Analyzing Mitochondrial DNA from Ancient Colonial Cattle

Meghan Elizabeth Cavanaugh

College of William and Mary

Follow this and additional works at: https://scholarworks.wm.edu/honorstheses

Recommended Citation
https://scholarworks.wm.edu/honorstheses/258

This Honors Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
Analyzing Mitochondrial DNA from Ancient Colonial Cattle

A thesis submitted in partial fulfillment of the requirement for the degree of Bachelor of Science with Honors in Biology from the College of William and Mary

by

Meghan Elizabeth Cavanaugh

Accepted for Honors

[Signatures]

Williamsburg, VA

April 2009
Abstract

Cattle were first brought to North America during colonization when settlers from Spain, France, the Netherlands and Great Britain brought livestock with them. Cattle could have also been introduced from the Caribbean islands if the vessels stopped there to restock supplies. In Virginia’s Jamestown Colony, historical records are incomplete as to the origin of its imported cattle. However, it is currently believed but not adequately documented that Virginia’s foundation cattle herd were Devons from Devonshire, England. On the other hand, because there were so many possible cattle origins in North America, historians and other researchers alike are curious to learn more about the ancestry of Virginia’s colonial cattle and the trade routes from which they came. The overall objective of this thesis research was to provide genetic evidence for the ancestry of Jamestown’s colonial cattle that will clarify and supplement historical records.

To discover the origin of colonial cattle, mitochondrial DNA (mtDNA) was analyzed from ancient bone samples. Because mtDNA is maternally inherited, it is often used to determine the ancestry of populations. Samples were taken from various sites that were occupied by early Virginia settlers throughout the James town region. Total DNA was extracted and a 218-bp piece of the most variable region of Bos taurus mtDNA, known as the displacement loop (D-loop), was amplified and sequenced. Ancient DNA is difficult to sequence, especially from samples that have been degrading in fluctuating environmental conditions for the past four hundred years. From the 22 bones analyzed, high quality sequences were obtained from 5 of them. DNA from the cheek cells of two living Devon cows was also sequenced for the basis of comparison. All sample sequences obtained were compared to the published Aberdeen Angus 218-bp fragment because it is the most common cow in Europe (Accession # V00654). All five sequences from bone samples matched exactly to the comparison sequence, and one of the Devon cow samples showed this exact match as well. The other Devon cow sequence consistently showed variation from the comparison sequence at positions 99 and 126.

It has been previously shown that, despite the variability of the D-loop region, Bos taurus breeds are differentiated into haplotypes based on as little as 1-2 nucleotide changes in the DNA base sequence from the common Aberdeen Angus comparison sequence. Some individuals that belong to different Bos taurus breeds have the exact same base sequence as the Aberdeen Angus. Within each breed of cattle, individuals often display mtDNA D-loop variation that is inconsistent between other members of that breed. Therefore, because there is not as much mtDNA diversity in Bos taurus cattle as once assumed and variation is often inconsistent within breeds, the origin of colonial cattle in Virginia cannot be determined with certainty. All of the successfully sequenced samples match that of one of the Devon cows, thus there is a possibility that the foundation herd in Jamestown was indeed made up of Devon cattle. However, this Devon sample also completely matched the Aberdeen Angus sequence. Because members of other breeds like the Kerry and Cuban Creole also match the Aberdeen Angus sequence exactly, these possibilities cannot be eliminated. More sequences from both ancient and Devon cow samples need to be obtained and analyzed before conclusions can be made about the ancestry of Jamestown’s colonial cattle.
# Table of Contents

I. Introduction 1

   a. Overview 1
   b. Possible Origins of Colonial Cattle 3
      i. Spanish and Caribbean Influence 3
      ii. French Influence 5
      iii. Dutch Influence 5
      iv. English and Irish Influence 6
      v. Feral Herd Influence 9
      vi. Summary of Likely Influences 10
   c. Breeds and Haplotypes 10
   d. Mitochondrial DNA 13
   e. Collagen and Bone Preservation 15
   f. Specific Aims for Research 16

II. Materials and Methods 17

   a. Sampling 17
   b. Bone Grinding 18
   c. DNA Extraction 19
   d. PCR 21
   e. DNA Gel Electrophoresis 25
   f. PCR Product Purification 25
   g. DNA Sequencing 26
   h. Sequence Analysis 26
   i. Collagen Extraction 27
   j. SDS-PAGE 28

III. Results 30

   a. Optimizing Procedures for Ancient DNA 30
      i. Modifications Made to the DNA Extraction Procedure 30
      ii. Modifications Made to the PCR Amplification Procedure 32
   b. Amount of Collagen Does Not Clearly Indicate Bone Preservation 35
   c. Molecular Evidence Suggests European Cattle 38

IV. Discussion 44

   a. Distinguishing Breeds Using mtDNA 45
   b. Future Directions 49

V. Acknowledgements 50
List of Figures and Tables

Figure 1—Possible Origins of Virginia Cattle from Importations 11
Table 1—Variable Positions in T Haplotypes 14
Figure 2—Arrangement of D-loop Target Region within Bos taurus mt Genome 23
Figure 3—Methods Involved in the Analysis of mtDNA and Collagen from Ancient Bone 29
Figure 4—Representative Optimal and Suboptimal Visualization of DNA 33
Figure 5—Relative Amounts of Collagen in B3, B4, J1-J4 Samples 37
Table 2—Successfully Sequenced Ancient Bone Samples 39
Figure 6—Chromatograms of Disputable Sequences in B1 and T1 Samples 40
Figure 7—Alignments of Representative Successful Samples with the 218 bp Comparison Sequence 42-43
Figure 8—Variable Positions in the 218 bp Fragment of the Aberdeen Angus Comparison Sequence 46
Table 3—Variable Positions in the 218 bp Fragment of Bos taurus Breeds 47
I. Introduction

a. Overview

Humankind has always pursued knowledge, whether it be about how the body works, what kind of practices other cultures have, or how cows came to reside in North America. When we cannot gain knowledge from our own experiences and observations, we must rely on the documented experiences of others present in historical records. However, historical records are not always solely adequate in showing us the entire picture. In regards to discovering the origin of colonial cattle in Virginia, much of what we accept as historical fact is based on assumptions and interpretations, and there is a great volume still to be learned in this area of study. Where historical records are incomplete, zooarchaeology and molecular biology hope to provide other kinds of evidence for the origin of colonial cattle. This kind of evidence has been used to show the ancestry and diversity of Colombian Creole cattle (Carvajal-Carmona et al., 2003). Molecular biology has also been used to determine the time and degree of divergence between Asian and Afro-European cattle lineages (Loftus et al., 1994b), and it has even been used to estimate when cattle were domesticated throughout the world (Bailey et al., 1996). Combining these three disciplines gives a more complete story of how cows came to North America.

Historically, it is widely accepted that cattle and other livestock did not exist in North America before European colonization because there is no
evidence indicating their presence at the arrival of the first colonists. According to G. A. Bowling (1942), there are four possible origins of cattle on this continent, all of which could be true. They could have come from the West Indies, or what is now known as the Caribbean Islands, to the Atlantic or Gulf coast, from Mexico to the American Southwest or California, from the French colonies in present-day Canada to the American Northeast, or from Europe to various American colonies. These introductions could have been due to direct importation, trade with other countries, colonies, or companies, or the escape of cattle from pens to form wild, migrating herds.

Because there were so many possibilities for cattle introduction in Virginia and historical records do not specify the breed or even place of origin of the cattle, historians and other researchers alike are curious to know more. The overall objective of this thesis research was to provide genetic evidence for the ancestry of Jamestown’s colonial cattle that will clarify and supplement historical records. Taking history, zooarchaeology, and molecular biology into account, this thesis research focused on determining possible breeds that were introduced and eliminating others.
b. Possible Origins of Colonial Cattle

i. Spanish and Caribbean Influence

During the sixteenth and seventeenth centuries, many European countries were colonizing specific regions of North America and importing their own cattle from various sources. The Spanish colonized the West Indies in 1493, bringing cattle with them. By the early 1500s, cattle stocks had grown and stabilized, which made the West Indies a good source of cattle and other livestock for the colonies. Not only are there records of cattle being purchased from Haiti in the late sixteenth and early seventeenth centuries in the colonies, but these Spanish cattle were also introduced into Florida and other southern colonies during Spanish colonization. Feral herds from Florida were documented to have reached as far north as Georgia and South Carolina, but there is no documentation of Spanish feral cattle herds reaching Virginia (Gray, 1958). If there was any Spanish cattle influence in Virginia, cattle would have come from the West Indies either through restocking livestock before arriving in Jamestown or through trade after settlements were established.

According to V.D. Anderson (2004), “ships bound for the Chesapeake sometimes procured livestock in the West Indies” whereas “vessels headed for New England typically tracked across the North Atlantic and rarely took so expensive a detour.” Ships often lost over half of their livestock during the journey (Anderson, 2004), so it was worthwhile for ships bound for more
southern colonies like Jamestown to restock in the West Indies. Trade was another possible way to bring cattle from the West Indies into Virginia. There is evidence of the Virginia Colony trading with the West Indies as early as 1633, but it is unknown whether or not cattle were imported from these islands (Bruce, 1896).

If cattle were brought to Virginia from the West Indies, they were probably of mixed origin. The West Indies, referred to today as the Caribbean Islands, consists of a series of small islands east of present-day Central America, including the Bahamas and other islands bordering the Caribbean Sea. Because this region includes nearly seven thousand isolated pieces of land, inevitably more than one country established colonies. The Spanish first discovered the islands in 1492 with Columbus’s first expedition, and they brought Spanish cattle with them on their second expedition in 1493. There is also evidence that the Portuguese brought cattle to Brazil and nearby islands in the 1530s (Anderson, 2004). Other European countries such as England, France, and the Netherlands also began to colonize the Caribbean islands, causing the Spanish to ship their cattle on the islands to their settlements in Florida and South America in the mid 1500s (Liron et al., 2006). In the 16th and 18th centuries, it is thought that West African cattle were also brought to the Caribbean Islands with the slave trade. As new cattle were introduced with colonization and cattle populations continued to propagate, these cattle formed a Creole population.
ii. French Influence

In the 1550s, the French brought cattle with them as they settled along the St. Lawrence River. However, the French came to the new world primarily to trade, so cattle served their dietary needs and had little economic importance. It was not until the French Jesuits came over with their Jersey, Brittany, and Normandy cattle that cows were likely moved farther south and west as the people sought religious converts in the mid 1600s (Bowling, 1942). The significance of this in the Virginia Colony is unknown because by this time, Jamestown’s foundation or original herd would have already been established. These French cows could only add to the genetic variety found in later generations through trade with New England colonies or through migrating feral herds.

iii. Dutch Influence

The Dutch established settlements in present-day New York by 1621, and shipments of cattle from the West India Company were documented beginning in 1625. It is thought that the Dutch imported few cattle directly from Holland because it was presumably easier to trade with the West Indies. It is also believed that the New England settlers observed that Dutch cattle might not be the best suited for the colonial climate. There was evidence of intercolonial trade in the New England colonies in the mid 1600s, and the Dutch preferred to cross
their cattle with English cattle to create a breed more adapted for the North American climate (Bowling, 1942). Because of trading between colonies, it is possible, but not very likely, that the few cattle imported from Holland made their way down to Virginia to diversify the gene pool of post-foundation herd generations.

iv. English and Irish Influence

The English colonized present-day Virginia and Massachusetts, bringing with them English cattle. Cattle were thought to have been imported to the colonies from various regions of England including Devonshire, Somerset, Gloucester, Herefordshire, Straffordshire, and others (Percy, 1979). There is specific documentation that in 1623, four Devon cows were imported from Devonshire into Massachusetts (Brown and Sorrells, 2004); however, it is difficult to determine exactly when the first cattle were introduced to the Virginia Colony, as well as what breed they were.

This difficulty is due to several possible reasons. One is that historical records are often indescript and incomplete. Another reason is that a breed of cattle was often solely defined by its color. For example, when a historical document mentioned red cattle residing in Norfolk and Suffolk regions of the Chesapeake, historians often assumed that these cattle were Devons because of their predominantly red color (Laing, 1954). However, cattle from Somerset and Gloucester also had red coloring (Percy, 1979). Additionally, even if
archaeologists did find a document that confirmed that cattle were brought into the colony at a specific time, the word “cattle” in sixteenth and seventeenth century English referred to all hoofed animals including horses, cows, sheep, goats, and pigs, not specifically bovines. Bovines were specified with terms like “neat cattle” or “kine” (Bowling, 1942). These issues only allow historians to speculate as to how these early bovines came to reside in Jamestown.

Records of early but short-lived settlements suggest the possibility of cattle being present before the founding of the Jamestown in 1607; however, the presence of cattle was not documented by Virginia settlers when they first arrived in Jamestown (Gray, 1958). In 1570, a Spanish Jesuit mission supposedly settled an area on the present-day York River and kept cattle. When Native Americans destroyed the settlement in 1571, it was assumed that any livestock present would also have been killed. However, there is also the possibility that cattle could have escaped and created a feral herd, although there is no documentation of this. It is unclear how many cattle if any were imported during the Jamestown settlement’s first three years. Records are nonspecific and famine struck the area, so any livestock that had existed were gone by 1610 when Lord De la Warr arrived from England, presumably bringing English cattle with him (Bowen and Andrews, 2007). According to historical records, it was only after De la Warr’s arrival that cattle were shipped to the Virginia Colony on a regular basis (Brown and Sorrells, 2004).

