Spring 5-13-2016

The Canine Genome: Discoveries, Applications, and Future Potential

Rachael Lander
Eastern Kentucky University, rachael_lander@mymail.eku.edu

EASTERN KENTUCKY UNIVERSITY

The Canine Genome: Discoveries, Applications, and Future Potential

Honors Thesis

Submitted

In Partial Fulfillment

of the

Requirements of HON 420

Spring 2016

By

Rachael Marie Lander

Faculty Mentor

Dr. Patrick J. Calie

Department of Biological Sciences

Abstract

The Canine Genome: Discoveries, Applications, and Future Potential

Thesis author: Rachael Marie Lander

Thesis mentor: Dr. Patrick J. Calie

Department of Biological Sciences

A frequent question that educators often encounter is: what is the value of learning? Does knowledge have an inherent value, or should there be an economic benefit? The results of the collaborative efforts to determine the nucleotide sequence of the canine genome were used as a platform to assess these thesis statements. A literature review, practical experience in the laboratory, and interviews with several genome scientists of the contributions of the efforts to determine the genome sequence of the domestic dog, and the biomedical contributions that have been made to both canine and human health demonstrate the inherent value of this scientific objective. The canine genome effort has led to the development of new genome science technologies, new discoveries regarding the cellular basis of many canine diseases and disorders, an understanding of the basis for similar disorders in humans, Most notably, several canine cancers have been shown to be homologous to human cancers, leading to the identification of the genes responsible for such disorders. The results of the canine genome effort provide full support for the unfettered pursuit of knowledge through scientific investigation, and highlight the relationships among different fields of science.

Key words: canine genome, canine evolution, canine and human disease, GWAS, SNPs.

Table of Contents

List of Figures Page

Figure 1. Examples of single nucleotide (point) mutations and their effects on the encoded protein 4

Figure 2. Schematic of the approach used to sequence a genome 8

Figure 3. An illustration of a SNP 9

Figure 4. An illustration of the GWAS approach 10

Figure 6. A GWAS plot of the association of the chromosome 15 SNP with the SCCD in dark poodles 29

List of Illustrations

Examples of canine diversity Frontispiece

Figure 5. Cranial bones of the Altai dog 13

Tables

Table 1. Summary of the major types of cancer associated with specific dog breeds 27

Topics

Overview 1

Introduction to the Canine Genome 2

Methodology 6

Geographic and temporal origin of the domestic dog 10

Evolution of behavior 22

Genetic basis of canine disease 25

Parallels between canine and human diseases 32

The Future 35

Final conclusion 36

Literature cited 37

Appendix 1 -Annotated Bibliography 45

Appendix 2 – Rachael Lander BIO598 Independent Research proposals for Fall 2015 and Spring 2016 53

Appendix 3 - Interviews with genome scientists conducted by Rachael Lander 60

Acknowledgements

Dr. Coleman

Fellow students in the HP and in your classes

Sorority sisters (support) SGA (speaking experience)

Mom and Dad

Mark, Dan, Busby and Elaine

When you insert text, make sure the page below is not "bumped down" I think that as you type the next page stays stable. I inserted a "Section" marker at the end of this page, so what you type here should not move successive pages down. Yes, it is working as I type this. Don't use the "return" key on this page, as it tends to delete the page below it. Insert your text and just remove this as you need to, and the formatting in the following pages should stay OK. But make sure you check first after you do your additions.

Frontispiece

Examples of canine diversity. The above breeds differ in as few as several specific gene variants (alleles) that give them their distinctive physical appearance. Clockwise from the left: the Bloodhound, the Chinese-crested, the Dandie Dinmont terrier, the Scottish deerhound, the long-haired Chihuahua, and the French bulldog. Many of these physical differences among breeds can be attributed to variants (alleles) of a single gene. From Shearin and Ostrander (2010).

Overview

Thesis statement I: The pursuit of knowledge in all forms is of value regardless of the perceived economic impact and the immediate relevance to the human condition.

Thesis statement II: Connections exist among all forms of knowledge, and discoveries in one field can lead to novel insights and can contribute to progress in other related fields.

Proposal: The Canine Genome Project provides an example of the value of research for the sake of gaining knowledge and understanding of a system, and illustrates the connections among different areas of life science, and applications of research relative to mankind .

A current theme in the public discussion of higher education is, what is the value of a liberal arts education? Should students focus on STEM (science, technology, engineering and mathematics) or business disciplines to gain employment after graduation, or should students be encouraged to explore majors in the liberal arts? Should students pursue knowledge for its own sake in college, or should they focus on "marketable skill sets", with an eye on future employment opportunities? How relevant are (at face value) unrelated disciplines to one's chosen field of study? This discussion has become more relevant given the current level of support for higher education in Kentucky and across the nation, and the calling into question the value of a diverse liberal arts education.

A related issue is, should federal research funds be expended on non-human model systems? The National Institutes of Health is the largest provider of funds for

biomedical research in the nation, and they do provide research funds for investigators utilizing non-human research systems such as mice, fruit flies (*Drosophila melanogaster*), roundworms (*Caenorabditis elegans*), yeast (*Saccharomyces cerevisiae*), and plants (thale cress, *Arabidopsis thaliana*). Some question the wisdom of providing funds for research on non-human systems, rather than exclusively on studies on human subjects. I maintain that this attitude is uninformed, and demonstrates a lack of understanding of the nature of scientific investigation and discovery.

My thesis topic, discoveries arising from the canine genome effort, will serve as a platform to address the above issues. In my discussion I will provide the following:

1) an overview of the major discoveries of the canine genome efforts, with regard to the origins of domesticated dogs and the various breeds;
2) a summary of the veterinary medical and medical insights that have arisen from investigations into the canine genome;
3) justification for pursuing biomedical investigations in non-human systems by providing specific examples of connections between canine and human health and disease; and
4) a defense for the pursuit of knowledge for the sake of learning and discovery.

Introduction to the Canine Genome

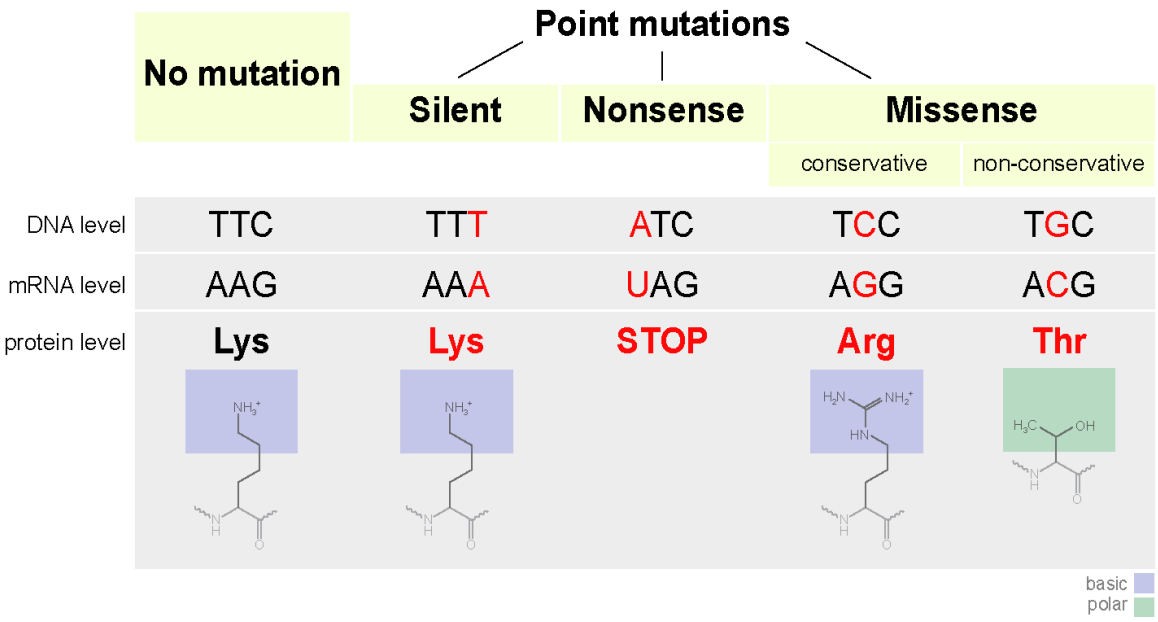
One might first ask "what is the canine genome? To address this question, I will define a genome in a broad sense. A genome is the collective genetic information (genotype) contained within the cells of a particular organism, that specify the structure and function of the biological molecules within that organism,

that in turn specify the physical traits (phenotype) of the organism. A copy of the genome is carried in every cell of the organism, both germ line and somatic. It is an organism's complete set of cellular instructions, composed of DNA. Each genome contains all of the information needed to build and maintain that organism. This information is encoded in the sequence of nucleotides that comprise specific genes, and determines the actual physical traits (phenotype) of the organism. In humans, a copy of the entire genome—more than 3 billion DNA base pairs distributed among 46 chromosomes—is contained in all cells that have a nucleus. The mitochondrion, the site of ATP production in the cell, contains a circular chromosome that encodes approximately 17 mitochondrial proteins and several transfer RNA genes. In canines, the nuclear genome contains 2.8 billion base pairs of DNA distributed among 39 pairs of chromosomes. There are 19,000 protein-coding genes in the canine nuclear genome, most of them with close counterparts (termed orthologs) in other mammals, including humans.

Changes in this cellular information, termed mutations, can occur, sometimes resulting in changes in the physical traits of the organism. We will focus on one type of mutation, single nucleotide substitutions, termed point mutations. Examples of different types of point mutations are shown in Figure 1. Mutations can be caused by various means, for example, ionizing radiation, exposure to mutagens, or infection by viral DNA, or they can occur spontaneously through errors in DNA replication.

Figure 1. Examples of single nucleotide (point) mutations and their effects on the encoded protein. The silent mutations have no effect on the protein; the nonsense and missense can alter the amino acid sequence of the protein.

https://upload.wikimedia.org/wikipedia/commons/6/69/Point_mutations-en.png



Germ line mutations occur in reproductive components such as the egg or sperm and are passed to the offspring. Somatic mutations occur in body cells and are not transmitted to the offspring. Over time mutations can accumulate, leading to differences in the phenotype, in turn leading to the appearance different strains, varieties, breeds, and even species. The phenotype can include both physical and behavioral traits, such as herding, pointing, and running.

The genetic information is encoded in the sequence of nucleotides found in DNA. Any variation in the order of adenine, thymine, cytosine, or guanine provides for varying traits in an organism. Not only does the set of genes in DNA change based on the order of nucleotides, but furthermore the physical expression of these

traits (the phenotype) corresponds to the order of nucleotides. Recent technological advances have allowed for the determination of all the nucleotide sequences of the chromosomes within a particular organism. These advances have led to major discoveries in evolutionary biology (e.g. phylogenetics), biomedical science (e.g. human genetic disorders), and even in ecology (e.g. microbial community structure).

But why the canine genome? What practical use could this have for understanding human health? What advantages did the canine genome offer as an experimental system?

The selective breeding (and inbreeding) of purebred dogs has led to a predominance of certain disorders in some lines, which will facilitate the discovery of those genes responsible. Many canine disorders have counterparts in humans, so studying the cause of these disorders in dogs could lead to improved detection methods in humans (Sutter and Ostrander, 2004). The inbreeding involved with developing certain dog breeds has also led to a loss of random genetic (nucleotide sequence) variation, so over time the genomes among different individuals of the same bred have become quite similar in sequence. When comparing groups of individuals for specific mutations, the lack of general sequence diversity facilitates the discovery of those mutations responsible for specific disorders (Ostrander and Wayne, 2005).

Dogs are the most diverse group of land mammals, exhibiting a wide range of traits and phenotypes. For a number of traits the actual genes responsible are being identified and they often turn out to be few in number. For example, a large genomic region has been identified that is responsible for canine skeletal size, one

gene being responsible is the insulin-like growth factor gene (IGF-1). Three genes are responsible for all the variation in coat color and texture among all dog breeds (Shearin and Ostrander, 2010).

I will now discuss the general methods used to determine the nucleotide sequence of the canine genome, the general ideas of the origin of the domesticated dog, highlights of discoveries of dog diseases and the counterparts in humans, and the basis for differences in behavior between wolves and dogs.

Methodology

Three technologies associated with the canine genome will be discussed, as these have been utilized in many of the papers reviewed in this thesis. These technologies are genome sequencing, detection of single nucleotide polymorphisms (SNPs), and Genome Wide Association Studies (GWAS).

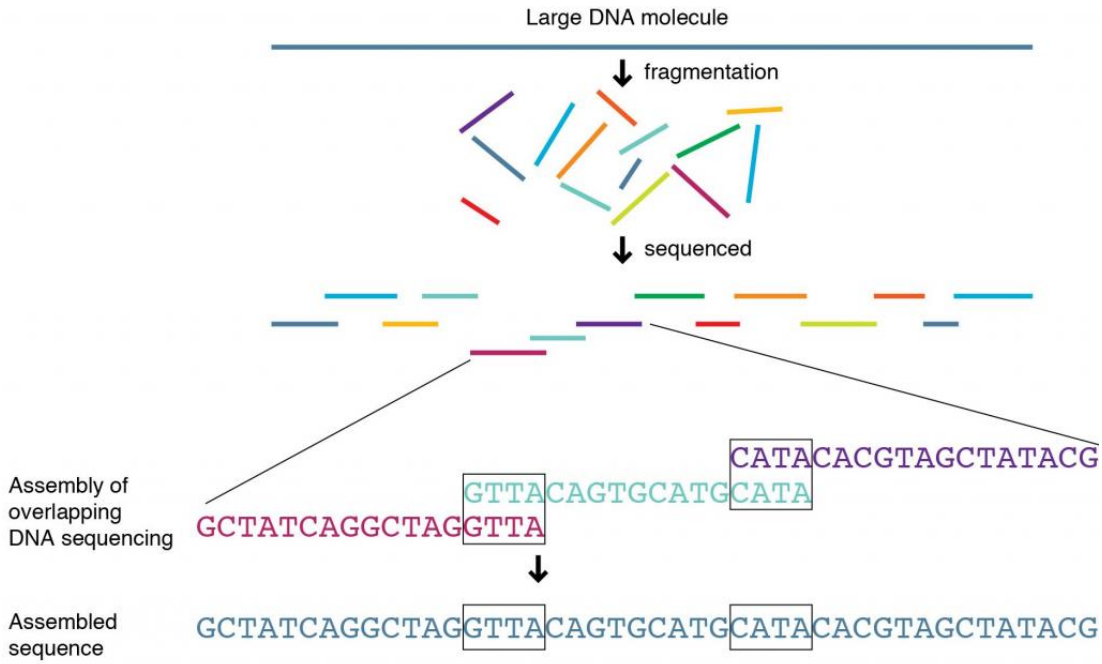
Genome sequencing

There are several platforms, or instruments, that are capable of generating the nucleotide sequence of a genome, utilizing different technical approaches. But all rely on the following basic strategy for mammalian genomes.

