ASSOCIATION OF MEASURES OF MUSCLE DENSITY AND SNPs IN HOMOLOGS OF MUSCLE FUNCTION GENES IN *C.ELEGANS*

by

Sevinc Alkan

B.S, Ankara University, Turkey, 2008

Submitted to the Graduate Faculty of

the Department of Human Genetics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Sevinc Alkan

It was defended on

August 1, 2012

and approved by

Robert E. Ferrell, Ph.D.
Professor, Department of Human Genetics
Graduate School of Public Health, University of Pittsburgh

Alfred L. Fisher. M.D., Ph.D.
Assistant Professor, Department of Medicine
School of Medicine, University of Pittsburgh

Thesis Director: Candace M. Kammerer, Ph.D.
Associate Professor, Department of Human Genetics
Graduate School of Public Health, University of Pittsburgh

ASSOCIATION OF MEASURES OF MUSCLE DENSITY AND SNPs IN HOMOLOGS OF MUSCLE FUNCTION GENES IN *C.ELEGANS*

Sevinc Alkan, M.S.

University of Pittsburgh, 2012

Sarcopenia, the loss of muscle mass and strength that occurs with aging, has a crucial role in development of frailty. Sarcopenia is recognized as one of the major public health problems affecting approximately 8-40% of individuals greater than age 60 years worldwide. Although sarcopenia and other muscle-related traits are heritable, the underlying genetic variation contributing to the development of sarcopenia is unclear. Previously, researchers identified 18 genes that will delay the onset of sarcopenia in an animal model for aging, *Caenorhabditis elegans* (Kashyap, 2010), of which 14 have a human homolog and SNP genotypes that were assayed in individuals from the Health, Aging, and Body Composition Study (Health ABC). The primary aim of this thesis is to assess whether there is a relationship between >700 single nucleotide polymorphisms (SNPs) total in 27 genes [14 genes identified from *C. elegans* and 13 genes in the mevalonate and ubiquinone pathways] and six traits related to sarcopenia using data from the HealthABC cohort. The 6 outcome variables are isokinetic leg muscle maximum torque(KCTMAX), thigh muscle density(THMUSD), thigh intermuscular fat area(THIMF), total lean mass (TOTLEAN), total percent fat(TOTPF) and thigh total muscle area(THMUS). The European and African American cohort were analyzed separately. After adjusting for the effects of gender, age, and ancestry, single SNP: single trait association tests were performed. Although no SNPs were statistically significant at the experiment-wide p-value ($p< 0.00001$), SNPs in the locus *MAGOHB* were associated with THIMF, THMUSD, and TOTPF in European Americans, whereas SNPs in *CYP3A5* were associated with these traits in African Americans ($p<0.01$ for all). Furthermore, SNPs in *GOLGA4* ($P<0.01$) and *RALGABP* ($P<0.001$) were associated with TOTPF in African Americans, and SNPs in

RALGABP were also associated with THIMF and TOTLEAN in African Americans. These results indicate that variation in some genes related to development of muscle-wasting in an animal model, C. elegans, may have pleiotropic effects on traits related to sarcopenia in older men and women. Follow-up studies of variation in these genes could elucidate new biological pathways involved in development of sarcopenia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to take this opportunity to express my appreciation to the many people who have supported and encouraged me in this project and have brought me thus far in my education.

First of all, I would like to express my sincere appreciation to my thesis advisor Candace Kammerer who made my thesis successful. I am sure that without her continuous encouragement, support, and wise advice, this work would not have been successful. Working on this project under her guidance has helped me to become a better scientist. Thank you for your time and thoughtfulness.

My special thanks go to my colleague Dr. Ryan Minster who helped me whenever I needed. Thanks to his assistance and intellectual engagement with my project, my project became successful. Also I would like to thank the members of my thesis committee, Drs. Alfred Fisher and Robert Ferrell for the original proposal of this study and their comments and insights.

I would like to thank all the participants of Health, Aging and Body Composition Study.

I would like to thank Turkish Ministry of Education for providing scholarship throughout my graduation study.

I would like to express my gratitude to my loving family for their unconditional support, help and guidance from the beginning of my career path. Although I may have been hundreds of miles away, they have been always with me with their love and encouragement.

Finally, I would like to thank my beloved husband, Ahmet Oguz Akcay who shares my burdens and my joys. I could not be successful without his support, love and extraordinary courage throughout this project and in every walk of life.

1.0 BACKGROUND AND SIGNIFICANCE

1.1 SARCOPENIA

1.1.1 Public Health Significance of Sarcopenia

Sarcopenia is characterized by a decline in muscle quantity and quality and it is also a progressive loss of muscle mass with advancing age (Evans, 1995). The prevalence of sarcopenia ranges from 8 to 40% of people over 60 years (Kan, 2009). Sarcopenia has a crucial role in the pathogenesis of frailty with advancing age (Morley, 2001). The decline in the muscle strength results in decreased mobility and increased mortality in older people (Janssen, 2002). The consequences of sarcopenia are decreased energy expenditure, decreased insulin sensitivity, reduction in muscle mass and strength, and increased physical disability, falls, and mortality. Sarcopenia has emerged as a highly important public health problem recently, because sarcopenia is associated with other common diseases or health problems including airway diseases (Sermet-Gaudelus et al., 2003), sarcopenic obesity (Roubenoff, 2000, 2004), osteoporosis (Szulc et al., 2005), cardiovascular diseases (Giles et al., 2006), and physical disability (Janssen, 2006). The number of older people who have health problems related to sarcopenia continues to increase all over the world. As a result, the economic burden of the sarcopenia-related health-care cost is large, and the health-care costs associated with sarcopenia were estimated to be $18.5 billion dollars in 2000 (Jansssen et al., 2004).

Age, race, and gender influence the prevalence of sarcopenia. In 1998 in the U.S., the prevalence of sarcopenia in men was as follows: 14% of men younger than 70 years, 20% of men 70-74 years, 27% of men 75-80 years, and 53% of men older than 80 years (Baumgartner et al., 1998).

Likewise, the prevalence in women was also high: 25% of women younger than 70 years, 33% of women were 70-74 years, 36% of women aged 75-80 years, and 43% of women aged greater than 80 years had sarcopenia (Baumgartner et al., 1998). According to Iannuzzi-Sucich et al. (2002), the sarcopenia prevalence in Caucasians aged 65 years or older was 22.6% in women and 26.8% in men, whereas among Caucasians, aged 80 years or older, it was 31.0% in women and 52.9% in men. On the other hand, French, Chinese and Mexican men and women have lower rates of sarcopenia compared to Caucasians, likely due to cultural effects. The prevalence of sarcopenia in Mexicans less than 70 years old was 13-24%, but greater than 50% among individuals over 80 years old (Baumgartner et al., 1998). The prevalence of sarcopenia in Chinese was low, 12.3% in men and 7.6% in women older than 70 years (Lau et al., 2005).Likewise the prevalence of sarcopenia in a French women group older than 80 years was also low and equaled 10.9%(Gillette-Guyenne et al., 2003). Thus, the prevalence of sarcopenia is high, and will become higher as the world-wide population ages.

1.1.2 Definition of Sarcopenia

I.H. Rosenberg gave the name sarcopenia, that comes from Greek roots, *sarx*(flesh) and *penia*(loss) which describes skeletal muscle loss (Vandervoort, 2002).There are many definitions of sarcopenia to date (Newman et al. 2003; Visser 2009). One definition is that an individual has sarcopenia if his/her gait speed less than 1m/s and his/her appendicular lean/fat mass ratio is less than two standard deviations below the mean of young adults (Evans 2010). The most common definition of sarcopenia is that the ratio of appendicular skeletal muscle mass to height squared is more than two standard deviations below the mean of young adults (Evans, 1997; Baumgartner et al., 1998). This definition of sarcopenia requires measurements obtained from dual-energy X-ray absorptiometry (DXA) methods. DXA has led to easy and cheap measurement of body composition and is a great tool for research on skeletal muscle mass.

The current diagnosis of sarcopenia requires a decrease in the whole body or appendicular fat-free mass combined with poor physical functioning (Newman, 2003). The current sarcopenia cut points are an appendicular fat lean mass/ht^2 ≤7.23 kg/m^2 in men and ≤5.67 kg/m^2 in women.

1.1.3 Risk factors of Sarcopenia

Sarcopenia has a complex and multifactorial etiology. The etiology of sarcopenia changes with age. Between ages 20 and 40 years, as a result of a decline in physical activity and a type II muscle fiber size, the sprinting capacity becomes lower. Between ages 40 and 60 years, as a result of the rapid decline in the motor units, decline in physical activity and increase in body fat, there is a decline in aerobic capacity and muscle protein synthesis. At the ages of 60-70 years, the physical activity is very low and body fat is very high, thus the declines in muscle protein synthesis are huge. After the age of 70 years, there is a further reduction in physical activity that causes increased muscle protein breakdown.

Skeletal muscle declines as a proportion of body mass with increasing age. At ages 21-30 years, skeletal muscle represents approximately 45% of total body mass, whereas among individuals greater than 70 years of age, skeletal muscle represents only 27% of total body mass (Tan and Culberson, 2003). Protein constitutes approximately 20% of muscle weight and myosin and actin are the major muscle proteins. There are three main muscle fiber categories including type I, type IIA and type IIX. Although type I fibers have slow contraction and high endurance, type IIX fibers have faster contraction and lower endurance. Type IIA fibers have faster contraction than type I fibers, whereas type IIA fibers have slower contraction than type IIX fibers (Barany 1967; Thomas et al. 2007). The muscle function depends on the proportions of type I, type IIA and type IIX fibers. There is a decline in the proportion of type IIA fiber in sarcopenia (Leger et al., 2008).

As a result of human and laboratory animal studies, there is some evidence about the possible mechanisms of sarcopenia including, decrease in muscle self-repair ability (Collins et al. 2005), inactivity,

decreased protein intake, loss of α-motor neurons, hormonal changes, inflammatory effects (increased IL-6), altered caloric intake, and changes in muscle cell physiology. However, the main mechanism of sarcopenia is still unclear (Volpi, 2004, Greenlund, 2003). This lack of knowledge regarding mechanisms also limits the effective treatment of sarcopenia. Thus, there is a need to develop new model systems to understand the mechanism of sarcopenia with subsequent follow-up studies in humans. Results from these studies could lead to the development of new pharmacological agents for the effective treatment of sarcopenia.

1.1.4 Genetic determination of Sarcopenia

Two most commonly studied sarcopenia-related phenotypes are muscle mass and muscle strength. Muscle mass and muscle strength have multiple factors that differ among individuals. Although the variation among individuals is mostly determined by genetic factors, the specific genes that control these phenotypes are mostly unknown (Tan et al., 2011). Whole body lean mass (LM), DXA-derived, is one of the most used indices for muscle mass (Kim et al.2004, 2006b). Fat-free mass (FFM) and muscle size (cross section area) are other commonly studied phenotypes that are related to sarcopenia. The heritability of body lean mass, fat-free mass and muscle size phenotypes is very high. The heritability of LM ranges from 52% to 80% (Deng et al., 2001; Hsu et al.2005). The genetic contributions of fat-free mass and muscle size are 45-65% (Abney et al, 2001) and 70-90% (Huygens et al., 2004c), respectively.

Results of segregation analyses have identified several loci that influence variation in whole body lean mass (Rice et al., 1993). Segregation analyses in Chinese (Liu et al.2004a) and other populations (Borecki et al., 1998; Karasik and Kiel 2008; Lecomte et al.1997) have confirmed that there are many genes with moderate effects that determine variations in whole body lean mass.

There is significant evidence that many traits related to muscle strength have high heritability (h^2). The heritability of hand grip strength(GS), quadriceps strength(QS) and lower limb muscle

strength(LLMS) were estimated as 30-52%(Arden and Spector 1997;Fredriksen et al., 2002), 31-78%(Thomis et al.,1998;Tiainen et al.,2004) and 42-64% (Frederiksen et al.,2002) based on twin studies.

With particular relevance to sarcopenia are reports that older people have showed significant heritability for muscle-related traits (Frederiksen et al., 2002; Mikkola et al., 2009). According to Prior et al, the heritability of appendicular lean mass was 20% among individuals older than 45 years (Prior et al, 2007). In addition, genes account for 14% of the variance in quadriceps strength and 31% of variance in knee extension strength (Tiainen et al., 2004). Furthermore, the change in muscle strength with older age is also heritable (Zhai et al, 2004).

To further investigate genetic basis of sarcopenia and muscle-related traits, additional genetic studies have been performed, including whole-genome linkage studies in humans, quantitative trait loci mapping in animal models, DNA microarrays and microRNA studies, genome-wide association studies, and candidate gene association studies of sarcopenia and sarcopenia-related skeletal muscle phenotypes. The results of some of these studies are described below.

As a result of genome-wide linkage search for genes associated with fat-free mass (FFM), three quantitative trait loci (QTLs) associated with FFM were detected (Chagnon et al. 2000). One QTL on chromosome 12 was linked to a CA repeat polymorphism within the insulin-like growth factor 1 receptor (IGF1R) gene(LOD=3.56).); LOD scores greater than 3.0 and 2.5 are considered significant and suggestive evidence, respectively, for linkage to a QTL. The second QTL showed evidence for linkage to a QTL on chromosome 18 (LOD=3.58).The last QTL showed moderate evidence for linkage to QTL on chromosome 7 (LOD=2.72) (Chagnon et al. 2000). Zha and colleagues performed a large-scale linkage study and reported evidence of a QTL on chromosome 5 (LOD=2.54) for lean mass (LM) (Zhao et al. 2007). In addition, the marker that had the highest LOD score (LOD=4.49) was located on chromosome 12q24.3 near the IGF1 gene (Zhao et al. 2007).

There is only one linkage study performed on older people. Investigators performed linkage analyses using 397 microsatellites genotypes in 217 twin pairs from the Finnish Twin Study who were 66-75 years of age. They detected a significant linkage between knee extensor isometric strength and a QTL on chromosome 8 (Tiainen et al.2008). Candidate gene studies of possible associations with sarcopenia have included a limited number of genes and have only analyzed a few SNPs within these genes. The results of the candidate gene studies have reported inconsistent results, probably because of population stratification, ethnic variability, sample size differences and gender differences (Tan et al.,2002).

The number of linkage studies for sarcopenia-related phenotypes is very low. Based on the results of two studies, only 7q21, 7q32, 7p15.3, 15p13, and 20q13 were significant (Tan et al.,2002).

There are a few genes identified that are related to variation in skeletal strength or skeletal mass. Based on the results of two or more studies, the ACE, ACTN3, MSTN, CNTF and VDR genes have been found to be associated with skeletal strength and skeletal mass (Tan et al.,2002). The strong genetic contribution of MSTN gene to variation in skeletal muscle phenotypes was confirmed by association studies, linkage studies, and expression studies. The IGF1 and IL-6 are also potential significant genes that were supported by linkage and association studies. Identification of genes that influence sarcopenia-related phenotypes could lead to development of new pharmacological interventions (Tan et al., 2002).

1.2 ANIMAL MODEL FOR MUSCLE LOSS: *CAERNOHABDITIS ELEGANS*

The nematode *C.elegans* has become a powerful model system in several areas of biology in recent years (Baumeister, 2002). *C. elegans* has a variety of experimental advantages including (1) it can be easily and cheaply cultured in the lab, (2) it has a short life span, (3) it has a completely sequenced genome , and (4) the ability to use RNA interference (Timmons,1998). Under laboratory conditions,

worms develop from egg to adult in 3 days, and reproduce at 4 or 5 days. The adult *C.elegans* has 12 to 18 days lifespan. *C.elegans* has also been used commonly for genetic studies of longevity (Glenn, 2004). Using genetic studies, investigators determined that longevity in *C. elegans* is regulated by genes involved in the insulin and insulin-like growth factor (IGF) signaling pathway (Kimura, 1997). The Insulin-like pathway includes the *daf-2* gene that encodes a protein similar to vertebrate insulin and IGF-I receptors (Kimura et al., 1997; Kenyon et al., 1993). Insulin-like pathways are found to be evolutionarily conserved in *C.elegans* (Tatar et al., 2003).

As in humans, *C. elegans* also develops sarcopenia that is characterized by muscle loss that starts in midlife, increases with advancing age, and results in decreased mobility (Herndon, 2002, Glenn 2004). The sarcopenia observed during aging has significant effects on worms. Because of the structural changes in the muscles, there is decreased mobility in the affected worms (Herndon et al.2002). Worms at young ages have increased physical activity that includes sinusoidal swimming in the agar plate. With aging of worms, there is an obvious physical decline in the physical activity (Driscoll, 2002). The declines in the mobility of worms were irreversible (Herndon et al., 2002).The decreased mobility is due to disorganized myofibrils and decreased myosin filaments (Herndon, 2002). Sarcopenia does not only affect the mobility of worms, but it also a major predictor of mortality during aging (Herndon et al., 2002).Investigators have shown that mutations in *daf-2* and *age-1,* two genes that are known to extend lifespan, also can lead to protection of animals from declines in movement during aging (Duhon, 1995; Morris et al., 1996; Larsen et al., 1995). In addition, there was also a delay in the locomotory declines in *C.elegans* that live long (Huang et al., 2004). These *daf-2* mutants not only showed increased longevity, they also showed increased resistance to oxidative stress, hypoxia, heat stress, heavy metals, and bacterial pathogens (Antebi, 2007). Herdon and colleagues reported that aging in *C.elegans* is associated with sarcopenia (Herndon et al., 2002). Other investigators reported that animals with specific mutations in *daf-2* pathway are resistant to the development of sarcopenia (Duhon, 1995, Herndon,

2002, Glenn, 2004). Using RNA interference in muscle, Kashyap and colleagues subsequently identified 18 genes which are required for the delay in the onset of sarcopenia in *daf-2* mutations (Kashyap, 2010).The identified genes included splicing factors, vacuolar sorting proteins, transcription factors, and metabolic enzymes.

1.3 SPECIFIC AIMS

The main aim of this study was to assess whether SNPs in total of 27 genes (including 14 that are human homologs of *C.elegans* muscle function genes and 13 that are mevalonate and ubiquinone pathway genes) influence muscle-related traits in the Health, Age and Body Composition (HealthABC) cohort (details of this cohort are described below). The specific genes are listed in Table 3 below. I also analyzed data available on the following muscle-related traits include isokinetic leg muscle maximum torque (KCTMAX), thigh intermuscular fat area(THMUS), thigh total muscle density (THMUSD), thigh total muscle density (THMUSD), total lean mass (TOTLEAN,) and total percent fat (TOTPF).

The following are the aims of this study:

1. Perform Single SNP-Single trait association using simple regression analysis
2. Use a few bioinformatic methods, including linkage disequilibrium among the SNPs within the genes, to facilitate identification of possible functional SNPs.

2.0 SUBJECTS AND METHODS

2.1 SUBJECTS

I used data from the Health, Aging and Body Composition Study (HealthABC). Health ABC is a longitudinal study of elderly European American and African American men and women who were healthy at the initial visit. The major objective of the HealthABC study is to evaluate changes in body composition with increasing age. Another aim of the Health ABC study is to identify genes that influence variation in body composition among an older age group. Briefly, the HealthABC study population consists of 3,075 well-functioning people from both sexes, who were 70-79 years old at baseline. Data on similar numbers of each sex by ancestry group were available, although slightly more women (N=1584) than men (N=1491) were recruited (Table 1). Likewise, more European (N=1794) than African American (N=1281) subjects were recruited. Study subjects were recruited from Pittsburgh, Pennsylvania (N=1527) and Memphis, Tennessee (N=1548) in March 1997- July 1998. The University of Pittsburgh and University of Tennessee institutional review board approved the protocol and all participants gave informed consent. European Americans and African Americans were recruited from all eligible ages in Pittsburgh and Memphis. Height was measured by a stadiometer and weight was measured by a calibrated balance-beam scale. Subjects' height and weight were measured while they were wearing a hospital gown and no shoes. Body mass index was also one of the body composition measures. Intermuscular fat area of the left thigh and right thigh were measured by CT. Total lean body mass is measured by DXA (Goodpaster et al. 2006) and thigh muscle cross-sectional area are measured by computed tomography (CT). (Goodpaster et al. 2006).

