UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

VIRAL METAGENOMICS AND ANTHROPOLOGY IN THE AMERICAS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ANDREW T. OZGA
Norman, Oklahoma
2015

VIRAL METAGENOMICS AND ANTHROPOLOGY IN THE AMERICAS

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF ANTHROPOLOGY

BY

_____
Dr. Cecil M. Lewis, Jr., Chair

_____
Dr. Katherine Hirschfeld

_____
Dr. Paul Spicer

_____
Dr. Kermyt G. Anderson

_____
Dr. Tyrrell Conway

Acknowledgements

This dissertation would not have been possible without the guidance and encouragement of a number of professors, colleagues, friends, and family members. First and foremost I'd like to thank my advisor and chair, Dr. Cecil Lewis. His guidance and critiques throughout the project design, execution, and drafting process were invaluable to my completion of this degree. I'd also like to thank my doctoral committee: Dr. Tassie Hirschfeld, Dr. KG Anderson, Dr. Paul Spicer, and Dr. Tyrrell Conway for their support and insight regarding my project design and manuscript draft. I must thank those that graciously provided fecal samples for this project including the Cheyenne & Arapaho Nation, University of Oklahoma colleagues and students, and those from the coastal and jungle regions of Peru. Additionally, I'd like to thank my LMAMR colleagues including Dr. Christina Warinner and Dr. Jiawu Xu for their encouragement and valuable knowledge. This includes Dr. Alexandra Obregon-Tito and Raul Tito for their relentless support of my pursuit of a Ph.D., their assistance in sample processing, and their assistance in the Peruvian IRB process and sample collection. I am indebted to both Dr. Krithivas Sankaranarayanan and Allison Mann for their invaluable input regarding my data analysis. Their patience with me during the bioinformatics and statistics steps of this dissertation should be commended. I would like to thank my fellow graduate student cohort for their support and friendship over these last six years. You guys are the best.

I need to finally thank my family and friends for their constant encouragement throughout my graduate career and the writing of this dissertation. Thank you specifically to my grandparents Nan and Pete, my sister Emily, and my dad, all of

whom encouraged me to keep working towards my academic (and life) goals. Lastly, thank you to my mom, who continues to be an unending source of love and support.

Table of Contents

Acknowledgements iv

List of Tables viii

List of Figures ix

Abstract xi

Chapter 1 - On Being Human: The Treehouse of Life 1

Our Bacterial Self 4

Our Viral Self 19

Other Treehouse Members 29

Final Thoughts on being Human 31

Chapter 2 - Viruses in Anthropology 34

A Brief History of Viruses (and Microbes) 35

Viruses as Infectious Agents 38

An Anthropological History 40

Prions 41

Double Stranded DNA Viruses 42

Single Stranded DNA Viruses 55

Double Stranded RNA Viruses 59

Positive Single Stranded RNA Viruses 61

Negative Single Stranded RNA Viruses 67

DNA Double Stranded Reverse-Transcribing Viruses 72

RNA Single Stranded Reverse-Transcribing Viruses 72

Concluding Statements 80

Chapter 3 - The Geographic Structure of Gut Viruses 81

Population Background and Sampling Strategies 82

Laboratory Methods 90

Predicted Proteins and Geographic Structure 94

Amino Acid Sequences and Geographic Structure 99

Conclusions 102

Chapter 4 - CrAssphage Genomes in the Americas 103

CrAssphage: Lytic or Lysogenic? 111

Potential Host Microbiota for crAssphage 113

CrAssphage Abundance and Host Metabolic Output 120

Final Thoughts on crAssphage 123

Chapter 5: Conclusions 124

References 129

Appendix 165

List of Tables

Table 1: Viruses from raw sewage 53

Table 2: Information regarding virus enriched samples 87

Table 3: Information regarding unenriched virus samples from Oklahoma 88

Table 4: Information regarding unenriched virus samples from Peru 90

Table 5: CrAssphage abundance from unenriched data 107

Table 6: CrAssphage abundance from enriched data set 109

Table 7: Contig reads that match crAssphage and other microbes 113

Table 8: CRISPR spacer hits 115

Table 9: CrAssphage abundance and microbial abundance 117

Table 10: CrAssphage abundance and metabolites 120

List of Figures

Figure 1: Almann's Tree of Life 3

Figure 2: Phylogenetic tree example 3

Figure 3: Double stranded DNA viruses 42

Figure 4: Eppstein-Barr Virus 46

Figure 5: Single stranded DNA viruses 56

Figure 6: Double stranded RNA viruses 60

Figure 7: Positive single stranded RNA viruses 62

Figure 8: Negative single stranded RNA viruses 67

Figure 9: Double stranded reverse-transcribing DNA viruses 72

Figure 10: Single stranded reverse-transcribing RNA viruses 73

Figure 11: Lentivirus 75

Figure 12: Map of Peru 84

Figure 13: Map of Oklahoma 84

Figure 14: Sample Collection Kits 86

Figure 15: Photograph of Bacteriophage 94

Figure 16: Shared predicted proteins between samples 97

Figure 17: PCoA (PC1 & PC2) 98

Figure 18: PCoA (PC1 & PC3) 98

Figure 19: Maximum likelihood tree based on proteins sequences of OTU 2209 101

Figure 20: CrAssphage genome 105

Figure 21: CrAssphage diagram 106

Figure A1: Viral associations to the field of anthropology 165

Figure A2: Detailed viral applications and future directions 166

Abstract

Viruses are the most abundant and diverse biological entities on earth. As humans, we could not exist without them. For billions of years they have been the unseen mediators of microbial and multicellular life. They are in the air, water, dirt and within every living thing on the planet. Although they are not technically alive since all viruses need a host cell to replicate, they outnumber our own cells an estimated 100 to 1 and therefore cannot be undervalued in their importance to human health and biology.

Viruses are in close contact with humans from the moment we are born and as anthropologists we have commonly examined the cultural, social and political implications and responses of human viral infections. Yet there is so much untapped research potential when it comes to viruses. There are essentially five subject areas to discuss when addressing viruses from an anthropological perspective (Figure A1). First is that of host identity. Viruses are an integral part of who we are. This human-virus connection extends back hundreds of thousands of years, as these infectious agents are deeply ingrained within our genomes. In the first chapter I integrate literature of microbiology, virology and many others to posit that the traditional visual metaphor of the tree of life is inadequate and we would be better suited to describe each branch as its own treehouse built upon microbial and viral foundations. A second subject area is that of novelty. In other words, what the discoveries of novel viruses and understanding of understudied viral ecosystems can contribute to our definition of the human superorganism. In the second chapter I address novelty and future research by examining all of the current viral taxonomic information and their associations within the field of anthropology. Next generation sequencing of organisms capable of zoonotic

transmission could identify viruses that may eventually be the next great health risk to humans.

The third subject area that intersects viruses and anthropology is host history. In the past, viruses have been used to study complex human migratory patterns but this technique has yet to be fully realized within a viral metagenomic dataset. In chapter three, I use gut viral information from four populations in Oklahoma (Cheyenne and Arapaho) and Peru (Matses, Chincha, Tambo de Mora) in order to address the question of whether gut viral proteins are geographically structured. I hypothesize that DNA viruses are unique enough to the host that their sequences are closely tied to the geographic origin of the host. A fourth subject to be addressed is that of human health. Viruses obviously have a tremendous influence on human cells and microbial cells, but there is still much to be uncovered. In chapter four I use gut information from Oklahoma (Cheyenne and Arapaho, Norman) and Peru (Matses) to examine a recently discovered abundant DNA bacteriophage known as crAssphage. I biologically characterize the phage within the three populations by discerning its method of replication, potential microbial hosts, and its association with gut metabolites. I hypothesize that the abundance of crAssphage sequences found within the Cheyenne and Arapaho is not a result of their geography, but a host health association due to a decreased gastrointestinal health state.

The final subject that links anthropology and viruses is the future, which is referred to throughout the document and summarized in the conclusion section. This dissertation provides a thorough literature review, novel data analysis and potential

future directions, all of which are intent to stress the unexplored connections between viruses and the field of anthropology.

Dissertation Keywords: Viral metagenomics, Biological anthropology, Human health and biology, Tree of life, Geographic structure, crAssphage.

Chapter 1

On Being Human and the Treehouse of Life

Viruses manipulate our physiology, and their shifts are often the first indicators of a change, such as the onset of disease. We cannot understand microbial evolution without consulting viruses.
Forest Rohwer (Rohwer and Youle 2011).

If Charles Darwin reappeared today, he might be surprised to learn that humans are descended from viruses as well as from apes.
Robin A. Weiss (Weiss 2006)

Every person on the planet is family. Every historical figure, every celebrity, every skeleton entombed in every grave in the world is at a minimum, your genealogical cousin. When we examine an even more distant past, every animal, plant, bacteria, slime mold, protozoa and maybe every virus has ancestral connections to humans. Viruses are all part of an extended genealogy that reaches back billions of years. Genetic data is often used to characterize this extended genealogy as a phylogenetic tree. Overall, phylogenetic trees provide a fair depiction of our shared common ancestry. But there is more to this story. Phylogenetic trees force species to form exclusive groups known as clades. Yet, life has a level of extraordinary interconnectivity that is missing from these traditional "Trees of life".

In the halls of the Missouri State History Museum in St. Louis there are a series of art pieces on display donated by Katherine Dunham, a revered activist and anthropologist. One of the pieces is a sculpture flattened and cut from a 55-gallon steel oil container. The piece, produced in the 1940's by Almann, a Haitian artist, is fittingly titled "Tree of Life" (Figure 1) and depicts humans and animals seamlessly connected

to a sturdy tree with a small child nestled in its trunk. Instead of the isolated branches of a traditional phylogeny, Almann's tree is a maelstrom of interwoven branches, a complex thicket of radiating sticks and twigs. This artistic depiction holds truths that are under-represented by traditional phylogenetic trees (Figure 2), in part, because genetic information can be exchanged between different species and such events can drastically impact genome size, structure and content. Thus, the ancestry of life is more complex, networked and interlaced than the linearity implied by the bifurcations and exclusive groupings depicted in phylogenetic trees. One interpretation of Almann's work is that the depiction of a child nestled within the tree of life conveys a connection beyond ancestry, hinting at the critical role that members of the tree of life play in our own development. While it is well established that human biology is as much a product of genes as it is of environment, this message has never been so meaningful and deep as what has been recently brought to light in the study of our microbial self: those bacteria, viruses and other organisms that not only inhabit us, but are in part, shaping us as humans, mammals, animals and eukaryotes.

Figure 1: Almann's "Tree of Life" (Missouri State History Museum)

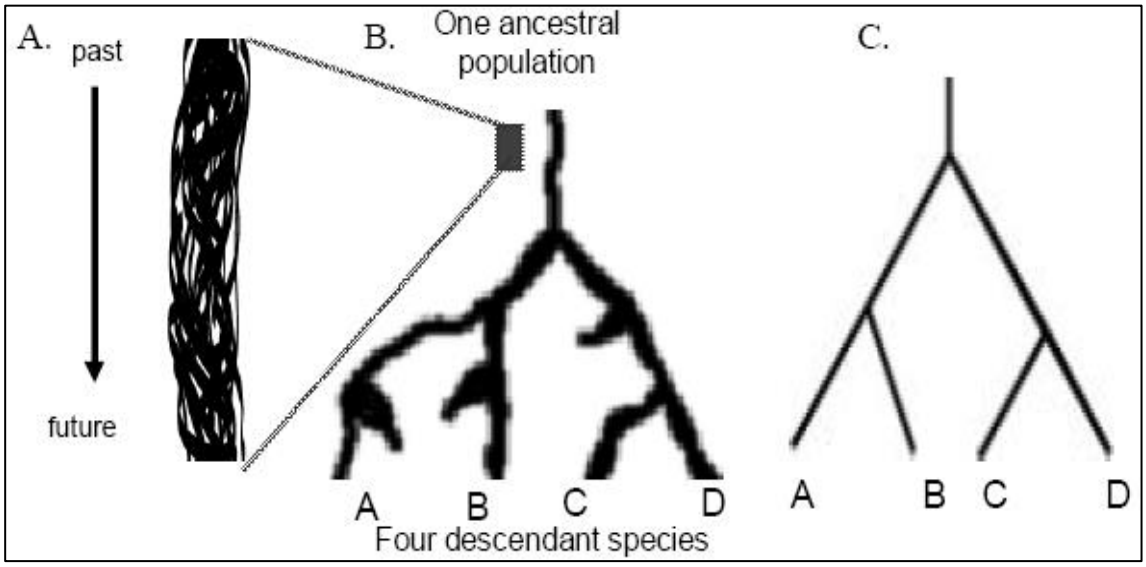

Figure 2: Example of a phylogenetic tree (Baum 2008)

Present here is a metaphor for human biology: the human treehouse of life, which well resembles Almann's tree. This metaphor is rooted not just in human metagenomics, but also panomics, which includes the genomics, proteomics and transcriptomics of all organisms on and within the human host. By drawing examples from fields of microbiology, virology and parasitology, we are able to paint a more holistic portrait of the unseen mediators of the human form. Here, I illustrate the many ways in which our neighboring tree branches influence and contribute to our own bodies. I begin the discussion with microbiota, followed by viruses and close with small eukaryotic coinhabitants. For each section, when possible, I discuss the densest human body site first, the gut, followed by the oral cavity and any other commonly studied sites. Essentially, we are superorganisms, covered head to toe, endoderm to ectoderm with bacteria viruses and microscopic eukaryotes. These microbes are not simply on us, they are part of us, deeply contributing to our own cells and genome. To think of it another way, the evolutionary branches of these organisms all contribute to the foundation, walls and roof of the human treehouse. In the same way that the tree in Almann's sculpture encapsulated organisms deeply within it, the human form is a mosaic of cellular and genetic information from countless organisms. Humans may be a separate species but they are never inseparable from the zig-zagging tree branch path of every organism on the planet.

Our Bacterial Self

All life as we know it has been classified into one of three domains, bacteria, archaea and eukaryote. Prokaryotes, both bacteria and archaea, are single celled

microorganisms that, unlike eukaryotes, lack a membrane bound nucleus. Prokaryotes initially evolved around 3.5 billion years ago, about a billion years after the earth was formed (Riding 2011). Their genomes are about 1,000 times smaller than those of eukaryotes and are organized into circular DNA rather than the multi-linear chromosomes common to eukaryotes. Prokaryotes are the most abundant life forms known today, making up arguably 50% of the earth's total biomass (Sleator 2010). They act as a huge reservoir of organic carbon in both aquatic and terrestrial ecosystems. The amount of carbon housed within prokaryotes on earth is estimated to be between 60 to 100% of the carbon found within all plant life (Whitman et al. 1998). This is in addition to the large amounts of nitrogen, phosphorous and other nutrients found within prokaryotes that are essential to life on the planet (Whitman et al. 1998). The variety and diversity of these prokaryotes is truly astonishing, so much so that the number of species cannot be properly estimated (there are currently around 8.7 million eukaryotic species) (Mora et al. 2011). Regardless of their total number of species, prokaryotes are plentiful and ubiquitous, inhabiting most every surface we come in contact with, every food that we eat, and every breath we take.

The tremendous amount of prokaryotic diversity is due partly to their ability to undergo horizontal gene transfer (HGT). First demonstrated in *Escherichia coli* in the 1940's (Tatum and Lederberg 1947), HGT involves the exchange of DNA between organisms. This process also occurs between eukaryotes such as parasites (Gilbert et al. 2010a), fungi (Richards et al. 2006), ducks (Kraus et al. 2012) and butterflies (Heliconius Genome 2012). There is growing evidence of eukaryote-prokaryote HGT through studies of photosynthetic sea slugs (Pierce et al. 2012), sea squirts (Nakashima

et al. 2004) and the unique Pacific White Shrimp, which receives genetic information both from bacteria and fungi (Yuan et al. 2013). Though the mechanisms are still unclear (Keeling and Palmer 2008) it seems that HGT is fairly ubiquitous spanning across the tree of life. It is not just the genes that transfer between organisms; HGT has a role in progressing speciation. It is an important diversification tool in the animal kingdom and is thought to be responsible for the rapid speciation seen during the Cambrian radiation (Syvanen 2012). However, genes are not the only thing being exchanged between living things, sometimes entire cells merge, and the fusion of prokaryotes is the likely cause of the emergence of eukaryotic cells.

Between 2.5 and 1.5 billion years ago, eukaryotic life forms evolved from prokaryotes (Davidov and Jurkevitch 2009). In fact, it is fair to say that our very cells were built up from prokaryotes. In this respect, the treehouse of life refers to more than just an inter-connected ancestry of microbes and humans but emphasizes the inseparability of the complex biology of all living things on the planet. To borrow a phrase, eukaryotes are chimeras (Martin and Mentel 2010). You do not have to look much further than your own cells to see this in action. In almost every human cell there are organelles known as mitochondria. Mitochondria generate much of the cell's energy, adenosine triphosphate (ATP) and contain genetic information not found within the nuclear genome of the cell. Although a few eukaryotic organisms such as *Giardia* spp. (Roger et al. 1998) and microsporidia (Hirt et al. 1999) have organelles that evolved from mitochondria known as mitosomes, in general, mitochondria are considered to be "just as defining and ubiquitous among eukaryotes as is the nucleus itself" (Martin and Mentel 2010). Mitochondria are small, around 0.5 to 1 micrometer in

size, and somewhat resemble bacteria. This is not a coincidence. Mitochondria evolved from free-living bacteria through symbiosis within the eukaryotic host cell (Margulis 1970). The hypothesis of such a relationship is attributed to Konstantin Mereschkowski, a Russian botanist (Martin and Kowallik 1999). Mitochondria were once free living prokaryotes, and around 2 billion years ago, they merged with an aerobic host forming an endosymbiotic relationship (Sagan 1967). Chloroplasts, the light harvesting organelles in plants, arose through a similar endosymbiotic relationship with cyanobacteria and a eubacterial cell (Schimper 1883). Endosymbiosis is still seen today, a recent case being *Paulinella chromatophora*, a unicellular eukaryote that developed photosynthetic properties derived from integrated cyanobacteria (Nakayama and Archibald 2012; Yoon et al. 2009). The merging of organisms that came along with endosymbiosis shook the tree of life to its core, completely transforming the evolutionary trajectory of living things. For one such example I revisit the human mitochondria, which fit within the Rickettsiaceae bacterial family. Interestingly, this family also includes *Rickettsia sp.*, which is a common genus in macroparasites that cause typhus, among other diseases, in humans (Emelyanov 2001). To emphasize mitochondria, which are an essential part of cells, also belong to the same evolutionary family as typhus, a disease that has afflicted and killed millions for five centuries.

The relationship between free living microbes and our mitochondria has remained surprisingly intimate over time. Consider for one, butyrate metabolism. Butyrate is also known as butyric acid, a short chain fatty acid (SCFA) found within the colon. Butyrate decreases colorectal cancer risk (Scheppach et al. 1995) and acts as an epigenetic mechanism stimulating T cell production (Furusawa et al. 2013).

Importantly, butyrate is essential for colon function through its regulation of energy metabolism and in its absence, the epithelial cells of the colon become starved and undergo self-digestion (Donohoe et al. 2011). In this way, butyrate is essential for survival. Yet, human cells are unable to generate butyrate. Butyrate is actually generated from the fermentation of dietary fiber by free living bacteria that reside within the gut. These bacteria feed on the fiber and convert it into butyrate (Ritzhaupt et al. 1998). While some bacteria can also metabolize the butyrate for further energy, the mitochondria, within the eukaryotic cells, also perform this task. Thus, there is a deep connection between human cells and microbes: a connection vital to human health.

The 2007 National Institutes of Health's Human Microbiome Project (HMP) (Group et al. 2009; Turnbaugh et al. 2007) highlighted the need to better understand the relationship between health and microbial ecologies. The HMP emphasized new molecular approaches to understanding microbial systems. The term *microbiome*, coined in 2001 (Lederberg and McCray 2001), originally referred to pathogenic and commensal microorganisms within a host, essentially indistinguishable from the phrases 'host microbiota' or 'host microbial ecology.' Today, the term microbiome more frequently refers to characterizing any microbial ecology using molecular methods, such as the Earth Microbiome Project (Gilbert et al. 2010b). Here I refer to microbiomes by their original definition, a host associated microbial ecology.

The number of microbes present on a single human host is astronomical. While the human body is composed of an astounding 37 trillion human cells (Bianconi et al. 2013), there are ten times as many bacterial cells (Group et al. 2009). That comes out to almost 400 trillion bacterial cells. That is a pretty daunting, difficult to conceptualize

number. We can use astronomy to help us. Think of the Milky Way galaxy. There are around 300 billion stars within our galaxy alone. That means it would take 1000 Milky Way galaxies in order for the number of stars to equal the number of bacteria in a typical healthy human body. If you were to line up all of the bacteria on a typical healthy human body (an average of 1 micrometer in size), the total length would be 400,000,000 meters, a little under 250,000 miles. That's long enough to wrap around the earth ten times!

If we use the call as a unit of analysis, our bodies are only 10% human and 90% bacteria. One might ask that since our bodies are 90% bacteria, why do we not look like a giant slimy biofilm? Fortunately, the surface of the human skin is less covered by bacteria than other body sites such as teeth, the GI tract and vagina, which in par prevents us from resembling swamp things. What also needs to be considered is that bacteria are very small, only 1 micrometer across on average. That is one millionth of a meter. Human cells are over ten times this size. Overall, bacteria make up roughly 2% of your total body mass (O'Hara and Shanahan 2006). This 2% is still sizable within your body, the equivalent size of your brain, liver or two pairs of lungs, and your microbiome is as important as those bodily organs because without these bacteria, you would be dead.

Human microbiomes are implicated in a wide range of metabolic, immunological and developmental processes including host metabolism (Tremaroli and Backhed 2012), vitamin production (LeBlanc et al. 2013), education of the immune system (Hooper and Macpherson 2010), and defense against infection (Brotman 2011). Microbiomes mitigate colon function (Hamer et al. 2008), impact inflammation

(Kamada et al. 2013) and cancer (Bultman 2014), and have been associated with obesity and diabetes (Devaraj et al. 2013), anxiety and depression (Foster and McVey Neufeld 2013), oral health (Aas et al. 2005), cardiovascular disease (Nakano et al. 2009) and many aspects of vaginal health (Brotman 2011). Microbiomes can be manipulated by antibiotic treatments (Panda et al. 2014), prebiotics, probiotics and diet (Chen et al. 2014; da Silva et al. 2013), time of the year (Davenport et al. 2014), a more industrialized Western or a more traditional hunter gatherer lifestyle (Obregon-Tito et al. 2015) and host global geography (Lin et al. 2013; Prideaux et al. 2013; Tyakht et al. 2013; Yatsunenko et al. 2012). In short, these microbes provide the walls, floor and ceiling of our treehouse. When our microbiome is out of balance from its core bacterial community, our treehouse is altered, shoving the human host into an overall dysbiotic state (Clemente et al. 2012).

Our dependence on our microbial coinhabitants begins shortly after birth (Dimmitt et al. 2010; Koenig et al. 2011). It was originally assumed that humans, like most mammals, are sterile at birth but recent look at the microbiome of the placenta show a microbial composition are more similar to the oral cavity of the mother than the skin or vagina (Aagaard et al. 2014). Our first big exposure to our treehouse coinhabitants is tied to the method of childbirth; a vaginal birth contributes to a richer infant microbiome when compared to a Caesarian section (Dominguez-Bello et al. 2010). Shortly after birth, breast-feeding or bottle-feeding activities will continue shaping the microbiome (Stark and Lee 1982). These early years are incredibly dynamic, usually until the age of three (Avershina et al. 2014; Yatsunenko et al. 2012), at which point the juvenile microbiome is most influence by the microbes of their

parents (Ley et al. 2008). Microbiome changes continue during the juvenile and adolescent years (Lan et al. 2013) until reaching an equilibrium in adulthood. At around age 60, the time at which the human immune system begins to decline, overall microbial diversity begins to decrease (Mariat et al. 2009). A decrease in species diversity at any age is a signature associated with inflammatory bowel disease (Cucchiara et al. 2009), anti-inflammatory drug usage (Makivuokko et al. 2010) and *Clostridium difficile* infection (Hopkins and Macfarlane 2002). Although the immune system efficiency is decreased at a later age due to the decrease in microbial types, it is important to note the role which microbes played in developing the first line of pathogen defense in humans.

As human bodies grow and mature, our microbial inhabitants help to educate our immature immune system (Hooper and Macpherson 2010; Lathrop et al. 2011). These early years are crucial to proper immune function and any severe perturbations in the gut microbiome are associated with the later development of obesity and atopic diseases including food allergies (Bisgaard et al. 2011), asthma (Gilstrap and Kraft 2013) and eczema (Nylund et al. 2013). Much of our knowledge of the impact of microbiota on human health comes from studies of the gut, so I will devote a small section to synthesizing some of the current knowledge. The human gut, mainly the colon, is the most studied site within human microbiome investigations and considered to be the most densely populated natural bacterial ecosystem (Frank and Pace 2008). The human gut contains somewhere in the realm of $10^{13\text{-}14}$ microorganism coinhabitants (Backhed et al. 2005), which are partitioned into at least 5,000 species (Handelsman et al. 1998), and contribute at least one million non-redundant bacterial genes to the host

(Tlaskalova-Hogenova et al. 2011). Despite this diversity, most of our gut bacteria, roughly 90%, belong to just two phyla: Firmicutes and Bacteroidetes (Ley et al. 2006). Within these phyla there is still a good deal of variation, so much so that everyone has their own 'core' microbiome upon reaching adulthood (Faith et al. 2013; Qin et al. 2010), even monozygotic and dizygotic twins (Turnbaugh et al. 2009; Turnbaugh et al. 2010). Currently, the adult 'core' microbiomes are partitioned into one of three gut 'enterotypes,' which are dictated by the overabundance of *Bacteroides* spp., *Prevotella* spp., or *Ruminococcus* spp. (Arumugam et al. 2011). The exact underpinnings of these enterotypes are still unknown, since they do not seem to be correlated with BMI, sex, age or geography (Arumugam et al. 2011). The differences in these abundances may actually be due to analysis methods (Koren et al. 2013; Lozupone et al. 2012) and some suggest an 'enterogradient' description better describes the differences in human gut flora (Faust et al. 2012). Alternatively, a four community type model, which is slightly more malleable, has been posited to replace the enterotype designation (Ding and Schloss 2014). Unlike the enterotypes hypothesis, this model suggests that a number of life-history characteristics, including associations with one's level of education, influence the overall gut community structure (Ding and Schloss 2014). This life-history applies not only to the gut but most other body sites, which may lend the way to personalized therapies (Ding and Schloss 2014). Although the appropriate categorization for gut microbiota is still a topic of debate, the consensus seems to be that the gut is at least moderately affected by host diet (De Filippo et al. 2010), both short term (David et al. 2013) and long term (Candela et al. 2012). The foods you consume during childhood are not just important to your own health, but the health of

your 400 billion gut microbes! Children that are not able to get enough food during adolescence may develop severe acute malnutrition syndromes such as marasmus and kwashiorkor (Smith et al. 2013), which too have an altered gut microbiome component.

The bacteria within the human gut truly are unseen mediators of human health. They play a role in host nutrition by producing vitamins such as folate, B_{12}, biotin and vitamin K (Bentley and Meganathan 1982). Bacteria help protect their host organism by reducing mucosal cell turnover, increasing digestive enzyme activity, reducing overall acidity and providing a protective barrier of the epithelial cells of the colon (Canny and McCormick 2008; Horz and Conrads 2011; Macdonald and Monteleone 2005). One of the beneficial roles of bacteria within the gut is the production of short chain fatty acids (Berg 1996). These SCFAs, such as butyrate, propionate and acetate are produced through a fermentation of dietary fiber within the small intestine, have anti-tumorigenic properties and assist in cell proliferation and differentiation within the colon (Albenberg and Wu 2014; Berg 1996; Frankel et al. 1994).

Although microbes are great overseers of our wellbeing, they can also wreak havoc on overall host symbiosis at any moment. A swath of diseases have been correlated with microbial gut diversity in recent years including diabetes (Creely et al. 2007), liver disease (Brun et al. 2007), rheumatoid arthritis in both children and adults (Toivanen 2003), and cancer (Scanlan et al. 2008). The role of bacteria in obesity and inflammation has been thoroughly examined (Cox et al. 2014). Reduction of certain microbial types, for example Bacteroidetes, was associated with obese mice (Ley et al. 2005). In humans, too, obese phenotypes are associated with lower Bacteroidetes and higher Firmicutes (Ley 2010). Obesity treatments such as gastric bypass surgery

tremendously alter the host gut microbiome. When the microbiome from this altered gut is placed into mice that did not undergo surgery, the mice show similar rapid weight loss (Liou et al. 2013). Although not the only factor in obesity and inflammation, microbes appear to play and underappreciated role in autoimmunity. Microbiome transplants for humans are now being seriously considered as the next step in aiding in host GI health.

One of the most concerning gut treehouse invaders to overall gut health is *Clostridium difficile*, which causes an infection so severe, that surgical intervention is sometimes necessary. *C. difficile* infection (CDI) is responsible for 14,000 U.S. deaths a year and instead of surgery, many are turning to fecal transplantation as the preferred method of treatment (Di Bella et al. 2013). Patients with CDI show an overall reduction of Bacteroidetes and Firmicutes and increased Actinobacteria presence (Chang et al. 2008; Khoruts et al. 2010). However, within just 2 weeks of a transplant, the bacterial ecosystem of the gut is completely normalized (Di Bella et al. 2013). Even more startling is the fact that fecal transplantation has a 91% success rate (van Nood et al. 2013), an example in which host microbiome transplantation is a viable option for disease intervention.

The gastrointestinal tract harbors a wide array of both beneficial and detrimental microbes existing in a delicate balance with one another. Any disturbance to this balance can transform a well-functioning human ecosystem into one with severe dysbiosis. So what is the 'best' bacterial gut ecosystem? Gut microbiomes vary between individuals so there are likely many different healthy and diseased states. Researchers have begun to look at fossilized fecal material, known as paleofeces or coprolites, as a

way of determining early human gut microbiome states. Several studies examined paleofeces (coprolites) from a cape in Durango, Mexico (Tito et al. 2008), and mummified remains from Casarones, Chile (Tito et al. 2012) in order to provide unique perspectives in ancient health. While Rio Zape (Mexico, 1,400 years B.P.) appeared similar to modern human feces, Hinds cave (Chile, 8,000 years B.P.) was starkly different than modern feces in its microbial makeup (Tito et al. 2012; Tito et al. 2008). This drastic shift is likely due to the a transformation of the human condition in cosmopolitan populations (Tito et al. 2012). Future studies of coprolites may be help us better understand the changing relationship we as humans have with our treehouse coinhabitants, and the underlying health implications those changes signify.

Next to the gastrointestinal tract, the oral cavity is the second most populated microbial site on the body. There are thought to be 10^{12} bacteria within the mouth (Tlaskalova-Hogenova et al. 2011) belonging to some 700 different species (Zarco et al. 2012). Bacteria can colonize every part of the oral cavity: the soft tissues like the cheek, soft palate and tonsils, and the hard surfaces of the teeth (Zaura et al. 2009). On a slightly grander scale than the gut, the oral cavity shows a large amount of interpersonal variation (Mason et al. 2013). Although it is sometimes difficult to specifically discern 'healthy' microbes from 'unhealthy' microbes, we do have knowledge of specific bacterial taxa being linked to increased disease risk (Aas et al. 2005). For example, poor oral hygiene can cause the buildup of *Streptococci* which can enter into the bloodstream and lead to endocarditis (Paik et al. 2005). Additionally, overall reduced dental care also linked to an increased risk of diabetes, cardiovascular disease and oral cancers (Zarco et al. 2012).

Dental plaque and its fossilized form dental calculus, are multispecies biofilms that act as rich sources of genetic information for both modern and ancient human samples (Adler et al. 2013; Parahitiyawa et al. 2010; Warinner et al. 2014). The microbes within modern dental plaque, like saliva, have been associated with both healthy and diseased states (Peterson et al. 2013). Plaque is ubiquitous across the globe and across the fossil record, leading to a unique snapshot of the health and diet (Henry et al. 2011) of ancient peoples. Much in the way coprolites can provide microbial information for the ancient human gut, calculus can provide information on periodontal disease and the presence of low-level antibiotic resistance genes in the oral cavity (Warinner et al. 2014). These approaches provide an unprecedented view of ancient diet and disease.

The microbes within the oral cavity are not the only ones that dictate the microbiome of the human respiratory tract (Morris et al. 2013). Though little is known regarding the complexity of the microbiome of the lungs, breathing tests and sputum examinations have shown a clear difference in microbiota between cases and controls for respiratory illnesses (Marsland et al. 2013). A number of other studies have associated lung microbe diversity with cystic fibrosis, idiopathic bronchiectasis, chronic obstructive pulmonary disease and HIV infections (Losada et al. 2011). Smoking has an influence on microbial composition, but does not significantly alter diversity (Morris et al. 2013).

We cannot ignore the largest organ on the body when discussing microbes: the skin. Within minutes of birth, the once sterile fetal skin is colonized by millions, possibly even billions of bacteria (Dominguez-Bello et al. 2010). Our microbes are

differentially distributed across our skin, with as many as 10^7 in the moist areas and substantially less in drier areas (Leyden et al. 1987). The same four common phyla that make up most of the oral and GI tract also make up most of the skin microbiome, just in different proportions: Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria (Sanford and Gallo 2013). Despite these common phyla, bacterial types between individuals are thought to be so unique that we may be identifiable based solely on our 'microbial fingerprint' (Fierer et al. 2010). In fact, the skin microbiota can tell more about a person's lifestyle than one may assume. The people (and pets) that you cohabitate with (Song et al. 2013), the occupation you choose (Meadow et al. 2013), your indoor environment (Kembel et al. 2012; Kembel et al. 2014; Meadow et al. 2014), and the makeup you apply (Holland and Bojar 2002) all influence your skin microbial communities. Your skin microbes may even be why mosquitoes are more attracted to some individuals over others (Verhulst et al. 2011). Skin microbiota are thought to be associated with atopic dermatitis, psoriasis, rosacea, acne (Chen and Tsao 2013) and even the development of breast cancer (Xuan et al. 2014). The microbiota on skin surfaces are less abundant on average than other body sites like the colon and oral cavity, but is still a dynamic system which impacts human health and disease.

The human microbiome is not only more diverse and dynamic than previously imagined, it is also present in body sites that we thought once to be sterile. For example, the urinary tract of both men and women showed significant microbial diversity between women (Siddiqui et al. 2011) and slightly less diversity between men (Lewis et al. 2013). Within men, we also see a difference in microbes based on circumcision (Mandar 2013), sexual health and age (Nelson et al. 2012). Although no microbial

fertility studies using both partners have ever been undertaken, bacterial composition of the seminal fluid appears diverse, sharing common microbiota with host urine and the vaginal cavity (Hou et al. 2013). The vaginal microbiome appears to be fairly stable over time (Ding and Schloss 2014) and host specific (Ravel et al. 2011), but can be altered by events such as pregnancy (Romero et al. 2014) and viral STDs (Nardis et al. 2013). Thus, there remains much to explore with respect to these urinary and reproductive microbiomes (Gonzalez et al. 2011).

Microbiomes may have roles that are not so easily tied to specific body sites, as we can see in several examples from non-human microbiome studies. Microbiota can protect bees from intestinal parasites (Koch and Schmid-Hempel 2011), affect mate choice in fruit flies (Sharon et al. 2010), differentiate hyena groups through odors (Theis et al. 2012), induce larval settlements in polychaetes (Huang et al. 2012) and even influence brain development and behavior in mice (Heijtz et al. 2011). Mice models have shown that the presence or absence of particular gut bacteria affect depression and anxiety by way of an increase in brain-derived neurotropic factor (BNDF) (Neufeld et al. 2011) and the production of certain proteins may be linked to altered feeding and emotion in humans with eating disorders (Tennoune et al. 2014). Mouse models aside, there is no direct evidence of such findings in humans until longitudinal studies herald more data (Foster and McVey Neufeld 2013). Nevertheless, it is biologically feasible that the human microbiome impacts human behavior.

Microbiomes do not reflect simple bystanders within the host ecosystem but are rent paying members of the human treehouse of life. We humans could not survive without our bacteria. Outside of our bodies bacteria purify our air, provide renewable

energy and destroy pathogens by way of antibiotics. Bacteria assist in producing our food: dairy products like yogurt and cheese, fermented foods like soy sauce and pickles and yeast products like bread, beer and wine. These microscopic critters are our immune system mediators, our trash decomposers, our nutrient providers. Bacteria have performed these thankless tasks for billions of years in practically every multicellular creature currently known. Thanks to advances in biotechnology, we now understand that they permeate the entire tree of life, not merely existing as a secluded branch, but as interwoven members that mediate much of eukaryote biology. But who mediates the mediators? The next section goes even deeper into the tree, examining the even more invisible players of eukaryotic and prokaryotic biology.

Our Viral Self

The exact origin of viruses on earth is still shrouded in mystery. Multiple origin hypotheses for viruses exist but to confound matters further, they are not classified as living things and cannot be placed in the tree of life. Despite their omission from the tree, viruses simply cannot be ignored in our discussion of the human treehouse. However, one thing is clear: viruses are everywhere. If we step back to look at the entire planet, there are an estimated 10^{31} viruses (Hendrix 2005). Ten quintillion! Of those, 10^{30} of them inhabit the ocean (Suttle 2007). If we zoom in to just ocean viruses, we see mostly bacteria infecting viruses known as bacteriophages (or phages). When laid end to end, these ocean phages would stretch for 100,000,000 light years, past the nearest 60 galaxies (2011; Suttle 2007). On both land and in the water, these viruses are extremely influential to the ecosystems in which they exist. In the same way that bacteria are

responsible for a large chunk of carbon in most ecosystems, it is estimated that ocean viruses can turn over as many as 150 gigatons of carbon per year, 30 times more than marine plankton (Suttle 2007). Common oceanic bacteriophages were shown to contribute photosynthetic proteins to ocean bacteria and, considering the percentage of O_2 in the atmosphere created by autotrophic oceanic bacteria, you can thank a virus for one out of every ten breaths you take (Suttle 2007; Zimmer 2011).

Viruses, like microbes, completely permeate our human ecosystem. As noted earlier, for every human cell there are roughly 10 times as many bacterial cells. This is enhanced with viruses. For every single bacteria within humans there estimated to be between 10 to 100 bacteriophages (Chibani-Chennoufi et al. 2004). One could then argue that based on this information, we are less than 1% human! We now know that bacteria have an important underlying role in human health, and it seems that bacteriophages and viruses too have a largely underestimated role (De Paepe et al. 2014). Whether it is the role of the phage in spreading bacterial antibiotic resistance genes or its role in bacterial carbohydrate utilization within the human gut (Reyes et al. 2012), viruses are influential to human hosts. Their complex role has just yet to be fully understood. In this section, I examine the current knowledge regarding viral diversity within our genome and within our bodies.

Viruses have existed alongside multicellular organisms for billions of years: replicating using their machinery, trading genetic material and, in some cases, destroying them. It has been suggested that they are responsible for first dividing of the three domains of life (Claverie 2006). Taxonomically, viruses can possess DNA or RNA genomes, which can be either double or single stranded. As referenced at the

beginning of the section, there are a number of hypotheses as to how viruses arose on earth including the progressive, regressive and virus first (Bamford et al. 2005; Koonin and Martin 2005; Wessner 2010). The first two hypotheses posit that viruses began as mobile genetic elements and entered host genomes or escaped from them to begin a parasitic free-living existence, while the third suggests that RNA viruses predate the last universal common ancestor between the three domains (Forterre 2010; Villarreal 2004). Regardless of the manner by which viruses arose, they have profound effects on all organisms they come in contact with, so much so that some scientists are adopting a 'virocentric view' of life's evolution (Koonin and Dolja 2013) which states that viruses have a key role in dictating the presence of bacteria in a given environment and the genomes of upper level organisms (Mills et al. 2013). Viruses have an even deeper role in dictating human health, through their integration into our own genomes.

Much like bacteria, the human genome was in part built upon viral foundations. When we consider the entire human genome, very little of it is 'uniquely human'. The human genome contains 3 billion base pairs, but only 1.5% of that, roughly 20,000 genes, actually code for proteins (2001). Of the other 98.5% of the human genome, 46% is comprised of transposable elements (with potential viral origins), while the origins of the other 52.5% still unknown (Cordaux and Batzer 2009). In fact, not only do humans have a potentially huge chunk of their genome arising from viruses, most eukaryotic genomes have substantial endogenous viral elements (EVEs) as part of their genomic composition (Feschotte and Gilbert 2012). These elements are ancient viruses that have been integrated into the genome as far back as the Late Cretaceous period (Belyi et al. 2010a). Certain viral types are more prone to integration than others. For example, in a

survey of vertebrates, 19 were shown to possess RNA viruses from the order Mononegavirales within their genomes, which likely arose as a result of pathogen resistance at least 40 million years ago (Belyi et al. 2010b). Endogenous retroviruses (ERVs) within humans are referred to as HERVs (human ERVs). There are an estimated 31 known families of HERVs present within the human genome, arising from 31 separate integration events (Katzourakis et al. 2005). The HERV origins predate our species, harkening back to the time of our primate ancestors some 30 million years ago (Emerman and Malik 2010).

The incorporation of viruses into the human genome has been shown in some cases, to be beneficial. The salivary amylase genes, which aid in the breakdown of starch into sugar in the mouth, were derived from the duplication of an ancestral pancreatic amylase gene, which arose from a retroviral insertion responsible for tissue-specific expression (Meisler and Ting 1993). Another classic example is the human placenta. The HERV-W family produces proteins known as syncytins, which help to form the placental synctiotrophoblast layer during early embryogenesis (Sugimoto and Schust 2009). Despite their importance in human development, ERVs are not always beneficial to the human host. An estimated 20% of cancers are due to exogenous retroviruses (Weiss 2001). Specifically, insertions from the HERV-K family play a role in transforming melanoma cells into malignant tumor cells (Serafino et al. 2009). Other HERVs may also be responsible for some autoimmune disorders like rheumatoid arthritis, multiple sclerosis and systematic lupus erythematosus (Katoh and Kurata 2013).