Leukocytes (white blood cells) are the usual cells collected, as they can be collected from blood samples with minimal harm to the animal. The cells are broken open by detergents in an aqueous (water) solution. The other biomolecules (proteins, lipids, carbohydrates) are removed from the sample by either chemical or enzymatic means, leaving DNA as the only intact molecule. The DNA is further purified through chemical approaches, examined through gel electrophoresis for quality and purity, then mechanically sheared into random fragments of a specific

size, typically between 500 and 1,000 nucleotides in length. The current technology is limited to sequencing short DNA fragments of this length. The DNA sample is loaded onto a DNA sequencing machine, and the unit will then generate the sequences (through chemical synthesis) of the separate fragments. The sequences of the separate fragments (these can number upwards of 50 -100 million separate sequences, termed "reads") are then input into a computer, and a software program then directs the matching up of identical overlapping ends of the fragments. This is done in a successive manner, adding fragments to each end one at a time. Eventually the entire set of chromosomes is rebuilt from the small collection of fragments, as shown in figure 1. The computer also reads the sequences of nucleotides on the chromosome and determines where specific genes are located, and which proteins are encoded by each gene. In comparative genomics, one can compare, side by side, the order and arrangement of genes on the chromosomes among different species. Closely related species will have highly similar gene order and arrangements; more distantly related species have more differences in their genes, both in sequence and in order. Thus one can either examine separate genes among individuals, or entire genomes.

Figure 2. Schematic of the approach used to sequence a genome. A collection of large DNA molecules, such as complete chromosomes, are randomly fragmented in to small pieces. The individual pieces are sequenced, and are assembled into larger pieces through computer analysis. This strategy is continued until the final complete chromosome has been fully reassembled. Figure courtesy of the National Human Genome Research Institute of the National Institutes of Health.



SNPs

Single nucleotide polymorphisms (SNPs) are mutations in the genome that involve single nucleotides (Figure 2). They are detected by aligning the sequences of two genomes and determining those positions at which the nucleotides differ between two individuals. The SNPs can serve as genetic markers, as they can be detected through DNA sequencing.

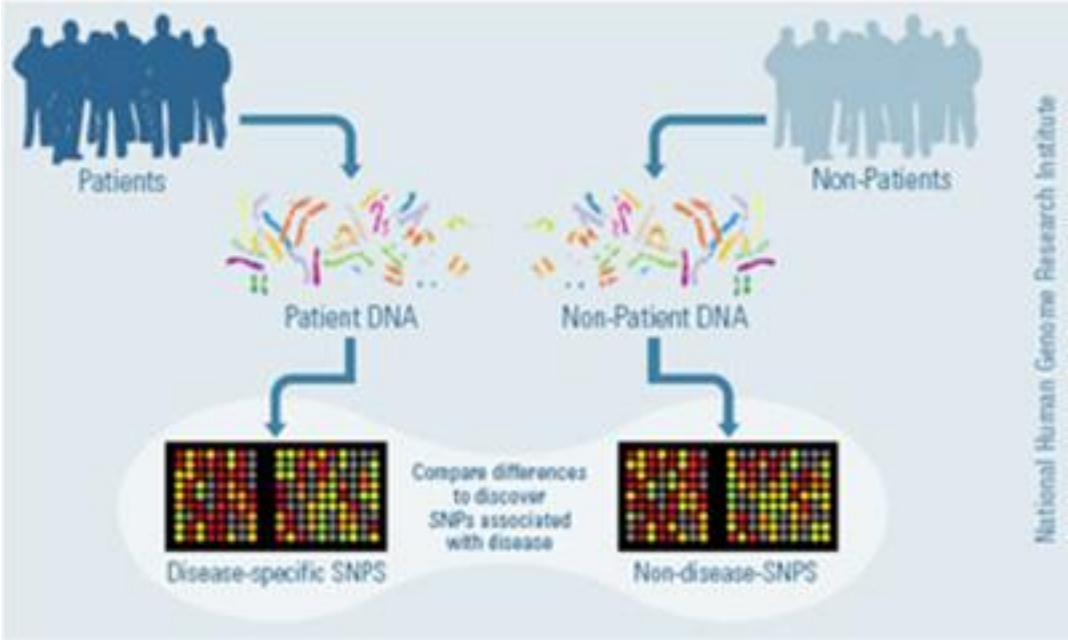
Figure 3. An illustration of a SNP. The nucleotides above and beneath the arrowheads are the SNPs in these two sets of sequences. Image courtesy of the Cregan Lab of the U.S.D.A.

http://bldg6.arsusda.gov/pberkum/Public/sarl/cregan/SNP.gif

GWAS

If a specific SNP is associated with a specific disorder or disease, the SNP can serve as a genetic marker for that disorder. The association of a series of SNPs with such a disorder is determined through an approach called the Genome Wide Association Study, or GWAS (Noorgard, 2008). In this approach one examines, throughout the genome, the SNPs found in two populations – one affected by a disorder, and one unaffected. By removing all the SNPs that are found in both samples, and examining the SNPs unique to the affected population, one could detect SNPs that are diagnostic for a particular trait, or disease (Figure 4).

Figure 4. An illustration of the GWAS approach. SNPs from both affected and unaffected individuals are compared in parallel. Those SNPs that are unique to the affected pool of individuals are examined further as candidates for possible association with the disease or disorder under question. Image from the National Human Genome Research Institute of the National Institutes of Health.



I will now begin a discussion of several of the major sets of discoveries made through canine genomics and genetics.

Geographic and temporal origin of the domestic dog

The geographic place of origin and the precise time in history for domestication of the modern dog has been a topic of investigation and disagreement for some time (Wayne, 1993; Wayne and Ostrander, 1999). We often think that science will provide one answer, but the reality is that there are often multiple answers for a particular question from different investigators. A number of explanations have been proposed based on various methods, and these explanations

sometimes do not agree. There are several reasons for this. One, the specific molecular system chosen in the study (e.g. autosomal, or non-sex chromosome, Y chromosome, mitochondrial DNA) could have different mutation rates, or different evolutionary histories, thus giving different stories. Second, the methods of phylogenetic and statistical analysis of morphological and genetic data often differ among investigators, which could lead to different conclusions. And third, the specific individuals or populations sampled might not represent the most informative or appropriate group to examine for the question being asked. These are issues with many efforts in evolutionary biology, and are not unique to the quest to determine the origin and timing of the domestication of modern dogs.

Two approaches to address this topic of the origin and time of domestication of the dog have been taken. The first was archeological, and the second genetic in nature. The archeological approach involved analyzing ancient human habitations for evidence of the presence of dogs, often in the form of canine bones or canine teeth marks on other animal bones, or examining actual fossil remains of canine ancestors. The inference was made that dogs that existed with humans were likely domesticated. The second approach involved genetics. Genetic approaches first involved either mitochondrial DNA sequences, or sequences from specific chromosomes. Advances in genome technology led to the utilization of the entire canine genome, or major portions of the canine genome, to address the question of domestication.

One conflict in the data summarized below is that some genetic evidence suggests the origin of canine domestication in central Asia approximately 15,000

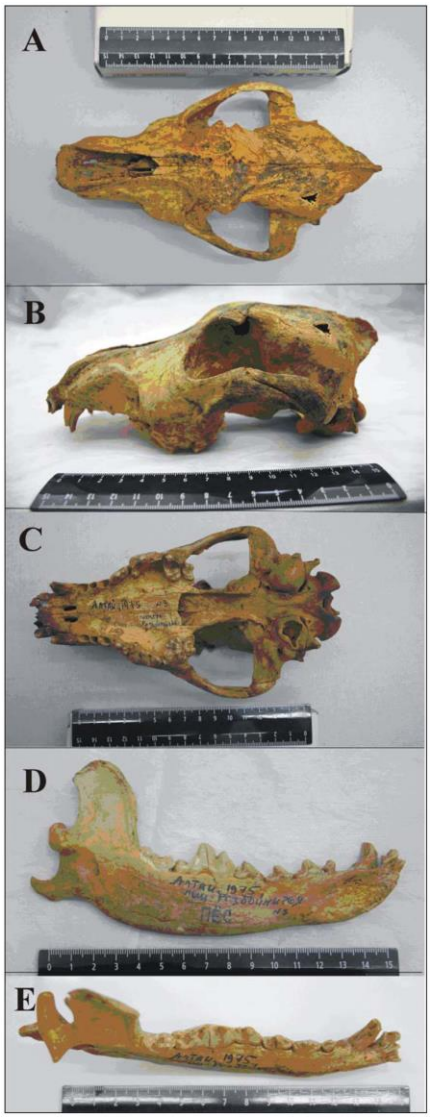
years ago, but the oldest dog-like fossils have been found in Europe and Siberia, and date to over 30,000 years ago (Ovodov et al., 2012; summarized in Thalman et al, 2013). The evidence for these two positions will now be presented.

Fossil evidence for ancient dog origins

The most reliable evidence for the origin and age of a particular group (often termed a "lineage") is the fossil record, that provides direct physical evidence, both geographic and temporal. Among the oldest remains found of dogs, or a canine ancestor, come from the Razboinichya Cave in the Altai Mountains of southern Siberia. Due to the quality of the preservation of the skull (Figure 5) morphological comparisons could be made to Pleistocene-era wolves, modern wolves, and prehistoric and modern domesticated dog lineages.

Radiocarbon dating, a means of determining the age of a biological specimen through radioactive decay of an isotope of carbon (the ^{14}C isotope to the ^{12}C isotope), dates the Altai remains at approximately 33,000 YBP (Years Before Present). Based upon measurements of the teeth, the Altai dog is most closely related to modern dogs from Greenland, a lineage that is about 1,000 years old, but is unlike ancient or modern wolves, or the canid found from the Eliseevichi I site in Russia (discussed below). This result supports the idea that there could have been multiple locations where dogs were first domesticated, an issue that will be explored later in this thesis. This is the oldest fossil of a dog yet discovered, and establishes the earliest possible date for the origin of a possible domestic dog.

Figure 5. Cranial bones of the Altai dog. A) aerial view, B) profile, C) palate, D) left mandible, E) left lower tooth row (scale on ruler in cm). (from Ovodov et al, 2012).



Archeological and genetic evidence – European origin of domestication

Among the earliest archeological evidence for domesticated dogs comes from Israel. The remains of two ancient dogs were found with three sets of human remains in an ancient burial site in northern Israel in a region known as the Hayonim Terrace. Archeological techniques (examination of human artifacts) date these sites as approximately 11,000 YBP (Tchernov and Valla, 1997). As in the case

of the Altai dog, the morphology of the skulls of these dog fossils is sufficiently different from wolves to consider them as true domesticated dogs. A report of a puppy being found buried with a human at this same location, and differences in the teeth of the canid remains provides additional support for dog domestication in this region (Davis and Valla, 1978). Prehistoric humans typically did not associate with wolves, due to competition for food sources between these two species. Further evidence for the existence of domesticated dogs from this time comes from central Russia, where dog remains dating from 13,000 – 17,000 YBP have been found (Sablin and Khlopachev, 2002). These authors also used radiocarbon dating to establish the age of these canine remains.

Paleolithic dog remains were found in Belgium and from two locations in the Ukraine, and using stable isotopes, dated to approximately 31,700 YBP. Mitochondrial DNA was isolated from the well-preserved bone marrow of these animals, and then sequenced. A comparison was done between these ancient samples and mtDNA sequences of modern dogs and modern wolves. The nucleotide sequence comparisons indicated these animals were genetically distinct from wolves, and likely represent early examples of dog domestication in Europe (Gemonpre et al., 2009).

These first genetic studies used partial genome sequences often single genes or portions of chromosomes. Mitochondrial DNA was the first molecule utilized, due to numerous copies per cell, its small size, and the ability to isolate it from the bone marrow of fossil bones in a form suitable for DNA sequencing. As the technology was developed, later efforts utilized nuclear genome sequences. In contrast to mtDNA, there are only two copies of each nuclear genome per cell.

The “control region” or “D-loop” is the origin of replication for the mtDNA molecule (the site where the DNA polymerase initiates replication). It was widely used for phylogenetic analysis among animals, due to the availability of primers for PCR amplification and DNA sequencing that worked across a wide range of animals. This region was examined from 162 wolf individuals from 27 geographic locations worldwide, and from 140 dogs representing 67 breeds. In this study (Vila et al., 1997) it was determined that, due to sequence similarity in this genomic region from canines and wolves, modern dogs evolved from wolves, and that dogs originated as a separate lineage (or group) from wolves approximately 100,000 YBP. As will be soon seen, this was an overestimate, due to improper estimation of the mutation rate in this region among different canids.

Thalman et al. (2013) isolated mtDNA from prehistoric fossil dog bones (18 individual dogs) from sites in Eurasia and North America. Rather than just the control region, they sequenced the entire mitochondrial genome and compared sequences of these ancient animals to the sequences of a set of modern dogs and different extant wolf individuals. The results indicated that modern dogs are most closely related to the ancient and modern European canids. By using a process called “molecular dating”, in which the rates of mutation are calculated for specific genomes, and a type of “clock” is then calibrated to determine the age of specific individuals, the origin of domestication of dogs in Europe is estimated at between 18,800 and 32,200 YBP, perhaps when ancient human hunter-gatherers began to first attempt to domesticate dogs for their use. An alternative hypothesis is that dogs were first domesticated by early practitioners of agriculture, which has

support from the Middle Eastern archeological results (Tchernov and Valla, 1997) and from a study of East Asian dogs (Pang et al., 2009).

Genetic evidence – Middle Eastern origin of domestication

In a more in-depth study, a survey of over 48,000 SNPs in a GWAS involving several dog breeds and several gray wolf individuals revealed a closer genetic relationship with Middle Eastern than with Asian gray wolves. This indicates that the domestic dog likely arose from an ancestral species in the Middle East. Through this effort, researchers found that dog breeds share a higher proportion of multi-locus haplotypes (conserved groups of genes) unique to grey wolves from the Middle East rather than wolves from East Asia. This finding is further supported by mtDNA sequence data (vonHoldt et al., 2010).

