Spina Bifida at a Pre-Columbian Cuban Site: A Molecular and Paleoepidemiological Perspective

by

Stephanie D. Armstrong

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF ARTS

Department of Anthropology

University of Manitoba

Winnipeg

Abstract

Health in archaeological populations needs to be investigated using a holistic approach. Molecular techniques, particularly multiplex PCR, can be used with paleopathology and dietary analysis to understand aspects of population health. This thesis demonstrates how spina bifida, a multi-factorial disease, can be investigated using this paleoepidemiological approach.

Based on skeletal evidence, spina bifida was present in a pre-Columbian Cuban population from the archaeological site of Canimar Abajo. Molecular techniques were employed to examine disease potential, examining individuals for five single nucleotide polymorphisms associated with spina bifida. It is postulated that the combined effect of these polymorphisms, as well as dietary factors, determines the risk of the population for spina bifida, and that these factors came together to create the observed high disease prevalence.

Therefore, this thesis demonstrates how the methods of molecular paleopathology, corroborated by dietary analyses, can be used within a paleoepidemiological framework to understand population health and disease.

Acknowledgments

First and foremost, I would like to thank my friend and supervisor, Dr. Mirjana Roksandic for her encouragement and support. I would like to acknowledge my committee, Dr. Rob Hoppa and Dr. Carney Matheson for their invaluable feedback and knowledge.

My research was only possible due to the assistance of many individuals who provided me with access to Canimar Abajo and its collection, as well as the work of past graduate students: Drs. Roberto Rodriguez Suarez, Carlos Arredondo, and Antonio Martinez, Yadira Chinique de Armas, and Elizabeth Campillo Álvarez of the University of Havana. I would also like to recognize the work and collaboration of the other investigators of the project who showed me the rewards of collaborative research: Drs. Bill Buhay, David Smith, Esteban Parra, Matthew Peros, and Ivan Roksandic.

I could not have even begun to pursue bioanthropology without the continued love, support, and motivation from my family and friends: Alistair, Mom, Dad, and Mike, Sara, and all of my UW and UM people, who are too many and too amazing to name.

I would like to acknowledge the generous funding provided by: the National Sciences and Engineering Research Council, the government of Manitoba, and the University of Manitoba (Faculty of Graduate Studies, Faculty of Arts, Department of Anthropology, Graduate Students' Association). Research costs were funded by the Social Sciences and Humanities Research Council Standard Research Grant #861-2009-0071.

Table of Contents

List of Tables

List of Figures

Chapter 1 - Introduction

Health in ancient populations is an important topic of study in biological anthropology, and yet it is difficult to study from the holistic perspective it deserves. Indeed, health is not one topic, but is rather the product of many biocultural forces: infectious and genetic diseases, diet and nutrition, exposure and sanitation, stress, reproduction, and so on. Most often, health is examined using approaches which reveal only one aspect; typically this approach is the study of macroscopic skeletal evidence of disease. Unfortunately, this approach does not necessarily reveal, for instance, how diet and nutrition contributed to the incidence of disease, nor does it aid in understanding the underlying risk to disease. Macroscopic paleopathology, stable isotope analyses, paleoethnobotany, paleodemography, molecular paleopathology and paleoepidemiology, are bioanthropological approaches that should be drawn together to help create an overall picture of the health of a population.

In particular, molecular paleopathology and paleoepidemiology are both important subdisciplines of biological anthropology. Moreover, these approaches have not seen application in a Caribbean, specifically pre-Columbian Caribbean, context. Therefore, for this project, my research objectives are to: 1) evaluate the disease of spina bifida, present in the skeletal assemblage from the pre-Columbian Cuban population of the archaeological site of Canimar Abajo; 2) determine the distribution of single nucleotide polymorphisms (SNPs) associated with spina bifida in the population using the techniques of molecular paleopathology; 3) discuss the distribution of these SNPs from a paleoepidemiological perspective, comparing this distribution to macroscopic

evidence of the disease and to dietary indicators; and 4) contribute to the understanding of how molecular paleopathology and paleoepidemiology can be used to its full potential within biological anthropology. This project is part of a larger multi-disciplinary, international study that examines the identity of pre-Columbian and contact period Cuban indigenous groups, and their interactions with the natural and created environment.

Paleopathology, a subdiscipline of bioanthropology that examines disease and abnormal variations in human remains, has encountered a number of problems and limitations (Roberts and Manchester 2005). Described collectively as the osteological paradox, these problems include the difficulties of interpreting selective mortality, individual risk and susceptibility (Roberts and Manchester 2005; Wood et al. 1992). Moreover, it has proven difficult to make any conclusions about the overall distribution of pathologies in a population due to problems shared with the study of paleodemography, namely non-stationary populations, issues of representative sampling, and so on (Roksandic and Armstrong 2011; Wood et al. 1992). Many of these issues are related to the limited manifestations and ambiguous interpretations of disease in bone tissue.

Molecular paleopathology is no longer a novel approach (Herrmann and Hummel 1998; Hofreiter et al. 2001; Hummel and Herrmann 1995). The major advantage of molecular versus macroscopic paleopathology lies in its ability to recognize the presence of diseases before they are observable in the skeleton. It avoids the difficulties of differential diagnosis arising from a limited number of disease manifestations in bone tissue, and provides unambiguous identification of pathogens. Moreover, it allows for recognition of active infection for certain diseases (Donoghue et al. 2005; Haas et al.

2000a). Consequently, molecular paleopathology avoids many of the difficulties associated with the osteological paradox (Wood et al. 1992). While previous studies encountered issues of DNA contamination and preservation, and inadequate laboratory methods, these problems are rapidly being rectified, making it possible for molecular paleopathology to see more widespread use (Donoghue et al. 2004; King et al. 2009; Matheson et al. 2009; Roberts and Ingham 2008). Using molecular techniques, a number of bacterial, viral, parasitic and genetic diseases can now be recognized, including tuberculosis (Salo et al. 1994), leprosy (Haas et al. 2000b), malaria (Nerlich et al. 2008), cystic fibrosis (Bramanti et al. 2003), and spina bifida (Armstrong et al. in press; Cloutier 2008).

Molecular paleopathology can only provide the necessary tools to recognize disease; it is by using a paleoepidemiological framework that we can reach an understanding of the health of a population. Paleoepidemiology is an interdisciplinary approach which considers how the prevalence of disease reflects health, lifestyle, and environment (Souza et al. 2003), and which aims to make statements about health and disease in a skeletal assemblage (Waldron 2007). Health is a holistic concept, in which infectious and genetic diseases play a small, but important role concurrently with diet, stress, reproduction, and overall quality of life and well-being (Bush and Zvelebil 1991). With these broad concepts in mind, molecular paleopathology will be used to investigate the health of a pre-Columbian population, centered on the interactions of genetics and diet.

The focus of this research will be on spina bifida, a genetic disease known to be present in high proportions in the population of Canimar Abajo. Neural tube defects are

common types of birth defects that involve the malformation of the central nervous system. Spina bifida, characterized by the midline separation of vertebrae, is one of these defects; it is a complex, multi-factorial disease ranging in severity, from asymptomatic to severe clinical presentations. Although considered a genetic disease, the incidence of spina bifida is also highly dependent on environmental factors, especially pre- and peri-conception maternal folate intake (Werler et al. 1993). Most of the mutations linked to the incidence of spina bifida focus on enzymes involved in the folate and homocysteine (an amino acid) metabolic pathway; mutations in these folate-dependent enzyme genes result in reduced enzymatic efficiency of homocysteine processing, thereby potentially contributing to the incidence of the birth defect (Cloutier 2008). Consequently, this disease is of particular interest from a paleoepidemiological perspective, as its prevalence can reflect both genetic and environmental factors.

Based on macroscopic paleopathological evidence in preserved sacra, it was previously uncovered that the disease was present in high proportions in a pre-Columbian Cuban population from the archaeological site of Canimar Abajo, Matanzas. Of the 21 preserved sacra from the site of Canimar Abajo, 14 demonstrated spina bifida, 67% of the sample population (Campillo Álvarez 2009). Though not directly comparable, of 273,346 births in Cuba from 1993 to 1998, there were 70 cases of spina bifida (birth prevalence x 10,000 of 2.56) (International Centre for Birth Defects 2003). For this period, the highest birth prevalence (x 10,000) was 15.25, observed in Mexico (International Centre for Birth Defects 2003). With such a high prevalence, I examined the distribution of five single nucleotide polymorphisms associated with spina bifida, including individuals both with and without macroscopic evidence of the disease; given the known association with diet,

I also chose to compare the genetic data with stable isotope, starch, and zooarchaeological evidence for the diet of the sample population. Thus, a study of spina bifida at Canimar Abajo provides a template for future paleoepidemiological investigations, combining skeletal, genetic, and dietary information to understand health in a particular population.

Chapter 2 - Literature Review

A paleoepidemiological study requires considerable background knowledge on the site and population, and the various types of evidence that will contribute to understanding health. This theoretical approach can be described as 'thick archaeology', building an interpretation from a sum of evidence using a variety of techniques (Carr and Case 2006). To contextualize the study and to understand the Canimar Abajo population, it is essential to include a discussion on Cuban archaeology in terms of terminology and assumptions regarding both Cuba as a whole, and the site of Canimar Abajo specifically. Moreover, to conduct such a holistic study of spina bifida at Canimar Abajo, it is necessary to review relevant past work, namely macroscopic paleopathology, stable isotope analysis, zooarchaeology, and paleoethnobotany. Only then is it possible to focus on spina bifida, including significant mutations associated with this genetic disease and how these underlying genetic mechanisms interact with the environment of the individual and the population. All of this information, multidisciplinary as it is, can be gathered together to discuss the paleoepidemiology of spina bifida in the Canimar Abajo population. This sum of information then provides a more colourful, and hopefully more accurate, picture of health at Canimar Abajo.

Paleoepidemiology

Before discussing the specific context of this study, it is necessary to review paleoepidemiology as an approach, and how molecular paleopathology can contribute to this approach. While modern epidemiology seeks to demonstrate the relationships

between risk factors and disease outcome (Pinhasi and Turner 2007), paleoepidemiology acts as a framework to understanding how disease was spread throughout an archaeological population, incorporating a variety of aspects related to the occurrence of disease to understand population health. Paleoepidemiology has been advocated as an approach to studying health in the past since the late 1960s (e.g. Angel 1966; Clark 1988; El-Najjar et al. 1975; Jurmain 1990), but has become less common in the post-'osteological paradox' bioarchaeological research environment. Basically, although paleoepidemiologists may seek the same type of information as modern epidemiologists, it is usually not possible to accomplish this in light of the limitations of working with an archaeological population. These limitations include those shared with paleodemography: non-stationary populations, non-representative sample populations, and the inability to apply cohort statistics in an archaeological context (Roksandic and Armstrong 2011). Instead, paleoepidemiology generally considers how the prevalence of disease, as observed in the skeletal evidence, reflects health, lifestyle, and the environment (Souza et al. 2003), and is thus a more qualitative approach than modern epidemiology. Data from molecular paleopathology can broaden this framework, allowing one to begin discussing the underlying heterogeneity of risk.

All archaeological research that attempts to study the population as a whole encounters the problem of the representativeness of samples. The sample population comes from a total population of unknown size and structure, and is never a truly random sample; the total living population has been reduced to a sample burial population by cultural burial practices, burial taphonomy, and archaeological methods (Souza et al. 2003). Moreover, the sample population may represent an extended period of time of

cemetery usage, rather than representing a group that can be treated as a stable, single generation cohort, as is the case in modern population studies (Roksandic and Armstrong 2011). Though bias in a sample population can be recognized (e.g. Jackes 1993; Jackes and Meiklejohn 2008), it is difficult to rectify this bias in order to conduct robust statistical analysis; the paleoepidemiologist must often resort to calculating the percentage of the total archaeological population afflicted by a disease without knowing how the archaeological population relates to the living population, or how many were afflicted by the disease without skeletal involvement (Waldron 2007). These problems make it difficult to draw quantitative conclusions regarding the distribution of disease in the original population. Nonetheless, paleoepidemiological research has value for understanding health and disease in the past, though interpretations must be drawn cautiously.

A long-standing issue in the study of the health of past populations is that of the 'osteological paradox', first outlined by Wood and colleagues (1992). In particular, the team deemed it impossible to uncover the underlying heterogeneity of risk that contributes to the problem of selective mortality. Selective mortality refers to the issue that an archaeological population represents individuals that died at a particular age of a particular cause, rather than those that were at risk of a particular cause at a particular age (Wood et al. 1992). An archaeological population instead reflects a biased sample of the sum result of all causes of death over an undefined period of time, making any interpretation of population health almost impossible. In this study, however, the distribution of disease in the population is not the focus; rather it is the underlying heterogeneity of risk (susceptibility) of a population to a specific non-infectious disease

that is of interest a subject that is not possible to discuss with macroscopic paleopathological methods, as discussed by Wood and colleagues (1992).

A population is composed of a mix of individuals of differing susceptibility to disease; this susceptibility represents the sum of genetic predisposition, socioeconomic status, environment, diet, exposure to disease, and so on (Wood et al. 1992). Using traditional macroscopic methods, disease in archaeological populations is ascertained by the limited manifestations of disease in bone, most of which are non-specific rather than pathognomic indicators. Briefly, though an entire population may be exposed to disease, these individuals all have varying degrees of susceptibility; only a few will experience sufficient illness to result in a bony outcome, while those who succumb to the disease quickly, those who recover quickly, and those who never experience disease will all have apparently 'healthy' skeletons, with no evidence of disease (Wood et al. 1992). The nature of susceptibility is slightly different when discussing a non-infectious disease, where risk is determined mainly by genetics or environment rather than exposure to an infectious organism, but the problem remains in that skeletal evidence may not be present in an affected individual. Thus, hidden heterogeneity is a major confounding factor in studying disease in the past by traditional macroscopic methods.

As noted by Souza and colleagues (2003) and by Waldron (1994; 2007), these problems, valid though they may be, should not prevent attempts to study paleopathology and paleoepidemiology. Instead, the limitations should be recognized, a systematic approach should be taken, and the available information and results should be considered as quantitatively and qualitatively as possible. In essence, paleopathology and paleoepidemiology still have value for reconstructing a general picture of health and

disease, and allows us to understand more than mere descriptions of pathological conditions in bones (Souza et al. 2003).

Molecular Paleopathology

The optimistic approach of Waldron and Souza and colleagues is encouraging, but it is still necessary to attempt to circumvent the limitations of current paleoepidemiological and traditional paleopathological approaches. Molecular paleopathology may be one approach that can help decipher the distribution of a disease or risk of disease in a population. One example of this is Aufderheide and colleagues' (2004) epidemiological study on Chagas' disease in Chile over 9,000 years; the team used molecular methods to study the prevalence of the disease, examining its prevalence diachronically in different parts of the population and revealing the role of human populations among the animal hosts of the disease pathogen. Studies like this reveal how paleoepidemiology can help to reconstruct past population health, revealing aspects of human behaviour that are rarely considered in traditional paleopathological studies.

Studying multi-factorial genetic diseases, rather than infectious diseases, can allow for an even more comprehensive picture. Multi-factorial diseases are so named because of the many factors involved in their outcome; though genetics lead to an individual's predisposition or risk, environmental and behavioural factors also play important roles in how or whether the disease manifests. Molecular paleopathology can elucidate this genetic predisposition, while other bioarchaeological methods can complement this data to understand the interaction of these various factors in the disease outcome.

In terms of methods, studying multi-factorial diseases avoids problems of skeletal sampling area that arise in the molecular study of infectious disease (Pinhasi and Turner 2007), both in terms of where pathogen DNA is likely to be found in the skeleton based on the etiology of the pathogen, and the concerns over taking samples from pathognomic areas that the bioarchaeologist would prefer to leave untouched. These are non-issues in the study of genetic diseases, where the best sampling area is the one that is most likely to have preserved DNA (areas that are less susceptible to taphonomic damage and diagenesis), as the goal is to recover human rather than pathogen DNA.

The use of nuclear DNA, required for the molecular study of most genetic diseases, is often perceived to be much more difficult because it is present in fewer copies than mitochondrial DNA and thus more susceptible to degradation. In addition, it may be more difficult to identify contamination in the study of nuclear DNA than pathogen DNA. However, it has been proven that nuclear DNA contamination is much less of an issue compared to mitochondrial DNA (Burger et al. 2007), and the amplification of very small DNA sequences, made possible by the study of single nucleotide polymorphisms, increases the success rate of working with nuclear DNA dramatically. Thus, while working with nuclear DNA certainly presents some difficulties, many of these can be overcome with repeated experiments, careful practices, and cautious interpretations.

One important limitation that remains, at least for the study of spina bifida, is that one cannot be sure whether those individuals who carried susceptibility alleles in fact displayed any phenotypic evidence of spina bifida or whether they suffered at all as a result. Spina bifida is a complex disease, and many factors contribute to its etiology. Here, macroscopic paleopathology can be helpful, since it may be possible to diagnose

different degrees of spina bifida in the sacrum or other vertebrae, but only in instances where these bones are preserved. In addition, evidence of bone atrophy of the lower limbs may suggest paralysis as a result of spina bifida, but this is also not always possible to observe. Overall, while macroscopic skeletal evidence can confirm that spina bifida itself was present in the population, and molecular and dietary evidence can provide information on the underlying risk to spina bifida, it is not possible to determine how many individuals were affected by the disease and whether they suffered as a result. Nevertheless, this type of paleoepidemiological study represents an interesting convergence of biological and cultural information, considering spina bifida from the perspective of genetic predisposition (which has implications for population reproduction patterns), disease outcome (in terms of macroscopic evidence of spina bifida), and dietary factors that interact with these two aspects.

Archaeological Context

The population studied here is from the pre-Columbian archaeological site of Canimar Abajo, Matanzas, Cuba. It is important to contextualize this population within the academic literature on pre-Columbian Cuban populations, which has often proven to be a contentious area of research.

Distinguishing different cultural and/or ethnic groups from prehistoric times is an endeavour fraught with difficulties, and one which must overcome the obstacles presented by colonial powers, modern populations, and contradictions in both the archaeological and ethnohistoric evidence. These are the issues that arise in the study of indigenous groups in Cuba before and immediately after the time of European contact in

1492. Names including Guanahatabey, Ciboney, Taíno (Western, Classic, etc.) and Arawak are used, often interchangeably, with little regard to whom they refer. Adding to the confusion is the use of terms by modern researchers to refer to particular migrations and/or ceramic and lithic technology. These will be discussed in turn in order to determine the best framework under which to operate for the purposes of further research in the area.

The earliest evidence of human habitation in Cuba is seen approximately 5,500 years ago, at the sites of Canimar Abajo (UBAR-170: 5030-4622 cal BP 2σ; UBAR-171: 5590-5300 cal BP 2σ) and Levisa I (GD-250: 6288-5584 cal BP 2σ) on the Western and Eastern parts of the island, respectively (Cooper 2010; Dacal Moure and Rivero De La Calle 1996). Although these radiocarbon dates are based on associated charcoal rather than from the skeletal material directly, dates from skeletal material support the antiquity of human habitation on the island (see Table 2.1; for a review of site chronology of pre-Columbian Cuba, see Cooper 2010). From this point onwards, the island saw a series of migrations that continued until well after European contact. Though it is difficult to discern the source of these early migrations, many have been suggested based on environmental patterns such as wind and water currents, ceramic and lithic typologies, dental morphology and most recently, ancient mitochondrial DNA (Callaghan 2003; Coppa et al. 2008; Dacal Moure and Rivero De La Calle 1996; Lalueza-Fox et al. 2003; Wilson 1997). Overall, however, the series of migrations seems to have left a number of pre-Columbian groups on the island at different periods of time, which have been assigned a variety of names and terms.

Groups in Pre-Columbian Cuba: Ethnohistoric Paradigm

One major classification system used in Cuba in reference to pre-Columbian populations is based largely on ethnohistoric evidence. This paradigm uses the terms found originally in ethnohistoric literature, to which modern research has conformed, and includes the Guanahatabey/ Ciboney/ Taíno differentiation.

The name 'Guanahatabey' is probably the most controversial and ambiguous in the modern Cuban archaeological literature, and the group this term refers to is the most puzzling. Generally, Guanahatabey is the term used to describe the mysterious people who lived in Western Cuba, who may or may not have existed or survived to the time of European contact, and who were more 'primitive' than their eastern neighbours (Reid 2009). They are often described as semi-mythical cave-dwellers because the accounts of those who claimed to have seen them are unreliable, and most of the information regarding their lifestyle seems to have come from the rather biased perspective of their neighbours to the east, the Taíno (Keegan 1989). The earliest encounter with this group may have been during Columbus's second voyage to the Americas, when he voyaged along the coastline of Cuba. This voyage is described by Andrés Bernaldez, who mentions people who lived in small groups, without permanent villages (Ramos 2008). Unfortunately, this description does not use the term 'Guanahatabey' specifically (Ramos 2008). A few years after this voyage had occurred, in 1516, Governor Diego Velázquez wrote a similar description, adding that these people did not practice agriculture and ate mainly fish and meat that they hunted (Ramos 2008). This information was later repeated by the friar Las Casas, who noted that their language was different from the groups on the

eastern part of the island (Granberry and Vescelius 2004), though some modern researchers have interpreted his writings to say that he described the western group as 'Ciboney', meaning 'cave-dweller,' rather than Guanahatabey (Dacal Moure and Rivero De La Calle 1996).

Unfortunately, all of the ethnohistoric accounts were written in the 16th century or later, making it difficult to assess the conditions at time of contact. Though these accounts do consistently describe a group of cave-dwellers who lived a hunter-gatherer lifestyle, these accounts also tend to consist of second-hand information. Some modern researchers, such as Keegan, believe that Las Casas and Velázquez never actually ventured to the western part of Cuba where the Guanahatabey were thought to reside, instead relying on descriptions of them by the Taíno (Keegan 1989). Furthermore, Keegan notes that shortly after time of contact, Panfillo de Narvaez did visit the area and found the people and culture to greatly resemble that of the east (Keegan 1989). By this time, there were five chieftainships with Taíno names – Guanahacabibes, Guaniguanico, Marien, Habana, and Hanabana, and hence the confusion between the names 'Guanahatabey' and 'Guanahacabibes' (Reid 2009). Unfortunately, it is unclear whether there is any relationship between these two names, or whether the language spoken at this time and in this area was a Taíno language. In any case, these accounts would imply that whether or not the Guanahatabey actually existed before contact, by the time ethnohistoric records properly demonstrate voyages into Western Cuba, Taíno culture had already spread and absorbed any remaining Guanahatabey culture.

Regardless, it would seem that whether they identified themselves as Guanahatabey or not, there were a group of people in Western Cuba at time of contact

that may have differed from the peoples to the East until shortly pre- or post-contact (Ramos 2008). Studies examining mitochondrial DNA and dental morphology have both indicated the existence of a separate group in the west as compared to the east of Cuba up until shortly before the time of contact, in 1400 AD (Coppa et al. 2008; Lalueza-Fox et al. 2003). Many early sites in Cuba are in this part of the island (Cooper 2010; Dacal Moure and Rivero De La Calle 1996), and the lithic technology here is referred to as Archaic or Casimiroid, including flaked stone blades, mortars, pestles, balls, and beads (Reid 2009). Though this technology spread across the Greater Antilles, it has been speculated that it persisted in Western Cuba until the 1600s, practiced by the Guanahatabey, unlike other Archaic traditions of the Greater Antilles, because of its distance from the Taíno and the Spanish (Granberry and Vescelius 2004). While it has been claimed that the Guanahatabey did not create ceramics, the archaeological evidence could be interpreted differently. Though quite rare, there have been Cuban sites with evidence of ceramics in 2000 BC, well before the arrival of the Ostionoid migration in 600 AD (Ramos 2008). Also in question is whether or not this group settled permanently, for though no permanent structures have been found, there are large shell middens which could imply long-term settlement. When this is considered with the island landscape and lack of seasonal changes, it would seem that there would be no necessity for continuous mobility, and consequently no reason to believe that the Guanahatabey did not settle at least semi-permanently (Ramos 2008). These small pieces of evidence are all that have been discovered regarding the material culture of the Guanahatabey so far, leaving many questions unanswered.

Overall, interpretations of the evidence have stated that the Guanahatabey were a

'relic population', meaning that they were part of one the initial migrations to Cuba who lasted longer than in other areas because of their geographic isolation (Rouse 1992). Based on similarities in the lithic technology, both the Yucatan peninsula and southern Florida have been suggested as possible sources for the earliest migrations into Cuba (Hayward et al. 2009; Kozlowski 1974), though the Yucatan is generally the more favoured opinion (Reid 2009). Computer simulations incorporating both lithic typology and environmental reconstructions have suggested the Yucatan and northern Venezuela of South America as possible sources; ancient mitochondrial DNA also suggests South America, though it cannot exclude the possibility of the Yucatan as being the source of the first wave of migrations (Callaghan 2003; Lalueza-Fox et al. 2003). The fate of this group is unknown: did this group persist until European contact, did they assimilate with the Taíno, or did they disappear by some other means?