The summary of phenotypes can be found in Table 2. Genotype data were available on a total of 1737 SNPs across 14 genes which are human homologs of *C. elegans* muscle function genes and 13 mevalonate and ubiquinone pathway genes.

Table 1: HealthABC Characteristics

		GENDER		
		Male	**Female**	**Total**
RACE	**White**	939	855	1794
	Black	552	729	1281
	Total	1491	1584	3075

2.2 PHENOTYPES EXAMINED IN THIS STUDY

Sarcopenia is loss of muscle mass and strength and increased muscle fat infiltration. Isokinetic leg muscle maximum torque (KCTMAX) is a measure of muscle strength and TOTLEAN, THMUS, and THMUSD were used for assessing the muscle mass changes. There is also an association between fat accumulation within skeletal muscle and muscle weakness (Goodpaster 2000), poor function, and increase in the mobility limitations (Visser 2002). Goodpaster previously reported that lower muscle attenuation values on computed tomography (CT), a marker of greater muscle fat content (Goodpaster 2000) is associated with lower muscle strength and muscle quality (Goodpaster 2001). A higher intermuscular fat area (IMF) is associated with several conditions related to aging such as insulin resistance (Goodpaster 2000, Greco 2002, Albu 2005) and type 2 diabetes mellitus (Ohkawa 2005). IMF is also associated with age (Ohkawa 2005). Based on these findings we included thigh intermuscular fat area as one of our phenotypes.

Goodpaster et al. reported that African American women from the HealthABC cohort had higher body fat and higher total body mass, and total fat mass than European American women (Goodpaster,

2009). African American men had a higher total thigh muscle area than European American men. Similarly, African American women had a higher total thigh muscle area than European American women (Goodpaster, 2009).

Table 2: HealthABC summary of phenotypes across both sexes and ancestral groups

Phenotypes	Min	1st Qu.	Median	Mean	3rd Qu.	Max	N
KCTMAX(nm)	12.30	78.91	100.60	106.30	132.40	461.70	2646
THIMF(cm sq)	0.100	6.048	8.955	10.310	13.030	91.250	3011
THMUS (cm sq)	47.77	89.93	109.00	111.40	130.50	226.20	3011
THMUSD (HU)	2.425	31.110	36.600	35.380	40.270	70.060	3011
TOTLEAN(gm)	247.50	406.60	483.90	89.50	565.40	839.80	3075
TOTPF (%)	13.10	28.90	34.60	34.99	41.20	56.00	3075
Weight (kg)	33.50	65.50	75.20	75.81	85.00	141.00	
Height(cm)	1370	1590	1658	1662	1732	2007	

Phenotypes examined in this study are;

1. KCTMAX: isokinetic leg muscle maximum torque (NM): 80-40deg
2. THIMF: Mean of right and left thigh intermuscular fat area (cm sq)
3. THMUS: Mean of right and left thigh total muscle area (cm sq)
4. THMUSD: Mean of right and left thigh total muscle density (HU)
5. TOTLEAN: total lean (gm)
6. TOTPF: total percent fat

2.2 GENOTYPING AND IMPUTATION

Whole genome genotyping of the HealthABC participants was performed by the Center for Inherited Disease Research (CIDR) using the Illumina Human 1M-Duo BeadChip system. The HapMap imputation and initial quality control was performed by Yongmei Liu and Kurt Lohman of Wake Forest University. The following quality control measures were implemented by the Wake Forest team: data on individual samples were removed from the data if the genotyping for the sample failed overall (<97% SNPs genotyped), if the chromosome sex did not match the reported sex or if the first-degree relatedness was detected using the single-nucleotide polymorphism (SNP) data. All data on specific SNPs were removed from the data file if the SNP had a minor allele frequency (MAF) ≤ 1%, was called with ≤97% success, or had a Hardy-Weinberg equilibrium (HWE) test *p* value ≤ 10^{-6}. A total of 1,151,215 autosomal SNPs were successfully genotyped in 1.663 European American individuals and were carried forward to imputation.

Imputation was performed using MACH 1.0.16 (Li et al. 2010b) and the HapMap II phased haplotypes (Frazer et al.2007) as the reference. Genotypes were available for 914,263 SNPs based on the HapMap Centre d'Étude du Polymorphisme Humain (CEPH) reference panel (rel. 22, b36). A total of 2,543,887 genotyped and imputed autosomal SNPs were ultimately available for analysis as part of the "HapMap SNP" set.

Quality control protocols for genotyping in African Americans were identical to European Americans; samples were excluded for low call rate (<97% SNPs were genotyped), mismatch between reported and chromosomal sex, and first degree relatedness. 1,151,215 SNPs were successfully genotyped in 1,139 African American individuals. SNPs with an MAF > 1%, a call rate > 97% and a HWE test *p* value > 10^{-6} were used for imputation. A total of 3,021,329 SNPs were ultimately available for analysis in the African Americans. Imputation was performed by Yongmei Liu and Kurt Lohman at Wake

Forest University using MACH 1.0.16 (Li et al.2010b) and the HapMap II phased haplotypes (Frazer et al.2007) as reference.

For the single SNP-single trait association analysis whole genome genotyping data were available on 1794 European American and 1281 African Americans. I included all SNPs (both assayed an imputed) that were located within the 27 genes of interest, plus all SNPs that were within ±5kb of the 5' and 3' untranslated regions of the 27 genes. The number of assayed plus imputed SNPs available (for the 27 gens of interest) was 736 for European Americans and 999 for African Americans. For the association analyses in the current study, SNPs with the minor allele frequencies less than 0.05 were excluded. After MAF less than 0.05 removals, there were 627 SNPs in European American and 778 SNPs in African Americans. The number of SNPs within each gene after removal of MAF and the results are given in Table 3.

Table 3: List of Candidate Genes (and number of SNPs for African Americans and European Americans) used in the current study

	African Americans **Total number of SNPs= 778**	**European Americans** **Total number of SNPs =627**
ADCK3	45	41
CFL2	6	5
COQ2	15	12
COQ3	24	23
COQ4	7	4
COQ5	17	16
COQ6	26	24
COQ7	11	9
CYP3A5	24	8
FDPS	7	7
GOLGA4	64	35
HMGCR	12	17
IDI1	23	23
JUP	22	18
MAGOHB	62	46
MVD	10	13
MVK	27	32
NPC1	47	40
PDSS1	62	53
PMVK	14	13
RALGAPB	81	30
SACM1L	74	73
SLC27A4	6	5
TBC1D24	9	9
TBP	34	34
TOP1	42	36
VPS33A	13	12

2.3 STATISTICAL METHODS

All statistical analyses were done using the R 2.11.1 [R Development Core Team 2008]. First, I analyzed the distribution of the six muscle-related traits to assess the outliers and deviations from normality. If necessary, the traits were transformed to reduce non-normality and values that were ±4sd were removed prior to subsequent analyses. Next, I did linear regression analyses to identify the effects of significant covariates, such as age, and sex. Because unknown admixture among individuals can bias results of population-based association studies, I also included measures of admixture using principle components of admixture derived by the Wake Forest team. Based on the results of these linear regression analyses, I removed the effects of significant covariates, including admixture. All subsequent analyses were performed using these residual, transformed traits. I also estimated correlation among traits using the residuals. Finally, because some of the genotypes contained imputed data, the genotype data was represented as "probable allele dosage". In other words, the genotype score ranged from 0 – 2. Single SNP-single trait analysis were done using PLINK (Whole genome association analysis toolset) 1.07 (Purcell, 2010). To control for multiple testing, results were not considered to be statistically significant unless the p-value was approximately $\leq 5\times10^{-5}$. Multiple testing is one of the most important concerns in association tests (Yongchao et al.2003). Single SNP-single trait association tests require many tests. Bonferroni method correction is one of the methods to decrease the concerns about multiple testing (Simes 1996). In our study, we used 6 traits and approximately 1400 SNPs and we are performing 8400 tests. Thus, by chance, at *p* value 0.05, we expect 410 false positive results. If we adjusted for multiple testing in our association study, we would need a p value= 0.05/8400 which is approximately 5×10^{-5}. With Bonferroni method, we reduced the number of false positives in our study. Effects of possibly significant SNPs on the traits were plotted using box and whiskers plots. Finally, for

genes with multiple, potentially significant SNPs, I used LocusZoom 1.1 (Pruim,2010) to assess the location of the significant SNPs and linkage disequilibrium among the SNPs.

3.0 RESULTS

3.1 PHENOTYPE CHARACTERIZATION

3.1.1 Distribution and outliers

Statistical analyses were performed in R 2.11.1 [R Development Core Team 2008]. All 6 phenotypes underwent quality control procedures including transformations to normality and outlier removal. KCTMAX (Figure 1 and Figure 2). THMUS, THMUSD and TOTLEAN were transformed by square roots. TOTPF was not transformed. THIMF was transformed by natural logarithms (Figure 3 and Figure 4). Observations that were greater than four standard deviations from the mean were classified as outliers and were removed. Values that were ±4sd were removed prior to subsequent analyses. In the HealthABC cohort, only one observation was removed from THIMF, THMUSD and KCTMAX. No other observation was removed from THMUS, TOTLEAN, and TOTPF.

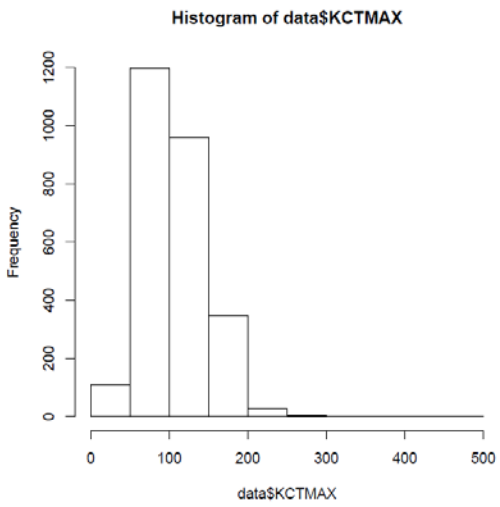

Figure 1: Histogram of KCTMAX

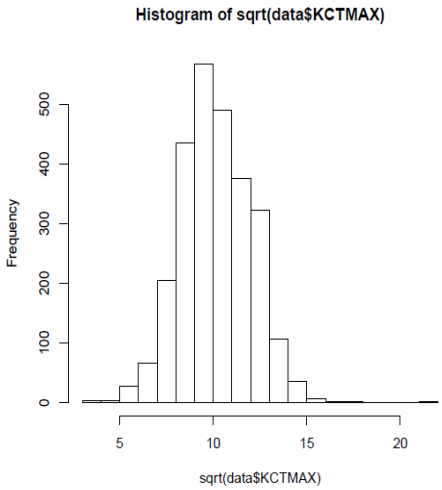

Figure 2: Histogram of sqrtKCTMAX

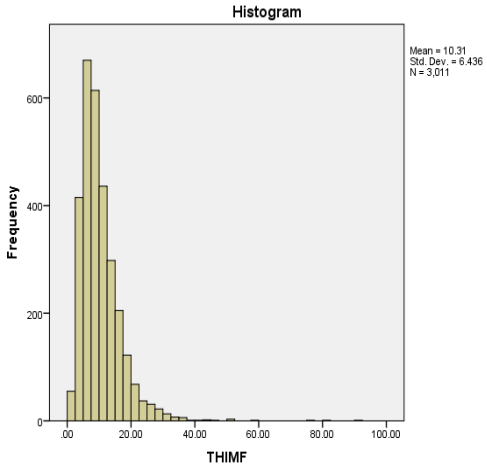

Figure 3: Histogram of THIMF

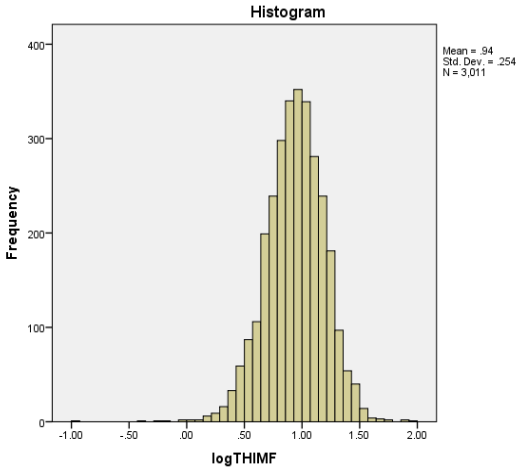

Figure 4: Histogram of logTHIMF

3.1.2 Effects of covariates

We analyzed effect of covariates (age, sex, and measures of admixture) and they were included in the regression models (Table 4). To be conservative, including these results as well as information from the

literature (See Introduction), I removed the effects of age, sex, and admixture in European and African Americans separately. The characteristics of residuals are given in table 5 and 6. These residuals were used in all subsequent analyses.

Table 4: Effects of covariates (beta-values) on the muscle-related traits in European and African Americans

	KCTMAX sqrt (nm)	TOTLEAN sqrt (gm)	THIMF log (cm sq)	THMUS Log (cm sq)	THMUSD sqrt (HU)	TOTPF (%)
European Americans						
Sex	-0.704**	-0.823**	-0.034(NS)	-0.789**	-0.197**	0.705**
Age	-0.130**	-0.056(NS)	-0.002(NS)	-0.082**	-0.089**	-0.051(NS)
PC1	0.083**	0.064(NS)	0.159**	-0.005(NS)	-0.136**	0.051(NS)
PC2	-0.037**	-0.001(NS)	0.010**	-0.005(NS)	-0.003**	-0.017(NS)
African Americans						
Sex	-0.613**	-0.704**	0.160**	-0.666**	-0.314**	0.746**
Age	-0.113**	-0.100(NS)	-0.039(NS)	-0.134**	-0.111**	-0.062(NS)
PC1	-0.057 (NS)	-0.102**	-0.043(NS)	-0.101**	-0.040(NS)	0.093**
PC2	-0.029(NS)	-0.055**	-0.025(NS)	-0.035**	-0. 019(NS)	0.035**
	** p< 0.001					

Table 5: Summary of residuals of transformed traits in European Americans

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max
SqrtKCTMAX(nm)	-5.186	-0.739	0.095	0.00	0.802	9.600
SqrtTOTLEAN(gm)	-39.69	-9.47	-0.58	0.00	8.74	56.58
LogTHIMF(cm sq)	-2.67	-0.33	0.03	0.00	0.36	2.37
LogTHMUS(cmsq)	-2.59	-0.53	0.01	0.00	0.53	2.85
SqrtTHMUSD(HU)	-3.22	-0.31	0.04	0.00	0.39	2.57
TOTPF (%)	-1374	-228.3	-14.61	0.00	234.5	1477

Table 6: Summary of residuals of transformed traits in African Americans

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max
SqrtKCTMAX(nm)	-5.34	-0.84	0.14	0.00	0.96	4.50
SqrtTOTLEAN(gm)	-52.51	-9.62	-0.11	0.00	9.99	45.21
LogTHIMF(cm sq)	-4.45	-0.36	0.006	0.00	0.41	2.20
LogTHMUS(cmsq)	-4.13	-0.55	0.05	0.00	0.58	3.18
SqrtTHMUSD(HU)	-4.03	-0.34	0.05	0.00	0.42	1.41
TOTPF (%)	-1365	-262.4	-6.59	0.00	235.4	1323

3.1.3 Correlation among traits

Phenotypic correlations were computed in order to assess the co-variation among the muscle-related traits. Traits with high correlations may show similar results from the subsequent association analyses. All analyses were done overall and within European Americans and African Americans. The highest

phenotypic correlations were between KCTMAX, TOTLEAN, and THMUS (Table 7).The correlations between KCTMAX and TOTLEAN, KCTMAX and THMUS, and TOTLEAN and THMUS were 0.72, 0.75 and 0.90 (all p-value < 00001), respectively. THIMF and THMUSD were significantly, but negatively correlated (r=-0.61, p< 0.0001). Although significant, correlations between the remaining phenotypes were moderate and low. The phenotypic correlations among muscle-related phenotypes for European Americans and African Americans were similar (Table 8 and Table 9), however, overall correlations were higher for European Americans. In African Americans, correlations between THIMF and THMUSD, THMUS-KCTMAX, THMUS-TOTLEAN, and KCTMAX-TOTLEAN were 0.61, 0.70, 0.87, and 0.634, respectively (all p-value <0.0001). Similarly, in European Americans, correlations between THIMF-THMUS, THMUS-KCTMAX, THMUS-TOTLEAN, and KCTMAX-TOTLEAN were 0.60, 0.79, 0.92, and 0.77, respectively (all p-value <0.0001).

Table 7: Phenotypic correlations (r) among muscle-related traits (transformed and adjusted for significant covariates)

Correlations

		logTHIMF	sqrtTHMUS	sqrtTHMUSD	sqrtKCTMAX	sqrtTOTLEAN	numericTOTPF
logTHIMF		1	.214**	-.614**	.039*	.275**	.465**
	Sig. (2-tailed)		.000	.000	.046	.000	.000
sqrtTHMUS		.	1	.061**	**.746****	**.901****	-.418**
	Sig. (2-tailed)			.001	.000	.000	.000
sqrtTHMUSD				1	.205**	.015	-.505**
	Sig. (2-tailed)				.000	.422	.000
sqrtKCTMAX					1	**.716****	-.464**
	Sig. (2-tailed)					.000	.000
sqrtTOTLEAN						1	-.415**
	Sig. (2-tailed)						.000
numericTOTPF							1
	Sig. (2-tailed)						

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 8: Phenotypic correlations (r) among muscle- related traits (transformed and adjusted for significant covariates) in European Americans

			logTHIMF	sqrtTHMUS	sqrtTHMUSD	sqrtKCTMAX	sqrtTOTLEAN	numericTOTPF
logTHIMF			1	.215**	**-.603****	.102**	.287**	.421**
	Sig. (2-tailed)			.000	.000	.000	.000	.000
sqrtTHMUS				1	.044	**.791****	**.922****	-.484**
	Sig. (2-tailed)				.064	.000	.000	.000
sqrtTHMUSD					1	.155**	.001	-.488**
	Sig. (2-tailed)					.000	**.**981	.000
sqrtKCTMAX						1	**.773****	-.493**
	Sig. (2-tailed)						.000	.000
sqrtTOTLEAN							1	-.479**
	Sig. (2-tailed)							.000
numericTOTPF								1
	Sig. (2-tailed)	.						

**. Correlation is significant at the 0.01 level (2-tailed).

a. RACE = European

Table 9: Phenotypic correlations (r) among muscle- related traits (transformed and adjusted for significant covariates) in African Americans

			logTHIMF	sqrtTHMUS	sqrtTHMUSD	sqrtKCTMAX	sqrtTOTLEAN	numericTOTPF
logTHIMF			1	.138**	**-.610****	-.048	.235**	.522**
	Sig. (2-tailed)			.000	.000	.121	.000	.000
sqrtTHMUS				1	.139**	**.706****	**.876****	-.375**
	Sig. (2-tailed)				.000	.000	.000	.000
sqrtTHMUSD					1	.280**	.059*	-.523**
	Sig. (2-tailed)					.000	.038	.000
sqrtKCTMAX						1	**.634****	-.435**
	Sig. (2-tailed)						.000	.000
sqrtTOTLEAN							1	-.355**
	Sig. (2-tailed)							.000
numericTOTPF								1
	Sig. (2-tailed)	.						