Genetic evidence – Asian origin of domestication

Some prior efforts (Vila et al., 1997) used the origin of replication (also called the "control region") of the mtDNA, which often does not provide adequate resolution of the evolutionary history of the group under examination due to a limited sequence data set. There were also problems with sampling too few individuals in prior studies. To address these issues, Pang et al. (2009) sequenced the entire mitochondrial genome (including the D-loop, a region on the mt genome) of 168 dogs, and the control region of 1,543 dogs from Old World populations. They found that the greatest level of genetic diversity was in southern China south of the Yangtzee River, the diversity decreasing across Eurasia. They concluded that the domestic dog originated in southern China approximately 16,300 years ago. This is

the approximate time of the origin of rice cultivation, and suggests that perhaps the early domestic dog was bred by rice farmers.

Likewise, another investigation of domestic dog origin was conducted utilizing whole genome sequences from 58 canids including 12 gray wolves, 27 primitive dogs from Asia and Africa, and 19 diverse breeds from across the globe. It was found that dogs from southern East Asia have higher degrees of genetic diversity compared to other populations of modern dogs. From these ancient Asian populations it is hypothesized that a subset of dogs started migrating to the Middle East, Africa, and Europe around 15,000 years ago, eventually evolving into our modern breeds. This study was helpful in understanding how dogs traveled across the globe to form populations in various locations different from the place of origin (Wang et al., 2015).

Village dogs are semi-feral animals that are free-ranging, breed freely, and contain genotypes that are more reflective of the local ancestral canine populations than purebred breeds. Using genomic data from specific autosomal chromosomes, the Y chromosome, and mtDNA, from 5,392 dogs from 161 breeds, including a population of 549 village dogs from 38 countries, Shannon et al. (2015) found strong evidence that dog domestication occurred in Central Asia, perhaps near Nepal and Mongolia. Furthermore, dogs in close proximity to this region (e.g. East Asia, India, and Southwest Asia) contain high levels of genetic diversity. This may also be due to the large population size associated with this region. Some Asian populations exhibit varying degrees of mixture with European populations; those from the Neotropics and the South Pacific are almost completely derived from

European populations, whereas those from Vietnam, India, and Egypt show little to no evidence of genetic mixing (interbreeding) with European populations.

In a supporting study Wang et al. (2015) employed complete genome sequences from 12 gray wolves, 27 "primitive" dogs from Asia and Africa, and a set of 19 breeds from across the world. They determined that the Southeast Asian dogs had higher genetic diversity than the other populations. The SE Asian dogs are, in an evolutionary sense, he closest group to modern gray wolves, indicating an ancient origin of domestic dogs in southern East Asia approximately 33,000 YBP. They estimated that approximately 15,000 YBP, a subset of ancestral dogs started migrating to the Middle East, Africa and Europe, arriving in Europe at about 10,000 YBP. These migrants could be responsible for the fossil remains in the Middle East, as previously noted.

Savolianen et al. (2002) examined mtDNA sequences from 654 domestic dog breeds representing the major worldwide dog populations. The East Asian populations had a higher level of genetic variation than other populations, and the pattern of genetic variation among the different geographic populations suggested an East Asian origin of the domestic dog, dating from approximately 15,000 years ago.

The results of this study, however, have been called into question. One criticism is that Savolianen et al. (2009) did not distinguish between two possibilities: that either the East Asian village dog populations they sampled either represent distinct, indigenous populations, or simply mixtures of different domestic breeds. Either possibility could lead to high genetic diversity. To address this issue

Boyko et al. (2009) sampled 318 village dogs from 7 regions in Egypt, Uganda, and Namibia, using as genetic markers the mtDNA D-loop, 300 SNPs, and 89 microsatellite markers (regions of repetitive DNA in the nuclear genome). They also sampled African breeds (Afghan hounds, Basenjis, Pharaoh hounds, Rhodesian ridgebacks, and Salukis), Puerto Rican street dogs, and mixed breed dogs from the United States. The level of genetic diversity within the mtDNA haplotypes was equal between the East Asian and African populations, calling into question the hypothesis regarding the East Asian origin of the domestic dog.

The exclusive use of nuclear genome data (as opposed to mitochondrial genome data) provided a different picture of canine domestication. Two lines of evidence supported the European origin of dog domestication in a study by Wayne and vonHoldt (2012). First, certain sets, or blocks, of genes termed "haplotypes", are more highly conserved between domestic dogs and European wolves, and are less identical to Asian wolves. Second, the sequence of a nuclear gene, IGF1, involved in the artificial selection of small breeds, has greater sequence identity to the gene in European than in Asian wolves.

To add to the confusion of origin Larson et al. (2012) analyzed 49,024 SNPs from 1,375 individuals representing 35 different breeds and 19 wolves. They also combined their data with that published by other researchers to obtain a larger data set. Their final data set contained the genetic profiles of 121 different breeds. This was compared with the archeological findings of domestic canine remains. Fourteen breeds of dog, among them the Shar Pei, Shiba Inu, Chow, Akita, Basenji, Siberian husky, Alaskan malamute, Afghan hound, Saluki, Pharoah hound, Ibizan hound, and

Norwegian Elkhound, were determined to be closest to the ancestral lineage and were termed "ancient" breeds. The current known origins of these "ancient" breeds were mapped on the prior archeological records of domestic dog origin, and surprisingly there was no correlation between the origin and current distribution of these ancient breeds and the archeological records of domesticated dogs. Second, three of the ancient breeds (Basenjis, Dingoes, and New Guinea singing dogs) are found in regions outside the natural range of the modern wolf, *Canis lupus*, and the range where dogs were introduced approximately 10,000 years after domestication. The conclusion of the paper was that modern dogs and their genomes contribute little to the understanding of the timing and origin of the domestic dog. This is due to the interbreeding among different groups of dogs throughout their evolutionary history, the likely interbreeding among wolves and partially domesticated dogs in prehistoric times, and the movement of dog groups by humans through different geographic regions.

Multiple origins of domesticated dogs

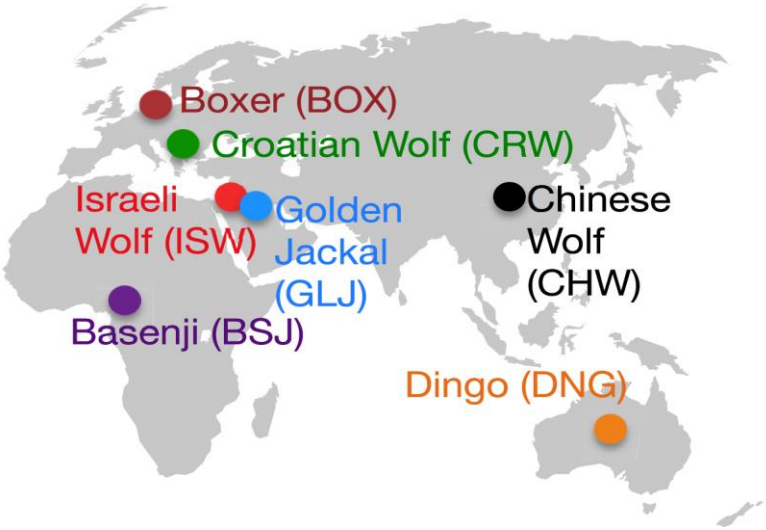
It is likely there are multiple points of origin of the domestic dog, several independent geographic locations. The Razboinichya Cave specimen appears to represent a lineage that went extinct soon after the last glaciation event (approximately 24,500 YBP). The earliest fossil dog remains from Western Europe (Goyet, Belguim) and Siberia (Razboinichya) are separated by thousands of kilometers, which suggests that domesticated dogs arose independently in several locations over time. This is in contrast to the view that there was a single place of origin, with later migration into other geographic areas (Ovodov et al., 2012).

Additional evidence in support of multiple geographic origins of domestic dogs comes from the sequenced genome of the Siberian Taymyr wolf, that lived approximately 34,900 YBP. The genome sequence was used to recalibrate the "molecular clock" used to previously establish the age of dog and wolf lineages, and also led to the conclusion that many higher-altitude dog breeds, such as those in northern Siberia and Greenland, owe their origins in part to extinct Siberian wolf-like ancestors (Skoglund et al., 2015).

Reasons for caution

Freedman et al. (2014) obtained genome sequences from three gray wolves, one from each of the proposed centers of dog domestication (Europe, the Middle East, and Asia), two basal (nearest to the ancestor) dog lineages, the Basenji and the Dingo, and a golden jackal as a comparison group (Figure 2). They determined that the domestic dog arose between 11,000 -16,000 YBP, before the rise of agriculture. Early human hunter-gathers were the most likely to have domesticated dogs, rather than early agriculturists. Their study indicates that the determination of the origin of the domestic dog is unclear, due to past genetic mixing (interbreeding) between different dog and wolf populations. If there is continuous breeding (exchange of genetic information) among different groups, the genetic markers that can be used to track evolutionary history will either become altered through a process known as genetic recombination, or will be shared among different groups, giving a false indication of genetic and evolutionary relationships.

Figure 6. Map of several of the sampled breeds from across the world in one study of dog domestication (Freedman et al., 2014).



The take home message is simple: different data sets, different groups sampled, different sample sizes, and different approaches to analysis of those data sets, can yield different results. Past events, such as interbreeding among different populations, can result in a mixing of genome sequences that can complicate results and interpretations. Science can give different answers to the same question, and we need to appreciate the fact that sometimes the truth is not easily obtained through a single study, but could take years of effort.

Evolution of behavior

We often think of domestication as involving selection of physical traits, such as coat length and color, or size, or speed. But during the domestication of the dog there has been selection by their human handlers for specific behavioral traits, such as obedience and docility. The biological basis for these traits has been determined for some behaviors, including those that separate wolves from domestic dogs. Remarkably, many of these changes involve single genes, and the level of activity

(termed expression) in specific areas of the canine brain. Examples of these will now be discussed.

Li et al. (2013) compared brain gene expression between a set of Chinese native dogs and wolves, and found differences in the activity of different brain regions. Specifically, genes associated with "negative defensive behavior" and aggressiveness were more active in the wolves than in the Chinese dogs, in which genes associated with "positive reactions", such as docility, were more highly expressed. The dogs had higher gene expression activity in the prefrontal cortex, a region that is associated with decision making, than the wolves. German Shepherd dogs had similar levels of gene expression activity as the Chinese dogs, indicating a higher level of function in these regions. The authors conclude that behavioral changes, and the associated neurological changes, were selected very early in the dog domestication process, perhaps earlier than other physical traits such as coat color and size.

Another study examined gene expression in three regions of the brain, the hypothalamus (associated with primitive responses and reflex actions), the amygdala (involved in emotional behavior), and frontal cortex (associated with decision making processes) among dogs, wolves, and coyotes. Between dogs and coyotes the hypothalamus had similar gene expression profiles, but these differed in dogs. Two neuropeptides were identified, *NPY* and *CALCB*, that have been implicated in other animals to be involved with energy control and feeding behavior. These two proteins could also play a role in anxiety and depression. The authors conclude that changes in the gene expression patterns in the hypothalamus

could have been a major factor in the domestication of dogs from wolf-like ancestors, and that these changes could have been accelerated by artificial selection (Saetre et al., 2004).

Modern efforts at domestication of foxes have also provided insights into the neurological changes associated with domestication. On a farm in Russia a population of silver foxes has been maintained for over 40 generations. These animals are non-aggressive, responsive to humans in a manner similar to dogs, and are able to interact with people. Differences in gene expression patterns were observed across different regions of the brains of these animals (Lindberg et al., 2005).

Behavioral differences between wolves and dogs were the focus of another study. A set of tasks was set up, with humans interacting with both wolves and dogs. The humans would provide cues to the animals, for example, finding hidden food, and the responses of each animal were recorded. It was observed that the dogs responded better to the human cues, such as pointing, and the dogs would look at the human handlers, but the wolves would not. It was concluded that the interaction between dogs and humans is the result of an evolutionary process, and cannot be achieved by wolves, even after attempts at training them (Miklosi et al., 2003). However, a later study of this situation came to a different conclusion. Udell et al. (2008) raised wolves in a domesticated environment in contact with humans. They determined that captive wolves would respond to human gestures in finding hidden objects, such as food. They attribute the different results in their study from those from Miklosi et al. (2003) to the experimental design, specifically the testing

environment (such as the use of a fence to separate the wolves, which was not present in the dog testing). The authors in this second study conclude that domestication is not required for canids to interact with humans, and suggest additional study is needed.

A review by Spady and Ostrander (2008) noted the association of many behaviors (such as pointing, herding, retrieving) with specific breeds, and speculated that many of these breed-specific behaviors have a genetic basis. Furthermore, these genetic traits were selected for during the development of these breeds though artificial selection. They point out that the specific genes responsible for specific behaviors have yet to be identified, but by comparing different breeds with different behaviors (such as greyhounds and Australian shepherd dogs) through GWAS or comparative genomics, the genes responsible for such behaviors might be identified.

Genetic basis of canine disease

Probably the biggest breakthrough in mapping canine diseases and traits has been through Genome Wide Association Studies (GWAS). The first major effort utilized approximately 27,000 SNPs and 40 dogs. Traits inherited in a mendelian fashion were readily assigned to specific chromosome regions, opening up an efficient approach for further mapping efforts (Karlsson et al., 2007).

There are a variety of diseases, disorders, and conditions that are exhibited among canines and are sometimes even associated with particular breeds. For example, canine compulsive disorders (CCDs) are repetitive behaviors that cause distress. Such disorders can arise from normal practices such as grooming, tail

chasing, blanket sucking, pacing or circling, among others. This condition has parallels with some obsessive-compulsive disorders in humans. In order to better understand the genetic reasoning behind CCDs, a GWAS were conducted using DNA from 92 affected Doberman pinschers and 68 unaffected animals. Out of 14,700 possible SNPs, only three SNPs located on canine chromosome 7 displayed significant correlation with the behavioral condition. The highly significant association of CCD with the *CDH2* region on chromosome 7 is the first genetic locus identified with any animal compulsive disorder (Dodman et al., 2010).