The second major wave of migration into Cuba seems to have originated from northern South America based on ancient mitochondrial DNA and dental morphology (Coppa et al. 2008; Lalueza-Fox et al. 2003). This migration into Cuba may or may not have been in conjunction with the Saladoid migration, also from northern South America, which brought an end to the Archaic populations and which brought ceramic and agricultural technology in the rest of the Caribbean islands (Curet 2003). While dental morphology would seem to indicate an ancestral relationship between the Taíno in Cuba between 1000 and 1500 AD, and the Saladoid people in Puerto Rico and the Virgin islands between 250 BC and 1500 AD (Coppa et al. 2008), most researchers believe that the Saladoid migration did not reach Cuba, often based on the fact that agriculture did not arise in Cuba until much later than the other Caribbean islands (Curet 2003). While many

agree that the Taíno developed from the Ostionoids present in Cuba, if they even make the distinction between Ostionoids and Taíno, it is unclear whether the Ostionoids developed from the Saladoids in Hispaniola, or whether it was a separate migration, though it is generally agreed that the Ostionoids were present in Cuba by 600 AD (Hayward et al. 2009; Stevens-Arroyo 2006). Some even deny any additional migrations, insisting rather that the Ostionoids developed from a group of Casimiroids already present in Cuba, which may be confirmed by the continuous production of rock art on the island (Reid 2009; Roe 2009). Still others suggest a mix of these two, offering the possibility that the Ostionoids migrated into Cuba from Hispaniola and that there was transculturation with the native Casimiroid technology, based on the sudden appearance of widespread ceramic technology in eastern and central Cuba resembling that of Hispaniola at this time, but with the retention of lithic technology as well (Wilson 2007).

In any case, from this group arose the Taíno, who were widespread across the Caribbean, from Cuba to Haiti and the Dominican Republic, to the Bahamas and the Virgin Islands, with many regional varieties. In Cuba, one variant was the group called 'Ciboney'. Las Casas, the main ethnohistoric source, describes the Ciboney as a local Taíno group, and writes that the term 'Ciboney' was used by the people themselves, but modern researchers have also used the term Ciboney interchangeably with Guanahatabey (Granberry and Vescelius 2004; Reid 2009). This confusion seems to have occurred fairly early on in Caribbean research, and became widespread with Lovén's seminal work *Origins of the Tainan Culture* (Lovén 1935). Cosculluela (1965 [1918]) seems to have been the first to use 'Ciboney' in lieu of 'Guanahatabey' after his excavations at Ciénaga de Zapata in 1918 (Ramos 2008). Nonetheless, Cosculluela later retracted this term in

1946, in favour of ‘Guanahatabey’, arguing that ‘Ciboney’ should be used the way Las Casas had intended, to refer to a subgroup of the Taíno (Ramos 2008). To add to this confusion, some have also used ‘Ciboney’ to describe a population ancestral to the Guanahatabey (Saunders 2005). In recent years, this confusion has led many to use the term ‘Western Taíno’ rather than ‘Ciboney’, using the term ‘Classic Taíno’ to refer to those in the far east of Cuba (Rouse 1992).

The Western Taíno resided in central Cuba and can be differentiated from their eastern neighbours archaeologically, as they used a slightly different style of ceramic decoration, called Meillacan, which used fine line incisions rather than rows of punctures as seen in the Chican style of the eastern Taíno (Granberry and Vescelius 2004). Otherwise, the Western Taíno were culturally similar to the eastern or Classic Taíno in terms of social organization and religion, though they may have spoken a different dialect of the Taíno language than those to the far east of the island (Granberry and Vescelius 2004). The Taíno residing on the eastern part of Cuba, which Rouse calls Classic Taíno (1992), are thought to have been a group from Hispaniola who fled to Cuba to escape the Spanish, only settling in Cuba approximately fifty years before Las Casas arrived (Granberry and Vescelius 2004).

The history and source of the term Taíno is unclear. Some say that the term ‘Taíno’ means good or noble, and was used by some of its members to Columbus’ crew to distinguish themselves from the Island Caribs who were thought to be violent and cannibalistic people, rather than a name to refer to themselves collectively (Hulme 1993; Reid 2009). Others say that the term was first used by Constantine Samuel Rafinesque in 1836 to refer to the language spoken in Haiti, and that the term was only later used to

describe natives in the northern Caribbean at contact (Reid 2009). Reid (2009) suggests the possibility that the term *nitayno*, which can be found in Columbus' diary entry for December 3, 1492 to refer to an important or noble person, similar to *cacique*, was perhaps bastardized into the term 'Taíno'. Unfortunately, Columbus' diary has only survived as paraphrased by Las Casas, making it unclear whether it is a reliable source (Hulme 1993). Furthermore, it is unclear whether the Taíno ever conceived of themselves as a cohesive group, given that the group we consider to be Taíno is actually composed of a large number of chieftainships and other variants who are culturally similar, both in Cuba and in other Caribbean islands (Reid 2009). It is more likely that each group referred to themselves by a local name, such as Ciboney in central Cuba, Borinquen in Puerto Rico, and Lucayo in the Bahamas (Rouse 1992). This problem has led some to suggest that the term 'Taíno' be discarded in order to better reflect this variation, but as Reid (2009) notes, these groups do tend to form a single Taíno language group, and for the most part, the term Taíno has been retained, often with the local names for specific groups. Indeed, it has been suggested that although many of these groups had their own dialects, or perhaps even their own languages, Taíno may have served as a *lingua franca*, or trading language (Granberry and Vescelius 2004; Reid 2009).

To further complicate matters, the term 'Arawak' has been used to describe the Taíno, beginning very early on in Caribbean research in 1871, due to the cultural and linguistic similarities to the Arawakan Indians of South America (Hulme 1993; Rouse 1992). More recently, the term 'Arawak' has been used to describe the larger language family to which the Taíno language belongs, encompassing a variety of dialects throughout the Caribbean islands (Granberry and Vescelius 2004). Although the

relationships between these languages is difficult to determine since most have died since the time of European contact, attempts have been made to understand these relationships based on ethnohistoric records and toponymic data (Granberry and Vescelius 2004).

Unlike the semi-mythical stories, and inadequate archaeological and ethnohistoric information which is all that is known of the Guanahatabey and their culture, a relative wealth of information is known concerning the Taíno of Cuba, including both the Classic and Western/Ciboney people. In addition, the Taíno have left a diverse material culture and are attributed with creating rock art, including anthropomorphic and geometric designs, and the first incised petroglyphs (Roe 2009). Clay figurines shaped to have female characteristics, masks (*caretonas*), and jewellery are also found in the Taíno creative repertoire (Dacal Moure and Rivero De La Calle 1996).

The Taíno, as a group, can be described to be permanently settled, agricultural people, with a complex social organization. Their subsistence practices included exploiting marine foods, including fish, turtles, and shellfish, certain birds and lizards and using methods such as slash-and-burn agriculture (Stevens-Arroyo 2006). Taíno agriculture, and furthermore, their economy, relied heavily on the exploitation of one crop by a community, such as yucca, squash or beans (Stevens-Arroyo 2006). This type of economy has been called a 'harvesting economy', which is to say that this reliance on one crop required the trading of commodities; for instance an inland community might rely on harvesting yucca and trade this product for fish from coastal communities (Stevens-Arroyo 2006).

It has been speculated that the development of this level of economic complexity led to the development of social inequalities and thereby the system of *caciques* (chiefs),

nitayno (ruling class, elite), and *naboría* (working class), as well as increased labour specialization, such as farmers and artisans (Stevens-Arroyo 2006). This organization was one of a complex chiefdom with chiefs (*caciques*) serving as community leaders and religious functionaries (Hayward et al. 2009). Their ceremonies, or *areytos*, occurred for major events such as a *caciques'* marriage or death and consisted of feasting, singing and dancing, mock battles, and ball games (Hayward et al. 2009; Stevens-Arroyo 2006).

Underlying these ceremonies was a complex belief system, which was based on the existence of supernatural beings (*cémis*), who were represented in small, usually anthropomorphic, three-point sculptures, also called *cémis*, created from a variety of material including stone, wood, and coral (Dacal Moure and Rivero De La Calle 1996; Hayward et al. 2009). The pantheon of gods consisted of a supreme male god Yúcahu Maórocoti, a fertility goddess Attabeira, and an assortment of lesser gods (Hayward et al. 2009). The creation of the *cémi* idols was itself a religious experience, involving discussion with the gods, as described by the friar Ramón Pané in his description of the Taíno of Hispaniola (Pané 1999). The religion of the Taíno included a complex mythology, as well as rituals for which the *cémi* idols played an important role, both of which have been studied at length by modern researchers (Dacal Moure and Rivero De La Calle 1996; Keegan 2007; Stevens-Arroyo 2006).

Overall, what can be gathered from the ethnohistoric sources, and from the archaeological evidence that has been gathered over time, is that little is known regarding the Western populations of prehistoric Cuba regarding their origins, subsistence, and other aspects. Indeed, the lack of evidence and the confusion over what little evidence there is, has led Cuban archaeologists to forego many of these terms, instead relying on

classification groups based on what is collected at individual archaeological sites.

Groups in Pre-Columbian Cuba: Economic Paradigm

Given the disparity between sources of evidence, and given the confusion outlined above, it may not be appropriate to use the term 'Guanahatabey', as it has implications for what is known about the Western group (or groups). Instead, it may be more appropriate to subscribe to the current framework used by Cuban archaeologists, which is a site classification system that does not rely solely on ethnohistoric sources or typological assumptions.

The system proposed here was first devised by Tabío and Rey (Tabío 1988; Tabío and Rey 1979), and is based on the presence or absence of evidence for agriculture and ceramic production: *preagroalfarero* (preagroceramacist); *protoagricola* (protoagriculturalist); and *agroalfarero* (agriculturalist and ceramic producers) (Cooper 2010). This system is often used in conjunction with another one based on the economic classification of artefacts: hunting; fishing and collecting; incipient agriculture; and agriculture. *Preagroalfarero* sites are characterized by the presence of shell artefacts and stone tools, and by an absence of any ceramics or artefacts associated with plant production or processing (Rodríguez 2007; Tabío 1988). Of course, *agroalfarero* sites are characterized by the opposite; these sites have few stone or shell artefacts, but have many ceramics, including those associated with cultivated plant processing such as *buréns*, described as ceramic griddles (Rodríguez Suárez and Pagán Jiménez 2008). In between these two extremes is the classification of *protoagricola*, which is described as a site with evidence of the onset of ceramic technology and incipient agriculture (La Rosa Corzo

2003), when plants were only just beginning to be cultivated. This category was created to classify sites with a stone tool technology similar to *preagroalfarero* sites, but with ceramics as well (Ramos et al. 2008). Though both of these systems have problems of their own (not least of which is the perhaps unrealistic association of pottery and agriculture, an assumption made in many regions of the world, see Ramos et al. 2008), they are the most widely used classifications in Cuba, and may be the most appropriate to apply to the study of western Cuba until more is understood about these populations.

Under this framework, the people of western Cuba are generally characterized as being *preagroalfarero* hunter-fisher-gatherers. Indeed, if one examines the distribution of sites using these two classification systems, there is a distinct predominance of these sites in the western part of the island, which then trends towards *agroalfarero,* full agricultural sites on the eastern part of the island, inciting many researchers to correlate the archaeological evidence with the ethnohistoric discussions of the Guanahatabey (in the West) and the Taíno (in the East) (Cooper 2010). Unfortunately, these assumptions may be exacerbated by targeted archaeological surveying and preconceived notions that skew the interpretation of archaeological evidence in order to conform to the current understanding of the distribution of site types.

It is only with more recent research, particularly at the site of Canimar Abajo, that these generalizations are beginning to be overturned (Chinique de Armas 2007; Chinique de Armas 2009; Chinique de Armas et al. 2008; Rodríguez 2007). This new research is indicating that plants had a much larger dietary role than previously understood, and implies that the subsistence system was more complex than assumed. An increased proportion of plants in the diet also has implications for the health of the western Cuban

groups generally, and the Canimar Abajo population specifically. Until this picture is clarified, however, Canimar Abajo remains classified as a *preagroalfarero* site, with a primarily hunter-fisher-gatherer population, although the importance of gathering and the possibility of low-level food production (*sensu* Smith 2001) remain to be determined.

Case study: Canimar Abajo

The population under study is an archaeological population from the site of Canimar Abajo, Matanzas, Cuba, which is an important, early pre-Columbian site on the southwestern bank of the Canimar River (23°02’15’’N, 81°29’41’’W; Rodríguez Suárez et al. 2006) (Figure 2.1). First discovered in the 1960s by campers, Cuban archaeologists from the Montané Anthropological Museum and the Faculty of Biology of the University of Havana have been excavating at the site intermittently since 1984 (Martínez López 2009). Although most of the excavated material is housed in the Montané museum, it is thought that, over time, some skeletal material has been lost to other institutions in the country based on unpublished site reports from Dr. Rivero de la Calle, one of the original excavators in the 1980s (Dumas León 2009). Studies so far have focused on: bioanthropological studies of human skeletal material (Arenas Laserna 2009; Arenas Laserna and Arredondo Antúnez 2008; Chinique de Armas and Arredondo Antúnez 2008; Morales and Arredondo Antúnez 2008; Morales Valdes 2009), mortuary practices (Cabrera 2008), taphonomy and skeletal preservation (Martínez López 2009), zooarchaeology (Arredondo 2004), macroscopic paleopathology (Campillo Álvarez 2009; Rivero de la Calle 1988), and paleodiet (Buhay et al. In press; Chinique de Armas 2007; Chinique de Armas et al. 2008; Rodríguez 2007).

More recently, since 2009 an interdisciplinary Canadian team of researchers has teamed up with the Cuban archaeologists to diversify the research at the site, given its importance in Cuban prehistory. This new research includes: further developing paleodietary studies using stable isotope and starch and phytolith analyses; radiocarbon dating of skeletal material; ancient mitochondrial, nuclear, and pathological DNA studies; comparisons of ancient DNA with modern DNA of Cuban populations; linguistic place-name analysis; studies of dispositional taphonomy and burial ritual; lithic residue analysis, and; geological studies of the site and surrounding area. Many of these studies have implications for understanding the health of the Canimar population.

Figure 2.1: Location of Canimar Abajo, Matanzas, Cuba (adapted from Google Maps).

Canimar Abajo, a sheltered site at the base of a karst outcrop, is primarily a burial site, consisting of at least 199 individuals in two cemeteries, separated by roughly 1,500 years. The first cemetery is located at a depth of 1.80-1.50 m, while the second is from

0.60 m to the surface level (Morales Valdes 2009). In between these two burial levels (1.50-0.60 m) is a habitation, ritual, or food-processing layer, with faunal remains and evidence of burning, described as a shell midden (Morales Valdes 2009). The site is roughly 40 m from the Canimar River, and the surrounding vegetation consists mainly of bushes, mangroves and, further off, semi-deciduous forest (Dumas León 2009). This vegetation, and the food sources available from the river itself, allow for a diversity of subsistence activities (Chinique de Armas et al. 2008), while the karst outcrop provided a protected rock shelter (Dumas León 2009) for the inhabitants of the area.

Radiocarbon dates from charcoal have provided dates as early as 5590-4622 BP (Cooper 2010; Rodríguez Suárez et al. 2006), while more recently acquired dates from skeletal material indicate that the site spans at least 3000-1250 BP (Table 2.1; Rodríguez Suárez et al. 2010). Although more dates are required, especially considering the difficulty of taphonomic interpretation of the site chronology due to bioturbation and the reuse of burial space, these dates make Canimar Abajo one of the oldest sites in Cuba, particularly Western Cuba (Cooper 2010; Dumas León 2009; Martínez López 2009; Rodríguez Suárez et al. 2006).

Table 2.1: Calibrated dates from Canimar Abajo (Rodríguez Suárez et al. 2010). Skeletal samples calibrated using IntCal09 (Reimer et al. 2009); charcoal samples calibrated using IntCal04 (Reimer et al. 2004).

Laboratory code	Sample	Material	Cal 2σ lower range	Cal 2σ upper range
UBAR-170		Charcoal	2950 BC	2750 BC
UBAR-171		Charcoal	3590 BC	3300 BC
AA89060	E-92 (38) 18	Human bone	AD 582	AD 660
AA89061	E-15 (6)	Human bone	1294 BC	1055 BC
AA89062	E-79 (31)	Human bone	AD 417	AD 622
AA89063	E-19	Human bone	1217 BC	1012 BC
AA89064	E-72 (13)	Human bone	AD 335	AD 554

The composition of the burial population is unique in terms of age distribution. As of the 2009 excavation season, including material excavated from the 1980s and the 2000s, the minimum number of individuals (MNI) at the site was determined to be 199 individuals. The age of these individuals was determined using standard methods of dental eruption, dental formation, and epiphyseal closure for juveniles, and analysis of the pubic symphysis and secondary coxal ossification for adults (Buikstra and Ubelaker 1994). More recently, these individuals have been aged according to the model of Roksandic and Armstrong (2011) in order to elucidate the age distribution by life history stage.

Although an in-depth analysis of the paleodemography of the population is not yet complete, it is clear that there are a particularly high number of infant burials. Studies on the taphonomy (Martínez López 2009) and mortuary practices (Cabrera 2008) at the site have more or less eliminated the likelihood that the prevalent infant burials are due to either differential burial or preservation, although sample bias due to missing adult skeletons has been proposed (Roksandic, pers. comm.). Even so, this potential of high infant mortality could have implications for the overall health of the population; while it is understood that high infant mortality often reflects higher fertility rather than lower life expectancy (Sattenspiel and Harpending 1983), it is also understood to be a marker for poor population health (Halcrow et al. 2008; Reidpath and Allotey 2003). In brief, however, the paleodemography of the Canimar population is an aspect that has yet to be investigated in full, though it will have serious implications for this study on the prevalence of spina bifida and its correlation with the nutritional status of the population

in terms of life expectancy.

Canimar Abajo: Diet

Stable Isotope Analysis

The question of diet and subsistence at Canimar Abajo has been addressed using zooarchaeological, starch, and stable isotope analyses. Stable isotope analysis of human bone tissue relies on the principle that individual variation in $\delta^{13}C$ and $\delta^{15}N$ levels is due to the consumed diet of plants and animals; bone collagen isotopes represent the protein part of the diet (Schwarcz and Schoeninger 2011). For carbon isotopes, dietary differences depend on the photosynthetic pathways used by the consumed plants. Paleodietary studies distinguish between C_3 (Calvin cycle; most plants in temperate, tropical, and subarctic environments) and C_4 (Hatch-Slack pathway; includes maize, millet, sugar cane) plants (Katzenberg 2000; Schwarcz and Schoeninger 2011). $\delta^{13}C$ values average -25‰ for C_3 plants and -12‰ for C_4 plants; these values tend to be +5% after incorporation into bone collagen (Katzenberg 2000; Schwarcz and Schoeninger 2011). Although the range of $\delta^{13}C$ for each plant type is discrete, these can be obscured by a mixed diet, especially when animal carbon contributions are considered. Marine and terrestrial animal protein sources will also contribute to the human collagen $\delta^{13}C$ levels; marine organisms use carbonate as their carbon source, which has a 0‰ $\delta^{13}C$ value, while terrestrial organisms use atmospheric CO_2 with a $\delta^{13}C$ value of -7‰ (Katzenberg 2000). Freshwater fish will also confuse the picture, since they have a large range of $\delta^{13}C$ values.

Nitrogen isotopes distinguish between trophic level, namely the means by which

an organism acquires its nitrogen. Legumes use their symbiotic relationship with *Rhizobium* bacteria to fix nitrogen from the atmosphere, while other plants use decomposed organics in the soil to acquire it (Katzenberg 2000). As a result, legumes have $\delta^{15}N$ values closer to atmospheric levels (-5 to 0‰), while other plants have higher levels (2-6‰) (Schwarcz and Schoeninger 2011). Herbivores that eat these plants will have $\delta^{15}N$ values of approximately 3‰ higher than the plants they consume, and carnivores eating herbivores have an even higher amount; marine animals have higher $\delta^{15}N$ values than terrestrial animals, which will also be reflected in human values if these are a dietary component (Schwarcz and Schoeninger 2011). Humans eating all of these organisms will have the highest $\delta^{15}N$ values, though if legumes form a large proportion of the diet, these will lower the $\delta^{15}N$ levels (Katzenberg 2000).

Given the omnivorous diet of humans, it can be quite difficult, if not impossible, to extricate all dietary information from stable isotope data. Nonetheless, it is generally possible to determine whether C_4 plants are an important part of the diet, and whether marine or terrestrial foods are favoured (Schwarcz and Schoeninger 2011).

Stable isotope analysis of human bone collagen has only recently been applied to a subset of samples from Canimar Abajo (Buhay et al. In press). Of the samples studied so far, most have been found to be viable samples for stable isotope analysis, with C: N ratios between the recommended 2.9-3.6, meaning that diagenetic alteration to the collagen is not significant (DeNiro 1985).

Based on the set of samples analyzed thus far, and using the food web constructed for the foods thought to be available to prehistoric Caribbean populations (Keegan and DeNiro 1988; Williams et al. 2009), the diet at Canimar Abajo overall appears to be quite

mixed (Figure 2.2). Residents at Canimar would have had a variety of protein sources available to them, from marine, riverine, and terrestrial environments. This mixed diet means that it is therefore not possible to distinguish between the relative contribution of C_3 and C_4 plants since these are obscured by the marine diet of the people (Buhay et al. In press). It is clear, however, that while marine foods may have made up the majority of the sample's diet, there was considerable plant food intake, particularly of legumes, observable by stable isotope analysis. The significant terrestrial plant intake of the Canimar population is surprising, since it has long been thought that marine molluscs were the most important protein source, while plant food collection may have been more opportunistic (Rodríguez 2007; Rouse 1992; Tabío and Rey 1979). These assumptions do not seem to be borne out by the stable isotope results.

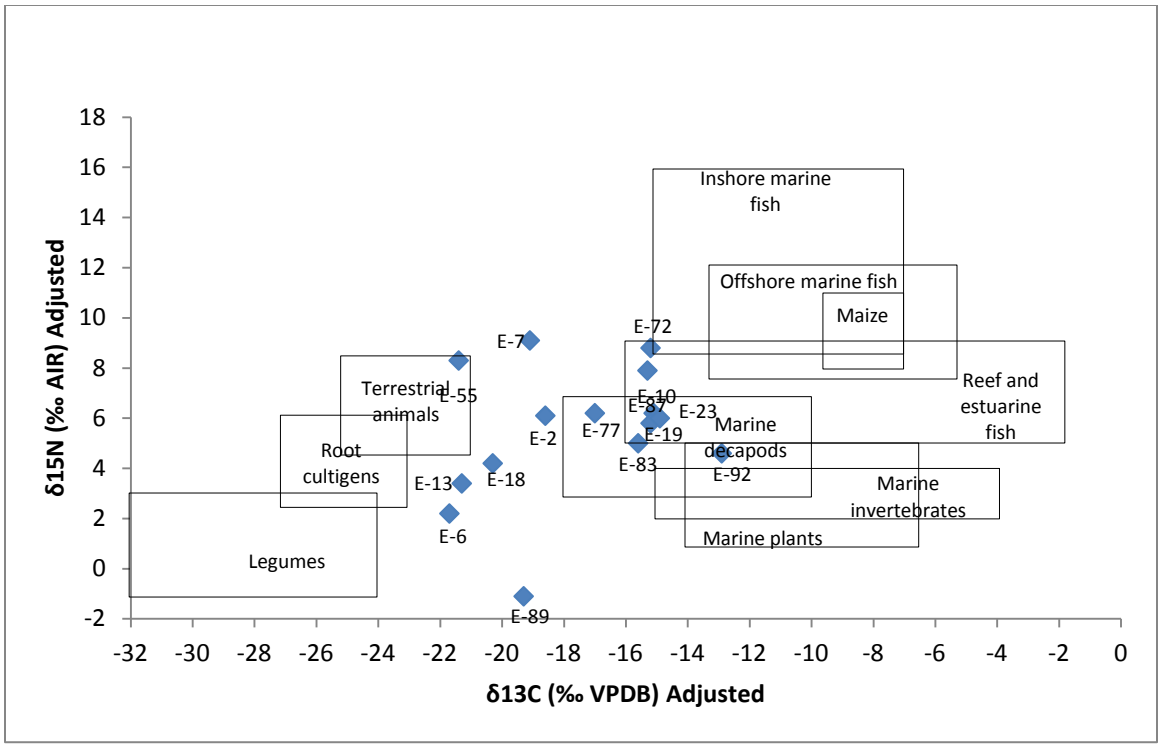

Figure 2.2: Adjusted carbon (±0.2 ‰) and nitrogen (±0.3 ‰) isotopic compositions of the Canimar Abajo remains. The food web was assembled using root cultigens (from Belize, Bahamas, Cuba) and terrestrial animals (hutia, iguana, riverine reptile, land crab from Jamaica, Turks and Cacaos, Bahamas, Belize, Cuba) isotopic values from Williams et al. (2009) and Keegan and DeNiro (1988). Adapted from Buhay et al. (In press), and including new data.

Zooarchaeology

A full zooarchaeological analysis has not been completed in terms of evidence of butchery and cooking (O'Connor 2008). Moreover, we do not necessarily expect food remains in a cemetery. Still, the distribution of faunal remains at the site allows us to understand the type of animals available to the Canimar population.