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

a. RACE = African

3.2 SINGLE SNP-SINGLE TRAIT ASSOCIATION RESULTS

The single SNP-single trait association analyses were performed using PLINK in the European American and African American cohort separately. Although a total of 1735 SNPs from 27 genes were initially available, 221 SNPs were excluded in African Americans and 109 SNPs were in European Americans due to low minor allele frequencies (MAF<0.05). A summary of the most significant results are given in Table 8 and 9. As described above, based on the number of tests, a significant result required $p \leq 5\times10^{-5}$. In Figures 5-16, the $-\log_{10}$p-value for each SNP in each gene is plotted for each trait. The SNPs are plotted by chromosome and base-pair position, and the horizontal line marks p= 0.01. As can be seen from Figures 3.5-3.16, none of the SNPs achieved experiment-wide significance with any trait in either European or African Americans. The most significant association (P <0.00015) was between SNP (rs805554) in *RALGABP* and TOTPF in African Americans (Figure 15 and Table 10).

Although no SNPs achieved experiment wide significance, several genes had multiple SNPs that were associated with a muscle-related trait at p<0.01. To investigate these results further, I tallied the number of SNPs within each gene had p-values 10^{-2} for each trait in European Americans and African Americans and also listed the most significant p- value for each gene(Table 10 and 11). Although several genes contained SNPs that were associated (at p< 0.01) with several traits, there was no overlap between European American and African American results. That is, the same SNPs did not affect the same traits in the two groups. In the next few sections, I will discuss results in each ancestry group separately.

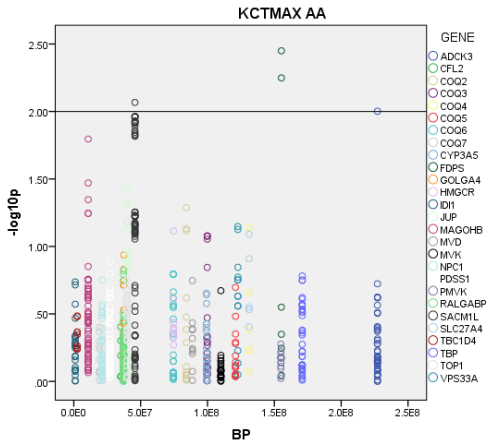

Figure 5: Plot of –(log$_{10}$p-value) and location (in bp) for KCTMAX

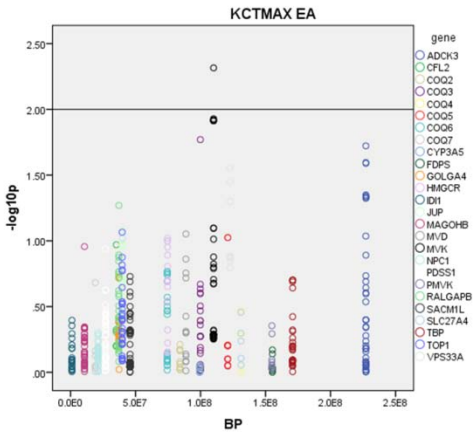

Figure 6:Plot of –(log$_{10}$p-value) and location (in bp) for KCTMAX

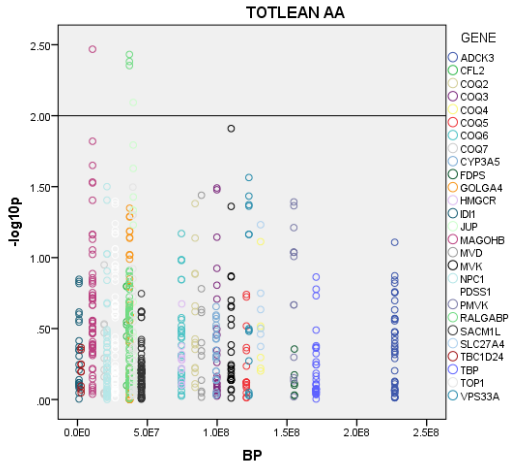

Figure 7: Plot of –(log$_{10}$p-value) and location (in bp) for TOTLEAN in AA

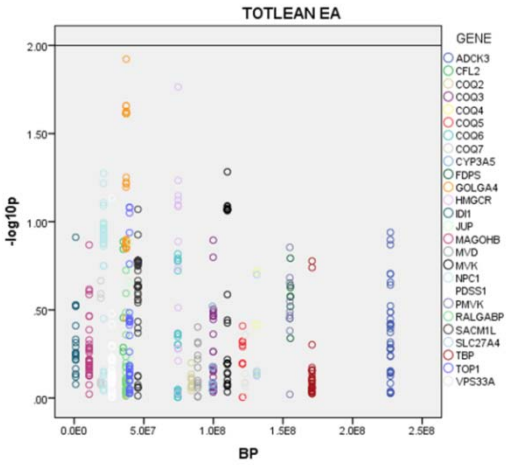

Figure 8: Plot of –(log$_{10}$p-value) and location (in bp) for TOTLEAN in EA

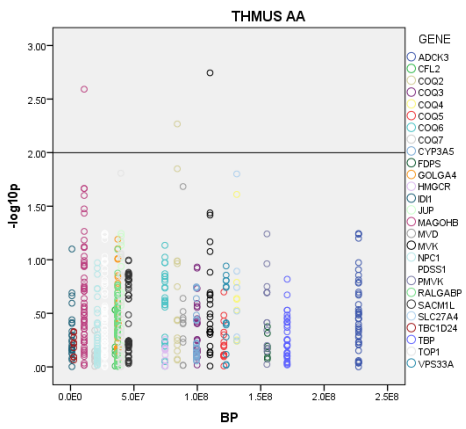

Figure 9: Plot of –(log $_{10}$p-value) and location (in bp) for THMUS in AA

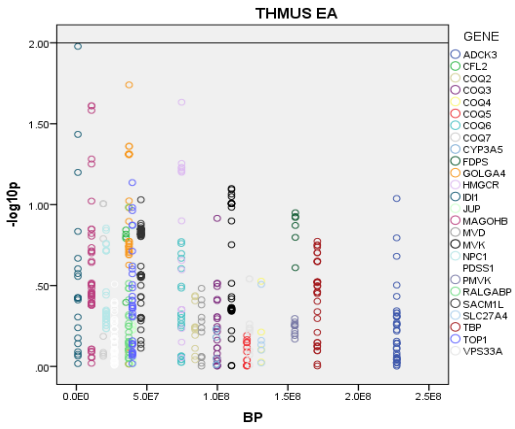

Figure 10: Plot of –(log $_{10}$p-value) and location (in bp) for THMUS in EA

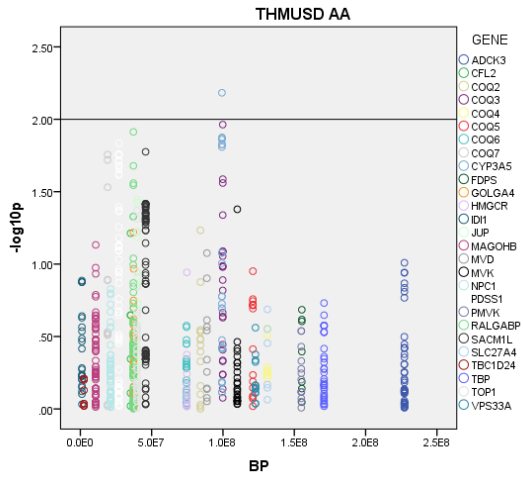

Figure 11: Plot of –(log $_{10}$p-value) and location (in bp) for THMUSD in AA

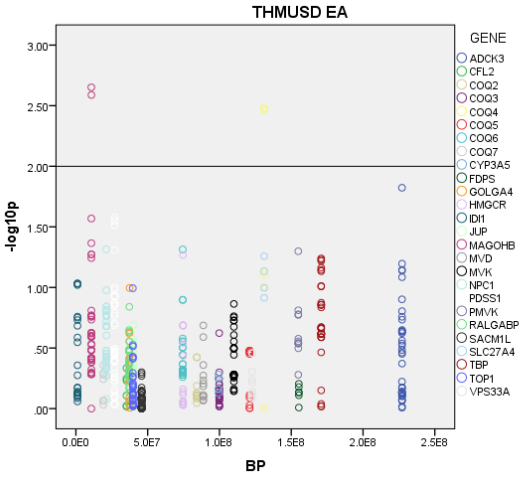

Figure 12: Plot of –(log $_{10}$p-value) and location (in bp) for THMUSD in EA

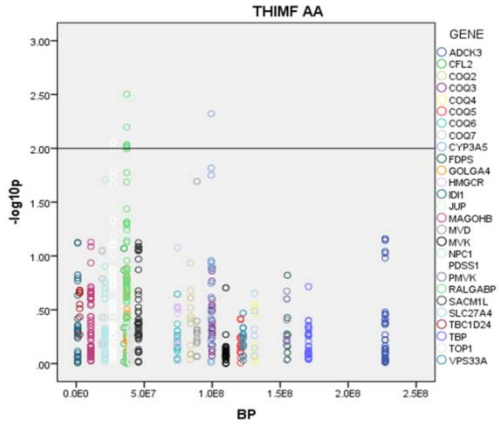

Figure 13: Plot of –(log $_{10}$p-value) and location (in bp) for THIMF in AA

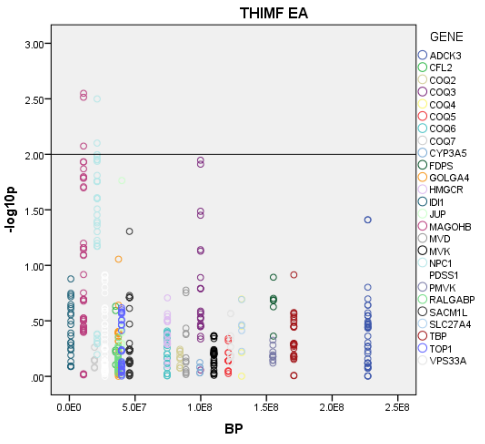

Figure 14: Plot of –(log $_{10}$p-value) and location (in bp) for THIMF in EA

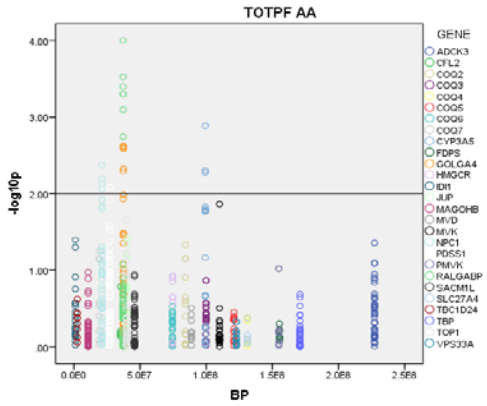

Figure 15: Plot of –(log$_{10}$p-value) and location (in bp) for TOTPF in AA

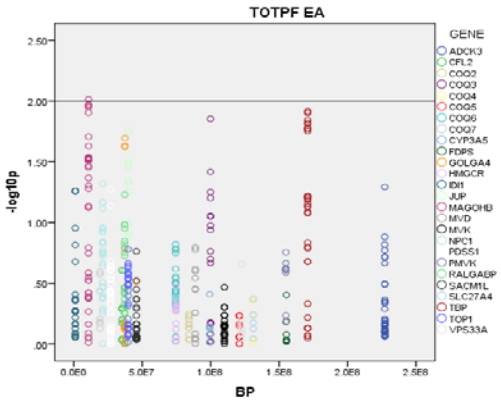

Figure 16:Plot of –(log$_{10}$p-value) and location (in bp) for TOTPF in EA

3.2.1 Association Results in European Americans

In European Americans, there was minimal clustering of associations within genes or across traits (Table 10). However, three SNPs (#rs11053877, #rs10845181, and #rs1026462) within MAGOHB locus on chromosome 12 are associated with THIMF (p= 3.062×10^{-3}, 8.433×10^{-3}, and 2.282×10^{-3}, respectively). In addition, 2 of these SNPs (#rs1026442 and #11053877, p= 2.237×10^{-3} and 2.582×10^{-3}, respectively) were associated with variation in thigh muscle density (THMUSD) and 1 SNP (rs12811247) was associated with total percent fat (TOTPF). These three traits are moderately correlated with each other (see table 7 and table 8: Phenotypic Correlations). These results may indicate that variation in *MAGOHB* has pleiotropic effects on multiple traits. There was no apparent clustering among the remaining results with p-values <0.01.

Table 10: Genes with SNP associations $p < 10^{-2}$ by trait in European Americans

	TRAITS									
GENE	KCTMAX	Total SNPs (#sig)	THIMF	Total SNPs (#sig)	THMUS	THMUSD	Total SNPs (#sig)	TOTLEAN	TOTPF	Total SNPs (#sig)
	-		2.8×10^{-3}	46(3)	-	2.2×10^{-3}	46(2)	-	9.7×10^{-3}	46(1)
MAGOHB										
MVK	4.8×10^{-3}	32(1)	-		-	-		-	-	
NPC1	-		3.1×10^{-3}	40(2)	-	-		-	-	
SLC27A4	-		-		-	3.3×10^{-3}	9(2)	-	-	

3.2.2 Association Results in African Americans

In African Americans, there were several clusters of p-values within genes and across traits (Table 11). As mentioned above, a SNP (rs805554) located in *RALGAPB (KIAA1219)* is significantly associated with total percent fat (TOTPF) in African Americans (p value= 1.517×10^{-4}). *RALGAPB* (is on chromosome 20 and codes for the Ral GTPase activating protein, beta subunit (non-catalytic).Five additional SNPs (rs1115600, rs1295992, rs12625459, rs805559, and rs805552) in *RALGAPB* are also associated with TOTPF (p= 8.404×10^{-3}, 5.133×10^{-3}, 3.33×10^{-3}, 5.305×10^{-3}, and 4.658×10^{-3} , respectively). In addition, eight SNPs (rs2185594, rs805554, rs130577, rs1315305, rs1303571, rs11699660, rs41812025, and rs2206783 in the *RALGAPB* locus are associated with thigh intermuscular fat, of which rs805554 is the same as the associated with total percent fat. Finally, two SNPs (rs9941745 and rs9941761) in the *RALGAPB* locus are associated with total lean mass.

Table 11: Genes with SNP associations $p < 10^{-2}$ by trait in African Americans

GENE	TRAITS KCTMAX	Total SNPs (#sig)	THIMF	Total SNPs (#sig)	THMUS	Total SNPs (#sig)	THMUSD	Total SNPs (#sig)	TOTLEAN	Total SNPs (#sig)	TOTPF	Total SNPs (#sig)
ADCK3	9.9×10^{-3}		-		-		-				-	
COQ2	-		-		5.4×10^{-3}	15(1)	-				-	
CYP3A5	-		4.7×10^{-3}	25(1)	-		6.5×10^{-3}	24(1)			1.3×10^{-3}	27(3)
FDPS	3.5×10^{-3}	7(2)	-		-		-				-	
GOLGA4	-		-		-		-				2.4×10^{-3}	68(6)
JUP	-		-		-		-		8.0×10^{-3}	23(1)	-	
MAGOHB	-		-		2.5×10^{-3}	62(1)	-		3.4×10^{-3}	59(1)	-	
MVK	-		-		1.8×10^{-3}	26(1)	-				-	
NPC1	-		-		-		-				4.2×10^{-3}	59(4)
PDSS1	-		8.7×10^{-3}	63(3)	-		-				8.2×10^{-3}	70(2)
RALGABP	-		3.1×10^{-3}	81(6)	-		-		4.1×10^{-3}	91(2)	1.5×10^{-4}	84(6)
SAC1ML	8.5×10^{-3}	75(1)	-		-		-				-	

Variation in total percent fat (TOTPF) in African Americans was also associated with SNPs in other genes including: (1) four SNPs (rs1652344, rs1631685, rs1652348, and rs1788781) within the NPC1 locus on chromosome 18 (p=8.776 $\times10^{-3}$, 7.433 $\times10^{-3}$, 6.498 $\times10^{-3}$, and 4.256 $\times10^{-3}$, respectively); (2) six SNPs (rs6800842 , rs11129756, rs6550470**,** rs7639447, rs4635655, and rs17266090) in the GOLGA4 locus on chromosome 3 (p= 5.149×10^{-3}, 2.606×10^{-3}, 2.438×10^{-3}, 2.528×10^{-3}, 2.488×10^{-3}, and 4.71×10^{-3}, respectively); and (3) three SNPS (#rs10242455**, #** rs4646457, and # rs776746) in the CYP3A5 locus on chromosome 7 (p= 1.325×10^{-3}, 5.09×10^{-3}, and 5.395×10^{-3}, respectively). Of particular interest, genotypes at SNP rs10242455 in CYP3A5 were also associated with both thigh intermuscular fat (p= 4.768×10^{-3}) and thigh muscle density (p=6.564×10^{-3}). Thus, this latter SNP may have pleiotropic effects on multiple traits.

SNPs in additional loci were associated (P< 0.01) with other muscle-related traits, but there were no apparent clusters among the remaining results.

3.3 LINKAGE DISEQUILIBRIUM AND MEAN EFFECTS OF SNPS IN GENES OF INTEREST

As described previously, single SNP-single trait analysis revealed several loci associated with specific phenotypes within each ancestry cohort, although none of these results achieved experiment-wide significance. However, to further assess the strength of some of the associations, I plotted regional significant association results from single gene-single trait association results using the LocusZoom 1.1 tool (Pruim RJ, 2010). LocusZoom provides visual display of regional strength association information based on genomic position, local linkage disequilibrium (LD), recombination hotspots and the gene positions in the significant regions. Estimates of recombination and LD are obtained from HapMap. According to Loos, SNPs within a gene that are several hundred kb may be still correlated to each other (Loos et al., 2008). Thus, multiple significant SNP associations may be due to one or a few 'true' signals. In addition, for specific SNP-trait pairs, genotypic means were calculated and phenotype by genotype box-plots were created to assess the effect of the variant on the specific muscle-related trait.

3.3.1 *MAGOHB* in European Americans

In European Americans, multiple SNPs within MAGOHB were associated with thigh intermuscular fat, thigh muscle density, and total percent fat. Figure 17 is a plot of the physical position of all tested SNPs within this chromosomal region by the $-\log_{10}$p values for association with thigh intermuscular fat (THIMF). Recombination hotspots (blue solid line) are at 10.63 Mb and 10.68 Mb. The purple diamond represents the SNP with the lowest *p* value (rs1026442). The colors of the other SNPs represent their correlation with the most significant SNP. Yellow circle shows the $0.6 < r^2 < 0.8$ and green circle shows

the $0.4 < r^2 < 0.6$. Blue circle shows the $0.2 < r^2 < 0.4$. As can be seen, the two SNPs with the most significant p-values, rs1026442 (purple diamond) and rs11053877 (red circle) are in high LD with each other. Thus, although these SNPs are 10.37kb apart, the significant associations with THIMF and THMUSD likely represent the same "true" signal. In addition, other SNPs within MAGOHB and KLRA1 show no evidence for association and are not correlated with the two most significant SNPs.