Jamestown obtained most of its provisions from the Virginia Company of London until 1624, when the company went out of business (Anderson, 2004).
The company required that imported cattle be of English breed. Despite this, there is still evidence that settlers imported cattle from Ireland as well because many people believed that Irish cattle were superior to English breeds (Bowling, 1942). Irish cattle were known to have been imported by one wealthy settler, Daniel Gookin, who was granted permission from the Virginia Company to import cattle from Ireland in exchange for tobacco in 1620. There is also evidence of the Virginia Colony trading with the West Indies as early as 1633, but it is unknown whether or not cattle were imported from these islands (Bruce, 1896).

Even though the Virginia Company went out of business, mass cattle importations from other colonies or the West Indies likely occurred for about another ten years, at which time they significantly decreased because of a booming and thriving cattle population in Jamestown. By 1649, there were over 20,000 cattle in the Virginia colony (Bowen and Andrews, 2007). Such large numbers of cattle being confined in pens only increased their chance for escape, and there was documentation of free-roaming, wild cattle in the colony in 1639 (Bowling, 1942). A colonist wrote in the mid 1600s that “from the variety of colors distinguishing the horned cattle entered in the appraisements, it would be inferred that there were no distinct breeds in the colony, the original ones having become by repeated crossings so confused in blood as to represent no separate types except in an extremely modified form,” meaning that cattle in the Virginia colony at this time were primarily of mixed breed (Bowling, 1942).

All of this evidence described so far leads many historians to believe that English Devons made up Jamestown’s foundation herd of cattle. Another piece
of striking support is that Colonial Williamsburg’s archaeological team discovered a cattle skull in a 1650 well in Portsmouth, Virginia that nearly exactly resembled a Devon skull (Brown and Sorrells, 2004). Therefore, historians still assume that Devons were the primary breed present in colonial Virginia. However, trade with other colonies and the West Indies and free-roaming cattle populations probably greatly diversified the bovine gene pool by the mid 1600s, so the possibility of other cattle breeds being present in early Virginia cannot be ignored.

vi. Feral Herd Influence

The colonists’ method of cattle husbandry presents many opportunities for genetic variation in colonial cattle populations. There is both zooarchaeological and written historical evidence that English settlers adopted a form of the woodland pasture system of keeping livestock, specifically cattle. This is when cattle were kept in very large fenced-in areas, sometimes up to one thousand acres, and were not fed anything but the grass that they obtained from grazing these areas. These woodland pastures were for communal use, which is a similar practice in British husbandry. Feral herds did form as a result of the freedom given to the cattle and the inevitability of breaching fenced areas, but historical records also reflect the colonist’s clear distinction between wild and domesticated herds. Domestic herds were able to be maintained within this woodland pasture system by enclosing cattle in smaller pens at night to protect them from Native American raids and natural predators (Bowen and Andrews,
By the mid 1600s, specific breeds of cattle could not be distinguished (Bowling, 1942), which was probably largely due to the open nature of the woodland pasture system and the genetic variety that feral populations brought to domestic herds through mating.

vii. Summary of Likely Influences

According to this evidence, there is a significant possibility that cattle brought to the Virginia Colony were not of English origin alone (Figure 1). The Virginia Company granted Daniel Gookin permission to import Irish cattle from Ireland in 1620 (Bruce, 1896). From evidence of trade with other colonies, Dutch cattle could have also been present in early colonial Virginia (Bowling, 1942). Because the long journey across the Atlantic often killed both its human and animal travelers, ships bound for Virginia might have stopped in the West Indies to restock (Anderson, 2004). Therefore, even Caribbean Creole cattle, with its Spanish and African genetic influences, were likely to have been brought to Virginia.

c. Breeds and Haplotypes

In some fields, the term “type” is used to refer to a group of organisms from a specific taxon that presumably share genotypic and phenotypic features, whereas the term “breed” refers to a type of organism from the same geographic
Figure 1—Possible Origins of Virginia Cattle from Importations
The four most likely origins of the foundation herd in Jamestown came from direct importations. The most widely accepted view is that cattle were imported directly from England (yellow), although it is documented that some cattle were also directly imported from Ireland (green) (Bruce, 1896). Because of the long, difficult journey from England, ships sometimes stopped to re-stock supplies and livestock in the West Indies (red) (Anderson, 2004). Trade between the Virginia Colony and the West Indies was also documented as early as 1633 (blue), but whether or not cattle were imported from the West Indies is unknown (Bruce, 1896).
location (Bowen, personal communication, 2009). For the purposes of this research, the term “breed” is used more generally and not necessarily applied to populations of animals of the same geographic region. Scientists theorize that all domestic cattle breeds were derived from one species of now extinct wild oxen, *Bos primigenius* (Bailey et al., 1996). Today, there are two distinct types of domestic cattle: *Bos taurus* and *Bos indicus*. *Bos taurus* cattle are humpless with shorter horns, whereas *Bos indicus* cattle have a hump in their spine at their shoulder blades and have longer horns. Humpless cattle are more typically found in Western Europe, and humped cattle are more commonly found in Asia. These two kinds are considered subspecies because they are interfertile, or capable of reproducing with each other. The domestication of taurine cattle supposedly took place eight to ten thousand years ago in the Near East or present day Southwest Asia and gave rise to all cattle breeds (Loftus et al., 1994b), and this domestication is supported by genetic evidence of European cattle being genetically similar to several populations of cattle in Southwest Asia (Troy et al., 2001). However, the genetic variety of several populations studied provides evidence for an independent domestication of humped cattle long before this ten thousand year time period (Bailey et al., 1996, Loftus et al., 1994b, Troy et al., 2001).

Taurine cattle can be divided into five predominate haplotypes, T, T1, T2, T3, and T4 based on conserved regions in mitochondrial DNA (mtDNA) (Edwards et al, 2004). The T haplotype is the ancestral sequence used for comparison. African cattle identify primarily as the T1 haplotype and can further
be differentiated into T1a and T1*. Caribbean Creole cattle brought over from Brazil were identified as the T1a type, and cattle brought over to America and the West Indies by the Spanish were identified as T1*. Cattle identified with the T2 haplotype are primarily from the Near East, and Western European breeds most commonly identify with the T3 group. Because the T3 haplotype often identifies identically with the T haplotype and Western European breeds are also commonly used as a basis for comparison, T/T3 is often used in the literature to refer to the Western European haplotype (Liron et al, 2006). Japanese cattle are the only group so far to be identified with the T4 haplotype (Edwards et al, 2004). Differences among these haplotype consensus sequences are found in Table 1.

d. Mitochondrial DNA

Mitochondrial DNA has been used for a variety of ancient cattle DNA studies. Using mtDNA, Troy et al. (2001) have shown the origin of European cattle to be from the Southwest Asia region, and Cymbron et al. (1999) have shown Portuguese cattle to contain African genetic influence. Mitochondrial DNA evidence is also commonly used to analyze relationships between breeds, and it has been used to reveal both European and African haplotypes in Caribbean Creole cattle, supporting the existing historical evidence of trade routes through the Caribbean (Magee et al., 2002). To be able to further trace lineages of cattle origin, the most highly variable region of this conserved segment of mtDNA called the displacement loop (D-loop) is used.
<table>
<thead>
<tr>
<th>Position in mt genome</th>
<th>42</th>
<th>49</th>
<th>50</th>
<th>51</th>
<th>57</th>
<th>58</th>
<th>74</th>
<th>82</th>
<th>84</th>
<th>85</th>
<th>93</th>
<th>102</th>
<th>109</th>
<th>110</th>
<th>113</th>
<th>116</th>
<th>117</th>
<th>119</th>
<th>121</th>
<th>161</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0/Ang comparison</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>—</td>
<td>C</td>
<td>T</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>C</td>
<td>T</td>
<td>—</td>
<td>—</td>
<td>T</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>T/T3 consensus</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1 consensus</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2 consensus</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4 consensus</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bos indicus</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Only variable positions are shown. The first row shows the position number in the mt genome (minus 18000) of each variable position in a region of the mitochondria used to distinguish haplotype by Edwards et al. (2003). Some of these variable positions were previously established in an Aberdeen Angus comparison sequence by MacHugh et al. (1999), shown in the second row. Variable positions not present in the Aberdeen Angus sequence are indicated with a dash. Consensus sequences for each taurine haplotype (T/T3, T1, T2, and T4) were taken from Edwards et al. (2003) and used to distinguish differences at these positions. The last row contains a Bos indicus sequence given by Bradley et al. (1998). The Aberdeen Angus comparison sequence used for sequence analysis in this research matches identically to the T/T3 haplotype, or Ancestral/European haplotype. Only 1-2 positions distinguish other haplotypes from one another. The Bos indicus sequence differs at many more positions than any of the taurine cattle.
Mitochondria are organelles present in the cell cytoplasm that provide energy for the cell. They contain their own circular, double stranded DNA, separate from the DNA found in the nucleus of the cell. Because mitochondria are inherited from the mother’s egg during fertilization, changes in mtDNA are often used to trace lineages (Allison, 2007). The D-loop is a region of 911 nucleotides where replication begins in mtDNA. This region has been shown to contain variable positions where mutations more frequently occur, thus providing more evidence of change over time (MacHugh et al., 1999). Researchers have compiled rather extensive databases for mtDNA sequences in many breeds of cattle, so the positions at which nucleotides vary between them are well established. By amplifying a piece of this variable D-loop region, sequencing it, and comparing it to breeds of cattle in the databases, one can speculate as to the cow’s origin.

e. Collagen and Bone Preservation

Collagen is the most common and important structural protein found in animal tissues and is made up of three peptide chains associated together in a helical structure. Type I collagen is found in skin and tendon fibroblasts as well as bone osteoblasts; however, the collagen found in bone has significantly different chemical properties than the collagen found in soft tissues (skin and tendons). Bone collagen is incorporated into a mineral matrix and therefore has a high denaturation temperature and is very resistant to degradation by
collagenases. Because of these chemical differences, collagen is still found in bone samples that are thousands of years old. Once it is separated from its mineral matrix, bone collagen is susceptible to denaturation using heat and acids (Semal and Orban, 1995). Once extracted, the amount of collagen present in ancient bone samples could thus serve as an indicator of the degree of their preservation.

f. Specific Aims for Research

The primary objective of this thesis research was to supplement incomplete historical records with concrete DNA evidence in discovering the ancestry of the Chesapeake region’s colonial cattle. Specifically, three basic goals were pursued in order to achieve the main objective:

1. To adapt and optimize methods for DNA extraction, PCR amplification, and collagen extraction from ancient colonial cow bones.
2. To determine the degree to which the amount of collagen present in ancient bone samples contributes to DNA preservation, if at all.
3. To provide molecular evidence for the origin of colonial cattle by determining the breeds of samples found throughout the region.
II. Materials and Methods

a. Sampling

Samples were obtained from the Faunal Lab of Colonial Williamsburg’s Department of Archaeological Research through Dr. Joanne Bowen, the head zooarchaeologist. Samples were taken from various sites in Virginia throughout the Hampton Roads area, from Colonial Williamsburg to Hampton. Several different elements were used including long bones, metatarsals, teeth, and other bones and were estimated to be from various time periods ranging from 1608 to 1650. Samples were first chosen based on availability. Bones from wells were also sought because MacHugh et al. (2000) showed that DNA was more effectively sequenced from samples that were from a stable environment, such as one that is always completely waterlogged. Although degree of preservation cannot be determined by visible bone condition, teeth as samples were later preferred because it is thought that the enamel and dentine outer layers might provide protection against degradation and weathering (MacHugh et al., 2000).

To prevent cross contamination, each of the following procedures was carried out with each bone individually before another bone was analyzed. An empty 1.5mL tube served as the negative (cross-contamination) control for each bone analyzed. This tube was open and kept in a tube rack in the fume hood during the bone grinding procedure, and it was taken through all the steps of the subsequent procedures as if it contained a DNA sample. In addition to these
negative (cross-contamination) controls, a positive control was also used during the analysis of each bone. This was a modern sample of *Bos taurus* DNA taken from the patella and surrounding tissue of a recently deceased cow. A diluted stock containing 0.5 ng/µL of DNA extract from this fresh DNA sample was used in PCR and gel electrophoresis to ensure that PCR was effectively completed.

In addition to the modern cow sample, cheek cell samples were also collected from two Devon cows living in Colonial Williamsburg as part of the Rarebreeds Program. Sequences obtained from these samples served as comparisons to sequences obtained from ancient samples. Because these cows were certainly of the Devon breed, comparing their mtDNA sequences was a very direct way to assess the breed of ancient samples.