Excavations have revealed that marine molluscs are most prevalent at the site, followed by fish and mammals (Arredondo 2004; González Herrera et al. 2005). Crustaceans, reptiles and birds are sparsely represented, although land crabs (*Cardisoma*) are an abundant exception. This pattern in the proportion of various faunal remains is

common in similar sites across Cuba (Rodríguez and Medina 2004).

Isognomon alatus, a species of bivalve mollusc, is the most abundant remain. The pattern of variation in shell size of these bivalves, along with the oyster species *Crassostrea rhizophorae*, is likely indicative of the collection of this species by humans (Arredondo 2004). Moreover, sea snail shells of the Muricidae family appear to have been intentionally broken in such a way as to allow for the removal of the meat (González Herrera et al. 2005). Similarly, two genera of terrestrial gastropods, *Zachryzia* and *Liguus*, have also been recovered, although they are relatively rare, and also show signs of intentional breakage (Arredondo 2004). Unlike other Cuban sites with a high prevalence of marine molluscs, there is a lack of remains from *Cittarium pica*, a large edible sea snail. It has been postulated that this is due to its particular reef habitat, which is not common in the Canimar area and is not necessarily worth investigating given the abundance of food sources in the sandy river area (González Herrera et al. 2005).

Fish remains (Class Osteichthyes) are also prevalent throughout the excavated area (González Herrera et al. 2005). Though these remains are difficult to identify to the species or even the genus level, the orders Tetraodontiformes, Mugiliformes, Perciformes, and Anguilliformes are present, representing a diversity of marine habitats and necessitating a range of capture strategies (Arredondo 2004; González Herrera et al. 2005).

Mammal remains are less abundant, but are still present throughout the site in relatively high numbers. These remains are mainly of rodent species, namely *Rattus rattus*, four species of hutia (Family Capromyidae: *Capromys pilorides, Capromys minimus, Mysateles prehensilis* and *Geocapromys columbianus*), and *Boromys offella*

(Arredondo 2004). Of these, *G. columbianus* and *B. offella* are now extinct, and are known to have been important in other pre-Columbian sites as food sources (Rodríguez and Medina 2004). Also included in the mammalian remains are two species of shrew (*Solenodon cubanus* and *Nesophontes micrus*, the latter now extinct) (González Herrera et al. 2005). It is important to note that many of these mammalian remains are represented only by single bones or fragments; it is therefore unlikely that these were actively hunted as food sources, but rather that acquisition was more casual and opportunistic (González Herrera et al. 2005).

Overall, the zooarchaeological analysis would indicate that molluscs, marine molluscs in particular, formed the largest proportion of the Canimar Abajo diet. Of course, this type of dietary analysis does not consider the contribution of plants, for which paleobotanical analysis is required.

Starch Grain Analysis

Starch grains are abundant in various plant organs (roots, seeds, tubers) and are often preserved in residues on archaeological artefacts (Piperno and Holst 1998). Once isolated, these starch grains can be compared to reference collections, since different plants produce starch grains of different forms that allow for species identification.

Starch grain analysis on Canimar Abajo material was initially performed on two stone grinding tools, one from the lower cemetery (1.70 m) and the other from the habitation level (0.90 m). Starch grains from maize (*Zea mays*), bean (*Canavalia sp*), sweet potato (*Ipomoea batatas*), an unidentified legume (Leguminosae – *Phaseolus*?),

and *Zamia* (a cycad plant) were identified on the older tool, while starch grains from maize, bean (Leguminosae), sweet potato, taro (*Xanthosoma sp*), and yam (*Dioscorea sp*) were uncovered on the other (Rodríguez 2007).

More recently, starch grain analysis has also been performed from dental plaque of four individuals from the Canimar Abajo site (Chinique de Armas 2009). Of these four individuals, three were from the later cemetery and one was from the earlier one. Represented in this sample were both hard and soft endosperm varieties of maize, the Fabaceae family (beans), including *Canavalia spp.*, *Phaseolus vulgaris* and other *Phaseolus* species, two species of Leguminosae, *Ipomoea batatas*, and *Zamia sp.* (Chinique de Armas 2009). Of particular importance is the presence of starch grains from *Zea mays*, *Canavalia*, and *Phaseolus sp.* in the dental plaque of skeleton E-19 (1217-1012 BC); this is the oldest archaeological evidence of maize in Cuba, and is among the oldest samples in the Caribbean, dating to well before the later Ostionoid migrations (Berman and Pearsall 2008; Roksandic et al. In prep).

While the presence of all of these varieties of plant is interesting, *Zamia* in particular implies rather sophisticated processing; in order to be edible and non-toxic, the stem must be steeped or fermented, a process that is recorded ethnohistorically by Las Casas (Sturtevant 1968). Similarly, *Canavalia spp.* must be cooked in a saline solution to remove lectin concanavalin A, a toxin (Zarrillo et al. 2008). Starch grains from *Zea mays*, including soft endosperm (flour) maize, also showed evidence of cooking in three of the four individuals (E10, E13, E19), namely gelatinization, indicative of cooking in water (Zarrillo et al. 2008).

The presence of maize and legumes early in the occupation of Canimar Abajo

would seem to indicate knowledge of the plant growth cycle, such as water requirements, if not deliberate cultivation by the population (Chinique de Armas 2009). It also acts as evidence against a strictly nomadic lifestyle; the growth cycle of some maize varieties can be up to three months, while for *Ipomoea batatas*, it can reach up to a year (Chinique de Armas 2009).

Overall, what can be concluded from the paleodietary studies conducted at Canimar Abajo thus far is that the diet of the population is much more diverse than the term '*preagroalfarero*' would imply. Though full-scale agriculture is not suggested, plants such as maize and legumes may have been deliberately cultivated, while molluscs, fish, and potentially some terrestrial fauna, made up the rest of the diet, if not the majority. This diverse diet has implications for the health of the population; in particular, the availability of dietary folate has implications for the incidence of spina bifida, which will be discussed in detail.

Canimar Abajo: Paleopathology

A study of macroscopic paleopathology in the adult Canimar population was conducted in 2009. Of the sample excavated at the time, 95 skeletal elements were observed to have pathologies. After examining the sample for all skeletally-observable indicators, 15 different types of pathologies were uncovered, eight of which are oral pathologies (Campillo Álvarez 2009). The remaining pathologies include fractures, cribra orbitalia (a common skeletal response to a variety of conditions; Walker et al. 2009), infectious pathologies such as periostitis and osteomyelitis, degenerative pathologies such as exostoses, vertebral osteophytes, and spina bifida, a genetic disease.

The prevalence of oral pathologies can be considered a reflection of diet. Dental attrition, which was observed in all 38 individuals whose teeth were examined in 2009, is indicative of hard objects in the diet, particularly since most of the wear was in the occlusal (biting surface) (Campillo Álvarez 2009; Roberts and Manchester 2005). These objects could include shell fragments or sand which may have been incorporated during food processing (Campillo Álvarez 2009). Dental caries were observed in 25 of the 38 individuals, often with multiple teeth affected and occurring primarily in the occlusal and labial surfaces (Campillo Álvarez 2009). Although the etiology of caries is complex, the presence of caries is considered to be an indicator for carbohydrate intake (Roberts and Manchester 2005), and the high prevalence again supports the evidence of at least a balanced diet, rather than a high-protein, low carbohydrate diet that would be expected of a hunter-gatherer population (Lieverse et al. 2007). There is, however, also a high prevalence of dental calculus, which can be associated with, but is not exclusive to, high protein in the diet (Lieverse 1999). Enamel hypoplasia was observed in 14 individuals of the 38 whose teeth were examined; though this defect is a general stress indicator, it is often associated with periods of nutritional deficiency during enamel development (Roberts and Manchester 2005). Unfortunately, there has not yet been a quantitative analysis on the distribution of these pathologies in the population, making any in-depth interpretation impossible.

Overall, however, there seems to be a relatively low amount of infection in the population, and certainly there have been no pathognomic indicators described for any particular infectious disease. However, museum records of the Canimar collection refer to the removal of approximately 100 adult individuals who may have had skeletal

evidence of treponemal disease. Preliminary molecular paleopathological work done on a sample population has supported the lack of infectious disease; though more research is necessary in this area, screening for a number of bacterial, parasitic, and viral diseases has not yet uncovered any positive results.

Spina Bifida at Canimar Abajo

Of particular importance, however, is the identification of spina bifida in the population. Based on macroscopic paleopathological evidence in preserved sacra, it was determined that spina bifida was present in high proportions in Canimar Abajo. Of the 21 preserved sacra from the site of Canimar Abajo, 14 demonstrated spina bifida, 67% of the sample population (Campillo Álvarez 2009). In this instance, spina bifida is defined as the non-fusion of at least two sacral vertebrae, and the 14 pathological sacra demonstrated a range of sacral involvement, from only a few unfused sacral vertebrae to a completely unfused sacrum (Campillo Álvarez 2009; Rivero de la Calle 1987). Though obviously not directly comparable, especially since it is unknown how the modern cases of spina bifida were diagnosed or defined, of 273,346 births in Cuba from 1993 to 1998, there were 70 cases of spina bifida (birth prevalence x 10,000 of 2.56) (International Centre for Birth Defects 2003). For this period, the highest birth prevalence worldwide (x 10,000) was 15.25, observed in Mexico (International Centre for Birth Defects 2003). Clearly, the prevalence of spina bifida in the Canimar Abajo population is very high.

All of the preserved sacra at Canimar were from adult individuals, and given that there is no other skeletal evidence, such as diminished lower limbs or hydrocephaly,

demonstrative of the more severe forms of spina bifida, the observed cases from Canimar can only be diagnosed as likely having spina bifida occulta, though it is not possible to completely eliminate the possibility of other forms of neural tube defects (Kumar and Tubbs 2011). This diagnosis likely means that the individuals did not suffer any serious symptoms. Such high levels of occurrence can be interpreted as being due to biological isolation, if not endogamous breeding (Kumar and Tubbs 2011; Rivero de la Calle 1987).

However, some problems have been noted regarding the diagnosis of spina bifida in skeletal remains. In particular, the non-fusion of the lower sacral vertebral neural arches (S4 and S5) has been noted as being due to normal population variation, and may not reflect any pathological condition (Kumar and Tubbs 2011). Molecular analyses may help to distinguish normal variation from pathology, particularly in the absence of other skeletal indicators of disease, though it may still not be possible to distinguish the degree of suffering in the absence of other indicators.

This unusually high prevalence of spina bifida at Canimar Abajo makes it an interesting population to study from a paleoepidemiological perspective. Furthermore, spina bifida is multi-factorial genetic disease, meaning that genetic mutations and environmental factors are both important to its phenotypic presentation in the individual. These factors will be examined in terms of how they contributed to the population's risk of spina bifida.

Spina Bifida

Neural tube defects are common types of birth defects that involve the malformation of the central nervous system. These defects occur during the neurulation

process of embryogenesis. This is a complex process involving multiple molecular regulators, making it prone to defects (Bronner-Fraser 2006). There are two main stages of neurulation: problems in the primary neurulation stage result in open defects that expose the unfused neuroepithelium (e.g. spina bifida aperta/ myelomeningocele, anencephaly), while problems during secondary neurulation result in closed defects (e.g. meningocele, spina bifida occulta), which allows the neural tube itself to close, but defects in the development of the axial mesoderm result in abnormal vertebral arch formation, particularly in the lumbar/sacral region (Greene and Copp 2006).

Spina bifida (split spine), characterized by the midline separation of vertebrae due to disturbances to vertebral arch development (Graham and Parsch 2009), is a blanket term for various forms of these congenital neural tube defects; it is a complex, multi-factorial disease ranging in severity, from asymptomatic to severe clinical presentations. Despite this range, all forms of spina bifida seem to have the same genetic basis and risk factors, demonstrating multi-factorial inheritance where genes combine with environmental factors to produce a phenotypic outcome (Graham and Parsch 2009; Kumar and Tubbs 2011; Rizk and Iskandar 2010).

Background and Classification

There have been a number of proposed systems for organizing the forms of spina bifida and other neural tube defects, resulting in a lot of confusion in the literature (Table 2.2).

Table 2.2: Neural tube defect classification systems. Compiled from Greene and Copp (2006), Moore (2006), Northrup and Volcik (2000), and Graham and Parsch (2009).

Classification system	Types	Conditions
Open (primary neurulation) vs. Closed neural tube defect (secondary neurulation)	Open	• Spina bifida aperta (aka myelomeningocele) • Anencephaly
	Closed	• Spina bifida occulta • Spina bifida cystica (aka meningocele) • Encephalocele
Neural tube defect vs. Spinal dysraphism	Neural tube defect	• Anencephaly • Spina bifida cystica: - Myelomeningocele - Myelocele - Meningocele • Encephalocele • Craniorachischisis • Iniencephaly
	Spinal dysraphism	• Spina bifida occulta • Spina lipomas • Tight filum terminale
Traditional: Spina bifida cystica vs. Spina bifida occulta (other defects classified separately)	Spina bifida cystica	• Myelomeningocele • Myelocele • Meningocele
	Spina bifida occulta	Any unfused vertebral arch without soft tissue involvement

The discrepancies between these various systems are based mainly on whether that researcher considers the various forms of spinal defect as a spectrum with the same or related underlying causes (open vs. closed) or as separate entities altogether (NTD vs. dysraphism). Often, researchers use a hybrid of these classification systems, depending on the design of their study. Classifying neural tube defects can be difficult even in the living individual since the criteria for diagnosing a form is not very well-established

(Kirby 2006). In archaeological contexts, researchers generally use the traditional system of spina bifida cystica (open) and spina bifida occulta (closed or hidden), since this is generally all that can be distinguished or surmised (Kumar and Tubbs 2011).

Spina bifida occulta (SBO) is one form that is particularly contentious. It is sometimes alternately classified as a spinal dysraphism, which refers to those spinal conditions that are viewed as being less detrimental than the more severe NTDs (Graham and Parsch 2009; Moore 2006). Indeed, 10-30% of the general population has some degree of spinal dysraphism, including SBO, with or without any other symptoms (Graham and Parsch 2009). This has led many medical practitioners and researchers to view SBO as being an abnormality of no consequence; however, SBO is associated with a number of conditions, including orthopedic deformities (Dahl and Ahlsten 2010), urodynamic problems (Ritchey et al. 1994; Yuan et al. 2008), posterior disc herniations (Avrahami et al. 1994), and lower back pain (Eubanks and Cheruvu 2009; Xeller et al. 2000). Furthermore, SBO that is asymptomatic at birth or in childhood, may still demonstrate symptoms later in life (Spacca and Buxton 2008).

Ultimately, many researchers now classify spina bifida occulta as one of many neural tube defects with the same genetic and environmental risk factors (Blom et al. 2006; Greene and Copp 2006; Johanning et al. 2000; Northrup and Volcik 2000; van der Put et al. 2001). In this study, I choose to follow the system of open (primary neurulation) and closed (secondary neurulation), since this system acknowledges the spectrum of neural tube defects, based on differences in embryonic development, rather than lethality or any preconceived notion of the seriousness of certain conditions over others (Greene and Copp 2006). However, I must also operate within the confines of the archaeological

evidence, which classifies the Canimar Abajo samples as demonstrating spina bifida occulta, which is only one form of closed neural tube defect.

SBO, under this classification system, is a defect of the lumbo-sacral region of the spine that results from defects during the embryonic development of the axial mesoderm (Greene and Copp 2006). SBO is characterized by unfused vertebral arches in this region, but with a skin covering that may or may not have any apparent signs, such as hairy tufts, dimples, or lipomas (Graham and Parsch 2009). Hydrocephaly, clubfoot, facial clefting, cardiac defects, limb reduction, renal abnormalities and genitourinary malformations are all conditions that are associated with SBO and other neural tube defects (Seaver and Stevenson 2006). Alternately, it could be entirely asymptomatic, or may involve neurological deterioration over time. Many of these symptoms are not readily visible in the archaeological record.

In sum, by considering NTDs as a spectrum, where spina bifida represents only one portion, the very term 'spina bifida' may be considered outdated and uninformative. However, it does represent a commonly used term in both the medical and archaeological literature. Moreover, it directly translates to 'split spine,' which is exactly what is observed archaeologically. Therefore, the term spina bifida will be used here, with the understanding that this does not fully encompass the range of NTDs.

Genetics of Spina Bifida and the Homocysteine-Folate Metabolic Pathway

Although many risk factors have been associated with the occurrence of spina bifida, including socioeconomic status and maternal age, many of these seem to relate to maternal folate status pre- and peri-conception, and to an underlying genetic

predisposition (Au et al. 2010). As early as the 1960s, the link between NTDs and folate was tentatively identified, since mothers of NTD-infants had high levels of formiminoglutamic acid in their urine, indicating impaired folate metabolism (Beaudin and Stover 2007). However, it was not until the early 1990s that the association between spina bifida and maternal folate status was established by epidemiological studies, resulting in the folic acid fortification of a variety of food products (Au et al. 2010; Watkins 1998). Now, for instance, all enriched cereals are fortified with folic acid in the United States, with the explicit purpose of preventing birth defects (Graham and Parsch 2009; Watkins 1998). This constant supplementation is important, since neural tube development occurs in the first three to four weeks post-conception, before most women are even aware that they are pregnant (Botto et al. 1999). Moreover, synthetic folic acid is much more readily bioavailable than naturally occurring folate (Botto et al. 1999), since it is a monoglutamate rather than a polyglutamate which must be processed before it can be used for cellular processes (van der Put et al. 2001).

Studies of the genetic etiology of spina bifida have proven much more difficult, and most of these studies have been limited to single individuals with myelomeningocele and their parents, an insufficient sample for genomewide linkage studies (Au et al. 2010). There has been a long-standing debate as to whether the same genes are implied in all forms of neural tube defects (Zemirline et al. 2012), but the recurrence risks for parents who have given birth to a child affected with one form does significantly increase the risk of having further children with other forms (Graham and Parsch 2009; Jährig and Wüchner 1985; Lorber and Levick 1967; Vannier et al. 1981). Furthermore, cases of SB in twins have demonstrated that the phenotypic outcome may vary in each twin, even

with the same genetic and intrauterine environment (Spacca and Buxton 2008). Overall, while the heredity of these conditions is complicated, the same genes seem to be involved to varying degrees depending on the condition (Relton et al. 2003).

Many genetic studies have found a correlation between the incidence of spina bifida and single nucleotide polymorphisms in genes associated with folate metabolism, though the mechanism of this association is not understood (van der Put et al. 1995; Wilson et al. 1999). Basically, while maternal folate intake is crucial, it is also necessary to have the genes necessary to transport, retain, and metabolize the folate for cellular processes. To date, there have been few studies of genes involved in folate uptake and retention that have demonstrated a strong correlation to spina bifida or other neural tube defects; it has been postulated that mutations in folate receptor genes may be too lethal to be maintained in a population (Gos and Szpecht-Potocka 2002). Genes associated with folate metabolism have had more success (Au et al. 2010).

Folate metabolism is a complex pathway involved in many biological processes, transporting single carbon units for, "... methionine metabolism, transsulfuration, synthesis of purines and pyrimidines, synthesis of serine/glycine, biomolecule methylation, membrane lipid synthesis, and drug metabolism," (Au et al. 2010) and failure in these processes can result in incomplete neurulation during embryogenesis (Morrison et al. 1998). In particular, folate is required for homocysteine metabolism (Figure 2.3); while homocysteine levels fall during a normal pregnancy, elevated levels of homocysteine have been reported among women who give birth to babies with spina bifida (Hague 2003; Mills et al. 1995; Steegers-Theunissen et al. 1995). Similarly, other dietary factors involved in the folate-homocysteine metabolic pathway, namely

methionine (Shaw et al. 1997) and vitamin B_{12} (Kirke et al. 1993), have all been established as dietary risk factors for neural tube defects, albeit less strongly than folate (Watkins 1998). Together this strengthens the evidence for the genes of the pathway being involved in the outcome of spina bifida.

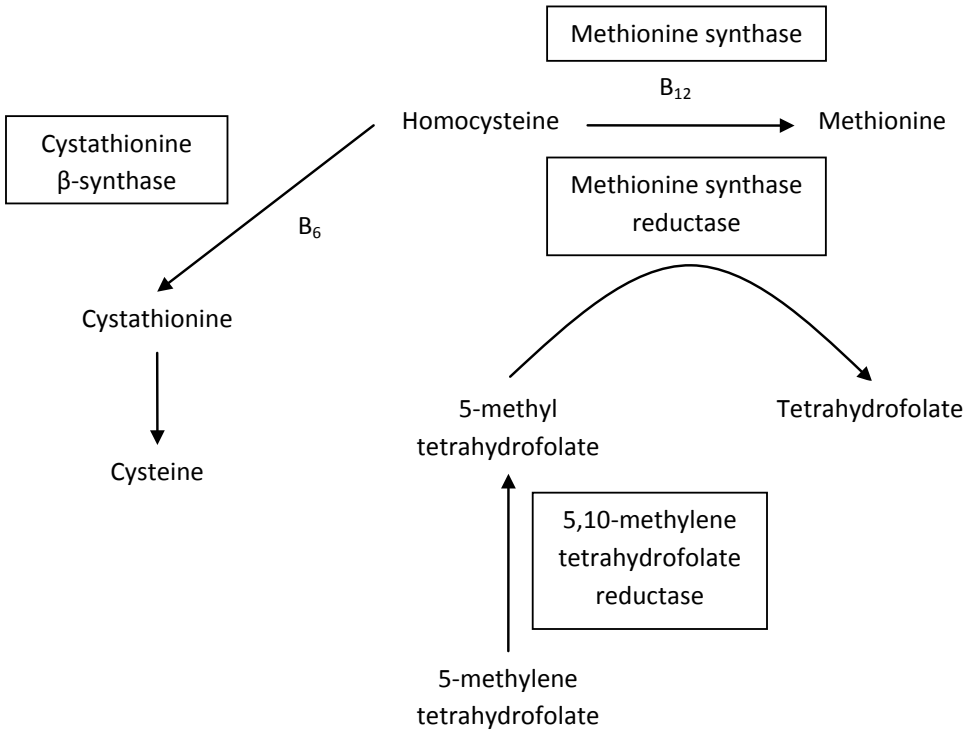

Figure 2.3: Homocysteine metabolism pathways.

Maternal diabetes and hyperglycemia are also known risk factors for spina bifida, and similar to folate, genes involved in glucose metabolism have been demonstrated to be associated with the disease, though not so conclusively as those associated with folate metabolism (Au et al. 2010). However, for archaeological populations, these conditions would be incredibly rare given their diet (Chakravarthy and Booth 2004; Milton 2000), and genes associated with these conditions were not investigated. Other candidate genes

which have been studied include those involved in oxidative stress, DNA repair, gene expression regulators and transcription factors, cell recognition, and planar cell polarity, but none of these genes have been associated so strongly with spina bifida as those involved in folate metabolism (Au et al. 2010).

Single Nucleotide Polymorphisms Associated with Spina Bifida

There are five SNPs in the folate-homocysteine metabolic pathway that have been variously reported to be significantly associated with increased NTD risk, either in the mother or in the child. These include SNPs in the methionine synthase (MS A2756G), methionine synthase reductase (MTRR A66G), and cystathionine β synthase (CBS 844ins68) genes, as well as two SNPs in the 5, 10 methylenetetrahydrofolate reductase gene (MTHFR C677T and A1298C). The frequencies of these SNPs vary between populations, as does their association with NTD risk. Discrepancies in various studies are likely related to the lack of investigation of maternal vs. NTD-child genotype (Doolin et al. 2002), and maternal nutrient status, particularly folate and vitamin B_{12} in the first month of pregnancy before most women are aware that they are pregnant (Rozen 2006). Gene-gene and gene-nutrient interaction studies are now being established, and these new studies are elucidating some of the reasons for these discrepancies (Botto and Mastroiacovo 1998; Christensen et al. 1999; Gos and Szpecht-Potocka 2002; Relton et al. 2004; van der Linden et al. 2006a; Yang et al. 2008). Study design and statistical measures are also very important variables; consideration of paternal, maternal, and individual genotype and nutrient status, transmission disequilibrium tests to test for preferential transmission of risk alleles, and calculation of odds ratios to test for

associations of various genotypes with disease outcome, are all measures that help to clarify the risk of neural tube defects.

At this stage in our understanding, we can say that neural tube defects follow a multigenic threshold model (Gos and Szpecht-Potocka 2002; Relton et al. 2004), in which environmental factors act on, and interact with, genetic factors to create a phenotypic outcome. Genetic predisposition is based on the interaction of a number of genes whose individual contributions to risk might be very small (Ziegler and König 2010). Those genetically predisposed to disease may mitigate their risk by adjusting their environment, and those who are not genetically predisposed can still succumb to disease if their environmental factors are sufficiently extreme. Basically, this combination of genetic and environmental factors must cross a threshold for disease to occur (Gos and Szpecht-Potocka 2002).