Another disorder found among certain dog breeds is canine hip dysplasia (CHD) and osteoarthritis. In order to narrow down the genetic basis responsible for this condition, researchers sampled 1,551 dogs by radiographically measuring their hip confirmations. The candidate *FBN2* gene, encoding the fibrillin protein that is associated with tendons, was sequenced from the DNA of 21 Labrador Retrievers and 2 greyhounds. Intron 30 (a non-coding region) of the *FBN2* gene was sequenced in 90 additional Labrador Retrievers and 143 unaffected dogs of 6 other breeds. It was found that Labrador Retrievers homozygous for a 10 base pair deletion (loss of bases) in intron 30 of *FBN2* had significantly worse case of CHD. Among 143 dogs of 6 other breeds, those homozygous for the same deletion has significantly worse CHD (as determined by clinical examinations). In general, the *FBN2* exon 30 (a coding region) 10 base pair deletion was associated with CHD, and could serve as a diagnostic marker for this disorder (Friedenberg et al., 2011).

The dog is an attractive system for study of cancer genetics for a variety of reasons such as their clinical presentations similar to human disorders. Table 1 presents a summary of different cancer types that are prevalent in certain breeds .

Table 1. Summary of the major types of cancer associated with specific dog breeds. From Shearin and Ostrander, 2010.

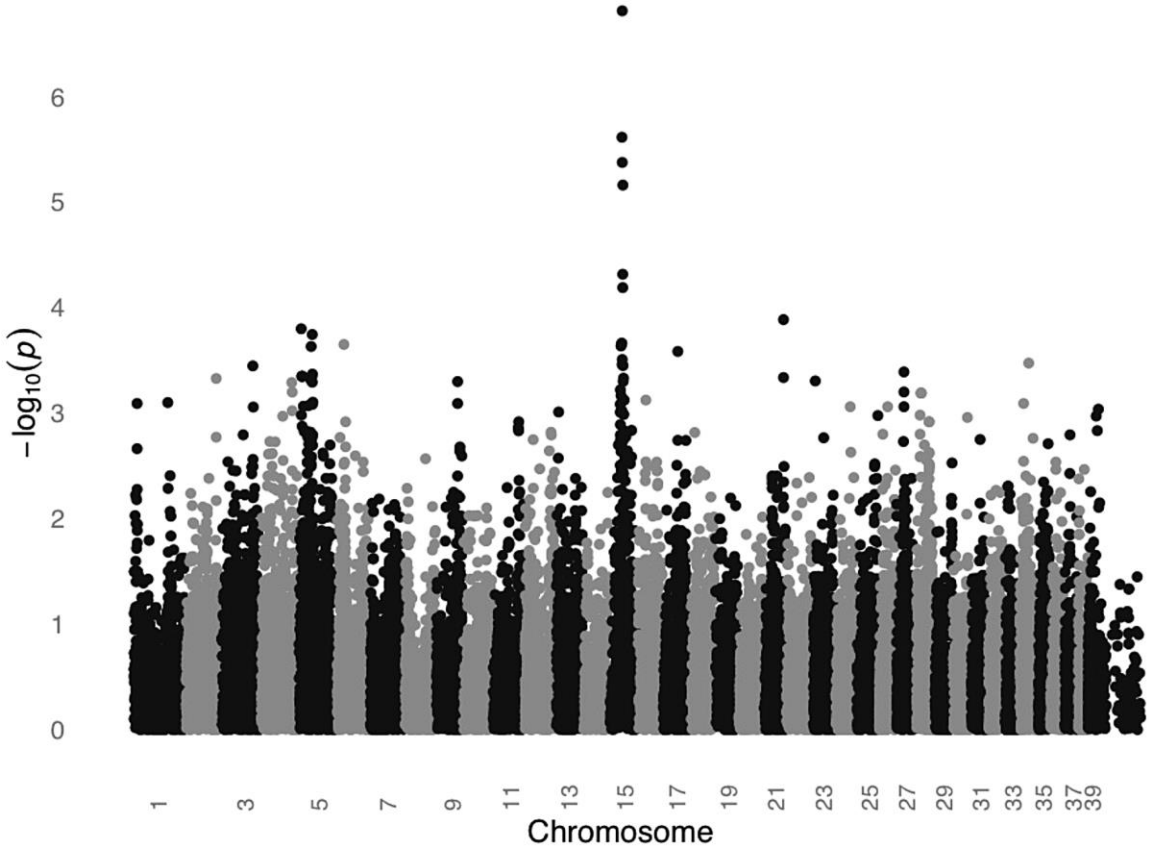
Disorder or disease	Dog breed
Gastric carcinoma	Chow, Belgian Shepherd
Hemangiosarcoma	Golden retriever, Boxer, German shepherd
Lymphoma	Boxer, Golden retriever
Malignant hisiocytosis/ Histolytic sarcoma	Bernese mountain dog, Rottweiler, Flat-coated retriever, Golden retriever
Mammary carcinoma	Doberman pinscher, English springer spaniel, Dachshund, Pointer
Mast cell tumor	Boxer, Boston terrier
Melanoma	Chow, Scottish terrier, schnauzer, Irish setter, Golden retriever, Doberman pinscher
Subungual malignant melanoma	Scottish terrier, Schnauzer, Irish setter, Rottweiler, Golden retriever
Osteosarccoma	Great Dane, Saint Bernard, German Shepherd, Irish setter, Rottweiler, Boxer, Greyhound, Scottish deerhound
Pancreatic carcinoma	Aiaredale, Boxer
Squamous cell carcinoma	Keeshond, schnauzer, Basset hound, Collie
Transitional cell carcinoma	Scottish terrier, West Highland white terrier, Shetland sheepdog, Beagle

Histiocytic sarcoma is one condition that is poorly understood among humans, but occurs in 15-25% of Bernese Mountain Dogs and therefore presents a valuable model for study. In order to study this condition among Bernese Mountain

Dogs, genomic DNA was collected from both affected and unaffected populations of BMD. Through GWAS efforts, researchers were able to identify the cancer-associated loci. In fact, through fine mapping efforts, the location was narrowed down to a single gene region. A single haplotype (a conserved cluster of nucleotides) spanning the *MTAP* gene and part of the *CDKN2A* gene is present among 96% of affected BMD, identifying this region as the possible cause of this cancer (Shearin et al., 2012).

Another cancer found among dogs, especially among Standard Poodles, is squamous cell carcinoma of the digit (SCCD). SCCD is an aggressive cancer that causes bone lesions, sometimes with multiple toe recurrence. This condition has been found to be associated almost entirely in dark coat color individuals while light colored Standard Poodles are rarely at risk. In order to identify the cause for this, a GWAS was performed comparing 31 SCCD cases to 34 unrelated black Standard Poodles. A diagnostic SNP for this disease was located on canine chromosome 15 (Figure 6).

Figure 6. A GWAS plot of the association of the chromosome 15 SNP with the SCCD in dark poodles. The "peak" of dots in the chromosome 15 column indicates a strong association of that SNP with the disorder. Figure from Karyadi et al. (2012).



Through additional mapping, the *KIT Ligand* locus was determined to be the region of interest. The locus was further narrowed down to a 144.9-Kb region after comparison of Standard Poodle cases to other at-risk breeds. It was also found that only the *MC1R* locus, which controls the coat color trait, was significantly different between datasets of black and light colored Standard Poodles. The mutation within

the *MC1R* locus is likely the reason light colored Standard Poodles are protected from this disease (Karyadi et al., 2013).

Another cancer, hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis (RCND) was mapped to canine chromosome 5. A mutation in exon 7 appears to be responsible for this cancer, which has a disease counterpart in humans (Lingass et al., 2003).

Canine transmissible venereal tumor (CTVT) is a cancer that is transmitted through the sexual transfer of malignant cells between animals. In order to better understand the mechanisms behind transmission of this cancer, a catalog of canine genome-wide variation was created and two CTVT genome sequences were compared. The analysis showed that, over time, CTVT has undergone evolutionary adaptation to the host (canine) niche. Mutations have occurred that provide the virus with the ability to avoid the host immune system, specifically with regard to antigen presentation and cell death. Thus, this effort provides the first insights into the specific genetic mutations that contribute to the persistence of this cancer in canids worldwide (Decker et al., 2015).

Another approach to mapping canine disease-causing genes is to examine multiple breeds that exhibit the same disorder in a comparative manner. Parker et al. (2007) divided 132 breeds of dog into five primary breed groups. They then used this comparative approach to fine-map the Collie eye anomaly (cea), a complex disorder of ocular development that was initially mapped to a 3.9 centiMorgan region on canine chromosome 37. The candidate gene region was then narrowed down to a 103-kb (kilobase) interval spanning four genes. Sequence analysis

revealed that all affected dogs share a deletion of 7.8 kb in the NHEJ1 (nonhomologous end joining factor 1) gene. The deletion is in an intron, spanning a highly conserved binding domain to which several developmentally important proteins bind, involved in RNA processing. This work establishes that the primary cea mutation arose as a single disease allele in a common ancestor of herding breeds and has been passed down through successive generations and breeds.

New genes can arise through a process called retrotransposition, in which a DNA-copy of a processed messenger RNA is inserted into the genome. The majority of retrotransposed genes acquire mutations that disrupt the coding sequence, thus disabling (or silencing) the gene. However, a small number become new genes that encode functional proteins. One example of this is fibroblast growth factor 4 (fgf4), a recently acquired retrogene in dogs. As noted in the previous study a comparative multibreed approach (involving both affected and unaffected dogs) was undertaken to demonstrate that expression of this gene is associated with chondrodysplasia, a short-legged phenotype that defines at least 19 dog breeds. This single gene acquisition event is one example illustrating the constraints and selection imposed by domesticated breeding practices (Parker et al., 2009).

A recent extensive study (Hayward et al., 2015) undertook a GWAS of 4,200 dogs, genotyping 180,000 diagnostic SNPs. Markers for hip dysplasia, elbow dysplasia, idiopathic epilepsy, lymphoma, mast cell tumor and granulomatous colitis were identified. Through computer-simulation studies the investigators determined that this large sample size of dogs and large set of SNPs will be sufficient to effectively map many complex canine diseases, while using fewer subjects than

needed to map human diseases. This paper allows us to transition into our final topic, the connection between human and canine disorders.

Parallels between canine and human diseases

There are several advantages to using dogs and the dog genome to study human diseases and disorders, and these are as follows. Among mammalian genomes, the canine genome has the highest similarity to the human genome, in terms of a disease model. Through selective breeding, different dog breeds have limited genetic variability among the individuals in that particular group, thus providing a more uniform genetic background in which disease-causing mutations can be detected. Dogs have similar diseases to humans, particularly complex diseases such as diabetes and cancer. The rate of occurrence of cancer in dogs is similar to that in humans, and the course of the disease in dogs is similar to that in humans. And finally, dogs and humans live in similar environments, and are exposed to many of the same agents that might be responsible for causing cancer (Ostrander and Franklin, 2012; Shearin and Ostrander, 2010).

Narcolepsy is a disabling sleep disorders characterized by sleepiness, abrupt transitions in sleep cycle, and abnormal sleep patterns. Researchers sought to determine the genomic reasoning behind this condition in both dogs and humans. In order to do this, molecular cloning was utilized to identify the autosomal recessive mutations responsible for narcolepsy. Disruption of the hypocretin receptor 2 gene was found to be the cause for narcolepsy among the individuals studied. Therefore, hypocretins are identified as major sleep-modulating neurotransmitters, leading to possible therapeutic approaches (Lin et al., 1999).

Some conditions are even more breed specific. For example, copper toxicosis in Bedlington terriers is genetic disease unique to this breed. Two copper carrier proteins have been identified as key components in copper homeostasis, more specifically, when dysfunctional, these proteins can cause either copper deficiency (Menkes disease) or copper accumulation among various tissues (Wilson disease). However, the main component affecting Bedlington terriers is the impairment of biliary excretion of copper of which these proteins are not directly associated with. In previous studies, copper toxicosis has been mapped to the chromosome region 10q26. In more recent investigations, localization of the copper toxicosis gene has been performed. The location has been confined to a region of <500 kb by linkage disequilibrium mapping. Furthermore, exon 2 of the *MURR1* gene was found to be deleted in both alleles of all affected Bedlington terriers. This finding has provided a new lead to understanding the complexities of copper metabolism in mammals (van der Sluis et al., 2002; Struehler et al., 2004).

Ichthyoses comprise a heterogeneous group of genodermatoses characterized by scale formation over the whole body. The genetic causes of several human forms remain unknown. However, the golden retriever breed served as a model of study for this condition. Golden retrievers are often affected by a lamellar icthyosis resembling that of human autosomal recessive congenital ichthyoses (ARCI). A GWAS identified a homozygous insertion-deletion mutation in the *PNPLA1* gene. This mutation leads to a premature stop codon in the coding sequence in the gene in all affected golden retrievers. Furthermore, one missense and one nonsense

mutation in the catalytic domain of human *PNPLA1* in six individuals were identified (Grall et al., 2012).

Canine invasive transitional cell carcinoma of the bladder (InvTCC) is a naturally occurring tumor that shares several clinical phenotypes with human muscle invasive bladder cancer. To identify the main components causing this condition, researchers used RNA sequencing (RNA-Seq)to determine the complete transcriptome for multiple tumors. All of the tumors contained a somatic mutation that is homologous to the human BRAF(V600E) mutation. The mutation was also detectable in urine sediments of all dogs tested with mutation-positive tumors. Just like human tumors, it is suggested that canine activating BRAF mutations stimulate the MAPK cell-signaling pathway. These findings have assisted in better understanding IvcTCC and BRAF-targeted therapies (Decker et al., 2015).

Among the more intensively examined cancers is that of the prostate. Prior to the use of genomic approaches more tedious methods were used to identify the genetic markers. By using a genome-wide screen from 70 affected families at risk for prostate cancer a specific marker locus had been located on chromosome 1 through classical pedigree analysis (Gibbs et al., 1999). This effort has been superseded by more recent developments. GWAS analysis has led to the identification of a number of diagnostic SNPs for the detection and prediction of recurrence of this disease. A set of forty SNPs was analyzed in a population of 553 affected men (prostate cancer) and 534 unaffected men. Three SNPs were associated with the recurrence and progression of the cancer (Holt et al., 2008). Chronic inflammation is partially responsible for the onset of this cancer. 143

candidate SNPs were examined in 16 candidate genes in a study involving several thousand men. From this set of mutations, ten SNPs in seven genes were found to be associated with both inflammation and the onset of prostate cancer (Kwon et al., 2011).