### b. Bone Grinding

The outermost surfaces of each ancient bone sample were initially decontaminated by treatment with ultraviolet light in a Spectrolinker XL-1500 UV Crosslinker for 10 minutes under the “optimal crosslink” setting (Spectronics Corporation). This outermost layer was then sanded away in a fume hood using medium-grain sandpaper. Some longer bones, such as metatarsals, were first cut by hand with a small, bleach-treated hacksaw to achieve more desirable samples of compact bone that would expose more internal surfaces as well as be small enough to fit in the coffee grinder in a later step. The newly exposed bone surface was also treated with UV light in the cross linker for 10 minutes. The
bone sample was placed in a manually controlled Cuisinart Grind Central coffee 
grinder (model DCG-12BC) and ground intermittently until most of the sample 
became a fine powder. Bones analyzed in Spring 2009 were ground using a 
KitchenAid BCG100ER blade coffee grinder.

Three hundred milligrams of fine bone powder was transferred to each of 
two 1.5mL tubes and used for the DNA extraction. The remaining bone powder 
was transferred to a 15mL tube and all bone powder was stored at room 
temperature in a lab drawer. Each part of the coffee grinder was cleaned with 
10% bleach solution to prevent cross contamination of subsequent bone 
samples.

c. DNA Extraction

Total DNA was extracted from the bone powder of ancient samples using 
the Geneclean Kit for Ancient DNA (Q biogene) according to the standard 
protocol developed by MacHugh et al. (2000). Bone cells were lysed by adding 
200µL of 10% SDS, 5µL of 0.5M EDTA, and 200µL Proteinase K to each of the 
three 1.5mL tubes (two samples and one control) and incubated in a Little Shot 
Hybridization Oven (Boekel Scientific, model 230500) at 37°C for 12-15 hours. 
Proteins were denatured by adding 1mL of DeHybernation Solution A (from 
Geneclean kit) to each tube and incubating in the rotator at 60°C for 2-4 hours. 
Tubes were then centrifuged in a Marathon 16KM microcentrifuge (Fisher 
Scientific) at 8000 rpm for 40 seconds to separate the mineralized bone and
cellular debris from the soluble DNA. To bind the DNA in the soluble layer, 1.2mL of DNA glassmilk (from Geneclean kit) and 3mL of DeHybernation Solution A were added to the supernatant and incubated in 15mL tubes in the rotator at 37°C for 2 hours. Tubes were spun in a Centra CL3R centrifuge (Thermo IEC) at 4000 rpm for 1 minute to separate the beads (with DNA bound to them) from the soluble layer. The soluble layer was discarded, the beads were resuspended with 500µL of Salton Wash #1 (from Geneclean kit), and the solution was transferred to a spin filter tube. Tubes were centrifuged in the same microcentrifuge at 8000 rpm for 1 minute and the flow through solutions were discarded. The beads were resuspended and washed with 500µL of Salton Wash #2 (from Geneclean kit), then 500µL of a 1:1 acetone ethanol solution, and twice with 500µL of Ancient DNA Alcohol Wash (from Geneclean kit) using the same procedure for centrifugation and discarding flow through solutions. The last alcohol wash was followed by two centrifugations at 8000 rpm, one for 1 minute and the other for 2 minutes, emptying the catch tube in between.

To elute the DNA from the spin filter column, each filter was placed in a fresh 2mL tube. The beads were resuspended with 200µL of DNA Elution Solution (from Geneclean kit) and tubes were centrifuged at 8000 rpm for 1 minute. These 2mL tubes contained the first elution of DNA extract. Filters were then placed in fresh 2mL tubes, beads were resuspended with 100µL of DNA Elution Solution, and tubes were centrifuged at 8000 rpm for 1 minute. These tubes contained the second elution of DNA extract, or any DNA that did not get eluted the first time. Total DNA concentrations in each of the four samples and
two controls were measured using the NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies) and then stored at -20°C. A typical yield was from 47 to 144 ng/µL.

DNA was extracted from the two Devon cow cheek cell samples using the standard DNA extraction procedure for tissue samples used in Dr. Allison’s laboratory. The cheeks of nine-month-olds Marietta and Zenia were swabbed using spatulas, collected in 50mL Falcon tubes, and stored at -20°C until DNA was extracted. Saliva was resuspended in 10mL of 0.9% NaCl and spun in a Centra CL3R centrifuge (Thermo IEC) at 3000 rpm for 10 minutes to form a pellet of cheek cells. The supernatant was discarded and the cheek cell pellet was resuspended in 500µL of 10% Chelex, incubated in boiling water for 10 minutes, and cooled on ice for 1 minute. Tubes were then centrifuged in a Marathon 16KM microcentrifuge (Fisher Scientific) at 10,000 rpm for 30 seconds. Two hundred microliters of the supernatant from each sample, containing the extracted DNA, were transferred to a new 1.5mL tube and stored at -20°C until PCR.

d. PCR

For all samples, a piece of the Displacement-loop (D-loop) hypervariable region was amplified using polymerase chain reaction (PCR) according to the standard ancient DNA protocol developed by MacHugh et al. (2000). Preparation for PCR was done in an isolated Class II BioSafety Cabinet (Lab
Conco) to provide a separate environment from the one in which bone grinding and extraction was conducted to prevent contamination. To each PCR tube, 41µL of ddH$_2$O, 8µL of PCR master mix, and 1µL of sample DNA template (0.5ng/µL) was added. According to the manufacturer’s recommendations for Turbo Pfu DNA Polymerase, 100-300 ng of DNA template should be added to each PCR for optimal amplification. This change was implemented when later samples were analyzed.

The master mix of PCR reagents was assembled in a 1.5mL tube obtained from this isolated room. For each bone sample analyzed, which included one modern cow sample positive control, four samples from the same ancient bone, and two negative (cross-contamination) controls, seven reactions of master mix were necessary and ten reactions were prepared. A water control was also done in some reactions to test for contamination of PCR reagents. The 10X master mix included 50µL of Pfu Buffer (Stratagene), 2.5µL of each of the three primers BOV-AN4 (F), BOV-AN1 (R), and BOV-AN3 (R) (Integrated DNA Technologies, diluted to 100µM stock in TE), 10µL PCR nucleotide mix (Promega, 10mM), 5µL Turbo Pfu DNA Polymerase (Stratagene), and 7.5µL ddH$_2$O. One forward primer, BOV-AN4 5’-GGTAATGTACATAACATTAATG-3’ and two reverse primers, BOV-AN1 5’-ACGCGGCATGGTAATTAAGC-3’ and BOV-AN3 5’-CGAGATGTCTTATTTAAGAGG-3’ were used as recommended by MacHugh et al. (1999). These are overlapping primers; when used together, two different size fragments are created. BOV-AN4-BOV-AN1 targets a smaller, 218 bp fragment, and BOV-AN4-BOV-AN3 targets a larger, 375 bp fragment (Figure
**Figure 2- Arrangement of D-loop Target Region within *Bos taurus* Mitochondrial Genome**

a) The *Bos taurus* mt genome is 16338 nucleotides in length. The entire D-loop region is 911 nucleotides in length and is found between positions 15792 and 364. Two short fragments of the D-loop were targeted for PCR. b) The D-loop is the most variable part of the *Bos taurus* mt genome. The 375 bp fragment within the D-loop region was targeted by primers BOV-AN4 (F) (binds at positions 15960 to 15981 on the mt genome) and BOV-AN3 (binds at positions 16315 to 16335 on the mt genome). The 218 bp fragment was targeted by primers BOV-AN4 (F) and BOV-AN1 (R) (binds at positions 16159 to 16178 on the mt genome). All three primers were used in PCR initially and yielded two fragments, but because the longer 375 bp fragment is more difficult to amplify in ancient samples (MacHugh et al., 2000), only the 218 bp fragment was targeted in subsequent samples. Only the 218 bp fragment was targeted for sequencing. c) Within the 218 bp fragment of the D-loop, the most variation in base sequence is found between positions 16032 and 16063, known as the hypervariable region (Bradley et al., 1996). There are ten positions where base variation has been documented between different *Bos taurus* breeds in this region. There are 21 other places of variation found throughout the rest of the 218 bp fragment (MacHugh et al., 1999). All these positions were analyzed in each sample sequence obtained through this research as well as in published sequences of known cattle breeds to determine haplotypes.
2). After nonspecific streaking was observed on gels of samples amplified using all three primers, BOV-AN3 was replaced with water, yielding more defined bands on gels.

A hot start at 95°C was used to ensure that the lid was already at the correct temperature when tubes were put in the machine. Samples were initially denatured for 3 minutes at 94°C. Each amplification cycle consisted of a 40 second denaturation step at 93°C, a 40 second annealing step at 55°C, and a 40 second extension step at 72°C, and this was repeated for a total of 50 cycles. A final extension for 4 minutes at 72°C concluded the PCR. Samples were stored at -20°C.

For Devon cow samples, PCR was carried out using the same protocol as was used for ancient samples, with a few changes. All three primers were used because samples were fresh and likely to be able to amplify both fragment sizes well. Five microliters of DNA extract (as recommended by the standard PCR procedure for DNA from cheek cells), 8µL of master mix, and 37µL ddH₂O were added to each reaction. Only 30 cycles were completed as recommended by the standard PCR procedure (Allison, personal communication, 2008). Cycle times and temperatures were kept the same as ancient samples because they were optimized to our Turbo Pfu DNA Polymerase.
e. DNA Gel Electrophoresis

Gel electrophoresis was performed to visualize amplified DNA of both ancient and Devon cow samples. Tubes were prepared with 6µL of PCR product, 3µL of ddH₂O, and 1µL of glycerol dye (0.2M EDTA, 50% glycerol, 0.25% bromophenol blue, 0.25% xylene cyanol) to yield 10µL total. Samples were loaded and run in a 2% agarose gel in 1X TBE (89mM Tris base, 2.7mM EDTA, 89mM boric acid) for 1 hour at 100V. The gel was stained for 5 minutes in a 0.01% ethidium bromide solution and destained in ddH₂O for 20-25 minutes. Gel appearance was captured using Polaroid 667 film under UV light. In Spring 2009, gels were captured using a BioRad Gel Doc XR documentation system with Quantity One analysis software (v4.6.1).

f. PCR Product Purification

Once clearly defined bands of the expected sizes 218 bp and 375 bp on the agarose gel confirmed the presence of amplified DNA, the PCR product was purified and prepared for sequencing. All samples were purified using QIAGEN PCR Purification Kit according to the manufacturer’s protocol for a sample volume of 50µL. This procedure was carried out for all bone samples. Samples were eluted in 50µL of Buffer EB, transferred to 1.5mL tubes, and stored at -20°C.
g. DNA Sequencing

Samples were sent to Yale DNA Analysis Facility on Science Hill to be sequenced and were first prepared according to their specifications. Samples were analyzed using the Nanodrop spectrophotometer to determine their DNA concentration after purification to ensure that 10-20ng of DNA was added to each tube. Eight-strip PCR tubes were used, each containing 1µL of 3.2pM Big Dye Sequencing primer (BOV-AN4 forward or BOV-AN1 reverse, accordingly), 10-20ng of DNA sample, and ddH₂O to total a volume of 12µL in each tube. Each sample was sent in duplicate so that one tube contained the forward primer and the other tube of that sample contained the reverse primer. Samples were sent at room temperature overnight via FedEx to Yale DNA Analysis Facility. In Spring and Fall 2006, samples were sequenced in-house using an ABI 3100-Avant Sequencer and analyzed using ABI Sequencing Analysis software (v5.1.1).

h. Sequence Analysis

Sequences were obtained electronically from Yale DNA Analysis Facility and imported into CLC Sequence Viewer 6.0. The primary sequence used for comparison was the 218 bp fragment extrapolated from a published *Bos taurus* complete mitochondrial genome sequence (Accession #V00654) (Anderson et al., 1982). The 218 bp fragment was obtained by aligning BOV-AN4 (F) and
BOV-AN1 (R) to this complete genome sequence and selecting the nucleotides in between and including the outer end of each primer. This 218 bp fragment showed 100% homology with the Aberdeen Angus consensus sequence, which is the most common cattle breed in Europe and is used in many different genetic studies as a comparison sequence (Bradley et al., 1996). Each sequence obtained from ancient samples was aligned with the 218 bp fragment and number and position of mismatches were recorded and analyzed. Each ancient sample was also aligned with each of the sequences obtained from the Devon cows and analyzed in the same way.

i. Collagen Extraction

Collagen was extracted from bone powder by acid and heat using a procedure adapted from Longin (1971), Brown et al. (1988), and Semal and Orban (1995). Samples were extracted in duplicate, and 200mg of powder in a 15mL tube were used for each extraction. Two milliliters of 2M HCl were added to each sample and rotated in the Little Shot Hybridization Oven (Boekel Scientific) at room temperature for 17.5 minutes. Samples were then spun using a Centra CL3R centrifuge (Thermo IEC) at 3000 rpm for 2 minutes, after which the supernatant was drawn off and discarded. Two milliliters of 0.2M HCl were added to each sample, and samples were incubated at 95-100°C for 50 minutes. During that incubation, tubes were vented and swirled every 15-20 minutes. After this incubation, samples were centrifuged using the same machine at 3000
rpm for 20 minutes. Supernatant from each sample was pushed through a 0.2 µm nylon filter using a 10-mL syringe into a fresh 15mL tube on ice. Samples were then placed in Snake Skin dialysis tubing with a molecular weight cut-off of 10,000 Da (Pierce Chemical Company) in 1600mL of ddH2O, being gently stirred overnight at 4°C. The next morning, 500 µL aliquots of samples were transferred to 1.5mL tubes and stored at -80°C until a sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) was run with multiple samples.

