
15


1
5
10
15

It was determined that the group of strains that generated DNA profiles similar to CK20-1 and CK20-4 and the group of strains that generated DNA profiles similar to CK21-1 could not be digested with *Not*I, but could be digested with *Xba*I. These DNA profiles had been generated with PFGE at approximately the same time. Marshall University's Potomac River Project's *Not*I documented gels of the two chicken samples had been generated 1 week before the *Xba*I documented gels of the same chicken samples. Of the two enzymes used for the same bacterial isolates, only the *Xba*I restriction endonuclease generated DNA profiles for CK20-1, CK20-4, and CK21-1.

When researching into a possible cause for this, it was discovered that Virginia Tech's laboratory had this same problem. DNA profiles for some strains were not being generated when digested with the *Not*I enzyme. They believe that there may be 'something' inherent to the strain, which does not allow the *Not*I enzyme to cut the chromosomal DNA (Sue Herbein, personal communication, February 2000). This is considered to be a common problem at the Virginia Tech laboratory. Using this enzyme could be detrimental to possible source identification of bacterial strains, since other possible unique DNA profiles could be lost for that particular animal source.

The *Not*I restriction enzyme from Promega was used for this study, the Marshall University Potomac River Project and the Biology Laboratory at Virginia Tech. This enzyme is insensitive to both dcm and dam methylation. This indicates that the restriction enzyme is not inhibited from cutting DNA that methylated cytosine residues or methylated adenine residues. This is an advantage because many strains of *E. coli* contain site-specific dam and dcm DNA methylases. Methylated DNA is not cut by many

restriction enzymes. Since *Not*I is insensitive to these methylases, dcm and dam methylases should not prevent the action of this enzyme.

It was determined, based on information from both Marshall University's Laboratory and Virginia Tech's Laboratory, that more rigorous methods are needed to isolate the DNA so that there wouldn't be any protein or divalent cations associated with the chromosomal DNA preventing the action of *Not*I. Secondly, Mg 2+ (magnesium ions), from the digestion buffer, might be lost to a chelating agent, thereby preventing proper digestion. Since certain strains were affected by improper digestion by *Not*I, it was determined that there was something inherently different with those strains (a possible protein or lysate) that prevented *Not*I from cutting.

INTERPRETING DNA PROFILES

Unfortunately, there have been few epidemiological studies published regarding the interpretation of chromosomal DNA profiles of various bacteria produced by PFGE. The Centers for Disease Control and Prevention (CDC) has published an article regarding the analysis of PFGE DNA profiles of bacteria. Richard V. Goering believes that isolates that differ by 3 or 4 bands are similar and should be classified as subtypes of the same strain, based upon the Dice Coefficient (1993). Thus, those strains containing more than a 4 band difference should be classified as a different strain of bacteria (Goering, 1993). According to Tenover et al. various DNA profiles of bacteria generated by pulsed field gel electrophoresis can be designated into the following four categories: (I) Indistinguishable, (II) Closely Related, (III) Possibly Related, and (IV) Unrelated, based upon criteria recorded by observation (1995). However, a sensitive numerical method to verify observational findings has yet to be determined for strain identification purposes.

The Virginia Tech Laboratory, under the supervision of Dr. George Simmons Jr., has a large database of approximately 2500 PFGE DNA profiles of *E. coli*. These DNA profiles have been analyzed with a tcl software program using a DNA band window of 10%. The combined DNA index service for convicted felons uses a DNA band window of 5%. Ideally, our database should have a tighter match window for more precise conclusions. The reproducibility of Virginia Tech's data was evaluated using the chi-square goodness of fit test. (G. Simmons Jr., personal communication, October 21, 1999).

For this study, the cladograms of all 6 divisions were analyzed to determine which divisions were capable of showing how related the strains were. Visual inspection of the DNA profiles was used to determine what was a true match. For all divisions, it was known that P7-8 1, P7-8 3, and P7-8 4 generated identical DNA profiles. In addition, all DNA profiles of A5-9 were included in the analysis of the divisions. This allowed for a better determination of what divisions were most reliable for grouping strains.

Division 2 (357.79 kb to 302.1 kb) produced a tree that was difficult to obtain any information from (Figure 81). Within this region, each DNA profile was observed to have between 1 and 3 bands. Thus, division 2 was determined to be an unreliable region to discriminate between strains. In addition, division 5 (185.29 kb to 102.0 kb) indicated there was a match between certain strains in this division when clearly there was none. For example, A5-9-2 and A5-9-4 have different banding patterns in this region but have been clustered together within this division. Also, P7-8-2 and P7-8-4 are clustered closely together within this division when observational data indicated that these strains are entirely different within this division. These errors in the cladogram indicate that this

Figure 81. Rectangular Cladogram of division 2, Phylip version 3.5c (Felsenstein, 1993)