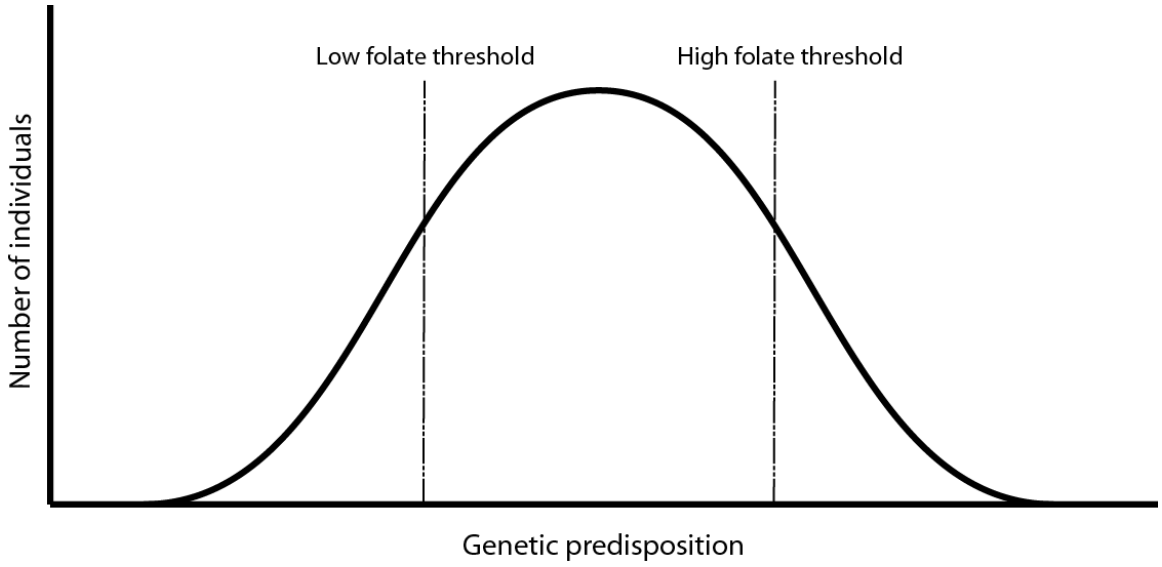

Figure 2.4: Multigenic threshold model in the context of neural tube defects. Genetic predisposition will be normally distributed throughout the population. Environmental factors (e.g. dietary folate availability) will act on this predisposition as thresholds, determining the degree to which an individual will demonstrate disease. Any individual whose predisposition crosses the appropriate threshold will demonstrate a defect.

This model, newly applied to the context of neural tube defect research, helps to explain many of the conflicting results found in earlier SNP studies. These conflicts are now being rectified, while meta-analyses are beginning to standardize the statistical approaches taken to data sets, particularly with the use of odds ratios which directly compare control and case groups (Beaudin and Stover 2007; van der Linden et al. 2006a).

5, 10 Methylenetetrahydrofolate reductase

Most genetic studies of spina bifida have focused on SNPs in the 5, 10 methylenetetrahydrofolate reductase (MTHFR) gene. MTHFR converts 5, 10 methylene tetrahydrofolate to 5 methyltetrahydrofolate (Rozen 2006). Two SNPs in the gene for this enzyme have been studied for their contribution to NTD risk, MTHFR C677T and A1298C.

The first NTD genetic risk factor to be identified was the MTHFR 677 C→T variant, which converts the amino acid sequence from an alanine residue to a valine, resulting in a thermolabile (temperature-sensitive) enzyme (Rozen 2006; Wilson et al. 1999). This SNP affects the catalytic domain of the MTHFR protein, resulting in decreased enzyme activity, increased homocysteine, and low plasma folate levels (De Marco et al. 2002). Studies in different populations and the meta-analyses that have compiled these studies have determined that the 677TT genotype is associated with an increased NTD risk. A maternal 677TT genotype increases risk by 50-70%, while the same genotype in the fetus increases risk by 80-90% (Beaudin and Stover 2007; Botto and Yang 2000; Christensen et al. 1999; Johanning et al. 2000; Relton et al. 2004). The

heterozygous form increases risk to a lesser extent, by 10% for the maternal genotype, and 30% for the fetus genotype (Beaudin and Stover 2007; Johanning et al. 2000; Relton et al. 2004). There is also some evidence that maternal-fetal 677 TT genotype interactions may further increase risk (Christensen et al. 1999). However, this increased fetal susceptibility is generally only observed with low maternal folate status (Beaudin and Stover 2007; Blom et al. 2006; Jacques et al. 1996), or risk increases with low maternal folate (Christensen et al. 1999).

The MTHFR 1298 A→C variant results in a conversion from a glutamate to an alanine residue (Rozen 2006; van der Put et al. 1998), occurring in the regulatory domain of the protein and decreasing enzyme activity in both heterozygous and homozygous genotypes (De Marco et al. 2002; van der Put et al. 1998). This variant is generally not considered as an independent risk for NTD (though see De Marco et al. 2002), though it is often implicated in cases of compound heterozygosity with the 677 SNP (Rozen 2006; van der Put et al. 1998). Compound heterozygosity of both MTHFR SNPs decreases MTHFR activity, increases plasma homocysteine, and decreases plasma folate to a level similar to that of MTHFR 677TT (van der Put et al. 1998). Maternal MTHFR 1298 AC and CC genotypes have also demonstrated an increased fetal NTD risk (De Marco et al. 2002).

These two SNPs are in strong linkage disequilibrium, meaning that co-homozygosity of both MTHFR SNPs is not likely (Rozen 2006). These mutations likely developed independently on different alleles, and are unlikely to crossover given the small chromosomal distance between them (van der Put et al. 1998). Co-heterozygosity, on the other hand, is quite common, and is observed in 15-20% of North American and

European populations (Rozen 2006; van der Put et al. 1998). Compound heterozygosity for both MTHFR SNPs decreases enzyme activity more than either SNP independently (De Marco et al. 2002).

Methionine synthase

Methionine synthase (MS) uses 5 methyltetrahydrofolate to methylate homocysteine to methionine, in a vitamin B_{12}-catalyzed reaction (Rozen 2006). The MS 2756 A→G variant converts a glycine residue to an aspartic acid residue in the active domain of the enzyme (Morrison et al. 1998; Rozen 2006). This gene was implicated early on in NTD-genetic research given its role in folate-homocysteine metabolism (Kirke et al. 1993; Mills et al. 1995; Steegers-Theunissen et al. 1995). Subsequent studies failed to demonstrate an association between this SNP and NTD risk (De Marco et al. 2002; Johanning et al. 2000; Morrison et al. 1998). Other studies have demonstrated the importance of compound effects of this SNP in the fetus with other folate-homocysteine pathway SNPs, particularly MTHFR 677CT or TT (Beaudin and Stover 2007; Morrison et al. 1998; Zhu et al. 2003).

More recently, it has become apparent that maternal genotype may be more important than the child's genotype (Doolin et al. 2002; Zhu et al. 2003). Maternal MS 2756AG/GG, in interaction with MTRR 66GG, increases NTD-risk threefold over those with MS 2756AA and MTRR 66AA/AG (van der Linden et al. 2006a). Mothers with variants in both MS 2756 and MTHFR 677 also have an increased risk of NTD pregnancy (Botto and Yang 2000).

Methionine synthase reductase

Methionine synthase reductase (MTRR) activates MS by regenerating methylated cobalamin cofactor levels from oxidized cob(II)alamin (O'Leary et al. 2005; Rozen 2006), and its gene is a housekeeping gene, meaning that it is expressed constantly at low levels in many tissues (Gos and Szpecht-Potocka 2002). The SNP of interest is at position 66, where an A→G variant converts isoleucine to methionine in the flavin mononucleotide binding domain, which would interfere with the MTRR-MS interaction (O'Leary et al. 2005; Rozen 2006). Unlike the other SNPs investigated, the MTRR SNP is of particular interest since the two nucleotide variants are present in nearly the same frequency, creating confusion as to which nucleotide was the 'wild-type' and which was the 'mutant' (Relton et al. 2004; van der Linden et al. 2006a; Wilson et al. 1999). However, evolutionary models have distinguished the A allele as the ancestral one (Wilson et al. 1999).

Studies of this SNP and NTD risk, like MS, have had mixed results. Though some studies observed increased risk of MTRR 66GG individuals without looking at nutrient status (Guéant-Rodriguez et al. 2003; Zhu et al. 2003), generally, this SNP increases NTD risk when combined with low maternal vitamin B_{12} (Wilson et al. 1999; Zhu et al. 2003) or may only increase risk in the presence of MTHFR 677CT or TT genotypes (Beaudin and Stover 2007; Botto and Yang 2000; Relton et al. 2004). Similarly, maternal MTRR 66GG genotype may be a more important risk factor than child genotype, though the heterozygous genotype does not seem to be a risk factor (Doolin et al. 2002; O'Leary et al. 2005; van der Linden et al. 2006a; Zhu et al. 2003). As in other SNPs, this risk increases with maternal MTHFR 677TT (van der Linden et al. 2006a).

Cystathionine β synthase

Cystathionine β synthase (CBS) is not involved in the homocysteine remethylation pathway, like the other genes investigated here, but is rather a key enzyme in the transsulfuration of homocysteine to cystathionine (Rozen 2006). Deficiencies in the CBS enzyme are the most commonly identified problem in cases of hyperhomocysteinaemia (Hague 2003), and one of these deficiencies is the 68 base pair insert in the coding region of exon 8 (Franco et al. 1998).

As in many of the folate metabolism SNPs, the reported link to NTDs depends on the study in question, again highlighting the importance of good study design. Morrison (1998) and Ramsbottom (1997) and colleagues did not find any increased risk in the presence of the 844 insert, while other studies have identified increased risks for homozygous mutant individuals, particularly as compound effects with MTHFR 677 (Botto and Mastroiacovo 1998; Botto and Yang 2000; Relton et al. 2004; van der Put et al. 2001). NTD individuals homozygous for MTHFR 677TT have also been shown to be significantly more likely to be heterozygous for the CBS insert, which is otherwise a very rare allele in most populations (Franco et al. 1998; Speer et al. 1999). Overall, CBS 844ins68 is still the least understood gene of those investigated in this study in relation to NTDs.

Chapter 3 – Materials and Methods

Samples

A sample of 62 individuals, including both teeth and bone, was selected for this analysis based on their position in the cemetery, relative dating according to burial context, burial practices, preservation, and sex and age indicators (Table 3.1). Such a large sample was deemed necessary, as my research is mainly exploratory and descriptive in nature. Moreover, this sample was also used by another graduate student, Yadira Chinique de Armas, to conduct an in-depth stable isotope analysis, with complete results forthcoming. This sample consists of adults only, with a similar study on children forthcoming. This has the potential consequence that more life-threatening forms of spina bifida, and indeed other diseases, may be missed in individuals who did not live into adulthood. However, limiting research to adults eliminates a number of additional variables that would make interpretation difficult. Only four individuals, all with the mildest classification of spina bifida (type 1, see Table 3.2), were included, although all types except for type 3 have been uncovered from Canimar Abajo (Campillo Álvarez 2009). The more severe cases of spina bifida lacked other skeletal elements for analysis. All of the individuals who presented with spina bifida were found in the later cemetery, though both cemeteries are represented in the sample.

As demonstrated in Table 3.2, even the non-fusion of S4 and S5 alone is considered as a mild form of spina bifida. This has been noted as being a standard non-metric trait representing normal population variation, and indeed this is the case (Graham and Parsch 2009). However, as discussed above, neural tube defects are a spectrum,

ranging from mild to severe demonstrations. Under the neurulation classification system used in this study and in other studies, even the non-fusion of S4 and S5 represents a defect that should be considered among a population with the entire spectrum of vertebral involvement, particularly in a population where more severe vertebral involvement is observed such as at Canimar Abajo (Greene and Copp 2006).

Table 3.1: Sample population from Canimar Abajo used in this study. Macroscopic paleopathology from Campillo Alvarez (2009). *: Material excavated in 2010 has not been fully catalogued, and the ID number is temporary.

ID	Excavation	Quadrant	Depth (m)	Age	Sex	Tissue	Macroscopic paleopathology
E-2	2004	C-118/119	0.4-0.5	Adult (~35)	M	Bone	Type 1 spina bifida in sacrum - unfusion that goes beyond fourth spinous process. Dental wear, caries, calculus, linear enamel hypoplasia
E-4	2004	C-96	0.1-0.2	Young adult (~24)	F?	Bone	
E-5	2004	C-96	0.2-0.3	Adult (27-30)	F?	Bone	
E-5*	2010					Bone	
E-6	2004	C-96/119	0.4-0.5	Young adult (~24)		Bone	
E-7	2004	C-96/119	0.4-0.5	Young adult (22-24)	M?	Bone	Type 1 spina bifida in sacrum - unfusion that goes beyond fourth spinous process.
E-7*	2010					Bone	
E-9	2004	C-118	0.45			Tooth	Dental wear, caries, calculus, linear enamel hypoplasia
E-10	2004	C-118	0.5-0.6	Adult (25-35)	F	Tooth	Type 1 spina bifida in sacrum - unfusion that goes beyond fourth spinous process. Osteomyelitis. Dental wear, caries, calculus, linear enamel hypoplasia
E-10*	2010	C-149/150	1.8-1.9			Bone	
E-11	2005	C-95	0.47-0.56	Adult (>22)	M?	Bone	
E-12	2005	C-95	0.61		M	Bone	
E-12*	2010					Bone	
E-13	2005	C-96	0.46-0.61	Adult (25-30)	M	Bone	Dental wear, caries, calculus, abscesses, linear enamel hypoplasia, antemortem tooth loss. Exostoses on coronoid fossas of humeri.

E-14	2005	C-153	1.6		M	Bone	
E-15	2005	C-153	1.65	Young adult (17-20)		Bone	
E-17	2005	C-153	1.7-1.8	Young adult (15-18)	F	Bone	
E-18	2005	C-119	0.67	Young adult (14-19)	F	Bone	Dental wear, calculus, linear enamel hypoplasia.
E-19	2005	C-118/119	1.7-1.8	Adult (>20)	M	Bone	Dental wear, caries, calculus, linear enamel hypoplasia.
E-20	2005	C-152	1.7-1.8	Adult (25-30)	M	Bone	Dental wear, caries, calculus, hypoplasia.
E-23	2005	C-94	0.6-0.7	Adult (>18)	F	Bone	Dental wear, caries, calculus, linear enamel hypoplasia.
E-24	2005	C-119	0.12	Adult (29-30)	F	Bone	
E-37	2007	C-115	0.85-0.9			Tooth	
E-55		C-119	1.8			Bone	
E-58		C-154	0.25-0.3			Bone	
E-68	2006	C-120	0.3-0.4	Young adult (20-23)	M	Bone	
E-69	2006	C-117	0.37	Adult (>22)	F	Bone	
E-70	2006	C-122	0-0.3	Young adult (22-25)	F?	Bone	
E-71	2006	C-120	0.52	Adult (30-34)		Bone	Dental wear, calculus
E-72	2006	C-120/121	0.52	Adult (25-27)	M	Bone	
E-73	2006	C-120	0.46	Adult (25-30)	F?	Bone	Periostitis
E-74	2006	C-116	0.2-0.3	Adult (35-55)	F	Bone	
E-75	2006	C-117	0.72	Young adult (18-23)	F	Bone	
E-77	2006	C-116	0.22-0.52	Adult (28-30)	M	Bone Tooth	Thoracic and lumbar vertebral osteophytes. Dental wear, caries, calculus, abscess, antemortem tooth loss.
E-78	2006	C-94/117	0.4-0.72	Adult (~25)	M?	Bone	
E-79	2006	C-97	0.37-0.5	Young adult (16-20)	F	Bone Tooth	Dental wear, caries, calculus, hypoplasias, antemortem tooth loss.
E-80	2006	C-98/121	0.7	Adult (~25)	M	Bone	Type 1 spina bifida in sacrum - unfusion that goes beyond fourth spinous process. Severe osteomyelitis.
E-81	2006	C-120	0.66	Adult (~30)		Bone	Sacral exostosis. Dental wear, caries, calculus, antemortem tooth loss.
E-83	2006	C-150/151	0.53-0.59	Adult (~35)	M	Bone Tooth	Lumbar vertebral osteophytes. Fibular exostosis. Dental wear, caries, calculus, antemortem tooth loss.
E-84	2006	C-150/151	0.65	Adult (~25)	F	Bone	
E-85	2006	C-154/155	1.5-1.6	Adult (>35)	F	Tooth	Dental wear, calculus, abscess, antemortem tooth loss.

E-86	2006	C-154/155	1.5-1.6	Young adult (23-25)	F?	Bone	
E-87	2006	C-120/154	1.6-1.7	Adult (~25)	F	Bone	
E-89	2006	C-112	1.54	Young adult (~20)	F?	Tooth	
E-90	2006	C-150	0.2-0.3	Adult (22-30)	M	Bone	Dental wear, calculus, dental cysts, linear enamel hypoplasia, antemortem tooth loss.
E-91	2006	C-150	0.3-0.4	Adult (~30)	M?	Bone	
E-92	2006	C-150/151	0.72	Adult (20-30)	F?	Bone Tooth	Dental wear, caries, calculus, abscess, dental cyst, antemortem tooth loss
E-93	2007	C-114	0-0.6	Adult (33-35)	F	Bone Tooth	
E-94	2007		0-0.6	Adult (35-39)	F	Bone Tooth	
E-95	2007		0-0.6	Adult (45-49)	F	Bone	
E-97	2007		0-0.6	Adult (20-34)	M	Tooth	
E-98	2007		0-0.6	Adult (30-34)	M	Bone Tooth	
MR1	1987-89					Bone	
MR2	1987-89					Bone	
MR3	1987-89					Bone	
MR4	1987-89					Bone	
MR5	1987-89					Bone	

Table 3.2: Diagnosis of spina bifida from the sacrum – classification system used by Campillo Alvarez (2009).

Type 1	Non-fusion of the vertebral spinous processes, beginning from S4 to S5
Type 2	Non-fusion of the vertebral spinous processes, beginning from S3 to S5
Type 3	Non-fusion of the vertebral spinous processes, beginning from S2 to S5
Type 4	Non-fusion of the vertebral spinous processes, beginning from S1 to S5
Type 5	Complete non-fusion of the sacrum

Sample Decontamination and Preparation

Given the nature of molecular work on archaeological samples, all analyses must be carried out under stringent conditions. Consequently, DNA extraction and purification was conducted at the Paleo-DNA laboratory located at Lakehead University, Thunder Bay, Ontario. The Paleo-DNA laboratory is a state of the art facility, and is one of the

few in the country that can conduct ancient DNA research, with separate ventilation systems for different work areas, and separate rooms for all stages of sample preparation, extraction, amplification, and post-amplification. It also provides an excellent venue for student training in molecular techniques, offering an intensive training program in DNA extraction, amplification, sequencing, and analysis. Samples were processed by me, under the supervision of Dr. Carney Matheson and the staff of the Paleo-DNA laboratory.

Many precautions were taken to prevent sample contamination, namely: two pairs of gloves were worn, with frequent changing; PCR preparation and analysis was performed in separate areas; disposable sterile laboratory instruments were used, and equipment was used exclusively in one work area; all areas were cleaned before and after use using 10% sodium hypochlorite and 70% ethanol, and all hoods were UV irradiated for a minimum of 30 minutes; and negative controls (including both extraction and reagent blanks) were included in all experiments.

Prior to DNA extraction, the samples were first subjected to surface decontamination by washing with a 6.0% (w/v) solution of sodium hypochlorite and 70% ethanol. The samples were then sprayed with sterile double distilled water and decontaminated by UV irradiation. The surface of the sample was sanded off, and then the sample was drilled to produce approximately 50 mg aliquots of bone powder (Matheson et al. 2009).

DNA Extraction and Purification – Guanidinium Thiocyanate, Silica Bead, and Size Exclusion Chromatography

DNA was extracted chemically and purified by using the guanidinium thiocyanate

and silica bead purification method modified from the protocol published by Boom and colleagues (1990). This method uses guanidibium thiocyanate (GuSCN - Fisher BP221-1), a chaotropic agent to release DNA from protein, followed by silica beads, which are positively charged and therefore bind the negatively charged DNA to isolate it from the protein solution.

Using this method, 900 µL of GuSCN solution (4M GuSCN, 0.1M Tris-HCl pH 6.4, 0.02M EDTA pH 8.0, 1.3% Triton X-100) and 20 µL of silica bead solution is added to each sample and control tube. These are then vortexed and chilled on ice for one hour, inverting every 15 minutes to resuspend the silica. After the hour, the tubes are centrifuged using a tabletop centrifuge (approximately 10,000 rpm) for one minute. The supernatant is removed and discarded, and 500 µL of working wash buffer (0.01M Tris-HCl pH 7.5, 0.05 M NaCl, 1 mM EDTA, 50% ethanol) is added. This is again vortexed and centrifuged for a minute, and again, the supernatant is removed and discarded. This washing may be repeated if the resulting pellet is discoloured. Otherwise, 150 µL of 100% cold ethanol is added and vortexed to resuspend the pellet. The tubes are cold centrifuged for one minute at 12,000 rpm. The supernatant is removed and discarded, and the pellet is left to dry. When dry, 50-100 µL ddH_2O is added (depending on the size of the pellet) and incubated for one hour at 56°C, mixing at a speed of 700 rpm. After incubation, the tubes are centrifuged for one minute at 12,000 rpm before proceeding.

The samples were further purified by size exclusion chromatography. This technique has been shown to eliminate metal ions (Matheson et al. 2009), humic substances (Matheson et al. 2010) and other environmental inhibitors that impede many DNA polymerases and prevent amplification during PCR reactions, especially of the

degraded DNA of archaeological samples. Size exclusion chromatography non-specifically removes all small inhibiting molecules in the sample as these are trapped in the pores of the column beads, while the target DNA (generally 50-400 base pairs for archaeological material) passes through; though some sample DNA may be lost, the removal of inhibition allows the remaining DNA to be amplified without requiring further dilution of the sample (Matheson et al. 2009). Micro Bio-Spin P-30® columns in sodium citrate buffer (BioRad 732-6202) were used according to the manufacturer's directions, involving saturating the chromatography column with buffer, and passing sample through the column by centrifugation before proceeding to prep the PCR reaction mix.

Multiplex PCR

The purified DNA was then subjected to multiplex polymerase chain reactions (PCR) to examine single nucleotide polymorphisms in five target genes associated with spina bifida (two loci in the 5,10-methylenetetrahydrafolate reductase enzyme, and one each in the methionine synthase, cystathionine-beta-synthase, and methionine synthase reductase genes). At the same time, these samples were also tested for a series of bacterial (cholera, tuberculosis, brucellosis, syphilis, Q-fever, anthrax, typhus, plague, leprosy, whooping cough, diphtheria, typhoid fever), viral (smallpox, chickenpox, hepatitis B, herpes, adenovirus, measles, rubella, poliomyelitis, yellow fever, dengue fever), and parasitic (malaria, schistosomiasis, leishmaniasis, trypanosomiasis, ascariasis) pathogens using six multiplex PCR reaction mixes for further study.

Multiplex PCR, first described in 1988 (Chamberlain et al. 1988) is used to

amplify multiple loci in one PCR reaction, provided that the primer and target sequences do not overlap and PCR conditions are optimized for all targets (Henegariu et al. 1997). This allows for screening for the five single nucleotide polymorphisms discussed in this study.

Primer Design and Optimization

Multiplex PCR must be carefully designed and optimized to ensure that all targets are amplified evenly and to reduce the chance of non-specific PCR products (Henegariu et al. 1997). Primer design (Table 3.3) and initial optimization was completed in another student thesis by Cloutier (2008), while further optimization for the Canimar samples was done by me. Briefly, primers were designed to meet a GC content of 40-60% and a length of a minimum of 20 base pairs (annealing temperature approximately 58°C), and were tested and aligned using a BLAST program (http://www.ncbi.nlm.nih.gov/BLAST) to predict any primer overlap or potential for primer interactions (Henegariu et al. 1997). Primers were purchased from Operon (Alameda, CA, USA).

Table 3.3: Multiplex PCR primers. *: Primers designed by Cloutier (2008). All other primers designed by Barbaux and colleagues (2000).

Gene	Forward Primer (5' → 3')	Primer Type	GeneBank Accession #	Position
CBS-1	TATTGGCCACTCCCATAATAGAA	Forward	AF042836	17974-17996
CBS-2	CGGCTCTGCGAGGATGGACCCTT	Reverse	AF042836	18101-18079
MS-A	TGTTCCCAGCTGTTAGATGAAAATC	Forward	AL359259	43622-43646
MS-B	AGTCACATTAAAAACAAGCAAAA	Reverse	AL359259	43762-43740
MTHFR-0	GAATGTGTCAGCCTCAAAGAAAAG	Reverse	AY338232.1	8790-

				8767
MTHFR-1	AGGGAGCTTTGAGGCTGACCTGAA	Forward	AY338232.1	8694-9717
MTHFR-E	GGAGCTGCTGAAGATGTGGGGGG	Forward	AY338232.1	10608-10630
MTHFR-D	GTAAAGAACAAAGACTTCAAAGAC	Reverse	AY338232.1	10677-10654
MTRR-F1*	GCTACACAGCAGGGACAGGC	Forward	AF121202	4175-4195
MTRR-R1*	GTAACGGCTCTAACCTTATCGG	Reverse	AF121202	4293-4272

These primers are then used in individual PCR reactions to amplify single products, and if this is successful, they are used collectively in a multiplex PCR in equimolar concentrations (Henegariu et al. 1997). Depending on the results of this test, parameters including the number of primer cycles, extension or annealing time, and template DNA, primer, $MgCl_2$, nucleotide, or buffer concentration are modified until all targets are successfully amplified (Henegariu et al. 1997; Markoulatos et al. 2002). Ultimately, the reaction mixture was as displayed in Table 3.4, initially prepared in a master mix and aliquoted to produce 10 µL reaction mixes; the reaction conditions were: initial hot start temperature of 105°C, initial denaturing at 94°C for two minutes, then 45 cycles of denaturing at 94°C for 30 seconds, annealing at 60°C for one minute, and extension at 72°C for one minute. The size of the amplification product from this multiplex PCR is displayed in Table 3.5. During the optimization process, 5 µL of PCR product was mixed with 2 µL 6X loading buffer and electrophoresed on a prepared 3% agarose gel (0.9 g low-melt agarose Fisher BP 1360-100, dissolved in 30 mL 1x TBE buffer) in 1x TBE buffer (10x TBE: 108 g Tris base, 55 g boric acid, 40 mL 0.5 M EDTA at ph 8.0, added to 700 mL ddH_2O, then make up to 1 L with ddH_2O) at 100 volts for one

hour. Products were stained for 20 minutes in an ethidium bromide wash (10 µL 1% ethidium bromide Fisher BP1302-10 in 100 mL 1x TBE buffer = 1 µg/mL) and were visualized under UV light, using a GeneRuler® low range DNA ladder (Fermentas SM1191) to determine amplicon size.