j. SDS-PAGE

An 8% poly-acrylamide gel was made according to the standard procedure (30% Acrylamide/Bis solution, 29:1, 3.3% C). One tube was prepared with 10 µl of BioRad Kaleidoscope pre-stained protein standard, 5 µl of ddH2O, and 5 µl of 2X SDS-PAGE sample buffer, which was made according to the standard protocol (125mM Tris, pH 6.8; 1% SDS; 5% glycerol; 0.005% bromophenol blue; 20mM DTT). Sample tubes were prepared with 5 µl sample, 5 µl ddH2O, and 10 µl 2X SDS-PAGE sample buffer. Sample tubes were denatured in a boiling water bath for 3 minutes prior to gel loading. The gel was run in 1X SDS-PAGE Running Buffer (192mM glycine; 25mM Tris; 3mM SDS, pH 8.2) at 150V for about 1 hour and 20 minutes, or until the dye was near the bottom of the plate. The gel was rinsed in 100mL ddH2O for 3 five-minute rinses, water being discarded each time. The gel was stained using 20mL of SimplyBlue SafeStain (Invitrogen) for 1 hour with gentle rocking, and then destained with
Figure 3—Methods Involved in the Analysis of mtDNA and Collagen from Ancient Bone

a) Bone samples were ground into a powder to be used for both DNA (b) and collagen (c) analysis. b) Total cellular DNA was extracted from bone powder, but only the 218 bp fragment of mtDNA was amplified using PCR. The PCR product was visualized on a gel using DNA gel electrophoresis. If bands of DNA were found on the gel, the PCR product was purified and sequenced. Sequences from multiple samples were then analyzed. c) Collagen was also extracted from bone powder and was visualized on a gel using SDS-PAGE. Relative amounts of collagen were compared between samples.
100mL of ddH$_2$O for at least 1 hour. The gel image was then captured using a BioRad Gel Doc XR documentation system with Quantity One analysis software (v4.6.1). A summary of the methods used for this research can be found in Figure 3.

III. Results

a. Optimizing Procedures for Ancient DNA

i. Modifications Made to the DNA Extraction Procedure

Ancient DNA is inherently difficult to work with because of factors that contribute to its degradation such as age, fluctuating environmental conditions, and contamination. Therefore, optimizing the DNA extraction and amplification procedures were crucial to isolating as much of our DNA of interest as possible. As mentioned in Materials and Methods, DNA was extracted from bones using a kit designed especially for ancient samples. However, a few changes were made to the manufacturer’s recommended procedure in order to minimize contamination and maximize DNA yield. The first few changes were added to purify the DNA more effectively. Before the ethanol wash during DNA isolation, a 1:1 acetone to ethanol wash was added to wash away contaminants without also washing away DNA. Humic and fulvic acids, which are found in decomposing plant material in the soil, are difficult to remove from bone samples because of
their ability to bind with multivalent ions such as calcium and iron in bone samples (Schnitzer, 1969). Contaminants like these as well as other environmental contaminants tend to inhibit PCR, which would make obtaining a sequence from such contaminated samples very difficult (Allison, personal communication, 2006). For this reason, an additional ethanol wash was added after the recommended one, as well another round of centrifugation after the ethanol was added, to minimize contamination and ensure that all ethanol was removed from the DNA before elution, as ethanol can inhibit PCR.

Despite these changes, some samples still did not yield sequenceable DNA, which is understandable given the age and condition of the samples. No quantifiable changes in DNA purity resulted even with these protocol modifications. The 260/280 value given by the spectrophotometer was used as an indicator of sample purity. The 260/280 value represents the ratio of DNA (nucleotide) to protein (or other contaminants that absorb 280 nm light) in the sample, which is based on the wavelength of light that is shone through the sample and the absorbance measured. The optical density 260/280 values were not any closer to being the optimal 1.9 ratio for DNA samples with these protocol changes, which was probably due to other factors that lower DNA purity, such as environmental contaminants present in the samples (see Appendix for DNA concentrations and purity values). Even though these changes did not result in increased DNA purity, they were continued as a means of precaution. In the case that they did increase the purity of the sample even though it was not reflected by the 260/280 value due to other contaminants, it was more beneficial
to include these steps than not. These modifications could only increase the rate of success or the number of samples from which sequences were obtained and were therefore permanently incorporated into this protocol for the remainder of samples analyzed.

ii. Modifications Made to the PCR Amplification Procedure

Changes in the PCR procedure reflect a more effective way to get quality copies of the fragment of interest. For the majority of samples analyzed, primers BOV-AN4 (F), BOV-AN1 (R), and BOV-AN3 (R) were used, as recommended by MacHugh et al. (1999). However, when attempting to amplify ancient samples, shorter fragments are more readily and effectively amplified than longer fragments (MacHugh et al., 2000). This is probably due to the poor condition of the DNA and the fragility of the bonds holding it together. As a result, many samples displayed streaking on the DNA electrophoresis gel when all three primers were used in PCR but showed clear bands if only the smaller 218 base pair (bp) fragment was targeted (Figure 4). After determining this, subsequently only BOV-AN4 (F) and BOV-AN1 (R) were used, resulting in more defined bands and more effective sequencing.

The amount of sample DNA included in each PCR was also modified throughout the course of this research. Originally, only 0.5 ng of DNA was added to each reaction; however, the yield of product was improved by increasing the amount to between 50 and 200 ng of sample DNA. This was probably because
Figure 4—Representative Optimal and Suboptimal Visualization of DNA

In lane 1 of all gels, standard Hyperladder II was used for size comparison. The modern cow sample (m1) served as the positive control during PCR and was always run in lane 2. The normal modern sample showed two bands, representing the products from BOV-AN4 (F) and BOV-AN3 (R) at an ideal length of 375 bp, and from BOV-AN4 (F) and BOV-AN1 (R) at an ideal length of 218 bp. In both a) and b), the BOV-AN3/4 fragment was shorter than 375 bp, which is typical of ancient samples (MacHugh et al., 2000).

a) This picture shows a representative optimal gel of samples B1 and B2. Lanes 3-6 contained DNA samples, which show faint bands at each expected size. Lanes 7-8 contained negative (cross-contamination) control samples, thus showing no bands of DNA.

b) This picture shows a representative suboptimal gel of sample B3. Lanes 3-6 contained DNA samples, and lanes 7-8 contained negative (cross-contamination) control samples. Lanes 3-8 all show streaks, which represent DNA of differently-sized fragments rather than DNA of a specific size. Streaking was likely caused by nonspecific interactions with the primers, polymerase, and DNA template due to suboptimal conditions such as contamination. This streaking often happened when all three primers were used in PCR and 50 cycles were completed during PCR. Despite the suboptimal visualization of DNA, some samples such as this one still yielded a good quality sequence.
this larger amount of DNA fell within the standard recommended range for Turbo Pfu Polymerase and therefore showed more effective amplification as visualized on a gel and also yielded greater sequencing success. When only 0.5 ng of sample DNA was added to each PCR reaction, it was simply not enough to provide the polymerase with a good quality full-length copy of the sequence of interest, and DNA amplification was highly unsuccessful.

Lastly, the number of cycles performed during PCR greatly affected the amount and quality of DNA templates after PCR. MacHugh et al. (1999) recommended to complete 50 cycles, which is more than what is recommended for standard PCR using modern DNA samples. Streaking on the DNA gel was observed when all three primers were used and 50 cycles of PCR were completed, but streaking was not present when all three primers were used but only 30-35 cycles were completed. It would seem that more cycles caused streaking because with the increased number of temperature changes during PCR and suboptimal conditions such as contamination, more nonspecific binding interactions occurred between the DNA template and both the primers and the polymerase. This would inhibit a good quality full-length copy of the fragment of interest from being successfully amplified. However, only completing 30-35 cycles did not seem to result in enough copies to be visualized on a gel, as most gels run with such PCR products showed no bands or streaks. This was probably because not enough copies of the fragment of interest were made with only 30-35 cycles. Therefore, there was a tradeoff between the amount and quality of the resulting DNA fragment. In the end, completing 50 cycles was
more effective at obtaining enough DNA to be sequenced, even if DNA fragments were shorter.

b. Amount of Collagen Does Not Clearly Indicate Bone Preservation

Because this objective was added to the project after many of the samples had already been analyzed and completely consumed, collagen data were not available for some samples. This analysis of the amount of collagen present in each sample was only intended to be a screening procedure to determine if a particular sample would likely yield sequenceable amounts of DNA. Because the degree of bone preservation is not always apparent from its outward appearance, the amount of collagen present could correlate more directly to the condition of the bone because of its structural role in connective tissues. Therefore, the more collagen that is present in a sample, the less DNA degradation that is thought to occur (Collins et al., 1995).

Extracting collagen from bones was first used for radiocarbon dating, and is still used today. The most common method used to extract pure collagen has been that of Longin (1971), which is based on the solubility of collagen in a hot, mild acid. Other researchers have adapted Longin’s procedure to improve the purity and quantity of collagen obtained from these extractions (Brown et al., 1988, Semal and Orban, 1995). Using these protocols as a foundation, we
developed a procedure to extract collagen from our ancient cow bone samples, which can be found in detail in Materials and Methods.

Because of the materials, equipment, and time available to us, our collagen extraction procedure was modified primarily from that of Semal and Orban (1995). Semal and Orban found that bone powder pretreatment with 2M HCl for 17.5 minutes at room temperature was optimal for maximizing inorganic compound dissolution and minimizing collagen degradation. Thus, we adopted this as our pretreatment step. Both Longin (1971) and Brown et al. (1988) suggest refluxing the collagen sample in a mild acid for 10-20 hours to solubilize collagen; however, because this equipment and time interval were not available, our procedure was modified. Instead of refluxing, we incubated our samples in an oven with intermittent agitation for a much shorter time period: 50 minutes. Semal and Orban (1995) suggested that shorter solubization times actually increase the yield of the larger peptide fraction of collagen (>10 kDa), which was desirable for this project.

Collagen was extracted and visualized using SDS-PAGE for samples B3, B4, and J1-J4. B3 was a bone from which sequence data was obtained, whereas we were unable to obtain a sequence from B4 or J1-J4 (Figure 5). Considering relative band brightness between samples B3 and B4, B3 did have the brighter band. However, J1, J2, and J4 also showed relatively brighter bands than J3, but no sequences were obtained from any of these four bones.
Figure 5—Relative Amounts of Collagen in B3, B4, J1-J4 Samples

a) Collagen extracted from samples B3 (lanes 2 and 3) and B4 (lanes 4 and 5) was separated by gel electrophoresis. Five microliters of each sample were loaded into lanes 2 and 4, and 10μL of each sample were loaded in lanes 3 and 5, hence the difference in brightness. Kaleidoscope protein standards were loaded in lane 1 for size comparison. Collagen is about 80kD, indicated by the red arrow. In all sample lanes, there is a much darker concentration of collagen at or slightly below this size marker. B3 lanes also contain darker bands than B4 lanes, indicating more collagen presence in the B3 sample. Interestingly, B3 yielded a good quality sequence whereas B4 did not, which could indicate that more collagen present in an ancient bone sample increases its preservation. b) Collagen extracted from samples J1 (lane 2), J2 (lane 3), J3 (lane 4) and J4 (lane 5) and separated by gel electrophoresis. Five microliters of each sample were loaded into lanes 1-5. J1, J2, and J4 show similar band brightness, indicating similar amounts of collagen. J3 shows a light smear, indicating less collagen. Despite these differences in brightness, none of the J samples yielded good quality sequences. All of the J samples show a concentration of collagen between the 50 and 37kD marks rather than the expected 80kD mark. This could be because these samples were stored at -80 degrees Celsius for several weeks before SDS-PAGE was done, whereas B3 and B4 samples were run out on a gel the day after collagen was extracted. The freezing and thawing process can damage proteins. Taking a) and b) together, more collagen presence in ancient samples does not indicate increased DNA preservation.
a. Molecular Evidence Suggests European Cattle

The primary objective of this thesis research was to obtain mtDNA sequences from ancient bone samples. By analyzing the differences between each sequence and a standard comparison sequence, conclusions potentially can be made with regards to which breed each sample belongs. The long-term goal of this project is to use these findings to determine the place(s) of origin of Jamestown’s first cattle and perhaps deduce possible routes for traveling here. Viable sequences were obtained from five different ancient specimens to date from four different archaeological sites, and duplicate sequences were obtained from both Devon cow samples (Table 2). Primers BOV-AN4 (F) and BOV-AN1 (R) were used in all sequencing reactions, which targeted the 218 bp fragment.