DIVISION 2

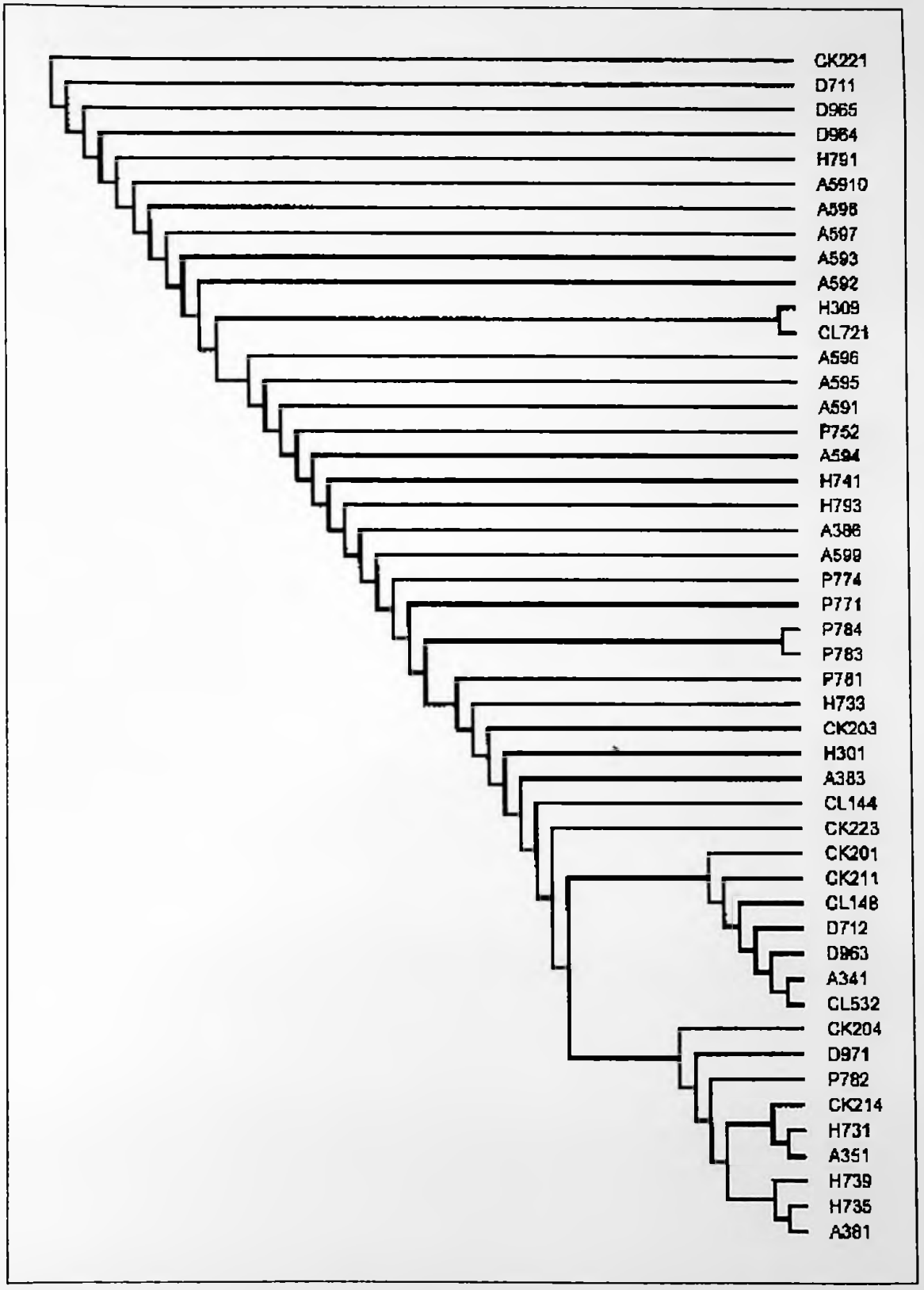
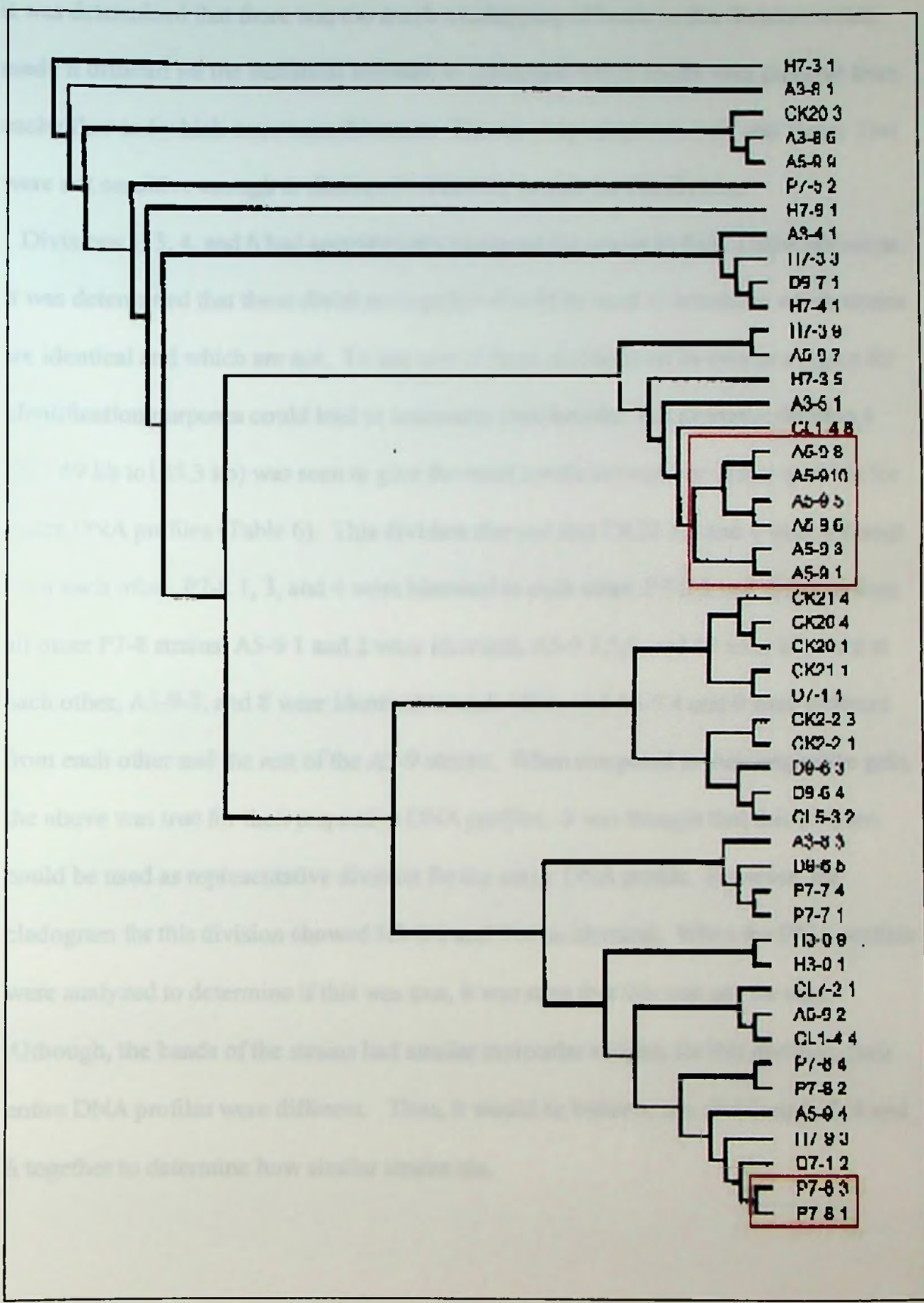