Table 3.4: Initial multiplex PCR reagent concentrations and volumes. Platinum Taq DNA polymerase and PCR buffer from Invitrogen 10966-034.

Reagent	Initial Concentration	Final Concentration	µL per reaction
PCR buffer	10X	1X	1
dNTP	10 mM	0.2	0.2
$MgCl_2$	50 mM	2 mM	0.4
CBS-1	10 µM	0.11 µM	0.11
CBS-2	10 µM	0.11 µM	0.11
MS-A	10 µM	0.1 µM	0.1
MS-B	10 µM	0.1 µM	0.1
MTHFR-0	10 µM	0.032 µM	0.032
MTHFR-1	10 µM	0.032 µM	0.032
MTHFR-E	10 µM	0.072 µM	0.072
MTHFR-D	10 µM	0.072 µM	0.072
MTRR-F1*	10 µM	0.16 µM	0.16
MTRR-R1*	10 µM	0.16 µM	0.16
Platinum®Taq	5 U/ µL	0.5 U	0.1
DNA extract/ negative extract/ ddH_2O	-	-	7.35

Table 3.5: Initial multiplex PCR reaction target amplicon size.

Polymorphism	Primers (F/R)	Amplicon Size (bp)
CBS 844ins68	CBS-1/2	128 (without insertion) 196 (with insertion)
MS A2756G	MS-A/B	141
MTHFR C677T	MTHFR-1/0	97
MTHFR A1298C	MTHFR-E/D	73
MTRR A66G	MTRR-F1/R1	119

Hemi-Nested PCR

Following the initial multiplex PCR, an additional hemi-nested PCR step is added to increase the sensitivity and amplification of the PCR product, often necessary when using minute amounts of archaeological DNA, and to increase the specificity of these products. Essentially, hemi-nested PCR reactions use the initial PCR product as a template, and uses one of the initial primers and a new second primer (hemi-nest) to amplify smaller fragments of DNA (Mullis et al. 1986). In archaeological samples, this second round of amplification serves to increase the total amount of target DNA sequences; more importantly, it exponentially reduces the probability that the initial primers amplified the wrong DNA sequences, i.e. sequences similar to the actual target (Mullis et al. 1986). Thus, hemi-nested PCR uses one of the initial (outer) primers, used in the first multiplex PCR, and another (inner) primer to target an area within the initial amplicon sequence, producing a smaller, more specific PCR product.

The hemi-nest primers were designed and optimized as in the initial multiplex PCR (Table 3.6). The reaction mixture (Table 3.7) was then prepared in a master mix and aliquoted to produce 25 µL reaction mixes; the reaction conditions were: initial hot start temperature of 105°C, initial denaturing at 94°C for two minutes, then 45 cycles of denaturing at 94°C for 30 seconds, annealing at 58°C for one minute, and extension at 65°C for two minutes. The size of the amplification product from the hemi-nested multiplex PCR is displayed in Table 3.8. During the optimization process, 5 µL of PCR product was mixed with 2 µL 6X loading buffer and electrophoresed on a prepared 3% agarose gel in 1x TBE buffer at 100 volts for one hour. Products were stained for 20 minutes in an ethidium bromide wash and were visualized under UV light, using a

GeneRuler® low range DNA ladder to determine amplicon size.

Table 3.6: Hemi-nested multiplex PCR primers. *: Primers designed by Cloutier (2008). All other primers designed by Barbaux and colleagues (2000).

Gene	Forward Primer (5' → 3')	Primer Type	GeneBank Accession #	Position
CBS-2	CGGCTCTGCGAGGATGGACCCTT	Reverse	AF042836	18101-18079
CBS-3*	GCTTTTGCTGGCCTTGAGCC	Forward	AF042836	18019-18036
MS-A	TGTTCCCAGCTGTTAGATGAAAATC	Forward	AL359259	43622-43646
MS-C*	CAAGCAAAATCTGTTTCTACCACTTAC	Reverse	AL359259	43748-43721
MTHFR-0	GAATGTGTCAGCCTCAAAGAAAAG	Reverse	AY338232.1	8790-8767
MTHFR-2*	GCTGACCTGAAGCACTTGAAGG	Forward	AY338232.1	8707-8738
MTHFR-D	GTAAAGAACAAAGACTTCAAAGAC	Reverse	AY338232.1	10677-10654
MTHFR-F*	TGGGGGGAGGAGCTGACC	Forward	AY338232.1	10624-10641
MTRR-F1*	GCTACACAGCAGGGACAGGC	Forward	AF121202	4175-4195
MTRR-R2*	GCAGAAAATCCATGTACCACAGC	Reverse	AF11202	4258-4236

Table 3.7: Hemi-nested multiplex PCR reagent concentrations and volumes.

Reagent	Initial Concentration	Final Concentration	µL per reaction
PCR buffer	10X	1.6X	4.0
dNTP	10 mM	0.2 mM	0.5
$MgCl_2$	50 mM	1.8 mM	0.9
CBS-2	10 µM	0.1 µM	0.25
CBS-3	10 µM	0.1 µM	0.25
MS-A	10 µM	0.12 µM	0.3
MS-C	10 µM	0.12 µM	0.3
MTHFR-0	10 µM	0.08 µM	0.2
MTHFR-2	10 µM	0.08 µM	0.2
MTHFR-D	10 µM	0.08 µM	0.2
MTHFR-F	10 µM	0.08 µM	0.2
MTRR-F1	10 µM	0.12 µM	0.3
MTRR-R2	10 µM	0.12 µM	0.3

Platinum®Taq	5 U/ µL	1.25 U	0.25
PCR product	-	-	5
ddH_2O	-	-	Brought up to 25

Table 3.8: Hemi-nested multiplex PCR reaction target amplicon size.

Polymorphism	Primers (F/R)	Amplicon Size (bp)
CBS 844ins68	CBS-3/2	83
MS A2756G	MS-A/C	126
MTHFR C677T	MTHFR-2/0	84
MTHFR A1298C	MTHFR-F/D	83
MTRR A66G	MTRR-F1/R2	54

Single Nucleotide Extension - SNaPshot® PCR

In the last stage of analysis, the hemi-nested PCR products were processed using multiplex single nucleotide extension (SNaPshot®) PCR. Single nucleotide extension (SNE), specifically SNaPshot® minisequencing, is a method in which a single fluorescent nucleotide probe is added to the 5' end of each locus sequence, which can then be detected by capillary electrophoresis, allowing for the identification of the nucleotide in that position and therefore the single nucleotide polymorphisms present in that individual for the genes of interest (Palacajornsuk et al. 2009).

Prior to the single nucleotide extension, the hemi-nested PCR product was further purified to remove any remaining nucleotides and primers from the previous PCR reactions. Five units of shrimp alkaline phosphatase (SAP; 1 U/µL; Promega) and six units of DNA exonuclease I (EXO1; 10 U/µL; BioLabs) were added to 15 µL of hemi-nested PCR product and incubated at 37°C for one hour at 300 rpm. The enzymes were then denatured at 80°C for 20 minutes, also at 300 rpm.

Single nucleotide extension was performed using a SNaPshot® multiplex kit (Applied Biosystems 4323159). 4.5 µL of SNaPshot® PCR Ready Reaction Mix, 1 µL of each SNE primer (Table 3.9), and 0.5 µL of purified PCR product were mixed and cycled 25 times under the following conditions: ten seconds at 96°C, five seconds at 50°C, and 30 seconds at 60°C. One unit of SAP was added to stop the reaction, and the mix was incubated at 37°C for one hour, followed by 80°C for 20 minutes to denature the enzyme.

After single nucleotide extension was complete, 9 µL HI-DI formamide and 0.1 µL of Genescan™ 120 LIZ™ size standard were added to 1 µL of sample, upon which the samples were heated to 95°C for 5 minutes. The samples were electrophoresed using the ABI Prism 3100 Genetic Analyser (Applied Biosystems). The amplicon sizes are outlined in Table 3.9.

Table 3.9: Single nucleotide extension primer sequences, from Cloutier (2008). MTHFR A1298C primer was the only reverse primer used.

Gene	Sequence	Length (bp)	GeneBank Accession #	Position
CBS 844ins68 (α-CBS)	TGCAGATCATTGGGGTGGATC	21	AF042836	18055-18075
MS A2756G (α-MS-F1)	GGAAATCATGGAAGAATATGAAGATATTAGACAGG	35	AL359259	43666-43700
MTHFR C677T (α-MTHFR-F1)	GAAGGTGTCTGCGGGAG	17	AY338232.1	8730-8746
MTHFR A1298C (β-MTHFR-R1)*	GAGGTAAAGAACAAAGACTTCAAAGACACTT	31	AY338232.1	10680-10650
MTRR A66G (α-MTRR (F2))	CAGGCAAAGGCCATCGCAGAAGACAT	26	AF121202	4190-4212

Data Analysis

Electropherogram Analysis

After electrophoresis, the products were analyzed using Peak Scanner™ Software version 1.0 (Applied Biosystems), and the genotypes were as described in Table 3.10. Briefly, each SNE primer is extended by one fluorescently labelled dNTP during single nucleotide extension, and therefore the length of the fragment is known, while each dNTP is of a known color (adenine: green, guanine: blue, thymine: red, cytosine: yellow/black; see Figure 3.1) (Palacajornsuk et al. 2009). Electrophoretic mobility means that each fragment may not appear at exactly the expected length, but it will be in the vicinity; for example, the MTHFR C677T genotype will be at approximately 18 base pairs, but mobility changes it so that it can migrate up to five base pairs (Applied Biosystems 2010). Therefore if there are two peaks in this area, one black (C) and one red (T), the individual is heterozygous for MTHFR 677 (CT). If there is only one peak of either colour, then the individual is a potential homozygote. This analysis was conducted for each individual at each locus.

Table 3.10: Genotype observed on electropherogram. *: Reverse primer used for this gene, so genotype was visualized on electropherogram as described in parantheses, though actual genotype was the one not in parantheses. CBS insert presence or absence was visualized on electropherogram as described in parantheses.

Gene	Wild type Homozygote	Heterozygote	Mutant Homozygote
CBS 844ins68	No insert -/- (CC)	+/- (AC)	Insert +/+ (AA)
MS A2756G	AA	AG	GG
MTHFR C677T	CC	CT	TT
MTHFR A1298C*	AA (TT)	AC (TG)	CC (GG)
MTRR A66G	AA	AG	GG

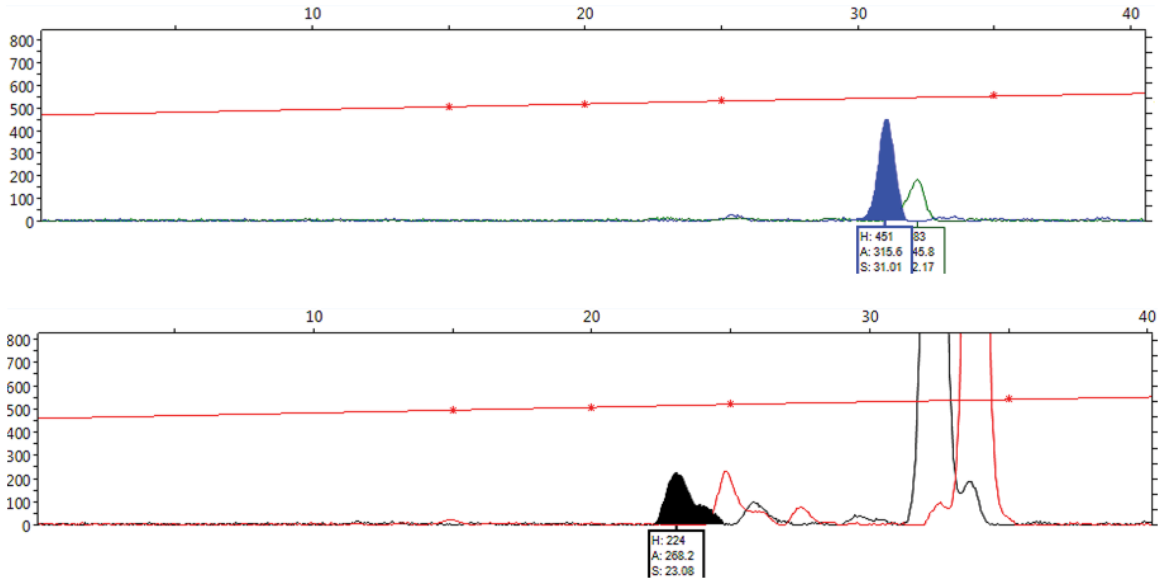

Figure 3.1: Sample electropherogram. Each peak represents a nucleotide at that position. Adenine and guanine are represented by green, and blue, and cytosine and thymine are represented by black and red, respectively. Large extraneous peaks, like the ones between 32 and 35bp are often caused by excess dNTPs or primers that should be removed by purification; in this case, it seemed like the enzymes used to purify were old and less efficient, resulting in a messier electropherogram. This makes interpretation difficult, since alleles in this area will be obscured. Shouldering can occur on peaks if dye was exposed to light during the experimental process, making the peaks small and shouldered as seen in the bottom diagram. Cytosine dyes are especially susceptible (Applied Biosystems 2009).

Quantitative Analysis

Once all of the electropherogram data was collected, statistical analysis of the SNP results could be conducted. Allele and genotype frequencies, fundamental calculations in population genetics, can be determined from the results of the SNP analysis. Allele frequency (F(x)) is calculated by determining the ratio of x allele for 2N alleles from N individuals (Botto and Yang 2000). Genotype frequency is similarly the proportion of a genotype (homozygous wild type, heterozygous, or homozygous mutant) of the total number of genotypes for all individuals. These frequencies were evaluated within the framework of Hardy-Weinberg equilibrium (HWE) for each SNP investigated, which elucidates the relationship between allele frequency and genotype frequency for a

randomly mating population at equilibrium (Gillespie 1998). While it is not expected that the Canimar sample fit the theoretical requirements of a Hardy-Weinberg population (i.e. no mutation, migration, or genetic drift), the deviations from HWE are informative regarding the processes acting on genotype frequency. Ideally, we would also calculate HW from the collection of SNPs, since these are not necessarily independent of each other, but the sample for this study is too small to allow for a reliable analysis of so many variables, as stated in the statistical theory of 'bias-variance tradeoff'. This theory acknowledges the compromise of analysis: to reduce bias, it is necessary to conduct a detailed analysis of all variables; however, this reduces sample size, increasing potential error and variance (Hastie et al. 2009).

Briefly, HWE predicts expected genotype frequencies from observed allele frequencies, where allele A_1 frequency (p) and allele A_2 frequency (q) add to one in a 2-allele system ($p + q = 1$). From here, the equation can be derived to calculate expected genotype frequency ($p^2 + 2pq + q^2 = 1$), where p^2 is the genotype frequency of A_1A_1, $2pq$ is the genotype frequency of A_1A_2, and q^2 is the genotype frequency of A_2A_2. These expected genotype frequencies are then compared to the observed frequencies from the SNP analysis to examine the goodness of fit to, or deviations from, HWE using the chi-square test (x^2) at one degree of freedom (number of genotypes minus number of alleles; significance at 0.05, $x^2 < 3.84$). This is calculated using the equation: $x^2 = (O - E)^2 / E$, where O equals the observed genotype frequency, and E is the expected genotype frequency; this is calculated for each genotype (homozygous wild type, heterozygous, homozygous variant), and the sum of these calculations represents the total chi-square.

To aid in understanding any excess homozygotes or heterozygotes, *f*-statistics can

be applied to these variables, using the equation $F = 1 - (O_{het} / E_{het})$ where 'het' equals the heterozygote genotype. This measure is also called the inbreeding coefficient, since inbreeding results in excess homozygosity, although in fact it reflects any mating system that deviates from completely random mating as defined by Hardy-Weinberg law (Templeton 2006).

In addition to these basic analyses, the frequency of 'risk' or variant alleles was calculated and determined according to sex, since maternal genotype is more often a NTD risk factor than paternal genotype. This frequency was calculated by determining the proportion of variant alleles among total alleles per individual, which was then averaged for all individuals of that sex. A two-tailed, equal variance, Student's t-test was applied to determine whether any differences between sexes were significant (*p*-value significance 0.05). Similar calculations were done with heterozygote and homozygote mutant genotype frequencies.

It has to be borne in mind when interpreting these results that this sample cannot be considered random; all results must be considered only in terms of the total sample population. As outlined by Waldron (2007), archaeological populations are never random, and it is irresponsible and misleading to ever consider it as such; this does not mean that statistical analysis of these samples is not useful, but rather that interpretation of the results of this statistical analysis must be done with caution before generalizing to the living population. Also, the perils of SNP analysis, particularly for ancient samples, means that any homozygote alleles cannot be distinguished from heterozygote signals that may have experienced allelic dropout, the preferential amplification of one allele (see Chapter 5). Nonetheless, all statistical analyses were applied to the samples considering

all potential homozygotes as true homozygotes. All other potential error was interpreted in the context of HWE; deviations from HWE can be informative, though results must be interpreted cautiously.

Chapter 4 - Results

Of the 62 samples, 18 individuals have completed results (Table 4.1). All controls were negative, indicating that no contamination of the samples occurred.

Table 4.1: Single nucleotide polymorphism results per individual

	MTHFR C677T	CBS 844ins68	MTRR A66G	MTHFR A1298C	MS A2756G
E-2	CT	- (C)	AG	A	
E-6	CT	- (C)	AG	A	
E-7 (2010)	CT	- (C)	AG	A	
E-10	CT	- (C)	AG	A	
E-13	CT	- (C)	AG	A	
E-18	CT	- (C)	AG	A	G
E-19	CT	- (C)	AG	A	
E-23	CT	- (C)	AG	A	
E-55					
E-72	CT	- (C)	AG	AC	
E-77	CT	- (C)	AG	A	
E-79	CT	- (C)	AG	AC	
E-83	CT	- (C)	A	A	A
E-87	T		G	A	
E-89	CT	- (C)	AG	C	
E-92	CT	- (C)	AG	AC	
E-98					
MR4	CT		A	A	

Two of the samples did not produce results (E-55 and E-98) and were excluded from analysis. An additional two samples were negative for the CBS signal, while only two were positive for the MS signal. Negative results are likely indicative of DNA degradation, although the difficulties with the MS signal may reflect the need for further method optimization for this locus.

Table 4.2: Hardy-Weinberg analyses from SNP data (x^2 = chi-square test; f = inbreeding coefficient; A_1 = allele 1, A_2 = allele).

	MTHFR C677T	CBS 844ins68	MTRR A66G	MTHFR A1298C	MS A2756G
Observed genotype count					
A_1A_1	0	14	2	12	1
A_1A_2	15	0	13	3	0
A_2A_2	1	0	1	1	1
Observed genotype frequency	**MTHFR C677T**	**CBS 844ins68**	**MTRR A66G**	**MTHFR A1298C**	**MS A2756G**
A_1A_1	0	1	0.125	0.750	0.5
A_1A_2	0.938	0	0.813	0.188	0
A_2A_2	0.063	0	0.063	0.063	0.5
Allele frequency	**MTHFR C677T**	**CBS 844ins68**	**MTRR A66G**	**MTHFR A1298C**	**MS A2756G**
p	0.469	1.000	0.531	0.844	0.500
q	0.531	0	0.469	0.156	0.500
Expected genotype frequency	**MTHFR C677T**	**CBS 844ins68**	**MTRR A66G**	**MTHFR A1298C**	**MS A2756G**
p^2 (A_1A_1)	0.220	1.000	0.282	0.712	0.250
$2pq$ (A_1A_2)	0.498	0	0.498	0.264	0.500
q^2 (A_2A_2)	0.282	0	0.220	0.024	0.250
Expected genotype count	**MTHFR C677T**	**CBS 844ins68**	**MTRR A66G**	**MTHFR A1298C**	**MS A2756G**
A_1A_1	3.52	14.00	4.52	11.39	0.50
A_1A_2	7.97	0	7.97	4.22	1.00
A_2A_2	4.52	0	3.52	0.39	0.50
x^2	**MTHFR C677T**	**CBS 844ins68**	**MTRR A66G**	**MTHFR A1298C**	**MS A2756G**
A_1A_1	3.516	0.000	1.401	0.033	0.500
A_1A_2	6.204	N/A	3.177	0.352	1.000
A_2A_2	2.737	N/A	1.800	0.951	0.500
Total	12.457	N/A	6.378	**1.335**	2.000
f	**MTHFR C677T**	**CBS 844ins68**	**MTRR A66G**	**MTHFR A1298C**	**MS A2756G**
	-0.882	N/A	-0.631	0.289	1

Hardy-Weinberg analyses from the SNP data indicated that only MTHFR A1298C met the requirements of Hardy-Weinberg equilibrium (Table 4.2). MS A2756G also met the goodness-of-fit requirements, but this is unreliable given that it is based on only two results. Chi-square analysis could not be fully applied to CBS 844ins68 since this locus had genotypes that were not represented, namely heterozygous and homozygous variants.

The inbreeding coefficient indicated excess heterozygosity of MTHFR C677T and MTRR A66G, slight excess homozygosity of MTHFR A1298C, and complete homozygosity of MS A2756G (though since only two individuals were represented, it is not possible to interpret this coefficient). The inbreeding coefficient could not be applied to the CBS locus because only homozygous wild types were recovered.

Table 4.3: Risk allele and genotype frequency by sex. *p*-value in bold represents significance (<0.05).

			Genotype frequency	
		Variant allele frequency	**Heterozygote**	**Homozygous variant**
Females	Mean	0.402	0.486	0.160
	Standard deviation	0.146	0.253	0.248
	p-value vs. Males	**0.042**	0.962	0.146
	p-value vs. indeter.	0.077	0.796	0.313
Males	Mean	0.246	0.492	0
	Standard deviation	0.087	0.174	0
	p-value vs. indeter.	0.682	0.682	N/A
Indeterminate	Mean	0.222	0.444	0
	Standard deviation	0.048	0.096	0

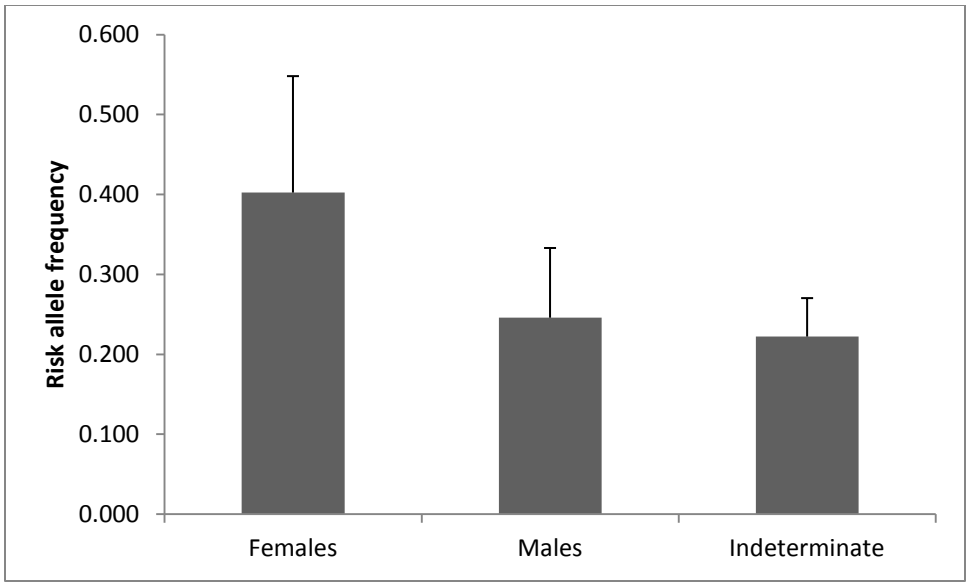

Figure 4.1: Risk allele frequency by sex.

Comparisons of variant allele frequency according to sex revealed significant differences between males and females ($p = 0.042$), where females were more likely to carry variant (aka risk) alleles (Table 4.3, Figure 4.1). However, there was no significant difference in genotype frequency, although only females were observed to have a homozygous variant genotype.