Sequences from samples A2 a1 (F), B3 d1 (F), T2 a1 (F), and T2 b1 (F) showed an exact match with the 218 bp comparison sequence (see Appendix for sequences, chromatograms, and alignments). The sequences from samples B1 b1 (F) and T1 a1 (F) showed exact matches as well, with the exception of a few nucleotide additions or eliminations that can be explained by the chromatograms (Figure 6). The sequence from sample B1 b1 (F) showed an extra adenine (A) at position 35 (position 103 in the 218 bp fragment). When the chromatogram was examined, however, the sequencer read two A’s for one peak, which really should have only been represented by one A. The sequence from sample T1 a1 (F) showed similar results. It matched exactly with the 218 bp fragment once discrepancies between the chromatogram and the base sequence were
Table 2—Successfully Sequenced Ancient Bone Samples

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Site Name</th>
<th>Site ID</th>
<th>Element</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2 a1 (F)</td>
<td>Martin’s Hundred</td>
<td>10064</td>
<td>1st phalange</td>
<td>1623-1645</td>
</tr>
<tr>
<td>B1 b1 (F)</td>
<td>Nansemond Fort</td>
<td>_____</td>
<td>4th tarsal</td>
<td>ca. 1650</td>
</tr>
<tr>
<td>B3 d1 (F)</td>
<td>Nansemond Fort</td>
<td>66-44SK192/3A</td>
<td>metacarpal</td>
<td>ca. 1650</td>
</tr>
<tr>
<td>T1 a1 (F)</td>
<td>Coffee House</td>
<td>17KD277</td>
<td>tooth</td>
<td>1755-1767</td>
</tr>
<tr>
<td>T2 a1 (F)</td>
<td>Wren Yard</td>
<td>16GA232</td>
<td>tooth</td>
<td>_____</td>
</tr>
<tr>
<td>T2 b1 (F)</td>
<td>Wren Yard</td>
<td>16GA232</td>
<td>tooth</td>
<td>_____</td>
</tr>
</tbody>
</table>

The sample ID column represents our system of labeling identifying each sample. The site name and site ID are used by the Department of Archaeological Research in Colonial Williamsburg to identify each element. The element column includes the anatomical origin of the sample. The date column shows the estimated time period from which the material from each of these archaeological sites originated. The abbreviation “ca.” stands for *circa*, which means “near.” The material from the Nansemond Fort site is estimated to originate somewhere around 1650.
Figure 6—Chromatograms of Disputable Sequences in B1 and T1 Samples

a) In sample B1 b1 (F), the only place it did not match identically to the Aberdeen Angus 218 bp comparison sequence was at position 103 in this comparison sequence, indicated by the red arrow. The sequence shows an extra A at this position when this peak more clearly represents one A rather than two. Sample T1 a1 (F) showed similar sequence discrepancies, as seen in figures b), c), and d). b) At position 82 in the 218 bp sequence, the chromatogram of T1 a1 (F) shows a peak that more closely resembles that of a double A than a single A peak. c) T1 a1 (F) also should have reflected an A at position 148 in the comparison sequence because of the triple A series seen in the chromatogram. d) An extra A is reflected in the T1 a1 (F) sequence at position 210 in the 218 bp sequence when there is clearly only one peak at that position. With all these discrepancies accounted for, both samples B1 b1 (F) and T1 a1 (F) would match identically to the Aberdeen Angus comparison sequence.
resolved. This sequence was missing an A between positions 10 and 11 (81 and 82 in the 218 bp fragment), which was a double peak that was only read as a single peak by the machine. There was also an A missing between positions 77 and 78 (147 and 148 of 218 bp fragment), which was a triple peak that was read as double peak. Lastly, at position 139 (210 in 218 bp fragment), there was an extra A as a result of the sequencer reading one peak as a double peak. Taking these factors into account, all six sequences from all five ancient samples showed an identical match to the 218 bp Aberdeen Angus fragment.

Two sequences were obtained from each Devon cow specimen, both of which matched identically with each other for each specimen along their regions of overlap. The sequences from Zenia showed an exact match with the 218 bp fragment, just like the ancient samples. However, the sequences from Marietta consistently showed two positions of difference from the 218 bp fragment (Figure 7). Interestingly, these two cows are half sisters from the same stud but different mothers. The difference in mtDNA reflects its maternal inheritance. Therefore, this technique could be useful to the Rarebreeds Program in Colonial Williamsburg to track familial relationships in the present-day herd.

One sequence from Marietta included a thymine (T) at position 3 (99 in 218 bp fragment) instead of a guanine (G), and a thymine at position 30 (126 in 218 bp fragment) instead of a cytosine (C). The other Marietta sequence only contained the second of the two variations because the sequence was shorter and did not include position 30 at all. Both positions 99 and 126 are known variable positions in this hypervariable region of the D-loop, but these specific
Figure 7—Alignments of Representative Successful Samples with the 218 bp Comparison Sequence

Ancient samples that yielded good-quality sequences included A2 a1 (F), B1 b1 (F), B3 d1 (F), T1 a1 (F), T2 a1 (F), and T2 b1 (F). Additionally, two sequences were obtained for each Devon cow cheek cell specimen (Marietta 1 & 2, Zenia 1 & 2). All these sequences were each aligned with the 218 bp comparison fragment and analyzed; however, only some were shown here. Samples B1 b1 (F) and T1 a1 (F) were not included because neither matched identically to the 218 bp sequence, differences for which were accounted by disputable chromatograms (see Figure 7). Only one sample from both T2 and Zenia was shown to avoid repetition. Places where the two sequences align are shown in red, and places where the two sequences do not align are shown in black. Samples shown in a), b), c), and f) all matched identically to the 218 bp sequence where the two sequences overlapped. Sequence Marietta 1 from a Devon cow (d) showed two places of difference (positions 99 and 126 in the 218 bp fragment), while another sequence from the same cow, Marietta 2 (e), only showed one of these variable positions (position 126). This is probably because Marietta 2 was shorter than Marietta 1 and did not contain position 99 in the sequence obtained.
Figure 7, cont.—Alignments of Representative Successful Samples with the 218 bp Comparison Sequence
differences are not both documented in any of the widely studied European breeds (MacHugh et al., 1999). Having no published Devon sequences to which to compare these, this could be considered a new specific haplotype, similar to the haplotypes published by MacHugh et al. (1999) and Bradley et al. (1996). However, because ancient samples identically matched a Devon sequence and an Aberdeen Angus sequence, it cannot be determined which breed the samples represent because of the lack of variation in the D-loop hypervariable region.

IV. Discussion

The purpose of this thesis research was to supplement incomplete historical records with concrete DNA evidence in discovering the ancestry of the Jamestown Colony’s foundation cattle population. Effective procedures were developed to extract and amplify mtDNA from ancient bone samples, as was a method for collagen extraction from ancient bone powder. The relative amount of collagen was not found to be indicative of the ability to obtain a good-quality sequence from a particular sample; more collagen did not imply better DNA preservation. Lastly, neither the breed(s) nor the specific origin of cattle present in the foundation herd can be determined from the D-loop hypervariable region mtDNA sequence of a particular sample because not as much variation was present in this region as originally thought. Because ancient samples identically matched a Devon sequence and an Aberdeen Angus sequence, the breed of each sample cannot be determined.
a. Distinguishing Breeds Using mtDNA

Based on historical records, most historians believe that Devon cows were the predominant breed in the earliest years of the Jamestown settlement. Regarding molecular techniques, it is widely accepted in the ancient cow DNA analysis community to use the D-loop to compare differences between species because it is the most variable region of the bovine mtDNA control region (Figure 8) (Loftus et al., 1994a, Bradley et al., 1996, MacHugh et al., 1999). Therefore, it is the optimal region to target because it distinguishes all bovines yet provides some means to differentiate breeds of phenotypically similar cows. However, differences both within each breed and between breeds were not completely consistent, making breed identification problematic. MacHugh et al. (2000) was also unable to distinguish between a Scandinavian or Irish origin of the population he studied because of “shallow levels of mtDNA diversity present in European taurine cattle populations.” The samples used in this research did not capture this variation in the D-loop region either, making breed distinctions virtually nonexistent.

To properly assess the variation of the sequence of interest, published D-loop sequences from eight different breeds of cattle found throughout Europe, South America, and the Caribbean were analyzed (sequences taken from Troy et al., 2001, Bradley et al., 1996, Komatsu et al., 2004, and Achilli et al., 2008) (Table 3). As would be expected in any population, there is genetic variation among individuals of the same breed from polymorphisms within the population,
Figure 8—Variable Positions in the 218 bp Fragment of the Aberdeen Angus Comparison Sequence

a) The 218 bp fragment was obtained by aligning primers BOV-AN4 (F) and BOV-AN1 (R) to a widely-used published sequence of the entire Bos taurus mt genome (Acc.# V00654) (Anderson et al., 1982). The alignment shown here includes this 218 bp fragment and the Aberdeen Angus 386 bp fragment published by MacHugh et al. (1999) (Acc.# US1816) that showed all the positions of variation within Bos taurus breeds. Only part of the Aberdeen Angus sequence is shown. Red indicates where the two sequences align and black indicates where they do not. The 218 bp fragment overlaps with the Aberdeen Angus sequence beginning at position 58 and identically matches it for the remainder of the sequence. b) The entire sequence of the 386 bp Aberdeen Angus fragment is shown. Within this fragment are 39 variable positions, 31 of which are present in the portion that overlaps with the 218 bp fragment. This overlap is shown with brackets, beginning at position 58 of the 218 bp fragment. Variable nucleotides are shown in red and numbered in green according to their position in the 218 bp fragment. The remainder of the variable positions that occur outside the 218 bp region are numbered in blue according to their position in the entire mt genome minus 16000, according to MacHugh et al. (1999). The 218 bp fragment used for sequence analysis throughout this research captured most of the variable positions published by MacHugh et al.(1999).
Table 3—Variable Positions in the 218 bp Fragment in *Bos Taurus* Breeds

<table>
<thead>
<tr>
<th>Position in genome</th>
<th>42</th>
<th>47</th>
<th>52</th>
<th>53</th>
<th>56</th>
<th>58</th>
<th>62</th>
<th>66</th>
<th>69</th>
<th>74</th>
<th>76</th>
<th>84</th>
<th>85</th>
<th>103</th>
<th>113</th>
<th>118</th>
<th>122</th>
<th>127</th>
<th>130</th>
<th>131</th>
<th>133</th>
<th>135</th>
<th>137</th>
<th>138</th>
<th>139</th>
<th>143</th>
<th>164</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position in fragment</td>
<td>84</td>
<td>89</td>
<td>91</td>
<td>92</td>
<td>95</td>
<td>96</td>
<td>98</td>
<td>100</td>
<td>104</td>
<td>110</td>
<td>115</td>
<td>116</td>
<td>120</td>
<td>127</td>
<td>130</td>
<td>132</td>
<td>135</td>
<td>143</td>
<td>149</td>
<td>160</td>
<td>161</td>
<td>164</td>
<td>172</td>
<td>173</td>
<td>177</td>
<td>178</td>
<td></td>
</tr>
<tr>
<td>Aberdeen Angus A</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aberdeen Angus B</td>
<td>G</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Devon</td>
<td>1</td>
<td>T</td>
<td>T</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jersey 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jersey 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friesian 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friesian 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kerry 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kerry 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bolivian Creole 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bolivian Creole 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cuban Creole 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cuban Creole 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Only variable positions are shown. The first row shows the position number of each of the variable positions documented by MacHugh et al. (1999) in the entire mt genome minus 16000. The second row shows the position of each of these places of variance within the 218 bp fragment targeted for this research. The third row shows the nucleotides at each of the variable positions in the Aberdeen Angus sequence used for comparison in MacHugh et al. (1999). The subsequent rows show nucleotide differences in two individuals from the same breed. The Devon sequence was obtained from this research, and the other sequences were obtained from published papers (except for the Bolivian and Cuban Creoles) (Achilli et al., 2008, Bradley et al., 1996, Komatsu et al., 2004, Troy et al., 2001). Breeds were chosen based on their place of origin. The Aberdeen Angus is from Scotland but is the most common breed in Europe. The Devon is from England, the Jersey is from France, the Friesian is from the Netherlands, and the Kerry is from Ireland. The Colombian Creole now exists in Colombia but was formed from cattle taken from the Spanish settlements in the West Indies. A similar situation exists for both the Bolivian and Cuban Creoles, which now inhabit Bolivia and Cuba, respectively. The Colombian Creole sequences both contain one additional place of variation at position 21, where they have a thymine (T) instead of an adenine (A). This position was not established in MacHugh et al. (1999) and was therefore not shown here. Only 1-3 nucleotides differ among the given breeds, if any. Some individuals from different breeds do not have any places of variation from the Aberdeen Angus comparison in the 218 bp fragment (indicated by *), revealing a lack of variability in this region of the D-loop.
which only amounts to 1-3 nucleotides, if at all. Perhaps unexpectedly, the variation between each of the breeds is the same, from 1-3 nucleotides. From this evidence alone, it seems as though this fragment of the D-loop is not as variable as it should be for being a reported means of differentiation of cattle breeds. Furthermore, we have also confirmed that the expected 218 bp sequence does indeed include most of the variable positions established by MacHugh et al. (1999), so variation was most definitely expected between cattle breeds as well as between ancient bone samples.

