UNDERSTANDING THE GENETICS OF AGING: A CANINE MODEL

A Dissertation

by

SARAH CHRISTINE CANTERBERRY

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2006

Major Subject: Genetics

UNDERSTANDING THE GENETICS OF AGING: A CANINE MODEL

A Dissertation

by

SARAH CHRISTINE CANTERBERRY

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Keith Murphy
Committee Members	Kimberly Greer
	Ann Kier
	Bhanu Chowdhary
Chair of Genetics Faculty:	James Wild

December 2006

Major Subject: Genetics

ABSTRACT

Understanding the Genetics of Aging: A Canine Model. (December 2006)

Sarah Christine Canterberry, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Keith E. Murphy

As life expectancy in the United States increases each year, the percentage of the population that is comprised of aged individuals rises also. Researchers expect the largest increase in population to occur in the segment consisting of individuals 85 and older. Thus, investigations of the aging process, with the goals of further extending average life expectancy and improving the quality of life for aged individuals, have become increasingly important to our society.

To better understand the genetics of aging, we elected to utilize another model organism, the domestic dog. The benefit to this work is that breeds exhibit extreme, natural variation in life expectancies. Here I report my contributions towards establishing the dog as another model organism for investigations of the aging process.

Multiple linear regression analysis was carried out to determine the association between life spans and breed size in the dog, based upon data derived from the American pet population. A negative correlation was observed between both height and longevity and between weight and longevity with weight being the significant predictor of life span.

Fifty-four genes implicated in the aging process were mapped to the canine genome. These genes were selected because of their demonstrated contribution to longevity in other organisms or based upon their proximity to a marker, D4S1564, on human chromosome 4.

Four genes that are associated with dwarf mice and extended life span were analyzed in nine dog breeds of varying sizes and life expectancies. Fifty-three polymorphisms were discovered in *Ghr*, *Ghrhr*, *Pit1*, and *Prop1*. Thirteen ancestral SNPs were discovered in which both alleles were found in every breed. In *Ghrhr*, a transition mutation was found that changes the amino acid sequence as well as the function of the protein and is statistically significant ($p=4.8 \times 10^{-6}$) when large dogs are compared to medium-sized breeds, but not when they are compared to small breeds ($p=0.001$). This SNP warrants further investigation in additional dogs and breeds.

DEDICATION

I would like to dedicate this dissertation to my parents, Susie and Al Zienko. Their support at every step in my journey of life, and the important lessons they taught me along the way, have shaped the person that I am. Also, for my husband, Bubba, who has helped me keep sight of my goals and maintain a steady course; without him I would be lost.

ACKNOWLEDGEMENTS

I would like to thank Dr. Keith Murphy, my advisor, for his investment of time and effort into my graduate career. I would like to thank the members of my committee, for providing me with invaluable advice, Dr. Ann Kier, Dr. Bhanu Chowdhary, and Dr. Kimberly Greer. I would like to especially thank Dr. Greer, who worked closely with me on the Canine Aging Project during her time in the Canine Genetics Research Laboratory. I also thank my fellow graduate students in both the Canine Genomics Laboratory and in the Program of Genetics at Texas A&M University, who all assisted me in many different ways. Last, but certainly not least, I would like to thank the owners and dogs that participated in this work; without their cooperation, it would not have been possible.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Aging and Longevity

Life expectancy in the United States is increasing each year; as a result, the percentage of the population that is comprised of aged individuals is on the rise as well, making investigations of the aging process highly beneficial to our society. The average life expectancy has climbed from less than 50 years at the onset of the 20th century to today's estimate of nearly 80. Researchers expect the largest increase in population to occur in the segment consisting of individuals over the age of 85 years followed by those aged 65 and older. Specifically, persons 85 and above are expected to number 19.4 million, or 4.8% of the total population, by the year 2050, while the portion of the population 65 and older will double by 2030 to 20% of the population, or 70.3 million individuals. Due to this rapid population growth in the elderly population, investigations of the aging process are necessary to postpone the onset of, and develop more effective treatments for, aging-associated diseases (National Institute on Aging, 2006).

Previous research has led to multiple theories of the aging process. While these theories are complex, they are not mutually exclusive. Each theory tries to address the many factors that contribute to aging, a process that, of course inevitably culminates in death. These theories can be classified as evolutionary, molecular, cellular and systemic (reviewed in Weinert and Timiras, 2003). Evolutionary theories include the mutation

---

This dissertation follows the style of Mammalian Genome.

accumulation theory, the disposable soma theory and the antagonistic pleiotropy theory (reviewed in Weinert and Timiras, 2003). The mutation accumulation theory states that mutations which only affect an organism after it has reproduced, will not be selected against and therefore remain in the population at a steady frequency (reviewed in Gavrilov and Gavrilova, 2003). The disposable soma theory indicates that once reproductive age has passed, the organism is disposable. This theory is supported by scientists studying organisms exhibiting diverse life spans that correlate with environmental hazards and exposure to predation (reviewed in Weinert and Timiras, 2003). The theory that some genetic differences necessary and beneficial early in life are the same factors that are, later in life, quite detrimental to longevity, is the basis for the antagonistic pleiotropy theory (reviewed in Gavrilov and Gavrilova, 2003). Many believe that these mutations are preserved in the genome because of their importance early in development of organisms. Gene regulation, codon restriction, error catastrophe, somatic mutation, and dysdifferentiation theories fall into the molecular theories of aging category (reviewed in Weinert and Timiras, 2003). Gene regulation theories suggest that as an organism ages, gene expression changes are responsible for the deterioration that accompanies aging (reviewed in Weinert and Timiras, 2003). Inaccuracies in translation and gene expression, which may lead to production of abnormal proteins, are the bases of codon restriction and error catastrophe theories, respectively (reviewed in Weinert and Timiras, 2003). The somatic mutation theory holds that as an organism ages, molecular damage accumulates, particularly to the genetic material. The dysdifferentiation theory postulates that as this damage is

accumulated, regulation of gene expression is altered (reviewed in Weinert and Timiras, 2003). Included in the cellular theories of aging are the cell senescence/telomere theory, free radical theory, wear-and-tear theory and apoptosis theory (reviewed in Weinert and Timiras, 2003). The cell senescence/telomere theory states that as an organism ages, the length of the telomeres shorten and/or the cell senesces due to stress. The free radical theory, or oxidative stress theory, suggests that free radicals, the by-products of metabolism in aerobic organisms, cause significant damage to cellular components such as lipids, proteins and genetic materials and that the accumulation of such damage leads to aging. The wear-and-tear theory that the accumulation of normal injury, and the apoptosis theory that programmed cell death are responsible for aging, are the remaining theories that are associated with cellular biology (reviewed in Weinert and Timiras, 2003). The system theories of aging are 1) that as an organism ages the neuroendocrine control of homeostasis is altered, 2) that an overall decline in immune function results in increased rate of autoimmune disorders and 3) that each organism has a fixed metabolic potential (reviewed in Weinert and Timiras, 2003). Many scientists believe that the key to solving the aging puzzle is to determine how these theories fit together. Only by understanding all aspects concerning the biology of aging as set forth by the various theories will we be able to see the big picture and fully dissect the aging process.

Genetics of Aging

Many of the above described theories are strongly anchored in genetic mechanisms that are involved in aging and longevity. Numerous studies have indicated

that there are significant genetic factors that contribute to longevity, and heritability of average life expectancy has been estimated to be as high as 30% (Ljungquist et al., 1998). Understanding how genes function in the aging process may enable progress to be made not only in the extension of life span but also in improving the quality of life of aged individuals. Studies of genes that govern aging have been conducted using several species, including the nematode, fruit fly, mouse, and human.

Caenorhabditis elegans

In normal laboratory settings, the average life span of *C. elegans* is three weeks, but there are mutants that live much longer (reviewed in Warner, 2003). Mutations that affect multiple systems have been found to extend the worm's life expectancy. Such mutations in genes that play roles in the insulin-signaling pathway, caloric restriction, oxidative stress, signal transduction and gene expression have been implicated in life span extension in the worm. Genes that have been found to lengthen life span which are involved in the insulin signaling pathway are *age-1*, *daf-2* and *daf-16*. Worms with mutations in *daf-2*, which is similar to insulin/IGF-1, live twice as long as wild-type (Kenyon et al., 1993; Larsen et al., 1995) and upon further manipulation of the reproductive system, these worms are as healthy and active at ages that correspond to that of 500 year old humans (Arantes-Oliveira et al., 2003). A gene that encodes a phosphatidylinositol-3-OH kinase-like protein, *age-1*, is another gene that may control life span in *C. elegans* (Morris et al., 1996). Additionally, mutations in *eat* that cause defects in pharyngeal function have been found to extend life span, possibly by mimicking caloric restriction (Lakowski and Hekimi, 1998). Mutations in *clk-1* render

worms unable to synthesize coenzyme Q_9 and confer a longer life expectancy than that of wild-type worms (Lakowski and Hekimi, 1996). These findings led investigators to the discovery that removing coenzyme Q_8 from the diet of non-mutant worms had similar effects possibly by decreasing the production of reactive oxygen species (Larson and Clarke, 2002). Increases in expression of *old* genes, homologous to receptor tyrosine kinases, extends life span and decreases susceptibility to stress by interruption of signal transduction (Murakami and Johnson, 1998). Alterations of gene expression can also be induced by mutations in *sir2*, a histone deacetylase (Tissenbaum and Guarente, 2001). These discoveries have not only proved that single genes can affect longevity, but also provided insight to the basic mechanisms of aging, and prompted studies of similar pathways and genes in more complex systems.

Drospophila melanogaster

Long lived mutants of *Drospophila melanogaster*, another common model organism, have been investigated in great detail. The first flies to exhibit extended life spans were those with mutations termed *methuselah* (Lin et al., 1998), a gene that is a member of the secretin family (reviewed in Helfand and Rogina, 2003). They lived approximately 35% longer and were more resistant to several forms of stress including oxidative stress, high temperatures and starvation (reviewed in Warner 2003). Mutations in *Indy*, a dicarboxylate transporter, allow flies to live almost twice as long as wild-type counterparts possibly by inducing a state of caloric restriction (Rogina et al., 2000). Two genes in the insulin-signaling pathway increase life span in female fruit flies, the insulin-like receptor (*Inr*) and the insulin-like receptor substrate protein (*chico*) (Clancy

et al., 2001; Tatar et al., 2001). Both males and females with mutations in these genes are dwarf (reviewed in Helfand and Rogina, 2003). There have been several genes found in *Drosophila* that are involved in stress resistance that can alter life expectancy. When the homolog to human *SOD* is overexpressed, the result is a 40% to 50% increase in life span (reviewed in Helfand and Rogina, 2003). Transgenic overexpressors of methionine sulfoxide reductase A are resistant to oxidative stress and remain active much longer than wild-type flies (Ruan et al., 2002). Also, transgenic fruit flies that overexpress *PCMT*, a gene responsible for repairing the damage that results from oxidative stress, exhibit extended life spans (Chavous et al., 2001). Increased life spans have also resulted from mutations in the *rpd3*, which encodes a histone deacetylace that is important in chromatin structure and regulation of gene expression (reviewed in Helfand and Rogina, 2003). While *Drosophila* is a much more complex model system than *C. elegans*, it is very far from the human on the evolutionary scale.

Mus musculus

The mouse is an exceptional mammalian model organism for genetic studies, and has been utilized by many scientists interested in aging. Currently there are sixteen strains of mice that have life spans that deviate from the average due to genetic alteration. Four of these strains are dwarfs. Ames mice are defective in *Prop1* and Snell dwarf mice harbor mutations in the *Pit1* gene; both strains are defective in pituitary development and exhibit 40-50% extension in life expectancy (reviewed in Quarrie and Riabowol, 2004). Mice which have mutations in *Ghrhr*, known as Little mice, live approximately 25% longer than wild-type, while Laron mice are *ghr*-knockouts that

have up to a 55% increase in life span (reviewed in Quarrie and Riabowol, 2004). These genes play roles in the insulin-signaling pathway (reviewed in Quarrie and Riabowol, 2004; reviewed in Warner, 2003). There are four additional strains that have mutations in genes included in the insulin-signaling pathway, 1) GH Ig mice overexpress GH and have decreased life span, 2) IGF1-R mice that exhibit an overall increase in life expectancy of 25%, although the extension of life span in not significant in the males, 3) $p66^{shc}$ mice have 30% longer lives, and 4) $IR_{adipose}$ mice have an 18% extension in life span (reviewed in Quarrie and Riabowol, 2004). Both p66shc and $IR_{adipose}$ mice are resistant to oxidative stress as are the following strains: Thioredoxin mice that live up to 35% longer and Bcl-2 tg DC mice that have increased longevity of their dendritic cells (reviewed in Quarrie and Riabowol, 2004). Alternatively, Peroxiredoxin mice are more susceptible to oxidative stress and have significantly reduced lives (reviewed in Quarrie and Riabowol, 2004). Three mutants that have significantly reduced life spans have been described. These are the Klotho mice, p53 mut mice and XPD TTD mice that have life spans shortened to less than 100 days, by as much as 17% and by approximately 50%, respectively (reviewed in Quarrie and Riabowol, 2004). Lastly, mutations in the genes, *UPA* and *SIRT1* extend the lives of mice by appetite suppression and mediation of caloric restriction (reviewed in Quarrie and Riabowol, 2004; reviewed in Warner, 2003). Although the mouse is far more similar to man than is the fruit fly or nematode, and can provide many useful insights, this model organism is still quite evolutionarily distant from the human.

Homo sapiens

The search for longevity genes in the human has been challenging for several reasons, two of which are long life expectancy, and highly heterogeneous populations. Thus, many genes that correlate with excessive old age in one population may not necessarily correlate in another population (De Benedictis et al., 2001). More genes have been discovered that are deleterious than have been found to extend human life span (reviewed in Lao et al., 2005). Numerous genes that are critical to diverse physiological processes, such as metabolism, cardiovascular health, oxidative stress, cancer susceptibility, coagulation, bone mineralization, and chromatin structure have been linked to varied life spans in multiple populations (De Benedictis et al., 2001; Heijmans et al., 2000). However, the results of these studies may indicate a decreased susceptibility to disease and not an extension of life span. A few protective alleles have been discovered, however. There are a few variants of the apolipoproteins E and C-III, and IGF-IR genes as well as HLA haplotypes that are found in higher frequency among centenarians (reviewed in Lao et al., 2005). Additionally, the mitochondrial genome is suggested to play a role in determining life span in humans because it is quite susceptible to oxidative damage (reviewed in Lao et al., 2005). While studies of single loci or haplotypes are most numerous, linkage studies have also been performed. One such study of siblings of centenarians identified marker D4S1564 on human chromosome 4, which correlated with a 4-fold increased ability of these siblings to live at least into their early nineties when compared to siblings of non-centenarians (Puca et al., 2001; Perls et al., 1998). It is estimated that a locus near D4S1564 is responsible for approximately

1.65-fold of the overall increased ability to reach extreme old age. Additionally, a second study conducted on male fraternal twins supported these findings (Reed et al., 2004). However, due to low statistical significance (Reed et al., 2004), the affect that this locus has on human longevity is still widely debated. While studies of human populations are directly applicable, there are numerous drawbacks to these investigations, such as long life span, a shortage of experimental controls and lack of extensive family histories. Due to these drawbacks, another model organism for investigations into the aging process could provide further information about the aging process.

Canis familiaris

Although the dog is not a well established model of aging, some preliminary data have been gathered. Diet restriction in Labrador Retrievers increased median life span by nearly two years, from 11.2 years for those dogs fed a greater amount to 13.0 years for dogs fed 75% of their pair-mate control (Kealy et al., 2002). Caloric restriction has been successfully used in several model organisms, from rodents to primates, to prolong life. In a study of Swedish dogs, researchers determined that dogs belonging to different breeds age at different rates, and that it is inappropriate to consider them as having equivalent biological age (Egenvall et al., 2005). This study utilizes pet insurance databases and while this provides an extensive amount of information, it also admittedly ignores dogs more that ten years of age, because such dogs are no longer eligible for insurance. Cellular proliferative capacity was shown to be inversely related to breed size by Li and colleagues (1996). Investigations based upon information contained within

the veterinary medical database (VMDB) revealed that larger breeds have truncated life spans when compared to smaller breeds (Deeb and Wolf, 1994; Patronek, et al., 1997). These investigators, similar to those in Sweden, omitted a portion of the canine population by using the VMDB. However, in this instance, the healthy pet population was not included because these dogs were not referred to a veterinary teaching hospital. Interestingly, cross breed dogs do not live longer than their pure breed counterparts, therefore these do not exhibit hybrid vigor as one might expect (Bronson, 1982). Additionally, telomere shortening has been observed in canine fibroblasts *in vitro*, and differences in telomere lengths between different breeds have also been reported (McKevitt et al., 2002). The preliminary investigations reveal that the dog does not radically deviate from previous model organisms and current theories of aging demonstrated by studies that involve use of caloric restriction to extend life span and analysis of telomere lengths in dogs of a variety of ages and breeds. In addition, the dog presents a unique opportunity to study aging in an organism that naturally exhibits a wide range of life expectancy.