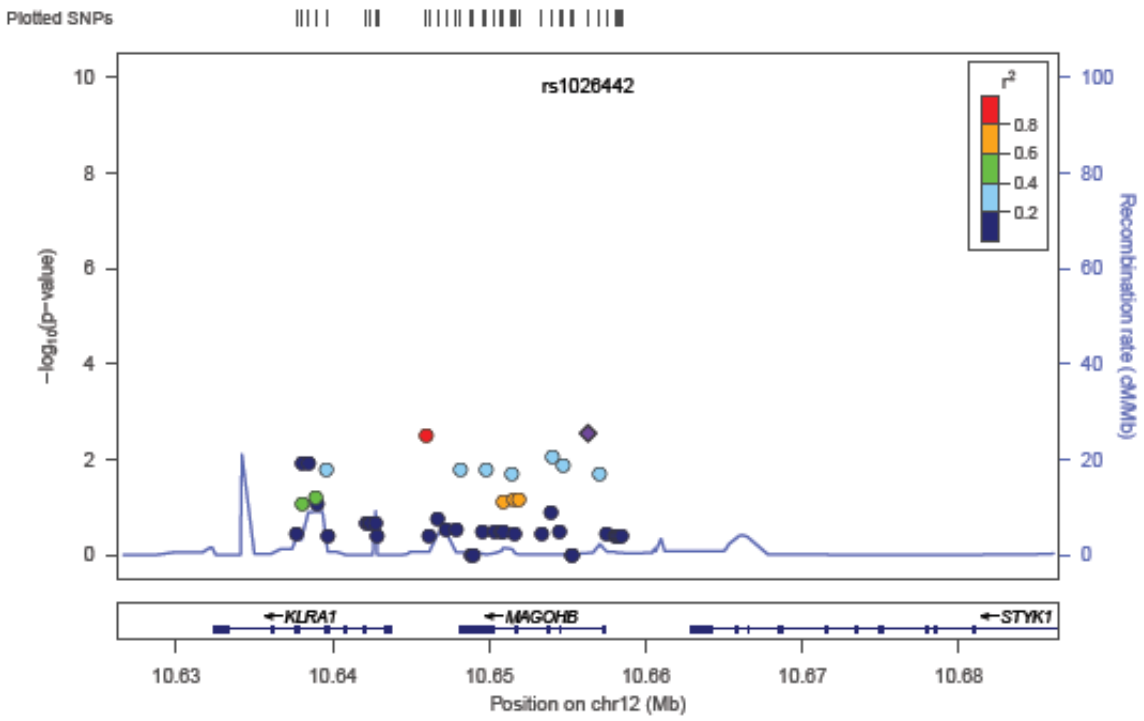

Figure 17: Association between THIMF and SNPs around *MAGOHB* (rs1026442) on 12p13.2 in European Americans

I next calculated the means of thigh muscle density (THMUSD) and intermuscular fat (THIMF) by each genotype for two SNPs in *MAGOHB*: rs11053877 and rs1026442 (Table 12). In the tables below, "22" represents the homozygote of the minor allele, "12" is the heterozygote, and "11" is the

homozygote of the major allele for each locus. The minor allele frequencies of both loci are similar and as can be seen in Table 12, the genotypic means of these two loci are similar for each trait. For example, for rs1026442, the mean of THIMF (using the residuals of the transformed data) for the homozygote (22) for the minor allele is higher than that of the other homozygote (11), whereas the mean of the heterozygote is closer to that of the minor allele homozygote mean. A similar effect is seen for rs11053877, that is, the minor allele genotype at each locus seems to have a recessive effect on THIMF. In contrast, for THMUSD, the mean of the homozygote (22) of the minor allele for rs1026442 is lower than the means of the other homozygote and the heterozygote man is intermediate. The results for rs11053877 are similar. Thus, for THMUSD, there appears to be more of a dosage effect of the minor allele. The genotypic effects of SNPs (rs1026442 and rs11053877) on the two traits are also illustrated using box and whisker plots (Figure 18, Figure 19, Figure 20, and Figure 21).

Table 12: Means (±s.e.) of Muscle-related traits (transformed and adjusted for covariates) by SNP genotypes in *MAGOHB* in European Americans

		SNP	
	Genotype	rs1026442	rs11053877
Alleles		A/G	A/G
MAF		0.1926	0.1939
THIMF	11	-0.03±0.01	-0.03±0.01
	12	0.05±0.02	0.06±0.02
	22	0.04±0.06	0.04±0.06
THMUSD	11	0.03±0.01	0.03±0.01
	12	-0.05±0.02	-0.05±0.02
	22	-0.08±0.07	-0.09±0.07

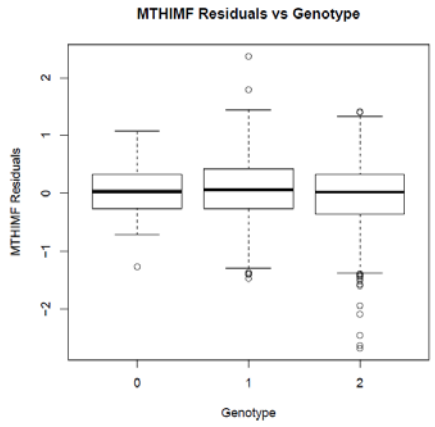

Figure 18: The genotype-THIMF plot of rs1026442

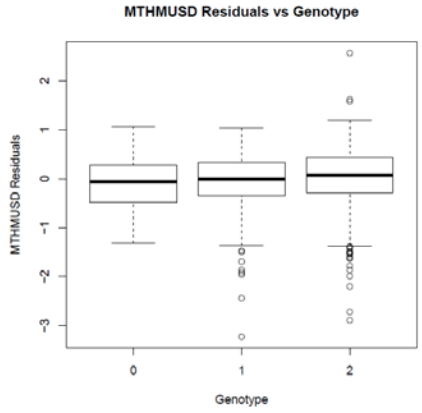

Figure 19: The genotype-THMUSD plot of rs1026442

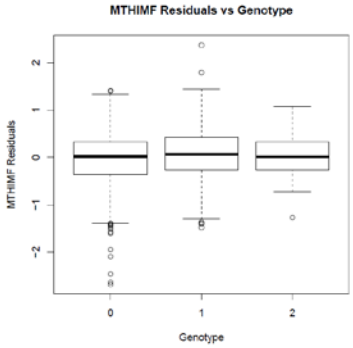

Figure 20: The genotype-THIMF plot of rs11053877

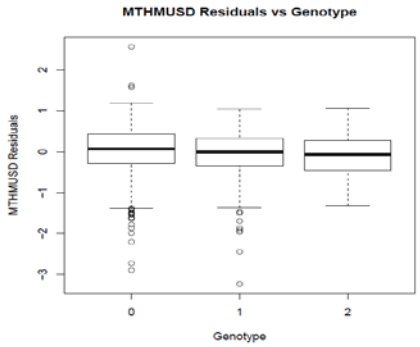

Figure 21: The genotype-THMUSD plot of rs11053877

3.3.2 *RALGAPB* in African Americans

In African Americans, multiple SNPs within *RALGPB* were associated with thigh intermuscular fat, total lean, and total percent fat. Figure 22 is a plot of the physical position of all tested SNPs within this chromosomal region by the $-\log_{10}$p values for association with total percent fat (TOTPF) using SNP rs805554 and rs805552 as the index SNPs. A similar result was obtained for association with thigh intermuscular fat (THIMF), so I just present the results for total percent fat (TOTPF). As can be seen, multiple SNPs across the *RALGAPB* locus are associated with the TOTPF. However, when rs805554 is used as the index SNP, several SNPs are not in high LD with this SNP which could indicate that variation

in more than one region of *RALGAPB* may influence the muscle-related traits. Although the "true" causal variant or region is not clear, when rs805552 is used as the index SNP, 4 SNPs across the gene are in high LD with the index SNP (These two SNPs are 4.234 kb apart from each other). These results may indicate that variation in possibly two regions within *RALGAPB* influence muscle-related traits. Figure 23 is a plot of the physical position of all tested SNPs within this chromosomal region by the $-\log_{10}$p values for association with total lean (TOTLEAN) using SNP rs9941761 as the index SNP. When rs9941761 (which is near the 3' end of *RALGABP*) is used as the index SNP, two SNPs across the gene are in high LD with the index SNP and one of these SNPs is near the middle of *RALGABP*. From all of these results it is not clear whether there is more than one 'true' signal for RALGABP, nor is it clear where the signal might be located.

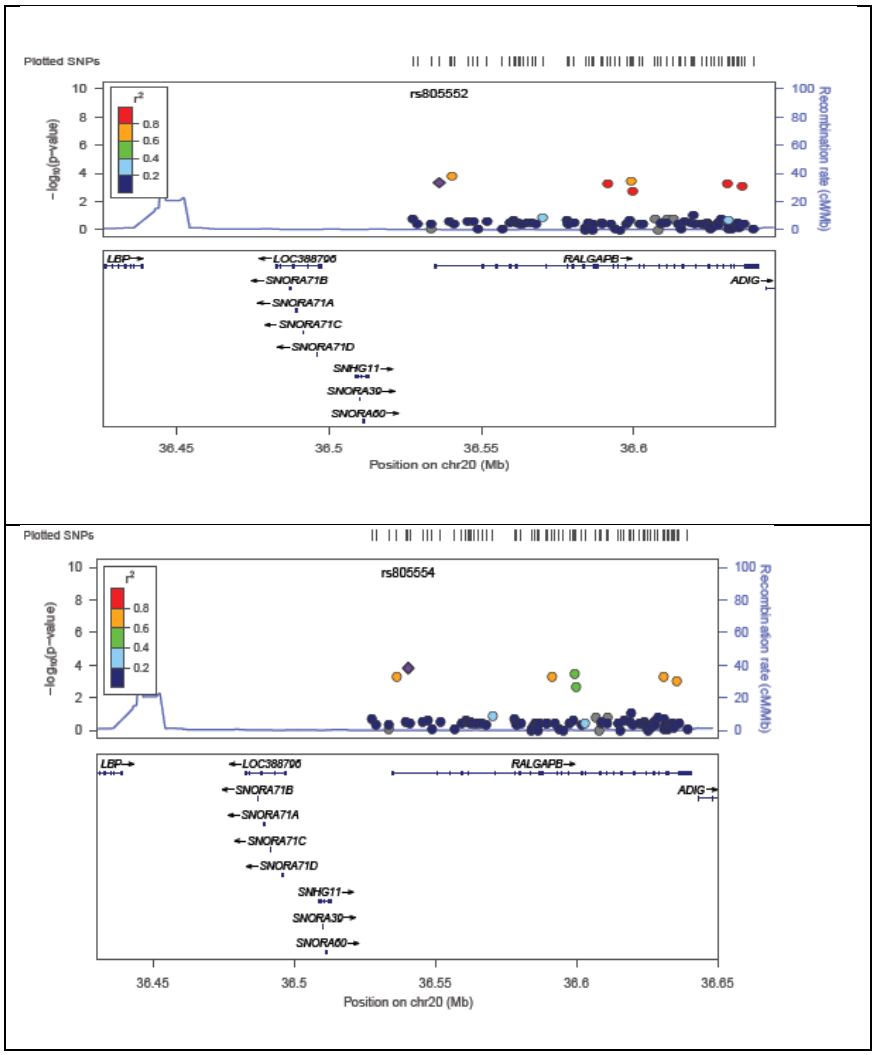

Figure 22: Association between TOTPF and SNPs near RALGBP in African Americans using SNPs rs805554 and rs805552 as the index SNP.

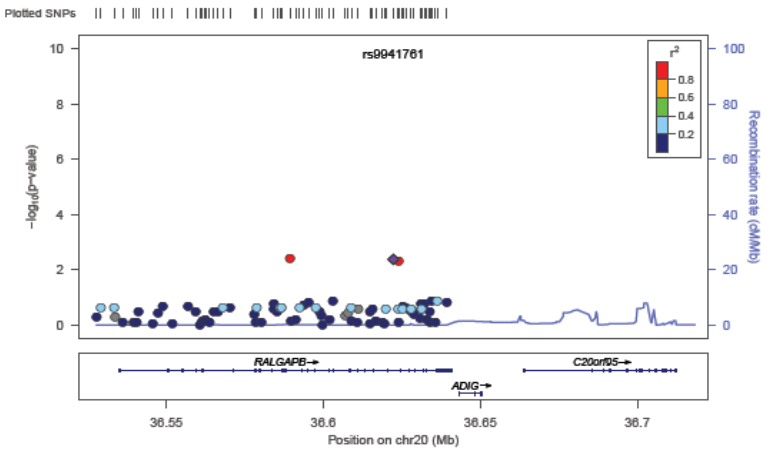

Figure 23: Association between TOTLEAN and SNPs near RALGBP in African Americans using SNPrs9941761 as the index SNP.

I next calculated the means of thigh intermuscular fat (THIMF), and total percent fat (TOTPF) by each genotype for rs805554 SNP in *RALGAPB*. (Table 13). As can be seen, the means of TOTPF and THIMF for the homozygote for the minor allele (22) are higher than the other homozygote (11), whereas the mean of the heterozygote is in between. Thus, there appears to be a dosage effect of the minor allele on each trait. The genotypic effects of SNP rs805554 on the TOTPF and THIMF are also illustrated using box and whisker plots (Figure 24 and Figure 25).

Table 13: Means (±s.e.) of Muscle-related traits (transformed and adjusted for covariates) by SNP genotypes in *RALGAPB* in African Americans

		SNP
	Genotype	rs805554
Alleles		A/T
MAF		0.1307
TOTPF	11	-27.22±13.86
	12	82.30±26.19
	22	120.62±87.12
THIMF	11	-0.02±0.02
	12	0.07±0.03
	22	0.19±0.10

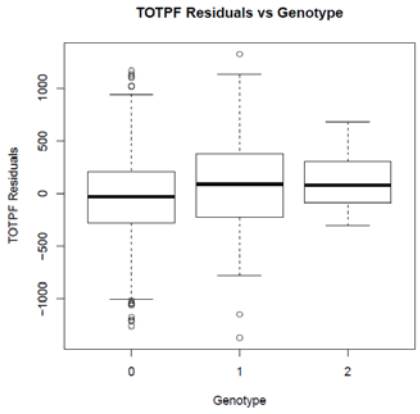

Figure 24: The genotype-TOTPF box-plot(rs 805554)

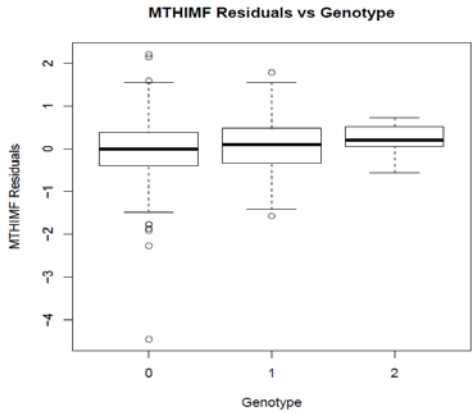

Figure 25: The genotype-THIMF box plot(rs805554)

3.3.3 *CYP3A5* in African Americans

In African Americans, multiple SNPs *within CYP3A5*were associated with thigh intermuscular fat (THIMF), thigh total muscle density (THMUSD), and total percent fat (TOTPF). Figure 26 is a plot of the physical position of all tested SNPs within this chromosomal region by the $-\log_{10}$p values for association with total percent fat (TOTPF) using SNP rs10242455 as the index SNP. When rs10242455 (which is in the 3'

UTR) is used as the index SNP, 2 SNPs across the gene are in high LD with the index SNP, one of which is near the 5' end. These results may indicate that variation in possibly this region within *CYP3A5* influences muscle-related traits, although the location of a 'true' causal variant is unclear.

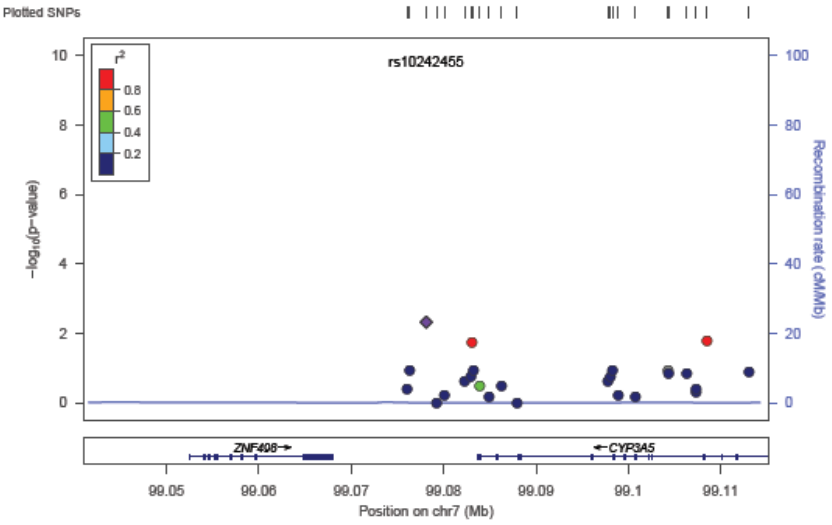

Figure 26: Association between THIMF and SNPs near *CYP3A5* in African Americans using SNP rs10242455 as the index SNP.

I also calculated the means of thigh THIMF, THMUSD and total percent fat by each genotype for three SNPs in *CYP3A5:* rs10242455 (Table 14). As can be seen, the means of THIMF and TOTPF for the homozygote (22) for the minor allele is lower than those of the other homozygote, whereas the means of the heterozygote are in between. On the other hand, the mean of THMUSD for the homozygote (22) for the minor allele is higher than those other homozygote (11), whereas the mean of heterozygote is in between. Again, the minor allele has a dosage effect of the traits. The genotypic effects of rs10242455 SNP on these traits is consistent with the phenotypic correlation among these traits (see Table 5). That is, the minor allele is associated with a decrease in TOTPF and THIMF, but is associated with an increase in THMUSD. The genotypic effects of SNPs (rs10242455) on the TOTPF, THIMF, and THMUSD are also illustrated using box and whisker plots (Figure 27, Figure 28, and Figure 29).

Table 14: Means (±s.e.) of Muscle-related traits by SNP genotypes in *CYP3A5* in African Americans

		SNP
	Genotype	rs10242455
Alleles		A/G
MAF		0.3143
THIMF	11	0.04±0.02
	12	-0.03±0.02
	22	-0.09±0.04
THMUSD	11	-0.04±0.02
	12	0.02±0.02
	22	0.10±0.04
TOTPF	11	36.33±18.07
	12	-20.30±18.82
	22	-83.22±33.59

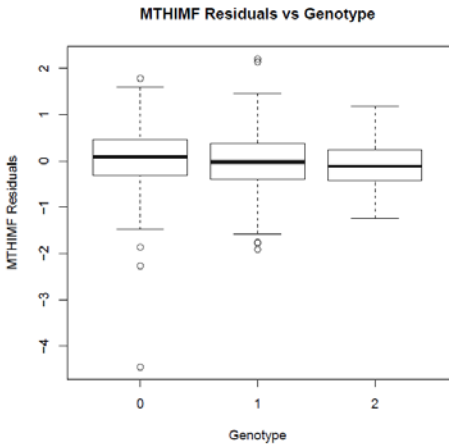

Figure 27: Genotype-THIMF box plot(rs10242455)

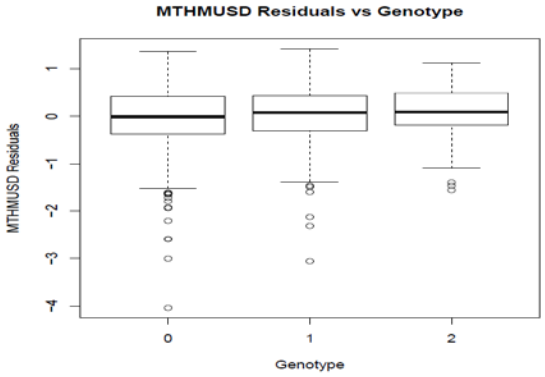

Figure 28: Genotype-THMUSD box plot(rs10242455)

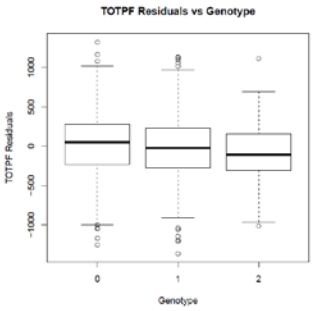

Figure 29: Genotype-TOTPF box-plot(rs10242455)

3.3.4 *GOLGA4* in African Americans

In African Americans, multiple SNPs within *GOLGA4* were associated with thigh total percent fat (TOTPF). Figure 30 is a plot of the physical position of all tested SNPs within this chromosomal region by the $-\log_{10}$p values for association with total percent fat (TOTPF) using SNP rs6550470 as the index SNP. When rs6550470 is used as the index SNP, 2 SNPs across the gene are in high LD with the index SNP, but others are not. These results may indicate that variation in possible multiple regions within GOLGA4 influences muscle-related traits.