Chapter 5 - Discussion

Multiplex and SNaPshot® PCR

This process of initial multiplex PCR, hemi-nested multiplex PCR and multiplex SNaPshot® minisequencing is relatively novel to bioarchaeology, and more specifically, paleopathology. Though there have been some studies which have successfully analyzed Y-chromosomal and mitochondrial SNPs to study ancient population diversity (Adachi et al. 2011; Bouakaze et al. 2007; Kuch et al. 2007; Lee et al. 2012) or used multiplex PCRs to confirm sex determination (Schmidt et al. 2003), these approaches have not come together in the context of paleopathology. It is therefore necessary to evaluate the potential problems and advantages of this PCR process for ancient specimens.

Potential Problems

Allelic dropout is the most significant issue for SNP analysis, where one allele of a heterozygote does not amplify during the PCR process due to the incidental preferential amplification of one allele, rather than amplification of each allele (Johnson and Haydon 2007). This issue, thought to be the most serious error in SNP analysis (Miller et al. 2002) is understood to occur particularly in samples with low amounts of template DNA, where primers and polymerases, by chance, bind to one allele copy more than the other (Miller et al. 2002). When two different alleles are present in one area of the SNP electropherogram, it is simple to identify a heterozygote. However, true homozygotes are more difficult to identify; these individuals will only show one fluorescent signal, which could indicate a true homozygote, or it could potentially be a heterozygote with allelic dropout.

The problem of allelic dropout, which is observed even in modern samples, is confounded by the problem of DNA degradation in ancient samples. This possibility can only be eliminated with many replications, which may or may not be possible with ancient samples. Alternatively, statistical modeling may be used to determine the likelihood of allelic dropout, but this generally requires previous knowledge of expected allelic frequencies, which is also often difficult to attain when dealing with material from past populations, and represents a circular argument when trying to ascertain if a population differs from others in their allele frequencies. If allelic dropout is not recognized, it could lead to interpretations of inbreeding, since decreased population heterozygosity results in an inbreeding coefficient approaching one (Johnson and Haydon 2007).

The SNP results reflect a high number of heterozygotes for the MTHFR C677T and MTRR A66G genes, so allelic dropout is likely not as much of a concern. However, allelic dropout is much more of a concern for the other genes, particularly CBS 844ins68, where only homozygous wild types were observed, and for MS A2756G, where only two individuals showed any SNP signal. The use of five different loci minimizes the effect of these dropouts on the overall interpretation, but repeated experiments are still necessary to counteract this occurrence and to authenticate data.

False alleles are also a potential problem in SNP analysis (Johnson and Haydon 2007). These generally produce unusual patterns that are easy to recognize in an electropherogram and are generally easily eliminated with replication. False alleles were observed in some of the sample electropherograms at the higher base pair ranges, though not in the range of the SNPs being studied. Though rare, these occurrences are problems

with SNP analysis that must be eliminated with multiple replications.

DNA contamination and degradation are ubiquitous problems with any ancient DNA study. The use of nuclear DNA is often much more difficult because it is much more susceptible to degradation, which again contributes to allelic dropout. However, it has been proven that DNA contamination is much less of an issue compared to mitochondrial DNA (Burger et al. 2007). In terms of DNA degradation, the amplification of very small DNA sequences, made possible by the study of SNPs, increases the success rate of working with nuclear DNA dramatically. Indeed, quantification studies (Alonso et al. 2004) have demonstrated the importance of using sensitive methods, methods of increasing template DNA, and methods of removing inhibitors such as metal ion contaminants, to improve the success of ancient nuclear DNA studies, all of which have been applied in this study.

DNA hydrolytic deamination is one type of damage that is especially problematic for SNP analysis. Hydrolytic deamination results in the misincorporation of nucleotides, generally resulting in C→T and A→G conversions (Gilbert et al. 2007; Pääbo et al. 2004), which could have significant effects on SNP analysis. Miscoding is difficult to identify and discern from PCR amplification error. Cloning and multiple replications are necessary to decrease likelihood of PCR error and to help identify hydrolytic damage.

The last problem with the process of initial, hemi-nest, and SNaPshot® multiplex PCR is the difficulty of interpreting electropherograms. Interpreting electropherograms is generally considered to be a simple matter in SNP analysis, since there are only two possible alleles for each position. However, electrophoretic mobility of smaller fragments can be up to 5 base pairs, which means that overlap with other loci can occur in a

multiplex electropherogram, making interpretation difficult (Applied Biosystems 2010). Although recommendations for multiplex SNP analyses in modern samples are for a four base pair difference in length of different target sequences (Sanchez et al. 2006), this should perhaps be extended to six or seven when working with very small target sequences that are more mobile, and when dealing with ancient samples whose electropherograms may be messier than modern samples. DNA quality, as visualized in electropherograms, also contributes to the problem of interpretation, resulting in non-uniform peak heights.

Overall, more authentication needs to be done to ensure replicability. The recurring theme with these problems is the need for replication to authenticate data. For this study, this has not yet been possible due to cost. Though no contamination was observed in the controls, replications are necessary to confirm the observed results.

Advantages

Despite these difficulties, the process of multiplex initial, nested, and SNaPshot® PCR could be a very valuable tool in bioarchaeology. Multiplexing PCR reactions uses the same amount of sample as a single PCR, minimizing the amount of material required, both in terms of DNA and reagents, and maximizing the DNA available by examining multiple loci at once (or multiple pathogens at once, for pathogenic DNA). This also reduces the likelihood of contamination, all of which are of the utmost importance in ancient DNA studies. The use of multiple loci strengthens any investigation, particularly an investigation of a complex genetic disease such as spina bifida, minimizing the effects

of a single negative result on the overall interpretation.

The stacking of PCRs increases the amount of template, while the inclusion of a hemi-nested PCR increases the specificity. These two important factors increase the success rate of working with ancient nuclear DNA and help to eliminate the likelihood of false positives. Moreover, studying single nucleotide polymorphisms is ideal for ancient DNA studies since it requires only small target sequences, which again helps to increase the chance of successful recovery from a sample with only very fragmentary DNA. SNPs are also ideal to study because, despite the difficulties of electropherogram interpretation discussed above, they are easy to analyze, while their population-dependent frequency and low mutation rates make them good population markers (Ziegler and König 2010). Overall, ancient DNA studies should take advantage of these methods to increase the success of their experiments, and to increase the breadth of their conclusions by maximizing on multiplex and SNaPshot® PCR methods. Furthermore, as demonstrated below, these methods, combined with dietary analysis, allow us to build more detailed paleoepidemiological studies.

Prevalence of SNPs and the Relevance of Hardy-Weinberg

Quantitative analysis of the SNP results includes allele and genotype frequency calculations and application of the Hardy-Weinberg equilibrium and its corollaries, the goodness-of-fit test and the inbreeding coefficient.

It is difficult to discuss allele frequency in terms of what is expected, since the population studies of these alleles are infrequent, and often limited to very specific populations. This limits discussion in this context, since frequencies vary widely between

different populations. More importantly, there is a discrepancy in how populations themselves are defined in medical literature versus the bioarchaeological paradigm. More specifically, medical studies define populations geographically or by how individuals classify themselves in terms of ethnicity or race. In comparison, bioarchaeological studies tend to define their populations more fluidly, relying on any aspects that can distinguish one population from another (e.g. mitochondrial DNA, culture history, skeletal characteristics and biological variation). This discrepancy can make it difficult to draw comparisons between these two bodies of literature. Here, however, modern studies will be referred to, describing the populations as they are delineated in their original studies, relying on the allele and genotype frequencies reported in these studies only as rough guidelines as to what might be considered 'normal', and discussing how the Canimar frequencies compare to these guidelines.

MTHFR C677T is the best-studied SNP of those used in the study, and a comprehensive study on the frequencies of this variant among modern populations in various geographical areas and among various ethnicities indicated a T allele frequency ranging from 12.6-57.0%, with the lowest frequency observed among African-Americans, and the highest frequency observed among Hispanics from Mexico (Wilcken et al. 2003). The frequencies observed at Canimar are within this range, though the T allele is near the higher end of the range at 53.1%.

Conversely, the MTHFR 1298C allele frequency at Canimar is lower than most modern populations, at 15.6%. The MTHFR A1298C SNP is much less studied, but the C allele frequency is between 25 and 40% among the Dutch (van der Put et al. 1998), Jewish Israelis (Friedman et al. 1999), South Indian Tamils (Angeline et al. 2004), and

Italians (De Marco et al. 2002).

The MS A2756G variant is present in lower frequencies elsewhere than at Canimar, with frequencies generally between 10-18% among the British (Skibola et al. 2002), Italians (De Marco et al. 2002), Sicilians (Bosco et al. 2003), a mixed Texan sample (Lin et al. 2004), and the Chinese (Zhang and Dai 2001). At Canimar, the G allele is technically present in a 50% frequency, but this is from a sample of two individuals.

The MTRR A66G variant is present in almost equal frequencies with the A allele, as noted above, so is generally approximately 50%. This has been observed among Sicilians (Bosco et al. 2003), the Dutch (van der Linden et al. 2006a), French Canadians (Wilson et al. 1999), and African-American, Caucasian, and Ashkenazi Jewish groups from Texas, although this frequency was only 28.6% among Hispanics from Texas (Rady et al. 2002). The frequency of each allele at Canimar followed the same pattern, with the A allele at 53.1% and the G allele at 46.9%.

The CBS variant 68 base pair insert is quite rare, with a frequency of only approximately 6-8% among Caucasians in three different studies (Franco et al. 1998; Morrison et al. 1998; Relton et al. 2004), and 4% among Native Americans (Franco et al. 1998). Interestingly, the CBS insert is present in higher frequencies among Africans (22.8%), and was not observed at all among a sample of 40 Japanese individuals (Franco et al. 1998). Similarly, the variant allele was not observed at Canimar, although this could be a function of small sample size.

All of these reported frequencies are for control groups, and are generally higher in NTD-cases and often in NTD mothers as well (Beaudin and Stover 2007; Botto and Yang 2000; Christensen et al. 1999; De Marco et al. 2002; Franco et al. 1998; O'Leary et

al. 2005; Rozen 2006; Speer et al. 1999; van der Linden et al. 2006a; van der Linden et al. 2006b; van der Put et al. 1998; Zhu et al. 2003). All that can be said at this point is that all of the observed allele frequencies in this study are maintainable, but that some of them are high for the mutant alleles, which would be expected of a population with a high prevalence of spina bifida. Notable is the lack of any observed CBS insertion alleles; this is a very low frequency allele, and does not seem to represent a risk factor for spina bifida in this population. The observed frequency of the MS 2756G allele is higher than expected, but is unreliable because it based on the results of only two individuals. MTHFR 677T is present in a very high frequency compared to most modern control populations, while MTHFR 1298C is at the lower end of variant allele frequencies observed in modern controls. MTRR 66G allele frequencies are within the range of modern populations.

Examining the allele frequencies, it is clear that some loci are not in Hardy-Weinberg equilibrium. Only the MTHFR A1298C and MS A2756G SNPs are in HWE, and the latter is represented by only two individuals, while the other three SNPs are not in HWE. However, their deviations from equilibrium may be informative in terms of how the distribution of the genotypes reflects reproductive practices or micro-evolutionary processes. Indeed, given the high prevalence of spina bifida in population, deviation of samples from HWE may reflect increased population disease susceptibility since genotypes are associated with different disease risks (Ziegler and König 2010). While heterozygotes are at an increased risk of disease relative to homozygous wild types, it is still less than homozygous mutants, and heterozygosity may have a selective advantage, or may be due to disassortative mating (choosing a mate with traits opposite to your own)

(Templeton 2006). This is why case-control studies in modern populations are important; it has been suggested that only the control group needs to be in HWE, since disease risk alleles in the case group are expected to deviate from HWE (Ziegler and König 2010). For SNPs, selection can be a very strong HWE-deviating force (Ziegler and König 2010), and this process may be at work in the Canimar sample.

Genetic drift, inbreeding, or chance sampling are all examples of HWE-deviating processes that would decrease heterozygosity of a sample population, as observed by the inbreeding coefficient (Gillespie 1998). Inbreeding has been suggested as a possible explanation for the high prevalence of spina bifida at Canimar Abajo (Rivero de la Calle 1987), but is not supported by the SNP analysis. Decreased heterozygosity which could reflect inbreeding is observed in MS A2756G, but this is not a reliable indicator since it reflects data for only two individuals. The inbreeding coefficient for MTHFR A1298C is approaching equilibrium, but is slightly positive; however, chi-square analysis reveals that it is still within HWE. It was not possible to calculate the inbreeding coefficient for CBS 844ins68 since only homozygous wild types were observed. In contrast, the last two genes, MTHFR C677T and MTRR A66G, have very negative inbreeding coefficients reflecting excess, not decreased heterozygosity. Again, this could be due to selection of a heterozygote trait, or could be a result of disassortative mating (Templeton 2006), though the latter seems unlikely.

In terms of observed genotype frequencies, it is important to remember that carriers of the variant allele, even in the heterozygous condition, have decreased enzymatic activity and increased NTD risk, either for that individual or as a maternal genotype. Thus, while there may be some unknown selective advantage to

heterozygosity, in terms of NTDs, these individuals reflect an increased risk of disease in the population. The fact that so many individuals are heterozygous could reflect the selective advantage of heterozygosity, while those individuals who are homozygous mutants may have perished early in life or may have been spontaneously aborted due to a serious defect. This is not possible to determine for Canimar Abajo, but is an interesting speculation given the high frequency of heterozygotes, particularly for MTHFR 677CT, the gene most often associated with NTD risk.

Paleoepidemiology of Spina Bifida at Canimar Abajo

Genetic Contribution to the Risk of Spina Bifida

The SNP analysis reveals a high prevalence of heterozygotes for the MTHFR C677T (94%) and MTRR A66G (81%) genes. The normal genotype frequency for MTHFR 677CT in other populations is generally between 30-50%, with a frequency of 1-32% for the TT condition, although this increases among those affected by neural tube defects (Scher et al. 2006; Weisberg et al. 1998; Wilcken et al. 2003). It is interesting that the heterozygous condition is maintained in the population at such high frequencies, since it reduces enzymatic activity by 30-40% (Botto and Yang 2000). If this is combined with low dietary folate, this could be a strong contributor to the high prevalence of spina bifida in the Canimar population (Beaudin and Stover 2007). One female individual was observed to be a homozygous mutant for MTHFR 677, which would decrease enzyme activity by 50-60% (Botto and Yang 2000) and increase NTD risk by 80-90%; as a

maternal genotype, this increases NTD risk in offspring by 50-70% (Beaudin and Stover 2007; Relton et al. 2004). All of these risks, however, are dependent on low maternal folate status (Christensen et al. 1999; Jacques et al. 1996).

Likewise, the high frequency of heterozygotes for the MTRR gene is a NTD risk factor, since carriers of the mutant allele in both the homozygous and heterozygous condition have been observed to be at a higher risk of neural tube defects, including spina bifida (Guéant-Rodriguez et al. 2003; Zhu et al. 2003). For the MTRR locus, homozygous mutant genotype frequencies range widely as well, from 10 to 30%, while the range of heterozygotes hovers around 47% in various North American populations (Rady et al. 2002). The effect of this variant is compounded with the MTHFR C677T effects, greatly increasing the risk for spina bifida in this population (Beaudin and Stover 2007; Botto and Yang 2000; Relton et al. 2004). Moreover, maternal genotype for this gene is important for conferring risk; although both sexes had high frequencies of heterozygotes, all females carried a variant allele, and the only homozygous mutant observed in the sample was a female. If these individuals bore children, their offspring would have an even further increased risk of NTD, particularly if they (Relton et al. 2004) or their mothers (van der Linden et al. 2006a) are also carrying the MTHFR 677 variant.

The genotype frequencies for the MTHFR A1298C SNP are within the range of modern North American control populations including African-Americans, Caucasians, and Hispanics (Rady et al. 2002). Therefore, while this in and of itself may not be a risk factor for the population, carriers of the variant allele would be at greater risk when combined with the heterozygous MTHFR C677T condition. Indeed, this variant is often

only observed to be a risk factor as a compound heterozygote with MTHFR C677T. Compound heterozygosity was observed in 19% of the sample of 16 individuals with usable results, slightly higher than the 15% average for modern control populations (Rozen 2006; Weisberg et al. 1998). The maintenance of co-heterozygosity therefore represents a risk for spina bifida at Canimar Abajo.

The results for CBS 844ins68 seem to indicate that this was not a risk factor for spina bifida in this population. The variant insert of this gene was not observed in the population. This is a very rare allele in many populations (Franco et al. 1998; Speer et al. 1999), and its relationship with NTDs is not as well-established as the other genes investigated here.

The results for MS 2756 are not reliable given that it is only reflective of two individuals with a great possibility for allelic dropout. Despite this low success rate, the variant allele was observed in one individual, in a potentially homozygous form. Moreover, this individual was a female, which has implications for maternal risk, particularly in combination with the MTHFR 677 variant (Botto and Yang 2000). Again, for this gene, maternal genotype may be more relevant to NTD risk than the child's genotype (Doolin et al. 2002; Zhu et al. 2003). Though no conclusions can be drawn from this sparse evidence, it can be said that the variant is present in the population.

It is unusual that the two genes with the most demonstrable risk for spina bifida in the population, namely MTHFR C677T and MTRR A66G, did not demonstrate any homozygous mutant individuals. This could be due to a number of factors, including primarily sample size, and also the chance that homozygous mutants, if affected, may not have lived to adulthood. The latter could also contribute to the deviation from HWE

observed, though clearly some must live to adulthood in order for heterozygosity to be maintained in the population, perhaps supported by some selective advantage. Even notwithstanding the lack of individuals homozygous for variant alleles, the overall genetic contribution to risk of spina bifida in the population is considerable, with a high frequency of variant alleles circulating in the population. Moreover, females carry considerably more risk alleles, many of which increase risk even further.

Dietary Contribution to Spina Bifida Risk

Genetics are only one facet of NTD risk; diet is also a strong contributor to the prevalence of spina bifida. For the study of spina bifida, dietary folate becomes an important variable, and it is necessary to evaluate the diet of the Canimar Abajo population in terms of how it contributes to maternal folate status pre- and peri-conception.

Folate, or vitamin B_9 is the form that occurs naturally in food, while folic acid is the synthetic form that is added to fortified foods (National Institutes of Health 2009). Folate is a key nutrient during pregnancy, as it is involved in the production and maintenance of cells, particularly in periods of rapid cell division such as neurulation (Kamen 1997). Although it was first recommended for pregnant women as a prevention for anemia (Chanarin et al. 1968), its connection with the incidence of neural tube defects was also established early on in epidemiological studies (Beaudin and Stover 2007), resulting in widespread folic acid fortification of foods (Au et al. 2010; Watkins 1998). Before this supplementation, which began in the 1990s, naturally occurring folate in

foods was the main source of this vitamin. Folate is a polyglutamate that must be processed before it can be absorbed intestinally and used for cellular processes, while synthetic folic acid is more readily bioavailable (Botto et al. 1999; Sanderson et al. 2003; van der Put et al. 2001). This discrepancy is significant, and folate levels in various food products are now reported as Dietary Folate Equivalents (DFEs), where 1 DFE is equal to one microgram of naturally occurring folate or 0.6 µg of folic acid (National Institutes of Health 2009).

The recommended dietary allowance for folate has been outlined by the National Institute of Health (2009) in the United States (Table 5.1). For infants, there is not enough information to establish a recommended allowance, but adequate intake levels have been established based on the amount consumed by healthy breastfed infants.

Table 5.1: Recommended dietary allowance for folate, in µg per day (National Institutes of Health 2009). Infant values are reported as Adequate Intakes (AI, see text).

Age	Males and Females	Pregnancy	Lactation
0-6 months (AI)	65		
7-12 months (AI)	80		
1-3	150		
4-8	200		
9-13	300		
14-18	400	600	500
19+	400	600	500

The amounts recommended for pregnant women are very high, and this may be difficult to achieve without the supplementation available today. Therefore, it is important to consider the Canimar diet from the perspective of folate sources. It is important to keep in mind that reported folate values are for modern foods, as established by the United States Department of Agriculture (USDA) nutrient database (2011). Therefore, these values

reflect years of selective breeding and other changes that may affect folate values. Nevertheless, these nutrient values serve as a useful baseline for establishing the relative folate contribution of various food sources known to be consumed at Canimar Abajo.

Stable isotope results revealed that the Canimar population took advantage of a variety of protein sources, including marine, riverine, and terrestrial sources. This analysis revealed a mixed diet of plant and animal protein, including legumes (Buhay et al. In press). Zooarchaeological analysis revealed a high prevalence of marine molluscs including *Isognomon alatus* and the oyster species *Crassostrea rhizophorae* (Arredondo 2004; González Herrera et al. 2005). Crab (*Cardisoma*) is also highly represented, as are a variety of fish remains. These are mainly from the orders Tetraodontiformes, Mugiliformes, Perciformes, and Anguilliformes (Arredondo 2004; González Herrera et al. 2005), which includes pufferfish, mullets, eels, tuna, and swordfish. Starch grain analysis, the most direct and specific form of dietary analysis, revealed a variety of plant remains, including maize (*Zea mays*), sweet potato (*Ipomoea batatas*), beans (*Canavalia* and *Phaseolus*), and taro (*Xanthosoma*).

Table 5.2: DFE values for foods consumed at Canimar Abajo (Values from USDA 2011).

	Dietary Folate Equivalent
Phaseolus (e.g. kidney, lima, mung)	62-230 µg/cup
Zea mays (maize)	57-103 µg/cup
Blue crab	50-58 µg/100 g
Assorted molluscs (e.g. oyster, clam)	16-50 µg/ 100 g
Sweet potato	9 µg/ potato
Shrimp (mixed)	24 µg/ 100 g
Perciformes (e.g. tuna, swordfish)	2-5 µg/100 g

When these assorted food sources are considered in terms of folate value (Table

5.2), it becomes clear that all of the foods represented here have relatively low amounts of folate, except perhaps some varieties of bean (*Phaseolus*), which are indeed good sources of folate. However, based on the stable isotope results, it seems that legumes did not form the majority of the diet, though the starch grain analysis reveals that they were eaten in some quantity. Comparing the fish and molluscs observed at Canimar to the USDA nutrient database is particularly difficult because the species represented at Canimar are not ones that tend to be eaten in North America, meaning that folate values are unavailable. Overall, however, these food groups tend to have very low amounts of folate, and this is reflected by the species that we do have folate values for, namely blue crab, assorted molluscs and shrimp, and various species of Perciformes (USDA 2011).

The best natural sources of folate are the leafy vegetables such as spinach (263 μg/ cup), asparagus (243 μg/ cup), and turnip greens (170 μg/ cup), as well as citrus fruits (USDA 2011), most of which were unavailable to the Canimar population. Indeed, it is not known what kinds of leafy vegetables may have been available to the Canimar population, though it is likely that there was some type of edible leafy vegetable available. Unfortunately, starch grains from leaves and stems, rather than in storage organs such as in seeds and tubers, are rarely studied; they are smaller and more difficult to identify, and may not preserve as well as storage starch grains, biasing research against the study of these types of plants (Haslam 2004). This is to say that it is possible that the population of Canimar Abajo was ingesting some leafy vegetable that may have had higher levels of folate, but this is not reflected in the dietary data to date. In the future, phytolith studies may have more success in identifying this portion of the diet, if there was one. Therefore, it would appear that the diet at Canimar Abajo was low in folate,

which would contribute to high prevalence of spina bifida. Diet was a risk factor for neural tube defects in this population.

Interaction of Genetic, Dietary, and Skeletal Evidence of Spina Bifida

With genetic and dietary risk factors discussed independently, it is now important to evaluate how these two domains interact to produce the observed skeletal evidence at Canimar Abajo. It is difficult to discuss individual risk for NTD because of the nature of the disease and its risk factors. Both diet and genetics are clearly risk factors, but both maternal and individual genotypes are in play, while only maternal rather than individual folate status is a risk factor. Moreover, the sample individuals are a mixed sample of those with varying degrees of spina bifida, and do not include fetuses, infants, or young children who may have perished from NTDs, in addition to individuals with no evidence of disease. Therefore, in terms of the interaction of risk factors, it is necessary to discuss risk at the population level. For instance, E-2 and E-10, the two sample individuals with spina bifida, are not genetically distinct from the other individuals in terms of the five investigated genes. However, their maternal genotypes are unknown, as is maternal folate status. In addition, these individuals did live to adulthood, and may not have been aware of their disease or suffered any additional symptoms as a result of their condition. Thus, it is necessary to discuss population risk to disease in terms of what risk alleles are maintained in the population, and what the dietary evidence implies in terms of folate availability, rather than discuss any direct causation.

Briefly, 67% of the total preserved sacra at Canimar showed spina bifida based on

non-fusion of sacral vertebrae in varying degrees. Of the 18 individuals examined thus far, two (E-2, E-10) demonstrate unfused sacral arches. Since diagnosis is based solely on the sacrum, individuals with unpreserved sacra may have also been afflicted with the disease.

The first question to address is whether the low folate from diet is sufficient to account for the high prevalence of disease, and whether the skeletal evidence can attest to folate deficiency. The amounts recommended for pregnant women are very high, and this may be difficult to achieve without folic acid supplementation, such as what is available today. Achieving the recommended amount of folate for pregnant women would be very difficult given the diet of the Canimar population as it is currently understood, and indeed, even achieving the regular recommended intake might have been difficult.