The Dog as a Model Organism

The domestic dog, *Canis lupus familiaris*, is a subspecies of wolf that diverged from its ancestor, *Canis lupus*, at least 15,000 years ago, and some evidence suggests a domestication event some 135,000 years ago (Leonard et al., 2006; Savolainen, 2006). Scientists are still working to further narrow this time frame. Since domestication of the dog, humans have utilized non-random breeding to develop more than 300 recognized

pure breeds. Most of these breeds have been developed within the past 250 years (Ostrander and Giniger 1997). Once breeds were established, inbreeding has been used extensively to fix the desirable physical and behavioral traits in dog breeds. Selective pressure has also presented itself through founder effects, popular sire effects, population bottlenecks, and surges in breed popularity which, in combination with intentional inbreeding, have resulted in over 450 hereditary diseases in the dog (OMIA 2006).

Approximately half of the aforementioned inherited diseases are also present in human populations. Historically, the dog has served as a useful model for many of the hereditary diseases found in both dog and human; often mutations in the homologous genes are causative in both species (OMIA 2006; Ostrander and Giniger 1997). Highly inbred populations, represented by breeds, are comprised of highly homogeneous individuals while comparisons between individuals from different breeds reveal sufficient heterogeneity to be useful in genetic investigations (Parker et al., 2004). Importantly, the canine genome was recently sequenced (Lindblad-Toh et al., 2005) and this revealed that the dog has a much higher level of sequence identity to the human than does the mouse (Jiang et al., 2005).

Animal husbandry practices commonly used by dog breeders also provide benefits to researchers. Due to breed registry practices, there is a limited amount of gene flow between dog breeds, resulting in highly homogeneous, inbred populations. Also, extensive pedigree information is readily available through owner records. The dog also has the additional benefit of a short generation time in which they most often produce litters rather than single offspring. Another advantage of using the dog as a model is that

sometimes it is not necessary to maintain a costly colony at the research institution. Instead, samples can often be obtained from the pet population through owner and breeder cooperation. These dogs also share a common environment with human owners, providing an automatic control for environmental exposures and alleviating concerns for both genetic drift of caged animals and for the overall welfare of the animals. Finally, it is critical to note that the dog enjoys medical surveillance second only to the human, with owners spending approximately $11.6 billion annually on veterinary related expenses for their dogs (Wise, et al., 2003).

Because of the close relationship humans have with their best friend, the dog, and the availability of extensive medical records, the clinical hallmarks of the aging process in the dog are well known. Additionally, the dog and its human caretakers exhibit many of the same ailments associated with aging. Not surprisingly then, the dog now provides a unique opportunity to study natural variation in life span within a single species to elucidate the genetic components of aging.

Specific Aims

Understanding the aging process is important not only from a scientific standpoint, but also from a social standpoint as the most rapidly growing segment of our population is that of the elderly. The ultimate goal of aging research is to extend life expectancy and improve the quality of life for aged individuals. To that end, this work is focused on establishing the dog as a model organism for aging research and determining genetic factors that play a significant role in life span of the dog, and by extension, the

human. To begin this line of investigation, it was necessary to first determine the extent of the inverse relationship between size and longevity in the pet dog population. Secondly, genes associated with the aging process in other systems were identified and mapped to the canine genome. Lastly, four genes that confer extended life spans in mice of smaller size were analyzed to determine if any sequence differences are responsible for the similar phenomenon which is observed in the dog.

CHAPTER II

STATISTICAL ANALYSIS REGARDING THE EFFECTS OF HEIGHT AND WEIGHT ON THE LIFE SPAN OF THE DOMESTIC DOG*

Overview

This study was undertaken to determine the association between life spans and breed size in the dog, based upon data derived from the pet population. Seventy-seven American Kennel Club (AKC) breeds were analyzed with data collected for more than 700 dogs. Multiple linear regression analysis was carried out with longevity as the dependent variable and height or weight as the independent variable. A negative correlation was observed between height and longevity ($r = -0.603$, $p < 0.05$), and between weight and longevity ($r = -0.679$, $p < 0.05$). Weight was the significant predictor of life span ($p < 0.001$), revealing that breeds smaller by weight generally live longer than heavier breeds. These data form the ground work for investigations of aging utilizing the dog as a model and provide owners with a quantitative method for predicting life span of dog breeds, thereby aiding in pet selection.

Introduction

Searching for susceptibility genes associated with multifactorial traits such as aging is problematic, and the criteria required for population dynamics and ideal

* Reprinted with permission from "Statistical analysis regarding the effects of height and weight on life span of the domestic dog" by Greer KG, Canterberry SC, Murphy KE, 2006. Research in Veterinary Science, in press. Copyright 2006 by Elsevier Ltd.

methodology by which to study them are still debatable (Freimer and Sabatti, 2004). Naturally, model systems have been employed to aid in the elucidation of factors contributing to human aging; however, many obstacles remain. For example, linkage analysis of human centenarians has led to investigation of certain chromosomal regions, but significant genetic heterogeneity, limited eligible participants, and lack of solid experimental controls reduce the effectiveness of this approach. Model organisms, including *Caenorhabditis elegans*, *Drosophila melanogaster*, *Mus musculus*, and non-human primates have also been studied, initially offering evidence of disparate mechanisms of aging, but it is apparent that there are indeed common molecular themes influencing longevity, even among invertebrate and vertebrate species. For example, longevity across these species is influenced by insulin signaling, stress resistance, the ability to repair cellular and macromolecular damage, chromosomal and nuclear structure, and caloric restriction (Warner, 2003). Although there are many differences in the way these organisms age, it is intriguing that key pathways are common to their aging processes despite their evolutionary distance. Collectively, examinations of these pathways in various model systems strongly suggest that genetic factors play a significant role in aging.

To determine the genetic factors that affect the process of aging, utilization of an alternative model organism, the dog, should contribute significantly to current knowledge regarding longevity and aging for several reasons. Genetic isolates, represented by breeds, are well-suited for identifying susceptibility loci because strict registration guidelines reduce gene flow among breeds. The resulting population

structure is highly independent and homogeneous, making population comparisons genetically informative. The breed structure is also useful because extensive pedigree information facilitates population-based genetic studies. Importantly, because medical surveillance of the dog is second only to that of the human (Wise et al., 2003), detailed medical records are also readily available. Furthermore, with the sequence assembly of the canine genome now available, it is known that the nucleotide sequence of the dog is more similar to the human than is the rapidly evolving murine nucleotide sequence (Kirkness et al., 2003). Lastly, in general, life expectancy of the dog is inversely related to body size (Li et al., 1996; Deeb and Wolf, 1994) and thus there is marked variation in life expectancy across breeds (Patronek et al., 1997). This natural variation is intriguing when compared to the study of longevity in other mammals, and access to genetically isolated populations with great differences in life spans provides an ideal and unique opportunity to study the genetic components critical to aging.

While data exist for a few breeds, information on longevity is lacking for the majority. Specifically, life tables were constructed for laboratory Beagles in 1981 (National Academy of Sciences) but not for other breeds. Another study addressed longevity and morbidity in a select number of giant and small breeds, but these authors had a primary interest in correlating specific diseases with dogs' age of death (Deeb and Wolf, 1994). That work and another completed in 1997 (Patronek et al.) utilized the Veterinary Medical Database (VMDB) which reflects mortality in veterinary teaching hospitals. Lastly, a Swedish study employed insurance claims as a method of data retrieval which admittedly ignores all dogs older than ten years because they are no

longer eligible for pet insurance (Bonnett et al., 1997). Therefore, unbiased data for life spans across breeds are lacking. This being so, the first step in evaluating the dog as a model for human aging and longevity was to gather data for pet dogs representing numerous breeds. All dogs were living as pets in owners' homes thereby reducing, if not eliminating, the aforementioned biases. Statistical analysis was performed on the height, weight, and longevity of the collected breeds to establish correlations therein.

Materials and Methods

Data collection

Height, weight, and medical information were collected on 718 individual dogs within 77 independent AKC breeds. Study participants were recruited at AKC dog shows and were restricted to those dogs eligible for AKC registration. This restriction was placed upon the participants to ensure known heredity for each dog in an attempt to eliminate extraneous factors that may influence or skew analyses. Owners were required to fill out detailed questionnaires for each dog, and the majority also contributed DNA via buccal swabs from the dogs. DNA was extracted from the collected cells utilizing a procedure previously described (Garcia-Closas et al., 2001), albeit with some minor modifications. Data from the questionnaire were entered into a computer database (Microsoft Access) for tracking, sorting, comparison, and analysis. Tables were assembled in Microsoft Excel and data exported to SPSS version 9.1 for statistical analysis (described below). Longevity information was collected from individual breed association records, generally the National breed club for the breed of interest. Because

this study is not a longitudinal study, the life spans of those dogs contributing data for height and weight were not required for participation; likewise, the heights and weights of those dogs contributing life span data were not required. Therefore, the data for height and weight utilized random sampling of the pet population and were collected independently of life expectancy data which was gathered from specific breed organizations. For analysis, the breed median life span value was assigned to all dogs belonging to that breed.

Statistical analyses

Statistical analyses were performed with SigmaPlot®, version 9.1 for Windows. Descriptive statistics were determined, followed by correlation analyses utilizing Spearman's rank order analysis (Belsley et al., 1980). Multiple linear regression analysis was utilized with a dependent variable, age, and two independent variables, height and weight (Weisberg, 1985) following forward stepwise regression analysis with independent variables height, weight, and breed group together with a dependent variable of age. A probability (p) value <0.05 was considered to be statistically significant.

Results

The 77 independent breeds have between 3 and 29 individuals per breed participating in the study. Generally, those breeds with higher numbers of participants are either more common breeds or have owners who are more interested in research. The dogs' heights at the withers range from 6.0 inches to 37.0 inches, while their weights range from 2.0 pounds to 196.0 pounds (Figure 1). As expected, the observations are not normally distributed as determined by the Kolmogorov-Smirnov test; their characteristics are shown in Table 1. The scatterplot matrix of age, median height, and median weight is depicted in Figure 2. Using Spearman's rank order analysis, a nonparametric test, age is most negatively related to weight ($r = -0.679$), and is also inversely related to height ($r = -0.603$). Both associations are statistically significant ($p < 0.05$). Additionally, and as would be expected, height and weight are significantly related to one another ($r = 0.919$, $p = 0.000$). Furthermore, the negative correlations are maintained when the individual dogs are separated into breed groupings (i.e. Herding, Hound, Non-sporting, Sporting, Terrier, Toy, Working), as would be expected since each group is composed of a size range of individual breeds (data not shown).

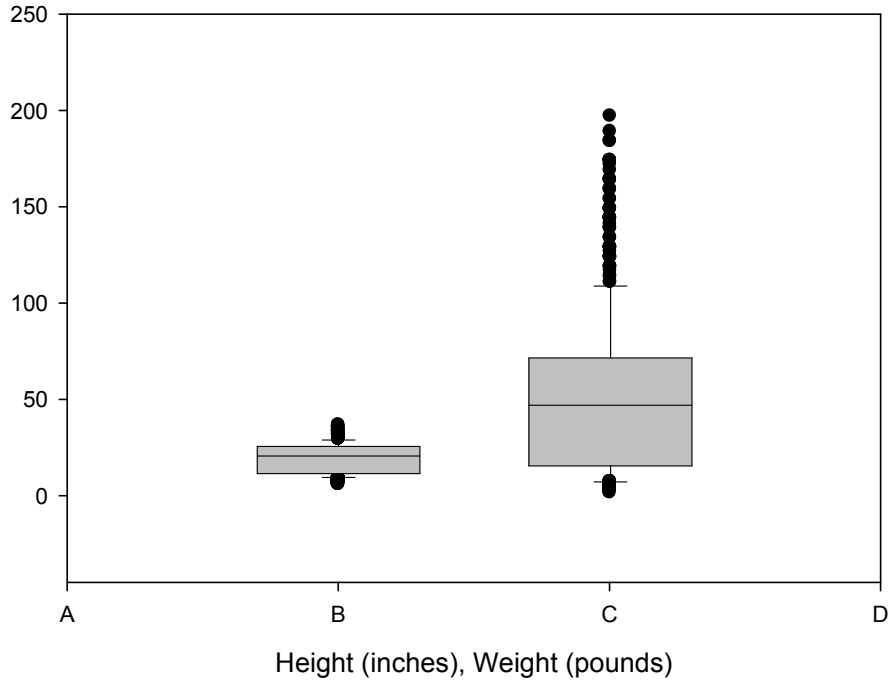

Figure 1: Heights and weights of sample population.
Box plot of the data set utilized for these analyses. Heights of the dogs at the withers range from 6.0 inches to 37.0 inches with N = 649. Weights of the dogs range from 2.0 pounds to 198.0 pounds with N = 701.

Table 1: Data description.

	Mean	Std Dev	Std Error	Range	Median	25%	75%
Height (in.)	19.669	7.677	0.301	31.000	21.000	12.000	26.000
Weight (lbs.)	51.708	40.701	1.537	196.000	48.000	16.500	72.250

Note: Collective information regarding the heights and weights of dogs in this study.

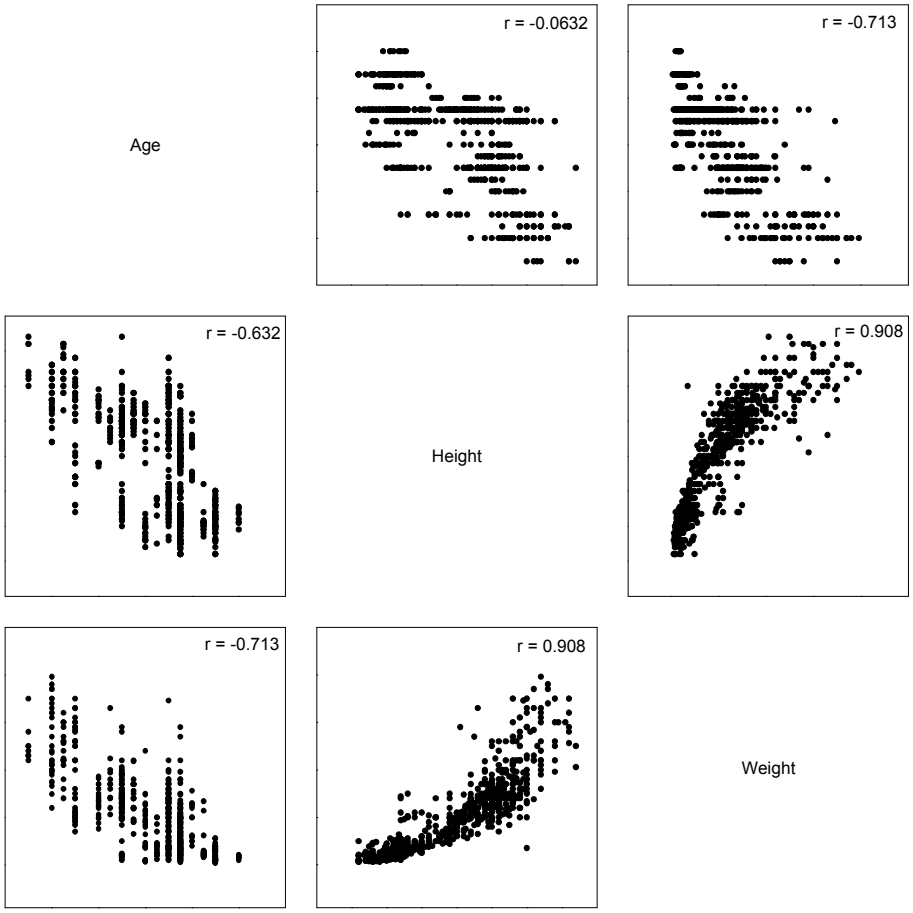

Figure 2: Correlation among heights, weights, and ages for 77 AKC breeds. Correlation coefficients, obtained by Spearman's rank order analysis, are shown in the upper right of each segment.