Virtually no variation in sequences obtained from ancient bone samples was found. Every ancient sample sequence and one Devon cow sequence was identical to the comparison sequence, which was of Aberdeen Angus breed. The Aberdeen Angus sequence was established and used for comparison by MacHugh et al. (1999) because this breed is the most predominant one in Europe. This lack of variability between samples could be due to one of several factors. One is that perhaps an insufficient number of specimens were analyzed and therefore did not capture the variability of the original cattle population. Although MacHugh et al. (1999) was able to show variability in the D-loop region between breeds that he analyzed, he also obtained sequences from fourteen individuals from the same breed, capturing more of the variability present in the gene pool. Another reason for lacking variability is that ancient samples taken from the same site or nearby sites might be genetically related, contributing to the lack of variability between samples. Another consideration is that there is not as much variation in this region of the D-loop as initially expected. Lastly, cattle
from these early colonial Virginia sites could all be of the same breed, Aberdeen
Angus, Devon, other breeds, or a combination of these.

From this molecular evidence, it cannot be determined with certainty what
breed(s) of cattle made up Jamestown's foundation herd. Because sequences
from each breed do not consistently differ at the same positions, we are not able
to make breed distinctions based on these sequences of mtDNA. Additionally,
the exact same sequence identifies its samples as both Devon and Aberdeen
Angus, which leads us to believe that this targeted region of the mtDNA is not
that variable after all. However, one Devon cow sequence presented a different
haplotype within that breed, which leads us to believe that if more Devon cow
samples were sequenced, perhaps we would discover additional different
haplotypes or places of variation, just as was observed by MacHugh et al.(1999)
and Bradley et al. (1996) in breeds like the Friesian, Jersey, and others.

b. Future Directions

Many more samples, both ancient and modern Devon cow alike, must be
analyzed in order to establish stronger molecular evidence for the presence of
one breed or another in the original cattle population. With more Devon
sequences, more haplotypes can be discovered and defined within the breed,
possibly expanding its range of matches and further contributing to finding the
breed and origin of early colonial cattle. Perhaps more regions of the
hypervariable region of the D-loop should be targeted for sequence analysis to
increase the likelihood of variability. D-loop segments that help characterize haplotypes (T1, T2, T3 and T4) using additional primers like BOV-AN2 and BOV-AN5 could be targeted in hopes of discovering the continent of origin of Jamestown’s cattle. Because haplotype characterization reveals the larger area of origin, more specific places of origin such as country or region could be supported or eliminated.

V. Acknowledgments

Over the past four years, numerous people have contributed to the progress of this project. First of all, I would like to thank Dr. Lizabeth Allison for giving me the opportunity to do research in her laboratory, guiding me with my project, answering all my many questions, and always being available to discuss ways to improve a procedure or fix a problem. Next I would like to thank Dr. Bowen and Steve Atkins for their collaboration, historical perspective, and bone-supplying efforts. And I greatly appreciate Dr. Forsyth and Dr. Wawersik for serving on my thesis committee and helping me make sense of my sequences.

From the Allison lab, I would first like to thank Vinny Roggero for never hesitating to drop whatever task he was engaged in to help me find a piece of equipment or talk about an approach to solve a problem. Wes Northam also helped immensely with PCR and sequencing-related questions as he is the lab expert in those areas. Matt Grespin and Chris Siebert also deserve my thanks for helping me refine my lab skills in my early years of research. Additionally, I
wish to thank all of those Allison lab members who have been a part of Team Moo Moo, especially Maura McAuliffe and Sarah Lehman. Maura created this project and gave me a solid background for conducting molecular biology research, and Sarah was my faithful lab partner who contributed many of our sample sequences. Of course, thanks also to the entire Allison lab for providing encouraging words and comedic relief throughout my years. I have been honored to work in quite obviously the best lab at William and Mary.

This research was financially supported by several organizations. The Howard Hughes Medical Institute provided us with grants every semester through the Undergraduate Biological Sciences Education Program at the College of William and Mary. Funds were also received from several Minor Research Grants through the College of William and Mary’s Roy E. Charles Center.

VI. References


VII. Appendix

<table>
<thead>
<tr>
<th>Date Received</th>
<th>Site Name</th>
<th>Sample ID</th>
<th>Site ID</th>
<th>Element</th>
<th>Date</th>
<th>Sequence?</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/1/05</td>
<td>Martin's Hundred</td>
<td>A1</td>
<td>10063</td>
<td>phalange 1</td>
<td>1623-45</td>
<td>no</td>
</tr>
<tr>
<td>9/1/05</td>
<td>Martin's Hundred</td>
<td>A2</td>
<td>10064</td>
<td>phalange 1</td>
<td>1623-45</td>
<td>yes</td>
</tr>
<tr>
<td>9/1/05</td>
<td>Martin's Hundred</td>
<td>A3</td>
<td>10065</td>
<td>phalange 2</td>
<td>1623-45</td>
<td>no</td>
</tr>
<tr>
<td>9/1/05</td>
<td>Martin's Hundred</td>
<td>A4</td>
<td>10066</td>
<td>phalange 1</td>
<td>1623-45</td>
<td>no</td>
</tr>
<tr>
<td>9/1/05</td>
<td>Martin's Hundred</td>
<td>A5</td>
<td>10067</td>
<td>phalange 2</td>
<td>1623-45</td>
<td>no</td>
</tr>
<tr>
<td>9/1/05</td>
<td>Martin's Hundred</td>
<td>A6</td>
<td>10068</td>
<td>phalange 1</td>
<td>1623-45</td>
<td>no</td>
</tr>
<tr>
<td>3/1/06</td>
<td>Nansemond Fort</td>
<td>B1</td>
<td></td>
<td>4th tarsal</td>
<td>ca.1650</td>
<td>yes</td>
</tr>
<tr>
<td>3/1/06</td>
<td>Nansemond Fort</td>
<td>B2</td>
<td></td>
<td>metatarsal</td>
<td>ca.1650</td>
<td>no</td>
</tr>
<tr>
<td>3/1/06</td>
<td>Nansemond Fort</td>
<td>B3</td>
<td>66-44SK192/3A</td>
<td>metacarpal</td>
<td>ca.1650</td>
<td>yes</td>
</tr>
<tr>
<td>3/1/06</td>
<td>Nansemond Fort</td>
<td>B4</td>
<td></td>
<td>metacarpal</td>
<td>ca.1650</td>
<td>no</td>
</tr>
<tr>
<td>3/29/07</td>
<td>Coffee House</td>
<td>CH</td>
<td>17KD277</td>
<td>calcaneus</td>
<td>1755-67</td>
<td>no</td>
</tr>
<tr>
<td>3/29/07</td>
<td>Wren Yard</td>
<td>T2</td>
<td>16GA232</td>
<td>tooth</td>
<td>no date</td>
<td>yes</td>
</tr>
<tr>
<td>3/29/07</td>
<td>Wren Yard</td>
<td>WY1</td>
<td>16GA96</td>
<td>femur</td>
<td>no date</td>
<td>no</td>
</tr>
<tr>
<td>3/29/07</td>
<td>Peyton Randolph</td>
<td>PR1</td>
<td>28H109</td>
<td>calcaneus</td>
<td>no date</td>
<td>no</td>
</tr>
<tr>
<td>3/29/07</td>
<td>Peyton Randolph</td>
<td>PR2</td>
<td>28H90</td>
<td>calcaneus</td>
<td>no date</td>
<td>no</td>
</tr>
<tr>
<td>3/17/08</td>
<td>Jamestown Well</td>
<td>J1</td>
<td>JR1101F</td>
<td>femur</td>
<td>ca.1624</td>
<td>no</td>
</tr>
<tr>
<td>3/17/08</td>
<td>Jamestown Well</td>
<td>J2</td>
<td>JR2158U</td>
<td>tibia</td>
<td>1608-17</td>
<td>no</td>
</tr>
<tr>
<td>3/17/08</td>
<td>Jamestown Well</td>
<td>J3</td>
<td>JR2158U</td>
<td>sacrum</td>
<td>1608-17</td>
<td>no</td>
</tr>
<tr>
<td>3/17/08</td>
<td>Jamestown Well</td>
<td>J4</td>
<td>JR2158P</td>
<td>scapula</td>
<td>1608-17</td>
<td>no</td>
</tr>
<tr>
<td>9/22/08</td>
<td>Nansemond Fort</td>
<td>NT1</td>
<td>44SK192/4/1B UB90</td>
<td>tooth</td>
<td>1636-1646</td>
<td></td>
</tr>
<tr>
<td>9/22/08</td>
<td>Nansemond Fort</td>
<td>NT2</td>
<td>44SK192/3A UB51</td>
<td>tooth</td>
<td>ca.1650</td>
<td></td>
</tr>
<tr>
<td>9/22/08</td>
<td>Nansemond Fort</td>
<td>NT3</td>
<td>44SK192/5 UB100</td>
<td>tooth</td>
<td>1636-1646</td>
<td></td>
</tr>
<tr>
<td>9/22/08</td>
<td>Nansemond Fort</td>
<td>NT4</td>
<td>44SK192/3 UB89</td>
<td>tooth</td>
<td>ca.1650</td>
<td></td>
</tr>
<tr>
<td>2/23/09</td>
<td>Nansemond Fort</td>
<td>NT5</td>
<td>44SK192/3A UB48</td>
<td>tooth</td>
<td>ca.1650</td>
<td></td>
</tr>
<tr>
<td>2/23/09</td>
<td>Nansemond Fort</td>
<td>NT6</td>
<td>44SK192/3A UB95</td>
<td>tooth</td>
<td>ca.1650</td>
<td></td>
</tr>
<tr>
<td>3/30/09</td>
<td>Hampton Univ.</td>
<td>HS</td>
<td>0142-44HT-216</td>
<td>skull</td>
<td>1630-1660</td>
<td></td>
</tr>
<tr>
<td>3/31/09</td>
<td>Hampton Univ.</td>
<td>HT1</td>
<td>44HT55-182 UB2923</td>
<td>up. premolar</td>
<td>1630-1660</td>
<td></td>
</tr>
<tr>
<td>3/31/09</td>
<td>Hampton Univ.</td>
<td>HT2</td>
<td>44HT55-182 UB2451</td>
<td>up. premolar</td>
<td>1630-1660</td>
<td></td>
</tr>
<tr>
<td>3/31/09</td>
<td>Hampton Univ.</td>
<td>HT3</td>
<td>44HT55-182 UB2525</td>
<td>up. premolar</td>
<td>1630-1660</td>
<td></td>
</tr>
<tr>
<td>3/31/09</td>
<td>Hampton Univ.</td>
<td>HT4</td>
<td>44HT55-182 UB2483</td>
<td>molar</td>
<td>1630-1660</td>
<td></td>
</tr>
</tbody>
</table>

Log of Ancient Samples Analyzed

DNA Concentrations After Extraction

Initials are given for the team member who performed the extraction.
MCM= Maura McAuliffe
SCL= Sarah Lehman
<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2 a1</td>
<td>37.6</td>
<td>1.42</td>
</tr>
<tr>
<td>A2 a2</td>
<td>-13.6</td>
<td>1.43</td>
</tr>
<tr>
<td>A2 b1</td>
<td>-33.4</td>
<td>1.43</td>
</tr>
<tr>
<td>A2 b2</td>
<td>32.3</td>
<td>1.64</td>
</tr>
<tr>
<td>A2 c1</td>
<td>-0.6</td>
<td>0.09</td>
</tr>
<tr>
<td>A2 c2</td>
<td>-3.2</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Sample ID | Concentration (ng/µL) | 260/280 |
-----------|----------------------|---------|
A2 a1      | 16.9                 | 1.77    |
A2 a2      | 47.0                 | Data not available |
A2 b1      | 7.0                  | Data not available |
A2 b2      | 11.2                 | 2.47    |
A2 c1      | -13.6                | Data not available |
A2 c2      | 4.7                  | 1.84    |

Sample ID | Concentration (ng/µL) | 260/280 |
-----------|----------------------|---------|
A3 a1      | 22.0                 | 1.81    |
A3 a2      | 0.9                  | -0.4    |
A3 b1      | 16.7                 | 1.85    |
A3 b2      | 7.0                  | 3.43    |
A3 c1      | -10.0                | 1.5     |
A3 c2      | -22.0                | 1.11    |

Sample ID | Concentration (ng/µL) | 260/280 |
-----------|----------------------|---------|
A5 a1      | 47.6                 | 2.62    |
A5 a2      | 7.5                  | 4.12    |
A5 b1      | 29.5                 | 2.67    |
A5 b2      | 17.3                 | 2.99    |
A5 c1      | -6.0                 | 1.56    |
A5 c2      | -9.3                 | 1.12    |

Sample ID | Concentration (ng/µL) | 260/280 |
-----------|----------------------|---------|
A6 a1      | 105.6                | 1.59    |
A6 a2      | 118.9                | 1.48    |
A6 b1      | 125.8                | 1.56    |
A6 b2      | 121.5                | 1.46    |
A6 c1      | 10.8                 | 1.89    |
A6 c2      | 8.1                  | 1.53    |

Sample ID | Concentration (ng/µL) | 260/280 |
-----------|----------------------|---------|
B1 a1      |                      |         |
B1 a2      |                      |         |
B1 b1      |                      |         |
B1 b2      |                      |         |
B1 c1      |                      |         |
B1 c2      |                      |         |
<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280*</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1 a1</td>
<td>135.9</td>
<td>Data not available</td>
</tr>
<tr>
<td>B1 a2</td>
<td>143.2</td>
<td>Data not available</td>
</tr>
<tr>
<td>B1 b1</td>
<td>109.4</td>
<td>Data not available</td>
</tr>
<tr>
<td>B1 b2</td>
<td>142.1</td>
<td>Data not available</td>
</tr>
<tr>
<td>B1 c1</td>
<td>16.2</td>
<td>Data not available</td>
</tr>
<tr>
<td>B1 c2</td>
<td>13.1</td>
<td>Data not available</td>
</tr>
</tbody>
</table>

*These values were not recorded.