The Future

The long-term goal of the dog genome effort is severalfold. First, to determine which alleles are responsible for the formation of specific dog breeds. For example, which alleles control various behaviors (such as herding), and which alleles control other phenotypes (such as skull shape). The effort is to develop a genotype/phenotype map, to be able to attribute the physical traits of each dog breed to specific gene combinations/alleles (Shearin and Ostrander, 2010). Second, to be able to screen breeding stock for specific markers that could be indicative of potential disorders in the offspring. In my interview with Dr. Ostrander, she noted that one source of her canine DNA samples was from dog breeders, who are interested in attempting to eliminate individuals from their breeding programs who might carry alleles that could cause disorders in the offspring. The eventual goal is to try to eliminate from different breeds as many deleterious alleles as possible so that future generations of dogs might carry less of a burden of specific disorders. The connection between human and canine disorders will be further developed. This will involve further mapping of canine disorders in the dog genome, separating environmental influences on disease from genetic causes, and developing potential therapies. And finally, the techniques and methods used to map specific genes in the canine genome could then be applied to mapping disease-causing genes in the more

complex and variable human genome (Sutter and Ostrander, 2004; Boyko, 2011). Dogs may indeed, on many levels, turn out to truly be "man's best friend".

Final Conclusion

I began this effort with the two statements outlined at the beginning of this thesis. The first that all knowledge is of value, and should be respected for its inherent value. The second is that there are connections among all fields of knowledge, and contributions to a specific area could come from a related discipline. Both of these thesis statements are supported by the canine genome effort. Clearly, outside of dog owners, canine health is not a major concern of most people, but many discoveries in the canine genome effort have led to insights into human diseases, in terms of the genetic basis and diagnostic approaches. The canine genome effort has also led to the development of new genetic approaches to examining genomes and detecting specific genes and alleles, which has applications to other fields, such as human genetics. Archeology has contributed to the study of the origin of the dog, and has offered support for some hypotheses. The physical science of radiocarbon dating has allowed for the accurate assignment of fossils to specific dates, thus shedding light on dog origins. It is clear that we should not judge the value of a particular field of study by its immediate economic impact, but need to be open-minded, as the future value or application of a particular field of study is often unknown in the present.

Literature Cited

Boyko, A.R. 2011. The domestic dog: man's best friend in the genomic era. Genome Biol. 12: 216-225.

Boyko, A.R., Boyko, R.H., Boyko, C.M., Parker, H.G., Castelhano, M., Corey, E.E., Degenhardt, J.D., Auton, A., Hedimbi, M., Kityo, R., et al. 2009. Complex population structure in African village dogs and its implication for inferring dog domestication history. Proc. Natl. Acad. Sci. 106: 13903-13908.

Buikstra, J.E., Druzhova, A., Graphodatsy, A.S, Ovodov, N.D., Wahlberg, N., Freedman, A.H., Schweizer, R.M., Koepfli, K.P., Leonard, J.A., Meyer, M., Krause, J., Paabo, S., Green, R.E., and Wayne, R.K. 2013. Compete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. Science 342: 871-874.

Davis, S.J.M. and Valla, F.R. 1978. Evidence for domestication of the dog 12,000 years ago in the Natufian of Israel. Nature 276: 608-610.

Decker, B., Davis, B.W., Rimbault, M., Long, A.H., Karlins, E., Jagannathan, V., Reiman, R., Parker, H.G., Drogemuller, C., Corneveaux, J.J., Chapman, E.S., Trent, J.M., Leeb, T., Huentelman, J.M., Wayne, R.K., Karyadi, D.M., and Ostrander, E.A. 2015. Comparison against 186 canid whole genome sequences reveals survival strategies of an ancient clonally transmissible canine tumor. Genome Res. 25: 1646-1655.

Decker, B., Parker, H.G., Dhawan, D., Kwon, E.M., Karlins, E., Davis, B.W., Ramos-Vara, J.A., Bonney, P.L., McNiel, E.A., Knapp, D.W., and Ostrander, E.A. 2015. Homologous mutation to human BRAF V600E is common in naturally

occurring canine bladder cancer - evidence for a relevant model system and urine-based diagnostic test. Mol. Canc. Res. 13: 993-1002.

Dodman, N.H., Karlsson, E.K., Moon-Fanelli, A., Galdzicka, M, Perloski, M., Schuster, L., Lindblad-Toh, K., and Ginns, E.I. 2010. A canine chromosome 7 locus confers compulsive disorder susceptibility. Mol. Psychiatry 15: 8-10.

Freedman, A., Schweizer, R.M., Gronau, I., Han, E., Vecchyo, D.O.D., Silva, P., Galaverni, M., Zhenxin, F., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vilà, C., Squire, K., Geffern, E., Kusak, J., Boyko, A.R., Parker, H., Lee, C., Tadigotla, V., Siepel, A., Bustamante, C., Harkins, T., Nelson, S.F., Ostrander, E.A., Marques-Bonet, T., Wayne, R.K., and Novembre, J. 2014. Genome sequencing highlights the dynamic early history of dogs. PLoS Genetics 10: e1004016.

Friedenberg, S.G., Lan, Z., Zhiwu, Z., van den Foels, W.B., Schweitzer, P.A., Wei, W., Fisher, P.J., Dykes, N.L., Corey, E.E., Vernier-Singer, M.A., et al. 2011. Evaluation of a fibrillin 2 gene haplotype associated with hip dyaplasia and incipient osterarthritis in dogs. Am. J. Vet. Res. 72: 530-540.

Germonpre, M., Sablin, M.V., Stevens, R.E., Hedges, R.E.M., Hofteiter, M., Stiller, M., and Despres, V.R. 2009. Fossil dogs and wolves from Palaeolithic sites in Belgium, the Ukraine, and Russia: osteometry, ancient DNA and stable isotopes. J. Archaeol. Sci. 36: 473-490.

Gibbs, M., Stanford, J.L., McIndoe, R.A., Jarvik, G.P., Kolb, S., Goode, E.L., Chakrabarti, L., Schuster, E.F., Buckley, V.A., Miller, E.L., Brandzel, S., Li, S., Hood, L., and

Ostrander, E.A. 1999. Evidence for a rare prostate cancer-susceptibility locus at chromosome 1p36. Am. J. Hum. Genet. 64: 776-87.

Grall, A., Guaguere, E., Planchais, S., Grond, S., Bourrat, E., Hausser, I., Hitte, C., Le Gallo, M., Derbois, C., Kim, G.-J., et al. 2012. PNPLA1 mutations cause autosomal recessive congenital ichthyosis in golden retriever dogs and humans. Nature Genetics 44: 140-147.

Hayward, J.J., Castelhano, M.G., Oliviera, K.C., Corey, E., Balkman, C., Baxter, T.L., Casal, M.L., Center, S.A., Fang, M., Garrison, S.J., Kalla, S.E., Korniliev, P., Kotlikoff, M.I., Moise, N.S., Shannon, L.M., Simpson, K.W., Sutter, N.B., Todhunter, R.B., & Boyko, A.R. 2015. Complex disease and phenotype mapping in the domestic dog. Nature Communications 7: 10460.

Holt, S.K., Karyad, D.M., Kwon, E.M., Stanford, J.L., Nelson, P.S., and Ostrander, E.A. 2008. Association of megalin genetic polymorphisms with prostate cancer risk and prognosis. Clinical Cancer Research 15: 3823-3831.

Jedrzejewski, W., Greco, C., Randi, E., Bannasch, D., Wilton, A., Shearman, J., Musiani, M., Cargill, M., Jones, P.G., Qian, Z., Huang, W., Ding, Z.L., Zhang, Y.P., Bustamante, C.D., Ostrander, E.A., Novembre, J., and Wayne, R.K. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464: 898-902.

Karlsson, E.K., Baranowska, I., Wade, C.M., Hillbertz, N.H.C., Zody, M.C., Anderson, N., Biagi, T.M., Patterson, N., Pielberg, G.R., Kulbokas, E.J. III, Comstock, K.E., Keller, E.T., Mesirov, J.P., von Euler, H., K|[auml]|mpe, O., Hedhammar, A., Lander, E.S., Andersson, G., Andersson, L., and Lindblad-Toh, K. 2007.

Efficient mapping of mendelian traits in dogs through genome-wide association. Nature Genetics 39: 1321-1328.

Karyadi, D.M., Karlins, E., Decker, B., vonHoldt, B.M., Carpintero-Ramirez, G., Parker, H.G., Wayne, R.K., and Ostrander, EA. 2013. A copy number variant at the KITLG locus likely confers risk for canine squamous cell carcinoma of the digit. PLoS Genetics 9: e1003409.

Kwon, E.M., Salinas, C.A., Kolb, S., Fu, R., Feng, Z., Stanford, J.L., and Ostrander, E.A. 2011. Genetic polymorphisms in inflammation pathway genes and prostate cancer risk. Cancer Epidemiol. Biomarkers Prev. 20: 923-933.

Lin, L., Faraco. J., Li, R., Kadotani, H., Rogers, W., Lin, X., Qiu, X., de Jong, P.J., Nishino, S. and Mignot, E. 1999. The sleep disoreder caninine narcolepsy is caused by a mutation in the hypocretin (oroexin) receptor 2 gene. Cell 98: 365-376.

Lindberg, J., Bjornerfeldt, S., Saetre, P., Svartberg, K., Seehuus, B., Bakken, M., Vila, C., and Jazin, E. 2005. Selection for tameness has changed brain gene expression in silver foxes. Current Biology 15: 915-916.

Lingaas, F., Comstock, K.E., Kirkness, E.F., Sørensen, A., Aarskaug, T., Hitte, C., Nickerson, M.L., Moe, L., Schmidt, L.S., Thomas, R., Breen, M., Galibert, F., Zbar, B., and Ostrander, E.A. 2003. A mutation in the canine BHD gene is associated with hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis in the German Shepherd Dog. Hum. Mol. Genet. 12: 3043-53.

Miklosi, A., Kubinyi, E., Topal, J., Gacsi, M., Viranyi, Z., and Csanyi, V. 2003. A simple reason for a big difference: wolves do not look back at humans, but dogs do. Current Biology 13: 763-766.

Noorgard, K. 2008. Genetic variation and disease: GWAS. Nature Education 1: 87.

Ostrander EA., and Franklin H. Epstein Lecture. 2012. Both ends of the leash: The human links to good dogs with bad genes. New Engl. J. Med. 367: 636-46.

Pang, J.-F., Kluetsch, C., Zou, X.-J., Zhang, A.-b., Luo, L.-Y., Angelby, H., Ardalan, H., Ekstrom, C., Skollermo, A., Lundberg, J., et al. 2009. mtDNA indicates a single origin for dogs south of the Yangtze River, less than 16,300 years ago, from numerous wolves. Mol. Biol. Evol. 26: 2849-2864.

Parker, H.G., Kukekova, A.V., Akey, D.T., Goldstein, O., Kirkness, E.F., Baysac, K.C., Mosher, D.S., Aguirre, G.D., Acland, G.M., and Ostrander, E.A. 2007. Breed relationships facilitate fine mapping studies: A 7.8 Kb deletion cosegregates with collie eye anomaly across multiple dog breeds. Genome Res. 17:1562-1571.

Parker, H.G., vonHoldt, B.M., Quignon, P., Margulies, E.H., Shao, S., Mosher, D.S., Spady, T.C., Elkahloun, A., Cargill, M., Jones, P.G., Maslen, C.L., Acland, G.M., Sutter, N.B., Kuroki, K., Bustamant,e C.D., Wayne, R.K., and Ostrander, E.A. 2009. An expressed Fgf4 retrogene is associated with breed-defining chondrodysplasia in domestic dogs. Science 325:995-8.

Sablin, M.V. and Khlopachev, G.A. 2002. The earliest ice age dogs: evidence from Eliseevichi I. Curr. Anthropol. 43: 795-799.

Saetre, P., Lindberg, J., Leonard, J.A., Olsson, K., Petterson, U., Ellegren, H., Bergstrom, T. F., Vila, C., and Janzin, E. 2004. From wild wolf to domestic dog: gene expression changes in the brain. Molecular Brain Research 126: 198-206.

Savolainen, P., Zhang, Y.-P., Luo, J., Lunderberg, J., and Leitner, T. 2002. Genetic evidence for an East Asian origin of domestic dogs. Science 298: 1610-1613.

Shannon, L.M., Boyko, R.M., Castelhano, M., Corey, E., Hayward, J.H., McLean, C., White, M.E., Said, M.A., Anita, B.A., Bondjengo, N.I., Calero, J., Galov, A., Hedimbi, M., Imam, B., Khalap, R., Lally, D., Masta, A., Oliveira, K.C., Pérez, L., Randall, J., Tam, N.M., Trujillo-Cornejo, F.J., Valeriano, C., Sutter, N.B., Todhunter, R.J., Bustamante, C.D., and Boyko, A.R. 2015. Genetic structure in village dogs reveals a Central Asian domestication origin. Proc. Natl. Acad. Sci. 112: 13639-13644.

Shearin, A.L,, Hedan, B., Cadieu, E,, Erich, S.A., Schmidt, E.V., Faden, D.L., Cullen, J., Abadie, J., Kwon, E.M., Gröne, A., Devauchelle, P., Rimbault, M., Karyadi, D.M., Lynch, M., Galibert, F., Breen, M., Rutteman, G.R., André, C., Parker, H.G., and Ostrander, E.A. 2012. The MTAP-CDKN2A locus confers susceptibility to a naturally occurring canine cancer. Cancer Epidemiol Biomarkers Prev, 21:1019-27.

Shearin, A.L., and E.A. Ostrander. 2010. Leading the way: canine models of genomics and disease. Disease models and mechanisms 3: 27-34.

Skoglund, P., Ersmark, E., Palkopoulou, E., and Dalen, L. 2015. Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. Curr. Biol. 25: 1515-1519.

Spady, T.C., and Ostrander, E.A. 2008. Canine behavioral genetics: Pointing out the phenotype and herding up the genes. Am. J. Human Genetics 82: 10-18.

Struehler, B., Juergen, R., Wolfgang, S., and Schaefer, M. 2004. Analysis of the human homologue of the canine copper toxicosis gene MURR1 in Wilson disease patients. J. Mol. Med. 82: 629 – 634.

Thalmann, O., Shapiro, B., Cui. P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpré, M.B., Sablin, M.V., López-Giráldez, F., Domingo-Roura, X., Napierala, H., Uerpmann, H.P., Loponte, D.M., Acosta, A.A., Giemsch, L., Schmitz, R.W., Worthington B, Buikstra, J.E., Druzhkova, A., Graphodatsky, A.S., Ovudov, N.D., Wahlberg, N., Freedman, A.H., Schweizer, R.M., Koepfil, K.-P., Leonard, J.A., Meyer, M., Krause, J., Paabo, S., Green, R.E., and Wayne, R.K. 2013. Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. Science 342: 871-874.