Multiple linear regression with age as the dependent variable, median height as an independent variable (x_H) and median weight as an independent variable (x_W), utilizes the equation: $y = 13.620 + (0.0702 * x_H) - (0.0538 * x_W)$ as determined by least squares estimates, with a residual standard deviation of 1.286, $R^2 = 0.585$, $R^{2adj} = 0.574$. Tests about the partial slope parameters yield an $F_{0.05} > 3.00$ ($p < 0.001$). The variance inflation factor for the equation is 4.662 and the coefficient of height has a t-value of 1.603 ($p = 0.113$) while the coefficient of weight has a t-value of -6.091 ($p < 0.001$). Plots of residuals (Figures 3 and 4) show an overestimation for the heavier and taller dogs. Therefore, the normal probability plot of residuals is presented (Figure 5) demonstrating that the residuals have a normal distribution.

Discussion

We assembled extensive data regarding the naturally diverse life spans across breeds of the domestic dog. This study extends the work of previous investigations that utilized the VMDB (Deeb and Wolf, 1994; Patronek et al., 1997) by examining life spans of pet dogs in association with their height and weight. Data pertaining to the general, healthy pet population expands upon previous investigations because dogs in the VMDB, by definition, have been referred to veterinary teaching hospitals, thereby biasing the median ages at death downward in comparison to the pet population (Patronek et al., 1997). Median ages were utilized here not only to offer a means of direct comparison to the work extrapolated from the VMDB, but also to allow classification of dogs by breed as a whole. Therefore, male life spans and female life spans are not investigated

independently, but as an entire breed group. Likewise, data for height and weight are also reported as median, and while the previous studies analyzed the relationship of body size to longevity, this work further dissects weight and height as independent factors of body size. The number of dogs contributing to the data for each breed varies between three and twenty-nine dogs, largely depending upon availability of each breed (*i.e.*, how common and/ or popular a particular breed is within the general population of dog owners) but also depending upon the owners' willingness to participate in ongoing scientific investigations by contributing a questionnaire. Therefore, as would be expected, the total population is not normally distributed, and has a preponderance of medium and small breeds, presumably due to their increased level of popularity. It is also quite possible that the financial burden of owning a dog is not such a consideration when owning small dogs versus larger ones, as it was casually observed that owners of small breeds tend to have multiple dogs while those who own large dogs often have only one. Due to the lack of normal distribution, nonparametric tests were employed throughout the analysis.

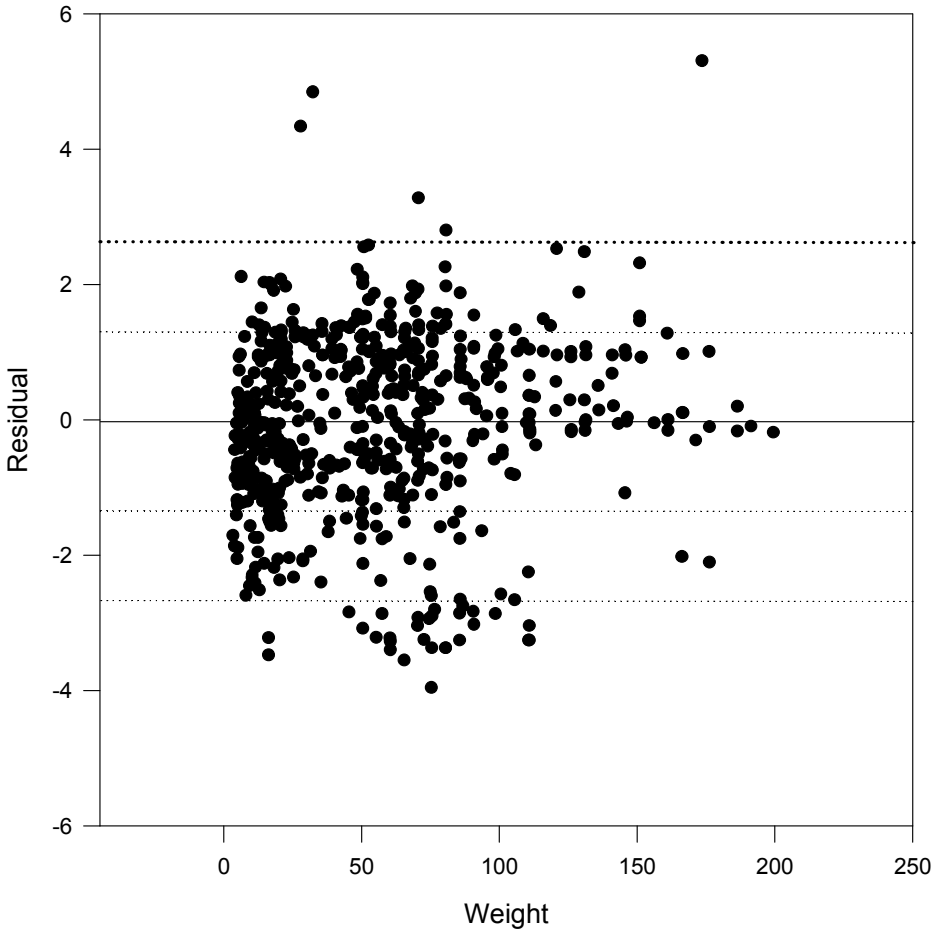

Figure 3: Plot of residuals for weights.
Plot of residuals versus weight in pounds of participating dogs. The graph reflects the mean, standard deviation and 2X standard deviation.

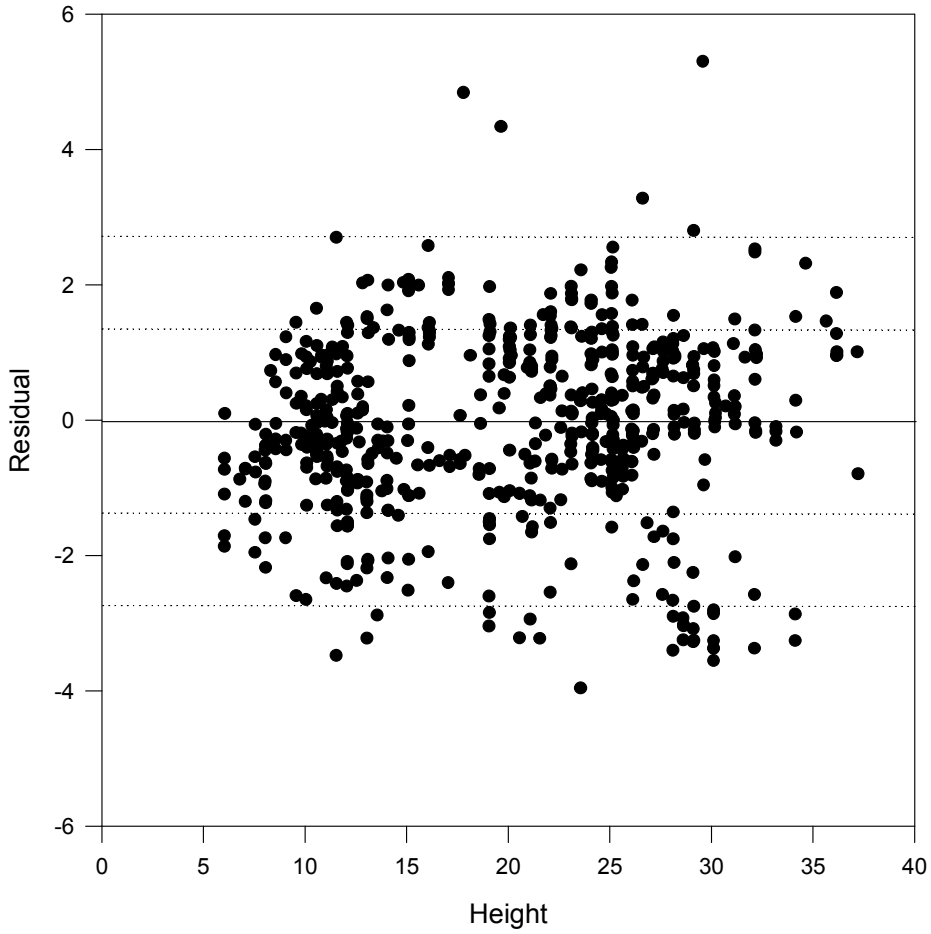

Figure 4: Plot of residuals for heights.
Plot of residuals versus heights in inches of participating dogs. The graph reflects the mean, standard deviation and 2X standard deviation.

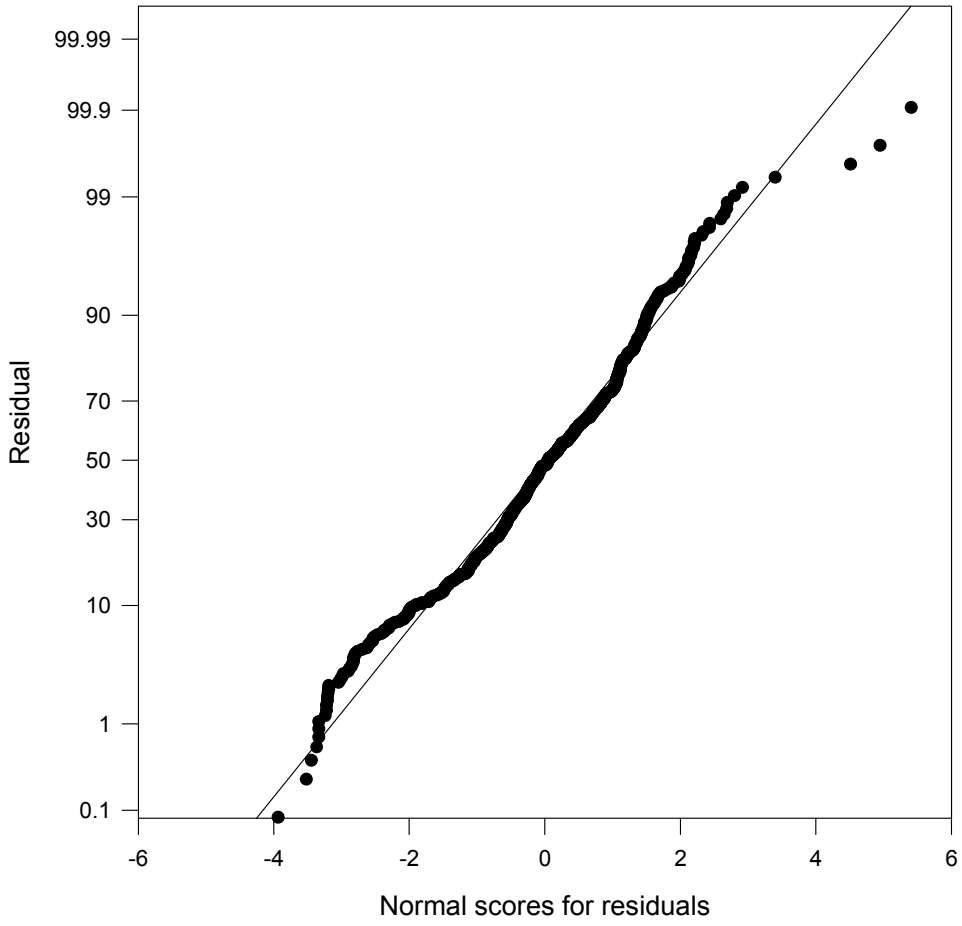

Figure 5: Probability plot of residuals.
Normal probability plot of residuals for $y = 14.016 + (0.0180 * x_H) - (0.0400 * x_W)$ as determined by least squares estimates.

The data collected from the general pet population corroborate those of previous studies derived from other select populations by indicating that longevity is inversely related to size (Deeb and Wolf, 1994; Patronek et al., 1997). This work demonstrates that life expectancy is related to both height and weight of breeds by utilization of the pet dog population previously noted by Patronek and colleagues (1997) as being unexamined. Although empirical data have offered these suggestions for many years,

there has been a lack of information for the generally healthy pet population, and specifically regarding regression and correlation data across breeds. The significant inverse relationship between life span and weight is -0.679 while the relationship between life span and height is -0.603. These results were as expected because it was evident that the heavier, short dogs, such as the English Bulldog, have unusually reduced life spans for their height and seem to fall in line more closely with the life spans of dogs with similar weight (i.e., Borzoi). Therefore, the individual relationship of each factor to longevity offered initial evidence that the two may be influencing life span unequally. Collectively, the coefficient of determination for this model indicates that 58.5% of variability in life span can be accounted for by variation in weight and height of a breed. This coefficient is interesting when considering that the difference between the life span of the shortest lived breed in this dataset, the Irish Wolfhound (median 7yrs) is approximately half that of the longest lived breed in this dataset, the Papillion (median 16 years).

A previous investigation determined that decreased life span of pure breed dogs compared with mixed breed dogs in all weight categories suggested that selective breeding of dogs over time for phenotypic traits such as body size had accelerated aging, independent of the effect of size alone (Patronek et al., 1997). Therefore, we introduced the independent variables of breed, height, and weight into a stepwise linear regression model, to get a better idea of which variables were affecting life span to the greatest degree. With life span as our dependent variable and breed, height and/or weight as independent variables, only weight emerged as a significant predictor of longevity. The

regression equation was: $y = 14.016 + (0.0180 * x_H) - (0.0400 * x_W)$ and an F test indicated evidence of an age-predictive value. As agreed, then, by these two studies, body size of the dog does have an affect on maximum life span; however, breed did not have an isolated effect on life span in this analysis.

A second question of interest, then, was whether dividing the study population into breed groups would explain a method by which the dogs' breed influenced its life span. Ideally, this type of division would separate the morphologically similar breeds of differing size into the same group and offer evidence of healthy and/ or unhealthy morphology and genetic health when aligned to longevity. The study population was divided into the American Kennel Club's defined breed groups: Herding, Hound, Non-sporting, Sporting, Terrier, Toy, Working, whose descriptive statistics are included in Table 1. The independent variables were introduced into a stepwise linear regression model. With life span as our dependent variable and breed group (x_{BG}), median height and median weight as independent variables, this analysis determined weight to be the primary and only significant predictor of longevity. The regression equation was: $y = 13.123 + (0.0111* x_{BG}) + (0.0710 * x_H) - (0.0528 * x_W)$ and an F test indicated evidence of an age-predictive value.

Although this would appear to be a difference between these conclusions and those of Patronek, we suggest that since the median weight of a dog, when examining a pure bred population, is not exclusive of its breed, these two parameters cannot be entirely separated from one another. Therefore, although these analyses indicate that breed and breed group do not have as great of an impact upon longevity as weight, it is

obvious that a given breed was not separated from its weight category in this study. While the data set was divided into breed groups, it should be noted that each breed group itself is composed of a wide range of sizes. For example, the Hound group contains both the Dachsund with a median weight of 12.0 pounds and a median height of 7.5 inches, as well as the Irish Wolfhound with a median weight of 122.5 pounds and a median height of 32.0 inches. Perhaps a follow-up investigation to differentiate breed and weight entirely could examine dogs falling significantly outside their breed standards; our interest, however, was to look at the typical height and weight of given dogs and therefore, this examination was limited to dogs generally representative of each breed.

With the consideration of a representative sample set, breed medians were also plotted on a scatterplot matrix which clearly indicates a relationship between age and weight, as well as between age and height, and of course, between height and weight. These negative relationships were expected, as mentioned, but raise the concern of collinearity (Belshley et al., 1980), which needs to be examined in greater detail. A scatterplot matrix of individual data points demonstrates more of a nonlinear than a linear association between height and weight. Their linearity was examined, however, because if one independent variable was highly collinear with the other independent variable, it could be expected to yield very large standard errors of partial slopes and inaccurate estimates of those slopes (Belsley et al., 1980). Further examination reveals that the variance inflation factor (VIF) is equal to 4.662. Although not entirely dismissive of a collinearity influence in this model, the number is much lower than may

have been generally expected. This dissociation may be due to those breeds whose healthy weight does not necessarily directly correlate with their heights. For example, the English Bulldog or the Borzoi, both breeds of which were included in the investigation, would be in this category. Our interest in including these breeds stems largely from the fact that these and several other breeds appear to be "disproportionate" in terms of either weight or height. Overall, however, the analysis does not suggest collinearity strong enough to interfere with regression analysis.

The t-value of 1.603 ($p = 0.113$) indicates that height does not have any additional predictive power in addition to weight with a t-value of -6.091 ($p < 0.001$). Plots of the residuals are presented, and while they seem to indicate an overestimation of the regression line (Slinker and Glantz, 1990), the normal probability plot of residuals demonstrates linearity. In general, therefore, given the inherent close association of height and weight, this regression model offers a reasonable estimation of expected longevity for any given healthy weight and height across breeds.