CK221
D711
D965
D964
H791
A5910
A598
A597
A593
A592
H309
CL721
A596
A595
A591
P752
A594
H741
H793
A386
A599
P774
P771
P784
P783
P781
H733
CK203
H301
A383
CL144
CK223
CK201
CK211
CL148
D712
D963
A341
CL532
CK204
D971
P782
CK214
H731
A351
H739
H735
A381

Figure 82. Rectangular Cladogram of division 5, Phylip version 3.5c (Felsenstein, 1993)

DIVISION 5


H7-3 1
A3-8 1
CK20 3
A3-8 6
A5-9 9
P7-5 2
H7-9 1
A3-4 1
H7-3 3
D9 7 1
H7-4 1
H7-3 9
AG 9 7
H7-3 5
A3-6 1
CL1 4 8
A5-9 8
A5-910
A5-9 5
AG 9 6
A5-9 3
A5-9 1
CK21 4
CK20 4
CK20 1
CK21 1
D7-1 1
CK2-2 3
CK2-2 1
D9-6 3
D9-6 4
CL5-3 2
A3-8 3
D9-6 5
P7-7 4
P7-7 1
H3-0 9
H3-0 1
CL7-2 1
AG-9 2
CL1-4 4
P7-8 4
P7-8 2
A5-9 4
H7 9 3
D7-1 2
P7-8 3
P7 8 1

division should not be used for determining what is a true match. Upon further analysis, it was determined that there was too much overlapping of bands in this division which made it difficult for the statistical software to determine which bands were different from each other and which ones were the same. The one way Anova analysis and Tukey Test were not sensitive enough to distinguish between strains for this division.

Divisions 1, 3, 4, and 6 had appropriately clustered the strains in their relative divisions. It was determined that these divisions together should be used to determine which strains are identical and which are not. To use one of these divisions on its own as a region for identification purposes could lead to erroneous conclusions. For example, division 4 (223.49 kb to185.3 kb) was seen to give the most consistent number of true matches for entire DNA profiles (Table 6). This division showed that CK20 1,3 and 4 were different from each other, P7-8 1, 3, and 4 were identical to each other, P7-8-2 was different from all other P7-8 strains, A5-9 1 and 2 were identical, A5-9 3,5,6, and 10 were identical to each other, A5-9-7, and 8 were identical to each other, and A5-9 4 and 9 were different from each other and the rest of the A5-9 strains. When compared to their respective gels, the above was true for their respective DNA profiles. It was thought that this division could be used as representative division for the entire DNA profile. However, the cladogram for this division showed H3-0-1 and 9 to be identical. When the DNA profiles were analyzed to determine if this was true, it was seen that this was not the case. Although, the bands of the strains had similar molecular weights for this division, their entire DNA profiles were different. Thus, it would be better to use divisions 1, 3, 4 and 6 together to determine how similar strains are.

Figure 83. Rectangular Cladogram of division 4, Phylip version 3.5c (Felsenstein, 1993)

DIVISION 4


D9-8 5
CK21 4
D7-1 2
D7-1 1
H7 9 1
CK20 4
A5-0 5
A5 8 3
H7-3 9
H7 3 6
A5-9 1
A5 8 2
H7-0 3
P7-7 4
A5-4 1
P7-8 9
P7 8 1

Figure 84. Rectangular Cladogram of division 1, Phylip version 3.5c (Felsenstein, 1993)