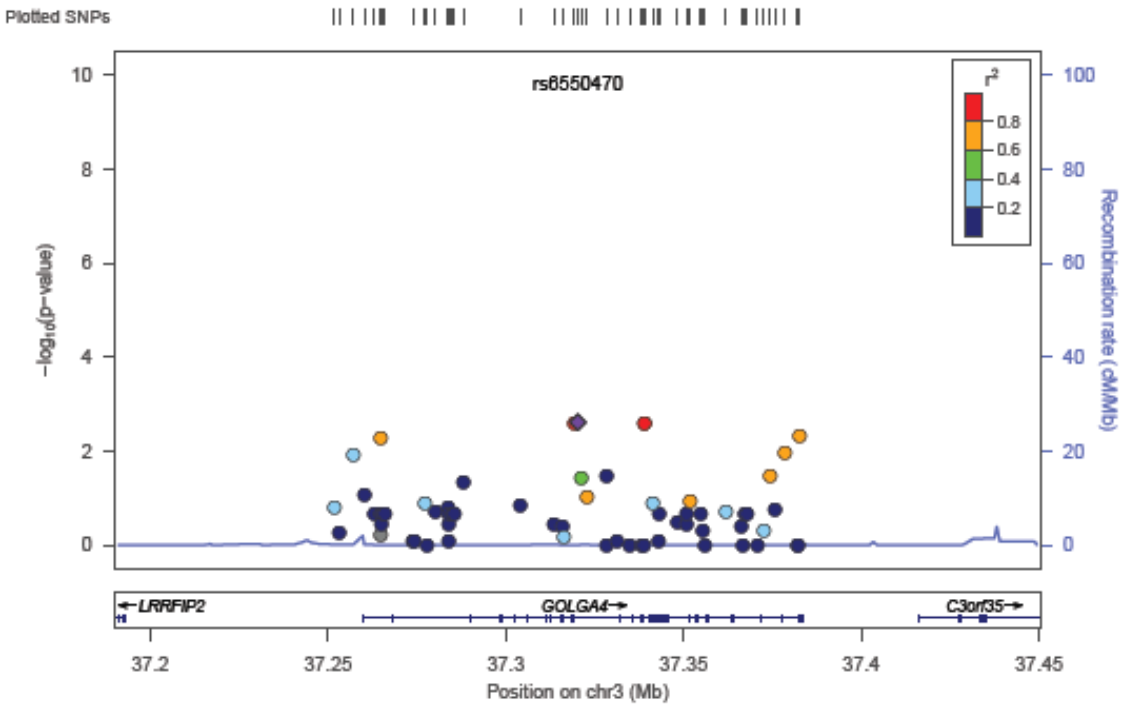

Figure 30: Association between TOTPF and SNPs near *GOLGA4* in African Americans using SNP rs6550470 as the index SNP.

Lastly, I calculated the mean of TOTPF by each genotype for the SNPs rs6550470 which is the most significant SNP with the p value 2.438×10^{-3} in *GOLGA4*. As can be seen, the mean of TOTPF for the

homozygote(22) for the minor allele is higher than those of the other homozygote(11), whereas the mean of the heterozygote is in between, showing that increasing dosage of the minor allele is associated with increasing TOTPF. (Table 15).

Table 15: Means (±s.e.) of TOTPF (transformed and adjusted for covariates) by SNP genotypes in *GOLGA4*in African Americans

		SNP
	Genotype	Rs6550470
Alleles		A/G
MAF		0.446
TOTPF	11	-41.11±30.17
	12	-2.02±16.23
	22	72.01±23.26

4.0 DISCUSSION

Advancing age results in major changes in body composition, including decreases in skeletal and muscle mass. Sarcopenia, the loss of muscle mass and strength that occurs with aging, plays a crucial role in frailty and results in increased morbidity and mortality among older individuals. The loss of muscle mass and function varies greatly among older individuals, however the underlying causes and mechanisms of this variation are unclear. Because sarcopenia and measures of muscle mass and strength are heritable, many genetic studies, including large genome-wide association consortium studies, have been done to identify genes that influence traits related to sarcopenia, such as total lean mass and obesity (e.g., BMI). These studies have identified several genes that influence lean mass and BMI, but these genes do not account for all of the known heritability of the trait. In addition, few studies have been performed in older individuals or using traits more strongly associated with sarcopenia, such as muscle function and muscle density. In order to better understand the genetic basis of sarcopenia, I performed the single SNP-single trait association analysis using genotype data on 778 SNPs in 27 candidate genes and phenotype data on 6 measures of muscle function and density (KCTMAX, TOTLEAN, TOTPF, THIMF, THMUS and THMUSD) from 3075 men and women aged 70-79 years in the HealthABC cohort. The muscle density measures of THIMF, THMUS, and THMUSD were calculated as the average of the measurements from left and right thigh. Analyses were performed separately in the European American (n=1794) and African American (n=1281) cohorts separately.

Overall, none of the associations in either ancestral group achieved the experiment-wide significance level ($p<10^{-5}$). The most significant result was an association In African Americans between total percent fat and SNPs in the *RALGAPB* locus (minimum p-value = 1.5×10^{-4}). I further investigated all

tests that achieved a more liberal significance level (p-value < 10^{-2}). Interestingly, even at this more liberal significance level, there was minimal overlap in results from analyses in the European Americans versus African American cohorts; only three genes showed an association <10^{-2} in both cohorts (*MAGOHB*, *MVK*, and *NPC1*) but the associations were between different SNPs and different traits (Tables 10 and 11). This result indicates that at least some of the variation in these muscle-related traits is influenced by different genes between the two groups. Such an observation is not surprising because of known phenotypic and genetic differences in body composition traits between individuals of predominantly African versus European ancestry (Shaffer, 2007; Goodpaster, 2002).

Although none of the results were statistically significant at the experiment-wise level, several associations were potentially interesting and I discuss two of these in more detail below.

4.1 ASSOCIATION WITH *RALGAPB*

One of most promising results was the association in African Americans between multiple SNPs within *RALGAPB* locus in African Americans and variation in total percent fat, thigh intermuscular fat, and total lean mass. The genomic location of *RALGAPB* in humans is on 20q11.23. The LD between two SNPs, rs805554 and rs805552, with the most significant association results is high and equals 0.797. These SNPs are also in moderate to high LD (0.6 < r^2 < 0.9) with other associated SNPs in *RALGAPB* (Figure 22). Because the LD among all of the SNPs is not uniformly high (e.g, LD>0.80), it is not clear whether one or more variants potentially influence variation in total percent fat, thigh intermuscular fat, or total lean mass. Furthermore, it is not clear where in *RALGAPB* the potentially 'causal' variants are located because associated SNPs are located across the gene. Both SNP rs805554 (T>A) and SNP rs805552 (C>G) are located in the intron 1 region of *RALGAPB*, and are unlikely to be causal. Variants in introns can have causal effects because introns have many short sequences which are significant in effective splicing. Acceptor and donor sites at both ends of the intron are important in proper splicing. It was found that

introns affect the expression level of host gene (Chorev 2012). Introns also enhanced the gene expression by intron-mediated enhancement (Jonsson,1992). However, we do not know whether these two SNPs are causal or whether they are in high LD with another, unassayed but causal variant.

The two most significant SNPs in *RALGAPB* were associated with variation in total percent fat and thigh intermuscular fat. Total percent fat and total intermuscular fat have a phenotypic correlation (r) equal to 0.47 (r^2 = 0.22). Total lean mass has a positive, but low correlation with thigh intermuscular fat (r = 0.28), and a negative, moderate correlation with total percent fat (r = 0.42). The mean effects of genotypes at the rs805554 locus on THIMF and TOTPF are consistent with the phenotypic correlation among the traits. Thus, the association results may indicate that polymorphisms within *RALGAPB* have pleiotropic effects on the two muscle-related traits.

RALGAPB, the homolog of *KIAA1219*, is a Ral GTPase activating protein, beta subunit (non-catalytic) and is a non-catalytic subunit of the heterodimeric RalGAP1 and RalGAP2 complexes that act as GTPase activators for the Ras-like small GTPases RALA and RALB. *RALGAPB* gene is the homolog of a putative gene D.2085.5 that is located on chromosome II of *C.elegans*. How variation in RALGAPB may influence muscle-related traits (and sarcopenia) is unclear. However, a RalGAP complex is involved in insulin–stimulated glucose trafficking in muscle and adipose tissue (Chen et al., 2011). As stated earlier, insuin-signalling pathways are associated with longevity, thus the involvement of Ral GAP and other members of this pathway in insulin signaling makes it a reasonable candidate for future study. Furthermore, another member of this pathway, the Ras-like small G protein, RalB, may be involved in the activation of the autophagosome in the presence of nutrient deprivation (Bodemann, et al. 2011). Thus, further study of this pathway may provide novel insights into the development of sarcopenia.

4.2 ASSOCIATION WITH *MAGOHB*

Another potentially interesting result is the association with SNPs within *MAGOHB* locus and thigh intermuscular fat, thigh muscle density, and total percent fat in European Americans. The chromosomal location of *MAGOHB* in humans is 12p13.2. The two SNPs, rs1026442 and rs11053877, with the strongest association results are in high LD with each other (LD= 0.97), thus the association between SNPs and MAGOHB and the muscle-related traits is likely due to a single, 'causal' variant. However, the location of the potentially causal variant is unclear because rs1026442 is located within the first intron whereas rs11053877 is located in the 3' untranslated region. As stated above, variants in introns may affect protein function and expression. Furthermore, 5' UTR and 3' UTR have roles in mRNA stability, translational efficiency and mRNA localization. 3' UTR influences regulation of transcript stability (Laroia 1999). In addition, SNPs on 3' UTR influence microRNA function (Zihua, 2011). To determine whether these associated SNPs may be Causal or whether they are in high LD with an unassayed variant will require additional analyses.

Both SNPs in *MAGOHB* were associated with variation in two correlated traits: THIMF and THMUSD density .Total percent fat and thigh intermuscular fat have a phenotypic correlation (r) equal to 0.47 (r^2 = 0.22), whereas thigh intermuscular fat is negatively correlated (r=-0.61) with thigh muscle density. The genotypic means of the three traits for SNP rs1026442, rs11053877, and rs12811247 in *MAGOHB* (Table 10) reflect the phenotypic relationship among the traits. Thus, variation in *MAGOHB* likely has pleiotropic effects on the correlated traits. *MAGOHB* (mago-nashi homolog B) is involved in mRNA splicing and in the nonsense-mediated decay (NMD) pathway (Roignant and Treisman, 2010) and has not been previously identified as potential candidate genes for development of sarcopenia, although control of RNA splicing could definitely have effects on muscle-related traits. Recent reports also indicate that *Magoh*, as a component of the exon junction complex, regulates division of neural stem

cells (Silver et a., 2010). Given the connection between peripheral nerve function and muscle strength (Strotmeyer et al, 2009), members of this pathway also are reasonable candidates for further study of the development of sarcopenia.

4.3 CONCLUSION

To identify genes that may influence development of sarcopenia, I performed association analyses between several muscle-related traits and SNPs in 27 candidate genes that were homologs of genes that influence the onset of sarcopenia in *C. elegans.* None of these candidate genes have been associated with lean mass or obesity (measured as BMI) in (mostly) middle-aged adults in studies from large consortia (http://www.genome.gov/26525384). However, none of the consortia studies have identified all of the genes that influence lean mass and BMI. Furthermore, lean mass and BMI in middle-aged adults are only crudely correlated with sarcopenia in older individuals and no large studies of sarcopenia or measures of muscle mass and function have been done, especially in older adults.

Although none of the results of this study were statistically significant at an experiment-wide level, several of the results were suggestive. In particular, variants in *MAGOHB* may influence fat infiltration into muscle among individuals with European ancestry, whereas variants in *RALGABP* may influence fat infiltration among individuals with African ancestry. Neither of these genes, or biological pathways, has been previously identified as playing a role in sarcopenia, but further studies may potentially lead to novel insights regarding development of sarcopenia. Replication in other populations is necessary, as well as additional bioinformatic and mechanistic studies, but these are beyond the scope of this thesis. However, these results indicate the potential usefulness of animal models to generate hypotheses for further study in men and women.

APPENDIX A

PLOTS OF RTHIMF, LTHIMF, RTHMUSD, LTHMUSD, RTHMUS, LTHMUS BY ANCESTRY GROUP

African Americans	European Americans
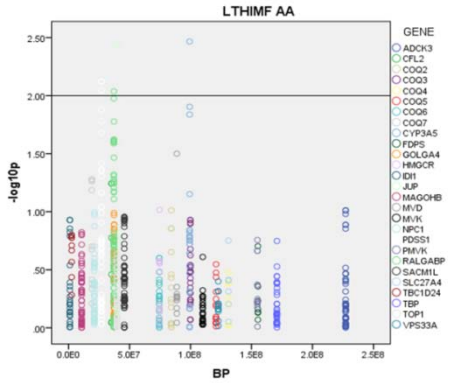	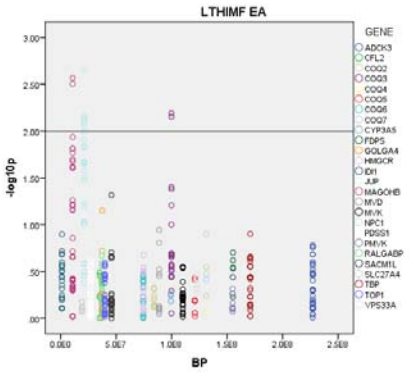
Figure A1:Plot of -log$_{10}$p and bp for LTHIMF in African Americans	Figure A2:Plot of -log$_{10}$p and bp for LTHIMF in European Americans
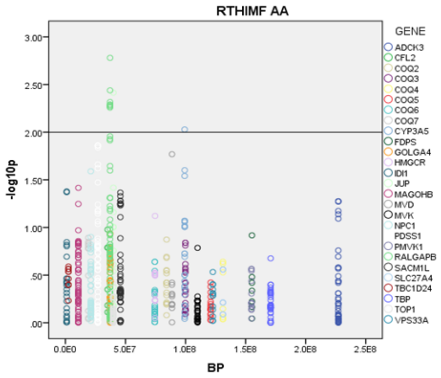	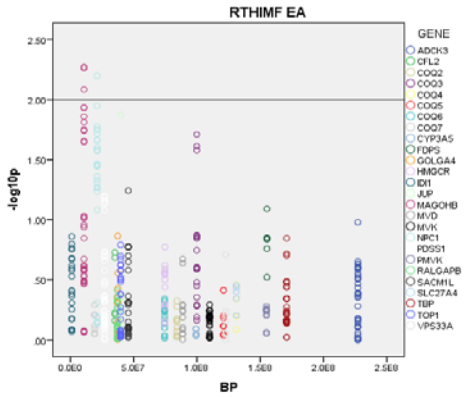
Figure A3: Plot of -log$_{10}$p and bp for RTHIMF in African Americans	Figure A4:Plot of-log$_{10}$p and bp for RTHIMF in European Americans