This work describes associations between longevity and size within the healthy pet population, demonstrating that a dog's weight is more predictive of life span than either height, breed, or breed group. While corroborating previous speculation and evidence of size in general, we utilized multiple linear regression to determine the factor contributing most significantly to life spans of 77 American Kennel Club (AKC) recognized breeds. With this parameter defined, and definitive populations within the species well established, it will be interesting to utilize this information as a basis for the study of mechanisms surrounding the genetics of longevity. The wide, natural variation

in life spans of the dog provides an excellent opportunity to examine wide ranges of life spans in a single species, which is critical because the research community has not taken full advantage of such natural variation to date (Austad, 1993). Furthermore, analysis of the canine nucleotide sequence should provide information not readily available in the past due to reliance upon rodents and other models not particularly close to the human in terms of nucleotide sequence. In this context, the dog seems an ideal candidate as a comparative model of aging (Patronek et al., 1997) and the work herein establishes a quantitative basis from which to further analyze the genetic components of longevity.

CHAPTER III

AGING-ASSOCIATED LOCI IN *Canis Familiaris*[*]

Overview

Although recent endeavors to discover the mechanisms of the aging process have been numerous and successful, there is still much to be learned. Genes implicated in the aging process were mapped to the canine genome and will serve as additional framework markers for the assignment of contiguous segments from the canine genome sequence to chromosomes. The 54 genes were selected because of their demonstrated contribution to longevity in other organisms or based upon their proximity to a marker, D4S1564, on human chromosome 4 (Puca et al., 2001). This effort lays the necessary groundwork for our utilization of the domestic dog as a model organism to define the genes that govern aging and longevity. Within the species, naturally diverse life expectancies and highly homogeneous populations create an ideal population structure for studying the genetic components of aging (Patronek et al., 1997).

Introduction

In recent years, many studies have been directed towards understanding the genetics and biochemistry of aging. These studies have involved genetic and mechanistic analyses of the complex process of aging by utilizing model systems,

[*] Reprinted with permission from "Aging associated loci in *Canis familiaris*" by Canterberry SC, Greer KG, Hitte C, Andre C, Murphy KE, 2005. Growth Development and Aging, 69, 101-113. Copyright 2005 by Growth Publishing Co., Inc.

including: *Caenorhabditis elegans*, *Drosophila melanogaster*, *Mus musculus*, and non-human primates. Investigations using these models initially offered evidence of disparate mechanisms of aging, but it has become apparent that there are indeed common molecular themes influencing longevity, even between invertebrate and vertebrate species. From these studies, it is clear that genetic factors are a primary component impacting an organism's longevity.

Towards the goal of utilizing the dog to determine the genes influencing aging, we selected 54 genes for mapping based on their various roles in aging and have placed 52 new genes on the current canine radiation hybrid map (Breen et al., 2004). Although the sequence of the canine genome is available (http://www.ensembl.org/), its assembly is ongoing and gaps remain. Importantly, because the accurate placement of contiguous segments (contigs) and scaffolds on the genome relies on the canine radiation hybrid (RH) map, the addition of 52 genes to the RH map provides an increased density of markers to which sequence scaffolds may be anchored. That these are gene based markers is of further benefit because the canine map is relatively sparse in such markers as compared to other organisms. Of the genes selected for this study, 26 were chosen due to their positive and/or negative contribution to longevity in other organisms, including *C. elegans*, *D. melanogaster*, *M. musculus*, and the human. The remaining 28 genes were selected based on their proximity to marker D4S1564 (HSA4), which was previously correlated with an increased ability for siblings of centenarians to achieve extreme old age by as much as 1.65 fold (Puca et al., 2001). Interestingly, several paralogs of the target genes were discovered in the dog. Additionally, it was found that

these paralogs have been evolutionarily conserved between the dog and human. Thus, this work provides a strong platform for additional studies of specific genes that control the aging process while concomitantly increasing the density of the RH map.

Study of the dog as an additional model of aging should contribute unique and valuable information to the current knowledge of common molecular themes influencing longevity. Studies across species have demonstrated that mechanisms influenced by insulin signaling, stress resistance, the ability to repair cellular and macromolecular damage, chromosome and nuclear structure, and caloric restriction (Warner, 2003) play contributory roles in aging and longevity. An intriguing link between these identified mechanisms which has attracted substantial attention is the effect of insulin and insulin-like growth factor I (*IGF-I*) signaling. Disruption of this signaling cascade, specifically by genetic mutation of genes resembling the human insulin/*IGF-1* pathway genes, can significantly extend life span across species. For example, *C. elegans* displays increased life span when harboring mutations of *daf-2*, *age-1*, and *daf-16* (Dorman et al., 1995; Morris et al., 1996; Kimura et al., 1997; Lin et al., 1997; Ogg et al., 1997). In mutant female flies, genetic manipulation of *InR* and *Chico* result in extension of life span (Clancy et al., 2001; Tatar et al., 2001). Mice that have been genetically altered in *Prop1* (Brown-Borg et al., 1996) and *Pit1* (Li et al., 1990; Flurkey et al., 2002) exhibit extension in life span. Although there are many differences in the way these organisms age, it is intriguing that these key pathways are common to their aging processes despite the organisms' evolutionary distance. Collectively, examinations of these pathways in various model systems strongly suggest that genetic factors play a significant role in

aging. When these analyses are applied to human centenarians, however, difficulties inherent to study of the human become apparent. These difficulties include long generation time, highly heterogeneous populations, limited availability of pedigree data, and poor experimental controls. As a result, these studies have relied primarily on linkage analysis to reveal regions of the genome associated with exceptional longevity (Puca et al., 2001). Specifically, linkage analysis revealed a marker on human chromosome 4 (HSA4) associated with familial longevity, albeit with limited statistical significance. The investigations pertaining to human longevity have, to date, been somewhat removed from the work in model organisms, but evidence indicates that understanding gene function in the aging process may augment various efforts to improve the quality of life for aging populations and also provide insight pertaining to extension of life spans.

Searching for susceptibility genes for multifactorial traits such as aging can prove to be quite problematic, and the criteria required for population dynamics, and ideal methodology by which to study them are still debatable (Freimer and Sabatti, 2004). Therefore, we propose the use of an alternative model to help unravel the genetic components of such complex traits as aging and longevity: the dog. Genetic isolates, represented by breeds, are well-suited for identifying susceptibility loci because strict registration guidelines reduce gene flow among breeds. The result is highly independent, homogeneous population structures that make comparisons between populations, or breeds, genetically informative. The breed structure of the dog is also useful because extensive pedigree information facilitates population-based genetic

studies. Importantly, because medical surveillance of the dog is second only to that of the human (Wise et al., 2003), detailed medical records are also readily available. Furthermore, when the first sequence assembly of the canine genome was made available, it was revealed that the nucleotide sequence of the dog is more similar to the human than is the rapidly evolving murine nucleotide sequence (Kirkness et al., 2003). Lastly, in general, life expectancy of the dog is inversely correlated to body size (Li et al., 1996; Deeb and Wolf, 1994; Greer, unpublished) thereby facilitating naturally occurring marked variation in life expectancy across breeds (Patronek et al., 1997). For example, a small breed such as the Chihuahua can be expected to live for as much as 15 years while the Saint Bernard is only expected to live a maximum of ten years (Greer, unpublished). This naturally extensive variation is an interesting component to the study of canine longevity as compared to the study of longevity in other mammals and access to many breeds, or genetically isolated populations, with different aging characteristics presents a unique opportunity to study the genetic components critical to aging.

Materials and Methods

Gene selection

The genes selected for this study were of two types: those previously linked to aging (Table 2), and those surrounding the marker D4S1564 (Table 3) on HSA4 (Puca et al., 2001). The genes selected from HSA4 cover this region with an average space between markers of 0.66 Mb, with the largest gap being 2.41 Mb and the smallest being 0.08 Mb. The human genome sequence data were utilized to select genes that would

provide optimal coverage of the region in the canine genome that exhibits conservation of synteny with the human.

Table 2: Genes previously associated with aging.

Gene	Human Location	Association to Aging	References
ACE	HSA17q23	Cardiovascular health	Reviewed in Panza *et al*., 2004; Blanche *et al*., 2000; Cambien *et al*., 1992
ApoA4	HSA11q23	Cardiovascular health	Reviewed in Panza *et al*., 2004; Merched *et al*., 1998; Pepe *et al*., 1998
ApoB	HSA2p23-24	Cardiovascular health	De Benedictis *et al*., 1998; De Benedictis *et al*., 1997
APP	HSA21q21.2	Oxidative stress response, neurodegeneration and AD	Poon *et al*., 2004; Moechars *et al*., 1999
ATM	HSA11q22-23	DNA damage response, oxidative stress and Ataxia-telangiectasia	Wong *et al*., 2003; Reviewed in Barzilai *et al*., 2002
Cyp2D6	HSA22q13.1	Cancer susceptibility	Reviewed in Agundez, 2004; Bathum, *et al*., 1998
F5	HSA1q23	Risk factor for thrombosis	Faure-Delanef *et al*., 1997; Mari *et al*., 1996
FGB	HSA4q28	Cardiovascular health and AD	Reviewed in Panza *et al*., 2004
FGF2	HSA4q26-27	Cellular proliferation	Cowan *et al*., 2003
FoxM1B	HSA12p13	Cellular proliferation and tissue repair	Kalinina *et al*., 2003; Kalinichenko *et al*., 2003; Krupczak-Hollis *et al*., 2003; Ly *et al*., 2000
FoxO1A	HSA13q14.1	Expression affected by CR	Furuyama *et al*., 2002
FoxO3	HSA6q21	Expression affected by CR, oxidative stress response	Furuyama *et al*., 2002; Kops *et al*., 2002
FoxO4	HSAXq13.1	Expression affected by CR, oxidative stress response	van der Horst *et al*., 2004; Furuyama *et al*., 2002; Kops *et al*., 2002

Note: Genes that have been previously associated to aging are listed here, along with their location in the human genome, the nature of the association with the aging process, and references. AD represents as associated risk for Alzheimer's disease, and CR denotes caloric restriction.

Table 2: Continued.

Gene	Human Location	Association to Aging	References
IGF1R	HSA15q25-26	Life span extension in mouse model and a human population	Reviewed in Quarrie and Riabowol 2004; Bonafe *et al*., 2003; Reviewed in Liang *et al*., 2003
Klotho	HSA13q12	Decreased life span in mouse model	Reviewed in Quarrie and Riabowol 2004
LEP	HSA7q31.3	Elevated levels in Snell dwarf mouse	Flurkey *et al*., 2001
Mthfr	HSA1p36.3	Cardiovascular health and AD	Reviewed in Panza *et al*., 2004
PAI1	HSA7q21.3-q22	Risk factor for thrombosis and AD, and increased expression in Klotho mouse	Reviewed in Panza *et al*., 2004; Takeshita *et al*., 2002; Kohler and Grant 2000
PAI2	HSA18q21.3	Increased expression in senescent cells	West *et al*., 1996
PARP	HSA1q41-42	DNA damage response	Grube and Burkle, 1992
Pik3C3	HSA18q12.3	Cellular proliferation and morphology	Matuoka *et al*., 2003
SHC1	HSA1q21	Oxidative stress response, and life span extension in mouse model	Migliaccio *et al*., 1999; Reviewed in Liang *et al*., 2003
SIRT1	HSA1q21.3	Mediates affects of CR in mice	Cohen *et al*., 2004
SOD2	HSA6q25.3	Oxidative stress response	Kokoszka *et al*., 2001
UPA	HSA10q24	Life span extension in mouse model	Miskin and Masos, 1997

Note: Genes that have been previously associated to aging are listed here, along with their location in the human genome, the nature of the association with the aging process, and references. AD represents as associated risk for Alzheimer's disease, and CR denotes caloric restriction.

Table 3: Genes on HSA4

Gene	Human Location	Canine Location	Closest Markers	LOD scores	Distance (cR)
SMARCAD1	95587588 –	CFA32	REN111K07	15.09	23.6
	95670837		D03908	14.64	24.2
BmpR1B	96255375 –	CFA32	BAC_374-A17	19.64	8.4
	96534470		REN111K07	17.58	13.9
COX7A3	98282346 –	CFA24	REN127K17	23.67	2.6
	98282666		FH3616	23.18	2.7
RAP1GDS1	99641014 –	CFA10	CFOR16F03	22.34	14.8
	99822211		BAC_376-O17	20.88	17.0
EIF4E	100259413 –	CFA28	FH2758	15.42	23.0
	100308675		BAC_286-F23	14.54	24.2
ADH7	100791876 –	CFA8	BAC_381-F17	14.97	18.5
	100814884		BAC_375-I4	14.90	18.6
MTP	100954375 –	CFA32	BAC_385-E11	20.10	8.1
	101003020		BAC_375-A15	19.54	8.3
DAPP1	101196384 –	CFA32	BAC_385-E11	18.11	13.5
	101249705		BAC_375-A15	16.64	16.7
H2AZ	101327636 –	CFA2	AHTH255REN	17.78	6.8
	101329827		EST29B7	16.62	10.2
PPP3CA	102402981 –	CFA32	BAC_375-A15	23.77	2.6
	102726747		EST26A4-T	22.20	2.8
BANK1	103170290 –	CFA32	BAC_375-A15	10.74	31.4
	103454289		BAC_286-J13	10.67	31.3
NfkB1	103880889 –	CFA32	FH4036	15.55	17.6
	103996878		BAC_417-L8	15.47	17.9
CENPE	104486024 –	CFA32	BAC_417-L8	17.78	11.7
	104578386		BAC_286-A17	17.78	11.7
TACR3	104969445 –	CFA32	BAC_286-A17	18.97	8.7
	105099793		BAC_286-B11	18.51	11.2
IDAX	105852150 –	CFA32	BAC_417-L8	16.12	13.0
	105871287		BAC_286-A17	16.12	13.0
KIAA	106622881 –	CFA32	BAC_417-L8	21.68	2.9
	106659778		BAC_286-A17	21.68	2.9
SCYE1	107696548 –	CFA32	BAC_286-A17	19.68	6.1
	107727738		BAC_417-L8	17.81	9.3
DKK2	108301779 –	CFA32	BAC_417-L8	21.68	2.9
	108417296		REN187G01	20.38	5.8

Note: Genes located on HSA4 surrounding marker D4S1564, their location on the Human Chromosome in megabases, the 2 closest markers, their respective Lod score values, their distance between the 2 markers and their CFA location are shown here.

Table 3: Continued.

Gene	Human Location	Canine Location	Closest Markers	LOD scores	Distance (cR)
D4S1564	108734149 - 108734386				
PAPSS1	108993642 –	CFA32	BAC_417-L8	14.85	22.0
	109100192		BAC_286-A17	13.36	27.8
LEF1	109427521 –	CFA29	BAC_381-H1	13.20	34.3
	109548398		BAC_375-M13	13.07	34.3
RPL34	110000543 –	CFA32	REN286D15	22.47	5.1
	110010388		BAC_417-L8	18.96	8.6
Col25A1	110203861 –	CFA32	EST12D6	16.29	9.9
	110682619		BAC_417-L8	12.10	21.7
CASP6	111068858 –	CFA32	REN286D15	16.82	16.1
	111083397		BAC_417-L8	15.48	17.5
EGF	111292870 –	CFA32	EST12D6	21.58	5.4
	111392239		REN286D15	18.43	13.2
PitX2	111997399 –	CFA3	REN296H17	11.77	23.1
	112017328		BAC_285-F22	10.40	30.5
ANK2	114429690 –	CFA32	BAC_382-G18	16.68	24.0
	114762088		BAC_416-D12	15.61	26.6
UGT8	115978730 –	CFA32	BAC_382-G18	28.56	3.8
	116056306		BAC_416-D12	25.43	7.8
NDST4	116207739 –	CFA32	BAC_382-G1	23.96	8.3
	116493852		BAC_374-E3	17.75	21.2

Note: Genes located on HSA4 surrounding marker D4S1564, their location on the Human Chromosome in megabases, the 2 closest markers, their respective Lod score values, their distance between the 2 markers and their CFA location are shown here.