DIVISION 1


CK20 4
CK20 1
H3-0 1
CK21 4
P7 5 2
A3-0 1
H7-3 3
A5-9 4
D9-6 5
A5-9 9
H7-8 3
A3-4 1
CL1-4 8
H7-3 9
CL7-2 1
A5 9 7
CK2-2 3
D9 6 3
A3-5 1
A5-9 8
A5-9 1
CK20 3
CL1-4 4
CK2 2 1
A5-9 6
A5-9 5
A5-910
A5-9 3
A5-9 2
H7-4 1
D7-1 1
H7-3 1
H7 9 1
CL5-3 2
D9 7 1
A3-8 3
P7-7 1
P7-7 4
CK21 1
D7-1 2
A3-8 6
D9-6 4
P7-8 2
H3-0 9
H7-3 5
P7-8 3
P7-8 4
P7-8 1

Figure 85. Rectangular Cladogram of division 3, Phylip version 3.5c (Felsenstein, 1993)

DIVISION 3

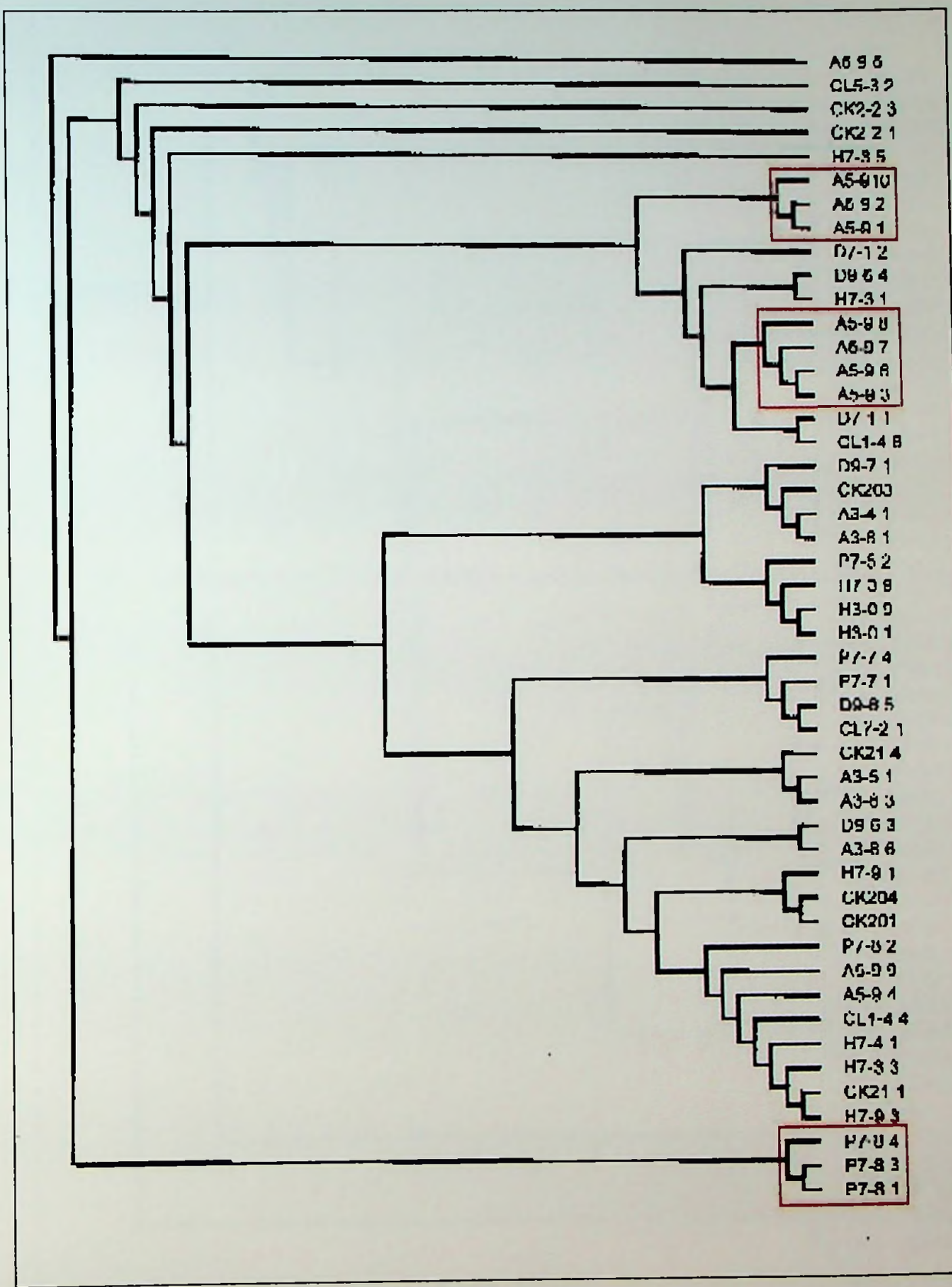

A5 9 6
CL5-3 2
CK2-2 3
CK2 2 1
H7-3 5
A5-9 10
A5 9 2
A5-9 1
D7-1 2
D9 6 4
H7-3 1
A5-9 8
A5-9 7
A5-9 6
A5-9 3
D7 1 1
CL1-4 8
D9-7 1
CK203
A3-4 1
A3-8 1
P7-5 2
H7 3 8
H3-0 9
H3-0 1
P7-7 4
P7-7 1
D9-8 5
CL7-2 1
CK21 4
A3-5 1
A3-8 3
D9 6 3
A3-8 6
H7-9 1
CK204
CK201
P7-5 2
A5-9 9
A5-9 4
CL1-4 4
H7-4 1
H7-3 3
CK21 1
H7-9 3
P7-8 4
P7-8 3
P7-8 1

Figure 86. Rectangular Cladogram of division 6, Phylip version 3.5c (Felsenstein, 1993)