Tchernov, E. and Valla, F.F. 1997. Two new dogs, and other Natufian dogs, from the southern Levant. J. Heredity 24: 65-95.

Udell, M.A., Dorey, N.R., and Wynne, C.D. 2008. Wolves outperform dogs in following human social cues. Animal Behavior 76: 1767 – 1773.

Vila, C., Savolainen, P., Maldonado, J.E, Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundberg, J., and Wayne, R.K. 1997. Multiple and ancient origins of the domestic dog. Science 276: 1687-1689.

Van de Sluis, B., Rothuizen, J., Pearson, P.L., van Oost, B.A., and Wijmenga, C. 2002. Identification of a new copper metabolism gene by positional cloning in a purebred dog population. Hum. Mol. Genet. 11:165-173.

Von Holdt, B.M., Pollinger, J.P., Lohmueller, K.E., Han, E. , Parker, H.G., Quignon, P., Degenhardt, J.D., Boyko, A.R., Earl, D.A., Auton, A., Reynolds, A., Bryc, K.,

Brisbin, A., Knowles, J.C., Mosher, D.S., Spady, T.C., Elkahloun, A., Geffen, E., Pilot, M., Jedrzejewski, W., Greco, C., Randi, E., Bannasch, D., Wilton, A., Shearman, J., Musiani, M., Cargill, M., Jones, P.G., Qian, Z., Huang, W., Ding, Z.-L., Zhang, Y.-P., Bustamante, C.D., Ostrander, E.A., Novembre, J., and Wayne, R.K. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464: 898-902.

Wang, G.D., Zhai, W., Yang, H-C., Wang, L., Zhong, L., Liu, Y.-H., Fan, R.-X., Yin, T.-T., Zhu, C.H., Poyarkov, A.D., Irwin, D.M., Hytönen, M.K., Loh, H., Wu, , C.-I., Savolainen, P., and Zhang, Y.-P. 2015. Out of southern East Asia: the natural history of domestic dogs across the world. Cell Research 26: 21-33.

Wayne, R.K. 1993. Molecular evolution of the dog family. Trends in Genetics 9: 218 – 224.

Appendix 1

Annotated Bibliography

General Overview

The following two review papers provide a general review of the scope and nature of the canine genome effort. Adam Boyko and Elaine Ostrander are two of the leaders in the efforts to sequence the canine genome, and they continue to investigate the connections between the dog and human genomes. These two papers summarize the justification and rationale for sequencing the dog genome, and the benefits of this research effort.

Boyko, A.R. 2011. The domestic dog: man's best friend in the genomic era. Genome Biol. 12: 216-225.

Ostrander, E.A, and R.K. Wayne. 2005. The canine genome. Genome Res. 15: 1706-1716.

Shearin, A.L. and E.A. Ostrander. 2010. Canine morphology: hunting for genes and tracking mutations. PLoS Biology 8: e10000310.

Sutter, N.B., and E.A Ostrander. 2004. Dog star rising: the canine genetic system. Nature Rev. Genet. 5: 900-910.

Geographic origin and domestication of dogs

The papers below describe pre-Canine Genome efforts at determining the evolutionary and geographic origin of domesticated dogs. These approaches were based primarily on archeological studies of ancient human settlements and associated canines. The Thalman et al. (2013) and Buikstra et al. (2013) papers suggest a European origin of dogs, using mitochondrial DNA sequences.

Buikstra, J.E., Druzhova, A., Graphodatsy, A.S, Ovodov, N.D., Wahlberg, N., Freedman, A.H., Schweizer, R.M., Koepfli, K.P., Leonard, J.A., Meyer, M., Krause, J., Paabo, S., Green, R.E., and Wayne R.K. 2013. Compete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. Science 342: 871-874.

Davis, S.J.M. and Valla, F.R. 1978. Evidence for domestication of the dog 12,000 years ago in the Natufian of Israel. Nature 276: 608-610.

Germonpre, M., Sablin, M.V., Stevens, R.E., Hedges, R.E.M., Hofteiter, M., Stiller, M., and Despres, V.R. 2009. Fossil dogs and wolves from Palaeolithic sites in Belgium, the Ukraine, and Russia: osteometry, ancient DNA and stable isotopes. J. Archaeol. Sci. 36: 473-490.

Ovodov ND, Crockford SJ, Kuzmin YV, Higham TFG, Hodgins GWL, et al. (2011) A 33,000-Year-Old Incipient Dog from the Altai Mountains of Siberia: Evidence of the Earliest Domestication Disrupted by the Last Glacial Maximum. PLoS ONE 6: e22821.

Sablin, M.V. and Khlopachev, G.A. 2002. The earliest ice age dogs: evidence from Eliseevichi I. Curr. Anthropol. 43: 795-799.

Tchernov, E. and Valla, F.F. 1997. Two new dogs, and other Natufian dogs, from the southern Levant. J. Heredity 24: 65-95.

Thalmann, O., Shapiro, B., Cui. P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpré, M.B., Sablin, M.V., López-Giráldez, F., Domingo-Roura, X., Napierala, H., Uerpmann, H.P., Loponte, D.M., Acosta, A.A., Giemsch, L., Schmitz, R.W., Worthington B, Buikstra, J.E., Druzhkova, A., Graphodatsky, A.S., Ovudov, N.D., Wahlberg, N., Freedman, A.H., Schweizer, R.M., Koepfil, K.-P., Leonard, J.A., Meyer, M., Krause, J., Paabo, S., Green, R.E., and Wayne, R.K. 2013. Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. Science 342: 871-874.

Vila, C., Savolainen, P., Maldonado, J.E, Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundberg, J., and Wayne, R.K. 1997. Multiple and ancient origins of the domestic dog. Science 276: 1687-1689.

Wayne, R.K. 1993. Molecular evolution of the dog family. Trends in Genetics 9: 218 – 224.

Wayne, R.K., and Ostrander, E.A. 1999. Origin, genetic diversity, and genome structure of the domestic dog. Bioessays 21: 247-257.

Wayne, R.K., and vonHoldt, B.M. 2012. Evolutionary genetics of dog domestication. Mamm. Genome 23: 3-18.

With the development of advanced genetic analysis techniques and genomic analyses, new evidence arose regarding the geographic origin of domestic dogs. The papers below provide evidence for the origin of dogs in Asia, not Europe or the Middle East. The determination of the geographic origin of dogs is controversial. Early archeological evidence put the origins of the dog in the Middle East; early genetic studies suggested a European origin for dogs. More recent genetic evidence establishes Asia as the most likely place of origin for canines.

Bustamante, C.D., and Boyko, A.R. 2015. Genetic structure in village dogs reveals a Central Asian domestication origin. Proc. Natl. Acad. Sci. 112: 13639-13644.

Freedman, A., Schweizer, R.M., Gronau, I., Han, E., Vecchyo, D.O.D., Silva, P., Galaverni, M., Zhenxin, F., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vilà, C., Squire, K., Geffern,, E., Kusak, J., Boyko, A.R., Parker, H., Lee, C., Tadigotla, V., Siepel, A., Bustamante, C., Harkins, T., Nelson, S.F., Ostrander, E.A., Marques-Bonet, T., Wayne, R.K., and Novembre, J. 2014. Genome sequencing highlights the dynamic early history of dogs. PLoS Genetics 10: e1004016.

Jedrzejewski, W., Greco, C., Randi, E., Bannasch, D., Wilton, A., Shearman, J., Musiani, M., Cargill, M., Jones, P.G., Qian, Z., Huang, W., Ding, Z.L., Zhang, Y.P., Bustamante, C.D., Ostrander, E.A., Novembre, J., and Wayne, R.K. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464: 898-902.

Pang, J.-F., Kluetsch, C., Zou, X.-J., Zhang, A.-b., Luo, L.-Y., Angelby, H., Ardalan, H., Ekstrom, C., Skollermo, A., Lundberg, J., et al. 2009. mtDNA indicates a single origin for dogs south of the Yangtze River, less than 16,300 years ago, from numerous wolves. Mol. Biol. Evol. 26: 2849-2864.

Savolainen, P., Zhang, Y.-P., Luo, J., Lunderberg, J., and Leitner, T. 2002. Genetic evidence for an East Asian origin of domestic dogs. Science 298: 1610-1613.

Shannon, L.M., Boyko, R.M., Castelhano, M., Corey, E., Hayward, J.H., McLean, C., White, M.E., Said, M.A., Anita, B.A., Bondjengo, N.I., Calero, J., Galov, A., Hedimbi, M., Imam, B., Khalap, R., Lally, D., Masta, A., Oliveira, K.C., Pérez, L., Randall, J., Tam, N.M., Trujillo-Cornejo, F.J., Valeriano, C., Sutter, N.B., Todhunter, R.J.,

Skoglund, P., Ersmark, E., Palkopoulou, E., and Dalen, L. 2015. Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. Curr. Biol. 25: 1515-1519.

Von Holdt, B.M., Pollinger, J.P., Lohmueller, K.E., Han, E. , Parker, H.G., Quignon, P., Degenhardt, J.D., Boyko, A.R., Earl, D.A., Auton, A., Reynolds, A., Bryc, K., Brisbin, A., Knowles, J.C., Mosher, D.S., Spady, T.C., Elkahloun, A., Geffen, E., Pilot, M., Jedrzejewski, W., Greco, C., Randi, E., Bannasch, D., Wilton, A., Shearman, J., Musiani, M., Cargill, M., Jones, P.G., Qian, Z., Huang, W., Ding, Z.-L., Zhang, Y.-P., Bustamante, C.D., Ostrander, E.A., Novembre, J., and Wayne, R.K. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464: 898-902.

Wang, G.D., Zhai, W., Yang, H-C., Wang, L., Zhong, L., Liu, Y.-H., Fan, R.-X., Yin, T.-T., Zhu, C.H., Poyarkov, A.D., Irwin, D.M., Hytönen, M.K., Loh, H., Wu, , C.-I., Savolainen, P., and Zhang, Y.-P. 2015. Out of southern East Asia: the natural history of domestic dogs across the world. Cell Research 26: 21-33.

The following paper discusses how analysis of existing African dog populations has provided insights into the evolutionary history of dog domestication.

Boyko, A.R., Boyko, R.H., Boyko, C.M., Parker, H.G., Castelhano, M., Corey, E.E., Degenhardt, J.D., Auton, A., Hedimbi, M., Kityo, R., et al. 2009. Complex population structure in African village dogs and its implication for inferring dog domestication history. Proc. Natl. Acad. Sci. 106: 13903-13908.

And this paper discusses how a single gene, that produces a protein that allows for the processing of starch in the diet, could have contributed to the domestication of the dog.

Axelsson E, Ratnakumar A, Arendt M-J, Maqbool K, Webster MT, et al. (2013) The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature 495: 360–364.

Canine Diseases and Disorders

The following papers describe canine disorders and their genetic basis. The primary approach taken was either GWAS (Genome Wide Association Study), utilizing SNPs (single nucleotide polymorphisms), or comparative genomics. These are some of the direct benefits of the canine genome effort, with a direct impact on dog breeding practices.

Decker, B., Davis, B.W., Rimbault, M., Long, A.H., Karlins, E., Jagannathan, V., Reiman, R., Parker, H.G., Drogemuller, C., Corneveaux, J.J., Chapman, E.S., Trent, J.M., Leeb, T., Huentelman, J.M., Wayne, R.K., Karyadi, D.M., and Ostrander, E.A. 2015. Comparison against 186 canid whole genome sequences reveals survival strategies of an ancient clonally transmissible canine tumor. Genome Res. 25: 1646-1655.

Dodman, N.H., Karlsson, E.K., Moon-Fanelli, A., Galdzicka, M, Perloski, M., Schuster, L., Lindblad-Toh, K., and Ginns, E.I. 2010. A canine chromosome 7 locus confers compulsive disorder susceptibility. Mol. Psychiatry 15: 8-10.

Friedenberg, S.G., Lan, Z., Zhiwu, Z., van den Foels, W.B., Schweitzer, P.A., Wei, W., Fisher, P.J., Dykes, N.L., Corey, E.E., Vernier-Singer, M.A., et al. 2011. Evaluation of a fibrillin 2 gene haplotype associated with hip dyaplasia and incipient osterarthritis in dogs. Am. J. Vet. Res. 72: 530-540.

Hayward, J.J., Castelhano, M.G., Oliviera, K.C., Corey, E., Balkman, C., Baxter, T.L., Casal, M.L., Center, S.A., Fang, M., Garrison, S.J., Kalla, S.E., Korniliev, P., Kotlikoff, M.I., Moise, N.S., Shannon, L.M., Simpson, K.W., Sutter, N.B., Todhunter, R.B., & Boyko, A.R. 2015. Complex disease and phenotype mapping in the domestic dog. Nature Communications 7: 10460.

Karlsson, E.K., Baranowska, I., Wade, C.M., Hillbertz, N.H.C., Zody, M.C., Anderson, N., Biagi, T.M., Patterson, N., Pielberg, G.R., Kulbokas, E.J. III, Comstock, K.E., Keller, E.T., Mesirov, J.P., von Euler, H., K|[auml]|mpe, O., Hedhammar, A., Lander, E.S., Andersson, G., Andersson, L., and Lindblad-Toh, K. 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. Nature Genetics 39: 1321-1328.

Karyadi, D.M., Karlins, E., Decker, B., vonHoldt, B.M., Carpintero-Ramirez, G., Parker, H.G., Wayne, R.K., and Ostrander, EA. 2013. A copy number variant at the *KITLG*

locus likely confers risk for canine squamous cell carcinoma of the digit. PLoS Genetics 9: e1003409.

Lingaas, F., Comstock, K.E., Kirkness, E.F., Sørensen, A., Aarskaug, T., Hitte, C., Nickerson, M.L., Moe, L., Schmidt, L.S., Thomas, R., Breen, M., Galibert, F., Zbar, B., and Ostrander, E.A. 2003. A mutation in the canine *BHD* gene is associated with hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis in the German Shepherd Dog. Hum. Mol. Genet. 12: 3043-53.