The mere presence of a particular virus or phage in humans does not translate to a health state: many viruses are shared between ill and healthy individuals (Pennisi 2011). One in 12 people, or 500 million worldwide are living with chronic viral infection (2011) and at any given time, the average human is infected with between 8 and 12 distinct viruses (Virgin et al. 2009). Some of these viruses, such as JC polyomavirus (JCV) (Shackelton et al. 2006), BK polyomavirus (BKV) (Zhong et al. 2009) and cytomegalovirus (CMV) (Stoner et al. 2000), are so prevalent that they are estimated to infect most of the global human population at any given time. These viruses along with others can cause persistent infections in humans and have evolutionary implications an important role in human evolution. Many persistent infectious viruses are known in humans: herpes simplex virus (HSV1, HSV2), Epstein-Barr virus (EBV), adenoviruses, human polyomaviruses, human papillomaviruses (HPV), and TT viruses (Villarreal 2004). Anthropologists have used these viruses as markers for tracking human migration, but they also have roles within host immunity. Studies have shown that AIDS occurs more slowly in patients that have been infected with the human cytomegalovirus (King et al. 2006) or non-pathogenic hepatitis G virus (Heringlake et al. 1998; Tillmann et al. 2001). There is so much more to learn about viruses within humans, especially when we consider that there an estimated 320,000 different viral types that infect mammals (Anthony et al. 2013).

This does not mean that every virus has a direct influence on the host, in fact, a majority of viruses within humans are temperate phages. If we were to look at specifically tailed bacteriophages, which make up a huge chunk of viral presence within humans, we would see that only 1% of them are virulent (Mathieu and Sonea 1995).

Our understanding of bacteriophages is still very preliminary, 95% of all documented phages belong to a single order: Caudovirales (Shen et al. 2012). Virulent phages tend to adhere to what is known as kill the winner dynamics, in which a single microbe with the highest density in an ecosystem is specifically targeted for infection (Thingstad and Lignell 1997). Most phages are not host microbe specific, for example certain cyanophages can infect multiple genera (Sullivan et al. 2003). Thus, it is possible that phages lyse nearby bacteria that are not their common host, a model known as kill the relative (De Paepe et al. 2014). During this attack, microbes do not sit idly by and allow themselves to be destroyed; there is an incredible evolutionary arms race between phages and bacteria in every ecosystem. We are getting a closer look at this constant microscopic struggle examination of CRISPRs (clustered regularly interspaced short palindromic repeats), their associated genes and their nearby spacers (Ishino et al. 1987; Jansen et al. 2002). CRISPR sequences relate to the immune system of microbes. When attacked by phages, the microbial hosts sometimes integrate phage DNA, which exist as spacers between repeating elements, and prevent future infection from that particular bacteriophage. This unique system also has applications for humans. CRISPRs can actually be used to fix a broken gene or activate a disabled one within our own cells (Pennisi 2013). In order to fully understand viral and phage applications to human disease, we must first learn more about the identity and roles of these treehouse compatriots.

This leads us to an exploration of the viral ecology within humans, typically referred to as the human virome. Viromes are recognized as an important counterpart to studies of the human microbiome. The first viral metagenomic studies of humans

focused on the gut, and yielded 1,200 viral genotypes identified mostly as prophages and siphophages, suggesting that human gut bacteria are heavily colonized by temperate bacteriophages (Breitbart et al. 2003). As with the microbiome, the densest ecosystem of viruses is the gut, making it a good starting point for virome discussion. During infancy and early childhood, humans undergo a period of viral volatility (Breitbart et al. 2008). During this time even healthy children test positive for pathogenic viruses such as human enteroviruses (HEV) (Krupovic et al. 2011) and human parechoviruses (HPeVs) (Cordaux and Batzer 2009). Into adolescence and adulthood, the gut virome communities become more stable with 80% of viral forms persisting over a 2.5 year period (Minot et al. 2013). Interestingly, overall virome diversity does not a show a strong association with genetic relatedness, even in monozygotic twins (Reyes et al. 2010). Overall, gut viral communities are more diverse than microbial communities and show a high degree of interpersonal variation and a malleability to host diet (Minot et al. 2011; Reyes et al. 2012). We acquire many transient viruses through our diet, for example the pepper mild mottle virus (Zhang et al. 2006) and many from the Circoviridae family, which are common insect viruses (Li et al. 2010) found within human feces.

A majority of the gut virome of an individual is dictated by two components: a persistent small portion of the common global virome and the rapid evolution of the other, longer term members (Minot et al. 2013). This was evidenced by the viral equivalent of microbial gut enterotyping (Arumugam et al. 2011; Ogilvie et al. 2013). The five viral enterotypes (A,B,C,D, and UC (unclassified)) are based on the relative abundance of *Bacteroidales*-like phages (Ogilvie et al. 2013). Viral enterotypes are

related to gut enterotypes because both *Bacteroides* and *Prevotella*, the dominant genera in the first two enterotypes, are regulated by *Bacteroides*-like phage types (Arumugam et al. 2011; Ogilvie et al. 2013). Although viral perturbations within the gut adversely affect the health of the host (Haynes and Rohwer 2011), studies of viral enterotypes and the overwhelming presence of CRISPR phages in the gut microbiota suggests a complicated relationship (Stern et al. 2012) that is likely related to host dysbiosis (Haynes and Rohwer 2011). This dysbiosis was recently examined in a gut virome study of sufferers of inflammatory bowel diseases such as Crohn's disease (CD) and ulcerative colitis (UC) (Norman et al. 2015). The authors found an abnormal enteric virome in CD and UC compared to healthy controls, pushed by the overabundance of Caudovirales bacteriophages (Norman et al. 2015). Bacteriophages clearly have a mediated role in gut health but determining exactly how differing abundances contribute to host dysbiosis is still an unanswered question.

The question of dysbiosis may be better addressed with an investigation of ancient gut viruses. Although difficult methodologically, researchers have been able to extract diverse bacteriophages from a 14^{th} century Middle Age Belgium coprolite (Appelt et al. 2014). A future study of coprolite CRISPRs would help us better understand the ongoing evolutionary arms race within our bodies. Yet the gut is not the only body site experiencing the evolutionary arms race between phage and microbe. Virome studies involving the oral cavity, the lungs and the skin have shown a plethora of viral diversity.

Viruses are plentiful within the human oral cavity, with an estimated 10^8 particles per milliliter of saliva (Pride et al. 2012a). The mouth is dominated by

bacteriophages, which are distinct from those in gut (Pride et al. 2012a). Similar to the gut virome, CRISPRs are incredibly abundant and shared between related individuals (Robles-Sikisaka et al. 2013). This signifies another body site of constant microbial and viral contact and interaction. Humans do not seem to share a 'core' salivary virome (Pride et al. 2012a). In fact, many of the viruses within the oral ecosystem are personalized and sex specific (Abeles et al. 2014), maintaining a stable presence for a several month time period (Pride et al. 2012b). In the oral cavity, viruses can also be influenced by what you eat (Willner et al. 2011). An example is the presence of the *Stretococcus mitis* phage SM1, which can be acquired through soy sauce, nicotine and white wine consumption (Willner et al. 2011). The oral cavity is also home to many pathogenic viruses including Human immunodeficiency viruses, papillomaviruses and Epstein-Barr viruses (Michaud et al. 2013; Pavesi 2004). The presence of some viruses in the mouth, EBV, is even associated with an increase in rick of oropharyngeal carcinomas (Michaud et al. 2013). When we examine subsections of the mouth, say supragingival and subgingival plaque from teeth, we see unique viruses, both temperate and pathogenic, when compared to the oral cavity (Ly et al. 2014). We also see significant differences in the presence of certain viral families between healthy individuals and those suffering from periodontal disease (Ly et al. 2014). We can observe viruses in the fossilized form of plaque as well. Viruses from the Anelloviridae family were detected in dental pulp samples from Napolean's great army (200 years B.P.) (Bedarida et al. 2011) and a human T cell lymphotropic virus type 1 (HTLV-1) was isolated from an Andean mummy (1,500 years B.P.) (Li et al. 1999) . This biofilm

can persist for thousands of years, allowing us to get a glimpse of ancient human migrations and extant infections.

The air we breathe is packed with viruses: between 1.6 million and 40 million per square cubic meter (Whon et al. 2012). As humans, we breathe .01 cubic meters of air each minute, which is the equivalent of inhaling about several hundred thousand viruses every sixty seconds (Whon et al. 2012; Zimmer 2011). But there is no need to don protective medical facemasks, these viruses are not like those seen in the movie *Outbreak* (Peterson 1995)but mostly derived from plants and insects (Whon et al. 2012). Granted, those airborne viruses that are virulent like rhinoviruses, influenza and respiratory syncytial viruses (hRSV), can contribute to host morbidity and mortality (Lysholm et al. 2012), they still make up only a small minority of the total airborne viruses. Yet variety of viral types is not necessarily a bad thing. Higher viral diversity was found within the lung of healthy subjects compared to their diseased counterparts (Lim et al. 2012). With all of these viruses coming in contact with human skin, it is easy to see why there is such complex viral flora on the skin of healthy individuals (Foulongne et al. 2012). Although there have been no complete virome studies of the urinary tract or reproductive system in humans, the continued examination of the human virome at various sites can only continue to increase our knowledge of these unseen mediators.

Viromes are a fundamental aspect of human biology, with both good and bad consequences. Certain viromes have been associated with an increased risk of obesity, rheumatoid arthritis, multiple sclerosis, Bell's palsy, systematic lupus erythematosus, brain tumors, and Crohn's disease (Perez-Brocal et al. 2013; Wylie et al. 2013).

Additionally, phages can actually modify bacteria through a process known as lysogenic conversion, in which toxin genes like Panton-Valentine, cholera, Shiga- and diphtheria are added to microbial genomes giving rise to new pathogenic strains (De Paepe et al. 2014). Hepatitis G is found within 1-4% of individuals in developing countries and 20% in developing regions but is not responsible for any known disease (Rydze et al. 2012). It is, however, thought to inhibit HIV replication within the body (Rydze et al. 2012). Additionally, latent infections of cytomegaloviruses are thought to induce immune response in mice (Cicin-Sain et al. 2012). Some gyroviruses even have the capability to code for a protein that is toxic to cancer cells and have been found on the skin and in the human gut (Los et al. 2009). Papillomaviruses are commonly found on the skin and were originally thought to simply be bystanders, but are now thought to have a role in proliferation of keratinocytes during wound healing (Lazarczyk et al. 2009). As illustrated in this section, viruses are ever present and likely play an important role in host health. There is one last group of treehouse inhabitants that are worth mentioning: parasites, fungi and those potential members not yet acquired by humans.

Other Treehouse Members

The importance of eukaryotic members of the human treehouse of life should not be undervalued. Humans have lived in close contact with parasites for thousands of years, sometimes in passing, sometimes in prolonged relationships. Ticks, leeches, malaria and the common yeast infection (*Candida albicans*) are just a few examples of these eukaryotes. With such a longstanding relationship, the question could be posed: did we ever benefit from our relationship with parasites? Helminthic therapy suggests

that we do benefit from parasites. Intestinal nematodes like those belonging to the *Necator, Trichuris* and *Hymenolepis* genera may be a treatment for autoimmune diseases like Crohn's, ulcerative colitis and even asthma (Elliott and Weinstock 2009). It is thought that the low incidence of autoimmune disorders in less developed nations is associated with their higher parasite load (Zandman-Goddard and Shoenfeld 2009). In other words, those nasty intestinal parasites may be making humans healthier. For over a century we've identified these macroparasites through microscopy in paleofeces (Vray 2002) and recent molecular techniques allow us to dive into even deeper investigations of parasite DNA (Cleeland et al. 2013; Gia Phan et al. 2013).

Many eukaryotic hitchhikers are merely passersby. Healthy scalps can contain high levels of Candida and Malasseria, the lungs can contain low levels of *Cryptococcus neoformans* and the oral cavity can contain Candida, *Aspergillus*, *Fusarium* and *Cryptococcus* (Huffnagle and Noverr 2013), all of which are asymptomatic under normal conditions. The problem comes in when these organisms bloom and enter the bloodstream, increasing morbidity and mortality of the host (Butler 2010). Fungi within the gut have also been shown to have disadvantageous host effects, such as *C. albicans*, which is associated with gastric ulcers (Kumamoto 2011) and many fungi recognized by the immune receptor *Dectin1* associated with ulcerative colitis (Iliev et al. 2012). Are these parasites longstanding members of the human treehouse of life? It may be possible to turn to ancient fecal material to address this question. A recent study of 2nd century coprolites from Puerto Rico where the paleofeces exhibited the presence of fish parasite (Cano et al. 2014) gives us insights into the interaction between parasite and diet. Similar to information gathered from microbial and viral

studies from human coprolites and calculus, next generation extraction and sequencing techniques are providing an even more detailed look at our rich co-history.

The last topic to mention is those treehouse invaders that are yet to be. In other words, those biological entities that have not yet made the jump to our species. Zoonotic diseases are caused by bacteria, viruses, protozoa or parasites that can jump from animal to human hosts. More than 60% of the recently emergent human pathogens have arisen this way (Hopkins and Wood 2013). Many of these zoonotic transmissions are caused by viruses, and estimates suggest that there are 320,000 mammalian viruses in nine families still to be discovered (Anthony et al. 2013). Animals like bats (He et al. 2013), pigs (Shan et al. 2011), birds (dos Santos et al. 2012; Phan et al. 2013), shrews (Sasaki et al. 2014), cows (Berg Miller et al. 2012), and rats (Firth et al. 2014; Phan et al. 2011) have had their viral loads sequenced. These organisms are constantly in close contact with humans and are potential zoonotic transmitters (Delwart 2012). Viral metagenomic studies have detected a number of transmissible human viruses in bats that have yet to make a leap (Dacheux et al. 2014). Insects are also now suggested to be harborers of novel viruses capable of infecting humans (Flanagan et al. 2012). Zoonotic transmissions are a great threat to human health and by identifying future treehouse invaders we as humans are more prepared to take on future outbreaks and epidemics.

Final Thoughts on Being Human

Katherine Dunham's take on the tree of life essentially continues the tradition of rearranging and reorganizing the metaphor. The tree of life is essentially a visual hypothesis, one which will continue to meld and transform over time with our increased

knowledge of the world around us (Zimmer 2012). It began around the time Ernst Haeckel first introduced the biological concept of a universal common ancestor in 1866 (Haeckel 1866). Much of our knowledge about biology has changed since then, but the core idea of all life being interconnected remains intact. Knowledge of ERVs, HGT, and human-microbe, human-virus and microbe-phage interactions have altered the current landscape on which the tree of life is rooted. Is the current tree of life the best representation of the relationships within and between the kingdoms of life? Likely not. As novel methods are introduced and developed, alternatives to the tree structure have been introduced, such as a star-like representation (Puigbo et al. 2009) and the ring of life (McInerney et al. 2014). Recently, the accuracy of particular microRNAs to create phylogenetic trees in animals like turtles and hookworms have been called into question (Thomson et al. 2014). Some suggest that the tree should reduce its trunk to only two domains (Williams et al. 2013), while others argue for the incorporation of large viruses as a fourth (Raoult et al. 2004). Others argue that the tree must be completely restructured through the addition of viruses in some way (Tom-Orme 1994), even suggesting that those large viruses like the Mimivirus exist as a fourth domain altogether (Raoult et al. 2004). Regardless of the origins of the tree, one thing is clear, our view of the human brand (and most other multicellular branches) need to be revised. Eukaryotes cannot exist without most of these microbes, viruses, fungi and small eukaryotes, and likely never will. These tree of life branches have been ignored for too long in a discussion of what it means to be human. We can no longer consider our biological self to be only human. It is intended intention that this review fosters a

reconstruction of the old “tree of life” adage to incorporate our microbial, viral and eukaryotic self.

Chapter 2

Viruses within Anthropology

The natural viral world encompasses the greatest genetic and biological diversity on Earth…Its continued exploration will undoubtedly unlock many more secrets that fundamentally change our understanding of the evolution and diversity of life on our planet.
Chris Suttle (Yong 2012)

Ignorance is not bliss. Just because we don't know what viruses are out there doesn't mean they're not dangerous.
John L. Gerin (Dalke 2003)

Viruses have existed on earth for billions of years. Although not living in the biological sense of the word they have had a crucial role in the evolution of all cellular life on earth. At its core, anthropology is the study of how humans have adapted over time to the world around them. Many of these adaptations have been in direct response to viruses and microbiota within and around all of us. One aspect of medical anthropology is the drive to explain human response to disease and the spread of those diseases across different landscapes. Yet the importance of viruses to the field continues to change. Since the 1990's and early 2000's, certain viruses and bacteriophages have been used as markers for the migration of human populations. In the decade since, next generation sequencing methods have allowed for a deeper understanding of human-virus interaction.

This chapter provides an overview of our current taxonomic understanding of DNA and RNA viruses (Hulo et al. 2011). Much of the descriptive information on viruses presented in this chapter appears in the regularly updated ViralZone, a

bioinformatics research portal (Hulo et al. 2011). Viral types change so rapidly, that some of the classifications and nomenclature will have changed by the completion of this dissertation. Both the International Committee on Taxonomy of Viruses (ICTV) and Baltimore classification system are the two schemes used to classify viruses, usually through phenotypic characteristics like morphology, replication strategy, and host organism. Everything discussed in this chapter is a result of substantial database searches through PubMed specific journals involved in the study of viral biological, medical, and molecular anthropology.

A Brief History of Viruses (and Microbes)

An understanding of viruses ties in closely with an understanding of microorganisms, which likely began when Antonie van Leeuwenhoek first identified bacterial "animalcules" in 17th century dental plaque (Gest 2004). This initial discovery marked the beginning of our understanding of microorganisms; critters existing within and shaping our world yet unseen by the naked eye. These organisms were originally thought to have arose through spontaneous generation from moist soil (Tortora et al. 2013). The idea of spontaneous generation remained a viable explanation as to the origins of microbial life until the 19th century despite Francesco Redi having disproven this hypothesis with studies of maggots on decaying meat many years before Leeuwenhoek's time (Redi 1668). A theory known as biogenesis eventually replaced the microbial origin story, which was put forth by Rudolf Virchow and Robert Remak (Virchow 1862) and built upon Theodor Schwann's ideas on cell theory (Schwann and Schleiden 1847). These ideas pushed 19th century innovations such as antiseptic surgery

(Lister 1909), preliminary antibiotics, and media for growing bacteria (Koch and Cheyne 1880). Innovations continued in the early 20th century as many found new ways to combat many harmful bacteria through penicillin (Alexander Fleming in 1928), sufonamide (Gerhard Domagk in 1932), and tyrothricin (Rene Dubos in 1939).

Virology as a discipline was not named until the early 19th century. It was Dimitrii Ivanovsky, a Russian botanist, who first discovered the presence of something smaller than bacteria while examining tobacco plants (Ivanovski 1899). Ivanovsky was able to infect tobacco plants even after he passed an 'element' through a filter, which he named *contagium vivum fluidum* ("soluble living germ") (Ivanovski 1899). Soon after, Martinus Beijerinck replicated Ivanovksy's experiments and confirmed the presence of the element, which he then named a virus (Beijerinck 1899). Virus, which means *poison* in Latin, actually first referred to a 'slime' which was thought to be neither dangerous nor infectious (Magner 2009). The first virus to be purified and crystallized was the Tobacco Mosaic Virus in 1935, which earned Wendell Meredith Stanley the Nobel Prize in Chemistry in 1946 (Stanley 1935). In another great leap forward, the electron microscope in 1939 first made it possible to visualize viruses in detail (Magner 2009).

The bacteriophage was discovered early in the 20th century when a century when a French microbiologist named Félix d'Hérelle noted an invisible microbe in a dysentery patient (d'Herelle 1917). He named it a bacteriophage (d'Herelle 1921), a virus that infects only bacteria. Although d'Hérelle is credited with the discovery and was been deemed the father of phage therapy (d'Herelle 1931), history shows that around this time Francis Twort (Twort 1915) independently discovered the virus but received little of the recognition. Bacteriophages played an important role in the

discovery of DNA as the material of inheritance (Hershey and Chase 1952), which was confirmed a year later with the experiments of James Watson and Francis Crick (1953). Revelations continued in the coming decades leading to the discovery of a method for DNA amplification known as polymerase chain reactions (PCRs) (Mullis and Faloona 1987). PCRs allowed researchers to examine the sequence the genetic material of bacteria, leading to the first full genome in 1995 (*Haemophilus influenze*) (Fleischmann et al. 1995). The first full viral genome was actually sequenced decades prior, Phi X 174 (ϕX174) (Sanger et al. 1977). In the 2000's, next generation sequencing analysis came to replace traditional PCR for large datasets, which promised high-throughput and low cost sequencing of full DNA and RNA genomes, transcriptomes, and much more. There are a number of Next Generation Sequencing options, the most popular currently being Illumina (Solexa), followed by Roche/454, Ion Torrent and SOLiD.

The first next generation sequencing study of human viruses examined those within the gut and showed bacteriophages to be the dominant viral type (Breitbart et al. 2003). This study emphasized the importance of these enteric viruses in the persistence of overall gut health. What is their role in regulating host bacteria? This idea that bacteriophages were important to microbial ecosystems was not new knowledge, in fact before the widespread use of antibiotics, d'Hérelle's phage therapy was an option for treating infections, especially during World War I (Fleming 1946). Strangely enough, bacteriophage therapy was largely unknown to the general public and was first introduced with the book Arrowsmith (Lewis 1925) and eventually Oscar nominated film of the same name (Ford 1931) which is the story of a physician who attempts to curtail a bubonic plague outbreak on a Caribbean island (Magner 2009). Today,

bacteriophage therapy continues to have uses and applications, especially in eliminating bacteria in the meat and poultry industry (Magner 2009).

Viruses as Infectious Agents

Although most viruses on earth are not actually virulent (disease causing), infectious viruses are more commonly examined within the medical anthropology field. Infectious diseases are a part of all human societies and anthropologists are fit to study both the populations affected and the agents doing the infecting (Inhorn and Brown 1997). There are four basic disease categories associated with viruses and microbes: "new" diseases (such as HIV), "old" diseases (such as malaria and tuberculosis), those microbial or viral agents with previously unknown etiologies (such as the recently discovered crAssphage), and those microbes or viruses that pose particular risk of infectious disease to previously unaffected populations (such as smallpox in the Americas) (Dutilh et al. 2014; Inhorn and Brown 1997). Many of the studies of the human virome have focused on the third category, examining unknown viral agents that have an unknown impact on humans.

Viral infections can be split into two groups: acute and chronic. Acute agents, like measles, poliovirus and smallpox either lead to immediate host death or immunity from future infections (Wolfe 2011). These acute viruses commonly affect larger human populations. Alternatively, chronic agents like HIV and hepatitis C, attach to their hosts for a lifetime, allowing them to thrive for a long time undetected, in smaller human populations (Wolfe 2011). Considering much of human existence was spent living and travelling in small groups, it is thought that early hominids only had a few viruses

(Burnet 1962). Viral diseases became a point of concern when urban centers first developed in the Near East around 3,000 B.C.E. (Armelagos and Barnes 1999). Diseases like cholera, typhus, measles and mumps were likely present in these early centers. These settlements, with as many as 50,000 inhabitants, would not likely have been large enough to maintain disease in an endemic form (Armelagos and Barnes 1999). Cockburn suggests that a population of one million is necessary to maintain something like measles as an endemic disease (Cockburn 1967). It is thought that these diseases that began infecting settlements so viciously arose from our close contact with domesticated animals.

Viruses like measles and mumps would be considered 'old' viruses by anthropological standards, whereas 'new' viruses tend to arise through zoonotic transmission (Wolfe et al. 2007). There is still much to be learned about the methods by which viruses make the jump to humans and how identification of future threats. Primates alone are the originators of 20% of diseases found within humans (Wolfe et al. 2007). Humans live in close contact with pigs, cows, chickens, bats and many other animals, and efforts have been put forth to characterize the viromes of these zoonotic transmission risk animals. Researchers estimate that there are 320,000 mammalian viruses left to be discovered and the cost to discover all those viruses is in the range of 6.3 billion dollars over 10 years (Anthony et al. 2013). This research not only betters human health through the understanding of potential zoonotic risks, but also uncovers phages and viruses that may have practical health applications. For example, certain non-human viruses are may be used for oncolytic therapy due to their antitumor efficiency (Koppers-Lalic and Hoeben 2011). Currently, anthropologists are more

focused on secondary transmission (human to human) but there is so much to learn from primary transmission (animal to human) (Brown and Kelly 2014).

An Anthropological History

For some time, anthropologists have understood the connection between a person's physical, sociocultural and political environment and their risk for infection (May 1958). Essentially, it is a discussion of maladaptation to the surrounding environment (Dubos 1965). During this time, anthropologists were discussing human disease from microbiota and viruses as one in the same. It was not until the until the 1970's, specifically 1975, when the WHO initiated a Training in Tropical Diseases program, that anthropologists began to seriously studying the effects of disease across the globe, many of which were viral in origin (Inhorn and Brown 1997). It was then that the specific framework for examining diseases within populations arose, including manners by which they were transmitted and persisted (Armelagos and McArdle 1976). One of the main talking points for many of the major diseases was role of asymmetrical relationships in perpetuating illnesses. These power struggles between the haves and the have-nots were underlying issues in both developing and industrialized nations across the globe (Morsy 1979) and the coming decades only saw an exacerbation of problems (Baer et al. 1997; Farmer 2001; Nguyen and Peschard 2003).

In the 1980's, the American Anthropological Association (AAAs) annual conference held a symposium devoted to the anthropology of infectious disease, which acted as a watershed moment for the field (Inhorn and Brown 1997). As a result of the HIV/AIDS pandemic and the Child Survival Initiative, medical anthropology continued

to gain momentum in the 1980's (Inhorn and Brown 1997). Although anthropologists have an obligation to examine 'priority' diseases like AIDS and most recently the causative agents of viral hemorrhagic fevers (VHFs) like Ebola, Marburg and Lassa (Brown and Kelly 2014), they also have the responsibility to address neglected diseases, those that affect individuals at a regional level. They can do so through methodological approaches including the documentation of individual ethnographies and examination of surrounding political and economic factors (Inhorn and Brown 1997). Granted, investigating infectious diseases should be part of medical anthropology, but there is more to be learned about human-virus relations outside of the study of pathogens. JBS Haldane was one of the first to suggest the importance of infectious diseases to the evolutionary process as a whole (Haldane 1949).

Prions

While not a virus, a brief mention of prions is warranted here as they are similar infectious particles which owe their discovery, in part, to anthropological research. Prion infections were first observed in the Fore people of the highlands of New Guinea (Lindenbaum 1979), which is similar in form to Mad Cow Disease which first appeared in England in the 1980's and reached a pinnacle in 1992 (Magner 2009). Prions, also known as transmissible spongiform encephalopathies (TSEs), can affect both humans and animals and are distinguished by long incubation periods within the host (Belay 1999). They have also been identified as causative agents for Creutzfeldt-Jakob and Gerstmann-Straussler-Scheinker syndromes (Prusiner 1998).

Double Stranded DNA Viruses

Double stranded DNA viruses are incredibly taxonomically diverse and complex. They fit into one of three named orders, Caudovirales, Herpesvirales, Ligamenvirales and a forth unassigned order. A breakdown of families and genera within those orders can be seen in Figure 3. Most viruses within the Caudovirales order are lytic and multiply within the cytoplasm, though there are some exceptions. The viruses within the Caudovirales are mainly phages and likely have a role in mediating the presence of particular microbes within humans. To date, anthropological studies related to this family are essentially absent, but a recent paper linked irritable bowel diseases with an overabundance of Caudovirales (Norman et al. 2015), suggesting an impact upon human health that would be worth exploring in future studies.

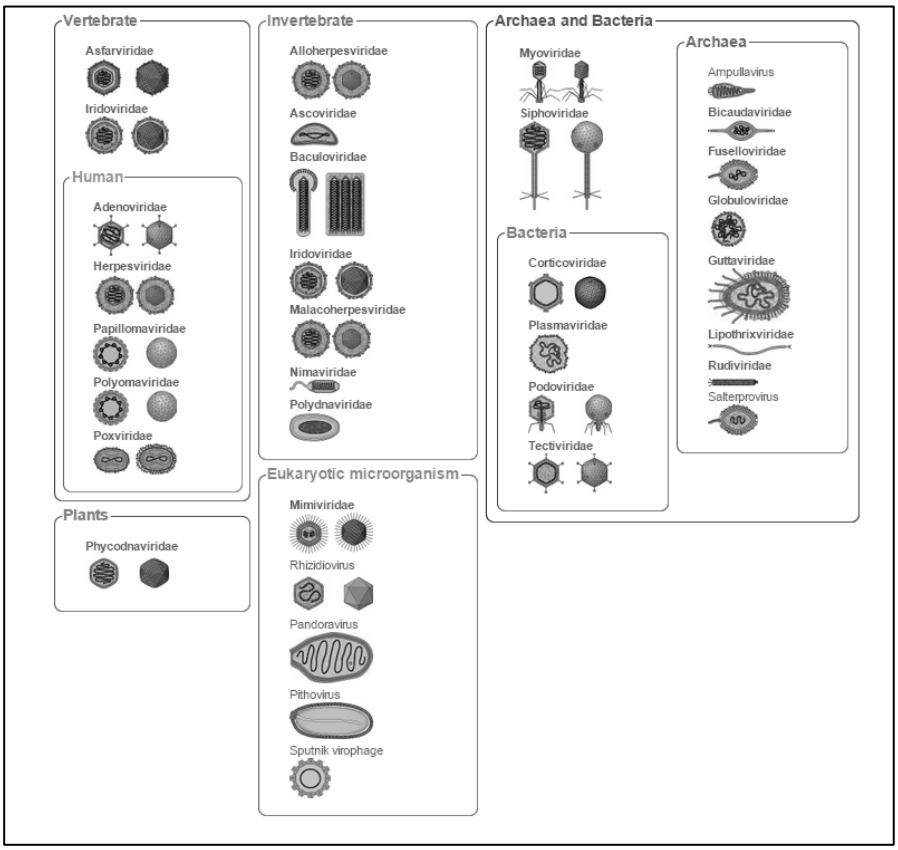

Figure 3: Double stranded DNA virus families (Hulo et al. 2011)

Caudovirales

The Caudovirales order is comprised of tailed bacteriophages and contains around 350 species. The Myoviridae family has a genome size of 33-244kb, encoding for 40-415 proteins. There are some 130 species of viruses within this family. Viruses belonging to Myoviridae are nonenveloped, comprised of a head, neck and tail. The term "myo" comes from the Greek word for "muscle" which refers to their contractile tail. When this order is mentioned in the virome literature it is usually in passing and has been briefly discussed in studies of bats (Wu et al. 2012) and freshwater sources (Roux et al. 2012). Presently, myoviruses have not been explored within anthropological research, but they have the potential to be used in bacteriophage therapy, and therefore have practical relevance to studies of human health.

The Siphoviridae family share traits with the Myoviridae: the head-tail structure and a linear genome. These viruses are on the whole, are a lot smaller (50kb and 70 genes) than Myoviridae. Prophages from this family have been known to infect Mycobacterium, Streptococcus and Lactobacilli (Ventura et al. 2006) and could be potential regulators of these bacteria within the human ecosystem. Lactobacilli are lactic acid bacteria common across many human body sites including the GI tract (Ventura et al. 2011) and the human vagina (Dicks et al. 2000). Their presence in the oral cavity is thought to be associated with dental caries (Badet and Thebaud 2008). Siphoviridae and Podoviridae are commonly detected in fermented foods (Kim et al. 2013; Park et al. 2011), consumption of which may impact the overall composition of the human gut virome.

Podoviridae possess a similar non-enveloped, head-tail structure as the other viruses in this order but do not possess a contractile tail. Genome size ranges from 40-42kb, with 55 genes and around 74 species. Viruses within the Podoviridae family are still largely unknown and considered very diverse. These viruses are so unknown, that a single study identified 93 complete phage genomes that separate into 12 distinct clusters and likely infect Firmicutes (genus *Bacillus*) (Grose et al. 2014). This analysis identified 4,922 protein families within these Bacillus phages, 19% of which have a known function. In other words, 81% of Podoviridae viruses have no known role in the human gut. To date, anthropology studies related to Siphoviridae and Podoviridae have been nonexistent, but their associations with microbiota commonly found within human body sites make them families of interest for future metagenomics investigations.

Herpesvirales

The Herpesvirales order consists of three families of large (124-195kbp) double stranded DNA viruses. For years, Herpesviridae was the only family in the order, but discoveries of novel Herpesvirales viruses in carp, goldfish and koi established the Annoherpesviridae family and a discovery in molluscs founded the Malacoherpesviridae family (Savin et al. 2010). The main point of focus for human studies is the Herpesviridae family, which split into three subfamilies: Alphaherpesvirinae, Betaherpesvirinae and Gammaherpesvirinae, around 200 million years ago (McGeoch et al. 1995). Herpesviruses encode for 100-200 genes and replicate in the nucleus rather than the cytoplasm of the host cell. There are 8 pathogenic strains currently known including human herpesviruses (1 & 2), cytomegaloviruses, Epstein-

Barr viruses and Varicella zoster viruses. Some of the herpes viruses have practical anthropological applications (HSV-1 is already used in human migration studies) mainly because they are a) ubiquitous across the world and b) stay in the host for life.

Herpes simplex 1 and 2 (HSV-1, -2) are common throughout the globe and can cause contagious oral or genital blisters. Their contagiousness makes them incredibly informative from an anthropological standpoint. For one, HSV-1 helped us better understand how genetic material is ejected from a virus. It actually uses internal genome pressure to eject DNA into the host cell which can be suppressed by external osmotic pressure (Bauer et al. 2013). HSV-1 is also one of several candidate viruses in human migration studies. Genetic information from HSV-1 show that the virus clusters into six distinct clades based primarily on geographic location (specifically continents) (Kolb et al. 2013). It was used to better understand the dispersion of humans across the islands of Japan as well (Ozawa et al. 2006). HSV-1 is such an excellent candidate for population structure studies due to its high mutation rate, which researchers utilize for tree building.

The Herpes virus 3, AKA Varicella zoster virus AKA chickenpox, is slightly less informative but just as ubiquitous. Common across most children but not incredibly dangerous, it becomes an issue when it reactivates later in adulthood and causes shingles. Herpesvirus 4, known as the Epstein Barr virus, has actually been used as a marker for psychological stress in humans (Kiecolt-Glaser et al. 1987). The chronic exposure to stressors induces an up regulation of the HPA (hypothalamic-pituitary-adrenal) axis, resulting in a reactivation of EBV and thus, EBV antibodies (Glaser et al. 1991). A more recent study in Siberia found a correlation between socioeconomic status

and an impaired immune system, which was linked to a higher abundance of EBV antibodies causing lower cell mediated immune function (Sorensen et al. 2009). It may not be only psychological stress but physical stressors as well, a study of Brazilian mothers showed that oxytocin, a neuroendocrine factor thought to be associated with stress, did not correlate with EBV infection (Rudzik et al. 2014). Epstein Barr was first discovered just over 50 years ago and we are only beginning to understand its role with cancers and other diseases (Figure 4) (Lieberman 2014).

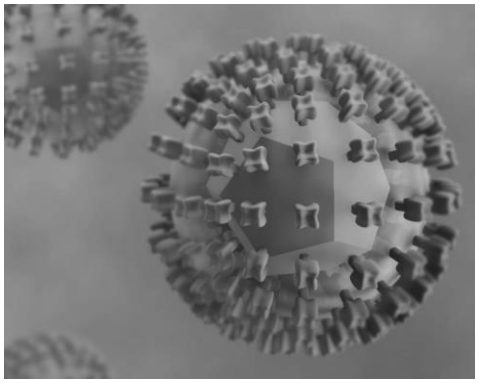

Figure 4: Eppstein-Barr Virus computer rendering (Lieberman 2014)

Cytomegalovirus (CMV), another name for herpesvirus 5, is a particularly good herpesvirus for anthropological studies. First, it is present in 70% of individuals over age 60 and remains in those people for life (Bate et al. 2010; Dowd et al. 2011). Second, since these traits are shared with other herpes viruses, the main draw of CMV is that it can be examined in ancient dried blood samples (compared to more difficult to examine RNA viruses). Third, CMV is more common in South America, Africa and Asia compared to Europe and North America (Cannon et al. 2010). Fourth, CMV has been suggested as a driving force behind age-associated alterations to T cell production (Pawelec et al. 2009) making it a unique exploratory pathway for examinations social and biological stress on immune function (Dowd et al. 2011). Specifically, CMV can

recruit CD4+ T cells and inhibit their antiviral function (Mason et al. 2012) and express a protein that can inhibit cell division (Song and Stinski 2002), both of which have adverse effects on the host. CMV seropositivity is more common among women than men (Cannon et al. 2010) and is also a great risk for children. According to the CDC, 1 in every 150 children is born with a congenital CMV infection and of those, 1 out of every 5 will suffer from developmental issues including permanent hearing loss and neurological impairment (Cannon et al. 2010).

Herpesvirus 6 infects T lymphocytes in humans and can affect humans and adults and reactivate as either an acute or chronic illness. Herpesvirus 6 along with CMV and HSV-1 and HSV-2, have an association with brain tumor development (Kofman et al. 2011). A lower abundance of herpes virus 7 was associated with AIDS patients versus control groups (Lucht et al. 1998) and a higher abundance was associated with connective tissue diseases like Still disease (Yoshikawa et al. 2005). However, HH7 is also present within healthy, asymptomatic adults, suggesting a potential use for population studies like HSV-1.

Human Herpesvirus-8 (HHV-8), also known as Kaposi's sarcoma-associated herpesvirus (KSHV), is present in all forms of Kaposi sarcoma (KS) (de Souza et al. 2007). In the U.S. and Western Europe HHV-8 is usually rare outside of high risk groups (de Souza et al. 2007). However, HHV-8 can be much higher in areas like Africa and the Mediterranean basin, sometimes reaching 100% of the population. HHV-8 has been studied in Amerindian and Non-Amerindian populations in the Brazilian Amazon in order to determine models for transmission (de Souza et al. 2007). The study showed

that seroprevalence of HHV-8 in Amerindians is much higher than non-Amerindians, which has also been shown in French Guiana (Kazanji et al. 2005)

Nucleocytoplasmic Large DNA Viruses

Nucleocytoplasmic Large DNA Viruses (NCLDV) were identified fairly recently (Yutin et al. 2009) and are split into seven families. They possess 47 core genes that infect the cells of eukaryotes and protists. A number of these viral families are rarely studied and have no known relationship to anthropology: Ascoviridae (which infect mostly insects), Iridoviridae (which infect amphibians, fish and invertebrates), and Phycodnaviridae (which infect algae). Phycodnaviridae actually plays an important role in the bloom of microalgae in marine ecosystems (Suttle 2005), a regulatory trait echoed in many marine virus studies.

Asfaraviridae possess only a single genus, Asfivirus, otherwise known as the African Swine Fever Virus (ASFV). First discovered in Kenya in 1921 (Eustace Montgomery 1921), it likely gained prominence during the 18th century in Europe (Michaud et al. 2013). It is considered 100% lethal to domestic pigs (Costard et al. 2009). Several outbreaks have occurred around the globe over the last decade and there is potential for a study of animal domestication and ASV evolution over time.

Both Marseilleviridae and Mimiviridae families possess enormous genomes for viruses: 372kbp and 1200kbp respectively. Little is known about these large viruses or what usefulness they may have to studies of evolution but they have been documented in oceans, human blood, insects and amoebas (Aherfi et al. 2014; Boyer et al. 2009; Martinez et al. 2014; Popgeorgiev et al. 2013). These two families have been used as

arguments for the placement of viruses into the tree of life and may have immense implications for studies of evolution,

The most commonly investigated viral family from the NCLDV order causes smallpox in humans (but affects many other animals): Poxviridae. Pox comes from the English word pock or pustule, referring to the skin lesions that result from the virus. These viruses are split up into two subfamilies: Entomopoxvirinae and chordopoxvirinae. The smaller of the two subfamilies, Entomopoxvirinae, has three associated genera and only infect insects. The chordopoxviridae subfamily, infecting humans and mammals, is the classification of the smallpox virus in humans. There are two forms in humans, the variola major (deadlier form) and the variola minor but there is a risk for zoonotic transmission of other nonhuman forms of the pox viruses (Shchelkunov 2013).

Smallpox is considered to be one of the most devastating viruses ever known, having caused an estimated 300-500 million deaths until its eradication on May 8th, 1990 (Koplow 2003). The virus kills 25% of those it infects but that number can reach as high as 50% in virgin populations (Dobyns 1983). Between 1520 and 1898, there were 41 separate epidemics in Native North Americans and by the end of the 19th century, smallpox had wiped out 90% of Native Americans (Tortora et al. 2013). The earliest known case of smallpox is from Ramses V, an Egyptian pharaoh who died in 1157 B.C.E. (Ruffer 1921)

Other zoonotic species of variola are the monkeypox virus (MPXV) vaccinia virus (VACV) and the camelpox virus (CMLV). The cowpox virus (CPXV) possesses the largest genome of the group, while the variola virus (VARV), with the smallest

genome, is also the most pathogenic (Shchelkunov 2013). The associations to anthropology would likely involve examining zoonotic risk. The orthopoxvirus genus has been a zoonotic risk for most of human history, with repeated emergences and exposures likely (Shchelkunov 2011). Some hypothesize that wildlife like horses and rodents that were repeatedly introduced to the Americas from Europe continued to bring over strains of pox viruses (Shchelkunov 2013). Though smallpox has been eradicated, it is important to closely monitor those animals that have the capability of transferring novel viruses, in order to the stem the progression of small outbreaks into full blown epidemics (Shchelkunov 2013).

Unassigned Order

There are 20 families of dsDNA viruses that are not assigned to any of the three established orders, infecting a range of organisms. The number of families and subfamilies are constantly shifting with new knowledge and there are currently several that do not belong to any family. In 2005, a new viral family, Ampullaviridae, was proposed for viruses found within Italian hot springs (Häring et al. 2005). Similarly, Bicaudaviridae viruses also infect archaea, particularly members of the Acidianus genera. There are a number of other archaea infecting viral families: Clavivirdae, Fuselloviridae, Globuloviridae, Guttaviridae Lipothrixviridae, Turriviridae, Nudiviridae and Rudiviridae. Baculoviridae is a family of invertebrate infecting viruses that have the ability to liquefy insect cadavers and alter feeding behavior in lepidopteran hosts (Hamblin and Tanaka 2013). Hytrosaviridae infect Diptera (flies), Polydnaviridae infect parasitoid wasps and Nimaviridae infect a wide range of organisms including

crustaceans, Corticoviridae infect gram-negative bacteria, Plasmaviridae infect mycoplasma and acholeplasma, while Tectiviridae infect a wide range of bacteria.

Those organisms that have not been put into a taxonomic family include the Sputnik Virophage, a unique satellite virus that actually infects amoeba cells that have already been infected by a virus (Mimivirdae). Another dsDNA virus that has not been placed into a taxonomic category is the recently discovered crAssphage (Dutilh et al. 2014).