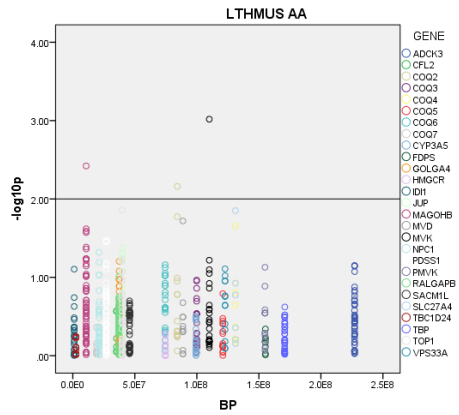

Figure A5: Plot of -log$_{10}$p and bp for LTHMUS in African Americans

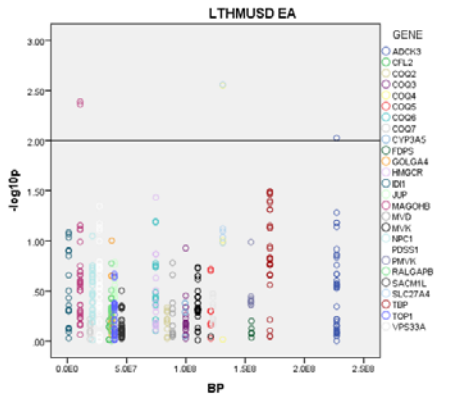

Figure A6 :Plot of –log$_{10}$p and bp for LTHMUSD in European Americans

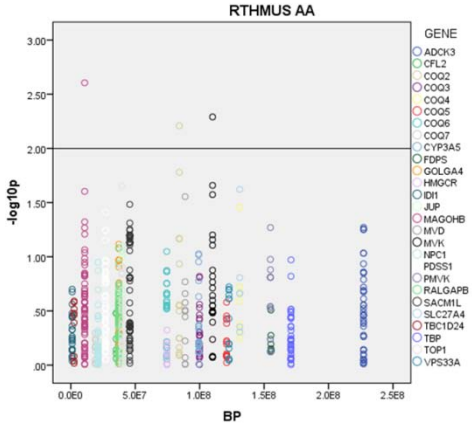

Figure A7:Plot of -log$_{10}$p and bp for RTHMUS in African Americans

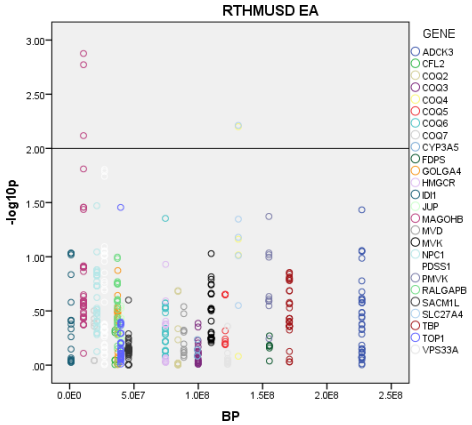

Figure A8:Plot of –log$_{10}$p and bp for RTHMUSD in European Americans

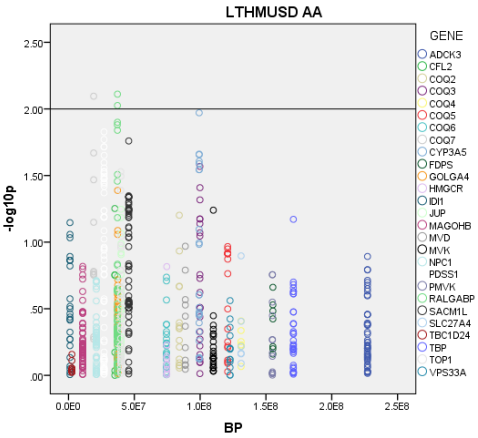

Figure A9: Plot of -log$_{10}$p and bp for LTHMUSD in African Americans

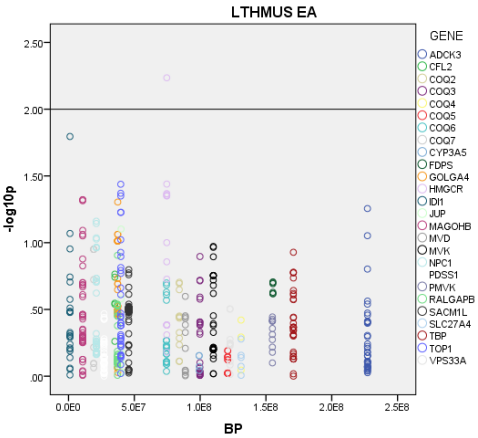

Figure A10:Plot of –log$_{10}$p and bp for LTHMUS in European Americans

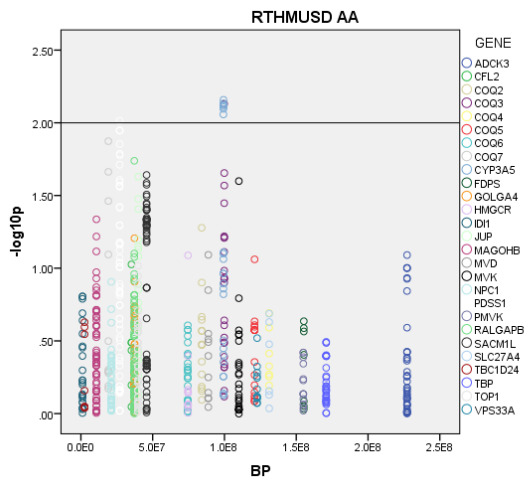

Figure A11:Plot of -$\log_{10}$p and bp for RTHMUSD in African Americans

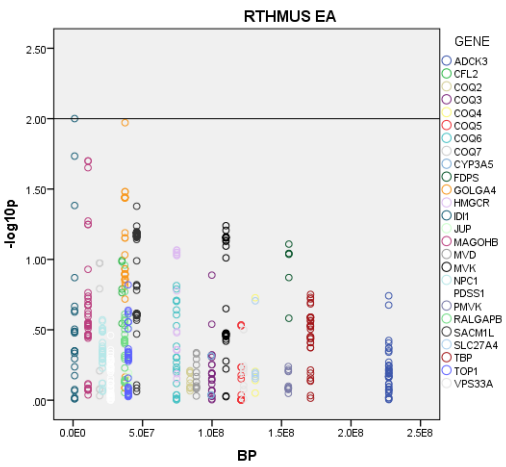

Figure A12:Plot of –$\log_{10}$p and bp for RTHMUS in European Americans

APPENDIX B

HISTOGRAMS OF TRAITS

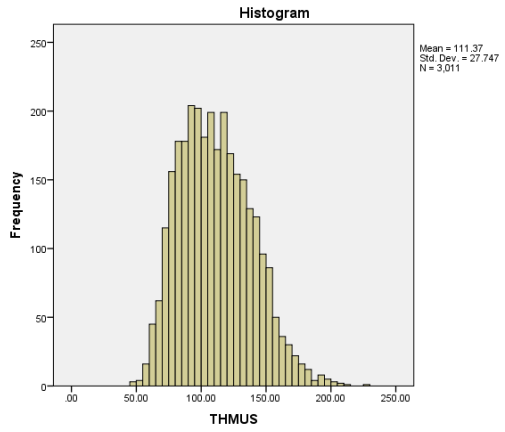

Figure B1: Histogram of THMUS

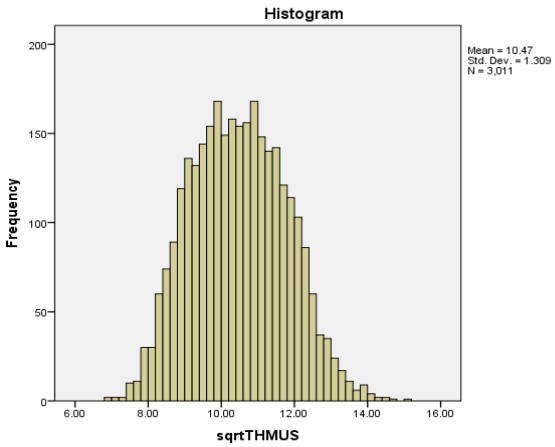

Figure B2: Histogram of sqrtTHMUS

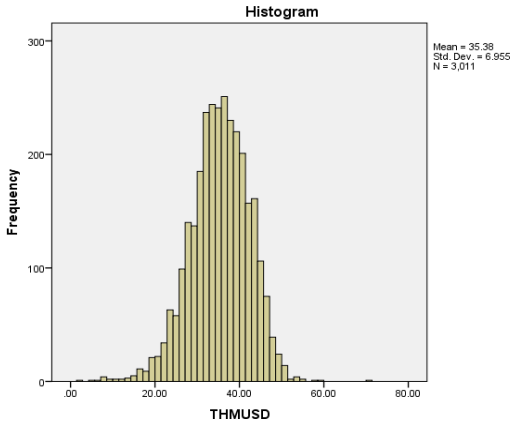

Figure B3: Histogram of THMUSD

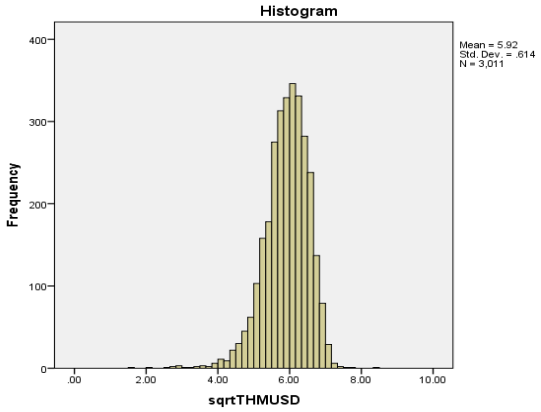

Figure B4 :Histogram of sqrtTHMUSD

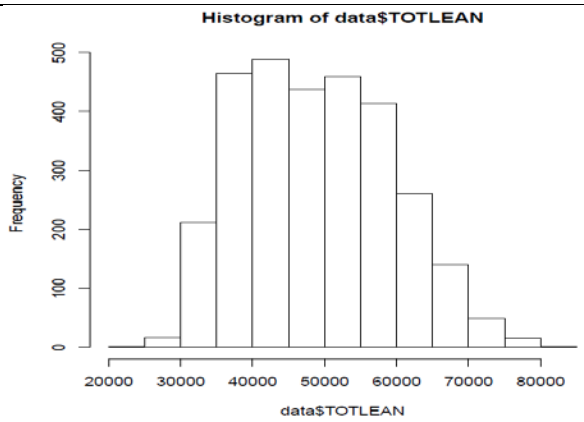

Figure B7: Histogram of TOTLEAN

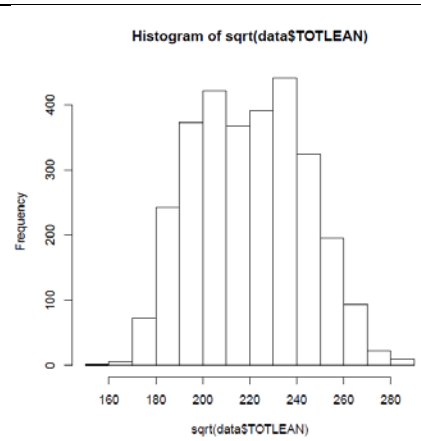

Figure B8: Histogram of sqrtTOTLEAN

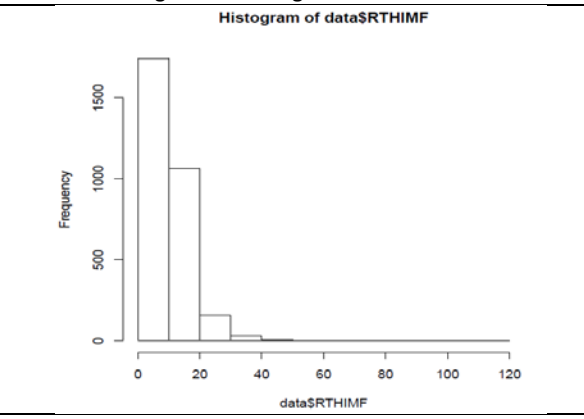

Figure B9: Histogram of RTHIMF

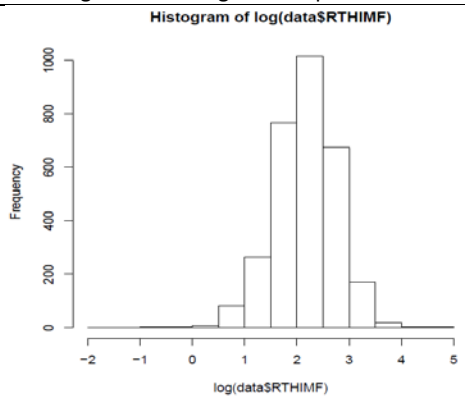

Figure B10: Histogram of sqrtRTHIMF

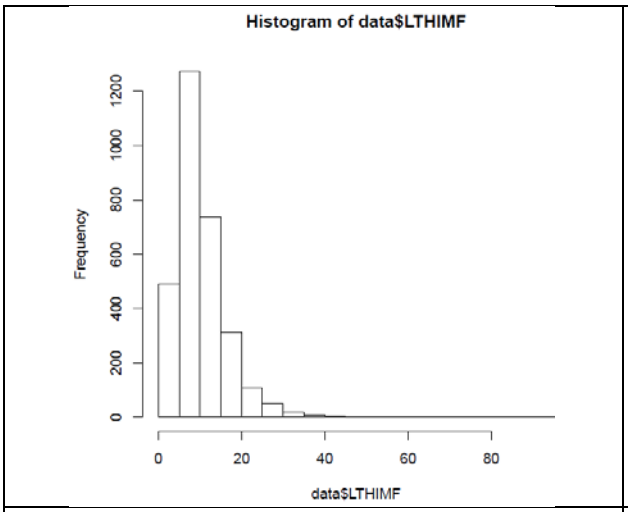

Figure B11: Histogram of LTHIMF

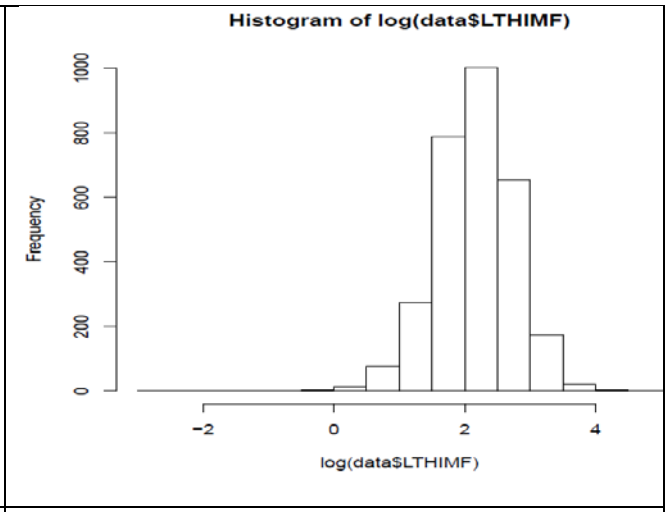

Figure B12: Histogram of loqLTHIMF

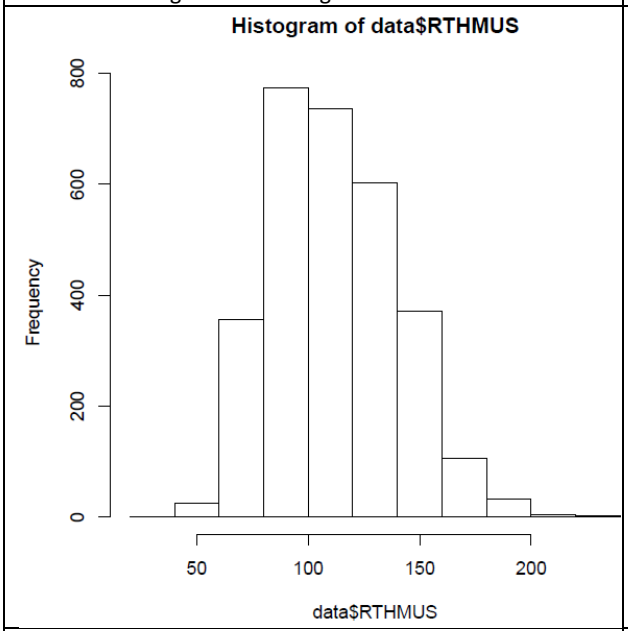

Figure B13: Histogram of RTHMUS

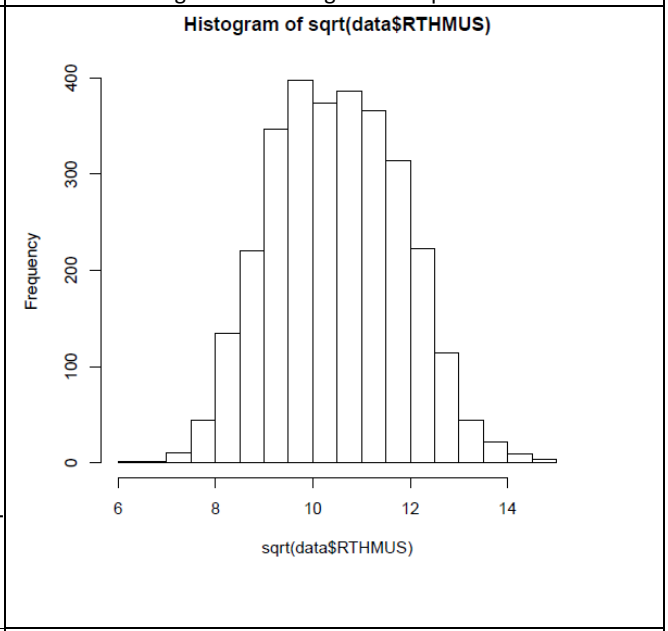

Figure B14: Histogram of sqrtRTHMUS

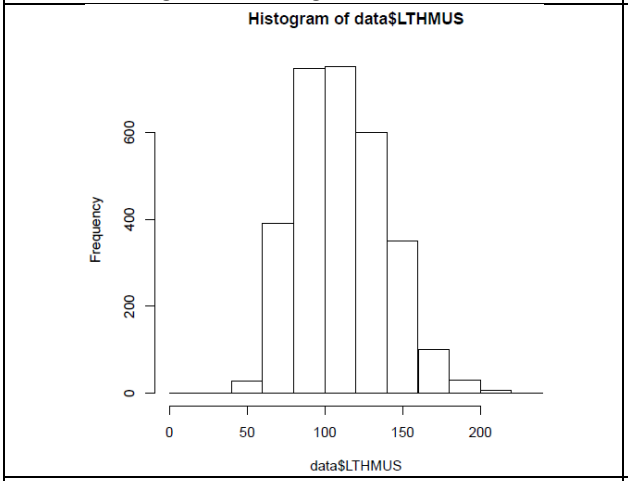

Figure B15: Histogram of LTHMUS

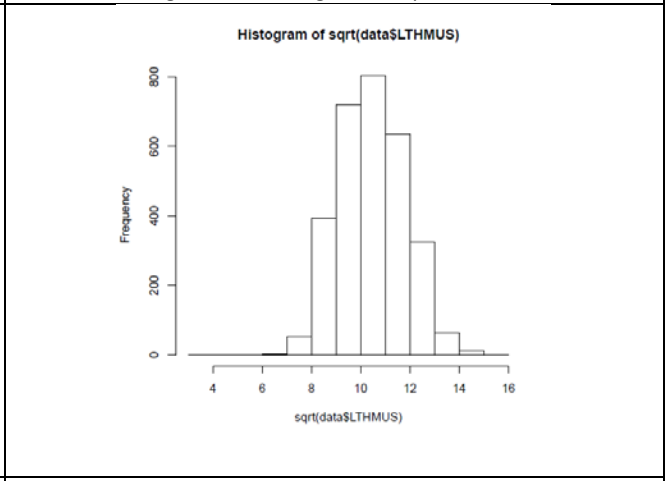