Lynch, M., Galibert, F., Breen, M., Rutteman, G.R., André, C., Parker, H.G., and Ostrander, E.A. 2012. The *MTAP-CDKN2A* locus confers susceptibility to a naturally occurring canine cancer. Cancer Epidemiol. Biomarkers Prev. 21:1019-27.

Parker, H.G., Kukekova, A.V., Akey, D.T., Goldstein, O., Kirkness, E.F., Baysac, K.C., Mosher, D.S., Aguirre, G.D., Acland, G.M., and Ostrander, E.A. 2007. Breed relationships facilitate fine mapping studies: A 7.8 Kb deletion cosegregates with collie eye anomaly across multiple dog breeds. Genome Res. 17:1562-1571.

Parker, H.G., vonHoldt, B.M., Quignon, P., Margulies, E.H., Shao, S., Mosher, D.S., Spady, T.C., Elkahloun, A., Cargill, M., Jones, P.G., Maslen, C.L., Acland, G.M., Sutter, N.B., Kuroki, K., Bustamante, C.D., Wayne, R.K., and Ostrander, E.A. 2009. An expressed *Fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *S*cience 325:995-8.

Shearin, A.L,, Hedan, B., Cadieu, E,, Erich, S.A., Schmidt, E.V., Faden, D.L., Cullen, J., Abadie, J., Kwon, E.M., Gröne, A., Devauchelle, P., Rimbault, M., Karyadi, D.M., Van de Sluis, B., Rothuizen, J., Pearson, P.L., van Oost, B.A., and Wijmenga, C. 2002. Identification of a new copper metabolism gene by positional cloning in a purebred dog population. Hum. Mol. Genet. 11: 165-173.

Shearin, A.L., and E.A. Ostrander. 2010. Leading the way: canine models of genomics and disease. Disease models and mechanisms 3: 27-34.

Van de Sluis, B., Rothuizen, J., Pearson, P.L., van Oost, B.A., and Wijmenga, C. 2002. Identification of a new copper metabolism gene by positional cloning in a purebred dog population. Hum. Mol. Genet. 11: 165-173.

Relationship of canine disease to human disorders

Decker, B., Parker, H.G., Dhawan, D., Kwon, E.M., Karlins, E., Davis, B.W., Ramos-Vara, J.A., Bonney, P.L., McNiel, E.A., Knapp, D.W., and Ostrander, E.A. 2015. Homologous mutation to human BRAF V600E is common in naturally occurring canine bladder cancer - evidence for a relevant model system and urine-based diagnostic test. Mol. Canc. Res. 13: 993-1002.

Gibbs, M., Stanford, J.L., McIndoe, R.A., Jarvik, G.P., Kolb, S., Goode, E.L., Chakrabarti, L., Schuster, E.F., Buckley, V.A., Miller, E.L., Brandzel, S., Li, S., Hood, L., and Ostrander, E.A. 1999. Evidence for a rare prostate cancer-susceptibility locus at chromosome 1p36. Am. J. Hum Genet. 64: 776-87.

Grall, A., Guaguere, E., Planchais, S., Grond, S., Bourrat, E., Hausser, I., Hitte, C., Le Gallo, M., Derbois, C., Kim, G.-J., et al. 2012. PNPLA1 mutations cause autosomal recessive congenital ichthyosis in golden retriever dogs and humans. Nature Genetics 44: 140-147.

Holt, S.K., Karyad, D.M., Kwon, E.M., Stanford, J.L., Nelson, P.S., and Ostrander, E.A. 2008. Association of megalin genetic polymorphisms with prostate cancer risk and prognosis. Clinical Cancer Research 15: 3823-3831.

Kwon, E.M., Salinas, C.A., Kolb, S., Fu, R., Feng, Z., Stanford, J.L., and Ostrander, E.A. 2011. Genetic polymorphisms in inflammation pathway genes and prostate cancer risk. Cancer Epidemiol. Biomarkers Prev. 20: 923-933.

Lin, L., Faraco. J., Li, R., Kadotani, H., Rogers, W., Lin, X., Qiu, X., de Jong, P.J., Nishino, S. and Mignot, E. 1999. The sleep disoreder caninine narcolepsy is caused by a mutation in the hypocretin (oroexin) receptor 2 gene. Cell 98: 365-376.
Ostrander, EA., and Franklin, H. Epstein Lecture. 2012. Both ends of the leash--The human links to good dogs with bad genes. New Engl. J. Med. 367: 636-46.

Ostrander, EA., and Franklin, H. Epstein Lecture. 2012. Both ends of the leash--The human links to good dogs with bad genes. *New Engl. J. Med.* 367: 636-46.

Shearin, A.L. and E.A. Ostrander. 2010. Leading the way: canine models of genomics and disease. Disease Models and Mechanisms 3: 27-34.

Struehler, B., Juergen, R., Wolfgang, S., and Schaefer, M. 2004. Analysis of the human homologue of the canine copper toxicosis gene MURR1 in Wilson disease patients. J. Mol. Med. 82: 629 – 634.

Van de Sluis, B., Rothuizen, J., Pearson, P.L., van Oost, B.A., and Wijmenga, C. 2002. Identification of a new copper metabolism gene by positional cloning in a purebred dog population. Hum. Mol. Genet. 11: 165-173.

Evolution of behavior

The specific genes involved in altering dog behavior from a wolf-like ancestor, and behaviors that differentiate wolves from dogs, are discussed in the papers below.

Li, Y., vonHoldt, B.M., Reynolds, A. Boyko, A.R., Wayne, R.K., et al. (2013) Artificial selection on brain expressed genes during the domestication of dog. Mol Biol Evol doi: 10.1093/molbev/mst088.

Lindberg, J., Bjornerfeldt, S., Saetre, P., Svartberg, K., Seehuus, B., Bakken, M., Vila, C., and Jazin, E. 2005. Selection for tameness has changed brain gene expression in silver foxes. Current Biology 15: 915-916.

Miklosi, A., Kubinyi, E., Topal, J., Gacsi, M., Viranyi, Z., and Csanyi, V. 2003. A simple reason for a big difference: wolves do not look back at humans, but dogs do. Current Biology 13: 763-766.

Saetre, P., J. Lindberg, J.A. Leonard, K. Olsson, U. Petterson, H. Ellegren, T. F. Bergstrom, C. Vila, and E. Janzin. 2004. From wild wolf to domestic dog: gene expression changes in the brain. Molecular Brain Research 126: 198-206.

Spady, T.C., and Ostrander, E.A. 2008. Canine behavioral genetics: Pointing out the phenotype and herding up the genes. Am. J. Human Genetics 82: 10-18.

Udell, M.A., Dorey, N.R., and Wynne, C.D. 2008. Wolves outperform dogs in following human social cues. Animal Behavior 76: 1767 – 1773.

Appendix 2

BIO598 Independent Research proposals for Fall 2015 and Spring 2016. These efforts provided Rachael Lander with practical experience and exposure to scientific research by setting up experiments and collecting data (Fall 2015) and the generation of specific DNA fragments through the Polymerase Chair Reaction (Spring 2016). The Fall 2015 effort was submitted for publication to the American Biology Teacher in December 2015. It was accepted, pending revision, in March 2016. A revised manuscript will be submitted with additional experiments in August 2016.

BIO 598
Independent Research
Fall 2015
Rachael M. Lander

Introduction

The effects of various alcohols has been documented in the literature for fruit flies (*Drosophila*), and we have carried out past experiments to assess the effects of isopropanol, ethanol, and acetic acid on different genetic stocks of fruit flies (this was an experiment that we used for several semesters in Genetics Lab). We now want to conduct parallel experiments in a plant system, to address the following two questions:

1) Do different alcohols have an effect on plant growth?
2) If so, do the different alcohols exert different effects (in terms of magnitude) on plant growth? And are any alcohols stimulatory in their effects on plant growth? Or do the alcohols exhibit an inhibitory effect?

The goal of this effort is to build a comparison with the animal system we previously investigated. After completion of the experiments, our intent is to write a manuscript for submission to The American Biology Teacher for review and eventual publication. We believe that this investigation would be suitable for a high school or college level introductory level class, providing students with an experience in the experimental manipulation of a plant system. This would address two needs:

1) to provide students with systems that they can experimentally manipulate, and
2) to provide a plant system as a model system for experimentation, filling a need for more plant investigatory systems in the high school and college curriculum (the majority of published experiences being animal in nature).

Experimental Approach

Representatives of two plant systems will be utilized – a dicot, narcissus (*Narcissus* sp.), and a monocot, maize (*Zea mays*). We want to determine if the effects of alcohol exposure, if any, are mirrored in these two representatives of the separate domains of flowering plants, or if the responses between these two taxa might differ.

Three alcohols (methanol, ethanol, and isopropanol, differing by the sequential addition of a methyl moiety) will be tested, in four separate concentrations – 2.5%, 5%, 7.5%, and 10%. These are the concentrations used in the past *Drosophila* experiments, and will allow for comparisons between the animal and plant experiments.

Ms. Lander's responsibilities will include:

a) setting up the growth containers with the plant materials,
b) preparing the different alcohol solutions
c) providing the plants with appropriate amounts of either water or alcohol solution during the course of the experiment, and
d) recording the growth of the plants at weekly intervals.

Learning Objectives

The LO's for Ms. Lander in this effort are as follows:

1) To gain experience in setting up a controlled experiment, with appropriate replicates.
2) To maintain living plants undergoing specific treatments, and to maintain the specific treatments throughout the duration of the experiment.
3) To gain experience in collecting data at regular intervals from the control and treated plants, and to maintain reliable records of the data.
4) To assist in manuscript preparation, specifically in the preparation of figures, graphs, and tables, that would be appropriate for the submitted manuscript.
5) To gain an understanding of the process of publication of scientific results, including manuscript preparation and revision prior to final acceptance.

This summer (2015) Ms. Lander conducted some preliminary experiments to determine the best way to maintain the plants in alcohol solutions. She first attempted a gravity feed system that was problematic due to the evaporation of alcohol solutions. She then settled on clear plastic drinking cups containing pebbles as a substrate, that would allow for the monitoring of alcohol solution levels without disturbing the plants, and would also minimize evaporation. She is now prepared to complete the experiment this semester. We anticipate submitting a manuscript for publication by November 2015.

BIO598
Independent Research
Ms. Rachael Lander
Spring 2016

Introduction

Arisaema is a cosmopolitan genus, represented by 181 described species. The center of diversity is Asia (primarily China and japan), with additional species found in east Africa and eastern North America. In the eastern United States the genus is represented by two species: *A. triphyllum* (Figure 1) and *A. dracontium* (the green dragon). There has long been speculation that those taxa included under the designation of *A. triphyllum* could represent a "species complex" (Trieber, 1980). Recent investigations by Dr. R.F. Naczi, the Arthur G. Cronquist Curator of North American Botany at the New York Botanical Garden, have indicated that there could be an undescribed, distinct species nested within the *A. triphyllum* complex. Dr. Naczi has gathered an extensive morphological data set that supports this contention (R. naczi, unpubl.). We now propose to generate an independent data set to evaluate this systematic hypothesis.

Figure 1. Image of *Arisaema triphyllum*, showing the inflorescence (spathe an spadix) that is diagnostic for the family.
https://upload.wikimedia.org/wikipedia/commons/8/83/Arisaema_triphyllum.jpg

Experimental Approach

We have entered into a collaboration with Dr. Naczi. He has provided us with 20 plant taxa (Table 1), from which we have isolated and purified genomic DNA, suitable for the PCR. We will generate nucleotide sequences for selected loci and utilize these as an independent data set in a phylogenetic context to evaluate Dr. Naczi's hypothesis.

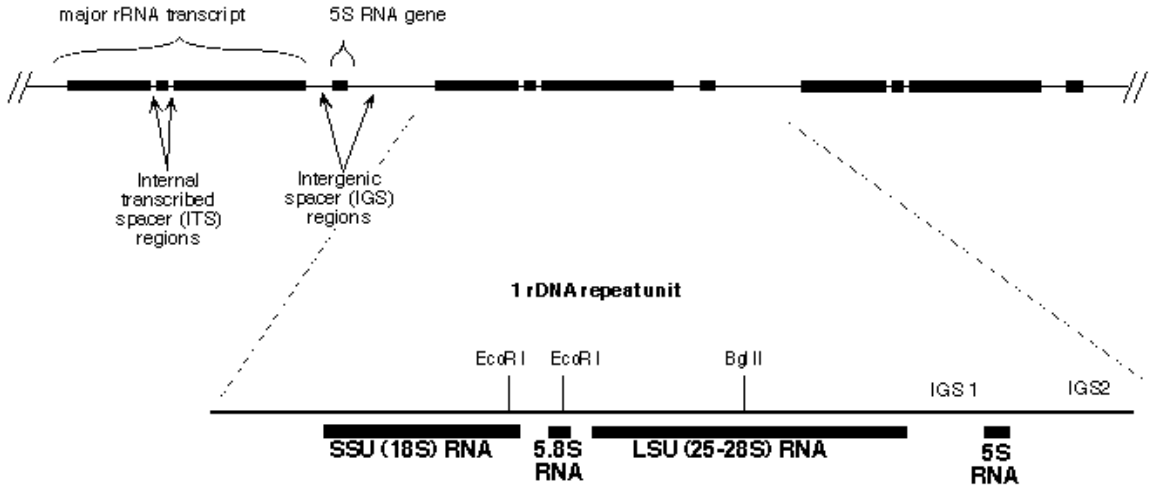
Table 1. List of the taxa provided to us by Dr. Naczi. The lab number is our designation; the RFCN number is Dr. Naczi's collection number.

#	RFCN#	taxon
A1	12986	*triphyllum*
A2	13019	*triphyllum* ssp. *pusillum*
A3	13058	*pusillum*
A4	12901	*triphyllum*
A5	12966	*triphyllum*
A6	*s.n.*	*arisaema* (non-glaucous)
A7	*s.n.*	*triphyllum* (glaucous)
A8	12957	*arisaema* (non-glaucous)
A9	13041	*triphyllum* (glaucous)
A10	*s.n.*	*stewardsonii*
A11	14103	*triphyllum*
A12	14141	*stewardsonii*
A13	14140	*stewardsonii*
A14	14128	*triphyllum* (glaucous)
A15	14139	*triphyllum*
A16	88324	*quinatum*
A17	73343	*quinatum*
A18	173 829	*quinatum*
A19	15479	no notation
A20	*s.n.*	*forgesii*

This effort will utilize two nuclear loci: the internal transcribed spacer, a region that separates the 16S and 23S ribosomal RNA genes, and the 26S ribosomal RNA gene (the LSU locus) (Figure 2). We have primers for these two loci (three sets in all), that have proved effective in our hands in other plant taxa. There are also a

number of archived sequences in the NCBI Nucleotide database for *Arisaema*, that could be utilized for testing phylogenetic hypotheses.