One of the most commonly discussed viral types from the unassigned order of DNA viruses is Adenoviridae, found in a number of vertebral hosts, mostly on account of some form of adenovirus. First isolated from Adenoid tissue in 1953, adenoviruses are linear, non-segmented dsDNA between 26 and 48kbp. Adenoviruses, which cause common colds and flus, are found not only within humans, but also reptiles, canines sea lions, birds, and many other large mammals. There are currently no vaccines available or antiviral treatments for adenovirus infections. It seems that adenovirus arose fairly recently not as a zoonotic transmission from chimps to humans (Wevers et al. 2011) but other nonhuman primates including marmosets (Yu et al. 2013). There are many avenues for future health research in which correlate seropositivity of certain strains (Ad36) to increased obesity and BMI (Trovato et al. 2009).

The Papillomaviridae viral family has hundreds of types, most of which are asymptomatic or cause benign tumors (warts). It infects most known mammals but rarely travels between species (Mistry et al. 2008). Many of these viruses can persist outside of the human host in environmental samples (as seen in Table 1) (Bibby and Peccia 2013; Cantalupo et al. 2011) and researchers are still trying to understand their

pathogenicity. The many types of human papillomaviruses are divided into five genera: Alphapapillomavirus, Betapapillomavirus, Gammapapillomavirus, Mupapillomavirus and Nupapillomavirus. HPV is a major cause of cervical cancer, for which there are 500,000 cases diagnosed globally each year (Munoz et al. 2006). It is considered the most common sexually transmitted infection worldwide and with over 200 strains, it infects a majority of sexually active men and women at some point during their lives (Baseman and Koutsky 2005). Not every strain is pathogenic; healthy individuals have shown the presence of certain types in their saliva (HPV16) (Turner et al. 2011). The main pathogenic type is HPV16 which can transform primary foreskin keratinocytes in men, giving them an increased risk for transmitting it to their female sexual partners, which can lead can then lead to cervical cancer (Niccoli et al. 2012). Anthropologists have chosen to approach the virus from a cultural opinion and prevention standpoint, as seen in studies related to responses to screening methods in Bogota, Columbia (Wiesner et al. 2012). Within the U.S. African American community women are at significantly higher risk of cervical cancer from HPV (Dunne and Sternberg 2007) but only 20% of eligible girls are vaccinated (Nelson 2010). Anthropologists are attempting to understand the reasons behind these decisions. Interviews with 30 African American women from the Midwest U.S suggest the reason behind the non-vaccination was not a religious one but a lack of knowledge regarding safety and cost of the vaccination (Thompson et al. 2012). With the recent backlash in America against the HPV vaccine, fueled by misinformation in the popular media (stemming from a since retracted Lancet article (Wakefield et al. 1998)) there is a good deal left for anthropologists to explore (Hopkins and Wood 2013).

TABLE 1 Human viruses present in raw sewage

Family	Species	Genome
Adenoviridae	Human adenovirus 41	dsDNA
Astroviridae	Astrovirus MLB1	ssRNA(+)
	Human astrovirus 1	ssRNA(+)
Caliciviridae	Norwalk virus	ssRNA(+)
	Sapporo virus	ssRNA(+)
Papillomaviridae	Human papillomavirus 112	dsDNA
Parvoviridae	Adeno-associated virus	ssDNA
	Human bocavirus 2	ssDNA
	Human bocavirus 3	ssDNA
Picobirnaviridae	Human picobirnavirus	dsRNA
Picornaviridae	Aichi virus	ssRNA(+)
	Human klassevirus 1/Salivirus NG-J1	ssRNA(+)
	Human parechovirus 1	ssRNA(+)
	Human parechovirus 3	ssRNA(+)
	Human parechovirus 4	ssRNA(+)
	Human parechovirus 7	ssRNA(+)
Polyomaviridae	Polyomavirus HPyV6	dsDNA

Table 1: Viruses present in raw sewage (Cantalupo et al. 2011)

Likely the most commonly referenced viral family from the unassigned order of dsDNA viruses is the Polyomaviridae family which includes two viruses used for human migration studies: Polyomavirus hominis type 1 better known as the BK (BKV) polyomavirus and type 2 JC polyomavirus (JCV) (Agostini et al. 2002). The sequences of the BK and JC viruses both show 75% homology, with 1-2.6% of the sequence being variable and helping to distinguish the eight major global genotypes (Agostini et al. 2002).

The JCV is about 5kbp in size and invades human cells through serotonin receptors called 5-HT2 receptors (Assetta et al. 2013) and persists within the kidneys. It was first isolated in 1971 (Padgett et al. 1971) and can cause a disease known as demyelinating disease called progressive multifocal leukoencephalopathry (PML) in

immunocompromised individuals (Astrom et al. 1958). The JC polyomavirus was first sequenced in 1984 but it was not until 1987 when some differences in Asian and European strains of JCV were noticed (Matsuda et al. 1987). Shortly thereafter, it was discovered that these strains are different genotypes, and can be used to track human populations (Ault and Stoner 1992; Yogo et al. 1991). The JC virus is a candidate genome for population migrations for several reasons: it is small enough to analyze easily, it is transmitted only by close contact, and it is usually ubiquitous and generally asymptomatic (Agostini et al. 2002; Kitchen et al. 2008). It has been associated with a rare but fatal disease involving demyelinating brain tissue known as progressive multifocal leucoencephalopahty (PML) (Stoner et al. 1988). Additionally, it is found in high levels in urine in adults (20-70%) so it can be easily studied using a urine sample (Pavesi 2004). JC virus replication may be induced by the presence of CMV, but more research is needed to understand their exact relationship (Heilbronn et al. 1993).

The JC virus is thought to have originated 50k to 100k years ago, coevolving with humans as they dispersed throughout the world (Agostini et al. 1997). Studies have investigated the serotypes from Asia (Guo et al. 2001; Ikegaya and Iwase 2004; Miranda et al. 2003), Africa (Chehadeh et al. 2013; Pavesi 2005; Takasaka et al. 2006), Europe (Ikegaya et al. 2005), the Pacific (Ryschkewitsch et al. 2000; Takasaka et al. 2004), and indigenous populations in the Americas (Cayres-Vallinoto et al. 2012; Fernandez-Cobo et al. 2002). Native American groups has shown that JCV infection predates contact with Europeans and likely came to the Americas across the Bering land bridge 12k to 30k years ago (Agostini et al. 1997). The first group to colonize the Americas brought in subtype 2A while Columbus brought in subtype 1 and 4 and the

African slave trade brought in subtype 3 and 6 (Stoner et al. 2000). Some suggest that the JC viruses mutates too quickly to be an accurate indicator of human history (Shackelton et al. 2006). The BK virus is studied less commonly than the JC virus (though they show 75% homology) (Padgett et al. 1971), but has four distinct serotypes and appears across the globe (Zhong et al. 2009). Like the JC virus, appears in the kidneys of affected peoples (Pastrana et al. 2013). Although, the BK virus may have fewer applications within anthropology when compared to JCV, considered many ancient populations have lost their intrinsic BKV lineages (Yogo et al. 2009). It now seems that utilizing BK and JC viruses for studies of geographic structure may be a thing of the past (Holtz et al. 2014). Anthropologists may be better suited to examine predicted proteins from gut and oral viruses in order to better understand human evolution, something that will be addressed in a forthcoming chapter.

Single Stranded DNA Viruses

Single stranded DNA viruses have been shown to infect humans, plants, archaea, invertebrates and bacteria. There is no formal order within ssDNA viruses, making all nine families unassigned (Figure 5). Despite this seeming lack of organization, "…their small size and lipid protein-coding capacity, the rapid evolution rates of single-stranded DNA (ssDNA) viruses have led to their emergence as serious plant and animal pathogens" (Rosario et al. 2012). On the whole, most human virome studies are dominated by dsDNA bacteriophages while marine studies are dominated by ssDNA virus families (Lopez-Bueno et al. 2009).

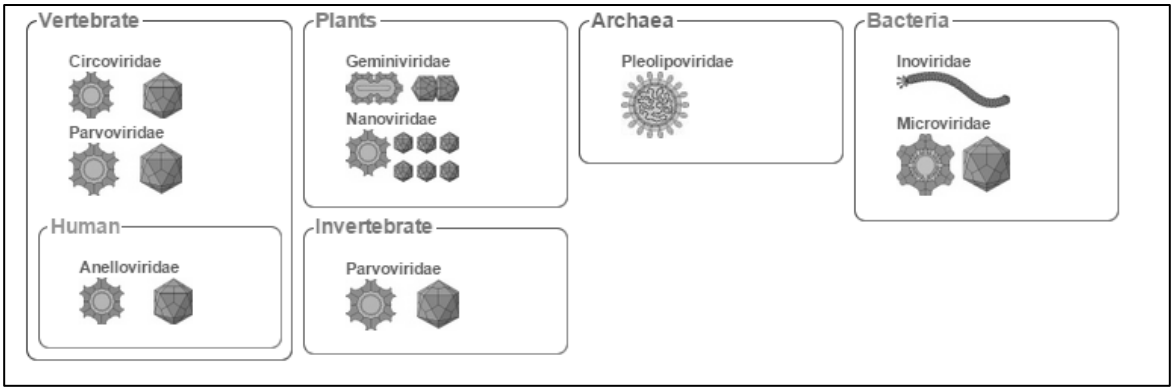

Figure 5: Single stranded DNA virus families (Hulo et al. 2011)

Viral families like Bidnaviridae infect only silkworms, Inoviridae infect only bacteria, and Pleolipoviridae infect only archaea. Geminiviridae and Nanoviridae viruses infect plants like bananas, beans, beets, maize and tomatoes. An unassigned virus known as Rhizidiovirus affects fungi. Circoviridae commonly infect insects but have been found in many other organisms including pigeons (Phan et al. 2013; Van Borm et al. 2013), rats (Phan et al. 2011), ferrets (Feher et al. 2014), dogs (Li et al. 2013), and humans (Gia Phan et al. 2013; Li et al. 2010). Although little is known about most of the viruses within the Circoviridae family, a recent investigation of the *Rep* gene (rolling circle replication initiator protein) established deep evolutionary roots for ssDNA viruses, predating the radiation of Eukaryotes (Delwart and Li 2012). There is no universal gene for studying viruses (unlike the 16S rRNA gene in bacteria) the discovery of the *Rep* gene (Rosario et al. 2009) has uncovered a wealth of viral diversity applicable to studies of viral emergence and evolution (Diemer and Stedman 2012).

Anelloviridae viruses affect a wide range of mammals and are considered fairly ubiquitous across the globe. These circular genomes are around 3.8kbp in size and have been found within the human GI tract (Okamoto et al. 1998). The most common viruses from this family are known as Torque Teno Viruses (TTV). TTV was first isolated in

1997 and there are a number of different variants including the mini and midi forms within humans (Nishizawa et al. 1997). TTV is incredibly widespread (Abe et al. 1999) and is prevalent in some Russian (Vasilyev et al. 2009) and Amazonian (Niel and Gomes 2002) populations at over 90%. The virus is very diverse across the globe, with certain genomes only showing 50% nucleotide similarity (Prescott et al. 1999). As a result, TTV genotypes have been informative enough to suggest a geographical distribution of genotypes (Prescott et al. 1999). For some time but the pathogenicity of the virus was poorly understood (Abe et al. 1999) but research shows that it is involved in the production of micro RNAs that allow it to evade host immune system detection (Kincaid et al. 2013). Considering TTV are fairly ubiquitous across the globe (despite their tremendous diversity) there may be opportunities within molecular anthropology to study genome variation across geography.

Viruses within the Microviridae family are only lytic and cannot lyse their host microbes (Brentlinger et al. 2002). Microviridae viruses hold an important place in next generation sequencing history. Not only was Phi X 174 (ϕX174) (Sanger et al. 1977) the first genome ever sequenced, but it is also used as a positive control in next generation sequencing due to its small genome size. Even with their small genome sizes, the Microviridae family is thought to have one of the higher substitution rates of all ssDNA genomes (Minot et al. 2013) and several independent studies have hinted to their very ancient origins (Kim et al. 2011; Minot et al. 2011). Microviridae have helped address questions as to the exact types of viruses present within the human gut. A longitudinal viral gut metagenomics study showed that short term members of the ecosystem are very diverse and those that are long term residents are capable of

evolving into new viral species over the course of the host's lifetime (Minot et al. 2013). In other words, every person has a personalized unique evolutionary viral history within them.

Parvoviridae are linear ssDNA viruses about 4kbp to 6kbp in size and split into two subfamilies including Parvovirinae and Densovirinae. Members of the Parvoviridae family are commonly found in both healthy (Kapoor et al. 2009) and diarrheal individuals (Finkbeiner et al. 2008). This family included the Human Parvovirus B19 is one of the smallest DNA viruses and primary infects children and causes a rash known as erythema infectiosum. These parvoviruses likely arose from an ancient zoonotic infection (Lukashov and Goudsmit 2001) and diversity within the family suggests a recent form arose in the 1950's (Norja et al. 2008). B19 is detected globally in human populations but there is debate as to whether it is driven by forces of positive (Barros de Freitas et al. 2007) or negative selection (Lukashov and Goudsmit 2001). The incredibly small size of the B19 virus and its even dispersion across the globe makes it a likely candidate for anthropological studies of genome evolution and associated evolutionary pressures.

A study of Chimpanzee stool showed an alarming number of circular DNA viruses (ChiSCV) that do not quite fit into the naovirus, circovirus or geminivirus categories (Blinkova et al. 2010). They found these viruses in the stool of wild and captive chimps from all around Africa. Categorically they do not fit into any taxonomic group aside from ssDNA and they have no known function.

Double Stranded RNA Viruses

There are twelve families under the classification of double stranded RNA viruses, all of which do not have an assigned order (Figure 6). dsRNA viruses are largely considered infectious agents and are found within a wide range of hosts. All RNA viruses are thought to have a common ancestor and the Flaviviridae family shares characteristics with both single and double stranded types (Strauss and Strauss 1988). RNA viruses are also very malleable; their genomes can accumulate change at a rate of 0.1%-1% per year (Strauss and Strauss 1988). Most of the families of viruses within the unassigned order we still know very little about. Some understudied families include: Amalgaviridae (plants like tomatoes, blueberries and rhododendrons), Birnaviridae (chickens, fish, insects), Cystoviridae (bacteria), Endornaviridae (plants, fungi, oomycetes), Partitiviridae (plants and animals), Totiviridae (fungi and protozoa), and four families that infect only fungi, Chrysoviridae, Hypoviridae, Megabirnaviridae and Quadviridae. Picobirnaviridae are small, non-enveloped viruses that can cause acute gastroenteritis in children (Picobirnaviruses (PBVs) specifically). These viruses are found to infect a wide range of mammals in a number of unique body sites (they were found in the lungs of pigs (Smits et al. 2012)). Animal models suggest that the acquisition of the virus occurs in early life and is followed by points of high viral activity and prolonged periods of silence for many years (Ganesh et al. 2012). These viruses may not be solely pathogenic, they have been observed in the GI tract of non-diarrheic healthy human hosts so understanding their relationship to human health is an important endeavor. This led some to posit that they can be considered enteric viruses

within some humans and may even be opportunistic (Giordano et al. 1998; Smits et al. 2011).

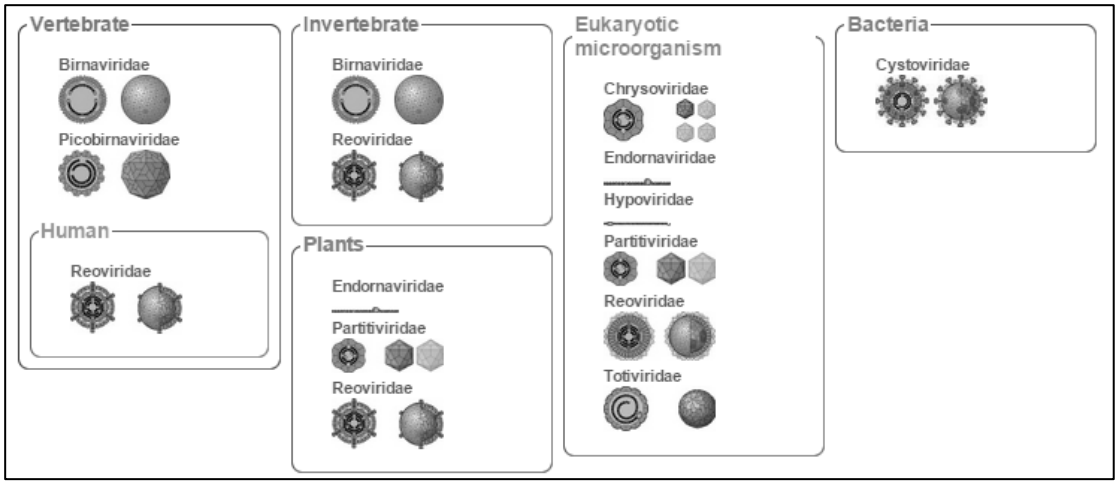

Figure 6: Double stranded RNA viruses (Hulo et al. 2011)

Reoviridae is derived from respiratory enteric orphan viruses, which refer to viruses that do not cause any known diseases. They are now known to cause a number of diseases in vertebrates and invertebrates, but the name is still in use. Transmitted through fecal-oral contact in humans this family contains the rotavirus, known to cause severe diarrhea among young children. Over 450,000 children are thought to die every year due to severe diarrhea caused by the rotavirus (Tate et al. 2012). Recently, public health officials responded to such viral devastation with requests for vaccines and rehydration therapy programs in high risk areas. In the past, anthropologists have been concerned with spreading health awareness of the causes of diarrhea to young mothers (Green 1985; Scrimshaw and Hurtado 1988). In addition to health interventions, there has been a desire to understand the unique cultural views as to the causes of diarrhea in children, like the belief in nyoka, a body purity guardian, in areas of Mozambique (Green et al. 1994).

Positive Single Stranded RNA Viruses

The positive single stranded RNA virus taxonomy is comprised of three orders and a fourth group of unassigned viral families (Figure 7). Like many of the viruses discussed in this chapter, the taxonomic organization of ssRNA+ viruses continues to change with the more information we gain about them. The stem loop structure found within several ssRNA+ families (from different orders) is considered to have been horizontally transferred between them suggesting a unique selective advantage for those possessing it (Tengs et al. 2013).

Nidovirales

Arteriviridae are made of linear 12-15kbp genomes and show pathogenesis in a wide range of vertebrates including pigs, horses, mice, and monkeys; but not humans. The Mesoniviridae and the Roniviridae families only infect mosquitoes and crustaceans (respectively). The only true relevance to humans within this order is the Coronaviridae family which includes severe acute respiratory syndrome coronavirus (SARS CoV) and other vertebrate associated coronaviruses. Coronoviruses are the largest RNA viruses ranging from 26.4kbp to 31.7kbp and are particularly prone to homologous recombination (Crossley et al. 2012). SARS CoV caused an outbreak in China and Hong Kong which led to almost 1,000 deaths (WHO 2003). It was thought to be caused by a species of fruit bat with an intermediary host in the palm civet (Li et al. 2005). The Middle East form of SARS (MERS-CoV) is also thought to have originated from bats and causes a 36% fatality rate in those humans infected (Yang et al. 2014). From an anthropological context, articles are primarily focused on media and regional level

responses to the disease in places like Toronto (Ali 2008; Galley 2009) and China (Chun 2004; Zhan 2005). From a genomics perspective, researchers have attempted to understand the inner workings of the virus and the cause of its virulence. Considering that coronaviruses possess a large genome size and have a potential for recombination due to RNA replication issues, they may not be the best candidates for studies of pathogenic virus evolution and dispersal.

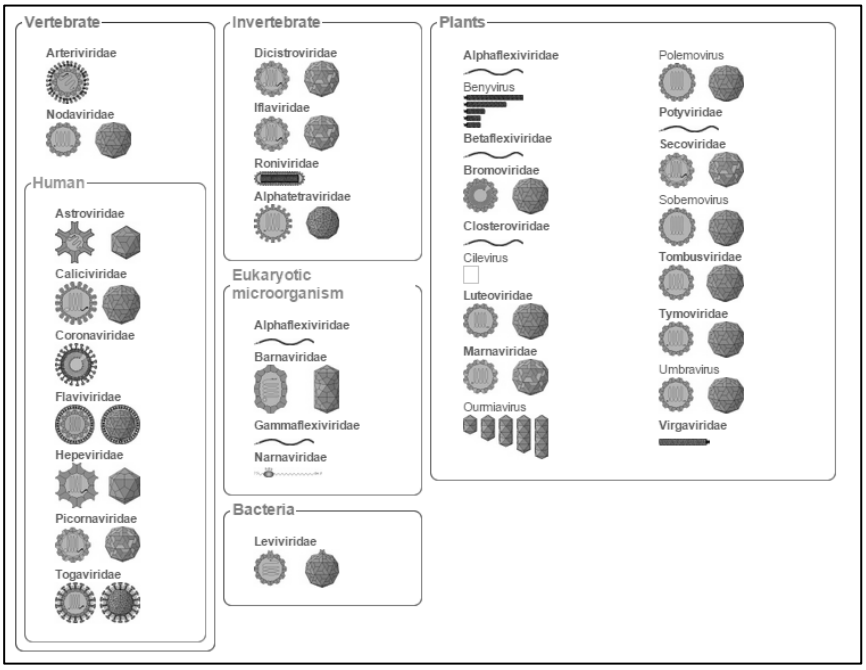

Figure 7: Positive single stranded RNA viruses (Hulo et al. 2011)

Tymovirales

The order Tymovirales include: Alphaflexiviridae and Betaflexiviridae which are both less than 10kbp in size and infect plants and fungi, and Gammaflexivirdae and Tyoviridae which are a comparable size and infect only fungi and plants (respectively). Outside of infecting many crops that humans consume (apples, grapes, potatoes) this order has not been investigated in any anthropological context and likely have little relevance to the field.

Picornavirales

There are five families within the Picornavirales order: Dicistroviridae, Iflaviridae, Marnaviridae, Picornaviridae and Secoviridae. Iflaviridae is considered to be particularly pathogenic to insects, Secoviridae affect a wide range of plants (cherries, cowpeas, parsnip), and Marnaviridae affect primarily seaweed. Dicistroviridae was considered associated with insects and has been found in high abundance in the gut of the (insect consuming) shrew (Sasaki et al. 2014). The important family to mention within this order is Picornaviridae. The genomes are smaller, 7.1kbp-8.9kbp but family is huge; at last count the number of associated genera was more than two dozen. In addition to many vertebrate associated viruses, those that commonly infect humans: rhinoviruses, the hepatitis A virus and the poliovirus. These viruses have been identified in the human GI tract previously (Pallansch and Roos 2007) and each is worth briefly mentioning in an anthropological context.

Human rhinoviruses cause the most common respiratory infection, resulting in the illness of millions every year (Jacobs et al. 2013). Currently, there have been investigations into the proper overlap of clinical research into pathogenic diseases (with an emphasis on Rhinoviruses) compared with medical research of diseases, discussing specific implications and challenges to the mixing of experimental and therapeutic healthcare (Greenhough 2012). Enteroviruses, the genera that include Rhinoviruses, are not surprisingly, thought to have infection rates that vary by season arising in summer and late fall and fading away in the winter (Center 2014). The Hepatitis A virus (HAV) infects the human liver and is more common in the developing world and other areas with poor sanitation and lack of clean water. Childhood infection rates, especially in the

third world, are high and studies have sought to investigate the overall prevalence and immunity rates in these areas (Brown et al. 1985; Hadjipanayis et al. 1999). Additionally, HAV infection rates have also been examined in adopted children from Africa and Asia (Abdulla et al. 2010). Lastly, the Poliovirus is the causative agent of Polio and is one of the more characterized and studied viruses in science. Although there is a mountain of research from an anthropological context on Polioviruses, they normally fall into one of three categories: person to person transmission (Blake et al. 2014), biological and cultural response (Anderson 2013; Dotzauer and Kraemer 2012), and strategies for global eradication (Kew et al. 2005).

Unassigned Order

There are a total of 19 viral families that have yet to be assigned to any order within the ssRNA+ taxonomy. Many of these are species specific and do not fit into the field of anthropology: Alphatetraviridae (moths and butterflies), Alvernaviridae (dinoflagellates), Barnaviridae (mushrooms), Carmotetraviridae and Permutotetraviridae (lepidopteran insects), Leviviridae (Enterobacteria, Caulobacter, Pseudomonas, Acinetobacter), Narnaviridae (fungi), and Nodaviridae (non-human vertebrates). There are also a half dozen other families of viruses that infect only plants: Bromoviridae, Closteroviridae, Luteoviridae, Potyviridae, Tombuviridae, Virgaviridae. Although it would take some time to mention all of the families of plant infecting viruses, it would be beneficial to address at least briefly, their potential importance to the anthropological field. To just use a single family as an example, Virgaviridae infect a wide range of domesticated fruits and vegetables: Furovirus (beets and wheat),

Hordeiviris (barley), Pecluvirus (peanuts), Pomovirus (beets and potatoes), Tobravirus (peas, peppers, and tobacco), and the most common genera Tobamovirus (cucumbers, hibiscus, peppers, tobacco, tomato, turnips and many more). These RNA viruses are capable of traveling through the entire GI tract unaltered and can provide information on host diet through the examination of modern and even ancient fecal material. The first ancient RNA virus genome (750 years old) was isolated from a barley grain and belonged to the Virgaviridae family (Smith et al. 2014). In terms of history, the tobacco and tomato mosaic viruses have greatly contributed to our understanding of viral genomes, being two of the early viruses studied by researchers (Bashir 2007). In terms of their effect on humans, it has been suggested that certain genera can induce immune response (Colson et al. 2010). One area of research that is still seemingly untapped is that of human response to viral infection of key regional crops. How do these impact overall yields? In what ways do these viruses diminish nutritional value? What areas of the world are most at risk for such viral infections in crops? There are currently many questions but few answers. Looking at it from another angle we can also inquire as to the way viruses have changed over the course of plants being domesticated, something that has already been suggested in a study of animal gut viromes (Lamendella et al. 2011).

Astroviruses belong to the Astroviridae family and are capable of causing diarrhea outbreaks in humans. In addition to the ssDNA Reoviridae family, Astroviruses have shown to be the cause of gastroenteritis in many young children across the globe (Victoria et al. 2007). Though no anthropological studies have directly looked at Astroviruses, certain developing nations like Brazil have taken it upon

themselves to identify novel strains responsible for youth deaths (Victoria et al. 2007). Viruses from the family can also infect other nonhuman organisms and there is no known anti-viral treatment.

Currently, no known anthropological associations with the Caliciviridae family have been made, but they have been used for studies of macroevolution within RNA viruses (Kitchen et al. 2011). Other viruses within the family include the Norwalk Virus (a norovirus, nicknamed the vomiting winter virus) which is responsible for a majority of nonbacterial gastroenteritis around the world and foodborne outbreaks in the United States. Togaviridae can infect a range of organisms and is responsible for both Rubella virus (which can cause congenital birth defects in many developing regions) and the Chikungunya Virus (transmitted by mosquitoes and has symptoms similar to Dengue). Within Togaviridae studies there is a need to keep those infectious viruses under control in at risk populations along with identifying future zoonotic viruses like the Ndumu virus, recently discovered in African pigs (Masembe et al. 2012).

The two final families within the Unassigned Order of ssRNA+ are Hepeviridae (Hepatitis E Virus) and Flaviviridae (Yellow Fever, West Nile Virus, Dengue Virus and Hepatitis C, among others). All of these diseases are commonly cited in medical anthropology texts and are each individually, unique cases of culture impacting pathogenic risks. Examples of such topics include disease strategies and vehicles of transmission for Hepatitis E (Mast and Krawczynski 1996), stagnant water presence and the risk of Yellow Fever (Singer and Baer 2011), the violent implications of West Nile Virus in Africa (Leopold 2005), and the ethnoecology of Dengue fever (Whiteford 1997)

Negative Single Stranded RNA Viruses

Viruses that are comprised of negative-sense single stranded RNA fall into either the Mononegavirales order or an order that is currently unassigned (Figure 8). Positive-sense ssRNA viruses are essentially mRNA and ribosomes can translate it into a protein easily. However, negative-sense ssRNA viruses must first be copied to the positive form by an RNA replicase enzyme that must be supplied by the virus itself. This classification contains many viruses that are pathogenic to humans and animals. Some of the pathogens infecting humans include the ebolavirus, influenzavirus and viruses that cause measles and mumps.

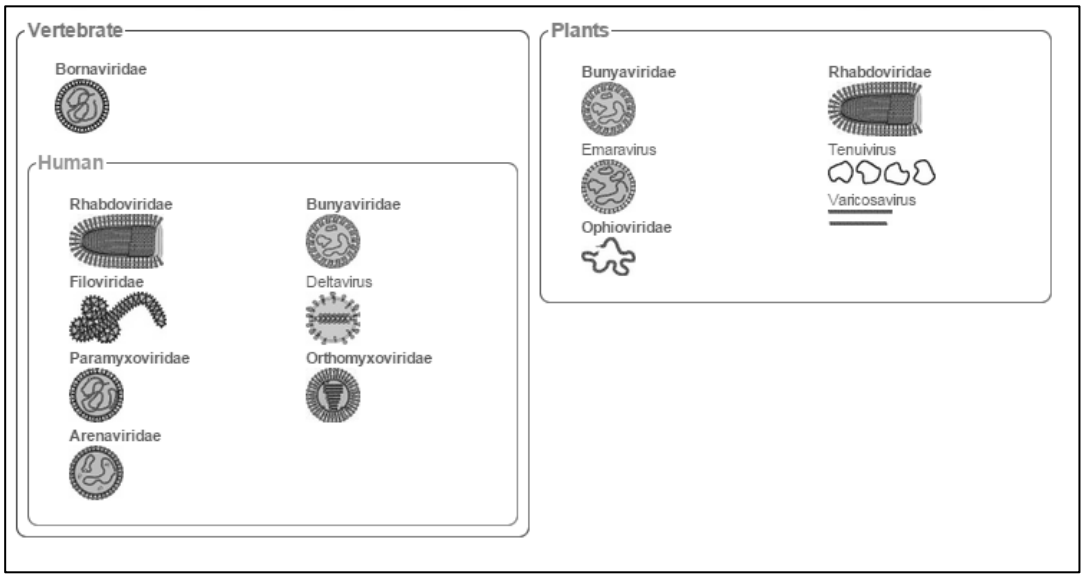

Figure 8: Negative single stranded RNA viruses (Hulo et al. 2011)

Mononegavirales Order

There are five families within the Mononegavirales order: Bornaviridae, Filoviridae, Nyamaviridae, Paramyxoviridae and Rhabdoviridae. Mononegavirales have a long history of integrating their genomes into the host. In fact, endogenous bornavirus-like elements (EBLs) have been shown to inhibit infection of exogenous bornaviruses in ground squirrels (*Ictidomys tridecemlineatus*) (Fujino et al. 2014).

Sequences like these are known as endogenous borna-like N (EBLN) elements and are found in a number of mammals today (Horie et al. 2010). Bornaviruses were thought at one time to cause diseases within humans but new research found now association between the presence of the Borna Disease Virus (BDV) and neurological disorders like schizophrenia, multiple sclerosis and many more (Hornig et al. 2012).

The Rhabdoviridae family of viruses have the capability of infecting a wide range of plants and animals. Novel genera within this family are discovered constantly, including some isolated in the Brazilian Amazon (Diniz et al. 2006). The lyssavirus genus contains within it the rabies virus. Rabies evolved around 1500 years ago (Nadin-Davis and Real 2011) and is transmitted through the saliva of a number of domesticated and wild animals like dogs, cats and bats. Rabies viruses are small (only 180 nanometers long and 75 nanometers wide and 12,000 base pairs in size) but incredibly powerful, with the ability to wreak havoc on the CNS of the host. Recent rabies studies show an interesting trend, mutation rates in the virus occur almost four times faster in tropical and subtropical climates compared to temperate ones (Streicker et al. 2012). According to the CDC, 40,000 humans in the U.S. receive post-exposure prophylaxis treatment for rabies every year. Thus, the anthropological relevance of this particular family of ssRNA+ is evident: not only because they cause much harm to humans but because their mutation rates in unexplored areas like the Brazilian Amazon make them ideal candidates for our understanding of viral evolution and zoonotic risks.

The family Paramyxoviridae is home to a number of human infectious diseases including measles and mumps. The measles virus, from the genus Morbillivirus, causes a highly contagious infection of the respiratory system and is one of few viruses with no

known animal hosts. Measles and mumps are still a problem in the United States, with 222 cases of measles reported in 2011 and 229 of mumps in 2012, but this is compared to 186,000 in 1967 before the vaccinations came about (CDC 2012). Many of the studies since the combination of the Measles-Mumps-Rubella vaccine into a single injection in 1971 have been related to mortality and exposure (Aaby et al. 1985). Unfortunately, a small group of anti-vaccination advocates who think that vaccines cause autism in children (concerns which are completely unfounded (DeStefano and Chen 2001)) have singlehandedly caused a resurgence in preventable diseases caused by those viruses in the Paramyxoviridae family. Anthropologists have taken to addressing these vaccine deniers and misinformation (Kata 2010) through panel discussions like "Measles and Mumps: Disease Comebacks Despite Vaccine Availability" by medical anthropologist Cameron Hay-Rollins of Miami University of Ohio.

The Filoviridae family is comprised of the deadly Ebolavirus and Marburgvirus. Since their discoveries in 1967 and 1980 (Preston 1995), Marburgvirus and Ebolavirus have been of particular interest to researchers. A PubMed search reveals over 1,500 results for both viruses combined. Though many examined the viruses' effect on tissues and their physical genomes, others draw attention to infection rates and viral hotspots. With the recent outbreak of Ebola around the world and the medical field scrambling to keep it under control, some have acknowledged the crucial missing piece in the Ebola solution: anthropologists (NPR Staff 2014). This is not news to our field, in fact whole journal issues have been dedicated to addressing the many issues with Ebola that anthropologists can address (Moran and Hoffman 2014) including the spread of Ebola

through funerals (Richards and Mokuwa 2014), government intervention in Liberia (Wesley 2014), and the impact of witchcraft (Bolten 2014). The Filoviridae family will continue to draw in anthropological research due mainly to its importation into the U.S. and Europe (Sarwar et al. 2011), the many cultural contexts (Hewlett and Amola 2003), and the need for infection control (Raabea and Borcherta 2012).

Unassigned Order

There are four families within the unassigned order, including a fifth group that are not assigned to any family. One of the viral families, Ophioviridae, only infect plants like lettuce and tulips. A second viral family, Bunyaviridae, is comprised of infectious human viruses including the Hantavirus. The Hantavirus is transmitted to humans through rodent droppings and can cause hemorrhagic fever with renal syndrome (Peters et al. 1999). Some controversy arose when researchers linked the Navajo Nation to Hantavirus Pulmonary Syndrome, leading the media to deem it the 'Navajo Disease' (Brugge and Missaghian 2006). Aside from a brief potential outbreak in Yosemite in 2012, the Hantavirus has remained out of the United States.

The sole genera within the Arenaviridae family, Arenavirus, can infect both rodents and humans. Members of the Arenavirus genus are classified as either Old World or New World in origin, but both likely arising from rodents (Briese et al. 2009). Viruses from this genus are known to cause hemorrhagic fevers including the Lassa virus which is common to Western Africa and the recently discovered Lujo virus (Briese et al. 2009; Brown and Kelly 2014). Anthropologists are attempting to identify

these hemorrhagic fever ‘hotspots’ and investigate permanent solutions to stop the spread of disease (Brown and Kelly 2014).

The largest family within ssRNA+ viruses is Orthomyxoviridae, comprised of Influenzaviruses A, B and C, and several other viruses. Influenza A can be broken down into 18 different subtypes and further broken down into many more strains. Influenza B does not have subtypes, but is merely divided into 2 lineages. This family of viruses is present within the human GI tract as well (Wootton et al. 2006). Influenza C is considered the mildest of the three, causing respiratory illness that is not associated with epidemics. The influenza pandemic during 1918 and 1919 killed between 24.7 and 39.3 million individuals worldwide (Patterson and Pyle 1991) and was actually thought to be caused by the avian H1N1 virus (Taubenberger 2006). The other pandemic influenza infections occurred in 1957 (H2N2) and 1968 (H3N2), the recent H2N3 which is a combination of both avian and swine forms (Ma et al. 2007). Influenzaviruses have been examined by anthropologists in a number of contexts including global response to disease emergence (Atlani-Duault and Kendall 2009; Singer 2009) and public health responsibilities (Kleinman et al. 2008). Researchers continue to use statistical models to determine the distribution and mortality rates of the Influenzaviruses across the globe (Chowell et al. 2011).

The final set of viral genera do not belong to any family and contain the Hepatitis D virus. Hepatitis D has the highest mortality rate of all the Hepatitis viruses and has recently been examined in relation to homosexual behavior (Prasetyo et al. 2014), young children (Abbas et al. 2014) and pregnant mothers (Rac and Sheffield 2014). Although co-infection is rare in the United States, many studies have examined it

(Negro 2014), and it is considered to be of particularly high prevalence in people of the Brazilian Amazon (Viana et al. 2005).

DNA Double Stranded Reverse-Transcribing Viruses

There are two families of viruses that are classified as DNA reverse-transcribing, both of which are not assigned to an order: Caulimoviridae and Hepadnaviridae (Figure 9). Caulimoviridae are 7-8kbp open circular genomes. Caulimoviridae is a virus that infects plants and insects including cauliflower, cassava, soybeans, rice, and tobacco. Hepadnaviridae are enveloped and spherical, infecting humans, apes and birds. Hepatitis B Virus is a member of the Hepadnaviridae family. Research shows that bats are likely an ancestral reservoir for primate hepadnaviruses including Hepatitis B (Drexler et al. 2013). There appears to be sex-related differences in hepatitis B infections in certain populations (Langendorfer et al. 1984), which may have applications to medical anthropology.

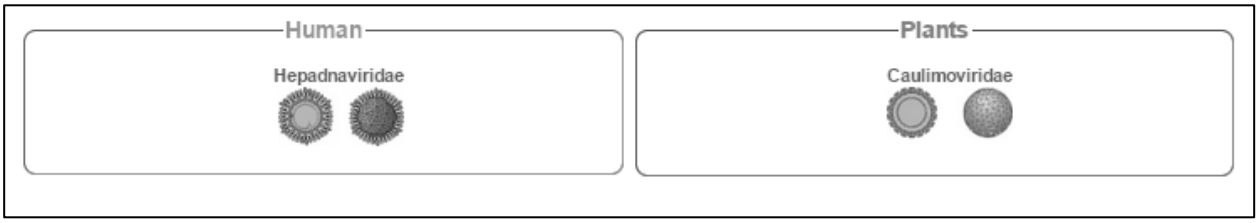

Figure 9: Double stranded reverse-transcribing DNA viruses (Hulo et al. 2011)

RNA Single Stranded Reverse-Transcribing Viruses

RNA single stranded reverse-transcribing viruses are comprised of the Retrotransposon order and the unassigned order (Figure 10). Viruses are contained within the host genome and go against the golden rule of biology in that they move from RNA to DNA to RNA then to a polypeptide. Retrotransposons are sequences

hidden within the genome of many organisms that can amplify themselves at random or when prompted. Vertebrate genomes have been completely transformed over time by the horizontal transfer of retrotransposons, including humans (Walsh et al. 2013). It is estimated 46% of the human genome is made up of retrotransposons (Cordaux and Batzer 2009). This includes LINEs (long interspersed elements) which make up 21%, SINEs which make up 13% (Lander et al. 2001), human endogenous retroviruses (HERVs) make up 9% (Feschotte and Gilbert 2012), and DNA transposons make up 3% (Cordaux and Batzer 2009). There are two families within this order, Pseudoviridae and Metaviridae, both of which are currently known to infect the Saccharomyces fungi and Drosophila.

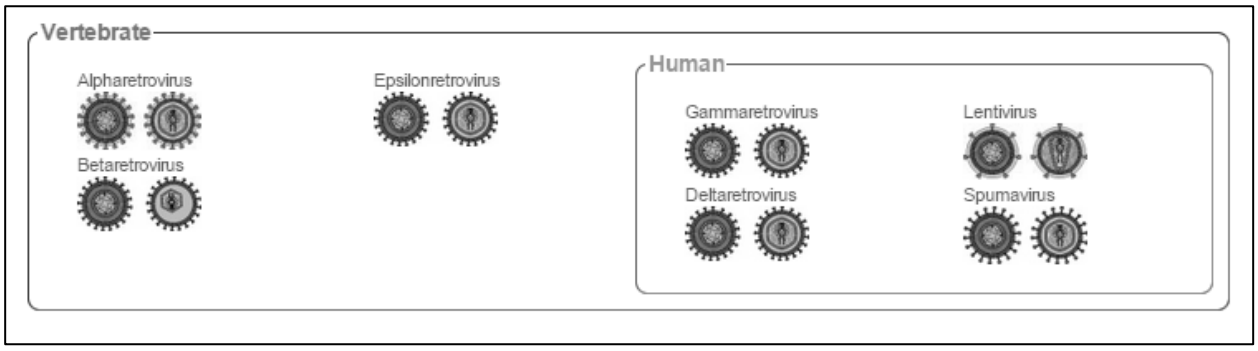

Figure 10: Single stranded reverse-transcribing RNA viruses (Hulo et al. 2011)

Unassigned Order

The unassigned order of ssRNA reverse transcribing viruses is comprised of only a single family, Retroviridae, which is split into two subfamilies, Orthoretrovirinae and Spumaretrovirinae. These viruses are normally between 7kbp and 12kbp in size and are thought to evolve very rapidly through mutations, recombination and genome re-assortment. There are essentially two types of retroviruses. Those exogenous ones, like HIV that reproduce within the host and spread through sexual intercourse and blood

contact, and those endogenous ones that have invaded the germlines of almost every species of vertebrate (Ryan 2004). Simian viruses including the nonhuman primate form of HIV belong to Retroviridae. Some other viruses to note in this family include the human T-cell lymphotropic virus (HTLV) which may be the cause of germline diseases like hemophilia and breast cancer (Deininger and Batzer 1999), and the Rous sarcoma virus, an avian retrovirus which was one of the first viruses discovered that affect the cell cytoskeleton (Taylor et al. 2011).