Unfortunately, most of the signs of folate deficiency cannot be observed skeletally, such as digestive disorders, behavioral disorders, and cardiovascular disease (Haslam and Probert 1998). Folate deficiency can also result in anemia and can slow growth rates in children. The most serious known effect of folate deficiency is the increased risk of pregnant women giving birth to infants with neural tube defects like spina bifida (Haslam and Probert 1998). There is no evidence of severe nutritional problems to suggest any type of chronic folate deficiency associated with spina bifida, which can sometimes manifest as porotic hyperostosis due to megaloblastic anemia (Walker et al. 2009). The high prevalence of enamel hypoplasia at Canimar Abajo could be associated with folate deficiency, but the non-specific nature of this indicator renders it impossible to know with any certainty. The lack of evidence, however, does not mean that there was not a folate deficiency; moreover, it is not necessary to have apparent

folate deficiency to increase the risk of NTDs (Beaudin and Stover 2007). Therefore, it is likely that while low maternal folate levels certainly increased the likelihood of NTDs, there was likely a genetic predisposition that these folate levels were acting on.

Next, does the genetic evidence account for the high prevalence of spina bifida in the population? This is a far more difficult question to address and quantify in this context. It is, however, possible to say that given the high skeletal prevalence of spina bifida, and knowing that other populations in the Caribbean of this time likely had similar diets to Canimar without demonstrating this high prevalence, that genetic predisposition is likely at work at Canimar Abajo.

The MTHFR 677T variant is especially likely to increase NTD risk in a low folate environment (Beaudin and Stover 2007; Blom et al. 2006; Christensen et al. 1999; Jacques et al. 1996). This variant is present in high frequencies in the sample (53.1%), particularly as a heterozygote genotype (93.8%). This genotype increases risk by 10% if present in maternal genotype, and 30% in individual genotype, before considering the cumulative effect of the other related genes or folate status (Beaudin and Stover 2007; Blom et al. 2006). This risk would increase even further in the event of the low folate status reflected in the current dietary evidence of Canimar.

Compound heterozygotes with the MTHFR 1298 SNP would have a further increased risk; these individuals have biochemical profiles similar to MTHFR 677TT individuals, though it is unclear whether this also means that co-heterozygosity also increases NTD risk to the same level as the 677TT genotype, which is 50-70% for maternal genotype and 80-90% for individual genotype (Beaudin and Stover 2007; Botto and Yang 2000). Likewise, compound effects with MTHFR 677 are important for the MS

2756 and MTRR 66 variants, and again these effects can be observed with maternal genotypes (Beaudin and Stover 2007; Botto and Yang 2000; Morrison et al. 1998; Relton et al. 2004; Zhu et al. 2003). The only investigated gene that does not seem to have contributed to the risk of NTDs is CBS 844ins68. This variant was not observed in the sample at all, and does not seem to be a risk factor at Canimar Abajo.

Collectively, two possible conclusions can be drawn from the current evidence, depending on the low versus high folate threshold in the population. First, the observed risk alleles in the Canimar population, combined with low folate, could certainly account for the high prevalence of spina bifida observed skeletally. Though this explanation considers the evidence of folate in the population thus far, it does not satisfactorily account for the genotype distribution, namely the high prevalence of heterozygotes and lack of homozygote variants.

Therefore, the second potential conclusion is that the evidence could indicate that there is an unknown folate source being ingested by the population that is not currently apparent by the dietary analysis, and that there is actually a high folate threshold acting on the population's genetic predisposition. By this reasoning, the genetic evidence can be interpreted as such: the lack of homozygous variant adults, with a high proportion of heterozygote adults and a high prevalence of mild types of spina bifida observed skeletally, could indicate that individuals homozygous for variant alleles may have demonstrated more severe forms of spina bifida and perished earlier in life. The high folate acted on those with heterozygous genotypes to diminish the detrimental effects of their less severe genetic predisposition, and the phenotypic outcome was a milder form of spina bifida. These heterozygous individuals were still able to reproduce, and the variant

alleles were maintained in the population.

A study of an infant and juvenile population is necessary to determine whether these individuals were more likely to have homozygous variant genotypes. In addition, folate needs to be measured in plants native to Cuba, which may reveal folate sources not recognized by the current dietary analysis.

In the case of either conclusion, the combination of genetic and dietary studies begins to reveal the underlying heterogeneity of risk and selective mortality, though a larger sample population including infants and juveniles is necessary to complete this picture. Though the genetics of the population reveal very little diversity among those who lived to adulthood, the high prevalence of variant alleles increases the overall risk of the population, which dietary factors would then act on, as described in the multigenic threshold model (Gos and Szpecht-Potocka 2002). Any homozygous mutants for the five single nucleotide polymorphisms investigated here, though not observed in this study, would have had a higher risk of disease and mortality.

In addition, the deviations of the genetic data from Hardy-Weinberg equilibrium hint at behavioural patterns in the population, namely the potential for disassortative mating. It also diminished the likelihood of inbreeding, which had been suggested by skeletal studies of spina bifida at Canimar Abajo (Rivero de la Calle 1987). Lastly, the deviations from HWE may indicate an unknown selective advantage to heterozygosity.

Taken as a whole, this study confirms the value of molecular and paleoepidemiological approaches to the study of health and disease in the past. Combining various forms of evidence to the study of spina bifida revealed aspects of reproductive, evolutionary, dietary, and of course, genetic processes at Canimar Abajo.

Chapter 6 - Conclusions and Future Considerations

Based on the evidence collected so far, it is possible to begin discussing the genetic versus dietary contributions to the prevalence of spina bifida in the population of Canimar Abajo. The diet of the population, based on stable isotope, starch grain, and zooarchaeological analyses, seems to be a mixed diet, but with insufficient sources of dietary folate. It is not possible to determine whether there was a folate deficiency sufficiently severe to result in pathological conditions in the population generally, but it was potentially serious enough to result in neural tube defects in infants of mothers not ingesting enough folate during conception and pregnancy. This is particularly true if there is an underlying genetic risk to spina bifida in the population.

There is certainly an underlying genetic predisposition, given the high frequency of variant alleles, particularly in MTHFR 677, the SNP with the most demonstrable risk for NTDs, and MTRR A66G. It is difficult to draw any conclusions so far given the small sample and the need for replication, but it seems likely that the underlying genetic predisposition greatly increases the risk for disease, particularly in a low folate environment.

These two components seem to have come together to create a high risk environment for the development of spina bifida, contributing to its high prevalence in the population. Given the nature of spina bifida, it is not possible to determine whether there were more individuals who demonstrated the spina bifida phenotype than what was demonstrated by macroscopic paleopathology, but it is possible, with molecular paleopathology and with dietary analysis, to get a better understanding of the underlying

risk to disease in the population.

In the future, this study should be replicated to confirm the SNP results. Additionally, this approach should be expanded to include infants and juveniles. This subpopulation may include homozygous variant individuals, those with more severe forms of spina bifida or those who did not survive the complications associated with spina bifida. The genetic and dietary results for this group may reveal more distinctly what processes were at play. Another route of investigation could be a deeper look into the availability of folate sources, and whether there was any differential access to these sources in the population. Lastly, it would be very interesting to investigate whether there was any change over time, both in terms of population genetics and diet. At this point, there are not enough individuals in the older cemetery at Canimar to discuss these types of changes, but with further work conducted at Canimar Abajo, this may be possible.

Overall, I believe I have achieved the research objectives: 1) evaluate the disease of spina bifida, present in the skeletal evidence from the pre-Columbian Cuban population from the archaeological site of Canimar Abajo; 2) determine the distribution of single nucleotide polymorphisms (SNPs) associated with spina bifida in the population using the methods of molecular paleopathology; 3) discuss the distribution of these SNPs from a paleoepidemiological perspective, comparing this distribution to macroscopic evidence of the disease and to dietary indicators; and 4) contribute to the understanding of how molecular paleopathology and paleoepidemiology can be used to its full potential within biological anthropology.

References

Adachi N, Shinoda K-i, Umetsu K, Kitano T, Matsumura H, Fujiyama R, Sawada J, and Tanaka M. 2011. Mitochondrial DNA analysis of Hokkaido Jomon skeletons: Remnants of archaic maternal lineages at the southwestern edge of former Beringia. American Journal of Physical Anthropology 146(3):346-360.

Alonso A, Martın P, Albarrán C, Garcıa P, Garcıa O, de Simón LF, Garcıa-Hirschfeld J, Sancho M, de la Rúa C, and Fernández-Piqueras J. 2004. Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies. Forensic Science International 139(2–3):141-149.

Angel JL. 1966. Porotic hyperostosis, anemias, malarias, and marshes in the prehistoric eastern Mediterranean. Science 153:760-763.

Angeline T, Jeyaraj N, Granito S, and Tsongalis GJ. 2004. Prevalence of MTHFR gene polymorphisms (C677T and A1298C) among Tamilians. Experimental and Molecular Pathology 77(2):85-88.

Applied Biosystems. 2009. DNA sequencing by capillary electrophoresis. Foster City, CA: Applied Biosystems.

Applied Biosystems. 2010. ABI Prism® SNaPshot™ multiplex kit protocol. Foster City, CA: Applied Biosystems.

Arenas Laserna JA. 2009. Estudio osteologico de los restos humanos subadultos del sitio arqueologico Canimar Abajo, Matanzas, Cuba. Havana: University of Havana.

Arenas Laserna JA, and Arredondo Antúnez C. 2008. Composición por edades (subadultos) en el sitio arqueológico preagroalfarero Canimar Abajo, Matanzas Cuba. IX Conferencia Internacional Antropología. Havana.

Armstrong SD, Cloutier LA, Arredondo C, Roksandic M, and Matheson CD. in press. Spina bifida in a pre-Columbian population: A paleoepidemiological study of genetic and dietary risk factors. International Journal of Paleopathology.

Arredondo C. 2004. Informe zooarqueológico sobre las excavaciones del año 2004 en el sitio Canímar Abajo. Havana: University of Havana.

Au KS, Ashley-Koch A, and Northrup H. 2010. Epidemiologic and genetic aspects of spina bifida and other neural tube defects. Developmental Disabilities Research Reviews 16(1):6-15.

Aufderheide AC, Salo W, Madden M, Streitz J, Buikstra J, Guhl F, Arriaza B, Renier C, Wittmers LE, Fornaciari G et al. . 2004. A 9,000-year record of Chagas' disease. Proceedings of the National Academy of Sciences of the United States of America 101(7):2034-2039.

Avrahami E, Frishman E, Fridman Z, and Azor M. 1994. Spina bifida occulta of S1 is not an innocent finding. Spine 19(1):12-15.

Barbaux S, Kluijtmans LAJ, and Whitehead AS. 2000. Accurate and rapid “multiplex heteroduplexing” method for genotyping key enzymes involved in folate/homocysteine metabolism. Clinical Chemistry 46(7):907-912.

Beaudin AE, and Stover PJ. 2007. Folate-mediated one-carbon metabolism and neural tube defects: Balancing genome synthesis and gene expression. Birth Defects Research (Part C) 81:183-203.

Berman MJ, and Pearsall DM. 2008. At the crossroads: Starch grain and phytolith analyses in Lucayan prehistory. Latin American Antiquity 19(2):182-204.
Blom HJ, Shaw GM, den Heijer M, and Finnell RH. 2006. Neural tube defects and folate: case far from closed. Nature Reviews Neuroscience 7(9):724-731.
Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM, and van der Noordaa J. 1990. Rapid and simple method for purification of nucleic acids. Journal of Clinical Microbiology 28(3):495-503.
Bosco P, Guéant-Rodriguez R-M, Anello G, Barone C, Namour F, Caraci F, Romano A, Romano C, and Guéant J-L. 2003. Methionine synthase (MTR) 2756 (A → G) polymorphism, double heterozygosity methionine synthase 2756 AG/methionine synthase reductase (MTRR) 66 AG, and elevated homocysteinemia are three risk factors for having a child with Down syndrome. American Journal of Medical Genetics Part A 121A(3):219-224.
Botto LD, and Mastroiacovo P. 1998. Exploring gene—gene interactions in the etiology of neural tube defects. Clinical Genetics 53(6):456-459.
Botto LD, Moore CA, Khoury MJ, and Erickson JD. 1999. Neural-tube defects. New England Journal of Medicine 341(20):1509-1519.
Botto LD, and Yang Q. 2000. 5, 10-Methylenetetrahydrofolate reductase gene variants and congenital anomalies: A HuGE review. American Journal of Epidemiology 151(9):862-877.
Bouakaze C, Keyser C, Amory S, Crubézy E, and Ludes B. 2007. First successful assay of Y-SNP typing by SNaPshot minisequencing on ancient DNA. International Journal of Legal Medicine 121(6):493-499.
Bramanti B, Chiarelli B, Herrmann B, and Hummel S. 2003. Ancient DNA analysis of the delta F508 mutation. Human Biology 75(1):105-115.
Bronner-Fraser M. 2006. Neural crest formation. In: Wyszynski DF, editor. Neural tube defects: From origin to treatment. Oxford: Oxford University Press. p 5-14.
Buhay WM, Chinique de Armas Y, Rodriguez Suárez R, Arredondo C, Smith DG, Armstrong SD, and Roksandic M. In press. A preliminary carbon and nitrogen collagen isotopic investigation on skeletal remains recovered from a pre-Columbian burial site, Matanzas Province, Cuba. Applied Geochemistry.
Buikstra JE, and Ubelaker DH. 1994. Standards for data collection from human skeletal remains. Fayetteville: Arkansas Archaeological Survey Press.
Burger J, Kirchner M, Bramanti B, Haak W, and Thomas MG. 2007. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. Proceedings of the National Academy of Sciences 104(10):3736-3741.
Bush H, and Zvelebil M. 1991. Pathology and health in past societies: An introduction. In: Bush H, and Zvelebil M, editors. Health in past societies: Biocultural interpretations of human skeletal remains in archaeological contexts. Oxford: Archaeological and Historical Associates. p 3-9.
Cabrera YC. 2008. Acercamiento a algunas problemáticas relacionades con el uso de los espacios funébres aborígenes. Sitio arqueológico Canimar Abajo, Matanzas, Cuba. IX Conferencia Internacional Antropología. Havana.
Callaghan R. 2003. Comments on the mainland origins of the preceramic cultures of the Greater Antilles. Latin American Antiquity 14(3):323-338.

Campillo Álvarez E. 2009. Paleopatologías en aborígenes cubanos del sitio arqueológico Canímar Abajo, Matanzas, Cuba. Havana: University of Havana.
Carr C, and Case DT. 2006. Gathering Hopewell: Society, Ritual, and Ritual Interaction. New York: Springer Science and Media Inc. 807 p.
Chakravarthy MV, and Booth FW. 2004. Eating, exercise, and "thrifty" genotypes: connecting the dots toward an evolutionary understanding of modern chronic diseases. Journal of Applied Physiology 96(1):3-10.
Chamberlain JS, Gibbs RA, Rainer JE, Nguyen PN, and Thomas C. 1988. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. Nucleic Acids Research 16(23):11141-11156.
Chanarin I, Rothman D, Ward A, and Perry J. 1968. Folate status and requirement in pregnancy. British Medical Journal 2:390-394.
Chinique de Armas Y. 2007. Estudios paleodietarios del sitio arqueológico Canímar Abajo, Matanzas, Cuba. Havana: University of Havana.
Chinique de Armas Y. 2009. Caracterización de los modelos de desarrollo económico de los aborígenes exhumados en el sitio arqueológico Canímar Abajo, Matanzas, Cuba. Havana: University of Havana.
Chinique de Armas Y, and Arredondo Antúnez C. 2008. Utilidad de los huesos de manos y pies para estudios osteológicos: Determinación del número mínimo de individuos adultos exhumados en el sitio arqueológico Canimar Abajo, Matanzas, Cuba. IX Conferencia Internacional Antropología. Havana.
Chinique de Armas Y, Rodríguez Suárez R, Arredondo Antúne C, Collazo O, Boza A, Sheila A, Alvarez M, Liva M, and Jiménez J. 2008. Estudio paleodietario en restos óseos aborígenes del sitio arqueológico Canímar Abajo, Matanzas, Cuba. Boletín Antropológico 2:131-148.
Christensen B, Arbour L, Tran P, Leclerc D, Sabbaghian N, Platt R, Gilfix BM, Rosenblatt DS, Gravel RA, Forbes P et al. . 1999. Genetic polymorphisms in methylenetetrahydrofolate reductase and methionine synthase, folate levels in red blood cells, and risk of neural tube defects. American Journal of Medical Genetics 84(2):151-157.
Clark GA. 1988. New method for assessing changes in growth and sexual dimorphism in paleoepidemiology. American Journal of Physical Anthropology 77(1):105-116.
Cloutier LA. 2008. Towards the optimisation of a detection system for markers linked to neural tube defects. Thunder Bay, ON: Lakehead University.
Cooper J. 2010. Pre-Columbian archaeology of Cuba: A study of site distribution patterns and radiocarbon chronologies. In: Fitzpatrick SM, and Ross AH, editors. Island shores, distant pasts: Archaeological and biological approaches to the pre-Columbian settlement of the Caribbean. Gainesville: University Press of Florida. p 81-107.
Coppa A, Cucina A, Hoogland MLP, Lucci M, Calderón FL, Panhuysen RGAM, María GT, Rojas V, and Vargiu R. 2008. New evidence of two different migratory waves in the circum-Caribbean area during the pre-Columbian period from the analysis of dental morphological traits. In: Hofman CL, Hoogland MLP, and van Gijn AL, editors. Crossing the borders: New methods and techniques in the study of archaeological materials from the Caribbean. Tuscaloosa: University of

Alabama Press. p 195-213.
Cosculluela JA. 1965 [1918]. Cuatro años en la Cienaga de Zapata. Havana: Comision Nacional Cubana de la UNESCO.
Curet LA. 2003. Issues on the diversity and emergence of Middle-Range societies of the ancient Caribbean: A critique. Journal of Archaeological Research 11(1):1-42.
Dacal Moure R, and Rivero De La Calle M. 1996. Art and archaeology of pre-Columbian Cuba. Pittsburgh: University of Pittsburgh Press.
Dahl M, and Ahlsten G. 2010. Spina bifida occulta in children: presentation and outcome. Cerebrospinal Fluid Research 7(0):1-1.
De Marco P, Calevo MG, Moroni A, Arata L, Merello E, Finnell RH, Zhu H, Andreussi L, Cama A, and Capra V. 2002. Study of MTHFR and MS polymorphisms as risk factors for NTD in the Italian population. Journal of Human Genetics 47(6):319-324.
DeNiro MJ. 1985. Postmortem preservation and alteration of in vivo bone collagen isotope ratios in relation to palaeodietary reconstruction. Nature 317(6040):806-809.
Donoghue HD, Marcsik A, Matheson C, Vernon K, Nuorala E, Molto JE, Greenblatt CL, and Spigelman M. 2005. Co–infection of *Mycobacterium tuberculosis* and *Mycobacterium leprae* in human archaeological samples: A possible explanation for the historical decline of leprosy. Proceedings of the Royal Society B: Biological Sciences 272(1561):389-394.
Donoghue HD, Spigelman M, Greenblatt CL, Lev-Maor G, Bar-Gal GK, Matheson C, Vernon K, Nerlich AG, and Zink AR. 2004. Tuberculosis: From prehistory to Robert Koch, as revealed by ancient DNA. The Lancet: Infectious Diseases 4:38-46.
Doolin M-T, Barbaux S, McDonnell M, Hoess K, Whitehead AS, and Mitchell LE. 2002. Maternal genetic effects, exerted by genes involved in homocysteine remethylation, influence the risk of spina bifida. The American Journal of Human Genetics 71(5):1222-1226.
Dumas León D. 2009. Historia y osteología de los restos óseos humanos exhumados entre 1984 y 1987 en Canímar Abajo, Matanzas, Cuba. Havana: University of Havana.
El-Najjar MY, Lozoff B, and Ryan DJ. 1975. The paleoepidemiology of porotic hyperostosis in the American Southwest: Radiological and ecological considerations. American Journal of Roentgenology 125(4):918-924.
Eubanks JD, and Cheruvu VK. 2009. Prevalence of sacral spina bifida occulta and its relationship to age, sex, race, and the sacral table angle: An anatomic, osteologic study of three thousand one hundred specimens. Spine 34(15):1539-1543.
Franco RF, Elion J, Lavinha J, Krishnamoorthy R, Tavella MH, and Zago MA. 1998. Heterogeneous ethnic distribution of the 844ins68 in the cystathionine β-synthase gene. Human Heredity 48:338-342.
Friedman G, Goldschmidt N, Friedlander Y, Ben-Yehuda A, Selhub J, Babaey S, Mendel M, Kidron M, and Bar-On H. 1999. A common mutation A1298C in human methylenetetrahydrofolate reductase gene: Association with plasma total homocysteine and folate concentrations. The Journal of Nutrition 129(9):1656-

1661.
Gilbert MTP, Binladen J, Miller W, Wiuf C, Willerslev E, Poinar H, Carlson JE, Leebens-Mack JH, and Schuster SC. 2007. Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. Nucleic Acids Research 35(1):1-10.
Gillespie JH. 1998. Population genetics: A concise guide. Baltimore: John Hopkins University Press.
González Herrera UM, Martínez López JG, Pereira Pereira O, Arredondo Antúnez C, Rangel A, Rodríguez Suárez R, and Hernández Godoy S. 2005. Informe de excavación: Sitio Canímar Abajo, Matanzas. Havana: University of Havana.
Gos M, and Szpecht-Potocka A. 2002. Genetic basis of neural tube defects. II. Genes correlated with folate and methionine metabolism. Journal of Applied Genetics 43(4):511-524.
Graham HK, and Parsch K. 2009. Neural tube defects, spina bifida, and spinal dysraphism. In: Benson M, Fixsen J, Macnicol M, and Parsch K, editors. Children's orthopaedics and fractures. London: Springer. p 265-286.
Granberry J, and Vescelius GS. 2004. Languages of the pre-Columbian Antilles. Tuscaloosa: University of Alabama Press.
Greene NDE, and Copp AJ. 2006. The embryonic basis of neural tube defects. In: Wyszynski DF, editor. Neural tube defects: From origin to treatment. Oxford: Oxford University Press. p 15-28.
Guéant-Rodriguez RM, Rendeli C, Namour B, Venuti L, Romano A, Anello G, Bosco P, Debard R, Gérard P, Viola M et al. . 2003. Transcobalamin and methionine synthase reductase mutated polymorphisms aggravate the risk of neural tube defects in humans. Neuroscience Letters 344(3):189-192.
Haas CJ, Zink A, Molńar E, Szeimies U, Reischl U, Marcsik A, Ardagna Y, Dutour O, Pálfi G, and Nerlich AG. 2000a. Molecular evidence for different stages of tuberculosis in ancient bone samples from Hungary. American Journal of Physical Anthropology 113(3):293-304.
Haas CJ, Zink A, Pálfi G, Szeimies U, and Nerlich AG. 2000b. Detection of leprosy in ancient human skeletal remains by molecular identification of *Mycobacterium leprae*. American Journal of Clinical Pathology 114(3):428-436.
Hague WM. 2003. Homocysteine and pregnancy. Best Practice and Research: Clinical Obstetrics and Gynaecology 17(3):459-469.
Halcrow SE, Tayles N, and Livingstone V. 2008. Infant Death in Late Prehistoric Southeast Asia. Asian Perspectives 47(2):341-404.
Haslam M. 2004. The decomposition of starch grains in soils: Implications for archaeological residue analyses. Journal of Archaeological Science 31:1715-1734.
Haslam N, and Probert CS. 1998. An audit of the investigation and treatment of folic acid deficiency. Journal of the Royal Society of Medicine 91(2):72-73.
Hastie T, Tibshirani R, and Friedman J. 2009. The elements of statistical learning: Data mining, inference, and prediction. New York: Springer.
Hayward MH, Atkinson L-G, and Cinquino MA. 2009. Introduction. In: Hayward MH, Atkinson L-G, and Cinquino MA, editors. Rock art of the Caribbean. Tuscaloosa: University of Alabama Press. p 1-12.

