

Danielle Johnsen

Intro to Scientific Computing

10 February 2023

The Dogs of *Tsenacomoco*: Jamestown Phylogeny Reconstruction

The goal of this project is to reconstruct Figure 2 from the unpublished manuscript of the relationship between ancient DNA and indigenous dogs at the Jamestown Colony. This phylogenetic tree illustrates the situation of Jamestown dogs in the 17th century among the indigenous dog population; a phylogenetic tree of the dogs indigenous to the Americas and those found in the Jamestown Colony.

by Ariane E. Thomas, Matthew E. Hill, Jr. and Andrew Kitchen Department of Anthropology,
University of Iowa, 114 Macbride Hall, Iowa City, IA 52242

Introduction

Due to the colonization of the Americas during and after Columbus' infiltration, like many things, there followed a drastic change in the preexisting environment. Specifically, this research interrogates the replacement of indigenous North American dogs with European dogs genetically between this initial contact and modern day. Because of a messy history with the specific utility of dogs in relation to human civilization, it is difficult to determine where the lineages overlap from just archeological evidence or historical record. Specifically, it is in question whether the dogs owned by the colonists at Jamestown were wholly European, Indigenous, or mixed. In order to gather data, Thomas and others collected and sequenced ancient mitochondrial DNA from 3 different archaeological dogs from between 1609 and 1617 AD, and found that the dogs from the Jamestown Colony had maternal ties to North American dogs. This creates a complexity of inquiry into social history and usages of dogs in colonialist environments. I chose this project as I don't have copious experience with coding or data to reproduce something wholly original without help or error.

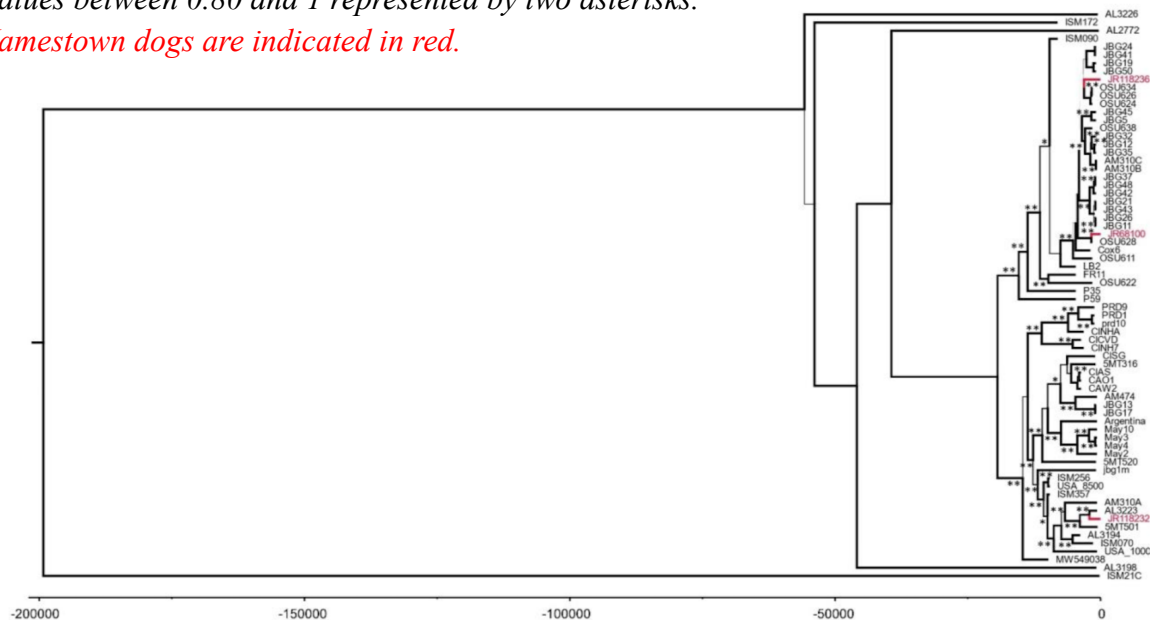
Historical accounts suggest many theoretical frameworks from which to view the interchange between indigenous peoples and European colonizers. The dogs found at Jamestown could represent a variety of different explanations for the existence of dogs at the colony. Should the dogs be taken from or given to the Europeans by Indigenous peoples, the dogs will show linkage

to North American dogs and clustered within their phylogenetic class. Should the settlers have brought their own dogs, they will be linked, at least maternally, to later recorded European dogs.