Retroviruses that reside within primates are helpful for a number of reasons. First, studies in Old World primates provide an in depth look at the evolution of endogenous retroviral sequences over time and allow us to create phylogenetic trees (Johnson and Coffin 1999). We are also able to gather insights as to the prevalence of these ERVs within and across species through primate studies (Gogarten et al. 2014). Lastly is the potential threat of novel zoonotic viruses, one example being the Simian Foamy virus of the Spumavirus genera. A major concern is that the foamy virus is transmitted to humans who consume bushmeat or are in close contact with infected primates (SFV is rampant across both captive and wild nonhuman primates, at prevalence rates of 75 to 100% (Betsem et al. 2011)). A study from Cameroon showed that a number of locals were infected with SFV after close contact with nonhuman primates (usually attacks against humans) but the pathogenicity and the possibility of human to human transmission of the virus is not yet known (Betsem et al. 2011). An anthropological take on this study involved examining the awareness of hunters in Sierra Leone to the zoonotic risks of bushmeat: 24% of them knew of any risk (Subramanian 2012).

The most studied viruses of the Retroviridae family belong to the Lentivrus genera and include the simian, feline and human forms of the immunodeficiency virus (Figure 11). These viruses are around 80-100nm in diameter with a genome of almost 10kbp. Non-human primate models of AIDS actually help to advance the field of knowledge of HIV progression substantially, of particular interest if the response of B cells in these animals (Schmitz and Korioth-Schmitz 2013). One of the more interesting recent discoveries, utilized metagenomic analysis investigated the enteric virome with SIV infected Rhesus monkeys (Handley et al. 2012). Researchers discovered 32 previously undescribed viruses from potentially pathogenic genera and suggest that the enteric virome helps to progressive SIV to AIDS by effecting immune activity and intestinal epithelial health (Handley et al. 2012). Further study of SIV within nonhuman primates could reveal even more information about the mysterious progression of this still present epidemic. The human form is considered to have arisen from cross-species contamination sometime during the early 20th century (Sharp and Hahn 2011).

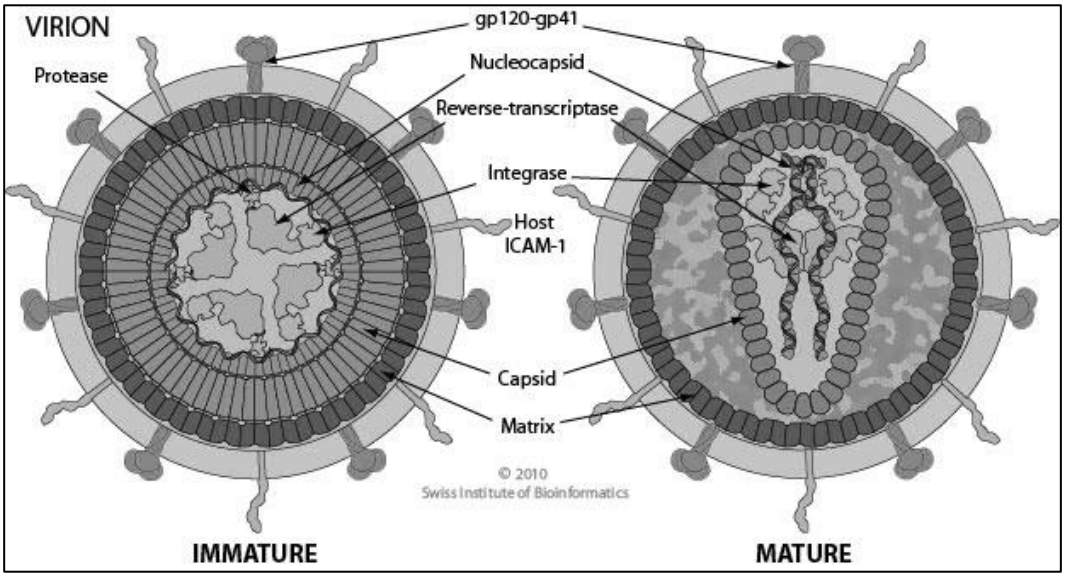

Figure 11: Lentivirus (HIV-1, HIV-2, etc.) (Hulo et al. 2011)

Lastly, and most importantly, is the discussion of HIV/AIDS to the field of anthropology. Although there is no way to synthesize everything into even an entire dissertation, some of the key components of the study of HIV will be mentioned here. The HIV genome was first discovered in 1983 and is now separated into two distinct subtypes: HIV-1, which is predominant across the globe and HIV-2, which is less easily transmitted and has a longer incubation period. HIV has several structural proteins including *gag* (group-specific antigen), *pol* (reverse transcriptase coding), and *env* (envelope coding). In short, the HIV virus infects and kills cells within the human immune system like CD+ T cells, macrophages, and dendritic cells. Once the virus enters the cell and integrates itself into the host genome, it may lie dormant for some time. The virus can remain undetectable for an HIV antibody test for a month or more and when it does start to replicate it does so faster than the immune system can halt it. As the host immune system weakens, any number of infections or illnesses can wreak havoc upon the host. HIV/AIDS is much more than just a viral infection and cannot be understood from a solely only from a biological perspective (Baer et al. 1997).

The spread of AIDS over the last several decades has exposed the "hidden vulnerabilities in the human condition" (Fineberg 1988). As illustrated by medical anthropologist Paul Farmer's research, HIV infection disproportionately affects the poor and socially deprived. Presently, there six million new HIV infections each year, with 40,000 of those occurring within the United States (Baer et al. 1997). Some 95% of AIDS cases appear in the world's poorest countries making it the fourth biggest killer worldwide (Baer et al. 1997). Blacks and Hispanics in the U.S. comprise 27% of the

population (in 1997) but also comprised 55.8% of AIDS cases (Baer et al. 1997). AIDS is a disease of society, created by its stratified and widely oppressive structure.

The current AIDS crisis in Africa is the result of societal, political, and economic orders that created a 'risk environment of violence and sexual exploitation' (Barnett and Whiteside 2002). The effects that this globalization and power has had on the HIV/AIDS atmosphere in Africa is illustrated in the collection: *Morality, Hope and Grief: Anthropologies of AIDS in Africa* (Dilger and Luig 2012). The AIDS crisis of the 1980's extended beyond Africa into the Caribbean and the United States. From an anthropological perspective, the first stateside study is considered to be interviews with gay men in New York City in 1985 (Feldman 1985). Though there are many research categories, the disease is primarily investigated from one of the following viewpoints, which have a good deal of overlap: epidemics outside of Africa, public knowledge and perception of the disease, personal experiences, prevention and response, and risk-associated behaviors (intravenous drug use, prostitution, homosexual intercourse) (Baer et al. 1997).

The outbreak of AIDS in Haiti in the 1980s was actually first attributed to voodoo (Moses and Moses 1983) and as Farmer notes, some even speculated cannibalism as a root cause of the spread (Farmer 1990). The real reason, a much more logical one, was due to the Haitian economy, which forced many into prostitution as a sole means of survival mainly in the red light district of the nation's capital Port-Au-Prince (Farmer 1990). The rest of the Caribbean was also affected in the 1990s including Trinidad and Tobago, the Bahamas and Jamaica. Some other largely populated countries like Cuba somehow avoided the epidemic, Cuba had a 0.9% HIV

infection rate in 1990 while Haiti's was 9.0% (Baer et al. 1997). Follow up anthropological investigations suggest that Cuba avoided the AIDS epidemic by having a highly educated population and a well-developed health education infrastructure (Hansen and Groce 2001). AIDS in India resulted in a high levels of discrimination and stigmatization in the early 1990s (Baer et al. 1997). The many sex workers and intravenous drug users in regions of Thailand saw a growing number of HIV infections (Baer et al. 1997). In the early 80's in the U.S., many considered it to be a disease of homosexuals but intravenous drug users were making up a majority of new AIDS cases in the Northeast at the time (Crimp 1988).

In India, HIV awareness significantly increased between 2003 and 2009 though there is still a general stigma of not speaking about sex across much of the population. Such societal views are affecting efforts to spread knowledge regarding safe sex and HIV/AIDS transmission (Bradley et al. 2011). Yet some investigations into HIV spread in South Africa show the opposite, focus group studies from Orange Farm suggest that culture is not a determining factor in HIV transmission (Saethre and Stadler 2009). The answer is likely somewhere in between. Models of the transmission of HIV must take into account the plasticity of human behavior and how it is altered based on certain contexts (Sattenspiel and Castillo-Chavez 1990).

In terms of identity, AIDS is considered to be analogous with leprosy, in that they are both "I am" diseases (Estroff 1993). Although the public discourse considered AIDS to be an "I have" disease, those infected with the virus would refer to it as an "I am" disease in casual conversation (Singer 1998). From a critical medical anthropology perspective, the sufferer experience is a social product constructed from both the

socially constituted categories and the political-economic forces that governs daily life (Baer et al. 1997)

HIV infection tends to be separated into two categories: pattern I, which involves same sex or bisexual partners and intravenous drug users, common in North America, Western Europe, Oceania, and parts of Latin America, while pattern II is associated with heterosexual contact in Africa, parts of Latin America, and the Caribbean. HIV has been studied in children around the globe (Epstein 1990). With an estimated 5.5 intravenous drug users in the world, stopping a some of Pattern I comes with its own challenges (Singer 1997). Singer speaks of how developing needle exchange programs and education may be more effective in HIV prevention than trying to stop individuals from using drugs altogether (Singer 1997). The ramifications of pattern II is the rise in pediatric AIDS. Anthropologists have children with AIDS co-infected with EBV and shown that a majority of children are immune to EBV and thus have a reduced risk of developing associated lymphomas (Petrara et al. 2014).

Lastly, HIV infection and understanding involves the social discussion of the topic. HIV infection can be a taboo subject, and gossip and rumors within communities establish a set ideas of acceptable behaviors, ideas and talking points surrounding infection (Smith et al. 1999). There are several common areas of blame once HIV infection has progressed into AIDS. Within Africa, elders see it as curable, ancient disease, while younger men and women see it as a modern, foreign disease(Stadler 2003). Another common blame assignment within indigenous societies is an association with witchcraft and sorcery (Thomas 2008). Such beliefs can have disadvantageous effects on prevention and treatment of HIV and AIDS (Thomas 2007).

Concluding Statements

The intention of this chapter was to not only describe the anthropological associations with viral studies, but stress all the missing information. So many viruses are not understood amidst humans. Are they harmful? Informative? Transient? Viruses and humans are inexorably linked through a long history of evolutionary circumstances. Likely the most important avenue of investigation is from a health perspective. There is still so much to understand about within and between species transmission. A recent study found novel viruses from six genera with capabilities of infecting humans in New York sewer rats (Firth et al. 2014). The next zoonotic transmission scare could belong to a ubiquitous street animal in the United States. There is a similar scenario with bats in China, in which many unknown potentially deadly viruses housed within a very common animal living in close quarters with humans (He et al. 2013). Even cats harbor a high level of viruses with asymptomatic infections (Ng et al. 2014). A pretty scary thought.

Humans are constantly in contact with viruses in our water, soil, air, fresh fruit and vegetables, seafood dairy and on almost every surface but we still know very little about what's there and what can harm us (Rzezutka and Cook 2004). That's not even including the viruses within and on every human: what role do they play? The anthropological field is in need of a rejuvenation of viral investigations, revising of the traditional concepts of medical anthropology, and integration of novel technologies to better understand genomic movers and shakers: viruses.

Chapter 3

The Geographic Structure of Gut Viruses

An alien visiting our planet, given a different sensory range that could directly detect viruses, would consider them the dominant form of life.
Forest Rohwer (Rohwer and Barott 2013)

How is that you keep mutating and can still be the same virus?
Chuck Palahniuk, *Invisible Monsters* (1999)

The field of viral metagenomics seeks to examine the full scope of viral DNA (and RNA) genetic sequences within a given environment. As illustrated in previous chapters, viruses have been examined using this shotgun methodology in a variety of ecosystems: soil, fresh water and deep oceans, the troposphere, and of course, the human body. Early human viral metagenomic studies revealed that most enriched virus sequences from the human gut did not match any genomes in current databases (Breitbart et al. 2003). Most of the viruses within us were, and continue to be, a complete mystery. Although databases have improved in the dozen years since the first human viral metagenomic study, the question still remains: how informative are the viruses of the gastrointestinal tract? This chapter uses shotgun metagenomic information to address the question of viral geographic structure within populations from North and South America. I test the hypothesis that dsDNA human gut viruses are structured based on host geographic region. In order to test this hypothesis, I collect viral data from 22 fecal samples representing human populations in North America (Cheyenne and Arapaho of Oklahoma) and South America (San Mateo, Tambo de Mora, and Chincha communities of Peru). Two different aspects of viral diversity are

presented: 1) the similarity of predicted viral proteins between populations and 2) the amino acid variation among the most commonly shared viral protein across all samples. In the former, the presence or absence of predicted viral protein types determines whether gut viruses adhere to a population specific pattern or occur randomly within the host GI tract. In the latter, I test whether the specific amino acid sequences within the most common predicted viral protein can accurately visualize host population structure.

Population Background and Sampling Strategies

There are three South American communities from Peru represented in the current study, here referred to as Tambo de Mora, Chincha, and San Mateo. Tambo de Mora and Chincha are from the coastal region of Peru (Figure 12). An Afro Peruvian community from Ica expressed desire to have its name withheld from microbiome and virome publications and will thus be referred to as Chincha in the current text. Tambo de Mora consider themselves to be Afro Mestizo in ethnicity. The Afro Peruvians have been a part of Peruvian society for almost five centuries and consider their identities to be a combination of African and Native South American culture (Bowser 1974).

The final Peruvian population included in this study is the San Mateo of the Amazon jungle. Within the region of Loreto, the province of Requeña, and the district of Yaquerana, San Mateo is comprised of 13 communities located on the border of the Galvez River in Northeast Peru. The people of San Mateo are of the Matses ethnicity and many speak the Matses language exclusively. There are an estimated 3,200 Matses tribal members, 2,000 of whom live on tribal land. The Matses subsist only on the flora

and fauna of the jungle and are geographically isolated from other Peruvian populations (Collis 2012; Svensson 2007).

Informed consent from each Peruvian participant was obtained in accordance with the Peruvian National Institute of Health and the University of Oklahoma Human Subjects Review Board. As part of our research engagements with these communities, all participants provided their perspective on microbes and microbiome studies. Subsequently, these samples were collected by researchers from the laboratories of Dr. Cecil M. Lewis at the University of Oklahoma. When possible, phenotypic information such as sex, age and BMI was collected from participants. Samples from each location were collected using polypropylene collection units and transported over four days on ice until reaching Lima, Peru. Soon after, the samples were shipped to the US and arrived at the University of Oklahoma for analysis.

The Cheyenne and Arapaho (C&A) were initially contacted to participate in a research investigation related to microbiomes in indigenous peoples. University of Oklahoma anthropologist, Dr. Paul Spicer, and his research team contacted representatives from five Oklahoma towns (Hammon, Clinton, Concho, Geary and Kingfisher (Figure 13) within the C&A Nation in order to recruit adults to participate in focus discussions related to knowledge of microbiomes. After four rounds of focus group discussions, a total of 40 individuals agreed to participate in fecal sample donation in order to study gut microbiomes, viromes, and proteomes.

Figure 12: A map of Peru. The three communities are highlighted along with the capital, Lima.

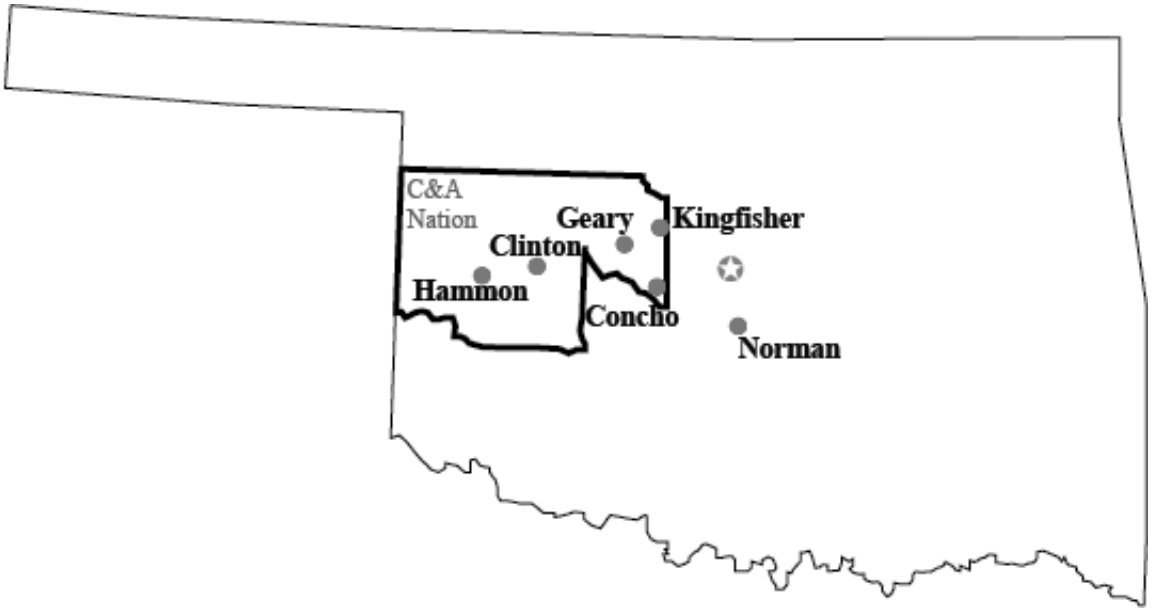

Figure 13: Map of Oklahoma. It shows the Cheyenne and Arapaho Nation and the five towns sampled within it. The Norman population and the capital city, Oklahoma City, are also noted on the map.

The Cheyenne and Arapaho have long closely been related, but in 1867, both groups were co-assigned the Western Oklahoma reservation under the Medicine Lodge Treaty. The C&A are a federally recognized tribe and one of 39 tribal nations within the state of Oklahoma. As of 2011, the C&A have 12,185 enrolled members (with at least ¼ Cheyenne or Arapaho blood quantum), 8,664 of whom live within tribal jurisdiction in Oklahoma (Oklahoma Indian Affairs Commission 2011). Type 2 diabetes risk across all American Indians and Alaska Natives is 182% higher than the rest of the U.S. population (IHS 2013). Diseases associated with obesity were a major health concern to the members of the C&A nation during our focus groups. Although no recent estimates for obesity among the C&A have been published, BMI results from our dataset show that 39 of the 40 participants were classified as overweight or obese and self-reported type 2 diabetes was documented in 19 of the 40 participants.

For sample collection, at a time arranged by the participants, I dropped off a sample collection kit containing a polypropylene fecal collection unit and a saliva collection unit (Figure 14). Upon arrival, participants were first consented and given a sample collection kit with necessary instructions. Anthropometric and phenotypic data were collected, and a pickup time and location was arranged. Anthropometric measurements were taken using a stadiometer (for height) and a standard bathroom scale (for weight). These values were used to calculate body mass index of the participants. In addition to anthropometric measurements, participants were asked several demographic questions such as sex and age. Participants were also questioned about their previous antibiotic or probiotic use, smoking habits and if they are affected by type 2 diabetes. Samples were transported back to the Stephenson Research and

Technology Center on the University of Oklahoma campus and placed in a -80°C freezer until extraction.

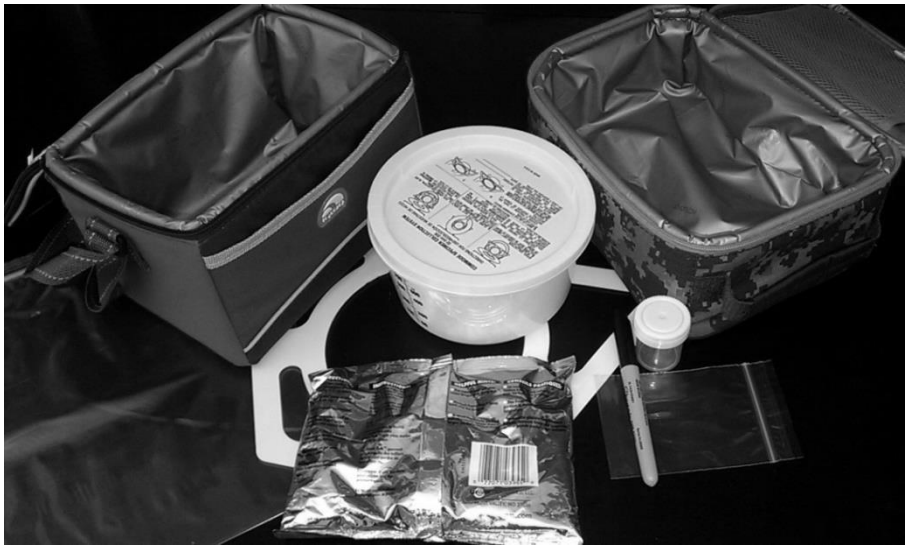

Figure 14: Sample collection kits

In addition to the native and Afro-Peruvian communities from North and South America, we included an outgroup population representing individuals from the University of Oklahoma in Norman, Oklahoma. Study volunteers primarily consisted of academics and their families. A majority of participants were of normal BMI and between the ages of 20 and 50. This cohort is used here to represent a non-native, metropolitan Oklahoma population.

The analyses within this chapter use an enriched viral dataset which includes 22 samples (5 C&A, 5 Chincha, 7 San Mateo, 5 Tambo de Mora). The unenriched viral dataset, used in chapter 4, includes 86 samples (24 San Mateo, 39 C&A, 23 Norman). The viral enriched dataset underwent specific laboratory methods intended to increase the total amount of viral DNA within the shotgun sequencing data. The second, unenriched dataset involves complete shotgun sequencing (not enriched for viruses) of San Mateo, C&A, and Norman samples. These data have a very low number of viral

reads and cannot be used for the analyses demonstrated in the current chapter. Information on all of the samples within the present dissertation can be seen in Tables 2, 3 and 4.

SAMPLE	POP	SEX	AGE	BMI
CAG6	C&A	F	54	26.06
CAH5	C&A	M	50-60	31.82
CAH8	C&A	M	50-60	44.01
CAH11	C&A	M	50-60	37.40
CAK2	C&A	F	44	33.44
CH02	CHINCHA	F	19	18.14
CH07	CHINCHA	F	9	24.30
CH10	CHINCHA	F	58	28.51
CH12	CHINCHA	F	16	19.60
CH13	CHINCHA	F	41	21.64
SM01	SAN MATEO	NA	NA	
SM23	SAN MATEO	M	7	15.66
SM28	SAN MATEO	F	52	19.47
SM29	SAN MATEO	F	50	17.90
SM34	SAN MATEO	M	4	15.62
SM40	SAN MATEO	F	18	NA
SM42	SAN MATEO	M	4	16.73
TM06	TAMBO DE MORA	F	31	29.76
TM11	TAMBO DE MORA	M	39	27.66
TM18	TAMBO DE MORA	F	13	23.19
TM23	TAMBO DE MORA	M	1	21.33
TM28	TAMBO DE MORA	F	32	27.01

Table 2: Information regarding virus enriched samples

SAMPLE	POP	SEX	AGE	BMI
C01	C&A	M	50-60	31.04
C02	C&A	F	50-60	40.77
C03	C&A	F	27	35.02
C04	C&A	M	30	33.13
C05	C&A	F	84	27.25
C06	C&A	F	51	44.21
C08	C&A	F	33	37.97
C09	C&A	F	29	48.25
G01	C&A	F	34	28.20
G02	C&A	M	45	39.10
G03	C&A	M	39	40.32
G04	C&A	M	69	36.26
G05	C&A	F	68	40.31
G06	C&A	F	54	26.06
G07	C&A	F	65	40.70
G08	C&A	F	41	26.22
H01	C&A	F	60+	24.66
H05	C&A	M	50-60	31.82
H06	C&A	M	60+	30.00
H07	C&A	M	56	42.11
H08	C&A	M	50-60	44.01
H09	C&A	F	50-60	36.33
H10	C&A	F	40-50	24.93
H11	C&A	M	50-60	37.40
H13	C&A	M	62	29.95
H14	C&A	F	44	43.70
K1	C&A	F	69	35.01
K2	C&A	F	44	33.44
K3	C&A	F	49	27.99
K4	C&A	M	66	29.55
K5	C&A	F	50	32.64
K6	C&A	F	29	47.40
K7	C&A	F	29	31.87
K10	C&A	M	55	35.07
K11	C&A	F	55	36.15
CN02	C&A	M	43	27.84
CN03	C&A	F	43	32.28
CN04	C&A	F	20	27.29

CN05	C&A	M	21	25.20
NO01	NORMAN	M	23	21.69
NO02	NORMAN	F	37	20.52
NO03	NORMAN	M	40	23.37
NO04	NORMAN	M	26	24.16
NO05	NORMAN	M	28	22.19
NO06	NORMAN	M	28	23.49
NO07	NORMAN	F	32	21.92
NO08	NORMAN	F	32	20.01
NO09	NORMAN	F	34	23.77
NO10	NORMAN	M	41	26.58
NO11	NORMAN	M	26	23.93
NO12	NORMAN	F	27	28.62
NO13	NORMAN	M	35	20.34
NO14	NORMAN	F	10	14.97
NO15	NORMAN	F	50	25.92
NO16	NORMAN	M	47	30.86
NO17	NORMAN	M	10	21.53
NO18	NORMAN	M	7	17.31
NO19	NORMAN	F	32	19.30
NO20	NORMAN	M	26	27.86
NO21	NORMAN	M	23	24.78
N022	NORMAN	M	26	30.22
NO23	NORMAN	F	26	26.53

Table 3: Information regarding unenriched virus samples from Oklahoma

SAMPLE	POP	SEX	AGE	BMI
SM01	SAN MATEO	NA	NA	NA
SM02	SAN MATEO	F	25	23.95
SM03	SAN MATEO	M	10	15.77
SM05	SAN MATEO	M	1	NA
SM11	SAN MATEO	F	4	15.78
SM18	SAN MATEO	F	36	26.44
SM20	SAN MATEO	F	20	21.27
SM23	SAN MATEO	M	7	15.66
SM24	SAN MATEO	M	2	17.85
SM25	SAN MATEO	F	2	19.33
SM28	SAN MATEO	F	52	19.47
SM29	SAN MATEO	F	50	17.90
SM30	SAN MATEO	M	4	18.26
SM31	SAN MATEO	M	30	22.60
SM32	SAN MATEO	F	21	22.14
SM33	SAN MATEO	F	5	16.45
SM34	SAN MATEO	M	4	15.62
SM37	SAN MATEO	M	12	19.84
SM39	SAN MATEO	F	40	28.93
SM40	SAN MATEO	F	18	NA
SM41	SAN MATEO	M	6	16.35
SM42	SAN MATEO	M	4	16.73
SM43	SAN MATEO	F	2	15.15
SM44	SAN MATEO	M	4	19.83

Table 4: Information regarding unenriched virus samples from Peru

Laboratory Methods

Metagenomic data were generated using two protocols. The protocols are very similar; however, one protocol includes a viral enrichment process that follows DNA extraction. Viral enrichment followed a previously published protocol (Phan et al. 2011). In a 2mL screw cap tube, between 0.25g and 0.50g of feces was placed with a 1.5:1 ratio of Hanks Buffered Saline Solution (HBSS). A total of 1.0mm zirconia/silica beads were placed into the stool solution and vortexed vigorously for 10min. A cut

pipet tip was used to pull 500µl of the stool solution and then placed into a new 1.5mL tube which was then centrifuged at 13,000 rpm for 5min. The supernatant was removed and placed into a fresh tube. If the volume was less than 400 µl, additional HBSS was added to the original stool solution and the previously mentioned steps were repeated. The final stool solution was then spun again at 13,000 rpm for 5min and the supernatant was once again pooled into a clean 1.5mL tube. These steps were repeated an additional time. Then 200 µl of the supernatant was removed and put through an Ultrafree-MC HV 0.45µm sterile filter (Millipore, UFC30HV0S). The filter tubes were centrifuged at 6,000rpm for 5min and 140 µl of the filtrate was transferred to a new screw cap tube. For the final step of enrichment, roughly 110µl of the filtered sample was combined with Turbo DNase (7µl), Baseline Zero DNase (3µl), Benzonase (3µl), RNase A (3µl) and 10xTurbo Buffer (14µl) and place at 37°C for 2hrs. Although the total number of viruses within metagenomic studies using these methods vary significantly, this enrichment procedure ensures that a higher number of viral sequences are represented and the contig sequences are longer when assembled, increasing the likelihood of matches to reference databases (Hall et al. 2014).

The DNA was then extracted using the QIAmp Viral RNA Mini Kit (QIAgen). A total of 560µl of the AVL buffer was pipetted into a 1.5ml tube with 140µL of the DNase/RNase treated solution and vortexed for several seconds. The solution was incubated at room temperature for 10min. After a brief centrifuge, 560µl of 100% ethanol is added to the sample and vortexed for 15sec followed by another centrifuge for several seconds. A total of 630µl of the solution was added to the QIAmp Mini column and centrifuged for 8,000rpm for 1min. The tube was discarded and the filter

was placed into a fresh tube. The previous step was repeated until all of the solution had been used. A total of 500µl of Buffer AW1 was added to the spin column and centrifuged at 8,000rpm for 1min. The spin column was placed in a new 1.5ml tube and 500µl of Buffer AW2 was added and centrifuged at 14,000rpm for 3min. The spin column was placed in a fresh 1.5ml collection tube and 60µL of RNase free water was added. The column was incubated for 1min and centrifuged at 8,000rpm for 1min. The spin column was discarded and the solution then underwent a REPLI-g Midi reaction (Qiagen) according to manufacturer protocol. The resulting 50µl of Repli-G solution was diluted with 50µl of Sigma Water and visualized in an agarose gel. Samples were prepared for shotgun sequencing at Yale University Genome Center using the TruSeq DNA PCR-Free Sample Preparation kit and sequenced using an Illumina HiSeq 2500 with paired-end 100bp reads. A total of 200ng for each sample was added to the pool by the sequencing center at Yale University.

The samples that were not enriched for viral sequences underwent slightly different laboratory methods. DNA from feces (39 C&A and 23 Norman participants) was extracted using the Power Microbiome RNA Isolation Kit (MoBio) according to manufacturer protocols (with slight amendments: bead beating and no use of RNase). An additional bead beating step was added in which approximately 250mg of fecal material was inserted into a 2mL tube filled with 0.1mL of 1.0mm Zirconia/Silica beads (Biospec) and 500mL of Lysis Buffer (MoBio). The extracts were sequenced at the Oklahoma Medical Research Foundation. The libraries were prepared using the Illumina Truseq DNA library preparation kit (Illumina) according to manufacturer's protocol. The DNA was then sheared, measured with a fluorometer, and pooled in

equimolar concentrations. Each pool of eight was sequenced on an Illumina Hiseq 2500 instrument with paired-end 100bp reads. The software Sickle (Joshi and Fass 2011) was used to filter out low quality reads (<30QC) and for the enriched virome dataset, 90% of reads passed initial quality control. A parallel genome assembler called Ray (21bp K-mer length specification) (Boisvert et al. 2010) was used to create the viral contigs. Those contigs that were below 100bp were removed from analysis. Open reading frames (ORFs) were identified using Fragenescan (Rho et al. 2010): this program is used to identify potential genes using short, error prone reads. These ORFs are predicted proteins. When necessary, basic local alignment search tool (BLAST) queries were performed using an e-value of 10^{-5}.

Presence of bacteriophages was confirmed through visualization using a Transmission Electron Microscope (ZEISS 10A). Sample G6 was selected for TEM due to its high viral content. After initial 0.45µm sample filtration, 3µl were placed onto a mesh copper grid for several minutes before being wicked off. The grid was then stained with 2% PTA in a 1% trehalose solution (pH=8.0) for 5min. The grids were visualized with the ZEISS 10A machine and photographs were taken. Figure 15 shows a bacteriophage that is likely a T4 phage at 50k magnification.

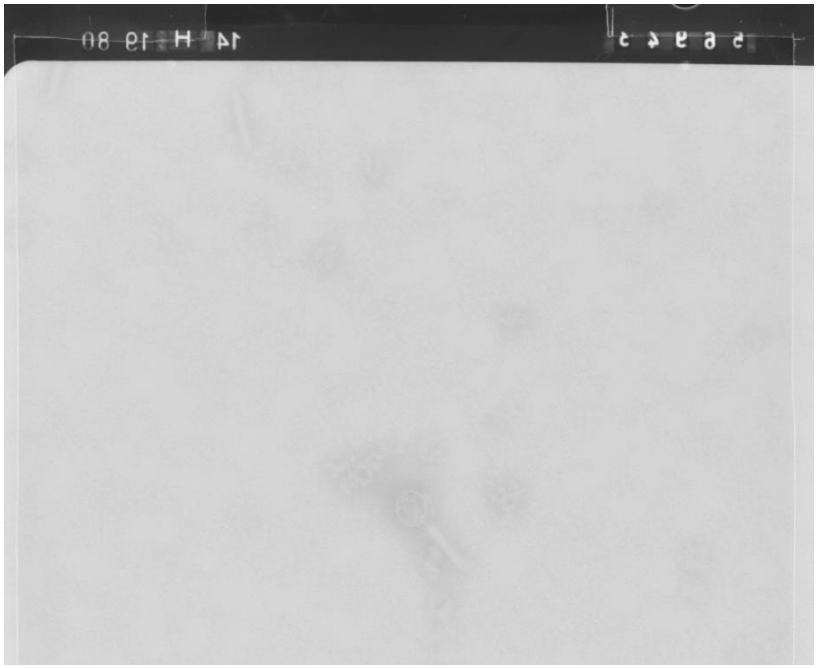

Figure 15: Photograph of bacteriophage (lower middle) (from sample CA G6 at 50,000x magnification)

Predicted Proteins and Geographic Structure

In order to address the question of the geographic structure of DNA gut viruses, I used the predicted protein sequences with the program UCLUST (Edgar 2010). Sequences were reordered based on size before UCLUST analysis which allows for the longest sequences to act as references for all subsequent ones (Edgar 2010). Those sequences that match the seed sequence (the longer contig sequence) at a 60% amino acid identity or higher are put into a uniquely numbered 'cluster'. Those ORFs that do not match any of the other sequences (singletons) are not useful for phylogenetic analysis and omitted. A total of 6,389 clusters were exhibited in the virome enriched ORF data with some 44,000 singletons. Using this information, I created an OTU table (in this case, OTU is synonymous with ORF) to produce a biom table, a format for the analytical software QIIME. This table provides relationships between the ORFs through presence/absence among each of the 22 samples. Within QIIME, I created a distance

matrix using the biom file which examines this presence/absence relationship using the binary jaccard method. MOTHUR (Schloss et al. 2009) was then used to visualize these relationships and a UMPGA tree through the Figtree algorithm. The results of these analyses can be seen in Figure 16.

The phylogenetic tree based on the similarity between predicted protein sequences in Peru and Oklahoma reveal a clustering pattern consistent with geographic structure. The longer branch lengths suggest a tremendous amount of overall diversity within the virome enriched sample set. Populations from Peru, for the most part, cluster together (aside from GU02). The differences between the three Peruvian communities are not distinguishable using this method.

Principal coordinates analysis (PCoA) plots are used to reduce the complexity of a dataset in order to easily visualize it in a scatterplot. The 22 samples are visualized across principal component 1 (accounting for 14.05% of variation) and principal component 2 (11.12%) (Figure 17). Given that each population shows less overall variation within compared to between populations, this plot is consistent with geographic structure. A similar conclusion can be arrived at when examining PC1 and PC3 (Figure 18). These plots provide compelling evidence that using predicted proteins, DNA gut viruses cluster based on geographic regions. Predicted proteins are more commonly shared between the Peruvian coastal populations (Chincha and Tambo de Mora) than the jungle San Mateo population and indigenous Cheyenne and Arapaho of Oklahoma. The observed structure may, in part, be attributed to different health variables and demographics of these populations. For example, a Spearman rank correlation coefficient between PC1 and BMI shows a statistical significance

(p=0.0008, corr=0.6977), as well as PC1 with age (p=0.01926; corr=0.5060). However, the points in the PC1 and PC2 scatterplots for age and BMI do not cluster tightly based on these variables, further suggesting that the strongest predictor of shared protein sequences is host geography.

It should be noted that in order to establish whether contamination interfered with any subsequent analysis, I examined the sequences present in the extraction blank (water) and PCR negative control which were included in the enriched shotgun pool. It was important to assess any potential contamination the extraction kit and Repli-G reaction may have introduced into the sample set that could modify conclusions. First, the predicted protein sequences for all samples, including the two blanks, were rarefied to 25,000 reads within QIIME. The sequences that showed matches for either the extraction or PCR negative control were removed from analysis except those containing 5 or fewer reads. Even removing 22,000 sequences from the negative control did not alter the sum of any of the sample reads by more than 100 sequences. In other words, the small amount of contamination that was present was not enough to alter any conclusions regarding geographic structure (most of the sequences in the blanks were likely singletons). As one last precaution, the filtered predicted protein sample data with extraction and PCR blank protein sequences omitted were converted into an OTU table and subsequently a PCA plot. These principal components looked nearly identical to those with the sequences from the blank included.

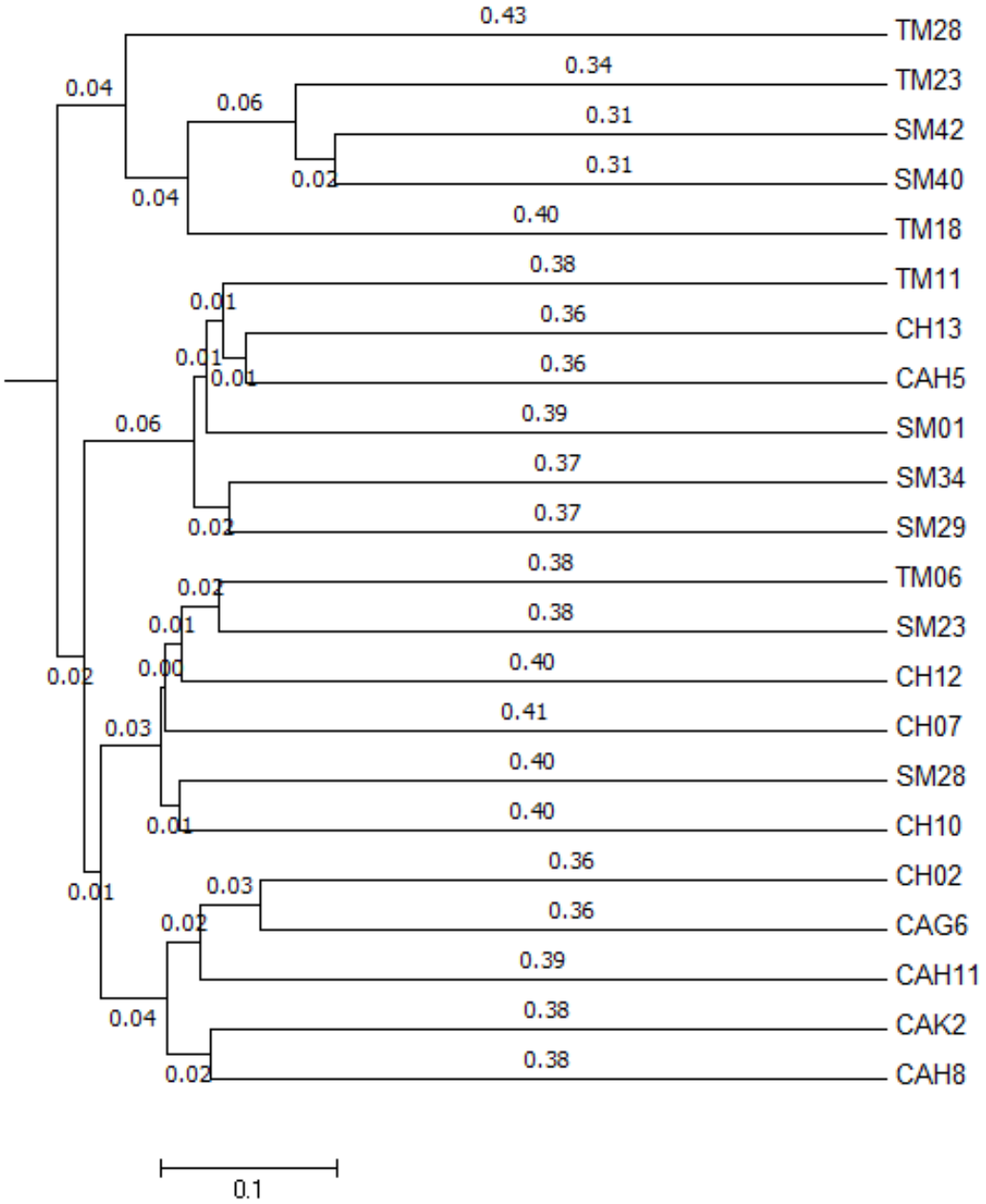

Figure 16: A phylogenetic tree based on shared predicted viral proteins

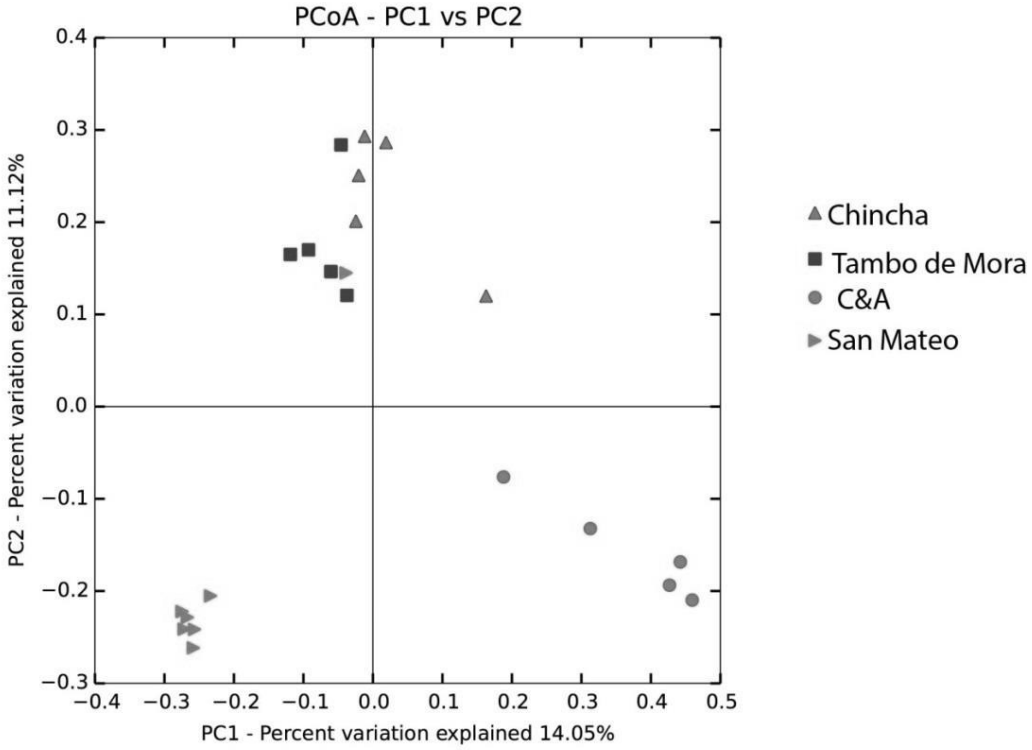

Figure 17: PCoA (PC1&PC2) in which samples are colored by population

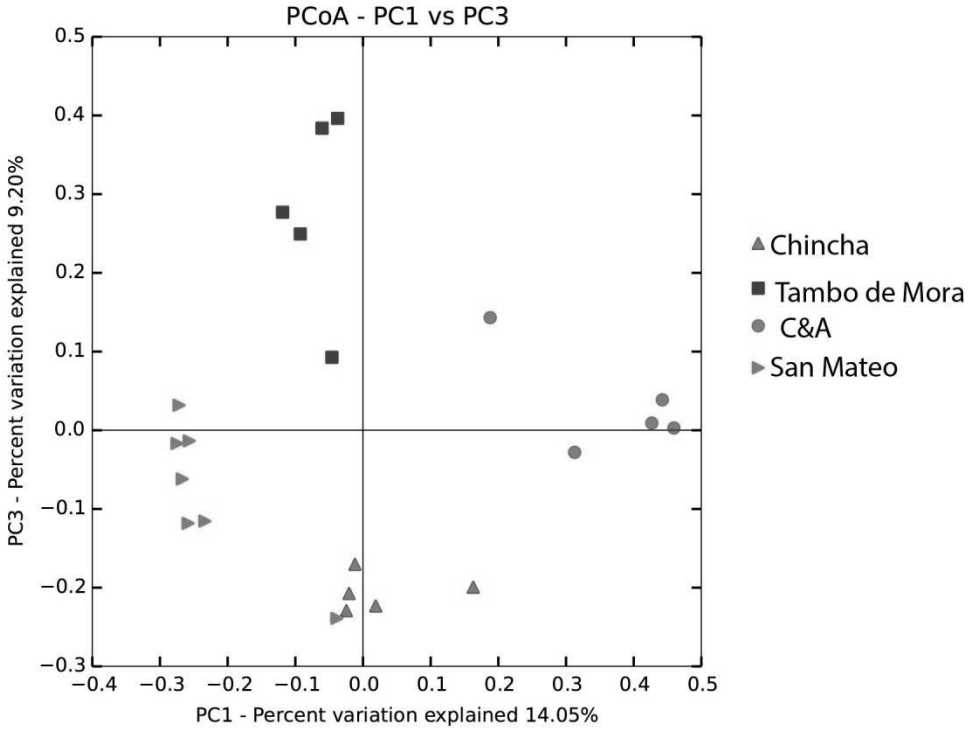

Figure 18: PcoA (PC1&PC3) in which samples are colored by population

Amino Acid Sequences and Geographic Structure

Predicted proteins from dsDNA viruses within the human gut adhere to a geographic structure, but in order to discern a deeper geographic structure between these populations, it is beneficial to examine the specific amino acid sequences within these shared proteins. The branch lengths within the shared protein phylogenetic tree are long, suggesting highly divergent protein types across all four populations. In order to further examine differences in shared proteins within the 22 samples, I selected the most common protein cluster from the UCLUST analysis and used a phylogenetic tree to determine if differences in amino acid sequences also resemble a geographic structure.