Henegariu O, Heerema N, Dlouhy S, Vance G, and Vogt P. 1997. Multiplex PCR: Critical parameters and step-by-step protocol. BioTechniques 23(3):504-511.
Herrmann B, and Hummel S. 1998. Palaeoepidemiology of infectious diseases using aDNA analyses: pros and cons. Bulletins et Mémoires de la Société d'anthropologie de Paris:195-200.
Hofreiter M, Serre D, Poinar HN, Kuch M, and Pääbo S. 2001. Ancient DNA. Nature Reviews Genetics 2(5):353-359.
Hulme P. 1993. Making sense of the native Caribbean. New West Indian Guide 67:189-220.
Hummel S, and Herrmann B. 1995. aDNA analysis in paleopathology: mini-review and prospects. Paleopathology Newsletter(91):6-9.
International Centre for Birth Defects. 2003. World atlas of birth defects. Geneva: World Health Organization.
Jackes M. 1993. On paradox and osteology. Curr Anthropol 34(4):434-439.
Jackes M, and Meiklejohn C. 2008. The paleodemography of central Portugal and the Mesolithic-Neolithic transition. In: Bocquet-Appel J-P, editor. Recent advances in palaeodemography: Springer Netherlands. p 209-258.
Jacques PF, Bostom AG, Williams RR, Ellison RC, Eckfeldt JH, Rosenberg IH, Selhub J, and Rozen R. 1996. Relation between folate status, a common mutation in methylenetetrahydrofolate reductase, and plasma homocysteine concentrations. Circulation 93(1):7-9.
Jährig K, and Wüchner U. 1985. Incidence of spina bifida occulta in relatives of children with myelodysplasia. Indications for high-risk screening. Zentralblatt für Gynäkologie 107(13):821-826.
Johanning GL, Tamura T, Johnston KE, and Wenstrom KD. 2000. Comorbidity of 5,10-methylenetetrahydrofolate reductase and methionine synthase gene polymorphisms and risk for neural tube defects. Journal of Medical Genetics 37(12):949-951.
Johnson PCD, and Haydon DT. 2007. Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. Genetics 175:827-842.
Jurmain R. 1990. Paleoepidemiology of a central California prehistoric population from CA-ALA-329: II. Degenerative disease. American Journal of Physical Anthropology 83(1):83-94.
Kamen B. 1997. Folate and antifolate pharmacology. Seminars in oncology 24(5 Suppl 18):S18-30-S18-39.
Katzenberg MA. 2000. Stable isotope analysis: A tool for studying past diet, demography, and life history. In: Katzenberg MA, and Saunders SR, editors. Biological anthropology of the human skeleton. New York: John Wiley & Sons. p 305-328.
Keegan WF. 1989. Creating the Guanahatabey (Ciboney): The modern genesis of an extinct culture. Antiquity 63(239):373-379.
Keegan WF. 2007. Taíno Indian myth and practice : The arrival of the stranger king. Gainesville: University Press of Florida.
Keegan WF, and DeNiro MJ. 1988. Stable carbon- and nitrogen-isotope ratios of bone

collagen used to study coral-reef and terrestrial components of prehistoric Bahamian diet. American Antiquity 53(2):320-336.
King C, Debruyne R, Kuch M, Schwarz C, and Poinar H. 2009. A quantitative approach to detect and overcome PCR inhibition in ancient DNA extracts. BioTechniques 47(5):941.
Kirby RR. 2006. Methodological issues in the study of neural tube defects. In: Wyszynski DF, editor. Neural tube defects: From origin to treatment. Oxford: Oxford University Press. p 103-108.
Kirke PN, Molloy AM, Daly LE, Burke H, Weir DC, and Scott JM. 1993. Maternal plasma folate and vitamin B12 are independent risk factors for neural tube defects. Quarterly Journal of Medicine 86(11):703-708.
Kozlowski JK. 1974. Preceramic cultures in the Caribbean. Krakow: PZeszyty Naukowe, Uniwerstytetu Jagiellonskiego.
Kuch M, Gröcke DR, Knyf MC, Gilbert MTP, Younghusband B, Young T, Marshall I, Willerslev E, Stoneking M, and Poinar H. 2007. A preliminary analysis of the DNA and diet of the extinct Beothuk: A systematic approach to ancient human DNA. American Journal of Physical Anthropology 132(4):594-604.
Kumar A, and Tubbs RS. 2011. Spina bifida: A diagnostic dilemma in paleopathology. Clinical Anatomy 24(1):19-33.
La Rosa Corzo G. 2003. La orientacion este de los entierros aborigenes en cuevas de Cuba: Remate de una fabula. Latin American Antiquity 14(2):143-157.
Lalueza-Fox C, Gilbert MTP, Martínez-Fuentes AJ, Calafell F, and Bertranpetit J. 2003. Mitochondrial DNA from pre-Columbian Ciboneys from Cuba and the prehistoric colonization of the Caribbean. American Journal of Physical Anthropology 121(2):97-108.
Lee EJ, Makarewicz C, Renneberg R, Harder M, Krause-Kyora B, Müller S, Ostritz S, Fehren-Schmitz L, Schreiber S, Müller J et al. . 2012. Emerging genetic patterns of the European Neolithic: Perspectives from a late Neolithic Bell Beaker burial site in Germany. American Journal of Physical Anthropology:n/a-n/a.
Lieverse AR. 1999. Diet and the aetiology of dental calculus. International Journal of Osteoarchaeology 9(4):219-232.
Lieverse AR, Link DW, Bazaliiskiy VI, Goriunova OI, and Weber AW. 2007. Dental health indicators of hunter–gatherer adaptation and cultural change in Siberia's Cis-Baikal. American Journal of Physical Anthropology 134(3):323-339.
Lin J, Spitz MR, Wang Y, Schabath MB, Gorlov IP, Hernandez LM, Pillow PC, Grossman HB, and Wu X. 2004. Polymorphisms of folate metabolic genes and susceptibility to bladder cancer: A case-control study. Carcinogenesis 25(9):1639-1647.
Lorber J, and Levick K. 1967. Spina bifida cystica: Incidence of spina bifida occulta in parents and in controls. Archives of Disease in Childhood 42(222):171-173.
Lovén S. 1935. Origins of the Tainan culture, West Indies. Göteborg: Elanders Boktryckeri Aktiebolag.
Markoulatos P, Siafakas N, and Moncany M. 2002. Multiplex polymerase chain reaction: A practical approach. Journal of Clinical Laboratory Analysis 16(1):47-51.
Martínez López JG. 2009. Caracterizacion tafonomica del sitio arqueologico Canimar

Abajo, Matanzas, Cub: Un estudio en restos osteos humanos. Havana: University of Havana.
Matheson C, Marion TE, Hayter S, Esau N, Fratpietro R, and Vernon KK. 2009. Technical note: Removal of metal ion inhibition encountered during DNA extraction and amplification of copper-preserved archaeological bone using size exclusion chromatography. American Journal of Physical Anthropology 140(2):384-391.
Matheson CD, Gurney C, Esau N, and Lehto R. 2010. Assessing PCR inhibition from humic substances. Open Enzyme Inhibition Journal 3:38-45.
Miller CR, Joyce P, and Waits LP. 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160:357-366.
Mills JL, Lee YJ, Conley MR, Kirke PN, McPartlin JM, Weir DG, and Scott JM. 1995. Homocysteine metabolism in pregnancies complicated by neural-tube defects. The Lancet 345(8943):149-151.
Milton K. 2000. Hunter-gatherer diets: A different perspective. The American Journal of Clinical Nutrition 71(3):665-667.
Moore CA. 2006. Classification of neural tube defects. In: Wyszynski DF, editor. Neural tube defects: From origin to treatment. Oxford: Oxford University Press. p 66-75.
Morales D, and Arredondo Antúnez C. 2008. Contribución a la determinación del sexo en esqueletos de individuos adultos de filiación preagroalfarera, Canímar Abajo, Cuba. Datos osteométricos. IX Conferencia Internacional Antropología. Havana.
Morales Valdes D. 2009. Caracterización osteológica de aborígenes adultos exhumados en Canímar Abajo, durante la campaña de excavación del 2006. Havana: University of Havana.
Morrison K, Papapetrou C, Hol FA, Mariman ECM, Lynch SA, Burn J, and Edwards YH. 1998. Susceptibility to spina bifida; an association study of five candidate genes. Annals of Human Genetics 62(5):379-396.
Mullis K, Faloona F, Scharf S, Saiki R, Horn G, and Erlich H. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symposium on Quantitative Biology 51(Pt 1):263-273.
National Institutes of Health. 2009. Folate. Office of Dietary Supplements, National Institutes of Health.
Nerlich AG, Schraut B, Dittrich S, Jelinek T, and Zink AR. 2008. *Plasmodium falciparum* in ancient Egypt. Emerging Infectious Diseases 14(8):1317-1319.
Northrup H, and Volcik KA. 2000. Spina bifida and other neural tube defects. Current Problems in Pediatrics 30(10):317-332.
O'Connor T. 2008. The Archaeology of Animal Bones: TAMU Press.
O'Leary VB, Mills JL, Pangilinan F, Kirke PN, Cox C, Conley M, Weiler A, Peng K, Shane B, Scott JM et al. . 2005. Analysis of methionine synthase reductase polymorphisms for neural tube defects risk association. Molecular Genetics and Metabolism 85(3):220-227.
Pääbo S, Poinar HN, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, and Hofreiter M. 2004. Genetic analyses from ancient DNA. Annual Review of Genetics 38:645-679.
Palacajornsuk P, Halter C, Isakova V, Tarnawski M, Farmar J, Reid ME, and Chaudhuri

A. 2009. Detection of blood group genes using multiplex SNaPshot method. Transfusion 49(4):740-749.
Pané FR. 1999. An account of the antiquities of the Indians (1571). Durham: Duke University Press.
Pinhasi R, and Turner K. 2007. Epidemiological approaches in palaeopathology. In: Pinhasi R, and Mays S, editors. Advances in human palaeopathology: Wiley and Sons. p 45-56.
Piperno DR, and Holst I. 1998. The presence of starch grains on prehistoric stone tools from the humid neotropics: Indications of early tuber use and agriculture in Panama. Journal of Archaeological Science 25(8):765-776.
Rady PL, Szucs S, Grady J, Hudnall SD, Kellner LH, Nitowsky H, Tyring SK, and Matalon RK. 2002. Genetic polymorphisms of methylenetetrahydrofolate reductase (MTHFR) and methionine synthase reductase (MTRR) in ethnic populations in Texas; a report of a novel MTHFR polymorphic site, G1793A. American Journal of Medical Genetics 107(2):162-168.
Ramos RR. 2008. From the Guanahatabey to the Archaic of Puerto Rico: The nonevident evidence. Ethnohistory 55(3):393-415.
Ramos RR, Babilonia E, Curet LA, and Ulloa J. 2008. The pre-Arawak pottery horizon in the Antilles: A new approximation. Latin American Antiquity 19(1):47-63.
Ramsbottom D, Scott JM, Molloy A, Weir DG, Kirke PN, Mills JL, Gallagher PM, and Whitehead AS. 1997. Are common mutations of cystathionine β-synthase involved in the aetiology of neural tube defects? Clinical Genetics 51(1):39-42.
Reid BA. 2009. Myths and realities of Caribbean history. Tuscaloosa: University of Alabama Press.
Reidpath DD, and Allotey P. 2003. Infant mortality rate as an indicator of population health. Journal of Epidemiology and Community Health 57(5):344-346.
Reimer PJ, Baillie MGL, Bard E, Bayliss A, Beck JW, Bertrand CJH, Blackwell PG, Buck CE, Burr GS, Cutler KB et al. . 2004. IntCal04 terrestrial radiocarbon age calibration, 0-26 cal kyr BP. Radiocarbon 46(3):1029-1058.
Reimer PJ, Baillie MGL, Bard E, Bayliss A, Beck JW, Blackwell PG, Ramsey CB, Buck CE, Burr GS, and Edwards RL. 2009. Intcal09 and Marine09 radiocarbon age calibration curves, 0-50,000 years cal BP. Radiocarbon 51(4):1111-1150.
Relton CL, Wilding CS, Pearce MS, Laffling AJ, Jonas PA, Lynch SA, Tawn EJ, and Burn J. 2004. Gene–gene interaction in folate-related genes and risk of neural tube defects in a UK population. Journal of Medical Genetics 41(4):256-260.
Relton CL, Wildling CS, Jonas PA, Lynch SA, Tawn EJ, and J. B. 2003. Genetic susceptibility to neural tube defect pregnancy varies with offspring phenotype. Clinical Genetics 64:424-428.
Ritchey ML, Sinha A, DiPietro MA, Huang C, Flood H, and Bloom DA. 1994. Significance of spina bifida occulta in children with diurnal enuresis. Journal of Urology 152:815-818.
Rivero de la Calle M. 1987. Reporte de la presencia de sacros con espina bífida en el sitio aborigen de Canímar Abajo, Matanzas, Cuba. Revista Biologia 1(1):75-83.
Rivero de la Calle M. 1988. Paleopatologia de los aborigénes de Cuba. Estudios de antropología biológica:(III Coloquio de Antropología Física Juan Comas, 1984)

81:475-498.
Rizk E, and Iskandar BJ. 2010. The role of folate supplementation in spina bifida occulta. In: Yamada S, editor. Tethered cord syndrome in children and adults. New York: Thieme Medical Publishers. p 142-144.
Roberts C, and Ingham S. 2008. Using ancient DNA analysis in palaeopathology: a critical analysis of published papers, with recommendations for future work. International Journal of Osteoarchaeology 18(6):600-613.
Roberts C, and Manchester K. 2005. The archaeology of disease. Ithaca, NY: Cornell University Press.
Rodríguez MP, and Medina AC. 2004. La zooarqueología en las comunidades con tradiciones Neolíticas incipientes de la región central de Cuba. VII Conferencia Internacional Antropologia 2004.
Rodríguez R. 2007. Canímar Abajo: No solo recolectores y pescadores. Primer Congreso Iberoamericano de Antropología La Antropología ante los desafíos del siglo XXI. Havana: Desoft s.a. Macromedia Inc. p 784-795.
Rodríguez Suárez R, Antúnez CA, Rivero AR, Hernández SHG, Hernández OL, Herrera UG, López JM, and Pereira OP. 2006. 5000 años de ocupación prehispánica en Canímar Abajo, Matanzas, Cuba. . Memorias del congreso Antropología.
Rodríguez Suárez R, Buhay B, and Roksandic M. 2010. New dates from Canimar Abajo. Biological and cultural landscapes of indigenous people in pre- and post-contact Cuba. Winnipeg.
Rodríguez Suárez R, and Pagán Jiménez JR. 2008. The burén in precolonial Cuban archaeology: New information regarding the use of plants and ceramic griddles during the late Ceramic Age of eastern Cuba gathered through starch analysis. In: Hofman CL, Hoogland MLP, and van Gijn AL, editors. Crossing the borders: New methods and techniques in the study of archaeological materials from the Caribbean. Tuscaloosa, Ala: University of Alabama Press. p 159-169.
Roe PG. 2009. The mute stones speak: The past, present, and future of Caribbean rock art research. In: Hayward MH, Atkinson L-G, and Cinquino MA, editors. Rock art of the Caribbean. Tuscaloosa: University of Alabama Press. p 198-240.
Roksandic M, and Armstrong SD. 2011. Using the life history model to set the stage(s) of growth and senescence in bioarchaeology and paleodemography. American Journal of Physical Anthropology 145:337-347.
Roksandic M, Arredondo C, Buhay WM, Smith D, Chinique de Armas Y, Armstrong SD, Roksandic I, Matheson C, Parra E, Peros M et al. . In prep. Evidence for early use of domesticated plant species in Cuba.
Rouse I. 1992. The Tainos: Rise & decline of the people who greeted Columbus. New Haven: Yale University Press.
Rozen R. 2006. Genetic risk factors for neural tube defects: Variants in folate metabolism. In: Wyszynski DF, editor. Neural tube defects: From origin to treatment. Oxford: Oxford University Press. p 176-184.
Salo WL, Aufderheide AC, Buikstra J, and Holcomb TA. 1994. Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. Proceedings of the National Academy of Sciences 91(6):2091-2094.
Sanchez JJ, Phillips C, Børsting C, Balogh K, Bogus M, Fondevila M, Harrison CD,

Musgrave-Brown E, Salas A, Syndercombe-Court D et al. . 2006. A multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis 27:1713-1724.

Sanderson P, McNulty H, Mastroiacovo P, McDowell IFW, Melse-Boonstra A, Finglas PM, and Gregory III JF. 2003. Folate bioavailabilility: UK Food Standards Agency workshop report. British Journal of Nutrition 90:473-479.

Sattenspiel L, and Harpending H. 1983. Stable populations and skeletal age. American Antiquity 48:489-498.

Saunders NJ. 2005. Archaic period. In: Saunders NJ, editor. The peoples of the Caribbean: An encyclopedia of archaeology and traditional culture. Santa Barbara: ABC-CLIO. p 13-16.

Scher AI, Terwindt GM, Verschuren WMM, Kruit MC, Blom HJ, Kowa H, Frants RR, van den Maagdenberg AMJM, van Buchem M, Ferrari MD et al. . 2006. Migraine and MTHFR C677T genotype in a population-based sample. Annals of Neurology 59(2):372-375.

Schmidt D, Hummel S, and Herrmann B. 2003. Brief communication: Multiplex X/Y-PCR improves sex identification in aDNA analysis. American Journal of Physical Anthropology 121(4):337-341.

Schwarcz HP, and Schoeninger MJ. 2011. Stable isotopes of carbon and nitrogen as tracers for paleo-diet reconstruction. In: Baskaran M, editor. Handbook of environmental isotope geochemistry. Berlin: Springer. p 725-742.

Seaver LH, and Stevenson RE. 2006. Syndromes with neural tube defects. In: Wyszynski DF, editor. Neural tube defects: From origin to treatment. Oxford: Oxford University Press. p 76-83.

Shaw GM, Velie EM, and Schaffer DM. 1997. Is dietary intake of methionine associated with a reduction in risk for neural tube defect-affected pregnancies? Teratology 56(5):295-299.

Skibola CF, Smith MT, Hubbard A, Shane B, Roberts AC, Law GR, Rollinson S, Roman E, Cartwright RA, and Morgan GJ. 2002. Polymorphisms in the thymidylate synthase and serine hydroxymethyltransferase genes and risk of adult acute lymphocytic leukemia. Blood 99(10):3786-3791.

Smith BD. 2001. Low-level food production. Journal of Archaeological Research 9(1):1-43.

Souza S, Carvalho D, and Lessa A. 2003. Paleoepidemiology: is there a case to answer? Memórias do Instituto Oswaldo Cruz 98:21-27.

Spacca B, and Buxton N. 2008. Spina bifida occulta and monozygotic twins. Journal of Neurosurgery: Pediatrics 2(4):258-260.

Speer MC, Nye J, McLone D, Worley G, Melvin EC, Viles KD, Franklin A, Drake C, Mackey J, George TM et al. . 1999. Possible interaction of genotypes at cystathionine β-synthase and methylenetetrahydrofolate reductase (MTHFR) in neural tube defects. Clinical Genetics 56(2):142-144.

Steegers-Theunissen RP, Boers GH, Blom HJ, Nijhuis JG, Thomas CMG, Borm GF, and Eskes TK. 1995. Neural tube defects and elevated homocysteine levels in amniotic fluid. American Journal of Obstetrics and Gynecology 172(5):1436-1441.

Stevens-Arroyo AM. 2006. Cave of the Jagua: The mythological world of the Taínos. Scranton: Scranton University Press.
Sturtevant WC. 1968. History and ethnography of some West Indian starches. In: Ucko PJ, and Dimbleby GW, editors. The domestication and exploitation of plants and animals. Chicago: Aldine Publishing Company. p 177-200.
Tabío EE. 1988. Introducción a la arqueologia de las Antillas. Havana: Editorial de Ciencias Sociales.
Tabío EE, and Rey E. 1979. Prehistoria de Cuba. Havana: Editorial de Ciencias Sociales.
Templeton AR. 2006. Population genetics and microevolutionary theory. Hoboken, NJ: Wiley-Liss.
USDA. 2011. USDA national nutrient database for standard reference, Release 18 - folate, DFE (µg) content of selected foods per common measure, sorted by nutrient content. Beltsville, MD: United States Department of Agriculture.
van der Linden I, den Heijer M, Afman L, Gellekink H, Vermeulen S, Kluijtmans L, and Blom H. 2006a. The methionine synthase reductase 66A>G polymorphism is a maternal risk factor for spina bifida. Journal of Molecular Medicine 84(12):1047-1054.
van der Linden IJM, Afman LA, Heil SG, and Blom HJ. 2006b. Genetic variation in genes of folate metabolism and neural-tube defect risk. Proceedings of the Nutrition Society 65(02):204-215.
van der Put NMJ, Gabreëls F, Stevens EMB, Smeitink JAM, Trijbels FJM, Eskes TKAB, van den Heuvel LP, and Blom HJ. 1998. A second common mutation in the methylenetetrahydrofolate reductase gene: An additional risk factor for neural-tube defects? The American Journal of Human Genetics 62(5):1044-1051.
van der Put NMJ, Trijbels FJM, van den Heuvel LP, Blom HJ, Steegers-Theunissen RPM, Eskes TKAB, Mariman ECM, den Heyer M, Frosst P, and Rozen R. 1995. Mutated methylenetetrahydrofolate reductase as a risk factor for spina bifida. The Lancet 346(8982):1070-1071.
van der Put NMJ, van Straaten HWM, Trijbels FJM, and Blom HJ. 2001. Folate, homocysteine and neural tube defects: An overview. Experimental Biology and Medicine 226:243-270.
Vannier JP, Lefort J, Cavelier B, Ledosseur P, Assailly C, and Feingold J. 1981. Spina bifida cystica families x-ray examination and HLA typing. Pediatric Research 15(4):326-329.
Waldron T. 1994. Counting the dead: The epidemiology of skeletal populations. Chichester: John Wiley & Sons.
Waldron T. 2007. Palaeoepidemiology: The Measure of Disease in the Human Past. Walnut Creek, CA: Left Coast Press.
Walker PL, Bathurst RR, Richman R, Gjerdrum T, and Andrushko VA. 2009. The causes of porotic hyperostosis and cribra orbitalia: A reappraisal of the iron-deficiency-anemia hypothesis. American Journal of Physical Anthropology 139(2):109-125.
Watkins ML. 1998. Efficacy of folic acid prophylaxis for the prevention of neural tube defects. Mental Retardation and Developmental Disabilities 4:282-290.
Weisberg I, Tran P, Christensen B, Sibani S, and Rozen R. 1998. A second genetic polymorphism in methylenetetrahydrofolate reductase (MTHFR) associated with

decreased enzyme activity. Molecular Genetics and Metabolism 64(3):169-172.
Werler MM, Shapiro S, and Mitchell AA. 1993. Periconceptional folic acid exposure and risk of occurrent neural tube defects. JAMA: The Journal of the American Medical Association 269(10):1257-1261.
Wilcken B, Bamforth F, Li Z, Zhu H, Ritvanen A, Redlund M, Stoll C, Alembik Y, Dott B, Czeizel AE et al. . 2003. Geographical and ethnic variation of the 677C>T allele of 5,10 methylenetetrahydrofolate reductase (MTHFR): findings from over 7000 newborns from 16 areas world wide. Journal of Medical Genetics 40(8):619-625.
Williams JS, White CD, and Longstaffe FJ. 2009. Maya marine subsistence: Isotopic evidence from Marco Gonzalez and San Pedro, Belize. Latin American Antiquity 20:37-56.
Wilson A, Platt R, Wu Q, Leclerc D, Christensen B, Yang H, Gravel RA, and Rozen R. 1999. A common variant in methionine synthase reductase combined with low cobalamin (vitamin B12) increases risk for spina bifida. Molecular Genetics and Metabolism 67(4):317-323.
Wilson SM. 1997. Introduction to the study of the indigenous people of the Caribbean. In: Wilson SM, editor. The indigenous people of the Caribbean. Gainesville: University Press of Florida. p 1-8.
Wilson SM. 2007. The archaeology of the Caribbean. Cambridge: Cambridge University Press.
Wood JW, Milner GR, Harpending HC, Weiss KM, Cohen MN, Eisenberg LE, Hutchinson DL, Jankauskas R, Česnys G, Katzenberg MA et al. . 1992. The osteological paradox: Problems of inferring prehistoric health from skeletal samples [and comments and reply]. Current Anthropology 33(4):343-370.
Xeller CF, Sanders M, Athari M, and Gibson R. 2000. The relationship of facet asymmetry, spina bifida occulta, and transitional vertebrae in the lumbar spine to backache. Journal of Neurological Orthopedia Surgery 20:82-85.
Yang Q-H, Botto LD, Gallagher M, Friedman J, Sanders CL, Koontz D, Nikolova S, Erickson JD, and Steinberg K. 2008. Prevalence and effects of gene-gene and gene-nutrient interactions on serum folate and serum total homocysteine concentrations in the United States: findings from the third National Health and Nutrition Examination Survey DNA Bank. The American Journal of Clinical Nutrition 88(1):232-246.
Yuan Z, Cheng W, Hou A, Wang W, Zhang S, Liu D, Gao F, Li H, and Wang W. 2008. Constipation is associated with spina bifida occulta in children. Clinical Gastroenterology and Hepatology 6(12):1348-1353.
Zarrillo S, Pearsall DM, Raymond JS, Tisdale MA, and Quon DJ. 2008. Directly dated starch residues document early formative maize (Zea mays L.) in tropical Ecuador. Proceedings of the National Academy of Sciences 105(13):5006-5011.
Zemirline A, Vincent J-P, Sid-Ahmed S, Le Nen D, and Dubrana F. 2012. Lumbo-sacral malformations and spina bifida occulta in medieval skeletons from Brittany. European Journal of Orthopaedic Surgery & Traumatology:1-5.
Zhang G, and Dai C. 2001. Gene polymorphisms of homocysteine metabolism-related enzymes in Chinese patients with occlusive coronary artery or cerebral vascular

diseases. Thrombosis Research 104(3):187-195.
Zhu H, Wicker NJ, Shaw GM, Lammer EJ, Hendricks K, Suarez L, Canfield M, and Finnell RH. 2003. Homocysteine remethylation enzyme polymorphisms and increased risks for neural tube defects. Molecular Genetics and Metabolism 78(3):216-221.
Ziegler A, and König IR. 2010. A statistical approach to genetic epidemiology: Concepts and applications. Verlag: Wiley-VCH.