DIVISION 6

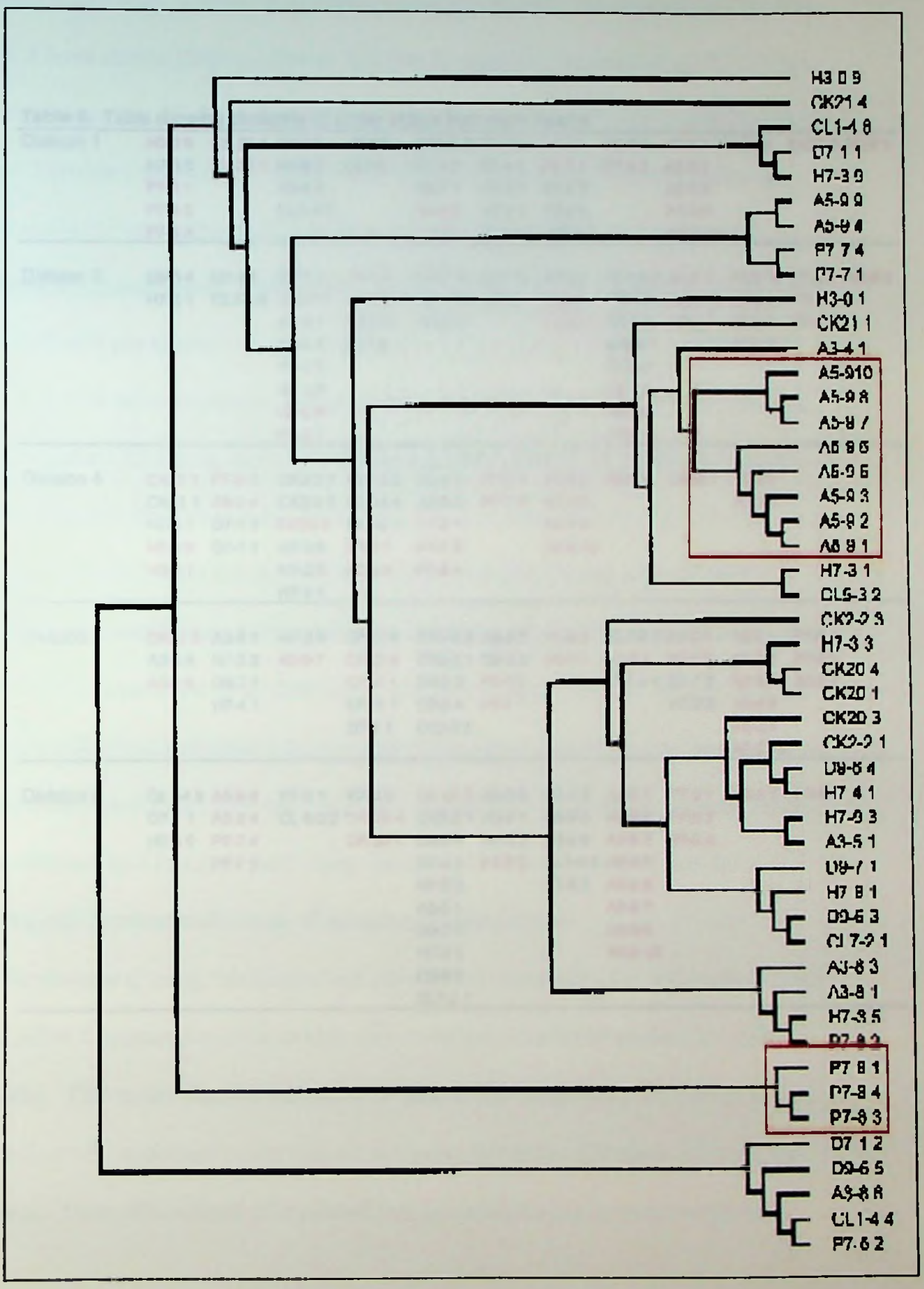

H3 0 9
CK21 4
CL1-4 6
D7 1 1
H7-3 9
A5-9 9
A5-9 4
P7 7 4
P7-7 1
H3-0 1
CK21 1
A3-4 1
A5-910
A5-9 8
A5-9 7
A5 9 6
A5-9 5
A5-9 3
A5-9 2
A5 9 1
H7-3 1
CL5-3 2
CK2-2 3
H7-3 3
CK20 4
CK20 1
CK20 3
CK2-2 1
D9-5 4
H7 4 1
H7-9 3
A3-5 1
D9-7 1
H7 8 1
D9-6 3
CL7-2 1
AJ-8 3
A3-8 1
H7-3 5
P7 8 1
P7-8 4
P7-8 3
D7 1 2
D9-6 5
A5-8 8
CL1-4 4
P7-6 2

Table 6. Table showing clustering of similar strains from each division.