Hypotheses suggest various accounts of the extent of indigenous and colonist contact. One theory suggests colonial site dogs show relation to their own European lineage of later dogs; that the dogs at the site were brought by the settlers to the Americas. Alternatively, DNA analysis showing linkage to indigenous dogs, at least maternally, implies dogs were either taken or given from indigenous peoples to the colonists. Should mitochondrial DNA show association with European dogs, one could expect a monophyletic tree separating them from the indigenous clade. In other words, they wouldn't be nested with such close linkage. Under the latter set of circumstances, the dogs would show mitochondrial linkage with Indigenous dogs would result in their being tightly clustered in relationship to the reference native dog genomes. The below phylogenetic tree shows the Jamestown Dogs nested within the mitochondrial lineages of pre-European contact North American Dogs. If they were indeed European dogs, they would show up in a monophyletic series; they would not cluster within the North American clade as shown in red.

Original Figure

*Figure 2. Bayesian phylogeny ancient North American canids.
Posterior probabilities greater than 0.60 are marked, with values between 0.6 and 0.79 represented by one asterisk and posterior values between 0.80 and 1 represented by two asterisks.
Jamestown dogs are indicated in red.*



Reproduced Figure

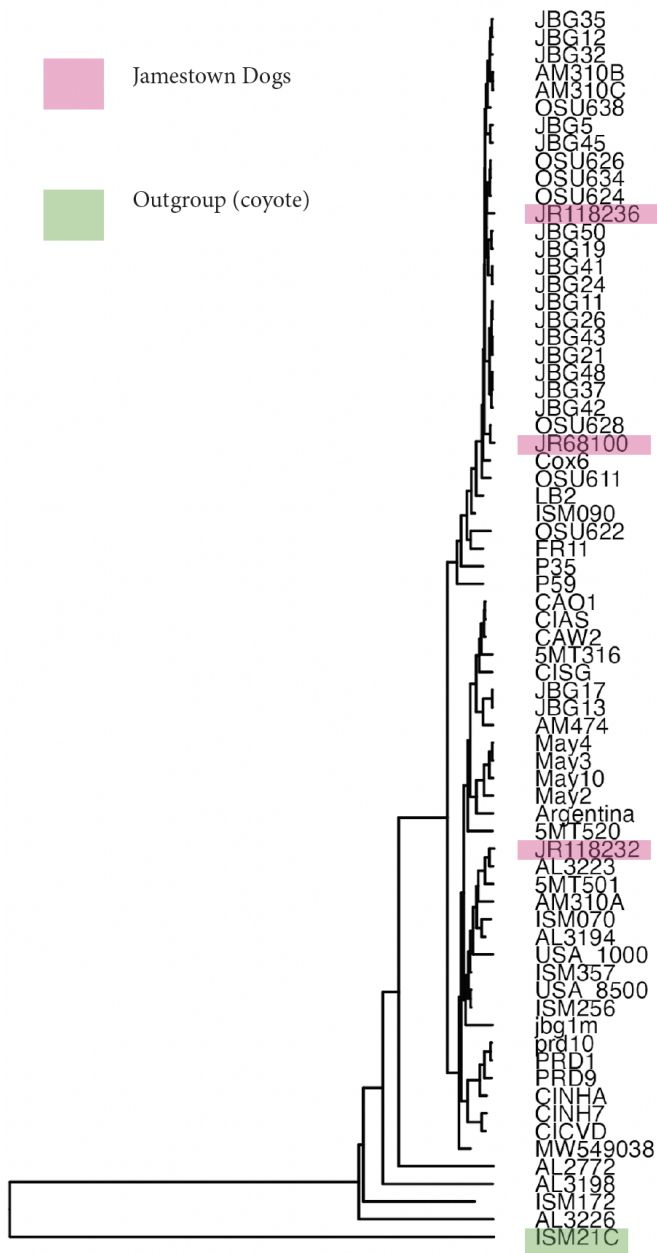


Figure 2(.2) ggplot/ggtree North American canids. This Figure shows the nested configuration of the Jamestown Colony dogs within Indigenous Dog clades via mitochondrial DNA. The outgroup species (coyote) are highlighted in green.

Discussion

The relationship pattern shown in the above tree proves indigenous ancestry, specifically with Hopewellian, Mississippian, and Late Woodland dogs, all native to eastern North America. The dogs found at the Jamestown Colony site (highlighted in pink) are clustered within the phylogenetic clade of pre-contact North American dogs. Certainty regarding the ancestry of these dogs are reasonably limited, because mitochondrial DNA is inherited maternally. The information obtained from nuclear

DNA should be the focus of consequent studies. The two plausible explanations for these findings are either that the Jamestown dogs have full indigenous dog ancestry or that they are hybrids of early English male dogs and Indigenous female dogs. The presence of indigenous dogs within Jamestown highlights pre existing evidence for an indisputable Indigenous presence at Jamestown during the early 17th century AD. The relationship between Jamestown and Indigenous communities varied, but food and supply subsistence from Native groups, (either given to or stolen by the English), were unquestionably part of their survival mechanism.