Primer design and gene amplification

Sequences from numerous organisms, including human, mouse and dog were downloaded into Microsoft Word from NCBI, Ensemble!, and DDBJ. If hamster sequence was available, it was downloaded for comparison as well. Most of the canine sequences were retrieved from the canine genome sequence database (http://www.ncbi.nlm.nih.gov/genome/seq/CfaBlast.html); however, if canine sequence was not available, bovine and/or porcine sequences were substituted. All sequences

were aligned with one another, utilizing ClustalW (http://www.ebi.ac.uk/clustalw/index.html), to determine base similarities and differences. Primers were designed within conserved sequence regions for canine specific annealing. Parameters for primer selection were set between 18 and 22 bp in length, with nearly 50% GC content, and Tm close to 58°C, yielding products from 150 bp to 1000 bp in length. Primers were synthesized by IDT (Coralville, IA) and Sigma Genosys (The Woodlands, TX) using standard parameters and were purified using standard protocols. Polymerase chain reaction (PCR) mixtures consisted of 10µl volumes containing: 1.5mM $MgCl_2$, 0.2mM each dNTP, 1.0µM each primer, 1 Unit Jumpstart Red Taq (Sigma-Aldrich, St. Louis, MO), and 50 ng genomic DNA. Prior to genotyping on the RHDF5000 radiation hybrid panel (as provided from Universite de Rennes1), conditions were optimized for mapping on dog-hamster radiation hybrid cell lines, and each primer pair was simultaneously amplified on canine genomic DNA from the MDCK cell line, on hamster genomic DNA (originating from the A2H cell line), and on a mixture consisting of both dog and hamster DNA (ratio 1:2). Reaction mixtures were subjected to an initial denaturation at 94°C for 1.0 minute, followed by 35 cycles of 94°C for 30 seconds, 58°C for 30 seconds, 72°C for 30 seconds, and a final extension at 72°C for 10 minutes. PCR amplicons were resolved on a 2% agarose gel (200ng/ml EtBr) and subsequently visualized by exposure to UV light. In the event of non-specific amplification, annealing temperatures were increased if there were multiple products from canine DNA, and if hamster DNA was amplified. Decreased annealing temperatures were used if amplification from canine DNA resulted in faint or nonexistent amplification products. Primer pairs yielding

spurious bands were rejected. Upon completion of the aforementioned optimization, all genes were specifically amplified at annealing temperatures ranging from 56° C to 68° C. Subsequently, only primer pairs amplifying a canine-specific fragment from genomic DNA were used to screen the RH panel in duplicate.

Gene sequencing

The identity of amplicons was confirmed by sequencing. Sequencing was carried out on 50ng template with Big Dye Terminator (Applied Biosystems, Warrington, UK) and separated on an ABI 377 automated sequencer (Applied Biosystems, Warrington, UK). For each of the 54 genes, sequences were analyzed using BLASTn (http://www.ncbi.nlm.nih.gov/blast/). Those sequences with ≥98% homology to the orthologous genes of the human, mouse, rat, cow, or pig were accepted. If homology of amplicons was <98%, canine-specific primers were re-designed and specific products re-sequenced before acceptance of primers for use in RH mapping experiments.

RH mapping

Typing on the canine radiation hybrid panel, RHDF5000-2 which consists of 118 hybrid cell lines, was completed by PCR on each cell line for every marker. The RH panel was screened in at least duplicate under the experimentally determined optimal conditions and results were scored manually. The RH vectors corresponding to the 54 genes markers studied here were computed using the two-point analysis of the rh_tsp_map2.0 package (Agarwala et al., 2000), on the 4249-marker version of the RH map (Breen et al., 2004).

Results

The 54 aforementioned genes were typed on the RHDF5000 panel (Vignaux et al., 1999), and resulting vectors were integrated through pairwise analysis to the 4249-marker RH map (Breen et al., 2004). This analysis revealed the closest markers and their distances from the new marker by calculating LOD scores. LOD scores greater than 6 indicate significant linkage between the 2 markers, and these values reflect the distance between the markers in such a way that higher LOD scores imply a decreased physical distance between markers, with 1 centiRay (cR) corresponding to 150 Kb. Of the 54 genes mapped in this study, the chromosomal locations of 45 genes correlated with known regions in which there is conservation of synteny between the canine and human genomes (Guyon et al., 2003), while 9 genes did not correlate in this manner.

Twenty-six genes previously shown to be associated with aging in other organisms (Table 2) were examined, with results listed in Table 4. The majority of these genes mapped to predicted canine chromosomal regions based on previous comparative analyses between the dog and human (Guyon et al., 2003); however, two genes, *FGF2* and *SOD2*, mapped to CFA08 and CFA19 respectively, and a BLAST search revealed paralogs at these locations. Two others, *ACE* (CFA9) and *APP* (CFA31), mapped to locations consistent with those indicated by previous work (Breen et al., 2004).

Table 4: Mapping results.

Gene	Human Location	Canine Location	Closest Markers	LOD scores	Distance (cR)
ACE	HAS17q23	CFA9	ACE	28.50	2.1
			BAC_133-B8	25.02	6.3
ApoA4	HAS11q23	CFA5	BAC_381-P23	12.52	22.1
			REN114G01	11.77	23.1
ApoB	HAS2p23-24	CFA17	FH1003	21.28	3.0
			EST24F7	19.84	6.0
APP	HAS21q21.2	CFA31	BAC_380-F20	24.96	4.5
			APP	22.32	9.0
ATM	HAS11q22-23	CFA5	BAC_286-H12	14.60	8.6
			BAC_417-J2	12.73	13.6
Cyp2D6	HAS22q13.1	CFA10	BAC_376-G15	11.84	30.9
			BAC_417-N14	11.76	30.8
F5	HSA1q23	CFA7	BAC_374-C17	14.67	16.8
			BAC_381-B6	14.00	17.6
FGB	HSA4q28	CFA15	FGA	17.44	11.8
			BAC_382-E15	17.02	14.3
FGF2	HSA4q26-27	CFA8	BAC_374-K1	7.20	60.0
			REN68M10	6.45	66.0
FoxM1B	HSA12p13	CFA27	BAC_372-K5	9.56	49.8
			REN100M16	9.30	50.7
FoxO1A	HSA13q14.1	CFA25	BAC_283-C1	22.27	14.8
			REN54E19	19.60	20.7
FoxO3	HSA6q21	CFA12	C12.406	26.38	2.3
			BAC_382-G21	24.77	4.6
FoxO4	HSAXq13.1	CFA39	EST15G11	11.11	24.4
			D04614	10.32	28.5
Ghrhr	HSA7p14	CFA14	BAC_283-G12	16.55	10.2
			BAC_382-E7	13.94	17.7
IGF1R	HSA15q25-26	CFA03	FH2984	18.45	8.8
			FH2320	16.38	12.2
Klotho	HSA13q12	CFA25	REN103F16	20.16	16.5
			BAC_282-I08	16.14	25.9

Notes: Genes that have been previously associated to aging and their location in the dog as determined by RH mapping are listed here, along with the 2 closest markers, their respective Lod score values, their distance between the 2 markers and their CFA. For example, ACE was already mapped on the RH map on CFA9, so the closest marker is of course ACE, with a very strong Lod score value (28.5) and a very small distance (2.1 cR).

Table 4: Continued.

LEP	HSA7q31.3	CFA14	EST5C4	21.15	9.6
			BAC_372-E16	19.59	12.3
Mthfr	HSA1p36.3	CFA2	NPPA	20.47	5.7
			EST18D6	19.82	8.4
PAI1	HSA7q21.3-q22	CFA6	BAC_382-E21	18.60	16.2
			BAC_372-O10	15.99	22.0
PAI2	HSA18q21.3	CFA1	FH3603	24.36	6.5
			AHTK338	22.16	9.1
PARP	HSA1q41-42	CFA14	BAC_373-A17	13.76	20.2
			EST19E2	12.90	25.5
Pik3C3	HSA18q12.3	CFA7	BAC_281-F22	8.93	38.4
			BAC_286-G18	8.34	40.5
SHC1	HSA1q21	CFA7	EST7B12	16.73	4.1
			PKLR	14.60	8.6
SIRT1	HSA10q21.3	CFA4	BAC_374-G23	17.75	15.6
			AHT120	17.25	16.1
SOD2	HSA6q25.3	CFA19	REN306J16	18.60	11.0
			BAC_375-F15	18.20	13.3
UPA	HSA10q24	CFA4	BAC_381-B21	13.39	15.8
			BAC_375-G11	13.39	15.8

Notes: Genes that have been previously associated to aging and their location in the dog as determined by RH mapping are listed here, along with the 2 closest markers, their respective Lod score values, their distance between the 2 markers and their CFA. For example, ACE was already mapped on the RH map on CFA9, so the closest marker is of course ACE, with a very strong Lod score value (28.5) and a very small distance (2.1 cR).

Genes with orthologs on HSA4 are listed together in Table 3. These genes were expected to map to CFA32 based on previous work (Guyon et al., 2003). Of the 28 genes selected for analysis, 21 mapped to CFA32 with an apparent conservation of order between the dog and human (Breen et al., 2004). Seven markers expected to be found on CFA32 mapped to other chromosomes. These are *COX7A*3, *RAP1GDS1, EIF4E*, *ADH7*, *H2AZ*, *LEF1*, and *PitX2*, and were mapped to CFA24, CFA10, CFA28, CFA08, CFA02, CFA29, and CFA03, respectively. As was the case for *FGF2* and *SOD2*, most

of the locations that in fact are incorrect were designated so due to the presence of paralogs, or in the case of *COX7A3*, a pseudogene.

Discussion

The genes selected for mapping are believed to play roles in aging and longevity. Of the 54 genes investigated, 28 were selected for their proximity to a marker on HSA4 (Table 3) shown to increase the ability of centenarians' siblings to reach extreme old age (Puca et al., 2001). The remaining 26 genes were chosen based on previous associations with aging and/or the aging process (Table 2). Genes with similar association, but excluded from this investigation due to the fact that they have been previously mapped in the dog are *F*7, *FGA*, *FGG*, *Pit1*, and *Prop1*. In those studies, these genes were localized to the following chromosomes: CFA22, CFA15, CFA15, CFA31, and CFA11 (Guyon et al., 2003). The positions of two previously mapped genes, *ACE* (CFA9) and *APP* (CFA31), were verified by this work.

ApoE (HSA19q3.2; CFA01) is a gene of particular interest because it has been associated with human aging, and allelic variation has been identified as a risk factor for diseases accompanying the aging process (Blanche et al., 2000; Gerdes et al., 2000; Jian-Gang et al., 1998; Schachter et al., 1994). Despite multiple attempts, however, we were not able to design canine-specific primers to amplify this gene. Further scrutiny of the canine ortholog of *ApoE* is warranted, however, due to its association to longevity in the human.

Of the 26 genes that had been associated with aging in other organisms, two mapped to unexpected regions of the canine genome. Specifically, *FGF2*, which lies near the evolutionary breakpoint of HSA4 (a region which maps to CFA32 and CFA19; (http://www-recomgen.univ-rennes1.fr/doggy.html) instead mapped to CFA08, which corresponds to HSA14 (Guyon et al., 2003). *SOD2*, predicted to be on CFA01, mapped to CFA19, in a region that has conservation of synteny with HSA2q21.1 (Guyon et al., 2003). The recent availability of the canine genome sequence has provided a method by which to verify these unexpected results and BLAST against the canine genome sequence implicated that paralogs of *FGF2* and *SOD2* were the cause of these results. The presence of these gene paralogs is believed to be a relatively frequent gene duplication event in evolution that, in turn, leads to the formation of gene families (Lynch and Conery, 2000). These paralogs can also decrease the accuracy with which orthologs are determined (Nembaware et al., 2002). Numerous gene paralogs that affect the aging process (Rikke et al., 2000) have been found in *C. elegans*. Therefore, although these gene paralogs were mapped inadvertently, the data may be pertinent to future studies of aging in the dog.

The region of HSA4 that has been associated with the ability to reach extreme old age (Puca et al., 2001) is a region of the genome that has been conserved, and corresponds to CFA32. Although gene order appears to have been conserved, there are several genes for which specific order could not be resolved in this study. *MTP* and *DAPP1* both map between the markers BAC_385-E11 and BAC_375-A15. Mapping to an identical location are the genes *PPP3CA* and *BANK1*, which co-localize with

BAC_286-J13 and EST26A4-T. Several genes, *CENPE*, *IDAX*, *KIAA*, *SCYE1*, and *PAPSS1* co-localize and map to a region that also houses the markers BAC_286-A17 and BAC_417-L8. Three genes that map to a similar region of CFA32, with exact order undetectable, are *RPL34*, *COL25A1*, and *CASP6*. *ANK2* and *UGT8* also co-localize between BAC_416-D12 and BAC_382-G18. Although exact order of some of the genes selected from HSA4 could not be detected using two-point analysis, it appears that overall gene order is conserved between the dog and human.

The genes *COX7A3*, *RAP1GDS1*, *EIF4E*, *ADH7*, *H2AZ*, *LEF1* and *PitX2*, on HSA4, were expected to be on CFA32 (Guyon et al. 2003), but instead mapped to CFA24, CFA10, CFA28, CFA08, CFA02, CFA29, and CFA03, respectively. The regions of these chromosomes correlate to HSA20p13, HSA12q13.2, HSA10q23, HSA14, HSA1p36.13, HSA8, and HSA15q26, respectively (Guyon et al., 2003). Due to these unexpected results, the sequences used in designing primers for these markers were subjected to a BLAST search of the canine genome. It was discovered that most of these genes (*RAP1GDS1*, *EIF4E*, *H2AZ*, and *LEF1*) have been localized to these unexpected regions due to the presence of gene paralogs at the aforementioned canine locations. The mapping of *COX7A3* to CFA24 was most likely the result of identification of *COX6CP2* which is in the corresponding region of the human genome. This hypothesis is supported by the fact that *COX7A3* is now known as *COX7AP2*, a pseudogene, which would have a decreased expected homology with canine sequence. Upon BLAST analysis of *ADH7* and *PitX2*, homology was found only to CFA32 as would be expected;

therefore the unexpected locations of the genes *ADH7* and *PitX2* cannot be explained as easily as the other genes.

Through this investigation, 52 new gene based markers have been added to the existing canine genome map (Breen et al., 2004) and the location of two gene based markers was verified. In total, genes believed to be involved in aging were localized to 19 of the 38 canine autosomes, as well as to the X chromosome. This effort has advanced the map of the dog as it applies to future studies pertaining to the genetics of aging and longevity. Indeed, all 52 mapped genes are positioned close to polymorphic markers in the latest RH map, at a minimum distance of less than 2 Mb, representing an important resource for genetic linkage studies or even genetic association studies because SNP markers will be soon available. Current experiments are focused on identification of SNPs in a subset of these genes using specific breeds that have marked differences in life spans, in order to better understand the inverse relationship between size and longevity. Identification of genes that govern aging in the dog may be of interest in studies of other organisms to determine the extent of evolutionary conservation of these mechanisms of aging. Utilization of the dog as a model organism for investigations into the aging process will provide additional information concerning the pathways that influence aging in mammals, and may therefore be of particular interest to those interested in human aging.

CHAPTER IV

ANALYSIS OF *Pit1*, *Prop1*, *Ghrhr*, AND *Ghr* IN NINE DOG BREEDS WITH VARYING LIFE EXPECTANCIES

Overview

As average life expectancy of humans increases, the importance of research to determine the mechanisms of aging also increases. It is believed that there are many genes involved in aging, a trait with an estimated heritability of up to 30%. While many model organisms exist, the dog provides a unique opportunity to study a mammal with a naturally diverse life expectancy that is inversely correlated to its size. This phenomenon is also seen in four strains of inbred mice with mutations in the genes *Pit1*, *Prop1*, *Ghrhr* and *Ghr*. We have analyzed the sequence of these genes in nine dog breeds with varying life expectancies. A total of 53 polymorphisms were identified. Of these, seven were located in coding regions of *Ghr* and *Ghrhr* and yielded statistically significant p-values upon analysis with Fisher's exact test.

Introduction

Average life expectancy for humans has steadily increased each year in the United States, from less than 50 at the turn of the twentieth century to an estimate of nearly 80 for 2006. Therefore, a better understanding of the aging process will be highly beneficial to our society in the very near future. The importance of studies pertaining to aging is especially critical now because the largest increase in population is expected for

those over the age of 85 years, closely followed by those aged 65 and older. As the population of aged individuals grows, and life expectancy is increased, better understanding of aging is crucial to postponing the onset of, and developing more effective treatments for, aging-associated diseases (National Institute on Aging, 2005).

It has been strongly suggested that genetic factors contribute to longevity, and heritability of average life expectancy has been estimated to be as high as 30% (Ljungquist et al., 1998). Identification of genes that govern aging, and elucidating how these genes function may enable progress to be made not only in the extension of life span but also in improving the quality of life of aged individuals. The genetics of several organisms such as the nematode, fruit fly, mouse and human have been scrutinized by scientists interested in genes that play significant roles in aging and longevity.