Division 1	H3-0 9	CK 20 4	A5-9 9	H7-3 9	CK2-2 3	H7-4 1	P7-7 4	D9-6 4	A5-9 2	A5-9 4	CK20 3	H3-0 1
	H7-3 5	CK 20 1	H7-9 3	CL7-2 1	D9-6 3	D7-1 1	P7-7 1	P7-8 2	A5-9 3			
	P7-8 1		A3-4 1		A3-5 1	H7-3 1	D7-1 2		A5-9 5			
	P7-8 3		CL1-4 8		A5-9 8	H7-9 1	A3-8 6		A5-9 6			
	P7-8 4				A5-9 1	CL5-3 2	CK21 1		A5-9 10			
Division 3	D9-6 4	D7-1 1	D9-7 1	P7-7 4	CK21 4	D9-6 3	H7-9 1	CL1-4 4	A5-9 1	A5-9 3	P7-8 1	A5-9 5
	H7-3 1	CL1-4 8	CK20 3	P7-7 1	A3-5 1	A3-8 6	CK20 4	H7-4 1	A5-9 10	A5-9 6	P7-8 3	
			A3-4 1	D9-6 5	A3-8 3		CK20 1	H7-3 3	A5-9 2	A5-9 7	P7-8 4	
			A3-8 1	CL7-2 1				CK21 1		A5-9 8		
			P7-5 2					H7-9 3				
			H7-3 9					A5-9 4				
			H3-0 9					A5-9 9				
			H3-0 1					P7-8 2				
Division 4	CK20 1	P7-8 2	CK2-2 1	CL5-3 2	A3-4 1	P7-7 1	A5-9 3	A5-9 9	CK20 4	A5-9 7		
	CK21 1	A5-9 4	CK2-2 3	CL1-4 4	A3-8 3	P7-7 4	A5-9 5			A5-9 8		
	H7-3 1	D7-1 2	CK20 3	CL7-2 1	P7-8 1		A5-9 6					
	H3-0 9	D7-1 1	H7-3 9	A5-9 1	P7-8 3		A5-9 10					
	H3-0 1		H7-3 5	A5-9 2	P7-8 4							
			H7-4 1									
Division 5	CK20 3	A3-4 1	H7-3 9	CK21 4	CK2-2 3	A3-8 3	H3-0 9	CL7-2 1	P7-8 1	A5-9 1	P7-8 4	
	A3-8 6	H7-3 3	A5-9 7	CK20 4	CK2-2 1	D9-6 5	H3-0 1	A5-9 2	P7-8 3	A5-9 3	P7-8 2	
	A5-9 9	D9-7 1		CK20 1	D9-6 3	P7-7 4		CL1-4 4	D7-1 2	A5-9 5	A5-9 4	
		H7-4 1		CK21 1	D9-6 4	P7-7 1			H7-9 3	A5-9 6		
				D7-1 1	CL5-3 2					A5-9 8		
										A5-9 10		
Division 6	CL1-4 8	A5-9 9	H7-3 1	H7-3 3	CK 20 3	A3-8 3	D7-1 2	A5-9 1	P7-8 1	H3-0 1	H3-0 9	
	D7-1 1	A5-9 4	CL 5-3 2	CK 20 4	CK2-2 1	A3-8 1	D9-6 5	A5-9 2	P7-8 3			
	H7-3 9	P7-7 4		CK 20 1	D9-6 4	H7-3 5	A3-8 6	A5-9 3	P7-8 4			
		P7-7 1			H7-4 1	P7-8 2	CL1-4 4	A5-9 5				
					H7-9 3		P7-5 2	A5-9 6				
					A3-5 1			A5-9 7				
					D9-7 1			A5-9 8				
					H7-9 1			A5-9 10				
					D9-6 3							
					CL7-2 1							

In addition, it was determined that similarity between strains should be defined as having a 2 band difference or less for each division and should cluster together (within 1 step of each other) in at least 3 of the 4 divisions. For example, P7-7-1 and P7-7-4 have similar DNA profiles except that in division 1 bacterial strain P7-7-1 has 2 extra bands that P7-7-4 does not have. The cladogram for division 1 indicated that these 2 strains were within 1 step of each other. This indicated that the strains are very similar to each other in this division. Also, they clustered together in divisions 3,4,and 6. In conclusion, it was determined that these strains are similar. Thus, P7-7-1 and P7-7-4 are similar.

Using the aforementioned rules, the following strains (Table 6) were determined to be similar: P7-8-1, 3, and 4 (4/4 division match), A5-9-1, 2 (3/4 division match), A5-9-3, 6 (4/4 division match), A5-9-3, 6,10 (3/4 division match), P7-7 1,4 [3/4 division match (2 band diff.)], and CK20 1,4 [3/4 division match (2 band diff.)]. Also, H3-0-1 and 9 were determined to be different (2/4 division match) and CK20-1 and 3 were determined to be different as well (0/4 division match). All other unique DNA profiles that had more than a 2 band difference in more than 2 of the 4 divisions were determined to be different. Thus, it was possible to distinguish between strains that differed by 4 bands or more using this method. Thus, it was possible to distinguish between individuals of animals and animal types.

Furthermore, using this method we were able to determine that, although, H7-3-3 and H7-4-1 appeared to look similar, they were not statistically similar (1/4 division match). This meant that the molecular weights of the bands were not similar enough to indicate by statistical means that there was no significant difference between the strains. Thus, this method of statistical analysis is dependent upon the molecular

weight data of the DNA profiles rather than the pattern as suggested by visual interpretation.