**B2 9/13/06 MEC**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280*</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2 a1</td>
<td>18.4</td>
<td>Data not available</td>
</tr>
<tr>
<td>B2 a2</td>
<td>6.7</td>
<td>Data not available</td>
</tr>
<tr>
<td>B2 b1</td>
<td>16.4</td>
<td>Data not available</td>
</tr>
<tr>
<td>B2 b2</td>
<td>4.0</td>
<td>Data not available</td>
</tr>
<tr>
<td>B2 c1</td>
<td>4.3</td>
<td>Data not available</td>
</tr>
<tr>
<td>B2 c2</td>
<td>3.1</td>
<td>Data not available</td>
</tr>
</tbody>
</table>

*These values were not recorded.

**B3 10/11/06 SCL & MEC**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280*</th>
</tr>
</thead>
<tbody>
<tr>
<td>B3 a1</td>
<td>44.7</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 a2</td>
<td>15.8</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 b1</td>
<td>115.6</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 b2</td>
<td>76.1</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 c1</td>
<td>32.9</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 c2</td>
<td>30.5</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 d1^</td>
<td>111.5</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 d2^</td>
<td>77.6</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 e1^</td>
<td>69.6</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 e2^</td>
<td>21.1</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 f1^</td>
<td>24.5</td>
<td>Data not available</td>
</tr>
<tr>
<td>B3 f2^</td>
<td>21.1</td>
<td>Data not available</td>
</tr>
</tbody>
</table>

*These values were not recorded.

^This sample was extracted twice for replication purposes.

**B4 2/8/07 SCL & MEC**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230*</th>
</tr>
</thead>
<tbody>
<tr>
<td>B4 a1</td>
<td>142.4</td>
<td>1.61</td>
<td>0.19</td>
</tr>
<tr>
<td>B4 a2</td>
<td>105.2</td>
<td>1.55</td>
<td>0.12</td>
</tr>
<tr>
<td>Sample ID</td>
<td>Concentration (ng/µL)</td>
<td>260/280</td>
<td>260/230</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>B4 b1</td>
<td>323.0</td>
<td>1.57</td>
<td>0.94</td>
</tr>
<tr>
<td>B4 b2</td>
<td>213.2</td>
<td>1.54</td>
<td>0.28</td>
</tr>
<tr>
<td>B4 c1</td>
<td>30.1</td>
<td>1.71</td>
<td>0.05</td>
</tr>
<tr>
<td>B4 c2</td>
<td>29.3</td>
<td>1.59</td>
<td>0.03</td>
</tr>
</tbody>
</table>

*These values were recorded for the first time here and always recorded hereafter.

**CH 4/21/07 MEC**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH a1</td>
<td>47.8</td>
<td>1.83</td>
<td>0.07</td>
</tr>
<tr>
<td>CH a2</td>
<td>48.3</td>
<td>1.74</td>
<td>0.07</td>
</tr>
<tr>
<td>CH b1</td>
<td>48.8</td>
<td>1.86</td>
<td>0.22</td>
</tr>
<tr>
<td>CH b2</td>
<td>67.9</td>
<td>1.77</td>
<td>0.10</td>
</tr>
<tr>
<td>CH c1</td>
<td>27.9</td>
<td>1.89</td>
<td>0.05</td>
</tr>
<tr>
<td>CH c2</td>
<td>34.6</td>
<td>1.56</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**T1 4/26/07 SCL**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 a1</td>
<td>80.0</td>
<td>1.87</td>
<td>0.37</td>
</tr>
<tr>
<td>T1 a2</td>
<td>51.1</td>
<td>1.85</td>
<td>0.09</td>
</tr>
<tr>
<td>T1 b1</td>
<td>64.8</td>
<td>1.77</td>
<td>0.10</td>
</tr>
<tr>
<td>T1 b2</td>
<td>44.6</td>
<td>1.62</td>
<td>0.05</td>
</tr>
<tr>
<td>T1 c1</td>
<td>22.1</td>
<td>1.53</td>
<td>0.03</td>
</tr>
<tr>
<td>T1 c2</td>
<td>18.7</td>
<td>1.46</td>
<td>0.06</td>
</tr>
</tbody>
</table>

**WY1 9/11/07 MEC**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>WY1 a1</td>
<td>333.9</td>
<td>1.59</td>
<td>1.26</td>
</tr>
<tr>
<td>WY1 a2</td>
<td>248.5</td>
<td>1.54</td>
<td>0.33</td>
</tr>
<tr>
<td>WY1 b1</td>
<td>345.4</td>
<td>1.58</td>
<td>0.95</td>
</tr>
<tr>
<td>WY1 b2</td>
<td>247.0</td>
<td>1.54</td>
<td>0.30</td>
</tr>
<tr>
<td>WY1 c1</td>
<td>28.8</td>
<td>1.47</td>
<td>0.03</td>
</tr>
<tr>
<td>WY1 c2</td>
<td>29.0</td>
<td>1.47</td>
<td>0.06</td>
</tr>
</tbody>
</table>

**T2 10/3/07 SCL**
<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2 a1</td>
<td>78.4</td>
<td>1.54</td>
<td>0.13</td>
</tr>
<tr>
<td>T2 a2</td>
<td>45.1</td>
<td>1.45</td>
<td>0.19</td>
</tr>
<tr>
<td>T2 b1</td>
<td>58.8</td>
<td>1.53</td>
<td>0.14</td>
</tr>
<tr>
<td>T2 b2</td>
<td>28.1</td>
<td>1.40</td>
<td>0.24</td>
</tr>
<tr>
<td>T2 c1</td>
<td>14.8</td>
<td>1.44</td>
<td>0.03</td>
</tr>
<tr>
<td>T2 c2</td>
<td>10.1</td>
<td>1.31</td>
<td>0.09</td>
</tr>
</tbody>
</table>

**WY2  10/9/07    MEC**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>WY2 a1</td>
<td>89.7</td>
<td>1.55</td>
<td>0.11</td>
</tr>
<tr>
<td>WY2 a2</td>
<td>63.6</td>
<td>1.42</td>
<td>0.14</td>
</tr>
<tr>
<td>WY2 b1</td>
<td>101.4</td>
<td>1.55</td>
<td>0.12</td>
</tr>
<tr>
<td>WY2 b2</td>
<td>68.8</td>
<td>1.45</td>
<td>0.12</td>
</tr>
<tr>
<td>WY2 c1</td>
<td>17.3</td>
<td>1.59</td>
<td>0.02</td>
</tr>
<tr>
<td>WY2 c2</td>
<td>14.8</td>
<td>1.46</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**PR1  10/29/07    MEC**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR1 a1</td>
<td>115.1</td>
<td>1.59</td>
<td>0.16</td>
</tr>
<tr>
<td>PR1 a2</td>
<td>57.0</td>
<td>1.51</td>
<td>0.08</td>
</tr>
<tr>
<td>PR1 b1</td>
<td>111.7</td>
<td>1.58</td>
<td>0.15</td>
</tr>
<tr>
<td>PR1 b2</td>
<td>93.9</td>
<td>1.52</td>
<td>0.11</td>
</tr>
<tr>
<td>PR1 c1</td>
<td>1.3</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>PR1 c2</td>
<td>20.8</td>
<td>1.60</td>
<td>0.09</td>
</tr>
</tbody>
</table>

--The Nanodrop did not give ratios for c1, probably because the DNA concentration was so low.

**PR2  10/31/07    SCL**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR2 a1</td>
<td>21.6</td>
<td>1.41</td>
<td>0.22</td>
</tr>
<tr>
<td>PR2 a2</td>
<td>9.4</td>
<td>1.33</td>
<td>0.28</td>
</tr>
<tr>
<td>PR2 b1</td>
<td>23.2</td>
<td>1.44</td>
<td>0.16</td>
</tr>
<tr>
<td>PR2 b2</td>
<td>12.9</td>
<td>1.39</td>
<td>0.21</td>
</tr>
<tr>
<td>PR2 c1</td>
<td>5.7</td>
<td>1.21</td>
<td>0.05</td>
</tr>
<tr>
<td>PR2 c2</td>
<td>3.0</td>
<td>1.03</td>
<td>0.07</td>
</tr>
</tbody>
</table>
### J1 3/27/08 MEC

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>J1 a1</td>
<td>28.6</td>
<td>1.63</td>
<td>0.04</td>
</tr>
<tr>
<td>J1 a2</td>
<td>25.3</td>
<td>1.52</td>
<td>0.05</td>
</tr>
<tr>
<td>J1 b1</td>
<td>26.6</td>
<td>1.66</td>
<td>0.04</td>
</tr>
<tr>
<td>J1 b2</td>
<td>22.0</td>
<td>1.55</td>
<td>0.05</td>
</tr>
<tr>
<td>J1 c1</td>
<td>13.1</td>
<td>1.59</td>
<td>0.03</td>
</tr>
<tr>
<td>J1 c2</td>
<td>12.4</td>
<td>1.42</td>
<td>0.08</td>
</tr>
<tr>
<td>J1 d1^</td>
<td>23.1</td>
<td>1.57</td>
<td>0.04</td>
</tr>
<tr>
<td>J1 d2^</td>
<td>19.2</td>
<td>1.53</td>
<td>0.09</td>
</tr>
<tr>
<td>J1 e1^</td>
<td>20.1</td>
<td>1.54</td>
<td>0.07</td>
</tr>
<tr>
<td>J1 e2^</td>
<td>11.8</td>
<td>1.40</td>
<td>0.11</td>
</tr>
<tr>
<td>J1 f1^</td>
<td>8.9</td>
<td>1.50</td>
<td>0.04</td>
</tr>
<tr>
<td>J1 f2^</td>
<td>5.7</td>
<td>1.49</td>
<td>0.06</td>
</tr>
</tbody>
</table>

^This sample was extracted twice for replication purposes.

### J4 5/1/08 SCL

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>J4 a1</td>
<td>170.4</td>
<td>1.49</td>
<td>0.29</td>
</tr>
<tr>
<td>J4 a2</td>
<td>149.2</td>
<td>1.41</td>
<td>0.52</td>
</tr>
<tr>
<td>J4 b1</td>
<td>162.1</td>
<td>1.51</td>
<td>0.30</td>
</tr>
<tr>
<td>J4 b2</td>
<td>132.0</td>
<td>1.44</td>
<td>0.43</td>
</tr>
<tr>
<td>J4 c1</td>
<td>13.2</td>
<td>1.43</td>
<td>0.05</td>
</tr>
<tr>
<td>J4 c2</td>
<td>10.5</td>
<td>1.34</td>
<td>0.11</td>
</tr>
</tbody>
</table>

### J2 5/9/08 MEC

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>J2 a1</td>
<td>238.6</td>
<td>1.47</td>
<td>0.28</td>
</tr>
<tr>
<td>J2 a2</td>
<td>354.6</td>
<td>1.40</td>
<td>0.49</td>
</tr>
<tr>
<td>J2 b1</td>
<td>280.7</td>
<td>1.47</td>
<td>0.33</td>
</tr>
<tr>
<td>J2 b2</td>
<td>326.3</td>
<td>1.41</td>
<td>0.49</td>
</tr>
<tr>
<td>J2 c1</td>
<td>18.1</td>
<td>1.54</td>
<td>0.03</td>
</tr>
<tr>
<td>J2 c2</td>
<td>17.2</td>
<td>1.55</td>
<td>0.05</td>
</tr>
</tbody>
</table>
J3  5/9/08  MEC

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>J3 a1</td>
<td>300.3</td>
<td>1.58</td>
<td>0.57</td>
</tr>
<tr>
<td>J3 a2</td>
<td>295.8</td>
<td>1.44</td>
<td>0.39</td>
</tr>
<tr>
<td>J3 b1</td>
<td>295.2</td>
<td>1.58</td>
<td>0.61</td>
</tr>
<tr>
<td>J3 b2</td>
<td>348.8</td>
<td>1.43</td>
<td>0.49</td>
</tr>
<tr>
<td>J3 c1</td>
<td>27.3</td>
<td>1.78</td>
<td>0.04</td>
</tr>
<tr>
<td>J3 c2</td>
<td>30.0</td>
<td>1.57</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Devon Cow Samples  11/3/08  MEC

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Concentration (ng/µL)</th>
<th>260/280</th>
<th>260/230</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marietta</td>
<td>109.6</td>
<td>1.65</td>
<td>0.30</td>
</tr>
<tr>
<td>Zenia</td>
<td>58.3</td>
<td>1.59</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Chromatograms and Sequences of Successful Samples

*m1 (modern, positive control)—sequenced when T2 was sequenced 199 bp

TATGTATATAGTACATTTAATTATATGCCCATGCATATAAGCAAGTACATGA
CCTCTATAGCAGTACATAATACATAATTATTGACTGTACATAGTACATTAT
GTCAAAATTCTTCTTGGATAGTATATCTATTATATTTCCTTACCATTAGATCAC
GAGCTTAATTACCATGCGCGTAAAAAAAAAAGCGACAGG
*Chromatogram not available for this sample