Identification of genes that affect human longevity has been challenging for several reasons; two of which are long life expectancy, and highly heterogeneous populations. For these reasons, many genes that are associated with excessive old age in one population may not necessarily be so correlated with another population (Benedictis, et al., 2000), and more genes that are deleterious, as opposed to genes that extend human life span, have been discovered (reviewed in Lao et al., 2005). Numerous genes that play roles in multiple activities, such as lipoprotein metabolism, cardiovascular health, oxidative stress, cancer susceptibility, coagulation, and chromatin structure have been linked to varied life spans in multiple populations (Benedictis, et al., 2000; Heijmans, et al., 2000). Linkage analysis of siblings of centenarians identified a locus on

chromosome 4 near the microsatellite marker D4S1564 that affects the ability to attain exceptional old age (Puca, et al., 2001). While results from studies of human populations are directly applicable, there are numerous drawbacks as well, such as long life span, poor experimental controls and lack of extensive pedigree information. Due to these drawbacks, another model organism for investigations into the aging process could provide valuable insight into genes that play important roles in aging and longevity.

While the dog is not a well established model of aging, some preliminary data have been gathered. Cellular proliferative capacity was shown to be inversely related to breed size by Li and colleagues (1996). Investigations based upon information contained within the veterinary medical database (VMDB) revealed that larger breeds have truncated life spans when compared to smaller breeds (Deeb and Wolf, 1994; Patronek, et al., 1997). Recently, this inverse correlation between size and longevity has been confirmed in the American pet population (Greer et al., in press). Interestingly, cross breed dogs do not live longer than their pure breed counterparts, therefore these do not exhibit hybrid vigor as one might expect (Bronson, 1982). Diet restriction in Labrador Retrievers increased median life span by nearly two years (Kealy et al., 2002). Additionally, telomere shortening has been observed in canine fibroblasts *in vitro*, and differences in telomere lengths between different breeds have also been reported (McKevitt et al., 2002). In a study utilizing pet insurance databases, researchers determined that dogs belonging to different breeds age at different rates (Egenvall et al., 2005). Therefore, the dog presents a unique opportunity to study aging in an organism that naturally exhibits a wide range of life expectancy.

Historically, the dog has served as an effective model organism for many investigations of hereditary diseases. Approximately half of the more than 450 inherited disorders in the dog are also present in human populations (OMIA 2006; Ostrander and Giniger 1997). While individuals within highly inbred populations, or breeds, are highly homogeneous, there is sufficient heterogeneity between individuals from different breeds to be informative in genetic studies (Parker et al., 2004). Importantly, the canine genome was recently sequenced (Lindblad-Toh et al., 2005) and this revealed that the dog has a much higher level of sequence identity to the human than does the mouse (Jiang et al., 2005).

In addition to the inherent genetic advantages the dog has as a model, the well established relationship of man and his best friend provides researchers with invaluable information as well. Extensive pedigree information is readily available through owner records. It is often easy to get samples from the pet population through owner cooperation. These dogs share a common environment with human owners, providing an automatic control for environmental exposures and alleviating concerns for the welfare of the animals involved in research and for genetic drift commonly observed in caged animals. Finally, it is critical to note that the dog enjoys medical surveillance second only to the human (Wise et al., 2003). Therefore, the clinical hallmarks of the aging process in the dog are well characterized.

We have taken advantage of this unique opportunity to study natural variation in life span within a single species to help determine the genetic components of aging. The inverse correlation between size and life expectancy that occurs naturally in the dog is

similar to what has been observed in four strains of dwarf mice with extended life spans. Ames mice are defective in *Prop1* and Snell dwarf mice harbor mutations in the *Pit1* gene; both strains are defective in pituitary development, are one-third normal size and exhibit 40-50% extension in life expectancy (reviewed in Quarrie and Riabowol, 2004). Mice which have mutations in *Ghrhr*, known as Little mice, are one-half the size and live approximately 25% longer than wild-type (reviewed in Quarrie and Riabowol, 2004). Also at one-half normal size are Laron mice, which are *ghr*-knockouts and have up to a 55% increase in life span (reviewed in Quarrie and Riabowol, 2004). The genes involved in Little and Laron mice play roles in the insulin-signaling pathway (reviewed in Quarrie and Riabowol, 2004; reviewed in Warner, 2003).

Because there is an inverse correlation between size and life span in these strains of mice, we chose to analyze the sequences of *Pit1*, *Prop1*, *Ghrhr*, and *Ghr* in nine dog breeds in search of single nucleotide polymorphisms (SNPs). Breeds to be included in this analysis were selected based on previous regression analysis of the pet population (Greer et al., in press). Utilizing this information, three breeds were chosen from each of three size categories, small, medium and large. Breeds were selected such that one breed would have a shorter than predicted life expectancy, one breed would have a life expectancy similar to that predicted by the linear regression information, and one breed would have a longer than predicted life expectancy. The breeds selected for SNP analysis in were the Pomeranian, Dachshund, Miniature Schnauzer, Basset Hound, Staffordshire Bull Terrier, Standard Schnauzer, Bloodhound, Great Pyrenees and Giant Schnauzer. In summary, the four selected genes were investigated in nine carefully

selected breeds to determine potential genetic variation that may be contributing to variation in life span of the dog.

Materials and Methods

Breed selection

Breed selection for this investigation was based on the relationship between size and life expectancy in the domestic dog (Greer et al., in press). Based on this analysis, the inverse correlation between size and longevity can be represented by the equation $y = 14.016 + (0.0180 * x_H) - (0.0400 * x_W)$. The heights and weights for dog breeds from the American Kennel Club (AKC) (as reported by breed clubs and AKC standards) were entered into this formula. These breeds were then divided into three categories based on their size as reported by breed clubs or AKC standards. Breeds less than 7 kilograms or 36 centimeters in height (as measured at the withers) were placed in the small category. The medium category consisted of breeds weighing in at 7 to 27 kilograms or being 36 to 60 centimeters in height. Breeds were classified as large if they weighed more than 27 kilograms or measured more than 60 centimeters at the withers. The output from the above equation was compared to the reported life expectancies of breeds, and utilizing this information, three breeds were chosen from each size category, one breed having a shorter than predicted life expectancy, one breed having a life expectancy similar to that predicted by the linear regression equation, and one breed having a longer than predicted life expectancy. The breeds selected in the small breed category were the Pomeranian, Dachshund and Miniature Schnauzer. Breeds included in this study from the medium

size category were the Basset Hound, Staffordshire Bull Terrier and Standard Schnauzer. Large breeds selected for analysis were the Bloodhound, Great Pyrenees and Giant Schnauzer. These breeds in each size class exhibit shorter than expected, expected and longer than expected life spans, respectively. The three Schnauzer breeds were included due to the fact that they exhibit similar life expectancies and have been derived from the same base population.

Sample collection

DNA samples were obtained from dogs in the pet population. Participation was solicited at dog shows, through direct contact with owners and breeders, and through publication in newsletters of multiple breed clubs and the Canine Health Foundation. Cytology brushes were used to collect buccal cells from dogs included in this investigation. Additionally, owners completed detailed questionnaires which included the dog's name, date of birth, sex, registration number, height and weight. The data from the questionnaires were entered into a database for future reference and DNA was extracted using standard protocols. In order to select the most informative samples to use in this study, pedigree information was acquired from the AKC databases based on information provided by owners. Pedigrees were analyzed to select ten unrelated dogs each from the chosen nine breeds. Dogs were considered unrelated if they shared no common grandparents (Parker *et al*., 2004) and were subsequently used for identifying SNPs.

Sequencing

Canine sequences for *Ghr* and *Pit1* mRNA were downloaded from Ensemble! Genome Browser (http://www.ensembl.org/index.html). Canine mRNA sequence for *Prop1* was retrieved from Entrez Gene (http://www.ncbi.nlm.nih.gov/entrez/query .fcgi?db=gene). Canine mRNA sequence was unavailable for *Ghrhr;* consequently, human mRNA sequences were downloaded instead. These sequences were used to conduct a BLAST search of the dog genome (http://www.ncbi.nlm.nih.gov/genome /seq/BlastGen/BlastGen.cgi?taxid=9615) which resulted in identification of the genomic sequence for each gene. One genomic contig with significant alignment to the human *Ghrhr* sequence was revealed. This contig represents a region of canine chromosome 14 that corresponds to the location of *Ghrhr*, as determined by radiation hybrid mapping (Canterberry et al., 2005). The following gene segments were selected for sequence analysis: 1000 base pairs (bp) upstream and downstream, all exons, and 100bp flanking each exon. Genomic sequence files and genomic DNA from the aforementioned ninety dogs were sent to Polymorphic DNA Technologies, Inc. (Alameda, CA) for re-sequencing.

Sequence analysis

Chromatogram files were provided by Polymorphic DNA Technologies, Inc. (Alameda, CA). These files were subjected to analysis utilizing Phred/Phrap/Consed software. Phred was used to call bases and assign quality values to each. Phrap assembled and aligned the reads into contigs. After running Phred and Phrap, Polyphred analysis was used to identify potential polymorphisms. Consed was then used to view

the sequence data and all potential polymorphisms were manually verified and the genotype of each dog was noted. Genotypes were entered into a database for each dog, and allelic frequencies were calculated for each breed as well as for this entire population.

SNP comparison

SNPs with overall allelic frequencies of 10% or greater were further analyzed using the Fisher's exact test (http://www.unc.edu/~preacher/fisher/fisher.htm). SNPs with allelic frequencies less than 10% were excluded from further scrutiny because alleles with such low frequencies are unlikely to be associated with a given phenotype. Comparisons were made between dogs classified as small, medium and large breeds and between dogs classified as having a longer life expectancy, an average life expectancy and shorter than predicted life expectancy. Observed differences were considered statistically significant when $p \leq 0.0001$.

Results

SNPs were identified in each of the genes as follows: seventeen SNPs were discovered in *GHR*, twenty-six were found in *Ghrhr*, five in *Prop1*, and twenty-five SNPs were identified in *Pit1*. Of these 53, those with an overall allelic frequency of at least 10% were selected for further analysis. These 31 polymorphisms were analyzed using the Fisher's exact test.

Fisher's exact test revealed that five polymorphisms in *Ghr* are in linkage disequilibrium (LD) in these dog breeds (Table 5). Two of these were in the coding

sequence. The first is located in exon 5, base 132, and results in a glutamic acid to lysine change in the amino acid sequence. This polymorphism is in linkage disequilibrium when small breeds are compared to both large and medium breeds, but not when large and medium breeds are compared. The second polymorphism within the coding region of *Ghr* that is in LD is in exon 9 at base position 887. This polymorphism is silent, however, and does not result in a change in amino acid sequence. Canine codon usage tables (http://mendel.berkeley.edu/dog/dogcod.html) also indicate that this mutation would have no effect on synthesis of appropriate quantities of the protein. The remaining three polymorphisms are in non-coding regions and are not within known promoter regions or regions that are highly conserved between species. Two of these are located 607 and 142 bases upstream from exon 1 which are statistically significant only when long lived dogs are compared to dogs with similar to predicted life expectancies. Finally, a SNP 79 bases upstream from exon 6 is significant when small breed dogs are compared to both large and medium breeds, but not when large and medium breeds are compared.

Table 5: P-values for *Ghr* polymorphisms.

Locus	Long LE vs. Average LE	Average LE vs. Short LE	Long Le vs. Short LE	Small vs. Medium	Medium vs. Large	Small vs. Large
Exon 1 Base -607	**p = 0.000035**	p = 0.0034	p = 0.13	p = 0.056	p = 0.0023	p = 0.15
Exon 1 Base -250	p = 0.16	p = 0.001	p = 0.031	p = 0.102	p = 0.50	p = 0.63
Exon 1 Base -142	**p = 0.000047**	p = 0.0041	p = 0.13	p = 0.14	p =0.0055	p = 0.11
Exon 4 Base -63	p = 0.058	p = 0.64	p = 0.049	p = 0.52	p = 0.0055	p = 0.003
Exon 5 Base 132	p = 0.28	p = 0.13	p = 0.028	**$p < 1x10^{-8}$**	p = 0.027	**$p < 1x10^{-8}$**
Exon 6 Base -79	p = 0.34	p = 0.60	p = 0.35	**$p < 1x10^{-8}$**	p = 0.068	**$p = 3.7x10^{-6}$**
Exon 9 Base 887	p = 0.35	p = 0.28	p = 0.13	**$p < 1x10^{-8}$**	p = 0.35	**$p < 1x10^{-8}$**

Note: Fisher's exact test p-values for polymorphisms with allelic frequencies of at least 10% in *Ghr* are recorded here. Bold numbers indicate statistically significant p-values.

In *Ghrhr*, seven polymorphisms were discovered that were in LD (Table 6), five of which were located within the coding sequence. Three of these, however, did not change the amino acid sequence and canine codon usage tables indicated that these polymorphisms would not likely affect protein synthesis by changing the codon to one of rare use in the dog. The SNP in exon 2 of *Ghrhr* results in the amino acid alanine being replaced with valine and allele frequencies are statistically different when large dogs are compared to small and medium sized breeds, but not when small and medium breeds are compared. A SNP in exon 4 at base position 70 replaces valine with methionine. These alleles are in LD only when medium breeds are compared to large breeds, however. The two remaining SNPs are in the non-coding region 6 bases upstream of exon 2, in LD when small and medium sized breeds are compared to large dogs, and 53 bases downstream of exon 10, significant only when medium sized breeds are compared to

large dogs. These are not located within known promoter regions or within regions that are highly conserved between species.

Table 6: P-values for *Ghrhr* polymorphisms.

Polymorphism	Long LE vs. Average LE	Average LE vs. Short LE	Long Le vs. Short LE	Small vs. Medium	Medium vs. Large	Small vs. Large
Exon 2 Base -6	p = 0.22	p = 0.077	p = 0.0084	p = 0.0021	**p = 0.000027**	**p = 8.9×10^{-6}**
Exon 2 Base 8	p = 0.00035	**p = 4.8×10^{-6}**	p = 0.24	p = 0.42	p = 0.027	p = 0.074
Exon 2 Base 43	p = 0.13	p = 0.17	p = 0.012	p = 0.53	**p = 5.6×10^{-6}**	**p = 3.5×10^{-6}**
Exon 2 Base 44	p = 0.11	p = 0.0083	**p = 0.000082**	p = 0.13	**p = 1.0×10^{-6}**	p = 0.00035
Exon 3 Base 17	p = 0.003	p = 0.26	p = 0.00022	p = 0.38	**p = 1.7×10^{-6}**	**p = 0.00002**
Exon 4 Base 70	p = 0.02	p = 0.002	p = 0.32	p = 0.12	**p = 4.8×10^{-6}**	p = 0.0013
Exon 6 Base -15 to -10	p = 0.046	p = 0.13	p = 0.36	p = 0.19	p = 0.0034	p = 0.059
Exon 9 Base 78	p = 0.073	p = 0.13	p = 0.45	p = 0.26	p = 0.0067	p = 0.059
Exon 9 Base +16	p = 0.073	p = 0.13	p = 0.45	p = 0.26	p = 0.0067	p = 0.059
Exon 10 Base -65	p = 0.199	p = 0.38	p = 0.078	p = 0.13	p = 0.50	p = 0.087
Exon 10 Base+53	p = 0.06	p = 0.035	p = 0.50	p = 0.0003	**p = 1.0×10^{-8}**	p = 0.013

Note: Fisher's exact test p-values for polymorphisms with allelic frequencies of at least 10% in *Ghrhr* are recorded here. Bold numbers indicate statistically significant p-values.

Two polymorphisms were found in *Pit1* that exhibited statistically significant changes in allelic frequency (Table 7). Both are located in the non-coding region 72 and 75 bases downstream of exon 3, a region that is not highly conserved across species. The SNP 72 bases downstream is in LD when medium-sized dogs are compared to small

and large dogs. When small and medium breeds are compared, the SNP found 75 bases downstream is in LD, but no other comparisons yielded statistically significant p-values.

Table 7: P-values for *Pit1* polymorphisms.

Polymorphism	Long LE vs. Average LE	Average LE vs. Short LE	Long Le vs. Short LE	Small vs. Medium	Medium vs. Large	Small vs. Large
Exon 1 Base -103	p = 0.097	p = 0.0022	p = 0.086	p = 0.065	p = 0.048	p = 0.196
Exon 3 Base +12	p = 0.29	p = 0.23	p = 0.068	p = 0.15	p = 0.496	p = 0.195
Exon 3 Base +72	p = 0.43	p = 0.0065	p = 0.014	**p = 2.2×10^{-8}**	**p = 0.000037**	p = 0.36
Exon 3 Base +75	p = 0.068	p = 0.0037	p = 0.14	**p = 0.000078**	p = 0.00034	p = 0.47
Exon 6 Base -16	p = 0.18	p = 0.13	p = 0.50	p = 0.0031	p = 0.00023	p = 0.29
Exon 6 Base 362	p = 0.13	p = 0.09	p = 0.50	p = 0.019	p = 0.0037	p = 0.34
Exon 6 Base 788	p = 0.03	p = 0.018	p = 0.50	p = 0.038	p = 0.0015	p = 0.16
Exon 6 Base 1022	p = 0.18	p = 0.18	p = 0.58	p = 0.0095	p = 0.00094	p = 0.295
Exon 6 Bases 1580 to 1583	p = 0.32	p = 0.37	p = 0.51	p = 0.0015	p = 0.00013	p = 0.34
Exon 6 Base 2150	p = 0.083	p = 0.22	p = 0.37	p = 0.011	p = 0.0042	p = 0.46
Exon 6 Base +52	p = 0.00069	p = 0.28	p = 0.011	p = 0.52	p = 0.22	p = 0.204

Note: Fisher's exact test p-values for polymorphisms with allelic frequencies of at least 10% in *Pit1* are recorded here. Bold numbers indicate statistically significant p-values.