The largest ORF cluster, ORF2209 possesses 326 predicted protein sequences containing at least one sequence from 20 of the 22 participants (C&A H11 and TM18 are not represented). Sequences within ORF2209 are largely associated with the Phage_F superfamily (PFAM02305). Proteins belonging to this family make up the capsid, or outer shell, of the DNA bacteriophage. Phages can have upwards of 60 capsid proteins within a single genome. Phage_F makes for an excellent candidate for an amino acid examination because it is common within many bacteriophages (and the family Microviridae) and common within the enriched dataset. Using BLAST, I compared random sequences from this OTU to the NCBI database and found sequence similarity to two single stranded DNA bacteriophage proteins: a structural protein and 1GFF, an atomic structure of degraded procapsid particle of bacteriophage G4. In order to examine amino acid differences within this protein cluster, they were first aligned using MUSCLE (Edgar 2004). MEGA5 (Tamura et al. 2011) was used to build a

maximum likelihood tree with 500 replications and was computed for a small segment of the total protein cluster (39 amino acids) that contained as many samples as possible (20 of 22 represented) (Figure 19). Unfortunately, the phylogenetic tree showed nothing that suggested geographic structure within a subset of the most common viral cluster. It is possible that examining other clusters or regions of OTU2209 may reveal amino acid level differences or even nucleotide level differences that cause sequences to cluster based on host geography. Thus, using the current methodology, it is impossible say whether amino acid level differences within viral proteins can be differentiated through host geography.

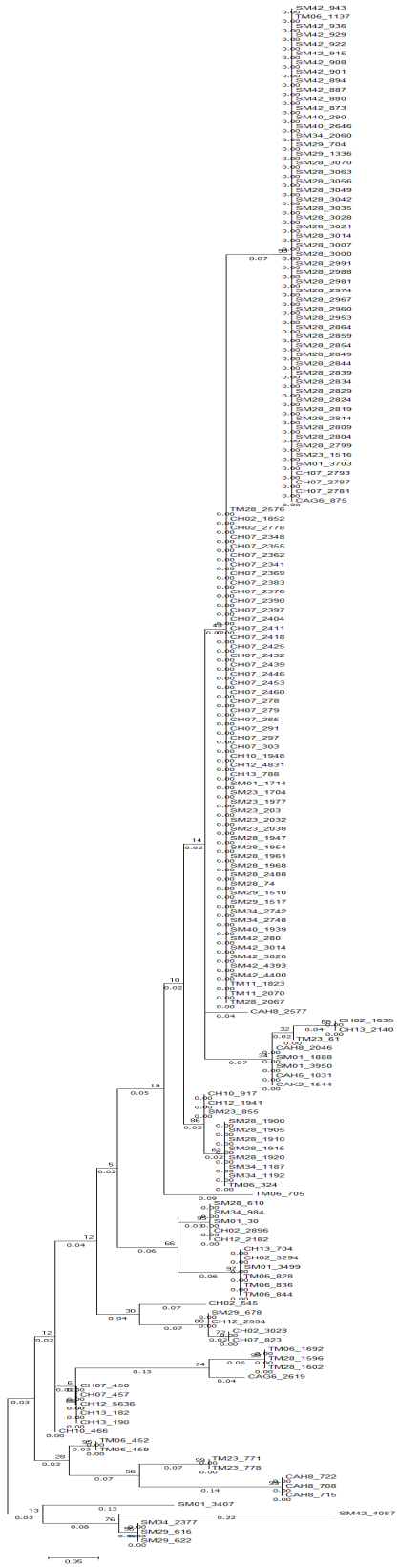

Figure 19: Maximum likelihood tree based on protein sequences from OTU2209

Conclusions

The data presented within this chapter suggest, for the first time, that DNA gut viral protein data are geographically structured. The phylogenetic tree constructed using shared predicted viral proteins allowed for a visual interpretation of the geographic structure belonging to the four different populations within this study: the C&A of Oklahoma, San Mateo of the Peruvian jungle and Tambo de Mora and Chincha of coastal Peru. The principal component analyses from these data even further illustrate the extent of their diversity and expose the strong tendency for predicted proteins of the coastal Peruvians to cluster together, apart from both the San Mateo and C&A communities. In order to determine whether this geographic structure extended to the sequence level, a single protein sequence was selected for further analysis: that of the fairly common Phage_F superfamily. A string of 39 amino acids yielded inconclusive results as to the extent of Phage_F amino acids are geographically structured. Despite showing no strong association at the level of amino acids, shared predicted viral proteins from dsDNA gut viruses are clearly indicative of host geographic region.

Chapter 4

CrAssphage Abundance in the Americas

All viruses are equal before human knowledge. No single discovery is more important than another, but we have to go beyond the comprehension of all the viral diversity and acquire a deeper understanding about the biological properties of viruses because the lack of information or, even worse, the diffusion of erroneous information might hamper future research.
(Canuti and van der Hoek 2014)

Viral diversity and novelty are staggering. Viruses really are the 'dark matter' of the biological universe and a rich source for discovery.
(Rohwer and Youle 2011)

In astrophysics, dark matter is thought to account for much of the matter in the entire universe. Conceived in the early 1930's but never truly postulated until the 1960's (Rubin and Ford Jr. 1970), the search for a type of matter that comprises over 90% of the universe continues to this day. As is the case with many space analogies, it extends to the study of genomics, particularly metagenomics. Forest Rohwer, a virologist from University of California, San Diego, took part in several of the early viral metagenomics studies and can be considered one of the fathers of viral metagenomics. He was the first to use the term 'dark matter' when referring to viral genomic sequences. Currently, many of the sequences from viral enriched datasets do not match anything in databases, so it is difficult to grasp the number of potential viruses within metagenomics samples. He stresses that with the current next generation sequencing technologies available, "anyone can find novel viruses simply by looking for them" (Rohwer and Youle 2011). Essentially, it is just a matter of mining unknown

genetic material for these novel viral sequences. Understanding what is within this dark matter is the first step in making sense of its roles within the human gut.

Recent investigations into viral dark matter have been fruitful. A study has revealed the presence of human gut 'viral enterotypes' which suggests that the gut viral ecosystem types can be organized by the abundance of Bacteroidales-like phages (Ogilvie et al. 2013), similar to what has been suggested for gut bacteria (Arumugam et al. 2011). No follow up studies have looked deeper into these somewhat unknown phages and their impact on host health, so there is still much to explore. Another recent study discovered a novel, large dsDNA virus from previously published gut metagenomic data which they deemed crAssphage.

The discovery of the crAssphage genome is a unique and exciting finding for the field of virome and microbiome research. First, the crAssphage genome is incredibly large in relation to most gut viruses; it possesses a 97,065bp dsDNA circular genome (Figure 20). Although this is not the largest viral genome ever discovered (the Mimivirus has a genome over 1,000,000bp) it is much larger than the average bacteriophage genome within the gut and almost 20 times larger than the first virus ever sequenced, PhiX-174 (Sanger et al. 1977). Secondly, Dutilh et al. found the phage to be particularly ubiquitous within the United States (Dutilh et al. 2014). Since the discovery of this phage occurred so recently, there have been no other studies published that attempt to understand its biological characteristics. I use a number of bioinformatic methods to characterize this recently discovered phage, which is found within the metagenomic samples from both North and South American populations. The Cheyenne and Arapaho, Norman, and San Mateo samples will provide a unique complimentary to

the current data set to Dutilh et al.'s reference genome (compiled from several individuals two sets of twins and their mothers from St. Louis) (Reyes et al. 2010).

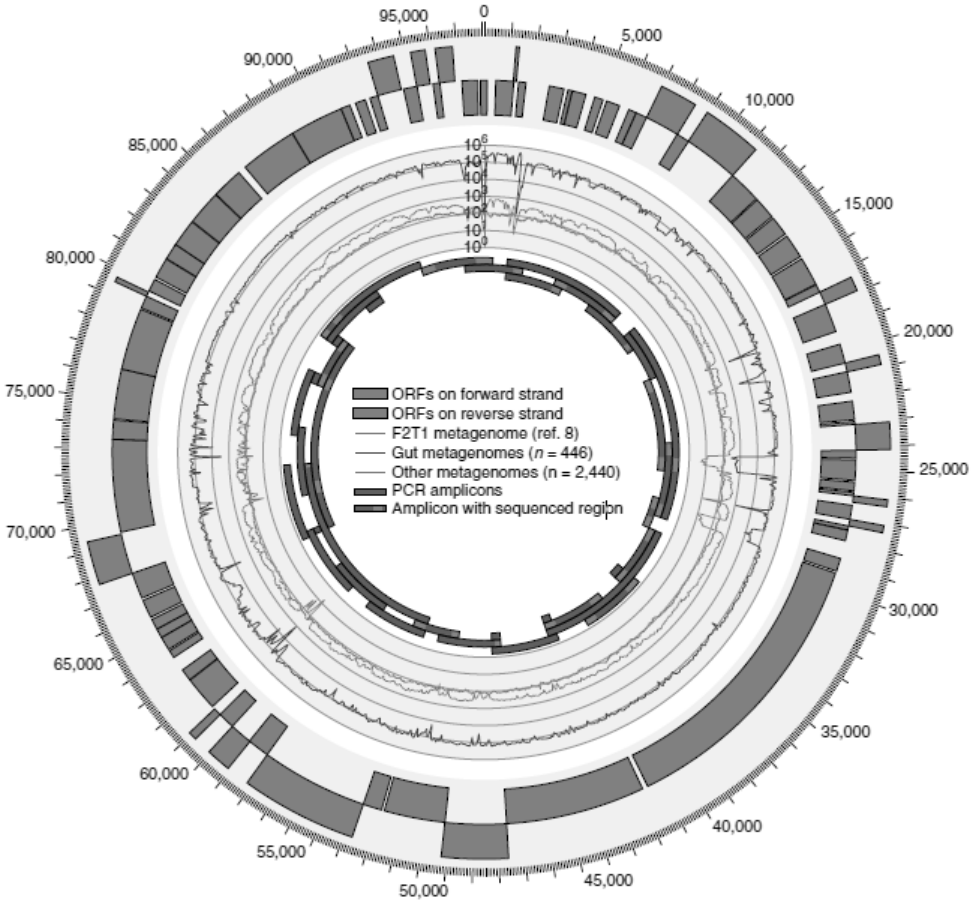

Figure 20: CrAssphage Genome (Dutilh et al. 2014)

In order to evaluate crAssphage, unenriched metagenomic information was utilized from the Norman (N=23), C&A (N=39) and San Mateo (N=24) communities. The laboratory methods for generating these shotgun data are outlined in the previous chapter. The Illumina HiSeq 2x100 paired-end chemistry was used for the sequencing of the unenriched sample set. With this method, the average read length is around 100bp (paired-end refers to the fragment being sequenced from both the forward and reverse directions). The average total number of raw paired-end reads for each population was 23,296,634 for Norman, 25,107,393 for C&A, and 29,931,882 for San Mateo. The raw

paired-end data was matched to the crAssphage reference using Bowtie (1.1.1) (Langmead et al. 2009) . Information on the participants from each population and the overall abundance of crAssphage sequences within each individual can be seen in Table 5 and Table 6.

The contrast between the three populations in regards to crAssphage abundance is startling. The average percentage of sequences within each population were 0.1355% for C&A (range: 0-1.20%), 0.0127% for Norman (range: 0-0.17%), and 0.0100% for San Mateo (range: 0-0.22%). The C&A population, on average, is 10 times as abundant for crAssphage when compared to the Norman population and almost 14 times as abundant compared to San Mateo (only two individuals had abundance above 0.01%). Using the crAssphage abundance data within the C&A and Norman populations I test three specific hypotheses regarding the biological characteristics of the virus (Figure 21).

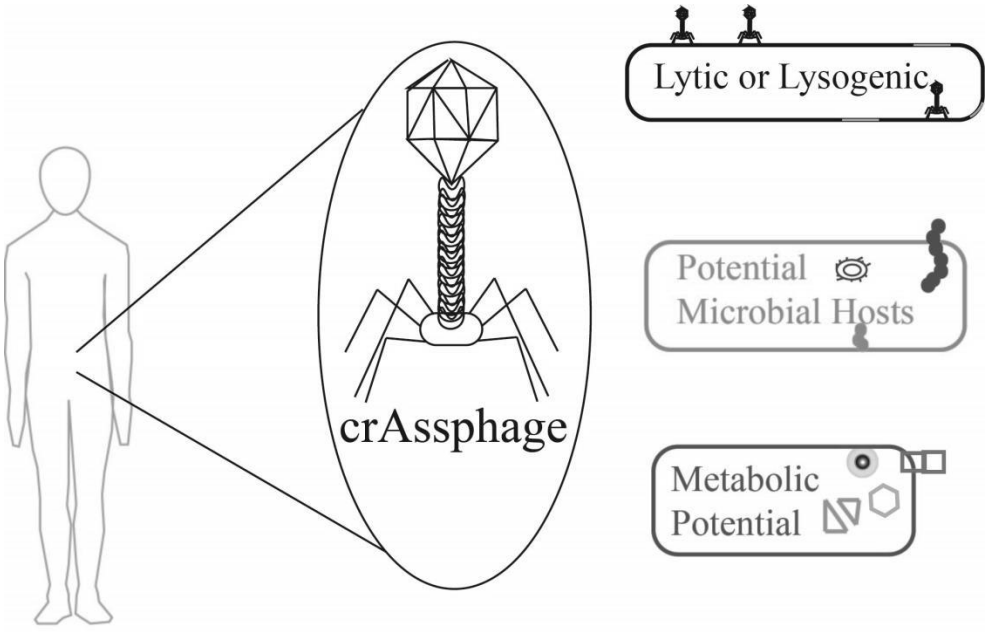

Figure 21: Biological structure of crAssphage

Sample Name	Total Paired Reads	Aligned 1 time	Aligned >1 time	Actual % from Bowtie output	Used for analysis
CAC01	21119162	188	0	0.000%	
CAC02	22977658	16	0	0.000%	
CAC03	27429081	8625	50	0.030%	
CAC04	23297393	17	0	0.000%	
CAC05	22454518	43427	273	0.190%	*
CAC06	22846992	33	0	0.000%	
CAC08	29005936	17	0	0.000%	
CACN02	22351717	98058	933	0.440%	
CACN03	23787715	12938	113	0.060%	*
CACN04	24331507	12624	26	0.050%	
CACN05	23971034	12709	17	0.050%	
CAG01	26787629	9639	61	0.040%	
CAG02	13196473	36	0	0.000%	
CAG03	27301012	56436	232	0.210%	*
CAG04	24025612	56	0	0.000%	
CAG05	22212917	21111	106	0.090%	
CAG06	21327680	236683	1663	1.120%	*
CAG07	23639072	72	0	0.000%	
CAG08	23171029	50	0	0.000%	
CAH01	17747631	55	0	0.000%	
CAH05	20690016	73489	598	0.360%	*
CAH06	23705100	164686	793	0.690%	*
CAH07	27034202	71	0	0.000%	
CAH08	22974190	73604	381	0.330%	*
CAH09	23112810	139	0	0.000%	
CAH10	22884207	40	0	0.000%	
CAH11	20154490	18744	147	0.090%	
CAH13	21080859	54	1	0.000%	
CAH14	20220520	54357	496	0.270%	*
CAK01	19604386	15268	115	0.080%	
CAK02	19640399	25594	144	0.130%	*
CAK03	30142343	12283	43	0.040%	
CAK04	26859150	4954	53	0.020%	
CAK05	30829874	19026	161	0.060%	
CAK06	25437467	6088	25	0.020%	
CAK07	27208566	183244	852	0.670%	*
CAK10	26345239	26540	108	0.100%	*

CAK11	12048843	524	4	0.010%	
NO01	20614384	8	0	0.000%	
NO02	25925960	11296	37	0.040%	
NO03	27096126	1	0	0.000%	
NO04	23787664	7	0	0.000%	
NO05	24149687	11611	85	0.050%	
NO06	19707190	43	0	0.000%	
NO08	21352814	6	0	0.000%	
NO09	18722553	1	0	0.000%	
NO10	23432059	4	0	0.000%	
NO11	24968807	31	0	0.000%	
NO12	24263287	5	0	0.000%	
NO13	22139866	13	0	0.000%	
NO14	23124993	10	0	0.000%	
NO15	20443938	45	0	0.000%	
NO16	28293876	747	0	0.000%	
NO17	21361734	6	0	0.000%	
NO18	24479585	14	0	0.000%	
NO19	23005555	5	0	0.000%	
NO20	22696895	39579	213	0.170%	*
NO21	21926322	7	0	0.000%	
NO22	24462802	3866	18	0.020%	
NO23	26569854	39	0	0.000%	
SM01	35399494	25	0	0.000%	
SM02	45023830	79	0	0.000%	
SM03	22216660	21	0	0.000%	
SM05	23030793	13	0	0.000%	
SM11	32137343	23	0	0.000%	
SM18	31111011	1858	13	0.010%	
SM20	29053562	12	0	0.000%	
SM23	29226654	64	0	0.000%	
SM24	32123359	41	0	0.000%	
SM25	30575527	14	0	0.000%	
SM28	26913134	79	0	0.000%	
SM29	29648053	76	0	0.000%	
SM30	27397584	55	0	0.000%	
SM31	29753792	31	0	0.000%	
SM32	34315106	63	0	0.000%	
SM33	28006919	48	0	0.000%	

SM34	26512933	46	1	0.000%	
SM37	26879818	36	0	0.000%	
SM39	33139273	71	0	0.000%	
SM41	29990550	80	0	0.000%	
SM42	27839736	621	1	0.000%	
SM43	29084962	63813	432	0.220%	
SM44	29053211	90	0	0.000%	

Table 5: crAssphage abundance from unenriched data set. Samples used for V4 and Metabolite analysis are starred in the final column (SM43 was omitted because no metabolite information was available).

Sample Name	**Total Paired Reads**	**Aligned 1 time**	**Aligned >1 time**	**Actual % from Bowtie output**
CAG6vi	7977627	130071	631	2.210%
CAH11vi	7443	41	0	0.230%
CAH5vi	5372982	10341	25	0.680%
CAH8vi	6353432	39657	132	0.770%
CAK2vi	5612698	8825	42	0.170%
CH02vi	4611304	0	0	0.000%
CH07vi	5581896	4	0	0.000%
CH10vi	6361095	1	0	0.000%
CH12vi	8000000	1873	18	0.040%
CH13vi	5712312	2	0	0.000%
SM01vi	5293371	0	0	0.000%
SM23vi	6688826	1	0	0.000%
SM28vi	5909638	4	0	0.000%
SM29vi	5116915	1	0	0.000%
SM34vi	5002451	2	0	0.000%
SM42vi	5601537	2	0	0.000%
TM06vi	6150314	0	0	0.000%
TM11vi	5121103	10	0	0.000%
TM18vi	5945652	0	0	0.000%
TM23vi	4803282	4	0	0.000%
TM28vi	4529424	0	0	0.000%

Table 6: crAssphage abundance from unenriched data set. These samples were not used for crAssphage analysis.

The first hypothesis is related to the replication strategy of crAssphage. Viruses and bacteriophages can adhere to a temperate (lysogenic) or lytic life cycle. Infection of a microbe by a bacteriophage begins with the entrance of the viral genome into the host bacterial cell. At this point, temperate (lysogenic) phages will integrate their own genome into the genome of the host cell. The expression of the virus, which is now within the host genome and referred to as a prophage, can be induced by a number of internal or external factors: UV light, temperature, or exposure to chemicals, just to name a few (Tapper and Hicks 1998). Alternatively, if the bacteriophage adheres to a lytic life cycle, the viral genome will be copied using the host cell's machinery and not integrated into the genome. Eventually the host cell will lyse, causing the bacteriophages to be released from the cell into the extracellular fluid, free to infect other microbiota. It is even possible for some viruses to express both life cycles or alternate between the two. The lytic hypothesis proposes that the overabundance of crAssphage seen within the C&A is attributed to a 'boom' phase of the lytic cycle, in which many phages are released from the host cell following a 'bust' cycle of inactivity. Alternatively, the phage is temperate and the abundance of crAssphage is related to the virus being detected within the bacterial genome in which it is integrated, and thus, it is not a free floating phage within the GI tract.

A second hypothesis is related to host microbiota of crAssphage. The authors that first assembled the crAssphage reference genome suggested microbiota from the genera *Bacteroides* as potential host organisms for crAssphage (Dutilh et al. 2014). Alternatively, the abundance of other microbial genera may also correlate with

crAssphage abundance in the C&A and Norman, suggesting another potential host for the virus.

Lastly, I investigate the potential role of crAssphage abundance in gut metabolic expression. Using metabolite count data, I address the null hypothesis that crAssphage abundance has no association with the presence of specific metabolites and biological compounds shown to impact a healthy GI tract. An alternative hypothesis is that the high abundance of crAssphage is associated with the abundance of host metabolites commonly found within a healthy or diseases human GI tract. If the null hypothesis is rejected, the abundance of crAssphage may have an association with host autoimmunity or obesity.

CrAssphage: Lytic or Lysogenic?

The first step in biologically characterizing crAssphage is testing whether the phage is lytic or lysogenic. In order to do so, I utilized contig (assembled raw reads) information that are similar in genetic sequence to the DNA around the crAssphage origin point (the point between 97,065bp and 0bp in Figure 20). If the crAssphage genome is integrated into the genome of a host microbe at this origin point, the contig information would partly match crAssphage and partly match the host microbe. I selected 500bp on either side of the crAssphage origin point as references and removed all of the contigs that partially matched those sequences using BLAST. BLAST, a Basic Local Alignment Search Tool within NCBI, identifies similar nucleotide sequences within a sample set and in this case, I compared complete contig data to the 500bp segments on either side of the crAssphage origin point (using an e-value of 10^{-5}). The

sequences were then aligned to the full crAssphage reference sequence using Sequencher (4.0). Those sequences that continued past the 500bp reference segments were further searched in the NCBI microbial genome database using BLAST. The BLAST results reveal that three uncultured gut microbiota have a greater than 80% identity: Uncultured organism clone 1041059766405, Uncultured organism clone 104105976836, and Uncultured organism clone VC1C642TF. These three organisms were not host microbes, but were in fact, part of the crAssphage genome which was evidenced by the fact that they could be reassembled into a circular crAssphage genome. These results do not indicate a lysogenic behavior in regards to crAssphage but indicate NCBI database error: all three genomes came from a 2009 unpublished gut metagenome study at the Pathogens Institute in Florida. These genomes, including many others from the unpublished 2009 study, were not microbial genomes but segments of the nearly 100kbp crAssphage viral genome. They have yet to be deleted from the NCBI microbial genome database. Thus, these data suggest that the crAssphage virus is a lytic bacteriophage.

The contig data suggest that crAssphage is a lytic phage, so the question becomes: do the raw data tell the same story? Since contigs are merely assembled raw sequence reads, they will likely provide the same information. However, if the crAssphage genome inserts itself into the host in pieces or somewhere outside of the origin point, looking closer at these raw reads could help to address the question of lytic or lysogenic replication. To do so, raw sequence data was first mapped to the crAssphage reference with an e-value of 10^{-5}. Those raw sequences that partially matched crAssphage at 70bp or less (and showed at least an 80% identity to the NCBI

bacterial genomes database at an e-value of 10^{-5}) were further examined. Data from only four participants matched these search criteria (Table 7). If crAssphage was a temperate virus, these sequences would contain genetic information from both crAssphage and a potential host, but only a small subset adhered to these criteria. Thus, this suggests that almost all of the raw sequence data from these samples that match to the crAssphage reference only match to the crAssphage reference and nothing else. The results from both contig and raw sequence data suggest that crAssphage is a lytic virus.

Sample	Raw Reads	NCBI Microbe BLAST hits
NO02	1	38
NO20	3	2
NO22	1	1
SM43	13	9

Table 7: Contig reads that partially matched crAssphage and other microbiota

Potential Host Microbiota for crAssphage

Dutilh et al. (2014) proposed that the host organism for crAssphage was *Bacteroides*. In order to examine what other microbes are potential crAssphage hosts I apply two analytical methods using the Peruvian and Oklahoma samples. The first technique uses the spaces between repeating elements in microbial genomes referred to as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) in order to identify potential host genomes. The second compares crAssphage abundance to gut microbial genera abundance, which is obtained through an analysis of the hypervariable region of the 16S rRNA gene, a ribosomal gene ubiquitous to all microbes.

CRISPR Spacer analysis

CRISPRs are part of the bacterial genomic locus and mediate the adaptive immune system. When initially discovered in *E. coli iap* in 1987, the reason for these CRISPR sequences was unclear (Ishino et al. 1987). When examined further, certain genes were found to be associated with these repeats (*cas* (CRISPR-associated) genes) (Haft et al. 2005). Between these CRISPRs (usually 24bp to 47bp) are spacers acquired from foreign nucleic acids, which allow the bacteria to continuously hone its adaptive immune system (Goren et al. 2012). Spacers are remnants of past viral infections which act to prevent future ones but phages continue to adapt in order to invade this microbial immune system (Swarts et al. 2012). Due to the evolutionary arms race, a rapid turnover of CRISPRs and spacers is common.

CRISPRs are fairly ubiquitous and thought to be present in 45% of bacteria and over 90% of archaea (Horvath and Barrangou 2010; Marraffini and Sontheimer 2010). Due to competition between microbes and viruses, new mutations in CRISPR-targeted regions are constantly arising, ones which microbial hosts have to combat (Makarova et al. 2011). CRISPR spacers can be used to better understand the relationship between microbiota and the bacteriophages that infect them.

Dutilh et al. (2014) utilized a program known as PILER-CR (Edgar 2007), which identifies CRISPR and spacer sequences in large datasets, in order to determine potential microbial hosts for crAssphage. They estimated that microbiota from the *Bacteroides* and *Prevotella* genera were potential hosts for crAssphage. I apply the PILER-CR program on the unenriched shotgun data from Norman, C&A, and Peru. After removing the CRISPR elements, I checked the spacer sequences using BLASTn

(e-value of 10^{-5}). If any of the spacer sequences matched to the crAssphage reference, it would be possible to go back to the original contig sequence and determine the host microbiota. Unfortunately, none of the spacer sequences matched crAssphage These results can be seen in Table 8. Thus, this method was not able to identify any crAssphage sequences within the spacer data and provided no information as to potential hosts for the phage itself.

Sample	Reference	Virus	Family	Genera	Virus present in enriched?	SEQ
CAC2	NC_020232.1	Megavirus lba	Mimiviridae	Unclassified	Yes	CTGTATTTCGCTCAACGCTATTAATGGTAAGCGTTG
CAC3	NC_020082.1	Edwardsiella phage MSW-3	Myoviridae	Unclassified	No	CAGTTCCAACCTGAGCGGTGCCGACCTGAGCGGGT
CAC5	NC_006151.1	Suid herpesvirus 1	Herpesviridae	Varicellovirus	No	GTCCAACGGAGAGTTTGACGCCTACGAGCGCGT
CACN5	NC_022644.1	East African cassava mosaic Malawi virus	Geminiviridae	Begomovirus	No	AATTGCATTGAAATTATTAAATTTGAGGGT
CACN5	NC_023423.1	Pithovirus sibericum	Unclassified ds	Unclassified	No	TGGCAAAAAGTAAAAAAGAAGTGATTGAAATAACAG
CAG2	NC_001575.2	Chlamydia phage 1	Microviridae	Chlamydiamicrovir	Yes	GATTCGCAGACTGTGGCAAAAGCCATAAGCGGCTTAAT
CAG3	NC_024382.1	Alcelaphine herpesvirus 2	Herpesviridae	Macavirus	No	AGCTTTGGAATTACTTTTTCCTCTTTTAGGGTTTC
CAG4	NC_020104.1	Megavirus lba	Mimiviridae	Unclassified	Yes	CACCGATATCAACAATATTTGATAAGTCCTCCGCAA
CAG5	NC_025443.1	Salmonella phage 9NA	Siphoviridae	Unclassified	No	CTTGTCAGTTGATCCGAAACCACCGTTACG
CAG6	NC_019909.1	Yersinia phage phiR1-RT	Myoviridae	Unclassified	No	GCCCGTTCATGCGCTTCATTAGCGCGCGTTCG
CAG6	NC_019909.1	Yersinia phage phiR1-RT	Myoviridae	Unclassified	No	CCGTTCATGCGCTTCATTAGCGCGCGTTCG
CAG8	NC_004303.1	Streptococcus phage O1205	Siphoviridae	Unclassified	No	TATTGTAAAAGTAGTCTCTTGTAATGTTTG
CAH11	NC_023607.1	Mycobacterium phage 40AC	Siphoviridae	Unclassified	No	TGACTAAGTAATACTACATCACCAACTTGACAGGCT
CAH5	NC_021794.1	Cellulophaga phage phi18:3	Podoviridae	Unclassified	No	CTATTGAACATTTGAGTAATAGCCAAATTGATGATA
CAH5	NC_023688.1	Aeromonas phage PX29	Myoviridae	T4likevirus	No	TCGTTACAGCTGCAAAAGAATCTTCTGTATACTCTTCAGCTTTCAAAGC
CAH6	NC_001664.2	Euproctis psedoconspersa nucleopolyhedr	Baculoviridae	Alphabaculovirus	No	GCTAATAGGTATAATGTTTCTTATAAC
CAH8	NC_008296.2	Synechoccus phage syn9	Myoviridae	T4likevirus	No	CCGGAAGAAATAGAAAATATGAAGAAAAAGGCA
CAH8	NC_023610.1	Erwinia phage PhiEaH1	Siphoviridae	Unclassified	No	GTGGTTCTTCTCCAGACGCTTCGTCACATTCT
CAH9	NC_020082.1	Edwardsiella phage MSW-3	Myoviridae	Unclassified	No	CAGTTCCAACCTGAGCGGTGCCGACCTGAGCGGGT
CAK11	NC_008720.1	Escherichia phage N4	Podoviridae	N4likevirus	No	TACCCAATTGCAGC
CAK11	NC_010191.1	Ostreococcus virus OsV5	Phycodnaviridae	Unclassified	No	CGCAAAATTGGTGA
CAK2	NC_002520.1	Amsacta moorei entomopoxvirus 'L'	Poxviridae	Betaentomopoxvir	No	AACTAATTCAGATCCTGAAAAATGCATCCTGCTTGACAACCCTAT
CAK4	NC_022644.1	East African cassava mosaic Malawi virus	Geminiviridae	Begomovirus	No	AATTGCATTGAAATTATTAAATTTGAGGGT
CAK5	NC_022644.1	East African cassava mosaic Malawi virus	Geminiviridae	Begomovirus	No	AATTGCATTGAAATTATTAAATTTGAGGGT
NO03	NC_021312.1	Phaeocystis globosa virus	Phycodnaviridae	Unclassified	No	TATGTAAAATAATAAGAAGAATTCACTTCTGGATTTT
NO05	NC_012639.1	Euproctis psedoconspersa nucleopolyhedr	Baculoviridae	Alphabaculovirus	No	AAATCAAAAAAAAGAAAACTACCTCAACTCCTAATAA
NO05	NC_012663.1	Lactococcus phage P087	Siphoviridae	Unclassified	No	CTTTTATTATGACAGGAAAAGTAAATTTTAATAAAG
NO05	NC_012663.1	Lactococcus phage P087	Siphoviridae	Unclassified	No	CTTTTATTATGACAGGAAAAGTAAATTTTAATAAAG
NO08	NC_015326.1	Lausannevirus	Marseillevirida	Unassigned	No	TTAAAAATCAAGAAGAAAATATTTTAAATAAAATAT
NO08	NC_020082.1	Edwardsiella phage MSW-3	Myoviridae	Unclassified	No	CAGTTCCAACCTGAGCGGTGCCGACCTGAGCGGGT
NO09	NC_021312.1	Phaeocystis globosa virus	Phycodnaviridae	Unclassified	No	TATGTAAAATAATAAGAAGAATTCACTTCTGGATTTT
NO10	NC_023503.1	Streptococcus phage 20617	Unclassified ph	Unclassified	Yes	CCAAAGAATGGACCATCTTAATGAGAATAT
NO11	NC_020082.1	Edwardsiella phage MSW-3	Myoviridae	Unclassified	No	CAGTTCCAACCTGAGCGGTGCCGACCTGAGCGGGT
NO16	NC_018277.1	Staphylococcus phage SpaA1	Siphoviridae	Unclassified	No	GTGGTTAGCATCACAGGCTGATATGTTAGCAGAAGAT
NO20	NC_020082.1	Edwardsiella phage MSW-3	Myoviridae	Unclassified	No	CAGTTCCAACCTGAGCGGTGCCGACCTGAGCGGGT
NO22	NC_001741.1	Chlamydia phage 1	Microviridae	Chlamydiamicrovir	Yes	TCATGATTATTTTACTTCTGCTCTTCCGTG
NO23	NC_019495.1	Cyprinid herpesvirus 2	Alloherpesviric	Cyprinivirus	Yes	TGAATATTGAATCTTCCAGTGAAGAAGAAGAAGAAGAA
SM20	NC_016166.1	Gordonia phage GTE7	Siphoviridae	Unclassified	No	TTTGCAGGCGCAGGTTGACAAAGATGTGCC
SM25	NC_025462.1	Acinetobacter phage vB_AbaM_Acibel004	Myoviridae	Unclassified	No	TAGCTCAGTTGGTAGAGCAGTAGGCTTTTAATCTATC
SM28	NC_020083.1	Serratia phage phiMAM1	Myoviridae	Unclassified	No	GAGTACCACCGTGGGCTGGTGATGTTTTGAAAGATGT
SM34	NC_021530.1	Synechoccus phage S-CAM8	Myoviridae	Unclassified	Yes	GCGCTGCTACTTATTCTTTTGAAAATATTA
SM39	NC_016072.1	Megavirus chiliensis	Mimiviridae	Unclassified	No	TAGTAAACATATTGAAATAATTGAGTATGA
SM41	NC_002642.1	Tanapox virus	Poxviridae	Yatapoxvirus	No	ATAAAGTAAAAAACATAAATATAAAAAAA
SM41	NC_021804.1	Cellulophaga phage phi39:1	Podoviridae	Unclassified	No	CTAAAATATTTATCAGAAATGAAATATAAA
SM42	NC_022339.1	Propionibacterium phage PHL037M02	Siphoviridae	Unclassified	No	ATATTTTAAACACTTAATAATTTAAACCATAGT
SM43	NC_005237.1	Seoul virus	Bunyaviridae	Hantavirus	No	GTTGAACTAAAAAAAATACTCTCTTTTTCCGG
SM43	NC_005237.1	Seoul virus	Bunyaviridae	Hantavirus	No	GTTGAACTAAAAAAAATACTCTCTTTTTCCGG
SM44	NC_020104.1	Moumouvirus	Mimiviridae	Unclassified	No	GATGATGTATTTATTTTTCTTGACCCTGAGAGGGA

Table 8: CRISPR spacer hits

16S rRNA V4 genera abundance information

The second method by which I examined potential crAssphage microbial hosts was through abundance correlations with 16S rRNA V4 data. The hypervariable region 4 of the 16S rRNA gene is a short, taxonomically useful segment that is found within all bacteria. Normalized genera counts were obtained for the Norman and C&A samples. The 16S rRNA V4 (~290bp) was generated through traditional PCR (Accuprime) using Illumina adapter barcoded primers. Triplicate PCRs were pooled and sequenced using an Illumina MiSeq (2x250 paired-end reads). The reads were filtered for quality (30<QC) and reads containing ambiguous bases ('N') were removed. Read pairs were merged to generate the full 16S V4 region and samples were identified by their unique barcode (de-multiplexed using QIIME). The sequences were assigned to operational taxonomic units (OTUs) within QIIME and rarefied to 10,000 reads per sample, the results of which were summarized at the genus taxonomic level (L6) in QIIME. Final genus counts were obtained by repeating rarefaction 100 times and taking the median value over those repetitions. These counts were compared to overall crAssphage abundance using a Spearman's rank correlation coefficient within the R statistical computing package. A correlation using samples with a 0.1% crAssphage or higher abundance (N=12) can be seen in Table 9 (p value<0.05 with no repeated number correction). Due to the large number of individuals who showed <0.01% crAssphage abundance (particularly in Norman), I chose to examine only the correlation results from individuals with >0.1% or higher abundance of crAssphage.

All Norman & C&A Samples (crAssphage abundance >0.1%) (N=12)		
Genus	**P Value**	**Coefficient**
p__Firmicutes, g__Streptococcus	0.0188	0.6783
p__Firmicutes, f__Gemellaceae, g__	0.0228	0.6478
p__Firmicutes, g__Dehalobacterium	0.0372	0.6048
p__Proteobacteria, g__Desulfovibrio	0.0378	0.6033

Table 9: CrAssphage abundance and microbial abundance

Four genera in particular had a strong association ($P<0.05$) and positive correlation with crAssphage abundance (>= 0.1%): *Streptococcus* (p=0.0188), an unknown genera from the Gemelleaceae family (p=0.0228), *Dehalobacterium* (p=0.0372) and *Desulfovibrio* (p=0.0378). According to 16S V4 genera counts, *Streptococcus* had an overall abundance of 1.98% in the population while the other three genera of statistical significance had an overall abundance of below 0.1%. Although lytic phages will eventually cause the host cell to rupture, the crAssphage abundance in the NGS data may be the result of the phages still existing within the host cell at the time of DNA extraction. Thus, all four of these genera are candidates for microbial hosts of crAssphage. Moreover, three of the four genera are from the Firmicutes phyla, which has long been associated with the obese phenotype (Turnbaugh et al. 2006).

The first Firmicutes genus, *Streptococcus*, is a fairly common microbe within humans and is the microbe most strongly correlated with overall crAssphage abundance. It is common within both the oral cavity (Lazarevic et al. 2009; Nasidze et al. 2009) and the GI tract (Human Microbiome Project 2012). The *Streptococcus* genus are not overly abundant in the Norman and C&A samples (1.9% of total V4 reads)

suggesting a unique association with crAssphage abundance. Increases in gut *Streptococcus* have been associated with non-alcoholic fatty liver disease (Jiang et al. 2015), decreased fiber intake and increased risk for colorectal cancer (Chen et al. 2013) and Celiac's disease (de Meij et al. 2013). Although there is no phenotypic information from these two populations regarding fatty liver disease, cancer, or Celiac's, these results could mean several things. Since *Streptococcus* is the most significant genus associated with crAssphage abundance, it is a candidate host for crAssphage. Alternatively, while not mutually exclusive, crAssphage abundance, like what has been demonstrated in regards to *Streptococcus* abundance, is an indicator for overall GI health. The overwhelming presence of crAssphage within the C&A coupled with their obese phenotype warrants further exploration of this bacteriophage in clinical studies of gut microbiota.

In addition to the statistical significance of *Streptococcus*, each of the three other potential microbial hosts for crAssphage is associated with aspects of gastrointestinal health but in very low overall abundance in our Oklahoma samples (<0.1%). *Desulfovibrio* is a genus of gram-negative microbiota found within the gut, which are known to be capable of reducing sulfate to sulphide, a compound that is toxic to epithelial cells (Pitcher et al. 1998). Although some studies have shown no link between *Desulfovibrio* abundance and obesity (Karlsson et al. 2012), others suggest its role as a sulfate reducer gives it a role in the progression of ulcerative colitis (Rowan et al. 2009). Also, an increase in Gemellaceae has been associated with Crohn's disease in several studies (Gevers et al. 2014; Haberman et al. 2014) and Celiac's disease in saliva (Francavilla et al. 2014), and *Dehalobacterium* abundance is thought to be impacted by

overall glucose intake of the host (O'Connor et al. 2014). In other words, each of these microbial genera have been implicated as indicators for host obesity, an issue affecting most C&A participants. Although the role of crAssphage within the GI tract is still unknown, the roles that these four microbial genera have in the progression of host GI disease states further support the hypothesis that crAssphage plays a role in the stability of the host immune system.