Figure B16: Histogram of sqrtLTHMUS

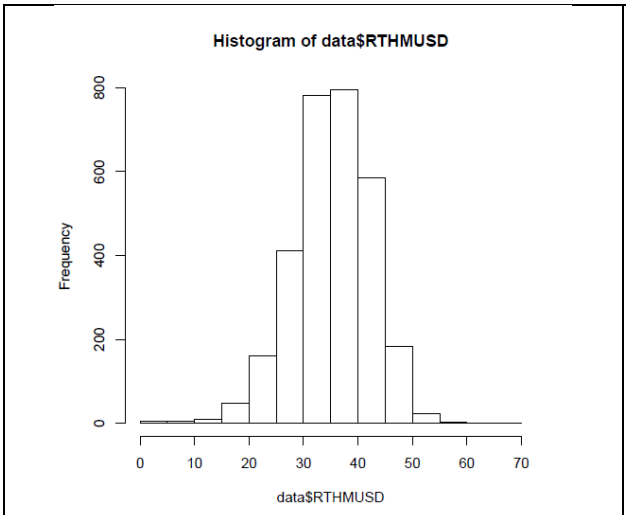

Figure B17: Histogram of RTHMUSD

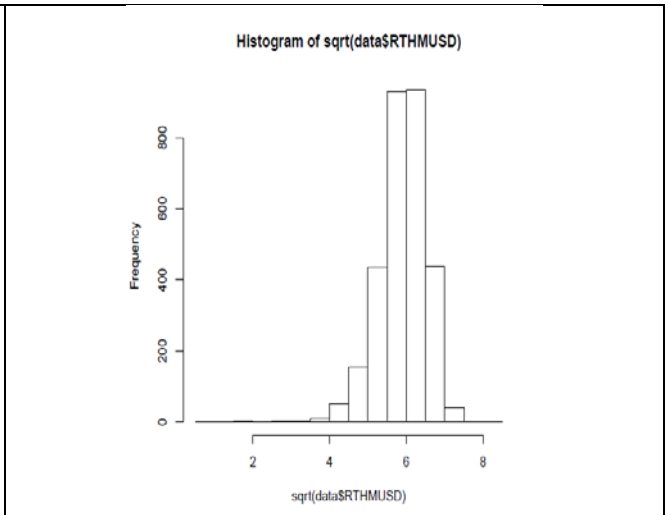

Figure B18: Histogram of sqrtRTHMUSD

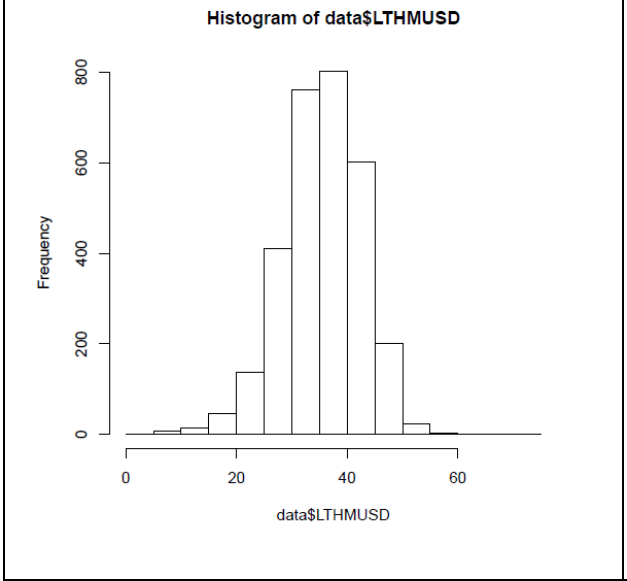

Figure B19: Histogram of LTHMUSD

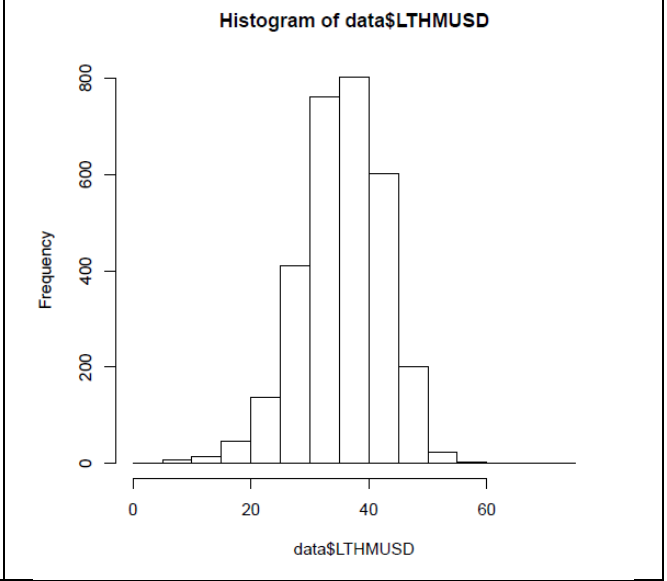

Figure B20: Histogram of untransformed LTHMUSD

APPENDIX C

LOCUS PLOTS OF PHENOTYPES

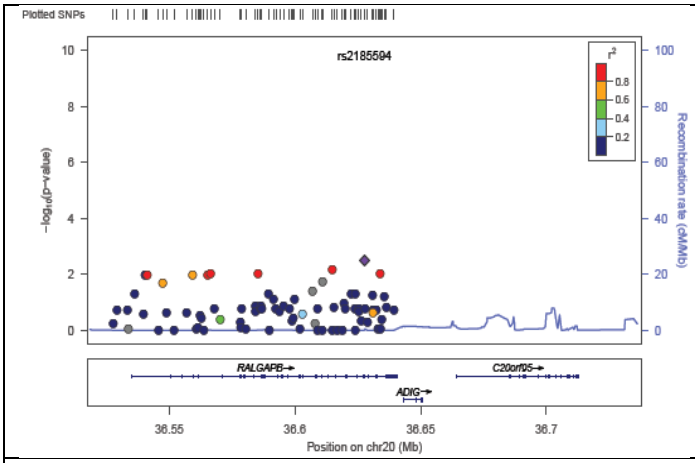

Figure C1 : Association between THIMF and SNPs around *RALGAPB* (rs2185594) on 20q11.13 in African Americans

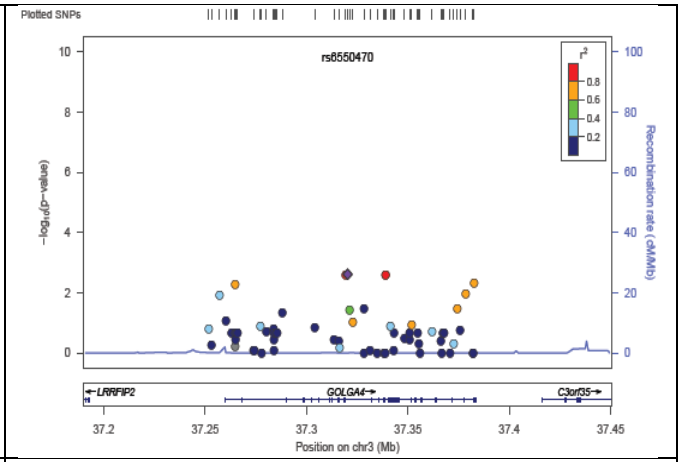

Figure C2: Association between TOTPF and SNPs around *GOLGA4*(rs6550470) on 3p22.2 in African Americans

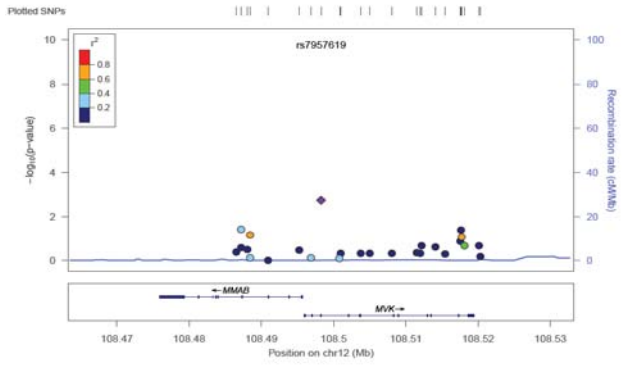

Figure C3: Association between THMUS and SNPs around *MAGOHB* (rs7957619) on 12p13.2 in African Americans

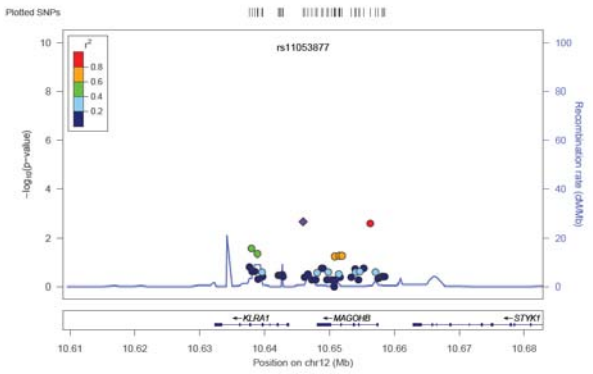

Figure C4: Association between THMUSD and SNPs around MAGOHB (rs11053877) on 12p13.2 in European Americans

APPENDIX D

RELATIONSHIP BETWEEN TRAITS AND AGE

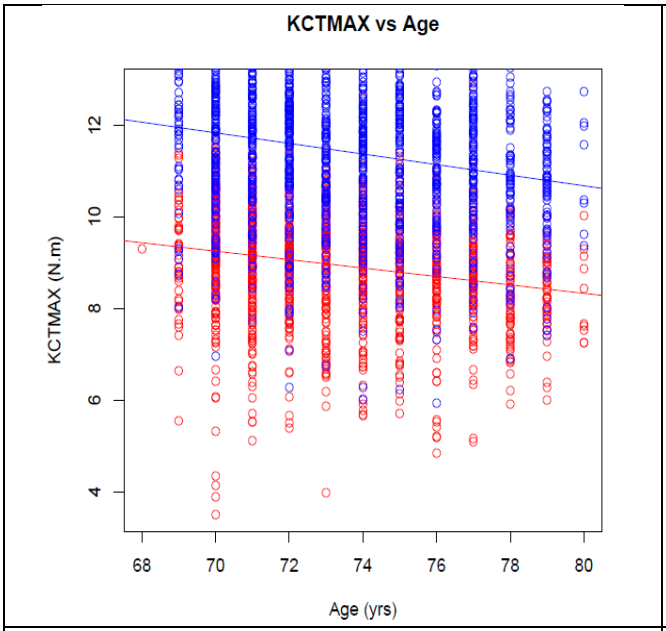

Figure D1: Relationship between KCTMAX and AGE (blue=M, red=F)

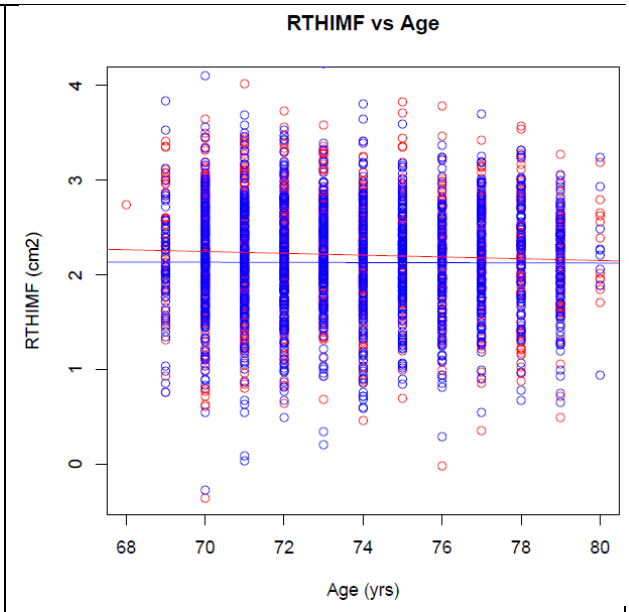

Figure D2: Relationship between RTHIMF and AGE(blue=M, red=F)

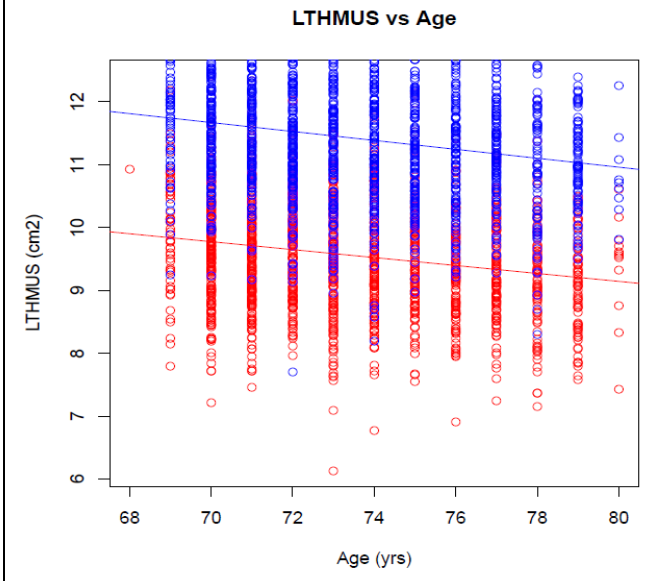

Figure D3: Relationship between LTHMUS and age(blue=M, red=F)

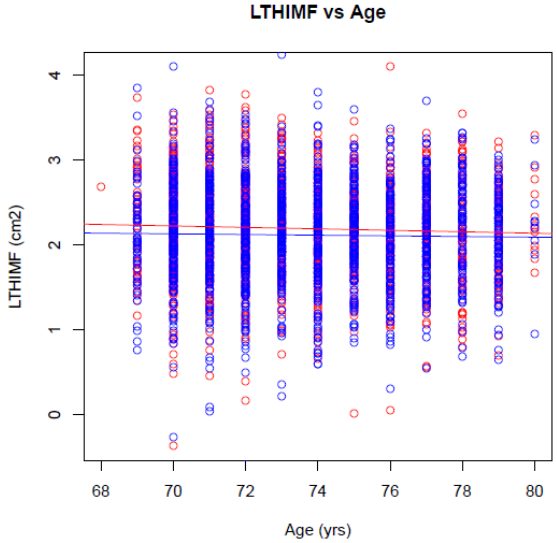

Figure D4: Relationship between LTHIMF and age(blue=M, red=F)

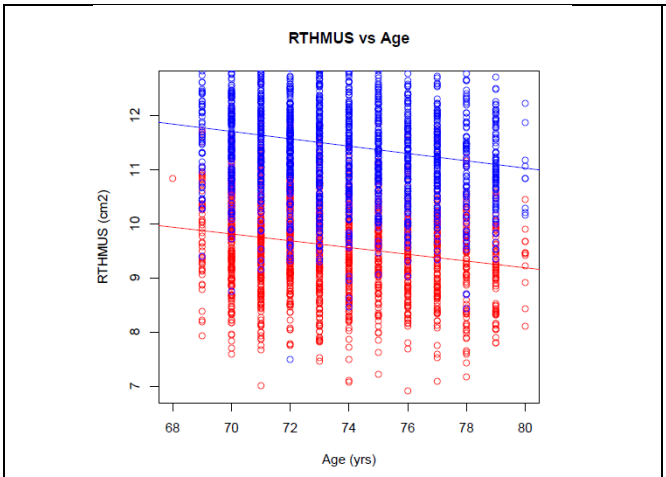

Figure D6: Relationship between RTHMUS and age(blue=M, red=F)

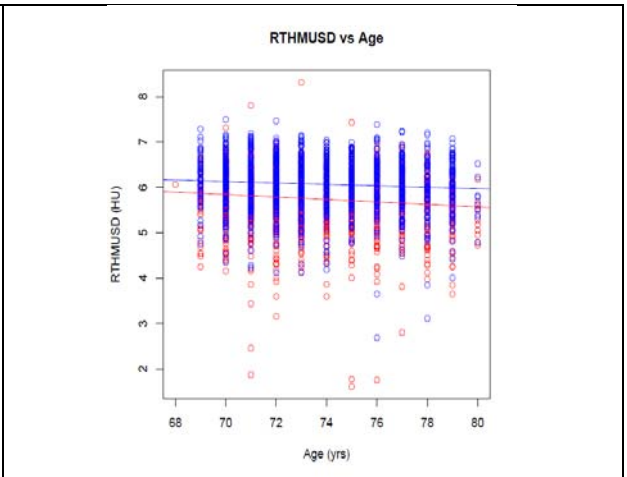

Figure D7: Relationship between RTHMUSD and age(blue=M, red=F)

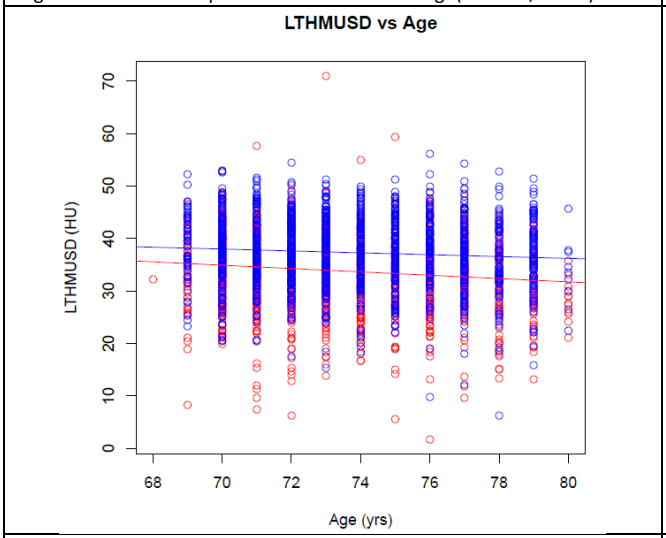

Figure D8: Relationship between LTHMUSD and age(blue=M, red=F)

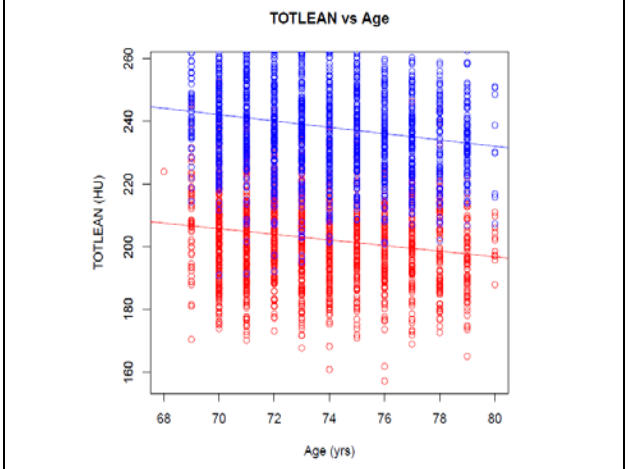

Figure D9: Relationship between TOTLEAN and age(blue=M, red=F)

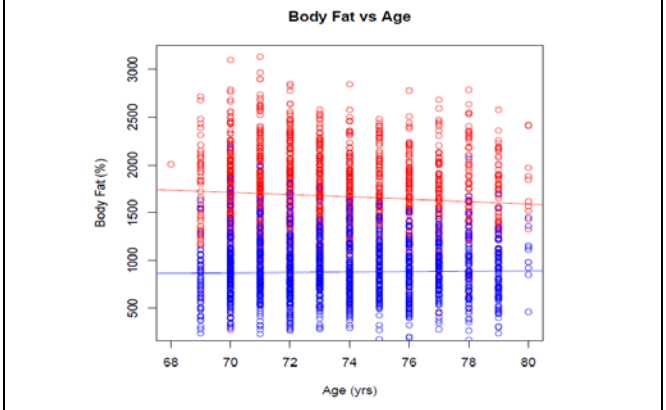

Figure D10: Relationship between TOTPF and age(blue=M, red=F)