Analysis of two SNPs within *Prop1* with allelic frequencies of 10% or more determined that both were in LD when dogs that have life spans similar to predicted are compared to dogs with shorter than expected life spans (Table 8). These polymorphisms

are located in the non-coding regions 898-902 and 917 bases upstream of exon 3, a region that is not highly conserved across species.

Table 8: P-values for *Prop1* polymorphisms.

Polymorphism	Long LE vs. Average LE	Average LE vs. Short LE	Long Le vs. Short LE	Small vs. Medium	Medium vs. Large	Small vs. Large
Exon3 Bases 898-902	p = 0.008	**p = 0.000023**	p = 0.06	p = 0.06	p = 0.03	p = 0.43
Exon3 Base 917	p = 0.008	**p = 0.000023**	p = 0.06	p = 0.06	p = 0.03	p = 0.43

Note: Fisher's exact test p-values for polymorphisms with allelic frequencies of at least 10% in *Prop1* are recorded here. Bold numbers indicate statistically significant p-values.

Discussion

This examination of *Ghr*, *Ghrhr*, *Pit1*, and *Prop1* in nine dog breeds of varying sizes and life expectancies led to the discovery of 53 polymorphisms. Seventeen were identified in *Ghr*, two of which were insertion/deletions (indels), and the remaining 15 were SNPs. Of the 26 found in *Ghrhr*, only one was an indel. Only five polymorphisms were discovered in *Prop1*, one being an indel. Lastly, 23 SNPs and two indels were identified in *Pit1*.

Thirteen ancestral SNPs were discovered in which both alleles were found in every breed. These polymorphisms were seen in *Pit1* (eight) and *Ghrhr* (five). Seven of the nine breeds included in this investigation belonged to an evolutionary breed cluster as described by Parker et al. (2004). These breeds were incuded in cluster four and are the three Schnauzer breeds, the Basset Hound, Bloodhound, Pomeranian and Dachshund. The Pomeranian and Dachshund also belong to clusters 2 and 3 (Parker et

al., 2004). The Great Pyrenees and Staffordshire Bull Terrier were not included in that preliminary investigation, but the Great Pyrenees is thought to be an ancestor of the Newfoundland (Crowley and Adelman, 1998), a breed found in cluster two (Parker et al., 2004). The Staffordshire Bull Terrier is descended from the Bulldog (Crowley and Adelman, 1998), a breed also included in cluster two (Parker et al., 2004). The presence of these ancestral polymorphisms, in conjunction with historical records, suggests that these two breeds may be added to one or both of these clusters, although additional data would be necessary.

Thirty one polymorphisms with allelic frequencies of at least 10% were selected for further analysis. These were compared using Fisher's exact test as follows: 1) breeds with life expectancies longer than predicted vs. breeds with average estimated life spans, 2) breeds with average estimated life spans vs. breeds with shorter than expected life spans, 3) breeds with life expectancies longer than predicted vs. breeds with shorter than expected life spans, 4) small breeds vs. medium-sized breeds, 5) medium-sized breeds vs. large breeds, and 6) small breeds vs. large breeds. Of these 31 polymorphisms, 16 presented with statistically significant p-values for at least one of the six comparisons made.

Of those that were statistically significant for at least one of the comparisons made, nine were in non-coding regions of the genes. These SNPs were not located in known regulatory regions, or in regions conserved between species. Although it is possible that these polymorphisms are in portions of the genome involved in as yet unknown regulatory mechanisms, we do not currently believe that these affect the

resulting protein and are, therefore, unlikely to be causative for the differences observed in size and longevity in these breeds.

The seven polymorphisms in coding regions which yielded statistically significant p-values were located in *Ghr* and *Ghrhr*. Two significant changes were seen in *Ghr*. At base position 132 of exon five an A to G transition mutation leads to an amino acid change from lysine to glutamic acid at residue 191. This change was evaluated with SIFT software, which considers values of 0.05 or less to indicate an intolerant change in the protein (Ng and Henikoff, 2001). The resulting SIFT value for this amino acid substitution was 1.0 and it was therefore determined that although this change results in an acidic amino acid being replaced by a basic one, this change is well tolerated, and does not impact the function of the protein. The second significant polymorphism in the coding region of *Ghr* is located in exon 9 at position 887, but does not affect the amino acid sequence. Additionally the canine codon usage table does not indicate that protein synthesis would be negatively affected by either of these SNPs. Although both of these mutations are in LD in small dogs when they are compared to medium and large breeds, because they do not seem to affect the function of the protein, it is unlikely that they are a determining factor for size in the dog. Five statistically significant SNPs in *Ghrhr* were within the coding region, three of which did not change the amino acid sequence. Two of these silent mutations were in exon 2 at base positions 8 and 44, and the third is in exon 3 at position 17. The mutation at base position 43 of exon 2, a C to T transition, results in alanine being replaced with valine. The SIFT value was calculated to be 0.43, a tolerated change for this protein. This mutation is observed

to be in LD in large breeds when compared to small and medium-sized breeds, but we do not believe this mutation to affect size in the dog because it does not appear to alter the protein function. Another transition of G to A, in exon 4 at base position 70, changes the amino acid sequence by replacing valine with methionine. The SIFT value was calculated at 0.01, indicating that this mutation may very well have an affect on this protein. It is interesting to note that this mutation is statistically significant ($p = 4.8 \times 10^{-6}$) when large dogs are compared to medium-sized breeds, but not when they are compared to small breeds ($p = 0.001$). Also, this mutation was found in the large breeds as follows: Great Pyrenees, four heterozygotes and three homozygotes; Giant Schnauzer, two heterozygotes and one homozygote; and in Bloodhounds two heterozygotes were discovered. The only other breed in which this allele was identified was the Dachshunds with three heterozygotes being found. Genotyping of additional dogs and/or breeds could lead to a significant change being discovered when large and small breeds are compared. Additionally, protein analysis to establish the effect of this mutation might show how this mutation may contribute to determining body size in the dog.

CHAPTER V

CONCLUDING REMARKS

The goal of the Canine Genetics Research Laboratory is to improve the health and quality of life for dogs and their human caretakers through a better understanding of canine genetics. The ultimate goal of aging research is to extend life expectancy and improve the quality of life for aged individuals. This investigation, in keeping with the goals of this laboratory and aging research in general, set three specific aims towards elucidating genetic factors involved in the aging process in the dog. These were 1) to determine the extent of the inverse relationship between size and longevity in the pet dog population, 2) to identify and map genes associated with the aging process in the canine genome, and 3) to analyze four genes that confer extended life spans in mice of smaller size to determine if any sequence differences are responsible for the similar phenomenon which is observed in the dog.

In order to accomplish the first aim, we collected an extensive data set from the general, healthy pet population. This data set included more than 700 dogs from 77 American Kennel Club-recognized breeds. This population had been previously noted by Patronek and colleagues (1997) as being unexamined with regards to life span and size. Upon analysis, results confirmed previous studies by indicating that life expectancy is inversely related to size (Deeb and Wolf, 1994; Patronek et al., 1997). Specifically, we determined that longevity is related to both height and weight. Although this trend has been suggested for many years, there has been a lack of information from the general pet dog population.

The significant inverse relationship between life span and weight is -0.679 ($p < 0.05$) while the relationship between life span and height is -0.603 ($p < 0.05$). This work also indicates that 58.5% of variability in life span can be accounted for by variation in weight and height of a breed. However, this analysis indicates that weight is the more significant predictor of longevity with a t-value of -6.091 ($p < 0.001$). The resulting regression equation is: $y = 14.016 + (0.0180 * x_H) - (0.0400 * x_W)$, which offers a reasonable estimation of expected longevity for any given healthy weight and height across breeds. This establishes a quantitative basis from which it is possible to further analyze the genetic components of longevity of the domestic dog.

Completion of the second aim, to identify and map genes associated with the aging process in the canine genome, resulted in 52 gene based markers being added to the existing canine genome map (Breen et al., 2004) as well as the verification of the position of two gene based markers already on the map. From the total of 54 genes investigated, 28 were selected for their proximity to the marker D4S1564 on human chromosome 4 (Puca et al., 2001), and 26 were chosen based on previous associations with aging and/or the aging process.

Of the 54 genes, 45 mapped as expected to regions of the canine genome that exhibit conservation of synteny with regions of the human genome containing these genes (Guyon et al., 2003). Determination of the canine genome sequence provided the tool necessary for analysis of these unexpected results, and led to the discovery of six gene paralogs and one pseudogene. Two of the unexpected results, however, are still unexplained.

Although these paralogs were mapped inadvertently, these data are important because several gene paralogs that affect the aging process have been found in *C. elegans* (Rikke et al., 2000). Additionally, the presence of gene paralogs is thought to be an evolutionary mechanism that results in the development of gene families (Lynch and Conery, 2000).

Genes implicated in the aging process were localized to 19 of the 38 canine autosomes, as well as to the X chromosome. Each gene was positioned close to polymorphic markers, at a minimum distance of less than 2 Mb. This effort to map genes associated with a specific biological mechanism resulted in the development of a valuable resource for genetic linkage studies or genetic association studies.

The third aim, examination of *Ghr*, *Ghrhr*, *Pit1*, and *Prop1*, utilized genomic DNA from nine dog breeds of varying sizes and life expectancies. Ten unrelated dogs from each breed were used to determine if any sequence differences in these genes are responsible for the inverse relationship between size and longevity in the dog. As a result, 53 polymorphisms were discovered. Seventeen were identified in *Ghr*, 26 were found in *Ghrhr*, only five polymorphisms were discovered in *Prop1*, and 25 were identified in *Pit1*.

Thirteen ancestral SNPs were discovered in which both alleles were found in every breed. The presence of these ancestral polymorphisms, in conjunction with historical records, suggest that the Staffordshire Bull Terrier and Great Pyrenees may be added to cluster one and/or two as described by Parker et al (2004) but only after collection of additional data.

Thirty-one polymorphisms with allelic frequencies of at least 10% were selected for further analysis with Fisher's exact test. Of these, 16 presented with statistically significant p-values. Nine of these SNPs were in non-coding regions of the genes and were not located in known regulatory regions, or in highly conserved regions of the genome. Therefore, we do not currently believe that these mutations are the cause of the observed differences in size and longevity. The remaining statistically significant polymorphisms were located in coding regions in *Ghr* and *Ghrhr*, and only three resulted in changes in the amino acid sequence. These changes were evaluated with SIFT software, a program with the criterion that values of 0.05 or less indicate an intolerant change (Ng and Henikoff, 2001). One mutation in *Ghrhr* led to valine being replaced with methionine. The SIFT value for this substitution was calculated to be 0.01, and may well affect the function of this protein. This mutation is statistically significant ($p = 4.8 \times 10^{-6}$) when large dogs are compared to medium-sized breeds, but not when they are compared to small breeds ($p = 0.001$). Genotyping of additional dogs and/or breeds could lead to a significant change being discovered when large and small breeds are compared. Additionally, protein analysis to establish the effect of this mutation might show how this mutation may contribute to determining body size in the dog.

In conclusion, these experiments have further established the dog as a model organism for investigations into the aging process. Statistical analysis of height, weight and life expectancy data revealed a distinct correlation between size and longevity. Mapping experiments enhanced the existing canine genome map, an essential tool for

future studies. Finally, four genes that confer extended life spans to mice of diminished size were analyzed in the dog. Based on this work, *Ghr*, *Ghrhr*, *Pit1* and *Prop1* are not thought to be significantly involved in determining life expectancy in the dog. However, many genes believed to be involved in the aging process remain to be investigated in the dog.

REFERENCES

1. Agarwala R, Applegate DL, Maglott D, Schuler GD, Schaffer AA (2000) A fast and scalable radiation hybrid map construction and integration strategy. Genome Res 3, 350-364

2. Agundez JAG (2004) Cytochrome P450 gene polymorphism and cancer. Curr Drug Metab 5, 211-224

3. Arantes-Oliveira N, Berman JR, Kenyon C (2003) Healthy animals with extreme longevity. Science 302, 611

4. Austad SN (1993) The comparative perspective and choice of animal models in aging. Aging 5, 259-267

5. Barzilai A, Rotman G, Shiloh Y (2002) ATM deficiency and oxidative stress, a new dimension of defective response to DNA damage. DNA Repair 1, 3-25

6. Bathum L, Anderson-Ranberg K, Boldsen J, Brosen K, Jeune B (1998) Genotypes for the cytochrome P450 enzymes CYP2D6 and CYP2C19 in human longevity. Eur J Clin Pharmacol 54, 427-430

7. Belsley DA, Kuh E, Welsch RE (1980) Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. (New York: John Wiley and Sons)

8. Blanche H, Cabanne L, Sahbatou M, Thomas G (2001) A study of French centenarians, are ACE and APOE associated with longevity? C R Acad Sci III 324, 129-135

9. Bonafe M, Barbieri M, Marchegiani F, Olivieri F, Ragno E, et al. (2003) Polymorphic variants of insulin-like growth factor I (IGF-I) receptor and phosphoinositide 3-kinase genes affect IGF-I plasma levels and human longevity, cues for an evolutionarily conserved mechanism of life span control. J Clin Endocrinol Metab 88, 3299-3304

10. Bonnett BN, Egenvall A, Olson P, Hedhammar A (1997) Mortality in insured Swedish dogs: rates and causes of death in various breeds. Vet Rec 141, 40-44

11. Breen M, Hitte C, Lorentzen TD, Thomas R, Cadieu E, et al. (2004) An integrated 4249 marker FISH/RH map of the canine genome. BMC Genomics 5, 65-75

12. Bronson R (1982) Variation in age at death of dogs of different sexes and breeds. Am J Vet Res 43, 2057-2059

13. Brown-Borg HM, Borg KE, Meliska CJ, Bartke A (1996) Dwarf mice and the ageing process. Nature 384, 33

14. Cambien F, Poirier O, Lecerf L, Evans A, Cambou JP, et al. (1992) Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. Nature 359, 641-644

15. Canterberry SC, Greer KG, Hitte C, Andre C, Murphy KE (2005) Aging associated loci in *Canis familiaris*. Growth Dev Aging 69, 101-113

16. Chavous DA, Jackson FR, O'Connor CM (2001) Extension of the *Drosophila* lifespan by overexpression of a protein repair methyltransferase. Proc Natl Acad Sci USA 98, 14814-14818

17. Clancy DJ, Gems D, Harshman LG, Oldham S, Stocker H, et al. (2001) Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein. Science 292, 104-106

18. Cohen HY, Miller C, Bitterman KJ, Wall NR, Hekking B, et al. (2004) Calorie restriction promotes mammalian cell survival by inducing the SIRT1 deacetylase. Science 305, 390-392

19. Cowan CM, Quarto N, Warren SM, Salim A, Longaker MT (2003) Age-related changes in the biomolecular mechanisms of calvarial osteoblast biology affect fibroblast growth factor-2 signaling and osteogenesis. J Biol Chem 278, 32005-32013

20. Crowley J, Adelman B (1998) The Complete Dog Book: Official Publication of the American Kennel Club 19th Edition. (New York: Howell Book House)

21. De Benedictis G, Falcone E, Rose G, Ruffolo R, Spadafora P, et al. (1997) DNA multiallelic systems reveal gene/longevity associations not detected by diallelic systems. The ApoB locus. Hum Genet 99, 312-318

22. De Benedictis G, Carotenuto L, Carrier G, De Luca M, Falcone E, et al. (1998) Age-related changes of the 3'APOB-VNTR genotype pool in ageing cohorts. Ann Hum Genet 62, 115-122

23. De Benedictis G, Tan Q, Jeune B, Christensen K, Ukraintseva S, et al. (2001) Recent advances in human gene-longevity association studies. Mech Ageing Dev 122, 909-920