Certain genera are surprisingly absent from the 16S rRNA V4 correlations, those that Dutilh and colleagues suggested may be potential hosts for crAssphage (Dutilh et al. 2014). Microbiota from both the *Bacteroides* and *Prevotella* genera (which Dutilh et al. (2014) said was a second potential candidate) were not significantly correlated with crAssphage abundance in the Oklahoma dataset. The results presented here suggest that crAssphage is a virus capable of infecting several different microbiota, even from different phyla. More importantly, it seems that crAssphage abundance is significantly correlated with four bacteria that have been implicated in gastrointestinal health. Though a Spearman's rank correlation was not significant for crAssphage abundance and host weight, the overall obesity rate within the C&A population was very high (37 of the 39 participants had a BMI that classified them as overweight or obese), not to mention average crAssphage abundance was 10 times higher in C&A versus Norman. Considering that crAssphage has yet to be cultured in a lab setting, there are still many more questions that need to be answered regarding its structure and function. The results presented here indicate that crAssphage has a currently unknown role in host health and gut microbiome dysbiosis.

CrAssphage Abundance and Host Metabolic Output

The final hypothesis related to the biological characterization of crAssphage involves the use of gut metabolite information. I utilize a Spearman's rank correlation coefficient test in order to determine the impact that crAssphage abundance on the host gut ecosystem. A company known as Metabolon Inc. offers a service that allows for the rapid and accurate quantification of a number of biomarkers from biological samples. Metabolites are the intermediates or products of the host metabolism and provide a unique glimpse into the inner workings of the human gastrointestinal tract. These data are expressed in metabolite counts which included 500 different metabolites from the human GI tract. Metabolite data was acquired for both Norman and C&A fecal samples, and act as an accurate indicator of metabolic pathways of the gastrointestinal tract. Individuals that showed an overall crAssphage abundance of 0.1% or higher were subject to a Spearman's rank correlation test, which can be seen in Table 10.

All Norman & C&A Samples (crAssphage abundance >0.1%) (N=12)		
Metabolite	**P Value**	**Coefficient**
gallate	0.0018	0.7988
gamma-CEHC	0.0153	0.6784
tryptamine	0.0208	-0.6550
phytanate	0.0256	0.6381
tyramine	0.0301	0.6364
myristate (14:0)	0.0324	0.6294

Table 10: CrAssphage abundance and metabolites

A total of 6 metabolites were significantly associated with crAssphage abundance in the Oklahoma samples: gallate (p=0.0018), gamma-CEHC (p=0.0153),

tryptamine (p=0.0208), phytanate (p=0.0256), and myristate (14:0) (p=0.0324). Gallate has the strongest association with crAssphage abundance. Almost nothing is known about gallate's role in microbial metabolism but in certain forms it is known to have antiproliferative, anti-inflammatory, and antioxidant effects (Högger 2013). The positive correlation with crAssphage abundance is surprising and somewhat unexpected considering crAssphage abundance was shown in the previous section to be associated with several microbial indicators of poor gut health. One of the other metabolites, gamma-CEHC, is a byproduct of Vitamin E metabolism and usually excreted through urine. The role crAssphage abundance may play in the presence of gamma-CEHC in the gut is unknown. Another significant metabolite is that of tryptamine, a neurotransmitter that occurs in the gut due to the release and subsequent breakdown of tryptophan by *Ruminococcus gnavus* and *Clostridium sporogenes* (Williams et al. 2014). Tryptophan induces the release of serotonin, which impacts host physiology and behavior. This is not the first time gut microbiota have been associated with host behavior (Cryan and O'Mahony 2011), and it is difficult to stay why tryptamine has a negative correlation with crAssphage abundance. Another metabolite, myristate, is significantly associated with crAssphage abundance but has an unknown role in the gut. Currently, forms of myristate are used in laboratory studies to induce cell growth. Additionally, phytanate, a branched-chain saturated fatty acid and an anion of phytanic acid, has no known association with microbiota of the gut.

The most intriguing significant metabolite with crAssphage is tyramine. Tyramine is a trace amine derived from tyrosine and is commonly produced through fermentation of foods. It can be found in many plants and animals eaten by humans,

especially chocolate, alcohol, cheeses, fruits, nuts and marinated meats. The level of tyramine in the gut could be the result of a number of factors including (primarily) host diet, inflammation and microbial groups present (Barry et al. 2010; Murray et al. 1986; Visentin et al. 2004). Tyramine is usually quickly broken down by the body, but individuals taking blood pressure medication may inhibit those enzymes (Bianchetti et al. 1982) and when left to accumulate it can cause panic attacks (Caston et al. 2002). Considering many of the C&A are classified as overweight or obese and likely suffer from type 2 diabetes (of which they are not aware) this metabolite could further indicate that crAssphage is the result of poor gut health.

Unfortunately, most of the metabolites that proved to be significantly correlated with crAssphage abundance did not provide any further clues as to its role within the human gut. There is a slight possibility that the significant p-values observed with these metabolites occurred by chance. No repeated number correction procedures were utilized in calculation of these p-values (542 metabolites were examined). If Bonferroni correction is used, none of the metabolites would reach statistical significance. Thus, the metabolites mentioned above have a risk of type 1 error.

In addition to metabolite information, I attempted to correlate crAssphage abundance with other phenotypic information provided by Oklahoma participants. Total crAssphage abundance did not correlate with age (p value=0.099) or body mass index (p value=0.725). It should be mentioned that no resting blood glucose tests were performed and type 2 diabetes was self-reported by C&A participants, many of whom may have type 2 diabetes but are unaware of it. As a result, type 2 diabetes is likely an unreliable indicator for autoimmune health of the C&A participants. Accurate blood

glucose information would be incredibly beneficial in an attempt to determine the role of crAssphage in gut health.

Final Thoughts on crAssphage

This chapter addressed three hypotheses regarding the biological characteristics of crAssphage. First, it utilized several methods to show that crAssphage abundance within the C&A and Norman communities are due to the presence of a lytic phage, not a temperate one. Secondly, 16S rRNA V4 data demonstrated the presence of several other potential host microbiota, all of which have been associated in the past with host gut health. Lastly, aside from the presence of tyramine, the metabolite data failed to show a strong link between host crAssphage abundance and poor gut health metabolites. Despite the lack of metabolic associations, the presence of obesity in the C&A along with significant associations between crAssphage and four microbiota identified as players in poor gut health, implicate crAssphage in the proper host immune system function, warranting further investigation. The logical next step in crAssphage analysis would be the further mining of already published databases. If crAssphage abundance is correlated with GI health, gut metagenomic data from clinical studies examining GI disorders such as Crohn's or Celiac's would be ideal datasets for reanalysis.

Chapter 5

Conclusions

Viruses are incredibly abundant everywhere on earth and are considered major players in the global ecosystem (Suttle 2007). Viruses are also considered major players in human health and biology. Throughout this dissertation I have stressed an idea that is on the surface, obvious, but has yet to be properly articulated within the field of anthropology: viruses are crucial to our understanding of humans. Over the course of this dissertation I have addressed four core concepts that intertwine viruses and the field of anthropology: human identity, novelty, host history, and health, with a fifth topic of future research directions being discussed along the way (Figure A2).

When we consider ourselves to be biologically human, we have to take into account those many viral and microbial foundations upon which our species lies. In the first chapter I investigated the classic visual metaphor that nicely organizes all living things on the planet: the tree of life. I illustrated that there is more to each organism on the tree than its sole genomic information, but an aggregate of microbial, viral and eukaryotic neighbors. Branches on the tree could even be described as treehouses; the floor, walls and roof of which are comprised of the coinhabitants in the human superorganism. This is not to say that the human, or any multicellular organism for that matter, treehouse is always properly assembled; harmful inhabitants can throw the host into dysbiosis. This is one of the reasons why understanding the many branches of organisms that make up a single species' treehouse is so important, and all of this can be accomplished through metagenomic investigations. Once we understand the role of

each of the separate branches that make up the house, we can take steps to maintain overall order and symbiosis within the host.

The second chapter addressed the incredible amount of taxonomic diversity within viruses and drew attention to their importance within the field of anthropology. Discussing these different viral phyla and families in a context that incorporates anthropology, biology, metagenomics, and genetics, allows for the introduction of unique research avenues, many of which involve novel viruses. Pathogenic viruses have been utilized within the field of medical anthropology through the documentation of cultural, political and economic responses, but the field can do so much more than that. Primatologists can use NGS techniques to examine more zoonotic viruses along with M.D./Ph.D's who can use NGS to quickly identify causes of infection in remote areas (Svraka et al. 2010). Paleogeneticists can detect viral DNA and RNA to study ancient pathogens and human diet (Warinner et al. 2014; Zhang et al. 2006). Forensic scientists may even be able to identify viruses specific to a single person, something that has been similarly shown in microbiome studies (Fierer et al. 2010). This chapter was intended to serve as both a comprehensive examination of the anthropological relevance of viruses and a discussion of implications for the future.

Chapter three addressed the ways in which gut viral genomes can provide information related to host geography. Viruses like JC and BK have been used in the past to identify human migratory patterns, but can a metagenomic assemblage of gut viruses provide enough resolution to make claims regarding the host? In this chapter I use DNA gut viral metagenomic information from Oklahoma and Peru to show that predicted gut viral proteins are in fact, geographically structured. Due to the rapid rate

at which gut viruses evolve, protein sequence similarity within populations was not significant but overall more viral types are shared between populations of similar geographic regions. These data only provide a small peek at the overall gut virome diversity and more populations are needed to understand the ways in which viruses evolve across the globe within humans.

The gut virome plays a key role in overall host health and dysbiosis (Glendinning and Free 2014). Chapter four selected a single, recently discovered virus known as crAssphage and biologically characterized it within populations from Oklahoma and Peru. The data presented in this chapter show that crAssphage is a lytic phage and abundance within the human gut is related to viral entities instead of the crAssphage genome being integrated into a microbial host. I compared crAssphage abundance within Oklahoma populations to microbial taxonomic counts in order to determine any potential host microbiota. One genus of *Streptococcus* from the phylum Firmicutes, was significantly associated with crAssphage abundance. Three other genera, two from Firmicutes (an unknown genus from Gemellaceae and *Dehalobacterium*) and one from Proteobacteria (*Desulfovibrio*) were significantly associated with crAssphage abundance but were in very low overall abundance in the GI tract of the Oklahoma populations. Microbiota from these significantly associated genera have been strongly implicated in previous studies of GI health, suggesting that crAssphage abundance may also act as an indicator for metabolic syndromes within the Cheyenne and Arapaho of Oklahoma. Additionally, crAssphage was significantly associated with several gut biomarkers including tyramine, which has been linked with host diet and inflammation. Unfortunately the other significant biomarkers revealed no

other information regarding its function in the human gut, thus necessitating future studies regarding the importance of crAssphage to the human gut environment.

Virome researchers, such as those within the microbiome field, must be careful not to attribute causation to such associations without further investigation (Hanage 2014). Currently, it is difficult to say whether the overabundance of a particular viral type, like crAssphage, is the cause or merely an unintended consequence of a complex human ecosystem like the GI tract. One way to address these concerns is through continued culturing work. Currently, crAssphage has yet to be cultured in a laboratory, thus many questions still exist surrounding its role in relation to microbiota and the human immune system. The field of metagenomics will always be closely involved with microbial and phage culturing. Hand in hand with culturing is the establishment of larger, more detailed databases with viral sequences. Only a small percentage of viruses are actually within reference databases like NCBI and as a result, upwards of 50% of viral metagenomics sequences are completely unknown (Breitbart et al. 2008). Another way to speed up the process of expanding databases is through the sampling of novel human and animal populations. These understudied populations likely have a multitude of novel families and genera. While still within its early stages, viral metagenomic studies can clearly provide a look at pathogens and hitchhikers in a detailed and informative manner.

To say that viruses are important to human identity and human health is a vast understatement. Viruses are both ingrained within our genome and within our cells; impacting our health and evolution in ways that we still do not understand. The questions, just like the number of viruses within all of us, are many, and will keep

researchers busy for years to come. The adventure can best be described with a quote from the Nobel Prize winning Peruvian author Mario Vargas Llosa: “Science is still only a candle glimmering in a great pitch-dark cavern” (1981). In other words, as we trek further into a vast cave, holding a candle that now burns brighter than ever before, we illuminate a world that is, not so surprisingly, full of viruses.

References

2001. Initial sequencing and analysis of the human genome. Nature 409(6822):860-921.

2011. Microbiology by numbers. Nat Rev Micro 9(9):628-628.

Aaby P, Bukh J, Hoff G, Leerhøy J, Lisse IM, Mordhorst CH, and Pedersen IR. 1985. High measles mortality in infancy related to intensity of exposure. The Journal of pediatrics 109(1):40-44.

Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, and Versalovic J. 2014. The Placenta Harbors a Unique Microbiome. Science translational medicine 6(237):237ra265.

Aas JA, Paster BJ, Stokes LN, Olsen I, and Dewhirst FE. 2005. Defining the normal bacterial flora of the oral cavity. Journal of clinical microbiology 43(11):5721-5732.

Abbas Z, Soomro GB, Hassan SM, and Luck NH. 2014. Clinical presentation of hepatitis D in Pakistani children. European journal of gastroenterology & hepatology 26(10):1098-1103.

Abdulla RY, Rice MA, Donauer S, Hicks KR, Poore D, and Staat MA. 2010. Hepatitis A in Internationally Adopted Children: Screening for Acute and Previous Infections. Pediatrics 126(5):e1039-e1044.

Abe K, Inami T, Asano K, Miyoshi C, Masaki N, Hayashi S, Ishikawa K, Takebe Y, Win KM, El-Zayadi AR et al. . 1999. TT virus infection is widespread in the general populations from different geographic regions. Journal of clinical microbiology 37(8):2703-2705.

Abeles SR, Robles-Sikisaka R, Ly M, Lum AG, Salzman J, Boehm TK, and Pride DT. 2014. Human oral viruses are personal, persistent and gender-consistent. The ISME journal.

Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, Bradshaw CJ, Townsend G, Soltysiak A, Alt KW et al. . 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nature genetics 45(4):450-455, 455e451.

Agostini HT, Jobes DV, and Stoner GL. 2002. Molecular Evolution and Epidemiology of JC Virus. Human Polyomaviruses: John Wiley & Sons, Inc. p 491-526.

Agostini HT, Yanagihara R, Davis V, Ryschkewitsch CF, and Stoner GL. 1997. Asian genotypes of JC virus in Native Americans and in a Pacific Island population: Markers of viral evolution and human migration. Proceedings of the National Academy of Sciences 94(26):14542-14546.

Aherfi S, La Scola B, Pagnier I, Raoult D, and Colson P. 2014. The expanding family Marseilleviridae. Virology 466-467C:27-37.

Albenberg LG, and Wu GD. 2014. Diet and the Intestinal Microbiome: Associations, Functions, and Implications for Health and Disease. Gastroenterology.

Ali SH. 2008. Stigmatized Ethnicity, Public Health, and Globalization. Canadian Ethnic Studies Journal 40(3).

Anderson C. 2013. When polio came home: How ordinary people overcame extraordinary challenges. Minneapolis, MN: Words & Deeds Incorporated dba CAB Press.

Anthony SJ, Epstein JH, Murray KA, Navarrete-Macias I, Zambrana-Torrelio CM, Solovyov A, Ojeda-Flores R, Arrigo NC, Islam A, Ali Khan S et al. . 2013. A strategy to estimate unknown viral diversity in mammals. mBio 4(5).

Appelt S, Fancello L, Le Bailly M, Raoult D, Drancourt M, and Desnues C. 2014. Viruses in a 14th-century coprolite. Applied and environmental microbiology.

Armelagos GJ, and Barnes K. 1999. The evolution of human disease and the rise of allergy: Epidemiological transitions. Medical anthropology 18(2):187-213.

Armelagos GJ, and McArdle A. 1976. Population, Disease and Evolution. Memoirs of the Society for American Archaeology 30:1-10.

Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM et al. . 2011. Enterotypes of the human gut microbiome. Nature 473(7346):174-180.

Assetta B, Maginnis MS, Gracia Ahufinger I, Haley SA, Gee GV, Nelson CDS, O'Hara BA, Allen Ramdial S-aA, and Atwood WJ. 2013. 5-HT2 Receptors Facilitate JC Polyomavirus Entry. Journal of virology 87(24):13490-13498.

Astrom KE, Mancall EL, and Richardson EP, Jr. 1958. Progressive multifocal leuko-encephalopathy; a hitherto unrecognized complication of chronic lymphatic leukaemia and Hodgkin's disease. Brain : a journal of neurology 81(1):93-111.

Atlani-Duault L, and Kendall C. 2009. Influenza, anthropology, and global uncertainties. Medical anthropology 28(3):207-211.

Ault GS, and Stoner GL. 1992. Two major types of JC virus defined in progressive multifocal leukoencephalopathy brain by early and late coding region DNA sequences. The Journal of general virology 73 (Pt 10):2669-2678.

Avershina E, Storro O, Oien T, Johnsen R, Pope P, and Rudi K. 2014. Major faecal microbiota shifts in composition and diversity with age in a geographically restricted cohort of mothers and their children. FEMS Microbiol Ecol 87(1):280-290.

Backhed F, Ley RE, Sonnenburg JL, Peterson DA, and Gordon JI. 2005. Host-bacterial mutualism in the human intestine. Science (New York, NY) 307(5717):1915-1920.

Badet C, and Thebaud NB. 2008. Ecology of lactobacilli in the oral cavity: a review of literature. The open microbiology journal 2:38-48.

Baer HA, Singer M, and Susser I. 1997. Medical Anthropology and the World System: A Critical Perspective: Praeger.

Bamford DH, Grimes JM, and Stuart DI. 2005. What does structure tell us about virus evolution? Current opinion in structural biology 15(6):655-663.

Barnett T, and Whiteside A. 2002. AIDS in the 21st Century. Disease and Globalization. Houndmills: Palgrave and Macmillan.

Barros de Freitas R, Durigon EL, Oliveira Dde S, Romano CM, Castro de Freitas MR, Linhares Ada C, Melo FL, Walshkeller L, Barbosa ML, Huatuco EM et al. . 2007. The "pressure pan" evolution of human erythrovirus B19 in the Amazon, Brazil. Virology 369(2):281-287.

Barry KA, Wojcicki BJ, Middelbos IS, Vester BM, Swanson KS, and Fahey GC, Jr. 2010. Dietary cellulose, fructooligosaccharides, and pectin modify fecal protein catabolites and microbial populations in adult cats. Journal of animal science 88(9):2978-2987.

Baseman JG, and Koutsky LA. 2005. The epidemiology of human papillomavirus infections. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology 32 Suppl 1:S16-24.
Bashir NS. 2007. Recent Developments in Design and Application of Plant Virus-based Gene Vectors. Recent Patents on DNA & Gene Sequences 1:214-226.
Bate SL, Dollard SC, and Cannon MJ. 2010. Cytomegalovirus seroprevalence in the United States: the national health and nutrition examination surveys, 1988-2004. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 50(11):1439-1447.
Bauer DW, Huffman JB, Homa FL, and Evilevitch A. 2013. Herpes virus genome, the pressure is on. Journal of the American Chemical Society 135(30):11216-11221.
Baum D. 2008. Reading a Phylogenetic Tree: The Meaning of Monophyletic Groups. Nature Education 1(1):190.
Bedarida S, Dutour O, Buzhilova AP, de Micco P, and Biagini P. 2011. Identification of viral DNA (Anelloviridae) in a 200-year-old dental pulp sample (Napoleon's Great Army, Kaliningrad, 1812). Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 11(2):358-362.
Beijerinck MW. 1899. Bemerkung zu dem Aufsatz von Herrn Iwanowsky uber die Mosaikkrankheit der Tabakspflanze. Centralblatt fur Bakteriologie, Parasitenkunde und Infektionskrankheiten Abt. I(5):310-311.
Belay E. 1999. Transmissible Spongiform Encephalopathies in Humans. Annual review of microbiology 53:283-314.
Belyi VA, Levine AJ, and Skalka AM. 2010a. Sequences from ancestral single-stranded DNA viruses in vertebrate genomes: the parvoviridae and circoviridae are more than 40 to 50 million years old. Journal of virology 84(23):12458-12462.
Belyi VA, Levine AJ, and Skalka AM. 2010b. Unexpected inheritance: multiple integrations of ancient bornavirus and ebolavirus/marburgvirus sequences in vertebrate genomes. PLoS pathogens 6(7):e1001030.
Bentley R, and Meganathan R. 1982. Biosynthesis of vitamin K (menaquinone) in bacteria. Microbiological reviews 46(3):241-280.
Berg Miller ME, Yeoman CJ, Chia N, Tringe SG, Angly FE, Edwards RA, Flint HJ, Lamed R, Bayer EA, and White BA. 2012. Phage-bacteria relationships and CRISPR elements revealed by a metagenomic survey of the rumen microbiome. Environ Microbiol 14(1):207-227.
Berg RD. 1996. The indigenous gastrointestinal microflora. Trends in microbiology 4(11):430-435.
Betsem E, Rua R, Tortevoye P, Froment A, and Gessain A. 2011. Frequent and recent human acquisition of simian foamy viruses through apes' bites in central Africa. PLoS pathogens 7(10):e1002306.
Bianchetti MG, Minder I, Beretta-Piccoli C, Meier A, and Weidmann P. 1982. Effects of tyramine on blood pressure and plasma catecholamines in normal and hypertensive subjects. Klinische Wochenschrift 60(9):465-470.
Bianconi E, Piovesan A, Facchin F, Beraudi A, Casadei R, Frabetti F, Vitale L, Pelleri MC, Tassani S, Piva F et al. . 2013. An estimation of the number of cells in the human body. Annals of human biology 40(6):463-471.

Bibby K, and Peccia J. 2013. Identification of Viral Pathogen Diversity in Sewage Sludge by Metagenome Analysis. Environmental science & technology.

Bisgaard H, Li N, Bonnelykke K, Chawes BL, Skov T, Paludan-Muller G, Stokholm J, Smith B, and Krogfelt KA. 2011. Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age. The Journal of allergy and clinical immunology 128(3):646-652 e641-645.

Blake IM, Martin R, Goel A, Khetsuriani N, Everts J, Wolff C, Wassilak S, Aylward RB, and Grassly NC. 2014. The role of older children and adults in wild poliovirus transmission. Proceedings of the National Academy of Sciences 111(29):10604-10609.

Blinkova O, Victoria J, Li Y, Keele BF, Sanz C, Ndjango J-BN, Peeters M, Travis D, Lonsdorf EV, Wilson ML et al. . 2010. Novel circular DNA viruses in stool samples of wild-living chimpanzees. Journal of General Virology 91(1):74-86.

Boisvert S, Laviolette F, and Corbeil J. 2010. Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. Journal of computational biology : a journal of computational molecular cell biology 17(11):1519-1533.

Bolten C. 2014. Articulating the Invisible: Ebola Beyond Witchcraft in Sierra Leone. Cultural Anthropology Online.

Bowser FP. 1974. The African Slave in Colonial Peru 1524-1650. California: Stanford University Press.

Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG et al. . 2009. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proceedings of the National Academy of Sciences of the United States of America 106(51):21848-21853.

Bradley J, Rajaram S, Moses S, Bhattacharjee P, Lobo AM, Ramesh BM, Washington R, and Alary M. 2011. Changes in HIV knowledge, and socio-cultural and sexual attitudes in South India from 2003-2009. BMC public health 11 Suppl 6:S12.

Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, Mahaffy JM, Mueller J, Nulton J, Rayhawk S et al. . 2008. Viral diversity and dynamics in an infant gut. Research in microbiology 159(5):367-373.

Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, and Rohwer F. 2003. Metagenomic analyses of an uncultured viral community from human feces. J Bacteriol 185(20):6220-6223.

Brentlinger KL, Hafenstein S, Novak CR, Fane BA, Borgon R, McKenna R, and Agbandje-McKenna M. 2002. Microviridae, a family divided: isolation, characterization, and genome sequence of phiMH2K, a bacteriophage of the obligate intracellular parasitic bacterium Bdellovibrio bacteriovorus. J Bacteriol 184(4):1089-1094.

Briese T, Paweska JT, McMullan LK, Hutchison SK, Street C, Palacios G, Khristova ML, Weyer J, Swanepoel R, Egholm M et al. . 2009. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. PLoS pathogens 5(5):e1000455.

Brotman RM. 2011. Vaginal microbiome and sexually transmitted infections: an epidemiologic perspective. The Journal of clinical investigation 121(12):4610-4617.
Brown H, and Kelly AH. 2014. Material proximities and hotspots: toward an anthropology of viral hemorrhagic fevers. Medical anthropology quarterly 28(2):280-303.
Brown P, Breguet G, Smallwood L, Ney R, Moerdowo RM, and Gerety RJ. 1985. Serologic Markers of Hepatitis A and B in the Population of Bali, Indonesia. American Journal of Tropical Medicine and Hygiene 34(3):616-619.
Brugge D, and Missaghian M. 2006. Protecting the Navajo People through tribal regulation of research. Science and engineering ethics 12(3):491-507.
Brun P, Castagliuolo I, Di Leo V, Buda A, Pinzani M, Palu G, and Martines D. 2007. Increased intestinal permeability in obese mice: new evidence in the pathogenesis of nonalcoholic steatohepatitis. American journal of physiology Gastrointestinal and liver physiology 292(2):G518-525.
Bultman SJ. 2014. Emerging roles of the microbiome in cancer. Carcinogenesis 35(2):249-255.
Burnet F. 1962. Natural History of Infectious Disease. Cambridge: Cambridge University Press.
Butler G. 2010. Fungal sex and pathogenesis. Clinical microbiology reviews 23(1):140-159.
Candela M, Biagi E, Maccaferri S, Turroni S, and Brigidi P. 2012. Intestinal microbiota is a plastic factor responding to environmental changes. Trends in microbiology 20(8):385-391.
Cannon MJ, Schmid DS, and Hyde TB. 2010. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. Reviews in medical virology 20(4):202-213.
Canny GO, and McCormick BA. 2008. Bacteria in the Intestine, Helpful Residents or Enemies from Within? Infection and immunity 76(8):3360-3373.
Cano RJ, Rivera-Perez J, Toranzos GA, Santiago-Rodriguez TM, Narganes-Storde YM, Chanlatte-Baik L, García-Roldán E, Bunkley-Williams L, and Massey SE. 2014. Paleomicrobiology: Revealing Fecal Microbiomes of Ancient Indigenous Cultures. PloS one 9(9):e106833.
Cantalupo PG, Calgua B, Zhao G, Hundesa A, Wier AD, Katz JP, Grabe M, Hendrix RW, Girones R, Wang D et al. . 2011. Raw sewage harbors diverse viral populations. mBio 2(5).
Canuti M, and van der Hoek L. 2014. Virus discovery: are we scientists or genome collectors? Trends in microbiology 22(5):229-231.
Caston JC, Eaton CL, Gheorghiu BP, and Ware LL. 2002. Tyramine induced hypertensive episodes and panic attacks in hereditary deficient monoamine oxidase patients: case reports. Journal of the South Carolina Medical Association (1975) 98(4):187-192.
Cayres-Vallinoto IMV, Vallinoto ACR, Azevedo VN, Machado LFA, Ishak MdOG, and Ishak R. 2012. Human JCV Infections as a Bio-Anthropological Marker of the Formation of Brazilian Amazonian Populations. PloS one 7(10):e46523.

CDC. 2012. Measles - United States, 2011. In: McClean H, editor: Centers for Disease Control and Prevention.

Center PSMSHM. 2014. Enterovirus D68 expected to fade away into winter. ScienceDaily: ScienceDaily.

Chang JY, Antonopoulos DA, Kalra A, Tonelli A, Khalife WT, Schmidt TM, and Young VB. 2008. Decreased diversity of the fecal Microbiome in recurrent Clostridium difficile-associated diarrhea. The Journal of infectious diseases 197(3):435-438.

Chehadeh W, Kurien SS, and Nampoory MR. 2013. Molecular Characterization of BK and JC Viruses Circulating among Potential Kidney Donors in Kuwait. BioMed Research International 2013:7.

Chen HM, Yu YN, Wang JL, Lin YW, Kong X, Yang CQ, Yang L, Liu ZJ, Yuan YZ, Liu F et al. . 2013. Decreased dietary fiber intake and structural alteration of gut microbiota in patients with advanced colorectal adenoma. The American journal of clinical nutrition 97(5):1044-1052.

Chen J, He X, and Huang J. 2014. Diet effects in gut microbiome and obesity. Journal of food science 79(4):R442-451.

Chen YE, and Tsao H. 2013. The skin microbiome: current perspectives and future challenges. Journal of the American Academy of Dermatology 69(1):143-155.

Chibani-Chennoufi S, Bruttin A, Dillmann M-L, and Brüssow H. 2004. Phage-Host Interaction: an Ecological Perspective. J Bacteriol 186(12):3677-3686.

Chowell G, Viboud C, Simonsen L, Miller MA, Hurtado J, Soto G, Vargas R, Guzman MA, Ulloa M, and Munayco CV. 2011. The 1918-1920 influenza pandemic in Peru. Vaccine 29 Suppl 2:B21-26.

Chun A. 2004. An Anthropologist Caught in SARS Hysteria. Anthropology News 45(2):10-11.

Cicin-Sain L, Brien JD, Uhrlaub JL, Drabig A, Marandu TF, and Nikolich-Zugich J. 2012. Cytomegalovirus Infection Impairs Immune Responses and Accentuates T-cell Pool Changes Observed in Mice with Aging. PLoS pathogens 8(8):e1002849.

Claverie JM. 2006. Viruses take center stage in cellular evolution. Genome biology 7(6):110.

Cleeland LM, Reichard MV, Tito RY, Reinhard KJ, and Lewis CM, Jr. 2013. Clarifying Prehistoric Parasitism from a Complementary Morphological and Molecular Approach. J Archaeol Sci 40(7):3060-3066.

Clemente JC, Ursell LK, Parfrey LW, and Knight R. 2012. The impact of the gut microbiota on human health: an integrative view. Cell 148(6):1258-1270.

Cockburn T. 1967. The Evolution of Human Infectious Diseases. In: Cockburn T, editor. Infectious Diseases: Their Evolution and Eradication. Springfield, IL: Charles C. Thomas. p 38-49.

Collis M. 2012. People of the Rainforest - The Matses.

Colson P, Richet H, Desnues C, Balique F, Moal V, Grob JJ, Berbis P, Lecoq H, Harle JR, Berland Y et al. . 2010. Pepper mild mottle virus, a plant virus associated with specific immune responses, Fever, abdominal pains, and pruritus in humans. PloS one 5(4):e10041.

Cordaux R, and Batzer MA. 2009. The impact of retrotransposons on human genome evolution. Nature reviews Genetics 10(10):691-703.
Costard S, Wieland B, de Glanville W, Jori F, Rowlands R, Vosloo W, Roger F, Pfeiffer DU, and Dixon LK. 2009. African swine fever: how can global spread be prevented? Philosophical Transactions of the Royal Society B: Biological Sciences 364(1530):2683-2696.
Cox AJ, West NP, and Cripps AW. 2014. Obesity, inflammation, and the gut microbiota. The lancet Diabetes & endocrinology.
Creely SJ, McTernan PG, Kusminski CM, Fisher f M, Da Silva NF, Khanolkar M, Evans M, Harte AL, and Kumar S. 2007. Lipopolysaccharide activates an innate immune system response in human adipose tissue in obesity and type 2 diabetes. American journal of physiology Endocrinology and metabolism 292(3):E740-747.
Crimp D. 1988. How to Have Promiscuity in an Epidemic. AIDS: Cultural Analysis, Cultural Activism. Cambridge, MA: MIT Press.
Crossley BM, Mock RE, Callison SA, and Hietala SK. 2012. Identification and characterization of a novel alpaca respiratory coronavirus most closely related to the human coronavirus 229E. Viruses 4(12):3689-3700.
Cryan JF, and O'Mahony SM. 2011. The microbiome-gut-brain axis: from bowel to behavior. Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society 23(3):187-192.
Cucchiara S, Iebba V, Conte MP, and Schippa S. 2009. The microbiota in inflammatory bowel disease in different age groups. Digestive diseases (Basel, Switzerland) 27(3):252-258.
d'Herelle F. 1917. Sur un microbe invisible antagoniste des bacilles dysenteriques. CR Acad Sci Paris 165:373-375.
d'Herelle F. 1921. Sur l'historique du bacteriophage. CR Soc Biol Paris 84:863-864.
d'Herelle F. 1931. Bacteriophage as a Treatment in Acute Medical and Surgical Infections. Bulletin of the New York Academy of Medicine 7(5):329-348.
da Silva ST, dos Santos CA, and Bressan J. 2013. Intestinal microbiota; relevance to obesity and modulation by prebiotics and probiotics. Nutricion hospitalaria 28(4):1039-1048.
Dacheux L, Cervantes-Gonzalez M, Guigon G, Thiberge JM, Vandenbogaert M, Maufrais C, Caro V, and Bourhy H. 2014. A preliminary study of viral metagenomics of French bat species in contact with humans: identification of new Mammalian viruses. PloS one 9(1):e87194.
Dalke K. 2003. The Human Virome.
Davenport ER, Mizrahi-Man O, Michelini K, Barreiro LB, Ober C, and Gilad Y. 2014. Seasonal Variation in Human Gut Microbiome Composition. PloS one 9(3):e90731.
David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA et al. . 2013. Diet rapidly and reproducibly alters the human gut microbiome. Nature.
Davidov Y, and Jurkevitch E. 2009. Predation between prokaryotes and the origin of eukaryotes. BioEssays : news and reviews in molecular, cellular and developmental biology 31(7):748-757.

De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, Collini S, Pieraccini G, and Lionetti P. 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proceedings of the National Academy of Sciences of the United States of America 107(33):14691-14696.

de Meij TG, Budding AE, Grasman ME, Kneepkens CM, Savelkoul PH, and Mearin ML. 2013. Composition and diversity of the duodenal mucosa-associated microbiome in children with untreated coeliac disease. Scandinavian journal of gastroenterology 48(5):530-536.

De Paepe M, Leclerc M, Tinsley CR, and Petit MA. 2014. Bacteriophages: an underestimated role in human and animal health? Frontiers in cellular and infection microbiology 4:39.

de Souza VAUF, Sumita LM, Nascimento M-C, Oliveira J, Mascheretti M, Quiroga M, Freire WS, Tateno A, Boulos M, Mayaud P et al. . 2007. Human Herpesvirus-8 Infection and Oral Shedding in Amerindian and Non-Amerindian Populations in the Brazilian Amazon Region. Journal of Infectious Diseases 196(6):844-852.

Deininger PL, and Batzer MA. 1999. Alu repeats and human disease. Molecular genetics and metabolism 67(3):183-193.

Delwart E. 2012. Animal virus discovery: improving animal health, understanding zoonoses, and opportunities for vaccine development. Current opinion in virology 2(3):344-352.

Delwart E, and Li L. 2012. Rapidly expanding genetic diversity and host range of the Circoviridae viral family and other Rep encoding small circular ssDNA genomes. Virus research 164(1-2):114-121.

DeStefano F, and Chen RT. 2001. Autism and measles-mumps-rubella vaccination: controversy laid to rest? CNS drugs 15(11):831-837.

Devaraj S, Hemarajata P, and Versalovic J. 2013. The Human Gut Microbiome and Body Metabolism: Implications for Obesity and Diabetes. Clinical chemistry.

Di Bella S, Drapeau C, Garcia-Almodovar E, and Petrosillo N. 2013. Fecal microbiota transplantation: the state of the art. Infectious disease reports 5(2):e13.

Dicks LM, Silvester M, Lawson PA, and Collins MD. 2000. Lactobacillus fornicalis sp. nov., isolated from the posterior fornix of the human vagina. International Journal of Systematic and Evolutionary Microbiology 50(3):1253-1258.

Diemer GS, and Stedman KM. 2012. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. Biology direct 7:13.

Dilger H, and Luig U, editors. 2012. Mortality, Hope and Grief: Anthropologies of AIDS in Africa. Oxford: Berghahn Books.

Dimmitt RA, Staley EM, Chuang G, Tanner SM, Soltau TD, and Lorenz RG. 2010. Role of postnatal acquisition of the intestinal microbiome in the early development of immune function. Journal of pediatric gastroenterology and nutrition 51(3):262-273.

Ding T, and Schloss PD. 2014. Dynamics and associations of microbial community types across the human body. Nature.

Diniz JA, Nunes MR, Travassos da Rosa AP, Cruz AC, de Souza W, Medeiros DB, Chiang JO, and Vasconcelos PF. 2006. Characterization of two new

rhabdoviruses isolated from midges (Culicoides SPP) in the Brazilian Amazon: proposed members of a new genus, Bracorhabdovirus. Archives of virology 151(12):2519-2527.
Dobyns H. 1983. Their number become thinned: Native American population dynamics in Eastern North America. Knoxville, TN: University of Tennessee.
Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, and Knight R. 2010. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proceedings of the National Academy of Sciences 107(26):11971-11975.
Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunger MK, and Bultman SJ. 2011. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell metabolism 13(5):517-526.
dos Santos HF, Knak MB, de Castro FL, Slongo J, Ritterbusch GA, Klein TA, Esteves PA, Silva AD, Trevisol IM, Claassen EA et al. . 2012. Variants of the recently discovered avian gyrovirus 2 are detected in Southern Brazil and The Netherlands. Veterinary microbiology 155(2-4):230-236.
Dotzauer A, and Kraemer L. 2012. Innate and adaptive immune response against picornaviruses and their counteractions: An overview. World Journal of Virology 1(3):91-107.
Dowd JB, Aiello AE, Chyu L, Huang YY, and McDade TW. 2011. Cytomegalovirus antibodies in dried blood spots: a minimally invasive method for assessing stress, immune function, and aging. Immunity & ageing : I & A 8(1):3.
Drexler JF, Geipel A, Konig A, Corman VM, van Riel D, Leijten LM, Bremer CM, Rasche A, Cottontail VM, Maganga GD et al. . 2013. Bats carry pathogenic hepadnaviruses antigenically related to hepatitis B virus and capable of infecting human hepatocytes. Proceedings of the National Academy of Sciences of the United States of America 110(40):16151-16156.
Dubos R. 1965. Man Adapting. New Haven, CT: Yale University Press.
Dunne EU, ER, and Sternberg M. 2007. Prevalence of HPV infection among females in the United States. Journal of the American Medical Association 297(8):813-819.
Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GGZ, Boling L, Barr JJ, Speth DR, Seguritan V, Aziz RK et al. . 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nature communications 5.
Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32(5):1792-1797.
Edgar RC. 2007. PILER-CR: fast and accurate identification of CRISPR repeats. BMC bioinformatics 8:18.
Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics (Oxford, England) 26(19):2460-2461.
Elliott DE, and Weinstock JV. 2009. Helminthic therapy: using worms to treat immune-mediated disease. Advances in experimental medicine and biology 666:157-166.
Emelyanov VV. 2001. Evolutionary relationship of Rickettsiae and mitochondria. FEBS letters 501(1):11-18.
Emerman M, and Malik HS. 2010. Paleovirology--modern consequences of ancient viruses. PLoS Biol 8(2):e1000301.

Epstein LG. 1990. Perspective: Human immunodeficiency virus infection in children. American Journal of Human Biology 2(4):365-372.
Estroff S. 1993. Identity, Disability and Schizophrenia: The Problem of Chronicity. In: Lindenbaum S, and Lock M, editors. Knowledge, Power and Practice: The Anthropology of Medicine in Everyday Life. Berkeley: University of California Press.
Eustace Montgomery R. 1921. On A Form of Swine Fever Occurring in British East Africa (Kenya Colony). Journal of Comparative Pathology and Therapeutics 34(0):159-191.
Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, Clemente JC, Knight R, Heath AC, Leibel RL et al. . 2013. The long-term stability of the human gut microbiota. Science (New York, NY) 341(6141):1237439.
Farmer P. 1990. The Exotic and the Mundane: Haitian Immunodeficiency Virus in Haiti. Human Nature 1:415-446.
Farmer P. 2001. Infections and Inequalities: University of California Press.
Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, and Huttenhower C. 2012. Microbial co-occurrence relationships in the human microbiome. PLoS computational biology 8(7):e1002606.
Feher E, Pazar P, Kovacs E, Farkas SL, Lengyel G, Jakab F, Martella V, and Banyai K. 2014. Molecular detection and characterization of human gyroviruses identified in the ferret fecal virome. Archives of virology.
Feldman D. 1985. AIDS and Social Change. Human Organization 44(4):343-348.
Fernandez-Cobo M, Agostini HT, Britez G, Ryschkewitsch CF, and Stoner GL. 2002. Strains of JC virus in Amerind-speakers of North America (Salish) and South America (Guaraní), Na-Dene-speakers of New Mexico (Navajo), and modern Japanese suggest links through an ancestral Asian population†. Am J Phys Anthropol 118(2):154-168.
Feschotte C, and Gilbert C. 2012. Endogenous viruses: insights into viral evolution and impact on host biology. Nature reviews Genetics 13(4):283-296.
Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, and Knight R. 2010. Forensic identification using skin bacterial communities. Proceedings of the National Academy of Sciences of the United States of America 107(14):6477-6481.
Fineberg H. 1988. The Social Dimensions of AIDS. Scientific American:128-134.
Finkbeiner SR, Allred AF, Tarr PI, Klein EJ, Kirkwood CD, and Wang D. 2008. Metagenomic analysis of human diarrhea: viral detection and discovery. PLoS pathogens 4(2):e1000011.
Firth C, Bhat M, Firth MA, Williams SH, Frye MJ, Simmonds P, Conte JM, Ng J, Garcia J, Bhuva NP et al. . 2014. Detection of Zoonotic Pathogens and Characterization of Novel Viruses Carried by Commensal Rattus norvegicus in New York City. mBio 5(5).
Flanagan ML, Parrish CR, Cobey S, Glass GE, Bush RM, and Leighton TJ. 2012. Anticipating the species jump: surveillance for emerging viral threats. Zoonoses and public health 59(3):155-163.
Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM et al. . 1995. Whole-genome random

sequencing and assembly of Haemophilus influenzae Rd. Science (New York, NY) 269(5223):496-512.
Fleming A. 1946. Penicillin, its practical application. Philadelphia,: Blakiston. 380 p.
Ford J. 1931. Arrowsmith.
Forterre P. 2010. Defining Life: The Virus Viewpoint. Orig Life Evol Biosph 40(2):151-160.
Foster JA, and McVey Neufeld KA. 2013. Gut-brain axis: how the microbiome influences anxiety and depression. Trends in neurosciences.
Foulongne V, Sauvage V, Hebert C, Dereure O, Cheval J, Gouilh MA, Pariente K, Segondy M, Burguiere A, Manuguerra JC et al. . 2012. Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing. PloS one 7(6):e38499.
Francavilla R, Ercolini D, Piccolo M, Vannini L, Siragusa S, De Filippis F, De Pasquale I, Di Cagno R, Di Toma M, Gozzi G et al. . 2014. Salivary microbiota and metabolome associated with celiac disease. Applied and environmental microbiology 80(11):3416-3425.
Frank DN, and Pace NR. 2008. Gastrointestinal microbiology enters the metagenomics era. Current opinion in gastroenterology 24(1):4-10.
Frankel WL, Zhang W, Singh A, Klurfeld DM, Don S, Sakata T, Modlin I, and Rombeau JL. 1994. Mediation of the trophic effects of short-chain fatty acids on the rat jejunum and colon. Gastroenterology 106(2):375-380.
Fujino K, Horie M, Honda T, Merriman DK, and Tomonaga K. 2014. Inhibition of Borna disease virus replication by an endogenous bornavirus-like element in the ground squirrel genome. Proceedings of the National Academy of Sciences 111(36):13175-13180.
Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, Nakanishi Y, Uetake C, Kato K, Kato T et al. . 2013. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature 504(7480):446-450.
Galley A. 2009. City of Plagues?: Toronto, SARS, and the anxieties of Globalization. vis-à-vis: Explorations in Anthropology 9(1):133-142.
Ganesh B, Banyai K, Martella V, Jakab F, Masachessi G, and Kobayashi N. 2012. Picobirnavirus infections: viral persistence and zoonotic potential. Reviews in medical virology 22(4):245-256.
Gest H. 2004. The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. Notes and records of the Royal Society of London 58(2):187-201.
Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M et al. . 2014. The treatment-naive microbiome in new-onset Crohn's disease. Cell host & microbe 15(3):382-392.
Gia Phan T, Phung Vo N, Sdiri-Loulizi K, Aouni M, Pothier P, Ambert-Balay K, Deng X, and Delwart E. 2013. Divergent gyroviruses in the feces of Tunisian children. Virology 446(1-2):346-348.
Gilbert C, Schaack S, Pace JK, 2nd, Brindley PJ, and Feschotte C. 2010a. A role for host-parasite interactions in the horizontal transfer of transposons across phyla. Nature 464(7293):1347-1350.