A2 a1 (F)—167 bp
Entire A2 a1 (F) Sequence

TAAATTATATGCCCCATGCATATAAGCAAGTACATGACCTCTATAGCAGTACATAATACATATAATTATTGACTGTACATAGTACATTATGTCAAATTCTTTGATAGTATATCTATTATATATTTCTTACCATTAGATCACGAGCTTAATTACCATGCCG

B1 b1 (F)—152 bp
Entire B1 b1 (F) Sequence

TGCATATAAGCAAGTACATGACCTCTCTATAGCAGTAACATAATACATATAATTA
TTGACTGTACATAGTACATTATGTCAAATTCATTCTTGATAGTATATCTATTAT
ATATTCCTTACCATTAGATCAGAGCTTAATTACCATGCGCGCTA

B3 d1 (F)—129 bp
Entire B3 d1 (F) Sequence

CCTCTATAGCAGTACATAATACATATAATTATTGACTGTACATAGTACATTAT
GTCAAATTCATTCTTGATAGTATATCTATTATATATTCCTTACCATTAGATCAC
GAGCTTAATTACCATGCCCGGT

T1 a1 (F)—150 bp
Entire T1 a1 (F) Sequence

GCATATAAGCAGTACATGACCTCTATAGCAGTACATAATACATATAATTATTG
ACTGTACATAGTACATTATGTCAATTCATTCTTGATAGTATATCTATTATATAT
TCCTTACCATTAGATCAGGAGCTTAATTACCAATGCGCGTAA

*T2 a1 (F)—203 bp
ATGTATATAGTACATTTAATTATATGCCCAARGCATATAAGCAAGTACATGAC
CTCTATAGCAGTACATAAATACATATATTATTTGACCTTAGATCATAGTACATTATG
TCAAATTCTTCTTGTAGTATATCTATTATATATATTCTTCTTACCATTAGATCGAG
AGCTTAATTACCATGCGCGTCTGAGATGTCTTTATTAAAGAGGT

*T2 b1 (F)—157 bp
CAGACTGCATATAAAGCAAGTACATGACCTCTATAGCAGTACATAATACATAT
AATTTGGACTGTACATAGTACATTATGTCAAATTCTTTGTAGATGTATATCT
ATTATATATCCCTTACCATTAGATCGAGCTTAAATTACCATGCGCGTAA

*Chromatograms were not available for T2 samples.

Devon Cheek Cell Samples

Marietta (1)—127 bp

Entire Marietta (1) Sequence
TATCAGTACATAATACATATAATTATTGATTGTACATAGTACATTATGTCAAAT
TCATTCTTGATAGTATATCTATTATATATTCTTCTTACCATTAGATCGAGCTTTA
ATTACCATGCGCGTAAAT

Marietta (2)—123 bp
Entire Marietta (2) Sequence

TACATATAATTATTGATTGTACATAGTACATTATGTCAAAATTCATTCTTGATAG
TATATCTATTATATATTCTCCTTTACCATTAGATCACGAGCTTAATTACCATGCCG
CGTAGCCCGGTCCCCG

Zenia (1)—128 bp
*Gray areas are trimmed regions and are not included in the analyzed sequence.

Entire Zenia (1) Sequence

TACATAATACATATAATTATTGACTGTACATAGTACATTATGTCAAATTTCATTC
TTGATAGTATATCTATTATATATTCCT
TACCATTAGATCACGAGCTTAATTACC
ATGCCGCGTAGAGCACCACCGAG

Zenia (2)—130 bp
*Gray areas are trimmed regions and are not included in the analyzed sequence.

**Entire Zenia (2) Sequence**

GTACATAATACATAATTATTGTGACTGTACATAGTACATTATGTCAATTTCTT
CTTGATAGTATATCTATTATATATTCTTACCATTAGATCAGGAGCTTAATTAC
CATGCCGCGTAGAGTGAGCTCC

**Sample Alignments with the 218 bp Comparison Sequence**

Red indicates where the two sequences align
Black indicates where the two sequences do not align

**m1 (modern, positive control)—sequenced when T2 was sequenced 199 bp**

```
<table>
<thead>
<tr>
<th></th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1 F (from T2) 218 bp frag Consensus</td>
<td>gggagcgt</td>
<td>tataaacta</td>
<td>atgaatta</td>
<td>gacataat</td>
<td>tataagta</td>
<td>catatta</td>
</tr>
<tr>
<td></td>
<td>140</td>
<td>160</td>
<td>180</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m1 F (from T2) 218 bp frag Consensus</td>
<td>attgacgt</td>
<td>cattacat</td>
<td>tagtcaat</td>
<td>tcctcttg</td>
<td>tagtatact</td>
<td>attatatt</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m1 F (from T2) 218 bp frag Consensus</td>
<td>ccctacatt</td>
<td>agatccagag</td>
<td>cttaattacc</td>
<td>atgccgcgt</td>
<td>aaaaaaaaaagc</td>
<td>gacagg</td>
</tr>
</tbody>
</table>
```

**B1 b1 (F)—152 bp**
**T1 a1 (F)—150 bp**

**T2 b1 (F)—157 bp**
T2 a1 (F) & T2 b1 (F) Alignment

Devon Cheek Cell Samples
Marietta 1 & 2 Alignment

Zenia (2)—130 bp

Zenia 1 & 2 Alignment
Published Sequences of Various *Bos taurus* Breeds and Alignments with the 218 bp fragment

**Aberdeen Angus 2—240 bp**

**Entire Aberdeen Angus 2 Sequence**

Acc.# AF336472- (Troy et al., 2001)

```plaintext
CCCCCATGCATATAAGCAAGTACATGAC
CTCTATAGCAGTACATAATACATAT
AATTATTGACTGTACATAGTACATTATGTCAAATTCATTCTTGATAGTATATCT
ATTATATATTCCTTACCATTAGATCACGAGCTTAATTACCATGCCGCGT
GAAACCAGCAACCCGCTAGGCAAGGATCCCCCTTCTCGCTCCGGGCCCAT
AAACCGTGGGGGTCGCTATTCAATGAATTTTACCA
```

**Jersey 1—436 bp**

```plaintext
```

74
Entire Jersey 1 Sequence
Acc.# AB079358- (Komatsu et al., 2004)

GTAATGTACATAACATAATGTAATAAAGACATAATATGTATATAGTACATTTAATTATGCCCACATGATATAAGCAAGTACATGACCTCTATAGCAGTACATAATACATATAATTATTGACTGTACATAGTACATTATGTCAAATTCATTCTTGATAGTATATCTATTACATATTCCTTACCATTAGATCAGACCTTAAATTACCATGCCCGTGAAACCAGCAACCCGCAGAGGGATCCCTCTTCTCGCTCCGGGCCATAAACCGTGGGGGTCGCTATCCAATGATTTTACCAGGCATCTGGTTCTTTTCAGGGCCATCTCATCTAAAACGGTCCATTCTTTCCTCTTAAATAAGACATCTCGATGGACTAATGGCTAATCAGCCCATGCTCACACATAACTGTGCTGTCATACATTTGGTATTT

Jersey 2—240 bp
Entire Jersey 2 Sequence
Acc.# AF336488- (Troy et al., 2001)

CCCATGCATATAAGCAAGTACATGACCTCTATAGTAGTACATAATACATATAATTATTGACTGTACATAGTACATTATGTCAAATTCATTCTTGATAGTATATCTATTATATCATTCCTTACCATTAGATCACGAGCTTAATTACCATGCCGCGTGAAA

Friesian 1—398 bp
Entire Friesian 1 Sequence
Acc.# U51818- (Bradley et al., 1996)

TAGTACATTTAATTATATGCCCCATGCATATAAGCAAGTACATGACCTCTATA
GCAGTACATAATACATAACAATTATTTGACTGTACATAGTACATTATGTCAAAATT
CATCTTTGATAGTATATCTATTATATATTCCTTACCATTAGATCACGAGCTTAA
TTACCATGCCGCGTGAAACCAGCAACCCGCTAGGCAGGGATCCCTCTCTC
GCTCCGGGCCCATAAACCGTGGGGGTCGCTATCCAGTGAATTTTACCAGGC
ATCTGGTTCTTTCTTCAGGGCCATCTCATCTAAAACGGTCCATTCTTTCCTCT
TAAATAAGACATCTCAGTGACCTATTGCTAATCAGCC

ACTCGTGCTGTACATACATTGTTATTTTTA
**Friesian 2—16341 bp**

Entire Friesian 2 sequence not shown because of length.

Acc.# EU177826- (Achilli et al., 2008)
Kerry 1—240 bp

Entire Kerry 1 Sequence
Acc.# AF336396- (Troy et al., 2001)

CCCCCATGCATATAAGCAAGTACATGACCTCTATAGCAGTACATAATACATATAATTATTGACTGTACATAGTACATTATGTCAAATTCATTCTTGATAGTATATCTATTATATTTTCCTTACCATTAGATCACGAGCTTAATTACCATGCCGCGTGAAAACAGCAACCCGCTAGGCAGGGATCCCTCTTCTCGCTCCGGGCCCATAAACCGTGGGGGTCGCTATCC

AATGAATTTTACCA

CCCCCATGCATATAAGCAAGTACATGACCTCTATAGCAGTACATAATACATATAATTATTGACTGTACATAGTACATTATGTCAAATTCATTCTTGATAGTATATCTATTATATTTTCCTTACCATTAGATCACGAGCTTAATTACCATGCCGCGTGAAAACAGCAACCCGCTAGGCAGGGATCCCTCTTCTCGCTCCGGGCCCATAAACCGTGGGGGTCGCTATCC

AATGAATTTTACCA
Kerry 2—240 bp

Entire Kerry 2 Sequence
Acc.# AF336404- (Troy et al., 2001)

CCCCATGCATATAAGCAAGCACATGACCTCTATAGCAGTACATAATACATAT
AATTATTGACTGTACATAGTACATTATGTCAAATTCATTCTTGATAGTATATCT
ATTATATAATTCCATTACATTAGATCAGAGCTTAATTACCAGCCGCGTGAAC
CCAGCAACCCGCTAGGCAGGGATCCCTCTTCTCGCTCCGGGCCCATAAAC
CGTGCCGCCTACCTAATGAATTTTACCA
### Entire Colombian Creole 1 Sequence

**Acc.# AY444492- (Carvajal-Carmona et al., 2003)**

```
CATAACATTATTTGTAATAGCAATAATATGTATATAGTTACATTAAATTATATG
CCCCATGCATATAAGCAAGTACATGACCTCTATAGCAGTACATAATACATAC
AATTATTGACTGTACATATGGAGTACATTATGTCAAATTCATTCTTGATAGTATATCT
ATTATATATTCCCTTGGCCATTAGATCAGCGCTTAATGCTATAGAGATGTATATCT
ATTATATATTCCCTTGGCCATTAGATCAGCGCTTAATGCTATAGAGATGTATATCT
```

*The position at which the first variation occurs (position 21) is not a variable position that is included in MacHugh et al. (1999).*
Colombian Creole 2 (Harton del Valle)—356 bp

Entire Colombian Creole 2 Sequence
Acc.# AY444429 - (Carvajal-Carmona et al., 2003)

CATACATATTGTAATAAAGACATAAATATGTAATATAGACATATTAAATTATAG
CCCCATGCTATATAAGCAAGTACATGACTTCTATAGCAGTGACATAAATACATATA
ATTATTTGACTGTACATAGTACATTTGCTATGTACAAAATCCTTTGACTATATCTA
TTATATAATTCACTTAGATACAGAGCTTTAATTTACCATAGCCCGCTGGAAAC
CAGCAACCCCGCTTGGCAGGAATCCTCTTCTCTCTCTCCTTCTCTCTCTCTCTCCT
GCCATCTCATCTAAACGCTTTCCCTTCCCTTCTTCTTAAATA
Bolivian Creole—240 bp

Entire Bolivian Creole Sequence
Acc.# EU131170- Submitted to NCBI’s GenBank only (unpublished)

CCCATGCATATAAGCAAGTACATGACCTCTATGGCAGTACATAATACATAT
AATTATTGACTGTACATAGTACATTATGTCAAATTCATTCTTGATAGTATATCT
ATTATATATTTCCCTTACCATTAGATCAGGAGCTTAATTACCATGCCCCGGTGA
CCAGCAACCCGGTAGGCAGGGATCCCTCTTTCTGCTCCGGGCCCATAAAC
CGTGGGGGTGCTGTATCCAATGAATTTTACCA
Cuban Creole—240 bp

Entire Cuban Creole Sequence
Acc.# FJ611981- Submitted to NCBI’s GenBank only (unpublished)

CCCATGCATATAAGCAAGTACATGACCTCTATAGCAGTACATAATAACATAT
AATTATTGACTGAGATCGATAGTACATTATGCTCAAAATTCAATTCTTGTAGATGATATC
ATTATATATTCTTACCATTAGATCAGAGCTTAATTACCATGCCTCGCGTGAAA
CCAGCAACCCCGCTAGGCAGGGATCCCTCTCGCTCCGGGCACAAAC
CGTGCGGGTGTCATTCAATGAAATTTTACCA