24. Deeb B, Wolf N (1994) Studying longevity and morbidity in large and small breeds of dogs. Vet Med 89, 702-713

25. Dorman JB, Albinder B, Shroyer T, Kenyon C (1995) The age-1 and daf-2 genes function in a common pathway to control the lifespan of *Caenorhabditis elegans*. Genetics 141, 1399-1406

26. Egenvall A, Bonnett BN, Hedhammar A, Olson P (2005) Mortality in over 350,000 insured Swedish dogs from 1995-2000: II. Breed-specific age and survival patterns and relative risk for causes of death. Acta Vet Scand 46, 121-136

27. Faure-Delanef L, Quere I, Zouali H, Cohen D (1997) Human longevity and R506Q Factor V gene mutation. Thromb Haemost 78, 1160

28. Flurkey K, Papaconstantinou J, Miller RA, Harrison DE (2001) Lifespan extension and delayed immune and collagen aging in mutant mice with defects in growth hormone production. Proc Natl Acad Sci USA 98, 6736-6741

29. Flurkey K, Papaconstantinou J, Harrison DE (2002) The Snell dwarf mutation Pit1(dw) can increase life span in mice. Mech Ageing Dev 123, 121-130

30. Freimer N, Sabatti C (2004) The use of pedigree sib-pair and association studies of common diseases for genetic mapping and epidemiology. Nat Genet 36, 1045-1051

31. Furuyama T, Yamashita H, Kitayama K, Higami Y, Shimokawa I, et al. (2002) Effects of aging and caloric restriction on the gene expression of Foxo1, 3, and 4 (FKHR, FKHRL1, and AFX) in the rat skeletal muscles. Micrsc Res Tech 59, 331-334

32. Garcia-Closas M, Egan KM, Abruzzo J, Newcomb PA, Titus-Emstoff L, et al. (2001) Collection of genomic DNA from adults in epidemiological studies by buccal cytobruch and mouthwash. Cancer Epidemiol Biomarkers Prev 10, 687-696

33. Gavrilov LA, Gavrilova NS (2003) The quest for a general theory of aging and longevity. Sci Aging Knowledge Environ 28, RE5 1-10

34. Gerdes LU, Jeune B, Ranberg KA, Nybo H, Vaupel JW (2000) Estimation of apolipoprotein E genotype-specific relative mortality risks from the distribution of genotypes in centenarians and middle-aged men, apolipoprotein E gene is a "frailty gene," not a "longevity gene". Genet Epidemiol 19, 202-210

35. Godfrey P, Rahal JO, Beamer WG, Copeland NG, Jenkins NA, et al. (1993) GHRH receptor of little mice contains a missense mutation in the extracellular domain that disrupts receptor function. Nat Genet 4, 227-231

36. Greer KA, Canterberry SC, Murphy KE (2006) Statistical analysis regarding the effects of height and weight on the life span of the domestic dog. Res Vet Sci, in press

37. Grube K, Burkle A (1992) Poly(ADP-ribose) polymerase activity in mononuclear leukocytes of 13 mammalian species correlates with species-specific life span. Proc Natl Acad Sci USA 89, 11759-11763

38. Guyon R, Lorentzen TD, Hitte C, Kim L, Cadieu E, et al. (2003) A 1-Mb resolution radiation hybrid map of the canine genome. Proc Natl Acad Sci USA 100, 5296-5301

39. Heijmans B, Westendorp R, Slagboom P (2000) Common gene variants, mortality and extreme longevity in humans. Exp Gerontol 35, 865-877

40. Helfand S, Rogina B (2003) Genetics of aging in the fruit fly, *Drospophila melanogaster*. Ann Rev Genet 37, 329-348

41. Jian-Gang Z, Yong-Xing M, Chuan-Fu W, Pei-Fang L, Song-Bai Z,et al. (1998) Apolipoprotein E and longevity among Han Chinese population. Mech Ageing Dev 104, 159-167

42. Jiang Z, Michal JJ, Melville JS, Baltzer HL (2005) Multi-alignment of orthologous genome regions in five species provides new insights into the evolutionary make-up of mammalian genomes. Chromosome Res 13, 707-715

43. Kalinichenko VV, Gusarova GA, Tan Y, Wang IC, Major ML, et al. (2003) Ubiquitous expression of the forkhead box M1B transgene accelerates proliferation of distinct pulmonary cell types following lung injury. J Biol Chem 278, 37888-37894

44. Kalinina OA, Kalinin SA, Polack EW, Mikaelian I, Panda S, et al. (2003) Sustained hepatic expression of FoxM1B in transgenic mice has minimal effects on hepatocellular carcinoma development but increases cell proliferation rates in preneoplastic and early neoplastic lesions. Oncogene 22, 6266-6276

45. Kealy RD, Lawler DF, Ballam JM, Mantz SL, Biery DN, et al. (2002) Effects of diet restriction in life span and age-related changes in dogs. JAVMA 220, 1315-1320

46. Kenyon C, Chang C, Genesh E, Fudner A, Tabtiang R (1993) A *C. elegans* mutant that lives twice as long as wild-type. Nature 366, 461-464

47. Kimura KD, Tissenbaum HA, Liu Y, Ruvkun G (1997) *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. Science 277, 942-946

48. Kirkness EF, Bafna V, Halpern AL, Levy S, Remington K, et al. (2003) The dog genome, survey sequencing and comparative analysis. Science 301, 1898-1903

49. Kohler HP, Grant PJ (2000) Plasminogen activator type 1 and coronary artery disease. N Eng J Med 342, 1792-1801

50. Kokoszka JE, Coskun P, Esposito LA, Wallace DC (2001) Increased mitochondrial oxidative stress in the Sod2 (+/-) mouse results in the age related decline of mitochondrial function culminating in increased apoptosis. Proc Natl Acad Sci USA 98, 2278-2283

51. Kops GJPL, Dansen TB, Polderman PE, Saarloos I, Wirtz KWA, et al. (2002) Forkhead transcription factor FOXO3a protects quiescent cells from oxidative stress. Nature 419, 316-321

52. Krupczak-Hollis K, Wang X, Dennewitz MB, Costa RH (2003) Growth hormone stimulates proliferation of old-aged regenerating liver through forkhead box m1b. Hepatology 38, 1552-1562

53. Lakowski B, Hekimi S (1996) Determination of life-span in *Caenorhabditis elegans* by four clock genes. Science 5264, 1010-1013.

54. Lakowski B, Hekimi S (1998) The genetics of caloric restriction in *Caenorhabditis elegans*. Proc Natl Acad Sci USA 95, 13091-13096

55. Lao JI, Montoriol C, Morer I, Beyer K (2005) Genetic contribution to aging: deleterious and helpful genes define life expectancy. Ann N Y Acad Sci 1057, 50-63

56. Larsen P, Albert PS, Riddle DL (1995) Genes that regulate both development and longevity in *Caenorhabditis elegans*. Genetics 139, 1567-1583

57. Larsen P, Clarke C (2002) Extension of life-span in *Caenorhabditis elegans* by a diet lacking coenzyme Q. Science 295, 120-123

58. Leonard JA, Vila C, Wayne RK (2006) From wild wolf to domestic dog. In The Dog and Its Genome, E.A. Ostrander, U. Giger, K. Lindblad-Toh, eds. (Woodbury, New York: Cold Spring Harbor Laboratory Press), pp. 95-117

59. Li S, Crenshaw BE III, Rawson EJ, Simmons DM, Swanson LW, et al. (1990) Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene pit-1. Nature 347, 528-533

60. Li Y, Deeb B, Pendergrass W, Wolf N (1996) Cellular proliferative capacity and life span in small and large dogs. J Gerontol Biol Sci 51A, B403-B408

61. Liang H, Masoro EJ, Nelson JF, Strong R, McMahan CA, et al. (2003) Genetic mouse models of extended life span. Exp Gerontol 38, 1353-1364

62. Lin K, Dorman JB, Rodan A, Kenyon C (1997) daf-16: An HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans*. Science 278, 1319-1322

63. Lin Y, Seroude L, Benzer S (1998) Extended life-span and stress resistance in the *Drosophila* mutant *Methuselah*. Science 282, 943-946

64. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, et al. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438, 803-819

65. Ljungquist B, Berg S, Lanke J, McClearn GE, Pedersen NL (1998) The effect of genetic factors for longevity: a comparison of identical and fraternal twins in the Swedish Twin Registry. J Gerontol A Biol Sci Med Sci 53, M441-M446

66. Ly D, Lockhart D, Lerner R, Schultz P (2000) Mitotic misregulation and human aging. Science 287, 2486-2492

67. Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science 290, 1151-1155

68. Mari D, Mannucci PM, Duca F, Bertolini S, Franceschi C (1996) Mutant Factor V (Arg506Gln) in healthy centenarians. Lancet 347, 1044

69. Matise TC, Perlin M, Chakravarti A (1994) Automated construction of genetic linkage maps using an expert system (MultiMap), a human genome linkage map. Nat Genet 6, 384-390

70. Matuoka K, Chen KY, Takenawa T (2003) A positive role of phosphatidylinositol 3-kinase in aging phenotype expression in cultured human diploid fibroblasts. Arch Gerontol Geriatr 36, 203-219

71. McKevitt TP, Nasir L, Devlin P, Argyle DJ (2002) Telomere lengths in dogs decrease with increasing donor age. J Nutr 132, 1604S-1606S

72. Merched A, Xia Y, Papadopoulou A, Siest G, Visvikis S (1998) Apolipoprotein AIV codon 360 mutation increases with human aging and is not associated with Alzheimer's disease. Neurosci Lett 242, 117-119

73. Migliaccio E, Giorgio M, Mele S, Pelicci G, Reboldi P, et al. (1999) The $p66^{SHC}$ adaptor protein controls oxidative stress response and life span in mammals. Nature 402, 309-313

74. Miskin R, Masos T (1997) Transgenic mice overexpressing Urokinase-Type Plasminogen activator in the brain exhibit reduced food consumption, body weight and size, and increased longevity. J Gerontol Bio Sci 52A, B118-B124

75. Moechars D, Lorent K, Van Leuven F (1998) Premature death in transgenic mice that overexpress a mutant amyloid precursor protein is preceded by severe neurodegeneration and apoptosis. Neuroscience 91, 819-830

76. Morris JZ, Tissenbaum HA, Ruvkun G (1996) A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in *Caenorhabditis elegans*. Nature 382, 536-539

77. Murakami S, Johnson T (1999) Life extension and stress resistance in Caenorhabditis elegans modulated by the tkr-1 gene. Curr Biol 9, R791

78. National Institute on Aging Well-Being Improves for Most Older People, But Not For All, New Federal Report Says http://www.nia.nih.gov/NewsAndEvents/ PressReleases/PR20000810WellBeing.htm (Accessed March 1, 2006)

79. Nembaware V, Crum K, Kelso J, Seoighe C (2002) Impact of the presence of paralogs on sequence divergence in a set of mouse-human orthologs. Genome Res 12, 1370-1376

80. Ng PC, Henikoff S (2001) Predicting deleterious amino acid substitutions. Genome Res 11, 863-874

81. Ogg S, Paradis S, Gottlieb S, Patterson GI, Lee L, et al. (1997) The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. Nature 389, 994-999

82. Online Mendelian Inheritance in Animals (OMIA) Summary of OMIA http://omia.angis.org.au/ (Accessed June 1, 2006)

83. Ostrander EA, Giniger E (1997) Semper fidelis: what man's best friend can teach us about human biology and disease. Am J Hum Genet 61, 475-480

84. Panza F, D'Introno A, Colacicco AM, Capurso C, Capurso S, et al. (2004) Vascular genetic factors and human longevity. Mech Ageing Dev 125, 169-178

85. Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, et al. (2004) Genetic structure of the purebred domestic dog. Science 304, 1160-1164

86. Patronek GJ, Waters DJ, Glickman LT (1997) Comparative longevity of pet dogs and humans, implications for gerontology research. J Gerontol Biol Sci 52A, B171-B178

87. Pepe G, Di Perna V, Resta F, Lovecchio M, Chimienti G, et al. (1998) In search of a biological pattern for human longevity, impact of apoA-IV genetic polymorphisms on lipoproteins and the hyper-Lp(a) in centenarians. Atherosclerosis 137, 407-417

88. Perls TT, Bubrick E, Wager CG, Vijg J, Kruglyak L (1998) Siblings of centenarians live longer. Lancet 351, 1560

89. Poon HF, Joshi G, Sultana R, Farr SA, Banks WA, et al. (2004) Antisense directed at the Aβ region of APP decreases brain oxidative markers in aged senescence accelerated mice. Brain Res 1018, 86-96

90. Puca AA, Daly MJ, Brewster SJ, Matise TC, Barrett J, et al. (2001) A genome-wide scan for linkage to human exceptional longevity identifies a locus on chromosome 4. Proc Natl Acad Sci USA 98, 10505-10508

91. Quarrie J, Riabowol K (2004) Murine models of life span extension. Sci Aging Knowledge Environ 31, RE5 1-11

92. Reed T, Dick DM, Uniacke SK, Foroud T, Nichols WC (2004) Genome-wide scan for a healthy aging phenotype provides support for a locus near D4S1564 promoting healthy aging. J Gerontol A Biol Sci Med Sci 59, 227-232

93. Rikke BA, Murakami S, Johnson TE (2000) Paralogy and orthology of tyrosine kinases that can extend the life span of *Caenorhabditis elegans*. Mol Biol Evol 17, 671-683

94. Rogina B, Reenan RA, Nilsen SP, Helfand SL (2000) Extended life-span conferred by cotransporter gene mutations in *Drosophila*. Science 290, 2137-2140

95. Ruan H, Tang XD, Chen ML, Joiner ML, Sun G, et al. (2002) High-quality life extension by the enzyme peptide methionine sulfoxide reductase. Proc Natl Acad Sci USA 99, 2748-2753

96. Savolainen P (2006) mtDNA studies of the origin of dogs. In The Dog and Its Genome, E.A. Ostrander, U. Giger, K. Lindblad-Toh, eds. (Woodbury, New York: Cold Spring Harbor Laboratory Press), pp. 119-140

97. Schachter F, Faure-Delanef L, Guenot F, Rouger H, Froguel P, et al. (1994) Genetic associations with human longevity at the APOE and ACE loci. Nat Genet 6, 29-32

98. Slinker BK, Glantz SA (1990) Missing data in two-way analysis of variance. Am J Physiol 258, R291-R297

99. Takeshita K, Yamamoto K, Ito M, Kondo T, Matsushita T, et al. (2002) Increase expression of Plasminogen Activator Inhibitor-1 with fibrin deposition in a murine model of aging, "*Klotho*" mouse. Sem Thromb Hemost 28, 545-553

100. Tatar M, Kopelman A, Epstein D, Tu MP, Yin CM, et al. (2001) A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function. Science 292, 107-110

101. Tissenbaum H, Guarente L (2001) Slow aging during insect reproductive diapause: Why butterflies, grasshoppers and flies are like worms. Exp Gerontol 36, 723-738

102. van der Horst A, Tertoolens LGJ, de Vries-Smits LMM, Frye RA, Medema RH, et al. (2004) FOXO4 is acetylated upon peroxide stress and deacetylated by the longevity protein hSir2^{SIRT1}. J Biol Chem 279, 28873-28879

103. Vignaux F, Hitte C, Priat C, Chuat JC, Andre C, et al. (1999) Construction and optimization of a dog whole-genome radiation hybrid panel. Mamm Genome 10, 888-894

104. Warner HR (2003) Subfield history, use of model organisms in the search for human aging genes. Sci Aging Knowledge Environ 12, 986-1002

105. Weinert BT, Timiras PS (2003) Invited review: theories of aging. J Appl Physiol 95, 1706-1716

106. Weisberg S (1985) Applied Linear Regression, 2nd ed. (New York: John Wiley and Sons)

107. West MD, Shay JW, Wright WE, Linskens MHK (1996) Altered expression of plasminogen activator and plasminogen activator inhibitor during cellular senescence. Exp Gerontol 31,175-193

108. Wise J, Heathcott B, Shepherd A (2003) Results of the 2002 AVMA survey of US pet-owning households regarding use of veterinary services and expenditures. JAVMA 222, 1524-1525

109. Wong KK, Maser RS, Bachoo RM, Menon J, Carrasco DR, et al. (2003) Telomere dysfunction and Atm deficiency compromises organ homeostasis and accelerates ageing. Nature 421, 643-648

VITA

Name:	Sarah Christine Canterberry
Permanent Address:	PO Box 173 Snook, TX 77878
Education:	Texas A&M University College Station, TX B.S., 2001, Animal Science
	Texas A&M University College Station, TX Ph.D., 2006, Genetics