Materials

DNA extraction from six Canis tooth specimens for DNA analysis shown in Table 1. Of the six specimens, two specimens were found in contexts other than those described above.

Mitochondrial DNA reads for 68100, 118232, and 118236 were aligned to a publicly available dog mitogenome (Kim et al. 1998). The mean coverage for specimen 68100 was 25.381x, 5.318x for specimen 118232, and 66.785x for specimen 118236. The sequences have been uploaded to GenBank*

**accession numbers not provided in manuscript*

Methods (Original Figure)

-Raw reads run through PALEOMIX (ancient DNA pipeline). PALEOMIX pipelines help to rapidly process HTS data (multiple genome alignments, in this case).

-Multiple Sequence Alignment using MUSCLE (140 sequences)

-Maximum Likelihood Tree RaxML (may try RAxML-ng

-Bayesian Phylogenetic Tree using BEAST v1. 10: --Using a fixed mutation rate 1.0×10^{-7} substitutions/site/year --Using a constant demographic model --Using an HKY substitution model with gamma distribution

Methods (Recreated Figure)

-Fasta file aligned in Sequotron

-ggtree created via Newick file in RStudio

Reflection

The robustness of this project was dependent on the ability for me to create a new tree under the same parameters as the original. Some errors in reproducibility was the inability to access the GenBank due to no sequence numbers being provided. Some errors aside from my lack of proficiency in coding thus far, included properly discussing the results without regurgitating what has been produced in the original manuscript. This project allowed me to become rather comfortable with an organized repository, such as GitHub.

Table 1.

Table 1. Canid specimens from Jamestown sequenced in this study.

	Sample	Element	Sampled Location (for DNA)	Structure	Layer	mtDNA Coverage
1	118294	Right mandible	Canine	145	JR1892D	4.14
2	118230	Left maxilla	Maxilla superior to the fourth permanent premolar (unerupted)	185	JR2718J	Failed quality control at PCR stage (was not submitted for sequencing)
3	68100	Left maxilla	Fourth premolar	183	JR2361C	25.38
4	118236	Right maxilla	Fourth premolar and a small amount from right first molar	185	JR2718W	66.79
5	72592	Right mandible	Not sampled (deciduous teeth)	185	JR2718J	Juvenile specimen; no suitable sample available
6	118231	Left maxilla	Fourth premolar	185	JR2718N	1.34
7	118232	Left maxilla	Fourth premolar	185	JR2718N	5.32