Figure 2. Diagram of the nuclear ITS locus and the 26SrDNA locus (LSU) that will be utilized as PCR targets n this investigation.



http://sites.biology.duke.edu/fungi/mycolab/primer5.gif

Ms. Lander's efforts will entail the following:

1) Data mine the NCBI nucleotide d-base for ITS and 26S nucleotide data files that would prove useful in our studies. She will consult with Dr. Naczi on the choice of informative taxa.
2) Download sequence files (fasta format) for downstream phylogenetic analyses, properly labeling the files and building a local database.
3) Generate, from Dr. Naczi's *Arisaema* DNAs, a set of PCR amplicons for the ITS and 26S loci.
4) If time allows, begin to edit, proofread and trim sequences (though the software program Sequencher v. 5.2) generated by her and our colleagues at EKU involved in this effort.

Anticipated Outcomes

Dr. Naczi's field investigations and prior morphometric analysis does indicate that there is a distinct (e.g. lineage-specific) taxon embedded within the *A. triphyllum* complex. We predict that the molecular data set, when input into a phylogenetic analysis pipeline, will provide congruence for Dr. Naczi's hypothesis. We will work with Dr. Naczi to bring this effort to publication in an appropriate journal.

Reference

Treiber, M. 1980. Biosystematics of the *Arisaema triphyllum* complex. Ph.D. dissertation, Univ. of North Carolina/Chapel Hill.

Appendix 3

Interviews with genome scientists conducted by Rachael Lander in 2015, as a component of her Honors Thesis. The individuals interviewed were Drs. Dan Howe of the Gluck Equine Research Center of the University of Kentucky, Mark Farman of the Department of Plant Pathology of the University of Kentucky, Ben Busby of the National Center for Biotechnology Information if the NIH, and Elaine Ostrander of the National Human Genome Research Institute of the NIH.

Dr. Daniel Howe
Gluck Equine Research Center
Department of Veterinary Science
University of Kentucky
http://www2.ca.uky.edu/gluck/HoweDK.asp

Dr. Howe studies a group of single-celled organisms called the Coccidia, which are primarily single celled parasites of humans and other mammals. These include the human parasite *Toxoplasma gondii* and other parasites, named *Eimeria*, *Neospora*, and *Sarcocystis*, that are responsible for some animal diseases. He is specifically studying *Sarcocystis neurona*, the parasite responsible for the disease equine protozoal myeloencephalitis (EPM). His training is in parasitology, and he is a Professor in the Gluck Equine Center. I interviewed Dr. Howe in October 2015 at the Gluck Center.

Q) What first attracted you to a career in research in animal parasites?
A) I was interested in parasitology as an undergraduate, and was involved in an undergraduate research project that really got me interested. I pursued a Master's degree at Western Illinois University (a lot like EKU), studying the flukes in eyes of chickens, and then earned my doctorate from Purdue University.

Q) What is the mission, or purpose, of the Gluck Equine Center? How was it first established?
A) The Center was established by a group of benefactors with the idea of helping find cures for equine diseases. Many of the benefactors were involved with the local thoroughbred industry. Our overarching goal is to conduct research to the

benefit of the health and welfare of the equine, and to train scientists in this area. We currently have 24-26 graduate students pursing the Ph.D. in Veterinary Health.

Q) Can you describe to me the clinical effects of equine protozoal myeloencephalitis, the disease that you study?
A) It is basically a neurological disease, and difficult to diagnose. An infected horse might exhibit a range of symptoms, from difficulty in walking, irregular behavior, to possible death, in rare cases. Infected animals will exhibit weight loss, and the owners could experience economic loss. So it is a real equine health concern.

Q) What sort of research efforts is your group conducting on this parasite?
A) We are trying to understand how the parasite recognizes the appropriate host, how it evades the host immune system, and how it is able to be transmitted from one animal to another. We are also interested in how the parasite survives outside the host animal in the natural environment, especially during times of environmental stress.

Q) How does genomics help you understand this parasite, and ways to prevent host infection by the parasite, and ways to better treat infected horses?
A) We are carrying out a series of genomic studies to better understand the genetic composition of the parasite, how it recognizes and infects its host, and evades the host immune system. We have generated an Expressed Sequence Tag library, a set of partial sequences of all the genes that are expressed in the parasite. We are trying to determine those genes that allow the parasite to recognize its animal host, and those genes that initiate the infection process. We are also developing diagnostic tests, using the proteins produced by the parasite, as ways to detect the presence of the parasite in equine blood and cerebrospinal fluid.

6) Does the Gluck Equine Center sponsor research externships for Veterinary Medicine students? Are there postdoctoral research opportunities for DVMs at the Gluck Equine Center?
A) Yes, visiting students are welcome for summer experiences. Many of our labs have opportunities for short-term research projects. And yes, there are post-DVM research opportunities as well. Close to half of our students in pursuit of the Ph.D. have the D.V.M. They are seeking careers in research, and want to couple their clinical experience with research opportunities.

Dr. Ben Busby
National Center for Biotechnology Information
National Institutes of Health
Bethesda, MD
http://www.ncbi.nlm.nih.gov/research/staff/busbybr/

Dr. Busby is a member of Dr. Eugene Koonin's group in Evolutionary Genomics at the NIH (http://irp.nih.gov/pi/eugene-koonin). Ben is a computational biologist, and his primary responsibility is to train scientists in using the computational tools needed to analyze genomes. He also develops software tools for biologists. Dr. Busby visited the University of Kentucky for a seminar in September 2015, and I had an opportunity to interview him.

Q) What are the major technical challenges facing the field of genomics at the present time? That is, what are the technical challenges facing scientists in this area?
A) The biggest issues are ones of statistical normalization. Sample size, accounting for variation, using appropriate statistical measures on one's data set. RNA sequencing is especially in need of more attention in this area. Batch effects are a real issue, and many people are not aware of the issues. For example, processing control and treatment samples together, rather than on separate days.

Q) What sort of training in genomics do you, and the NCBI, do for scientists and other investigators who are new to the field?
A) Our biggest effort is the Fundamentals of NCBI course. This takes our students through the basics of the available tools on the NCBI website, and provides them with some familiarity with the platform. It is a guided tour, so to speak.

Q) What genomics resources are available at the NCBI?
A) A lot – probably our most utilized is the "Genomes" page. Also, our 1000 Genomes Browser is quite popular. These are expanding and changing rapidly, so you really need to keep up with developments. These are also among our more popular resources.

Q) What technical advances in genomics do you see occurring over the next five years or so?
A) Technically, sequence reads are going to get longer, so the problem with sequence assembly won't be as onerous. The Pac Bio system is really revolutionizing the field. Pac Bio uses an advanced technology to sequence single molecules, which eliminates a lot of the technical variation from amplification steps, that you see with Illumina.

Q) In terms of medical science, and veterinary science, what are the major contributions (diagnostics, treatments) that genomics is currently providing?
A) Determining diagnostic SNPs through GWAS studies is probably the biggest contribution. Then, through comparative genomics, determining those mutations responsible for specific cancers, such as prostate and lung. And from this, determining better therapeutic approaches. For example, using prostate cancer drugs on brain cancers, if the brain cancer has the same mutation as the prostate cancer. Some major advances are taking place, in terms of improved therapeutic approaches.

Q) The One Health Initiative – how invested is the NIH in this? Are genomics making an impact or making a contribution to this effort?
A) Oh, very much so! Zoonotic diseases are a major concern – which animal populations are serving as reservoirs for pathogens? The NIH also has a Veterinary Diagnostic Lab, to examine animals for possible diseases, pathogens, and parasites.

Dr. Mark Farman
Department of Plant Pathology
University of Kentucky
Lexington, KY
http://www2.ca.uky.edu/agcollege/plantpathology/FARMAN/farman.htm

Dr. Mark Farman is a fungal biologist, who studies the plant pathogen *Magnaporthe oryzae*, a fungus that infects rice plants throughout the world. Dr. Farman uses genomic approaches to study the origin and evolution of different strains of *M. oryzae*, and tries to determine why some strains are more virulent (deadly to the plant host) than others. He is also interested in which genes in the fungus are involved in recognizing the host species, and which genes allow the fungus to avoid host defense mechanisms. Dr. Farman is former Director of the Advanced Genetics Technology Center at the University of Kentucky, and is currently Associate Director of UK- HealthCare Genomics. I interviewed Dr. Farman in the Department of Plant Pathology in April 2015.

Q) Why would one want or need to sequence an entire genome? Why not look at individual, specific genes of interest, rather than all the genes? It would seem more practical to focus on specific genes, rather than all the genes in a genome.
A) In the past that is what one would do, that is, sequence individual genes rather than entire genomes, due to limitations in the DNA sequencing technology. But that was a very slow process, one gene at a time. With genomics, one can sequence all the genes, and then do comparisons between and among strains and

species, and try to see patterns of association between specific genes and specific traits. Such as virulence.

Q) In terms of disease, what could the sequence of a gene tell you regarding that specific disorder, such as cancer? Are there mutations that are specific for specific diseases?

A) Yes there are. So the challenge is finding out which mutations are responsible for specific diseases, such as cancer. Many mutations in our genome are harmless, and do not have a negative effect on the phenotype. But a few can be serious, and even fewer can be the causes of specific diseases. These are the ones that we are most interested in.

Q) How can you tell if a gene is mutated in such a way that it might be responsible for a disease, such as cancer? How do you associate a specific mutation with a specific disease?

A) So this is where it gets really tricky. You need to examine a large pool of patients with the disorder, and scan their genomes for those mutations that are common to all those individuals. Then you need to determine if the mutations occur in or around genes that could be responsible for cancer, or the disease of interest. This process can take quite a long time, and is not at all easy.

Q) How different are the same genes from different people? Are all human genes of the same type pretty much all the same sequence, or are there differences among different people?

A) There is a remarkable amount of variation in the sequences among different individuals. This is what makes working with human genomes so difficult – the variation among them. It is difficult to tease apart the random variation from those disease-causing variants.

Q) How does one analyze a genome, for example, how does one determine the sequence and location of specific genes?

A) Computers – it is all driven by sophisticated algorithms and software programs. The program "reads" the nucleotide sequence of the gene, and then determines which protein they encode.

Dr. Elaine Ostrander
Chief and NIH Distinguished Investigator
Cancer Genetics and Comparative Genomics Branch
National Institutes of Health
Bethesda, MD
https://www.genome.gov/12513335/ostrander-group/

Dr. Elaine Ostrander is a Distinguished Investigator and Chief of the Cancer Genetics Branch of the National Human Genome Research Institute of the National Institutes of Health. Dr. Ostrander is one of the leaders in the field of canine genomics, and her laboratory is currently focusing on the genes involved in prostate cancer, in both dogs and humans. Her goal is to use the canine genome as a model system for examining the genes and mutations responsible for a set of human disorders, such as cancer. I interviewed her by phone in Dr. Calie's office in October 2015.

Let me first ask you some general questions, if I may:

Q) What first motivated you to pursue the sequencing of the canine genome?
A) We knew from earlier studies that the dog genome was approximately the same size as the human genome, so it wouldn't be too difficult to manage. Dogs and humans share many similar disorders, and mapping the responsible gene in one species could lead to mapping the homologous gene in the other. Human genomes are difficult to work with, so the dog genome could give us a bit of a "short-cut" in gene mapping.

Q) Why the dog? What made this specific animal group so attractive to your team's research interests?

A) Several things. First, the highly inbred lines made this an attractive system due to reduced genetic variation. Second, humans and dogs share many of the same disorders and diseases, so the cause of a disease in dogs could be similar to the cause in humans. And third, we had many generous offers from owners and breeders for DNA samples, which made our task much easier.

Q) What questions did you and your team intend to address by sequencing the canine genome?

A) Well, aside from detecting specific disease causing genes, what makes a dog a dog? For example, what genes contribute to various breeds' behavior? What genetic changes have led to the instinct to retrieve, or the instinct to herd? How many genes are responsible for the breed differences, such as height, jaw structure, and coat color?

Now, I'd like to ask about the research itself:

Q) I know that scientific discovery can be full of surprises. What were the major surprises that came to you from this effort?

A) Perhaps among the biggest surprises were how few genetic differences there are among different breeds. It can be as little as a single gene, or allele, difference that separated one breed from another. The phenotypic differences are huge, but the genotypic differences are often very small.

Q) What sort of impact of the Canine Genome effort do you see in the future for Veterinary Science and Medicine? Will your discoveries lead to improved diagnostic methods or treatments for dogs? For other animal species?

A) Well, I doubt that we will begin to use diagnostic tests on dogs, due to the costs and the lack of insurance coverage. Just too expensive. But we have been approached by different breeders who want to determine which animals are the best breeding stock, and which are best not used for breeding purposes, due to their carrying specific undesirable alleles.

Questions about the methodology:

Q) How large a team of investigators was needed to carry out the genome sequencing?

A) This effort did involve several labs around the world, each group taking on a specific region of the genome. The draft of the human genome gave us a good reference genome to work with, in terms of sequence assembly.

Q) How long did the genome assemblies take? Were there any issues unique to the canine genome (such as repetitive elements) that complicated genome assembly?

A) We used a boxer as our model system, as this is the most inbred dog breed and has very low variability. There were some issues with repetitive sequences, but these were of a level typical for a mammalian genome. So no real issues there.

A bit about you, if I may:

Q) What first attracted you to science?

A) I was always interested in learning about the world around me. I enjoyed the discovery process about science, and genetics always appealed to me with the logic and intuitive process. It seemed to be a natural fit for me.

Q) What was your path to the NIH like? How did you get to where you want to be as a scientist, especially in your years as a student?

A) I trained at the University of Washington, which has an excellent program in human genetics. I then did my postdoc at Harvard, where I really gained great experience in using some of the tools of genome mapping. I then returned to Seattle for my faculty position, and then onto the NIH where I now am.

A big question – where do you see genome science going in the next 5 to 10 years? What impacts will it have on Vet Medicine?

A) I see us developing better drugs for specific canine ailments, better breeding programs that will hopefully reduce the occurrence of many of the common maladies associated with many dog breeds. Hopefully we will see, over time, an improvement in the genetic health of dogs, with less burden to their owners of specific ailments. Time will tell.