FUTURE CONSIDERATIONS FOR CONSTRUCTING A DATABASE

There were many variables that were not accounted for that lead to a large amount of variation in the DNA profiles within animals or animal types. These global considerations were geographical, temporal, and diet. The samples came from a selected area that was too large. The sampling area should be narrower for more precise results. The change in the DNA profiles over time was not studied. Samples from the same animal should be analyzed over a long period of time (at least 1 year). This helps to account for possible temporal changes due to a change in animal habits. In addition, a change in diet for the animals was not taken into consideration. Diet is a key variable that needs to be accounted for. A change in diet will generate different DNA profiles for an animal (Segelken, 1998). These factors need to be taken into account so that more precise results could be obtained. Although one tried to reduce the number of variables in this study, sampling was a major problem that led to less precise results. A study that should be done to determine the validity of the DNA profiles generated would be to take samples from one animal over the course of a year. This would determine the variation of DNA profiles that can be observed from this animal because of changes in intestinal flora. After constructing this primary study, a small database should be developed for the same type of animal on one particular farm for a certain amount of time. This type of animal should give the same results as based upon one of these individual animals. Any deviations of DNA profiles that were found to be different from the individual animal should be noted. The same research should be repeated for all potential animal sources on this farm.

After establishing this baseline of study, similar research on a different farm should be done to determine if DNA profiles from one farm can be distinguished from DNA profiles from another farm in the same area. Wildlife sites should also be analyzed in the same way. Afterward, the database can be made larger and larger using the same criteria until all farms and wildlife sites have been studied in that area. Ideally, one wants to saturate the database so that all possible DNA profiles can be determined for each animal. *Xba*I should be the restriction enzyme to use for this type of study. *Not*I has the following disadvantages: it is unable to differentiate between unique DNA profiles, it is unable to digest certain strains, and a *Not*I database cannot be compared to the C.D.C.'s database. The C.D.C.'s database consists of pathogenic *E. coli* DNA profiles that have been generated using *Xba*I. It would be useful to compare an *E. coli* database of known sources to the C.D.C.'s database in order to monitor water quality and prevent possible outbreaks.

FUTURE METHODS OF STATISTICAL ANALYSIS

One Way Anova analysis generated cladograms is a method that could determine which DNA profiles are similar. The larger the database the greater the chances of higher division by division matches and the more precise the clustering of similar strains. Ideally, if more DNA profiles had been analyzed, P7-7-1 and P7-7-4 would be regarded to be different and not the same. Similar results would have been obtained for CK20-1 and CK20-4. Thus, better discrimination between DNA profiles would be seen since only identical DNA profiles would be clustered together.

The One Way Anova analysis, though a valid method, is a slow and laborious method to determine matching between bands. An uncertainty window needs to be established so rapid analysis can occur to determine which DNA profiles are similar and which are not. An uncertainty window is a large enough window that

encompasses similar bands yet prevents different bands from falling into that window. An amount determined by the precision of the PFGE instrument is added to and subtracted from the band size determines the uncertainty window, $X \pm \alpha X$, where X is the measurement of the band size and α is the uncertainty value (National Research Council, 1998, p. 140). To establish an uncertainty window, for this study, the standard deviation of similar bands (α) needs to be determined. To do this, 100 DNA profiles of the same strain needs to be generated. Gels, buffers, and conditions for electrophoresis will give added variation to molecular weights of a band. The DNA profile for this test should be uniformly distributed along the gel.

The Scanalytics software can be used to analyze these DNA profiles using this pre-established uncertainty window. When comparing two strains, if two uncertainty windows do not overlap then there is no match and if the windows overlap then a match is declared. For more precise matching criteria, a match window for each of the 4 divisions in this study should be established. Divisions 1, 3, 4, and 6 should have their own uncertainty windows since bands have been observed to be closer together in division 3, whereas, they appear to be further apart in divisions 1, 4, and 6. This has been observed for all generated DNA profiles in this study.

CHAPTER IV.

SUMMARY

A small database of 190 *Xba*I DNA profiles was generated using PFGE. Of these profiles screened, 39 were determined to be unique. These unique strains were cut with a second restriction enzyme, *Not*I. It was determined that, for databasing purposes, the *Xba*I restriction enzyme was a better enzyme to use. *Not*I does not allow for better differentiation of strains within animals and is unable to cut certain strains. All gels were documented and analyzed.

The *Xba*I DNA profiles were compared to determine similarity of strains within animals. It was determined that global considerations needed to be taken into account to obtain more precise results. A statistical method to determine similarity of strains was established using a One Way Anova Analysis of raw data, Tukey Test for pairwise comparisons, and the generation of rectangular cladograms from presence/absence (1/0) reports. It was determined that 4 of the 6 divisions should be used as good identifier regions for rapid analysis.

Thus, using observational and statistical data of PFGE generated DNA profiles, it was possible to differentiate between animal types and individuals of the same animal type.

LITERATURE CITED

Arbeit, R.D., M. Arthur, R. Dunn, C. Kim, R.K. Selander, and R. Goldstein. (1990). Resolution of Recent Evolutionary Divergence among *Escherichae coli* from Related Lineages: The Application of Pulsed Field Electrophoresis to Molecular Epidemiology. The Journal of Infectious Diseases. 161: 230-235.

Barrett, T.J., H. Lior, J.H. Green, R. Khakria, J.G. Wells, B.P. Bell, K.D. Greene, J. Lewis, and P.M. Griffin. (1994). Laboratory Investigation of a Multistate Food-Borne Outbreak of *Escherichiae coli* O157:H7 by Using Pulsed-Field Gel Electrophoresis and Phage Typing. Journal of Clinical Microbiology. 32(12): 3013-3017.

Becton Dickinson Microbiology Systems, (1995). BBL Oxidase Reagent Droppers instructions for use. p.1.