Gilbert JA, Meyer F, Jansson J, Gordon J, Pace N, Tiedje J, Ley R, Fierer N, Field D, Kyrpides N et al. . 2010b. The Earth Microbiome Project: Meeting report of the "1 EMP meeting on sample selection and acquisition" at Argonne National Laboratory October 6 2010. Standards in genomic sciences 3(3):249-253.

Gilstrap DL, and Kraft M. 2013. Asthma and the host-microbe interaction. The Journal of allergy and clinical immunology 131(5):1449-1450 e1443.

Giordano MO, Martinez LC, Rinaldi D, Guinard S, Naretto E, Casero R, Yacci MR, Depetris AR, Medeot SI, and Nates SV. 1998. Detection of picobirnavirus in HIV-infected patients with diarrhea in Argentina. Journal of acquired immune deficiency syndromes and human retrovirology : official publication of the International Retrovirology Association 18(4):380-383.

Glaser R, Pearson GR, Jones JF, Hillhouse J, Kennedy S, Mao HY, and Kiecolt-Glaser JK. 1991. Stress-related activation of Epstein-Barr virus. Brain, behavior, and immunity 5(2):219-232.

Glendinning L, and Free A. 2014. Supra-organismal interactions in the human intestine. Frontiers in cellular and infection microbiology 4:47.

Gogarten JF, Akoua-Koffi C, Calvignac-Spencer S, Leendertz SA, Weiss S, Couacy-Hymann E, Kone I, Peeters M, Wittig RM, Boesch C et al. . 2014. The ecology of primate retroviruses - an assessment of 12 years of retroviral studies in the Tai national park area, Cote dIvoire. Virology 460-461:147-153.

Gonzalez A, Stombaugh J, Lozupone C, Turnbaugh PJ, Gordon JI, and Knight R. 2011. The mind-body-microbial continuum. Dialogues in clinical neuroscience 13(1):55-62.

Goren M, Yosef I, Edgar R, and Qimron U. 2012. The bacterial CRISPR/Cas system as analog of the mammalian adaptive immune system. RNA biology 9(5):549-554.

Green EC. 1985. Traditional healers, mothers and childhood diarrheal disease in Swaziland: the interface of anthropology and health education. Social science & medicine (1982) 20(3):277-285.

Green EC, Jurg A, and Djedje A. 1994. The Snake in the Stomach: Child Diarrhea in Central Mozambique. Medical anthropology quarterly 8(1):4-24.

Greenhough B. 2012. Room with a rhinovirus? Blurring the boundaries between research and therapeutic space. Transactions of the Institute of British Geographers 37(3):402-417.

Grose J, Jensen G, Burnett S, and Breakwell D. 2014. Genomic comparison of 93 Bacillus phages reveals 12 clusters, 14 singletons and remarkable diversity. BMC genomics 15(1):855.

Group TNHW, Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA et al. . 2009. The NIH Human Microbiome Project. Genome research 19(12):2317-2323.

Guo Z, Zheng SP, Sugimoto C, Wang YL, Zheng H-Y, Takasaka T, Kitamura T, Guo J, and Yogo Y. 2001. JC Virus Genotypes in Northwestern China

Implications for Its Population History. Anthropological Science 109(3):203-212.

Haberman Y, Tickle TL, Dexheimer PJ, Kim M-O, Tang D, Karns R, Baldassano RN, Noe JD, Rosh J, Markowitz J et al. . 2014. Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. The Journal of clinical investigation 124(8):3617-3633.

Hadjipanayis A, Hadjichristodoulou C, Kallias M, Sava K, Petsa A, Demetriadou K, Christodoulou C, Constantinou A, and Sidera M. 1999. Prevalence of antibodies to hepatitis A among children and adolescents in Larnaca area, Cyprus. European journal of epidemiology 15(10):903-905.

Haeckel E. 1866. Generelle Morphologie der Organismen.

Haft DH, Selengut J, Mongodin EF, and Nelson KE. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS computational biology 1(6):e60.

Haldane JBS. 1949. Disease and Evolution. La Ric Sci Suppl 19:68-76.

Hall RJ, Wang J, Todd AK, Bissielo AB, Yen S, Strydom H, Moore NE, Ren X, Huang QS, Carter PE et al. . 2014. Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. J Virol Methods 195(0):194-204.

Hamblin S, and Tanaka M. 2013. Behavioural manipulation of insect hosts by Baculoviridae as a process of niche construction. BMC evolutionary biology 13(1):170.

Hamer HM, Jonkers D, Venema K, Vanhoutvin S, Troost FJ, and Brummer RJ. 2008. Review article: the role of butyrate on colonic function. Alimentary pharmacology & therapeutics 27(2):104-119.

Hanage WP. 2014. Microbiology: Microbiome science needs a healthy dose of scepticism. Nature 512(7514):247-248.

Handelsman J, Rondon MR, Brady SF, Clardy J, and Goodman RM. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chemistry & Biology 5(10):R245-R249.

Handley SA, Thackray LB, Zhao G, Presti R, Miller AD, Droit L, Abbink P, Maxfield LF, Kambal A, Duan E et al. . 2012. Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome. Cell 151(2):253-266.

Hansen H, and Groce NE. 2001. From quarantine to condoms: Shifting policies and problems of HIV control in Cuba. Medical anthropology 19(3):259-292.

Häring M, Rachel R, Peng X, Garrett RA, and Prangishvili D. 2005. Viral Diversity in Hot Springs of Pozzuoli, Italy, and Characterization of a Unique Archaeal Virus, Acidianus Bottle-Shaped Virus, from a New Family, the Ampullaviridae. Journal of virology 79(15):9904-9911.

Haynes M, and Rohwer F. 2011. The Human Virome. In: Nelson KE, editor. Metagenomics of the Human Body. New York: Springer Science+Business Media. p 63-77.

He B, Li Z, Yang F, Zheng J, Feng Y, Guo H, Li Y, Wang Y, Su N, Zhang F et al. . 2013. Virome profiling of bats from myanmar by metagenomic analysis of tissue samples reveals more novel Mammalian viruses. PloS one 8(4):e61950.

Heijtz RD, Wang S, Anuar F, Qian Y, Björkholm B, Samuelsson A, Hibberd ML, Forssberg H, and Pettersson S. 2011. Normal gut microbiota modulates brain development and behavior. Proceedings of the National Academy of Sciences 108(7):3047-3052.

Heilbronn R, Albrecht I, Stephan S, Bürkle A, and zur Hausen H. 1993. Human cytomegalovirus induces JC virus DNA replication in human fibroblasts. Proceedings of the National Academy of Sciences 90(23):11406-11410.
Heliconius Genome C. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487(7405):94-98.
Hendrix RW. 2005. Bacteriophage HK97: Assembly of the Capsid and Evolutionary Connections. In: Polly R, editor. Advances in virus research: Academic Press. p 1-14.
Henry AG, Brooks AS, and Piperno DR. 2011. Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets (Shanidar III, Iraq; Spy I and II, Belgium). Proceedings of the National Academy of Sciences of the United States of America 108(2):486-491.
Heringlake S, Ockenga J, Tillmann HL, Trautwein C, Meissner D, Stoll M, Hunt J, Jou C, Solomon N, Schmidt RE et al. . 1998. GB virus C/hepatitis G virus infection: a favorable prognostic factor in human immunodeficiency virus-infected patients? The Journal of infectious diseases 177(6):1723-1726.
Hershey AD, and Chase M. 1952. INDEPENDENT FUNCTIONS OF VIRAL PROTEIN AND NUCLEIC ACID IN GROWTH OF BACTERIOPHAGE. The Journal of General Physiology 36(1):39-56.
Hewlett BS, and Amola RP. 2003. Cultural contexts of Ebola in northern Uganda. Emerg Infect Dis 9(10):1242-1248.
Hirt RP, Logsdon JM, Jr., Healy B, Dorey MW, Doolittle WF, and Embley TM. 1999. Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proceedings of the National Academy of Sciences of the United States of America 96(2):580-585.
Högger P. 2013. Nutrition-derived bioactive metabolites produced by gut microbiota and their potential impact on human health. Nutrition and Medicine 1(1).
Holland KT, and Bojar RA. 2002. Cosmetics: what is their influence on the skin microflora? American journal of clinical dermatology 3(7):445-449.
Holtz LR, Cao S, Zhao G, Bauer IK, Denno DM, Klein EJ, Antonio M, Stine OC, Snelling TL, Kirkwood CD et al. . 2014. Geographic variation in the eukaryotic virome of human diarrhea. Virology 468-470:556-564.
Hooper LV, and Macpherson AJ. 2010. Immune adaptations that maintain homeostasis with the intestinal microbiota. Nature reviews Immunology 10(3):159-169.
Hopkins MJ, and Macfarlane GT. 2002. Changes in predominant bacterial populations in human faeces with age and with Clostridium difficile infection. Journal of medical microbiology 51(5):448-454.
Hopkins TG, and Wood N. 2013. Female human papillomavirus (HPV) vaccination: Global uptake and the impact of attitudes. Vaccine 31(13):1673-1679.
Horie M, Honda T, Suzuki Y, Kobayashi Y, Daito T, Oshida T, Ikuta K, Jern P, Gojobori T, Coffin JM et al. . 2010. Endogenous non-retroviral RNA virus elements in mammalian genomes. Nature 463(7277):84-87.
Hornig M, Briese T, Licinio J, Khabbaz RF, Altshuler LL, Potkin SG, Schwemmle M, Siemetzki U, Mintz J, Honkavuori K et al. . 2012. Absence of evidence for bornavirus infection in schizophrenia, bipolar disorder and major depressive disorder. Mol Psychiatry 17(5):486-493.

Horvath P, and Barrangou R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science (New York, NY) 327(5962):167-170.
Horz HP, and Conrads G. 2011. Methanogenic Archaea and oral infections - ways to unravel the black box. Journal of oral microbiology 3.
Hou D, Zhou X, Zhong X, Settles ML, Herring J, Wang L, Abdo Z, Forney LJ, and Xu C. 2013. Microbiota of the seminal fluid from healthy and infertile men. Fertility and sterility 100(5):1261-1269.
Huang Y, Callahan S, and Hadfield MG. 2012. Recruitment in the sea: bacterial genes required for inducing larval settlement in a polychaete worm. Sci Rep 2.
Huffnagle GB, and Noverr MC. 2013. The emerging world of the fungal microbiome. Trends in microbiology.
Hulo C, de Castro E, Masson P, Bougueleret L, Bairoch A, Xenarios I, and Le Mercier P. 2011. ViralZone: a knowledge resource to understand virus diversity. Nucleic acids research 39(Database issue):D576-582.
Human Microbiome Project C. 2012. Structure, function and diversity of the healthy human microbiome. Nature 486(7402):207-214.
IHS. 2013. Indian Health Disparities.
Ikegaya H, and Iwase H. 2004. Trial for the geographical identification using JC viral genotyping in Japan. Forensic science international 139(2–3):169-172.
Ikegaya H, Zheng HY, Saukko PJ, Varesmaa-Korhonen L, Hovi T, Vesikari T, Suganami H, Takasaka T, Sugimoto C, Ohasi Y et al. . 2005. Genetic diversity of JC virus in the Saami and the Finns: implications for their population history. Am J Phys Anthropol 128(1):185-193.
Iliev ID, Funari VA, Taylor KD, Nguyen Q, Reyes CN, Strom SP, Brown J, Becker CA, Fleshner PR, Dubinsky M et al. . 2012. Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. Science (New York, NY) 336(6086):1314-1317.
Inhorn MC, and Brown PJ. 1997. Introduction. In: Inhorn MC, and Brown PJ, editors. The Anthropology of Infectious Disease: International Health Perspectives. Atlanta, GA: Emory University. p 3-29.
Ishino Y, Shinagawa H, Makino K, Amemura M, and Nakata A. 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol 169(12):5429-5433.
Ivanovski DI. 1899. Ueber die Mosaikkrankheit der Tabakspflanze. Centralblatt fur Bakteriologie, Parasitenkunde und Infektionskrankheiten Abt. II(5):250-254.
Jacobs SE, Lamson DM, St George K, and Walsh TJ. 2013. Human rhinoviruses. Clinical microbiology reviews 26(1):135-162.
Jansen R, Embden JD, Gaastra W, and Schouls LM. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Molecular microbiology 43(6):1565-1575.
Jiang W, Wu N, Wang X, Chi Y, Zhang Y, Qiu X, Hu Y, Li J, and Liu Y. 2015. Dysbiosis gut microbiota associated with inflammation and impaired mucosal immune function in intestine of humans with non-alcoholic fatty liver disease. Sci Rep 5.

Johnson WE, and Coffin JM. 1999. Constructing primate phylogenies from ancient retrovirus sequences. Proceedings of the National Academy of Sciences 96(18):10254-10260.
Joshi N, and Fass J. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software].
Kamada N, Seo SU, Chen GY, and Nunez G. 2013. Role of the gut microbiota in immunity and inflammatory disease. Nature reviews Immunology 13(5):321-335.
Kapoor A, Slikas E, Simmonds P, Chieochansin T, Naeem A, Shaukat S, Alam MM, Sharif S, Angez M, Zaidi S et al. . 2009. A newly identified bocavirus species in human stool. The Journal of infectious diseases 199(2):196-200.
Karlsson CL, Onnerfalt J, Xu J, Molin G, Ahrne S, and Thorngren-Jerneck K. 2012. The microbiota of the gut in preschool children with normal and excessive body weight. Obesity (Silver Spring, Md) 20(11):2257-2261.
Kata A. 2010. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. Vaccine 28(7):1709-1716.
Katoh I, and Kurata S-i. 2013. Association of endogenous retroviruses and long terminal repeats with human disorders. Frontiers in Oncology 3.
Katzourakis A, Rambaut A, and Pybus OG. 2005. The evolutionary dynamics of endogenous retroviruses. Trends in microbiology 13(10):463-468.
Kazanji M, Dussart P, Duprez R, Tortevoye P, Pouliquen JF, Vandekerkhove J, Couppié P, Morvan J, Talarmin A, and Gessain A. 2005. Serological and Molecular Evidence That Human Herpesvirus 8 Is Endemic among Amerindians in French Guiana. Journal of Infectious Diseases 192(9):1525-1529.
Keeling PJ, and Palmer JD. 2008. Horizontal gene transfer in eukaryotic evolution. Nature reviews Genetics 9(8):605-618.
Kembel SW, Jones E, Kline J, Northcutt D, Stenson J, Womack AM, Bohannan BJM, Brown GZ, and Green JL. 2012. Architectural design influences the diversity and structure of the built environment microbiome. The ISME journal 6(8):1469-1479.
Kembel SW, Meadow JF, O'Connor TK, Mhuireach G, Northcutt D, Kline J, Moriyama M, Brown GZ, Bohannan BJM, and Green JL. 2014. Architectural Design Drives the Biogeography of Indoor Bacterial Communities. PloS one 9(1):e87093.
Kew OM, Sutter RW, de Gourville EM, Dowdle WR, and Pallansch MA. 2005. VACCINE-DERIVED POLIOVIRUSES AND THE ENDGAME STRATEGY FOR GLOBAL POLIO ERADICATION*. Annual review of microbiology 59(1):587-635.
Khoruts A, Dicksved J, Jansson JK, and Sadowsky MJ. 2010. Changes in the composition of the human fecal microbiome after bacteriotherapy for recurrent Clostridium difficile-associated diarrhea. Journal of clinical gastroenterology 44(5):354-360.
Kiecolt-Glaser JK, Fisher LD, Ogrocki P, Stout JC, Speicher CE, and Glaser R. 1987. Marital quality, marital disruption, and immune function. Psychosomatic medicine 49(1):13-34.

Kim MS, Park EJ, Roh SW, and Bae JW. 2011. Diversity and abundance of single-stranded DNA viruses in human feces. Applied and environmental microbiology 77(22):8062-8070.
Kim MS, Whon TW, and Bae JW. 2013. Comparative viral metagenomics of environmental samples from Korea. Genomics & informatics 11(3):121-128.
Kincaid RP, Burke JM, Cox JC, de Villiers E-M, and Sullivan CS. 2013. A Human Torque Teno Virus Encodes a MicroRNA That Inhibits Interferon Signaling. PLoS pathogens 9(12):e1003818.
King CA, Baillie J, and Sinclair JH. 2006. Human cytomegalovirus modulation of CCR5 expression on myeloid cells affects susceptibility to human immunodeficiency virus type 1 infection. The Journal of general virology 87(Pt 8):2171-2180.
Kitchen A, Miyamoto MM, and Mulligan CJ. 2008. Utility of DNA viruses for studying human host history: case study of JC virus. Molecular phylogenetics and evolution 46(2):673-682.
Kitchen A, Shackelton LA, and Holmes EC. 2011. Family level phylogenies reveal modes of macroevolution in RNA viruses. Proceedings of the National Academy of Sciences 108(1):238-243.
Kleinman AM, Bloom BR, Saich A, Mason KA, and Aulino F. 2008. Avian and Pandemic Influenza: A Biosocial Approach. Journal of Infectious Diseases 197(Supplement 1):S1-S3.
Koch H, and Schmid-Hempel P. 2011. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. Proceedings of the National Academy of Sciences 108(48):19288-19292.
Koch R, and Cheyne WW. 1880. Investigations into the etiology of traumatic infective diseases. London,: New Sydenham society. xiii, 74 p. p.
Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, and Ley RE. 2011. Succession of microbial consortia in the developing infant gut microbiome. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1:4578-4585.
Kofman A, Marcinkiewicz L, Dupart E, Lyshchev A, Martynov B, Ryndin A, Kotelevskaya E, Brown J, Schiff D, and Abounader R. 2011. The roles of viruses in brain tumor initiation and oncomodulation. J Neurooncol 105(3):451-466.
Kolb AW, Ané C, and Brandt CR. 2013. Using HSV-1 Genome Phylogenetics to Track Past Human Migrations. PloS one 8(10):e76267.
Koonin EV, and Dolja VV. 2013. A virocentric perspective on the evolution of life. Current opinion in virology.
Koonin EV, and Martin W. 2005. On the origin of genomes and cells within inorganic compartments. Trends in genetics : TIG 21(12):647-654.
Koplow DA. 2003. Smallpox : the fight to eradicate a global scourge. Berkeley: University of California Press. ix, 265 p. p.
Koppers-Lalic D, and Hoeben RC. 2011. Non-human viruses developed as therapeutic agent for use in humans. Reviews in medical virology 21(4):227-239.
Koren O, Knights D, Gonzalez A, Waldron L, Segata N, Knight R, Huttenhower C, and Ley RE. 2013. A guide to enterotypes across the human body: meta-analysis of

microbial community structures in human microbiome datasets. PLoS computational biology 9(1):e1002863.
Kraus RH, Kerstens HH, van Hooft P, Megens HJ, Elmberg J, Tsvey A, Sartakov D, Soloviev SA, Crooijmans RP, Groenen MA et al. . 2012. Widespread horizontal genomic exchange does not erode species barriers among sympatric ducks. BMC evolutionary biology 12:45.
Krupovic M, Prangishvili D, Hendrix RW, and Bamford DH. 2011. Genomics of Bacterial and Archaeal Viruses: Dynamics within the Prokaryotic Virosphere. Microbiology and Molecular Biology Reviews 75(4):610-635.
Kumamoto CA. 2011. Inflammation and gastrointestinal Candida colonization. Current opinion in microbiology 14(4):386-391.
Lamendella R, Domingo JW, Ghosh S, Martinson J, and Oerther DB. 2011. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. BMC microbiology 11:103.
Lan Y, Kriete A, and Rosen GL. 2013. Selecting age-related functional characteristics in the human gut microbiome. Microbiome 1(1):2.
Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W et al. . 2001. Initial sequencing and analysis of the human genome. Nature 409(6822):860-921.
Langendorfer A, Davenport W, London WT, Blumberg BS, and Mazzur S. 1984. Sex-related differences in transmission of hepatitis B infection in a Melanesian population. Am J Phys Anthropol 64(3):243-254.
Langmead B, Trapnell C, Pop M, and Salzberg S. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10(3):R25.
Lathrop SK, Bloom SM, Rao SM, Nutsch K, Lio CW, Santacruz N, Peterson DA, Stappenbeck TS, and Hsieh CS. 2011. Peripheral education of the immune system by colonic commensal microbiota. Nature 478(7368):250-254.
Lazarczyk M, Cassonnet P, Pons C, Jacob Y, and Favre M. 2009. The EVER Proteins as a Natural Barrier against Papillomaviruses: a New Insight into the Pathogenesis of Human Papillomavirus Infections. Microbiology and Molecular Biology Reviews 73(2):348-370.
Lazarevic V, Whiteson K, Huse S, Hernandez D, Farinelli L, Osteras M, Schrenzel J, and Francois P. 2009. Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. J Microbiol Methods 79(3):266-271.
LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D, and Ventura M. 2013. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. Current opinion in biotechnology 24(2):160-168.
Lederberg J, and McCray AT. 2001. 'Ome sweet 'omics - A genealogical treasury of words. Scientist 15(7):8.
Leopold M. 2005. Inside West Nile: Violence, History, and Representation on an African Frontier: SAR Press.
Lewis DA, Brown R, Williams J, White P, Jacobson SK, Marchesi JR, and Drake MJ. 2013. The human urinary microbiome; bacterial DNA in voided urine of asymptomatic adults. Frontiers in cellular and infection microbiology 3:41.
Lewis S. 1925. Arrowsmith. New York: Grosset & Dunlap.

Ley RE. 2010. Obesity and the human microbiome. Current opinion in gastroenterology 26(1):5-11.
Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, and Gordon JI. 2005. Obesity alters gut microbial ecology. Proceedings of the National Academy of Sciences of the United States of America 102(31):11070-11075.
Ley RE, Lozupone CA, Hamady M, Knight R, and Gordon JI. 2008. Worlds within worlds: evolution of the vertebrate gut microbiota. Nature reviews Microbiology 6(10):776-788.
Ley RE, Peterson DA, and Gordon JI. 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 124(4):837-848.
Leyden JJ, McGinley KJ, Nordstrom KM, and Webster GF. 1987. Skin microflora. The Journal of investigative dermatology 88(3 Suppl):65s-72s.
Li H-C, Fujiyoshi T, Lou H, Yashiki S, Sonoda S, Cartier L, Nunez L, Munoz I, Horai S, and Tajima K. 1999. The presence of ancient human T-cell lymphotropic virus type I provirus DNA in an Andean mummy. Nature medicine 5(12):1428-1432.
Li L, Kapoor A, Slikas B, Bamidele OS, Wang C, Shaukat S, Masroor MA, Wilson ML, Ndjango JB, Peeters M et al. . 2010. Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. Journal of virology 84(4):1674-1682.
Li L, McGraw S, Zhu K, Leutenegger CM, Marks SL, Kubiski S, Gaffney P, Dela Cruz FN, Jr., Wang C, Delwart E et al. . 2013. Circovirus in tissues of dogs with vasculitis and hemorrhage. Emerg Infect Dis 19(4):534-541.
Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Crameri G, Hu Z, Zhang H et al. . 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science (New York, NY) 310(5748):676-679.
Lieberman PM. 2014. Virology. Epstein-Barr virus turns 50. Science (New York, NY) 343(6177):1323-1325.
Lim YW, Schmieder R, Haynes M, Willner D, Furlan M, Youle M, Abbott K, Edwards R, Evangelista J, Conrad D et al. . 2012. Metagenomics and metatranscriptomics: Windows on CF-associated viral and microbial communities. Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society.
Lin A, Bik EM, Costello EK, Dethlefsen L, Haque R, Relman DA, and Singh U. 2013. Distinct distal gut microbiome diversity and composition in healthy children from bangladesh and the United States. PloS one 8(1):e53838.
Lindenbaum S. 1979. Kuru Sorcery: Disease and Danger in the New Guinea Highlands. Palo Alto, CA: Mayfield.
Liou AP, Paziuk M, Luevano J-M, Machineni S, Turnbaugh PJ, and Kaplan LM. 2013. Conserved Shifts in the Gut Microbiota Due to Gastric Bypass Reduce Host Weight and Adiposity. Science translational medicine 5(178):178ra141.
Lister J. 1909. The collected papers of Joseph Baron Lister microform. Oxford: Clarendon Press.
Llosa MV. 1981. The War of the End of the World: Picador.

Lopez-Bueno A, Tamames J, Velazquez D, Moya A, Quesada A, and Alcami A. 2009. High diversity of the viral community from an Antarctic lake. Science (New York, NY) 326(5954):858-861.
Los M, Panigrahi S, Rashedi I, Mandal S, Stetefeld J, Essmann F, and Schulze-Osthoff K. 2009. Apoptin, a tumor-selective killer. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1793(8):1335-1342.
Losada L, Ghedin E, Morris A, Chu HW, and Nierman WC. 2011. The Human Lung Microbiome. In: Nelson KE, editor. Metagenomics of the Human Body. New York: Springer.
Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, and Knight R. 2012. Diversity, stability and resilience of the human gut microbiota. Nature 489(7415):220-230.
Lucht E, Brytting M, Bjerregaard L, Julander I, and Linde A. 1998. Shedding of Cytomegalovirus and Herpesviruses 6, 7, and 8 in Saliva of Human Immunodeficiency Virus Type 1—Infected Patients and Healthy Controls. Clinical Infectious Diseases 27(1):137-141.
Lukashov VV, and Goudsmit J. 2001. Evolutionary relationships among parvoviruses: virus-host coevolution among autonomous primate parvoviruses and links between adeno-associated and avian parvoviruses. Journal of virology 75(6):2729-2740.
Ly M, Abeles SR, Boehm TK, Robles-Sikisaka R, Naidu M, Santiago-Rodriguez T, and Pride DT. 2014. Altered Oral Viral Ecology in Association with Periodontal Disease. mBio 5(3).
Lysholm F, Wetterbom A, Lindau C, Darban H, Bjerkner A, Fahlander K, Lindberg AM, Persson B, Allander T, and Andersson B. 2012. Characterization of the viral microbiome in patients with severe lower respiratory tract infections, using metagenomic sequencing. PloS one 7(2):e30875.
Ma W, Vincent AL, Gramer MR, Brockwell CB, Lager KM, Janke BH, Gauger PC, Patnayak DP, Webby RJ, and Richt JA. 2007. Identification of H2N3 influenza A viruses from swine in the United States. Proceedings of the National Academy of Sciences 104(52):20949-20954.
Macdonald TT, and Monteleone G. 2005. Immunity, inflammation, and allergy in the gut. Science (New York, NY) 307(5717):1920-1925.
Magner LN. 2009. A History of Infectious Diseases and the Microbial World. Westport, CT: Praeger Publishers.
Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF et al. . 2011. Evolution and classification of the CRISPR-Cas systems. Nature reviews Microbiology 9(6):467-477.
Makivuokko H, Tiihonen K, Tynkkynen S, Paulin L, and Rautonen N. 2010. The effect of age and non-steroidal anti-inflammatory drugs on human intestinal microbiota composition. The British journal of nutrition 103(2):227-234.
Mandar R. 2013. Microbiota of male genital tract: impact on the health of man and his partner. Pharmacological research : the official journal of the Italian Pharmacological Society 69(1):32-41.
Margulis L. 1970. Origin of eukaryotic cells. New Haven, CT: Yale University.

Mariat D, Firmesse O, Levenez F, Guimaraes V, Sokol H, Dore J, Corthier G, and Furet JP. 2009. The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. BMC microbiology 9:123.
Marraffini LA, and Sontheimer EJ. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nature reviews Genetics 11(3):181-190.
Marsland BJ, Yadava K, and Nicod LP. 2013. The airway microbiome and disease. Chest 144(2):632-637.
Martin W, and Kowallik K. 1999. Annotated English translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche'. European Journal of Phycology 34(03):287-295.
Martin W, and Mentel M. 2010. The Origin of Mitochondria. Nature Education 3(9):58.
Martinez JM, Swan BK, and Wilson WH. 2014. Marine viruses, a genetic reservoir revealed by targeted viromics. The ISME journal 8(5):1079-1088.
Masembe C, Michuki G, Onyango M, Rumberia C, Norling M, Bishop R, Djikeng A, Kemp S, Orth A, Skilton R et al. . 2012. Viral metagenomics demonstrates that domestic pigs are a potential reservoir for Ndumu virus. Virology Journal 9(1):218.
Mason GM, Poole E, Sissons JGP, Wills MR, and Sinclair JH. 2012. Human cytomegalovirus latency alters the cellular secretome, inducing cluster of differentiation (CD)4+ T-cell migration and suppression of effector function. Proceedings of the National Academy of Sciences 109(36):14538-14543.
Mason MR, Nagaraja HN, Camerlengo T, Joshi V, and Kumar PS. 2013. Deep Sequencing Identifies Ethnicity-Specific Bacterial Signatures in the Oral Microbiome. PloS one 8(10):e77287.
Mast MD, M.P.H, Eric E., and Krawczynski MD, Ph.D, Krzysztof. 1996. HEPATITIS E: An Overview1. Annual Review of Medicine 47(1):257-266.
Mathieu LG, and Sonea S. 1995. A powerful bacterial world. Endeavour 19(3):112-117.
Matsuda M, Jona M, Yasui K, and Nagashima K. 1987. Genetic characterization of JC virus Tokyo-1 strain, a variant oncogenic in rodents. Virus research 7(2):159-168.
May JM. 1958. The Ecology of Human Disease. New York: MD Publications.
McGeoch DJ, Cook S, Dolan A, Jamieson FE, and Telford EA. 1995. Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. Journal of molecular biology 247(3):443-458.
McInerney JO, O'Connell MJ, and Pisani D. 2014. The hybrid nature of the Eukaryota and a consilient view of life on Earth. Nat Rev Micro 12(6):449-455.
Meadow JF, Altrichter AE, Kembel SW, Kline J, Mhuireach G, Moriyama M, Northcutt D, O'Connor TK, Womack AM, Brown GZ et al. . 2014. Indoor airborne bacterial communities are influenced by ventilation, occupancy, and outdoor air source. Indoor Air 24(1):41-48.
Meadow JF, Bateman AC, Herkert KM, O'Connor TK, and Green JL. 2013. Significant changes in the skin microbiome mediated by the sport of roller derby. PeerJ 1:e53.
Meisler MH, and Ting CN. 1993. The remarkable evolutionary history of the human amylase genes. Critical reviews in oral biology and medicine : an official publication of the American Association of Oral Biologists 4(3-4):503-509.

Michaud V, Randriamparany T, and Albina E. 2013. Comprehensive Phylogenetic Reconstructions of African Swine Fever Virus: Proposal for a New Classification and Molecular Dating of the Virus. PloS one 8(7):e69662.
Mills S, Shanahan F, Stanton C, Hill C, Coffey A, and Ross RP. 2013. Movers and shakers: influence of bacteriophages in shaping the mammalian gut microbiota. Gut Microbes 4(1):4-16.
Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, and Bushman FD. 2013. Rapid evolution of the human gut virome. Proceedings of the National Academy of Sciences of the United States of America.
Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, Lewis JD, and Bushman FD. 2011. The human gut virome: inter-individual variation and dynamic response to diet. Genome research 21(10):1616-1625.
Miranda JJ, Sugimoto C, Paraguison R, Takasaka T, Zheng H-Y, and Yogo Y. 2003. Genetic diversity of JC virus in the modern Filipino population: Implications for the peopling of the Philippines. Am J Phys Anthropol 120(2):125-132.
Mistry N, Wibom C, and Evander M. 2008. Cutaneous and mucosal human papillomaviruses differ in net surface charge, potential impact on tropism. Virology Journal 5(1):118.
Mora C, Tittensor DP, Adl S, Simpson AGB, and Worm B. 2011. How Many Species Are There on Earth and in the Ocean? PLoS Biol 9(8):e1001127.
Moran M, and Hoffman D. 2014. Introduction: Ebola in Perspective. Cultural Anthropology Online.
Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L et al. . 2013. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. Am J Respir Crit Care Med 187(10):1067-1075.
Morsy S. 1979. The missing link in medical anthropology: The political economy of health. Reviews in Anthropology 6(3):349-363.
Moses P, and Moses J. 1983. Haiti and the Acquired Immune Deficiency Syndrome. Annals of Internal Medicine 99(4):565.
Mullis KB, and Faloona FA. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods in enzymology 155:335-350.
Munoz N, Castellsague X, de Gonzalez AB, and Gissmann L. 2006. Chapter 1: HPV in the etiology of human cancer. Vaccine 24 Suppl 3:S3/1-10.
Murray KE, Adams RF, Earl JW, and Shaw KJ. 1986. Studies of the free faecal amines of infants with gastroenteritis and of healthy infants. Gut 27(10):1173-1180.
Nadin-Davis SA, and Real LA. 2011. Chapter 11 - Molecular Phylogenetics of the Lyssaviruses—Insights from a Coalescent Approach. In: Alan CJ, editor. Advances in virus research: Academic Press. p 203-238.
Nakano K, Nemoto H, Nomura R, Inaba H, Yoshioka H, Taniguchi K, Amano A, and Ooshima T. 2009. Detection of oral bacteria in cardiovascular specimens. Oral microbiology and immunology 24(1):64-68.
Nakashima K, Yamada L, Satou Y, Azuma J, and Satoh N. 2004. The evolutionary origin of animal cellulose synthase. Development genes and evolution 214(2):81-88.

Nakayama T, and Archibald J. 2012. Evolving a photosynthetic organelle. BMC biology 10(1):35.
Nardis C, Mosca L, and Mastromarino P. 2013. Vaginal microbiota and viral sexually transmitted diseases. Annali di igiene : medicina preventiva e di comunita 25(5):443-456.
Nasidze I, Li J, Quinque D, Tang K, and Stoneking M. 2009. Global diversity in the human salivary microbiome. Genome research 19(4):636-643.
Negro F. 2014. Hepatitis D virus coinfection and superinfection. Cold Spring Harbor perspectives in medicine 4(11):a021550.
Nelson B. 2010. Lagging HPV vaccination rates dampen outlook in U.S.: cloudy prospects for overcoming multiple barriers. Cancer Cytopathology 118(3):113-114.
Nelson DE, Dong Q, Van der Pol B, Toh E, Fan B, Katz BP, Mi D, Rong R, Weinstock GM, Sodergren E et al. . 2012. Bacterial communities of the coronal sulcus and distal urethra of adolescent males. PloS one 7(5):e36298.
Neufeld KM, Kang N, Bienenstock J, and Foster JA. 2011. Reduced anxiety-like behavior and central neurochemical change in germ-free mice. Neurogastroenterology & Motility 23(3):255-e119.
Ng TF, Mesquita JR, Nascimento MS, Kondov NO, Wong W, Reuter G, Knowles NJ, Vega E, Esona MD, Deng X et al. . 2014. Feline fecal virome reveals novel and prevalent enteric viruses. Veterinary microbiology.
Nguyen VK, and Peschard K. 2003. Anthropology, Inequality and Disease: A Review. Annual Review of Anthropology 32:447-474.
Niccoli S, Abraham S, Richard C, and Zehbe I. 2012. The Asian-American E6 variant protein of human papillomavirus 16 alone is sufficient to promote immortalization, transformation, and migration of primary human foreskin keratinocytes. Journal of virology 86(22):12384-12396.
Niel C, and Gomes SA. 2002. TT virus infection in Brazilian Amazon Indians and comparison with non-Indian subjects. Vox sanguinis 82(3):162.
Nishizawa T, Okamoto H, Konishi K, Yoshizawa H, Miyakawa Y, and Mayumi M. 1997. A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. Biochemical and biophysical research communications 241(1):92-97.
Norja P, Eis-Hubinger AM, Soderlund-Venermo M, Hedman K, and Simmonds P. 2008. Rapid sequence change and geographical spread of human parvovirus B19: comparison of B19 virus evolution in acute and persistent infections. Journal of virology 82(13):6427-6433.
Norman Jason M, Handley Scott A, Baldridge Megan T, Droit L, Liu Catherine Y, Keller Brian C, Kambal A, Monaco Cynthia L, Zhao G, Fleshner P et al. . 2015. Disease-Specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. Cell 160(3):447-460.
NPR Staff. 2014. The Experts The Ebola Response May Need: Anthropologists.
Nylund L, Satokari R, Nikkila J, Rajilic-Stojanovic M, Kalliomaki M, Isolauri E, Salminen S, and de Vos WM. 2013. Microarray analysis reveals marked intestinal microbiota aberrancy in infants having eczema compared to healthy children in at-risk for atopic disease. BMC microbiology 13:12.

O'Hara AM, and Shanahan F. 2006. The gut flora as a forgotten organ. EMBO reports 7(7):688-693.
O'Connor A, Quizon PM, Albright JE, Lin FT, and Bennett BJ. 2014. Responsiveness of cardiometabolic-related microbiota to diet is influenced by host genetics. Mammalian Genome 25(11-12):583-599.
Obregon-Tito AJ, Tito RY, Metcalf J, Sankaranarayanan K, Clemente JC, Ursell LK, Zech Xu Z, Van Treuren W, Knight R, Gaffney PM et al. . 2015. Subsistence strategies in traditional societies distinguish gut microbiomes. Nature communications 6.
Ogilvie LA, Bowler LD, Caplin J, Dedi C, Diston D, Cheek E, Taylor H, Ebdon JE, and Jones BV. 2013. Genome signature-based dissection of human gut metagenomes to extract subliminal viral sequences. Nature communications 4:2420.
Okamoto H, Akahane Y, Ukita M, Fukuda M, Tsuda F, Miyakawa Y, and Mayumi M. 1998. Fecal excretion of a nonenveloped DNA virus (TTV) associated with posttransfusion non-A-G hepatitis. Journal of medical virology 56(2):128-132.
Oklahoma Indian Affairs Commission. 2011. 2011 Oklahoma Indian Nations Pocket Pictoral Directory.
Ozawa S, Eda H, Hayashi K, Cooperation Group for HSVRVS, Yoshino K, and Yanagi K. 2006. Geographical distribution of the herpes simplex virus type 1 BgKL variant in Japan suggests gradual dispersion of the virus from Shikoku Island to the other Islands. Journal of clinical microbiology 44(6):2109-2118.
Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, and Dessel BH. 1971. Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. Lancet 1(7712):1257-1260.
Paik S, Senty L, Das S, Noe JC, Munro CL, and Kitten T. 2005. Identification of Virulence Determinants for Endocarditis in Streptococcus sanguinis by Signature-Tagged Mutagenesis. Infection and immunity 73(9):6064-6074.
Palahniuk C. 1999. Invisible Monsters. New York: W.W. Norton & Company.
Pallansch M, and Roos R. 2007. Enteroviruses: polioviruses, coxsackieviruses, echoviruses, and newer enteroviruses. In: Knipe D, and Howley P, editors. Fields Virology. 5th ed. Philadelphia: Lippincott Williams and Wilkins. p 839-894.
Panda S, El Khader I, Casellas F, Lopez Vivancos J, Garcia Cors M, Santiago A, Cuenca S, Guarner F, and Manichanh C. 2014. Short-term effect of antibiotics on human gut microbiota. PloS one 9(4):e95476.
Parahitiyawa NB, Scully C, Leung WK, Yam WC, Jin LJ, and Samaranayake LP. 2010. Exploring the oral bacterial flora: current status and future directions. Oral diseases 16(2):136-145.
Park EJ, Kim KH, Abell GC, Kim MS, Roh SW, and Bae JW. 2011. Metagenomic analysis of the viral communities in fermented foods. Applied and environmental microbiology 77(4):1284-1291.
Pastrana DV, Ray U, Magaldi TG, Schowalter RM, Cuburu N, and Buck CB. 2013. BK polyomavirus genotypes represent distinct serotypes with distinct entry tropism. Journal of virology 87(18):10105-10113.
Patterson KD, and Pyle GF. 1991. The geography and mortality of the 1918 influenza pandemic. Bulletin of the history of medicine 65(1):4-21.