Cody, S.H., M.K.Glynn, J.A. Farrar, L.Cairns, P.M. Griffin, J. Kayobayashi, M. Fyfe, R. Hoffman, A.S. King, J.H. Lewis, B. Swaminathan, R.G. Bryant, D.J. Vugia. (1999, February 2). An Outbreak of *Escherichiae coli* O157:H7 Infection from Unpasteurized Commercial Apple Juice. Annals of Internal Medicine. Retrieved on November 29, 1999 from the World Wide Web: http://www.acponline.org/journals/annals/02feb99/ecoli.htm

EPA Environmental News. (1998, February 20). EPA Finalizes Pollution Limits for Seven Rivers in West Virginia. Retrieved July 25, 1999 from the World Wide Web: http://www.epa.gov/region03/r3press/pr98-118.htm

Goering, R.V. (1993). Molecular Epidemiology of Noscomial Infection: Analysis of Chromosomal Restriction Fragment Patterns by Pulsed-Field Gel Electrophoresis. Infection Control and Hospital Epidemiology. 14(10): 595-600.

Kariuki, S., C. Gilks, J. Kimari, A. Obanda, J. Muyodi, P. Waiyaki and C.A. Hart. (1999). Genotype Analysis of *Escherichia coli* Strains Isolated from Children and Chickens Living in Close Contact. Applied and Environmental Microbiology. 65(2): 472-476.

Kennedy, C. and L.S. Pierson III. (1997). Microbial Genetics 428-Spring 1999 Lecture Notes. Retrieved September 8, 1999 from the World Wide Web: http://latin.arizona.edu/~pipweb/micgen_98/Lect3/Lect_3.htm

Marino, R.P. and J.J. Gannon. (1991). Survival of Fecal Coliforms and Fecal Streptococci in Storm Drain Sediment. Wat. Res. 25(9): 1089-1098.

Maslow, J.N., M.E. Mulligan, and R.D. Arbeit. (1993). Molecular Epidemiology: Application of Contemporary Techniques to the Typing of Microorganisms. Clinical Infectious Diseases. 17: 153-64.

Matushek, M.G., M.J.M. Bonten, M.K. Hayden. (1996). Rapid Preparation of Bacterial DNA for Pulsed-Field Gel Electrophoresis. Journal of Clinical Microbiology. 34(10) 2598-2600.

Micro Conslnts. (1998). Bacteriology 302-Lab 19: Bacteriological Water Testing. Retrieved on July 1, 1998 from the World Wide Web: file:///C/MU Webpage/bacti labs/lab 19.htm

National Research Council. (1998). The Evaluation of Forensic DNA Evidence (2nd ed.). Washington, D.C.: National Academy Press.

NVPDC's Bacteria Research. (1999, July 19). What's Bugging our Streams? Retrieved July 19, 1999 from the World Wide Web: http://www.nvpdc.state.va.us/4MileRun/bacteria.htm

Reuter Medical News. (1999). Possible E.coli Contamination at Second New York Fair Reported. Retrieved on September 21, 1999 from the World Wide Web: http://id.medscape.com/reuters/prof/1999/09/09.13/ep09139b.html

Salamon, H., M.R. Segal, A. Ponce de Leon, and P.M. Small. (1998). Accommodating Error Analysis in Comparison and Clustering of Molecular Fingerprints. Emerging Infectious Diseases. 1(2): 159-168.

Segelken, Roger. (Sept. 17, 1998). CU and USDA: Cattle feeding change could cut *E. coli* risk. Cornell Chronicle. Retrieved on July 19, 1999 from the World Wide Web: http://www.news.cornell.edu//Chronicle/98/9.17.98/cattle_feeding.html

Simmons, G.M., Jr., S.A. Herbein and S.L. Myers. (1998). Solving Nonpoint Introductions of Fecal Coliforms. Retrieved on June 9, 1998 from the World Wide Web: http://atlantic.evsc.virginia.edu/auto_docs/esi_nrs96962812858.html

Tenover, F.C., R.D. Arbeit, R.V. Goering, P.A. Mickelsen, B.E. Murray, D.H. Persing, and B. Swaminathan. (1995). Interpreting Chromosomal DNA Restriction Patterns Produced by Pulsed-Field Gel Electrophoresis: Criteria for Bacterial Strain Typing. Journal of Clinical Microbiology. 33(9): 2233-2239.

Tortora, G.J., B.R. Funke, and C.L. Case. (1989). Microbiology: An Introduction, (3rd Edition). The Benjamin/Cummins Publishing Company, Inc. 196pp.

Tynkkynen, S., R. Satokari, M. Saarela, T. Mattila-Sandholm, and M. Saxelin. (1999). Comparison of Ribotyping Randomly Amplified Polymorphic DNA Analysis, and Pulsed-Field Gel Electrophoresis in Typing of *Lactobacillus rhamnosus* and *L. casei* Strains. Applied and Environmental Microbiology. 3908-3914.

U.S. Environmental Protection Agency Region III. (February 20, 1998). Fecal Coliform Total Maximum Daily Loads for the South Fork South Branch Potomac River including Lunice Creek, Mill Creek and North Fork and Anderson Run, West Virginia. 2-1pp.

Wiggins, B.A., R.W. Andrews, R.A. Conway, C.L. Corr, E.J. Dobratz, D.P. Dougherty, J.R. Eppard, S.R. Knupp, M.C. Limjoco, J.M. Mettenberg, J.M. Rinehardt, J. Sonsino, R.L. Torrijos, and M.E. Zimmerman. (1999). Use of Antibiotic Resistance Analysis to Identify Nonpoint Sources of Fecal Pollution. Applied and Environmental Microbiology. 65(8): 3483-3486.