Pavesi A. 2004. Detecting Traces of Prehistoric Human Migrations by Geographic Synthetic Maps of Polyomavirus JC. Journal of molecular evolution 58(3):304-313.
Pavesi A. 2005. Utility of JC polyomavirus in tracing the pattern of human migrations dating to prehistoric times. The Journal of general virology 86(Pt 5):1315-1326.
Pawelec G, Derhovanessian E, Larbi A, Strindhall J, and Wikby A. 2009. Cytomegalovirus and human immunosenescence. Reviews in medical virology 19(1):47-56.
Pennisi E. 2011. Microbiology. Going viral: exploring the role of viruses in our bodies. Science (New York, NY) 331(6024):1513.
Pennisi E. 2013. The CRISPR craze. Science (New York, NY) 341(6148):833-836.
Perez-Brocal V, Garcia-Lopez R, Vazquez-Castellanos JF, Nos P, Beltran B, Latorre A, and Moya A. 2013. Study of the viral and microbial communities associated with Crohn's disease: a metagenomic approach. Clinical and translational gastroenterology 4:e36.
Peters M, C. J., Simpson M, PhD, MPH, Gary L., and Levy M, PhD, H. 1999. SPECTRUM OF HANTAVIRUS INFECTION: Hemorrhagic Fever with Renal Syndrome and Hantavirus Pulmonary Syndrome1. Annual Review of Medicine 50(1):531-545.
Peterson SN, Snesrud E, Liu J, Ong AC, Kilian M, Schork NJ, and Bretz W. 2013. The dental plaque microbiome in health and disease. PloS one 8(3):e58487.
Peterson W. 1995. Outbreak.
Petrara MR, Penazzato M, Massavon W, Nabachwa S, Nannyonga M, Mazza A, Gianesin K, Del Bianco P, Lundin R, Sumpter C et al. . 2014. Epstein-Barr virus load in children infected with human immunodeficiency virus type 1 in Uganda. The Journal of infectious diseases 210(3):392-399.
Phan TG, Kapusinszky B, Wang C, Rose RK, Lipton HL, and Delwart EL. 2011. The fecal viral flora of wild rodents. PLoS pathogens 7(9):e1002218.
Phan TG, Vo NP, Boros A, Pankovics P, Reuter G, Li OT, Wang C, Deng X, Poon LL, and Delwart E. 2013. The viruses of wild pigeon droppings. PloS one 8(9):e72787.
Pierce SK, Fang X, Schwartz JA, Jiang X, Zhao W, Curtis NE, Kocot KM, Yang B, and Wang J. 2012. Transcriptomic evidence for the expression of horizontally transferred algal nuclear genes in the photosynthetic sea slug, Elysia chlorotica. Molecular biology and evolution 29(6):1545-1556.
Pitcher MC, Beatty ER, Harris RM, Waring RH, and Cummings JH. 1998. Sulfur metabolism in ulcerative colitis: investigation of detoxification enzymes in peripheral blood. Digestive diseases and sciences 43(9):2080-2085.
Popgeorgiev N, Boyer M, Fancello L, Monteil S, Robert C, Rivet R, Nappez C, Azza S, Chiaroni J, Raoult D et al. . 2013. Giant Blood Marseillevirus recovered from asymptomatic blood donors. The Journal of infectious diseases.
Prasetyo AA, Ariapramuda R, Kindi EA, Dirgahayu P, Sari Y, Dharmawan R, and Kageyama S. 2014. Men having sex with men in Surakarta, Indonesia: demographics, behavioral characteristics and prevalence of blood borne pathogens. Southeast Asian J Trop Med Public Health 45(5):1032-1047.

Prescott LE, MacDonald DM, Davidson F, Mokili J, Pritchard DI, Arnot DE, Riley EM, Greenwood BM, Hamid S, Saeed AA et al. . 1999. Sequence diversity of TT virus in geographically dispersed human populations. The Journal of general virology 80 (Pt 7):1751-1758.

Preston R. 1995. The Hot Zone. New York: Anchor Books. 422 p. p.

Pride DT, Salzman J, Haynes M, Rohwer F, Davis-Long C, White RA, 3rd, Loomer P, Armitage GC, and Relman DA. 2012a. Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. The ISME journal 6(5):915-926.

Pride DT, Salzman J, and Relman DA. 2012b. Comparisons of clustered regularly interspaced short palindromic repeats and viromes in human saliva reveal bacterial adaptations to salivary viruses. Environ Microbiol.

Prideaux L, Kang S, Wagner J, Buckley M, Mahar JE, De Cruz P, Wen Z, Chen L, Xia B, van Langenberg DR et al. . 2013. Impact of ethnicity, geography, and disease on the microbiota in health and inflammatory bowel disease. Inflammatory bowel diseases 19(13):2906-2918.

Prusiner SB. 1998. Prions. Proceedings of the National Academy of Sciences 95(23):13363-13383.

Puigbo P, Wolf YI, and Koonin EV. 2009. Search for a 'Tree of Life' in the thicket of the phylogenetic forest. Journal of biology 8(6):59.

Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T et al. . 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464(7285):59-65.

Raabea VN, and Borcherta M. 2012. Infection control during filoviral hemorrhagic Fever outbreaks. Journal of global infectious diseases 4(1):69-74.

Rac MW, and Sheffield JS. 2014. Prevention and Management of Viral Hepatitis in Pregnancy. Obstetrics and gynecology clinics of North America 41(4):573-592.

Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, and Claverie JM. 2004. The 1.2-megabase genome sequence of Mimivirus. Science (New York, NY) 306(5700):1344-1350.

Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO et al. . 2011. Vaginal microbiome of reproductive-age women. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1:4680-4687.

Redi F. 1668. Esperienze intorno alla generazione degl'insetti. Firenze,: All'insegna della Stella. 3 p. l., 228 p. p.

Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, and Gordon JI. 2010. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 466(7304):334-338.

Reyes A, Semenkovich NP, Whiteson K, Rohwer F, and Gordon JI. 2012. Going viral: next-generation sequencing applied to phage populations in the human gut. Nature reviews Microbiology 10(9):607-617.

Rho M, Tang H, and Ye Y. 2010. FragGeneScan: predicting genes in short and error-prone reads. Nucleic acids research 38(20):e191.

Richards P, and Mokuwa A. 2014. Village Funerals and the Spread of Ebola Virus Disease. Cultural Anthropology Online.

Richards TA, Dacks JB, Jenkinson JM, Thornton CR, and Talbot NJ. 2006. Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms. Current biology : CB 16(18):1857-1864.
Riding R. 2011. The Nature of Stromatolites: 3,500 Million Years of History and a Century of Research. Advances in Stromatolite Geobiology: Springer Berlin Heidelberg. p 29-74.
Ritzhaupt A, Ellis A, Hosie KB, and Shirazi-Beechey SP. 1998. The characterization of butyrate transport across pig and human colonic luminal membrane. The Journal of physiology 507 (Pt 3):819-830.
Robles-Sikisaka R, Ly M, Boehm T, Naidu M, Salzman J, and Pride DT. 2013. Association between living environment and human oral viral ecology. The ISME journal.
Roger AJ, Svard SG, Tovar J, Clark CG, Smith MW, Gillin FD, and Sogin ML. 1998. A mitochondrial-like chaperonin 60 gene in Giardia lamblia: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. Proceedings of the National Academy of Sciences of the United States of America 95(1):229-234.
Rohwer F, and Barott K. 2013. Viral information. Biol Philos 28(2):283-297.
Rohwer F, and Youle M. 2011. Consider something viral in your research. Nat Rev Micro 9(5):308-309.
Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, Galuppi M, Lamont RF, Chaemsaithong P, Miranda J et al. . 2014. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. Microbiome 2(1):4.
Rosario K, Duffy S, and Breitbart M. 2009. Diverse circovirus-like genome architectures revealed by environmental metagenomics. The Journal of general virology 90(Pt 10):2418-2424.
Rosario K, Duffy S, and Breitbart M. 2012. A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. Archives of virology 157(10):1851-1871.
Roux S, Enault F, Robin A, Ravet V, Personnic S, Theil S, Colombet J, Sime-Ngando T, and Debroas D. 2012. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. PloS one 7(3):e33641.
Rowan FE, Docherty NG, Coffey JC, and O'Connell PR. 2009. Sulphate-reducing bacteria and hydrogen sulphide in the aetiology of ulcerative colitis. The British journal of surgery 96(2):151-158.
Rubin VC, and Ford Jr. WK. 1970. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. Astrophysical Journal 159:379.
Rudzik AE, Breakey A, and Bribiescas RG. 2014. Oxytocin and Epstein-Barr virus: Stress biomarkers in the postpartum period among first-time mothers from Sao Paulo, Brazil. American journal of human biology : the official journal of the Human Biology Council 26(1):43-50.
Ruffer MA. 1921. Pathological Note on the Royal Mummies of the Cairo Museum. In: Moodie RL, editor. Studies in the Palaeopathology of Egypt. Illinois: University of Chicago Press.

Ryan FP. 2004. Human endogenous retroviruses in health and disease: a symbiotic perspective. Journal of the Royal Society of Medicine 97(12):560-565.
Rydze RT, Bhattarai N, and Stapleton JT. 2012. GB virus C infection is associated with a reduced rate of reactivation of latent HIV and protection against activation-induced T-cell death. Antiviral therapy 17(7):1271-1279.
Ryschkewitsch CF, Friedlaender JS, Mgone CS, Jobes DV, Agostini HT, Chima SC, Alpers MP, Koki G, Yanagihara R, and Stoner GL. 2000. Human polyomavirus JC variants in Papua New Guinea and Guam reflect ancient population settlement and viral evolution. Microbes and Infection 2(9):987-996.
Rzezutka A, and Cook N. 2004. Survival of human enteric viruses in the environment and food. FEMS microbiology reviews 28(4):441-453.
Saethre E, and Stadler J. 2009. A Tale of Two "Cultures": HIV Risk Narratives in South Africa. Medical anthropology 28(3):268-284.
Sagan L. 1967. On the origin of mitosing cells. Journal of theoretical biology 14(3):225-IN226.
Sanford JA, and Gallo RL. 2013. Functions of the skin microbiota in health and disease. Seminars in immunology 25(5):370-377.
Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes CA, Hutchison CA, Slocombe PM, and Smith M. 1977. Nucleotide sequence of bacteriophage phi X174 DNA. Nature 265(5596):687-695.
Sarwar UN, Sitar S, and Ledgerwood JE. 2011. Filovirus emergence and vaccine development: a perspective for health care practitioners in travel medicine. Travel medicine and infectious disease 9(3):126-134.
Sasaki M, Orba Y, Ueno K, Ishii A, Moonga L, Hang'ombe BM, Mweene AS, Ito K, and Sawa H. 2014. Metagenomic Analysis of Shrew Enteric Virome Reveals Novel Viruses Related to Human Stool-Associated Viruses. The Journal of general virology.
Sattenspiel L, and Castillo-Chavez C. 1990. Environmental context, social interactions, and the spread of HIV. American Journal of Human Biology 2(4):397-417.
Savin KW, Cocks BG, Wong F, Sawbridge T, Cogan N, Savage D, and Warner S. 2010. A neurotropic herpesvirus infecting the gastropod, abalone, shares ancestry with oyster herpesvirus and a herpesvirus associated with the amphioxus genome. Virology Journal 7:308.
Scanlan PD, Shanahan F, Clune Y, Collins JK, O'Sullivan GC, O'Riordan M, Holmes E, Wang Y, and Marchesi JR. 2008. Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. Environ Microbiol 10(3):789-798.
Scheppach W, Bartram HP, and Richter F. 1995. Role of short-chain fatty acids in the prevention of colorectal cancer. European journal of cancer (Oxford, England : 1990) 31A(7-8):1077-1080.
Schimper A. 1883. Über die Entwicklung der Chlorophyllkörner und Farbkörper. Bot Zeitung 41:105-114.
Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ et al. . 2009. Introducing mothur: open-source, platform-independent, community-supported software for

describing and comparing microbial communities. Applied and environmental microbiology 75(23):7537-7541.
Schmitz JE, and Korioth-Schmitz B. 2013. Immunopathogenesis of simian immunodeficiency virus infection in nonhuman primates. Current opinion in HIV and AIDS.
Schwann T, and Schleiden MJ. 1847. Microscopical researches into the accordance in the structure and growth of animals and plants. The Sydenham Society. London: The Sydenham Society,.
Scrimshaw SC, and Hurtado E. 1988. Anthropological involvement in the Central American diarrheal disease control project. Social science & medicine (1982) 27(1):97-105.
Serafino A, Balestrieri E, Pierimarchi P, Matteucci C, Moroni G, Oricchio E, Rasi G, Mastino A, Spadafora C, Garaci E et al. . 2009. The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. Experimental cell research 315(5):849-862.
Shackelton LA, Rambaut A, Pybus OG, and Holmes EC. 2006. JC virus evolution and its association with human populations. Journal of virology 80(20):9928-9933.
Shan T, Li L, Simmonds P, Wang C, Moeser A, and Delwart E. 2011. The fecal virome of pigs on a high-density farm. Journal of virology 85(22):11697-11708.
Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I, and Rosenberg E. 2010. Commensal bacteria play a role in mating preference of Drosophila melanogaster. Proceedings of the National Academy of Sciences 107(46):20051-20056.
Sharp PM, and Hahn BH. 2011. Origins of HIV and the AIDS pandemic. Cold Spring Harbor perspectives in medicine 1(1):a006841.
Shchelkunov SN. 2011. Emergence and reemergence of smallpox: the need for development of a new generation smallpox vaccine. Vaccine 29 Suppl 4:D49-53.
Shchelkunov SN. 2013. An increasing danger of zoonotic orthopoxvirus infections. PLoS pathogens 9(12):e1003756.
Shen PS, Domek MJ, Sanz-García E, Makaju A, Taylor RM, Hoggan R, Culumber MD, Oberg CJ, Breakwell DP, Prince JT et al. . 2012. Sequence and Structural Characterization of Great Salt Lake Bacteriophage CW02, a Member of the T7-Like Supergroup. Journal of virology 86(15):7907-7917.
Siddiqui H, Nederbragt AJ, Lagesen K, Jeansson SL, and Jakobsen KS. 2011. Assessing diversity of the female urine microbiota by high throughput sequencing of 16S rDNA amplicons. BMC microbiology 11:244.
Singer M. 1997. Needle exchange and aids prevention: Controversies, policies and research. Medical anthropology 18(1):1-12.
Singer M. 1998. Articulating Personal Experience and Political Economy in the AIDS epidemic: The Case of Carlos Torres. In: Singer M, editor. The Political Economy of AIDS. Amityville, NY: Baywood Publishing. p 61-73.
Singer M. 2009. Pathogens Gone Wild? Medical Anthropology and the “Swine Flu” Pandemic. Medical anthropology 28(3):199-206.
Singer M, and Baer HA. 2011. Introducing Medical Anthropology: A Discipline in Action: AltaMira Press.

Sleator RD. 2010. The human superorganism - of microbes and men. Medical hypotheses 74(2):214-215.
Smith LC, Lucas KJ, and Latkin C. 1999. Rumor and gossip: Social discourse on HIV and AIDS. Anthropology & Medicine 6(1):121-131.
Smith MI, Yatsunenko T, Manary MJ, Trehan I, Mkakosya R, Cheng J, Kau AL, Rich SS, Concannon P, Mychaleckyj JC et al. . 2013. Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. Science (New York, NY) 339(6119):548-554.
Smith O, Clapham A, Rose P, Liu Y, Wang J, and Allaby RG. 2014. A complete ancient RNA genome: identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus. Scientific reports 4:4003.
Smits SL, Poon LL, van Leeuwen M, Lau PN, Perera HK, Peiris JS, Simon JH, and Osterhaus AD. 2011. Genogroup I and II picobirnaviruses in respiratory tracts of pigs. Emerg Infect Dis 17(12):2328-2330.
Smits SL, van Leeuwen M, Schapendonk CM, Schurch AC, Bodewes R, Haagmans BL, and Osterhaus AD. 2012. Picobirnaviruses in the human respiratory tract. Emerg Infect Dis 18(9):1539-1540.
Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, Caporaso JG, Knights D, Clemente JC, Nakielny S et al. . 2013. Cohabiting family members share microbiota with one another and with their dogs. eLife 2.
Song Y-J, and Stinski MF. 2002. Effect of the human cytomegalovirus IE86 protein on expression of E2F-responsive genes: A DNA microarray analysis. Proceedings of the National Academy of Sciences 99(5):2836-2841.
Sorensen MV, Snodgrass JJ, Leonard WR, McDade TW, Tarskaya LA, Ivanov KI, Krivoshapkin VG, and Alekseev VP. 2009. Lifestyle incongruity, stress and immune function in indigenous Siberians: the health impacts of rapid social and economic change. Am J Phys Anthropol 138(1):62-69.
Stadler J. 2003. Rumor, Gossip and Blame: Implications for HIV/AIDS Prevention in The South African Lowveld. AIDS Education and Prevention 15(4):357-368.
Stanley W. 1935. Isolation of a crystalline protein possessing the properties of tobacco-mosaic virus. Science (New York, NY) 81:644-645.
Stark PL, and Lee A. 1982. The microbial ecology of the large bowel of breast-fed and formula-fed infants during the first year of life. Journal of medical microbiology 15(2):189-203.
Stern A, Mick E, Tirosh I, Sagy O, and Sorek R. 2012. CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. Genome research 22(10):1985-1994.
Stoner GL, Jobes DV, Fernandez Cobo M, Agostini HT, Chima SC, and Ryschkewitsch CF. 2000. JC virus as a marker of human migrationto the Americas. Microbes and Infection 2(15):1905-1911.
Stoner GL, Walker DL, and Webster HD. 1988. Age distribution of progressive multifocal leukoencephalopathy. Acta neurologica Scandinavica 78(4):307-312.
Strauss JH, and Strauss EG. 1988. Evolution of RNA Viruses. Annual review of microbiology 42(1):657-683.

Streicker DG, Lemey P, Velasco-Villa A, and Rupprecht CE. 2012. Rates of viral evolution are linked to host geography in bat rabies. PLoS pathogens 8(5):e1002720.
Subramanian M. 2012. Zoonotic disease risk and the bushmeat trade: assessing awareness among hunters and traders in sierra leone. EcoHealth 9(4):471-482.
Sugimoto J, and Schust DJ. 2009. Review: Human Endogenous Retroviruses and the Placenta. Reproductive Sciences 16(11):1023-1033.
Sullivan MB, Waterbury JB, and Chisholm SW. 2003. Cyanophages infecting the oceanic cyanobacterium Prochlorococcus. Nature 424(6952):1047-1051.
Suttle CA. 2005. Viruses in the sea. Nature 437(7057):356-361.
Suttle CA. 2007. Marine viruses--major players in the global ecosystem. Nature reviews Microbiology 5(10):801-812.
Svensson B. 2007. Indigenous People of the Amazon Native Amazonian Tribe.
Svraka S, Rosario K, Duizer E, van der Avoort H, Breitbart M, and Koopmans M. 2010. Metagenomic sequencing for virus identification in a public-health setting. The Journal of general virology 91(Pt 11):2846-2856.
Swarts DC, Mosterd C, van Passel MW, and Brouns SJ. 2012. CRISPR interference directs strand specific spacer acquisition. PloS one 7(4):e35888.
Syvanen M. 2012. Evolutionary implications of horizontal gene transfer. Annual review of genetics 46:341-358.
Takasaka T, Kitamura T, Sugimoto C, Guo J, Zheng HY, and Yogo Y. 2006. Phylogenetic analysis of major African genotype (Af2) of JC virus: Implications for origin and dispersals of modern Africans. Am J Phys Anthropol 129(3):465-472.
Takasaka T, Miranda JJ, Sugimoto C, Paraguison R, Zheng H-Y, Kitamura T, and Yogo Y. 2004. Genotypes of JC virus in Southeast Asia and the western Pacific: implications for human migrations from Asia to the Pacific. Anthropological Science 112(1):53-59.
Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular biology and evolution 28(10):2731-2739.
Tapper MA, and Hicks RE. 1998. Temperate viruses and lysogeny in Lake Superior bacterioplankton. Limnology and Oceanography 43(1):95-103.
Tate JE, Burton AH, Boschi-Pinto C, Steele AD, Duque J, Parashar UD, and Network WH-cGRS. 2012. 2008 estimate of worldwide rotavirus-associated mortality in children younger than 5 years before the introduction of universal rotavirus vaccination programmes: a systematic review and meta-analysis. The Lancet Infectious diseases 12(2):136-141.
Tatum EL, and Lederberg J. 1947. Gene Recombination in the Bacterium Escherichia coli. J Bacteriol 53(6):673-684.
Taubenberger JK. 2006. The origin and virulence of the 1918 "Spanish" influenza virus. Proceedings of the American Philosophical Society 150:86-112.
Taylor MP, Koyuncu OO, and Enquist LW. 2011. Subversion of the actin cytoskeleton during viral infection. Nature reviews Microbiology 9(6):427-439.

Tengs T, Kristoffersen AB, Bachvaroff TR, and Jonassen CM. 2013. A mobile genetic element with unknown function found in distantly related viruses. Virology Journal 10:132.

Tennoune N, Chan P, Breton J, Legrand R, Chabane YN, Akkermann K, Jarv A, Ouelaa W, Takagi K, Ghouzali I et al. . 2014. Bacterial ClpB heat-shock protein, an antigen-mimetic of the anorexigenic peptide [alpha]-MSH, at the origin of eating disorders. Transl Psychiatry 4:e458.

Theis KR, Schmidt TM, and Holekamp KE. 2012. Evidence for a bacterial mechanism for group-specific social odors among hyenas. Scientific reports 2:615.

Thingstad T, and Lignell R. 1997. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. Aquatic Microbial Ecology 13(1):19-27.

Thomas F. 2007. 'Our Families are Killing Us': HIV/AIDS, Witchcraft and Social Tensions in the Caprivi Region, Namibia. Anthropology & Medicine 14(3):279-291.

Thomas F. 2008. Indigenous Narratives of HIV/AIDS: Morality and Blame in a Time of Change. Medical anthropology 27(3):227-256.

Thompson V, Arnold L, and Notaro S. 2012. African American Parents' HPV Vaccination Intent and Concerns. Journal of Health Care for the Poor and Underserved 23(1):290-301.

Thomson RC, Plachetzki DC, Mahler DL, and Moore BR. 2014. A critical appraisal of the use of microRNA data in phylogenetics. Proceedings of the National Academy of Sciences 111(35):E3659-E3668.

Tillmann HL, Heiken H, Knapik-Botor A, Heringlake S, Ockenga J, Wilber JC, Goergen B, Detmer J, McMorrow M, Stoll M et al. . 2001. Infection with GB virus C and reduced mortality among HIV-infected patients. The New England journal of medicine 345(10):715-724.

Tito RY, Knights D, Metcalf J, Obregon-Tito AJ, Cleeland L, Najar F, Roe B, Reinhard K, Sobolik K, Belknap S et al. . 2012. Insights from characterizing extinct human gut microbiomes. PloS one 7(12):e51146.

Tito RY, Macmil S, Wiley G, Najar F, Cleeland L, Qu C, Wang P, Romagne F, Leonard S, Ruiz AJ et al. . 2008. Phylotyping and functional analysis of two ancient human microbiomes. PloS one 3(11):e3703.

Tlaskalova-Hogenova H, Stepankova R, Kozakova H, Hudcovic T, Vannucci L, Tuckova L, Rossmann P, Hrncir T, Kverka M, Zakostelska Z et al. . 2011. The role of gut microbiota (commensal bacteria) and the mucosal barrier in the pathogenesis of inflammatory and autoimmune diseases and cancer: contribution of germ-free and gnotobiotic animal models of human diseases. Cell Mol Immunol 8(2):110-120.

Toivanen P. 2003. Normal intestinal microbiota in the aetiopathogenesis of rheumatoid arthritis. Annals of the rheumatic diseases 62(9):807-811.

Tom-Orme L. 1994. Traditional beliefs and attitudes about diabetes among Navajos and Utes. In: Young JR, and Young RS, editors. Diabetes as a Disease of Civilization: The Impact of Culture Change on Indigenous Peoples. New York: Mouton de Gruyter. p 271-292.

Tortora GJ, Funke BR, and Case CL. 2013. Microbiology : an introduction. Boston: Pearson. xviii, 818, [121] p. p.

Tremaroli V, and Backhed F. 2012. Functional interactions between the gut microbiota and host metabolism. Nature 489(7415):242-249.

Trovato GM, Castro A, Tonzuso A, Garozzo A, Martines GF, Pirri C, Trovato F, and Catalano D. 2009. Human obesity relationship with Ad36 adenovirus and insulin resistance. International journal of obesity (2005) 33(12):1402-1409.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP et al. . 2009. A core gut microbiome in obese and lean twins. Nature 457(7228):480-484.

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, and Gordon JI. 2007. The human microbiome project. Nature 449(7164):804-810.

Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, and Gordon JI. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444(7122):1027-1031.

Turnbaugh PJ, Quince C, Faith JJ, McHardy AC, Yatsunenko T, Niazi F, Affourtit J, Egholm M, Henrissat B, Knight R et al. . 2010. Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. Proceedings of the National Academy of Sciences of the United States of America 107(16):7503-7508.

Turner DO, Williams-Cocks SJ, Bullen R, Catmull J, Falk J, Martin D, Mauer J, Barber AE, Wang RC, Gerstenberger SL et al. . 2011. High-risk human papillomavirus (HPV) screening and detection in healthy patient saliva samples: a pilot study. BMC oral health 11:28.

Twort FW. 1915. An Investigation on the Nature of Ultra-Microscopic Viruses. The Lancet 186(4814):1241-1243.

Tyakht AV, Kostryukova ES, Popenko AS, Belenikin MS, Pavlenko AV, Larin AK, Karpova IY, Selezneva OV, Semashko TA, Ospanova EA et al. . 2013. Human gut microbiota community structures in urban and rural populations in Russia. Nature communications 4:2469.

Van Borm S, Rosseel T, Steensels M, van den Berg T, and Lambrecht B. 2013. What's in a strain? Viral metagenomics identifies genetic variation and contaminating circoviruses in laboratory isolates of pigeon paramyxovirus type 1. Virus research 171(1):186-193.

van Nood E, Vrieze A, Nieuwdorp M, Fuentes S, Zoetendal EG, de Vos WM, Visser CE, Kuijper EJ, Bartelsman JF, Tijssen JG et al. . 2013. Duodenal infusion of donor feces for recurrent Clostridium difficile. The New England journal of medicine 368(5):407-415.

Vasilyev EV, Trofimov DY, Tonevitsky AG, Ilinsky VV, Korostin DO, and Rebrikov DV. 2009. Torque Teno Virus (TTV) distribution in healthy Russian population. Virology Journal 6:134.

Ventura M, Canchaya C, Bernini V, Altermann E, Barrangou R, McGrath S, Claesson MJ, Li Y, Leahy S, Walker CD et al. . 2006. Comparative genomics and transcriptional analysis of prophages identified in the genomes of Lactobacillus gasseri, Lactobacillus salivarius, and Lactobacillus casei. Applied and environmental microbiology 72(5):3130-3146.

Ventura M, Sozzi T, Turroni F, Matteuzzi D, and van Sinderen D. 2011. The impact of bacteriophages on probiotic bacteria and gut microbiota diversity. Genes & nutrition 6(3):205-207.
Verhulst NO, Qiu YT, Beijleveld H, Maliepaard C, Knights D, Schulz S, Berg-Lyons D, Lauber CL, Verduijn W, Haasnoot GW et al. . 2011. Composition of Human Skin Microbiota Affects Attractiveness to Malaria Mosquitoes. PloS one 6(12):e28991.
Viana S, Parana R, Moreira RC, Compri AP, and Macedo V. 2005. High prevalence of hepatitis B virus and hepatitis D virus in the western Brazilian Amazon. The American journal of tropical medicine and hygiene 73(4):808-814.
Victoria M, Carvalho-Costa FA, Heinemann MB, Leite JP, and Miagostovich MP. 2007. Genotypes and molecular epidemiology of human astroviruses in hospitalized children with acute gastroenteritis in Rio de Janeiro, Brazil. Journal of medical virology 79(7):939-944.
Villarreal LP. 2004. Can Viruses Make Us Human? Proceedings of the American Philosophical Society 148(3):296-323.
Virchow RLK. 1862. Die Cellularpathologie in ihrer Begrundung auf physiologische und pathologische Gewebelehre. Berlin: A. Hirschwald. xviii, 457 p. p.
Virgin HW, Wherry EJ, and Ahmed R. 2009. Redefining chronic viral infection. Cell 138(1):30-50.
Visentin V, Prevot D, De Saint Front VD, Morin-Cussac N, Thalamas C, Galitzky J, Valet P, Zorzano A, and Carpene C. 2004. Alteration of amine oxidase activity in the adipose tissue of obese subjects. Obesity research 12(3):547-555.
Vray B. 2002. [Relationships of molecular biology with paleoparasitology]. Vesalius : acta internationales historiae medicinae 8(1):45-52.
Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, Berelowitz M, Dhillon AP, Thomson MA, Harvey P et al. . 1998. RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. The Lancet 351(9103):637-641.
Walsh AM, Kortschak RD, Gardner MG, Bertozzi T, and Adelson DL. 2013. Widespread horizontal transfer of retrotransposons. Proceedings of the National Academy of Sciences 110(3):1012-1016.
Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J, Radini A, Hancock Y, Tito RY, Fiddyment S et al. . 2014. Pathogens and host immunity in the ancient human oral cavity. Nature genetics 46(4):336-344.
Watson JD, and Crick FHC. 1953. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature 171(4356):737-738.
Weiss RA. 2001. Retroviruses and Cancer. Current Science 81:528-534.
Weiss RA. 2006. The discovery of endegenous getroviruses. Retrovirology 3:67.
Wesley PJ. 2014. Liberia's Ebola Epidemic: Did the Government Fall Asleep at the Wheel? Cultural Anthropology Online.
Wessner DR. 2010. The Origins of Viruses. Nature Education 3(9):37.
Wevers D, Metzger S, Babweteera F, Bieberbach M, Boesch C, Cameron K, Couacy-Hymann E, Cranfield M, Gray M, Harris LA et al. . 2011. Novel adenoviruses in wild primates: a high level of genetic diversity and evidence of zoonotic transmissions. Journal of virology 85(20):10774-10784.

Whiteford L. 1997. The Ethnoecology of Dengue Fever. Medical anthropology quarterly 11(2):202-223.
Whitman WB, Coleman DC, and Wiebe WJ. 1998. Prokaryotes: the unseen majority. Proceedings of the National Academy of Sciences of the United States of America 95(12):6578-6583.
WHO WHO. 2003. Summary table of SARS cases by country, 1 November 2002-7.
Whon TW, Kim MS, Roh SW, Shin NR, Lee HW, and Bae JW. 2012. Metagenomic characterization of airborne viral DNA diversity in the near-surface atmosphere. Journal of virology 86(15):8221-8231.
Wiesner C, Acosta J, Diaz Del Castillo A, and Tovar S. 2012. Social representations of human papillomavirus in Bogota, Colombia. Medical anthropology 31(1):77-92.
Williams Brianna B, Van Benschoten Andrew H, Cimermancic P, Donia Mohamed S, Zimmermann M, Taketani M, Ishihara A, Kashyap Purna C, Fraser James S, and Fischbach Michael A. 2014. Discovery and Characterization of Gut Microbiota Decarboxylases that Can Produce the Neurotransmitter Tryptamine. Cell host & microbe 16(4):495-503.
Williams TA, Foster PG, Cox CJ, and Embley TM. 2013. An archaeal origin of eukaryotes supports only two primary domains of life. Nature 504(7479):231-236.
Willner D, Furlan M, Schmieder R, Grasis JA, Pride DT, Relman DA, Angly FE, McDole T, Mariella RP, Jr., Rohwer F et al. . 2011. Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1:4547-4553.
Wolfe N. 2011. The viral storm : the dawn of a new pandemic age. New York: Times Books. 304 p. p.
Wolfe ND, Dunavan CP, and Diamond J. 2007. Origins of major human infectious diseases. Nature 447(7142):279-283.
Wootton SH, Scheifele DW, Mak A, Petric M, and Skowronski DM. 2006. Detection of human influenza virus in the stool of children. The Pediatric infectious disease journal 25(12):1194-1195.
Wu Z, Ren X, Yang L, Hu Y, Yang J, He G, Zhang J, Dong J, Sun L, Du J et al. . 2012. Virome Analysis for Identification of Novel Mammalian Viruses in Bat Species from Chinese Provinces. Journal of virology.
Wylie KM, Weinstock GM, and Storch GA. 2013. Virome genomics: a tool for defining the human virome. Current opinion in microbiology.
Xuan C, Shamonki JM, Chung A, Dinome ML, Chung M, Sieling PA, and Lee DJ. 2014. Microbial dysbiosis is associated with human breast cancer. PloS one 9(1):e83744.
Yang Y, Du L, Liu C, Wang L, Ma C, Tang J, Baric RS, Jiang S, and Li F. 2014. Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. Proceedings of the National Academy of Sciences 111(34):12516-12521.
Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP et al. . 2012. Human gut microbiome viewed across age and geography. Nature 486(7402):222-227.

Yogo Y, Iida T, Taguchi F, Kitamura T, and Aso Y. 1991. Typing of human polyomavirus JC virus on the basis of restriction fragment length polymorphisms. Journal of clinical microbiology 29(10):2130-2138.
Yogo Y, Sugimoto C, Zhong S, and Homma Y. 2009. Evolution of the BK polyomavirus: epidemiological, anthropological and clinical implications. Reviews in medical virology 19(4):185-199.
Yong E. 2012. A Parasite's Parasites. The Scientist.
Yoon HS, Nakayama T, Reyes-Prieto A, Andersen RA, Boo SM, Ishida K, and Bhattacharya D. 2009. A single origin of the photosynthetic organelle in different Paulinella lineages. BMC evolutionary biology 9:98.
Yoshikawa T, Ihira M, Taguchi H, Yoshida S, and Asano Y. 2005. Analysis of Shedding of 3 β-Herpesviruses in Saliva from Patients with Connective Tissue Diseases. Journal of Infectious Diseases 192(9):1530-1536.
Yu G, Yagi S, Carrion R, Jr., Chen EC, Liu M, Brasky KM, Lanford RE, Kelly KR, Bales KL, Schnurr DP et al. . 2013. Experimental Cross-Species Infection of Common Marmosets by Titi Monkey Adenovirus. PloS one 8(7):e68558.
Yuan JB, Zhang XJ, Liu CZ, Wei JK, Li FH, and Xiang JH. 2013. Horizontally transferred genes in the genome of Pacific white shrimp, Litopenaeus vannamei. BMC evolutionary biology 13:165.
Yutin N, Wolf YI, Raoult D, and Koonin EV. 2009. Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. Virology Journal 6:223.
Zandman-Goddard G, and Shoenfeld Y. 2009. Parasitic infection and autoimmunity. Lupus 18(13):1144-1148.
Zarco MF, Vess TJ, and Ginsburg GS. 2012. The oral microbiome in health and disease and the potential impact on personalized dental medicine. Oral diseases 18(2):109-120.
Zaura E, Keijser BJ, Huse SM, and Crielaard W. 2009. Defining the healthy "core microbiome" of oral microbial communities. BMC microbiology 9:259.
Zhan MEI. 2005. Civet Cats, Fried Grasshoppers, and David Beckham's Pajamas: Unruly Bodies after SARS. American anthropologist 107(1):31-42.
Zhang T, Breitbart M, Lee WH, Run JQ, Wei CL, Soh SW, Hibberd ML, Liu ET, Rohwer F, and Ruan Y. 2006. RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol 4(1):e3.
Zhong S, Randhawa PS, Ikegaya H, Chen Q, Zheng H-Y, Suzuki M, Takeuchi T, Shibuya A, Kitamura T, and Yogo Y. 2009. Distribution patterns of BK polyomavirus (BKV) subtypes and subgroups in American, European and Asian populations suggest co-migration of BKV and the human race. Journal of General Virology 90(1):144-152.
Zimmer C. 2011. A Planet of Viruses. Chicago, IL: University of Chicago Press.
Zimmer C. 2012. Redrawing the Tree of Life. Phenomena: National Geographic.

Appendix

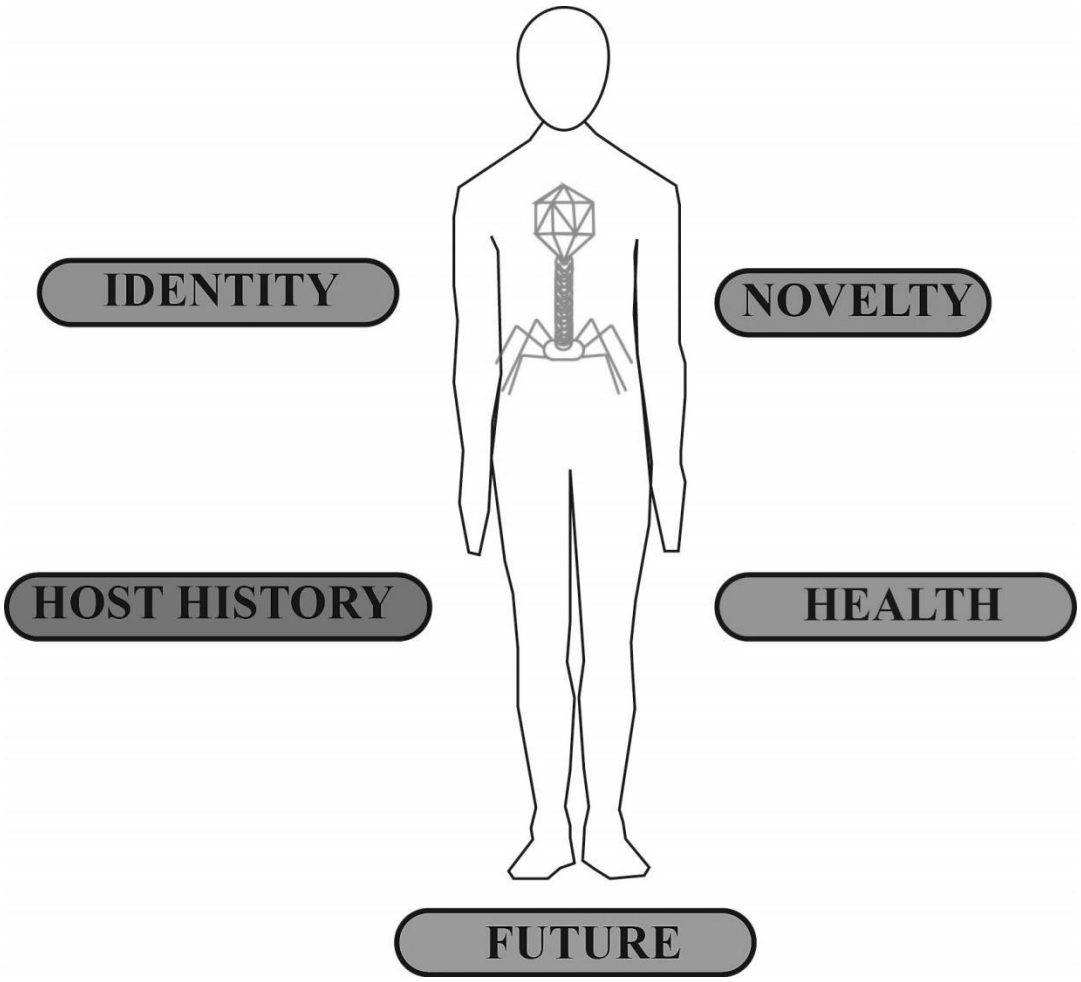

Figure A1: Viral associations to the field of anthropology

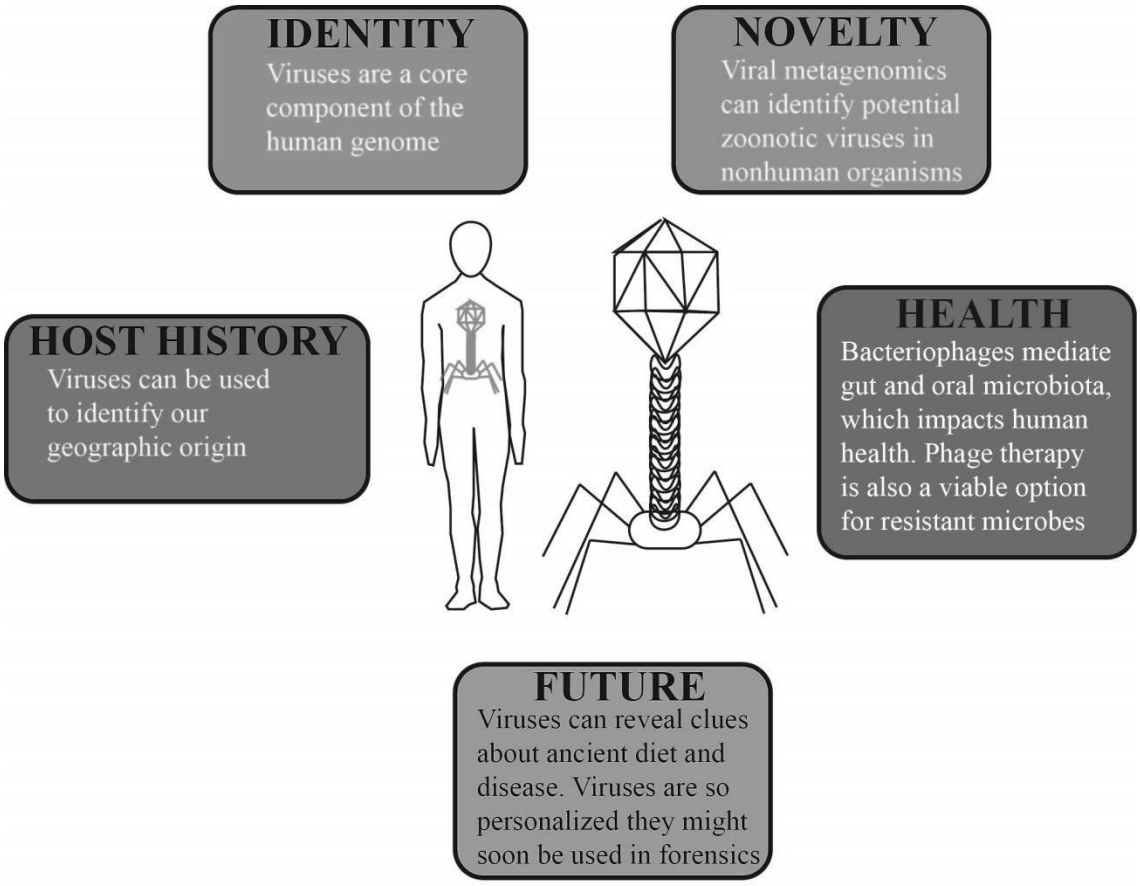

Figure A2: Detailed viral applications and future directions