APPENDIX E

RELATIONSHIP BETWEEN TRAITS AND RACE

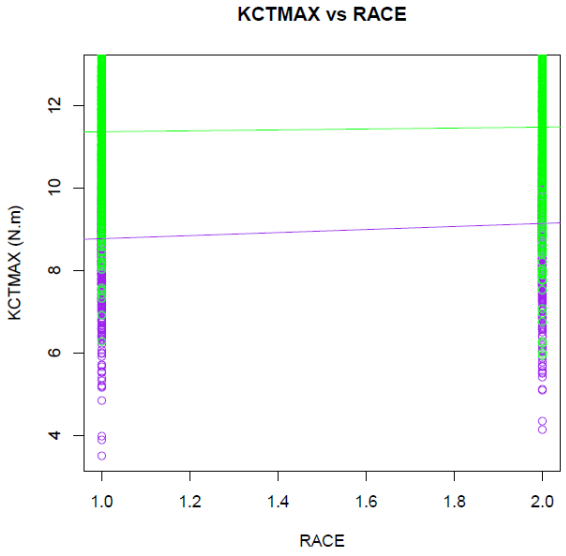

Figure E1: Relationship between KCTMAX and race,green=M,purple=F

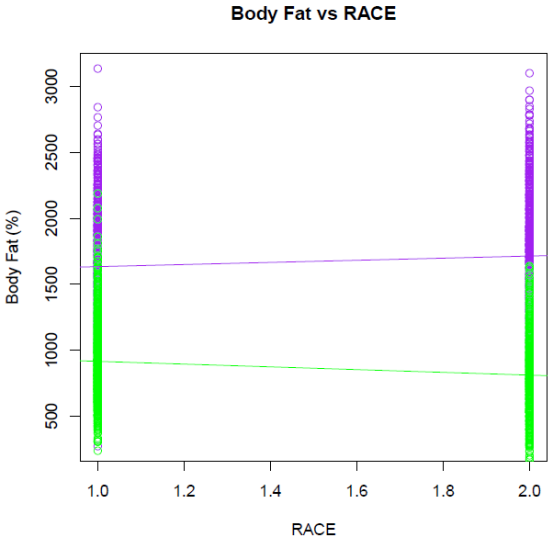

Figure E2: Relationship between TOTPF and race ,green=M,purple=F

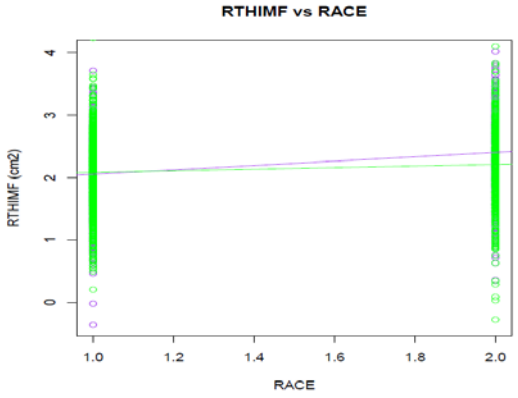

Figure E3: Relationship between RTHIMF and race ,green=M,purple=F

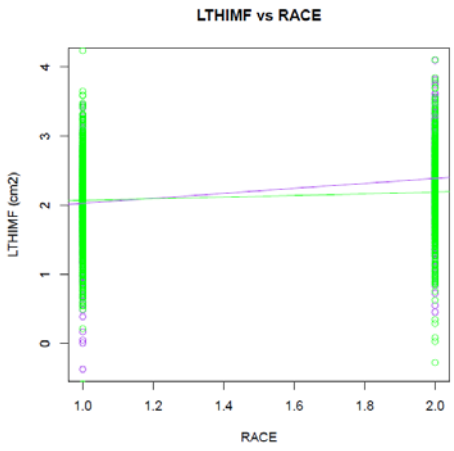

Figure E4: Relationship between LTHIMF and race green=M,purple=F,

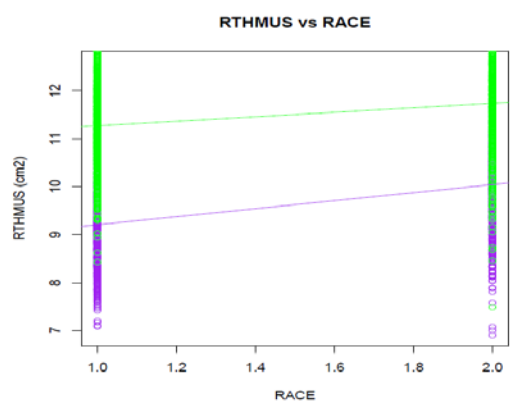

Figure E5: Relationship between RTHMUS and race,green=M,purple=F

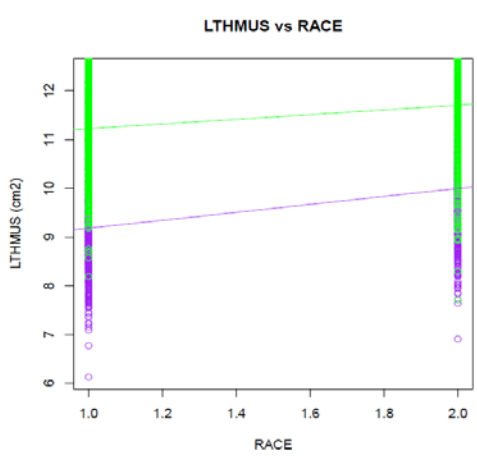

Figure E6: Relationship between LTHMUS and race green=M,purple=F,

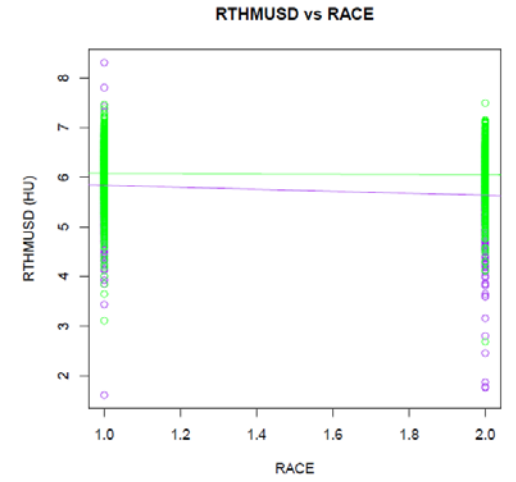

Figure E7: Relationship between RTHMUSD and race,green=M,purple=F

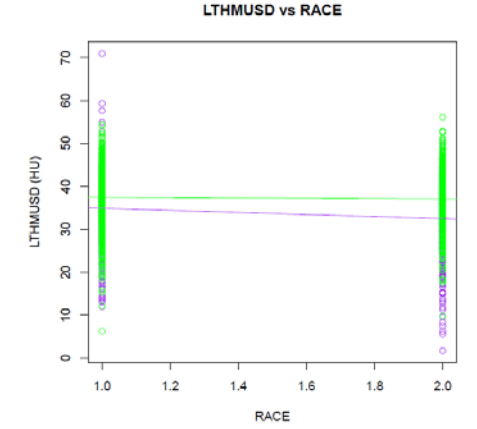

Figure E8: Relationship between LTHMUSD and race, green=M,purple=F

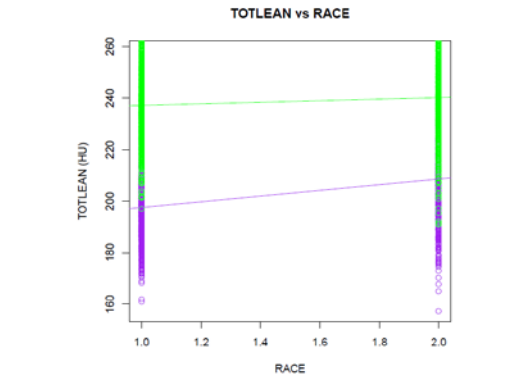

Figure E9: Relationship between TOTLEAN and race green=M,purple=F,

APPENDIX F

GENOTYPE-PHENOTYPE BOX AND WHISKER PLOTS IN EUROPEAN AMERICANS

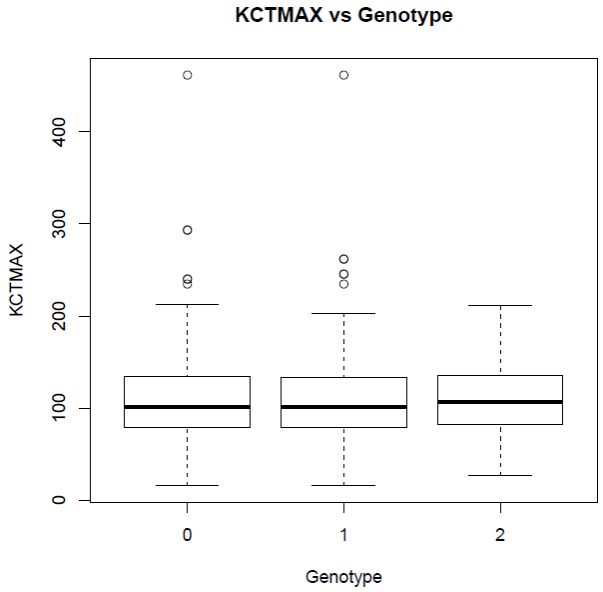

Figure F1: Genotype-KCTMAX box plot(rs10850436)

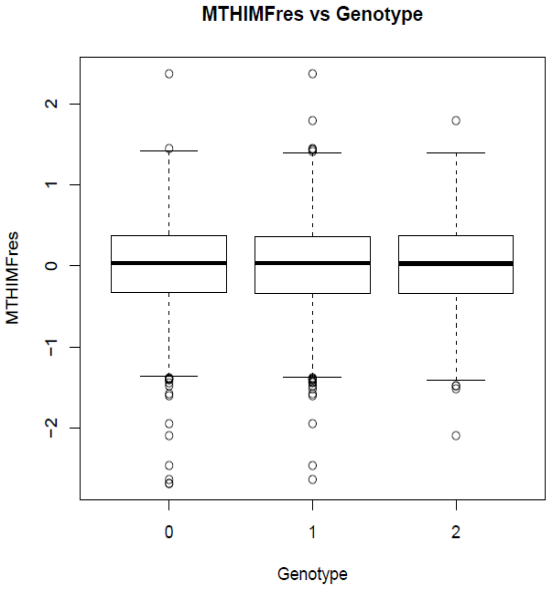

Figure F2: Genotype-THIMF plot(rs1805081)

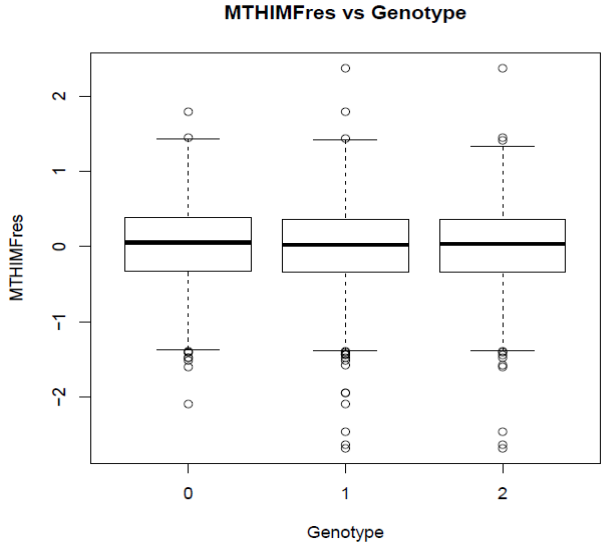

Figure F3: Genotype-THIMF plot(rs1808579)

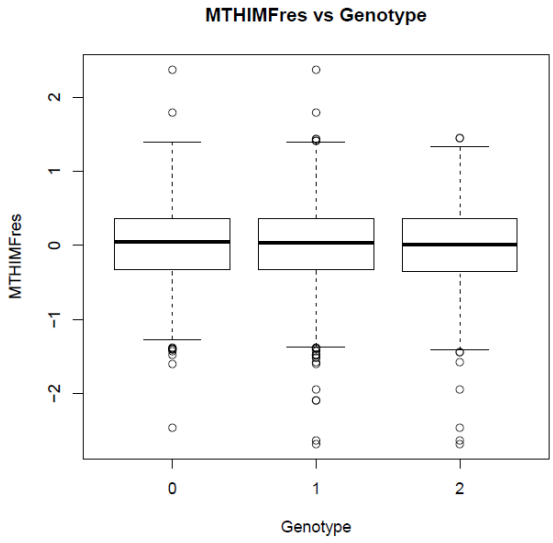

Figure F4: Genotype-THIMF plot(rs10845181)

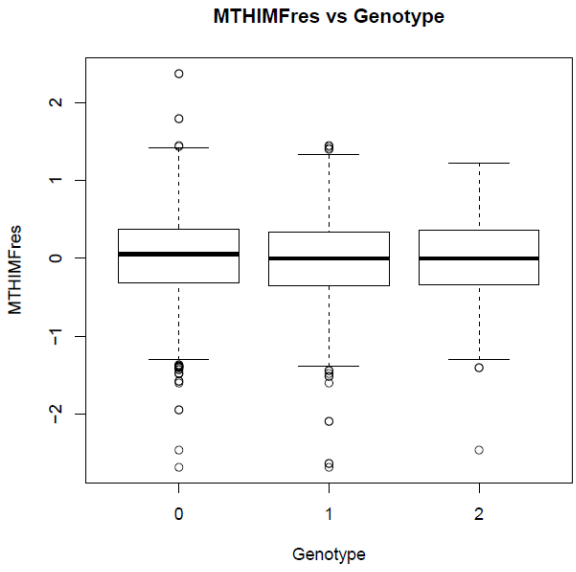

Figure F5: Genotype-THIMF plot(rs11053877)

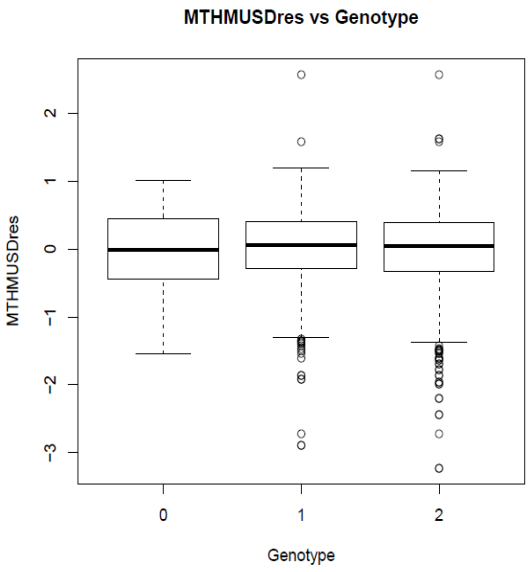

Figure F6: Genotype-MTHMUSD plot(rs1026442)

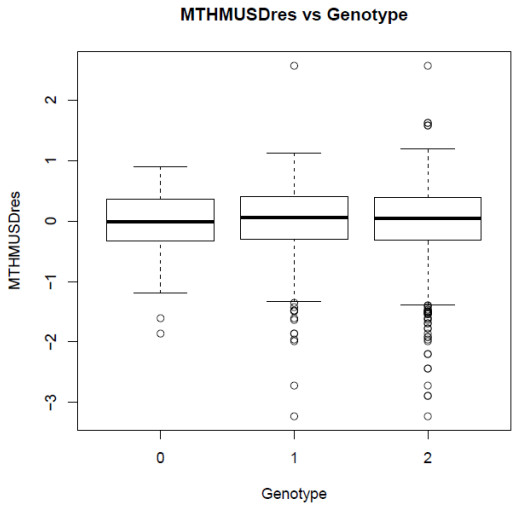

Figure F7: Genotype-THMUSD plot(rs7852856)

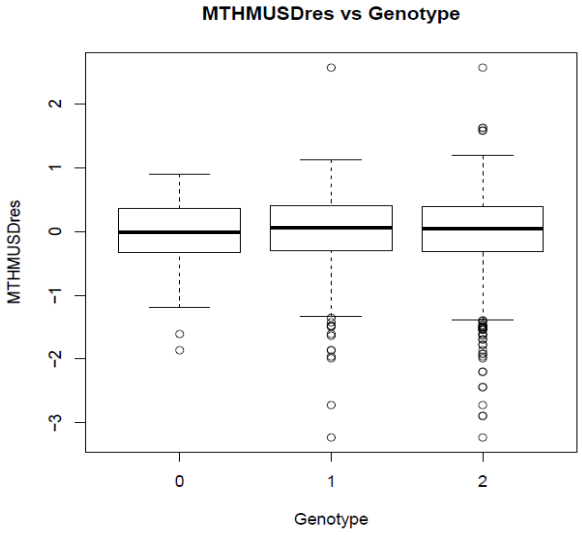

Figure F8: Genotype-THMUSD plot(rs7869223)

APPENDIX G

GENOTYPE-PHENOTYPE BOX AND WHISKER PLOTS IN AFRICAN AMERICANS

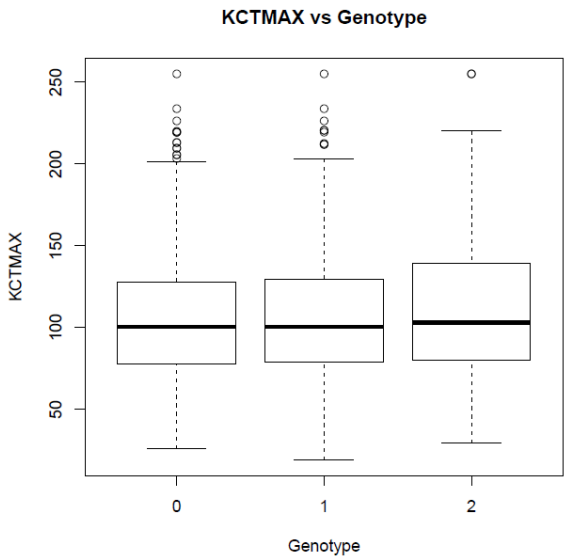

Figure G1: Genotype-KCTMAX plot(rs11264359)

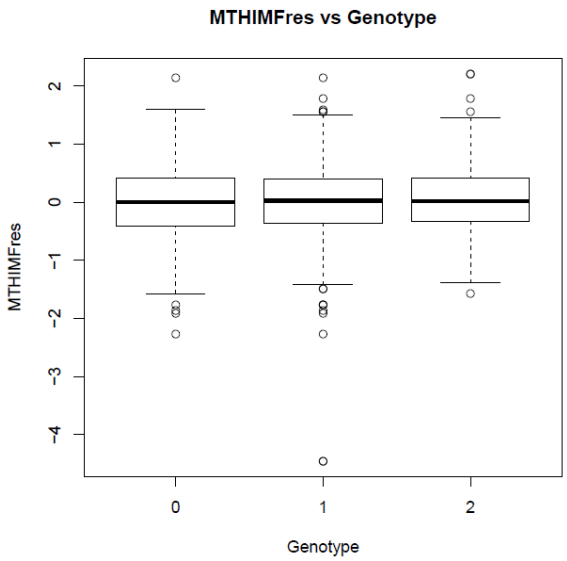

Figure G2: Genotype-THIMF plot(rs17801180)

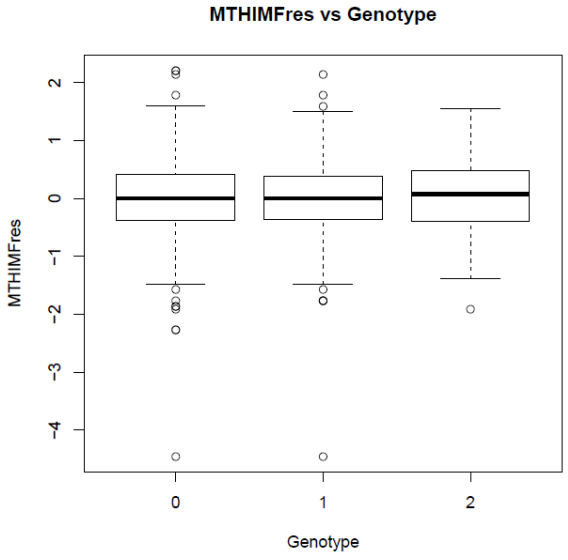

Figure G3: Genotype-THIMF plot(rs2185594)

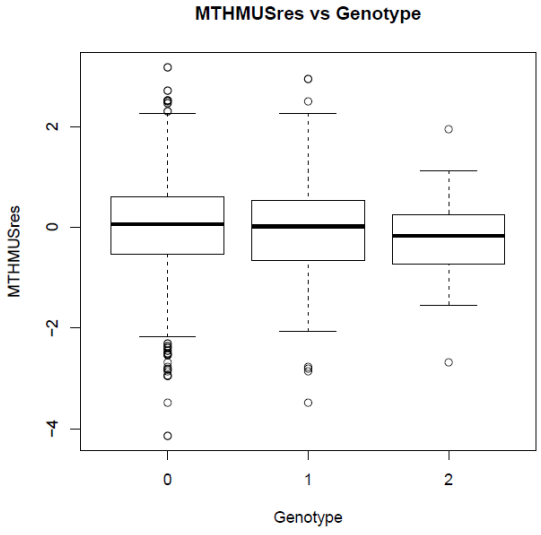

Figure G4: Genotype-THMUS plot(rs7957619)

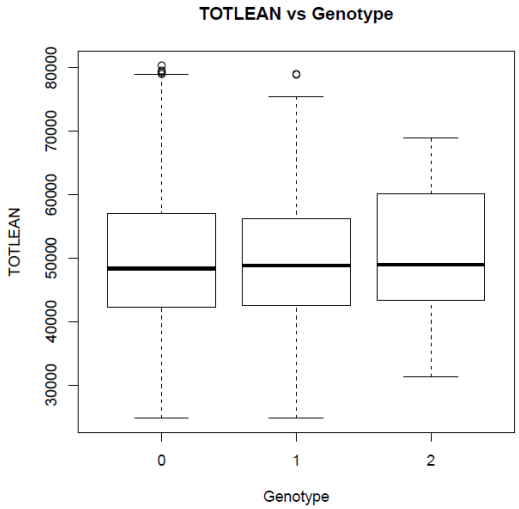

Figure G5: Genotype-TOTLEAN plot(rs11053877)

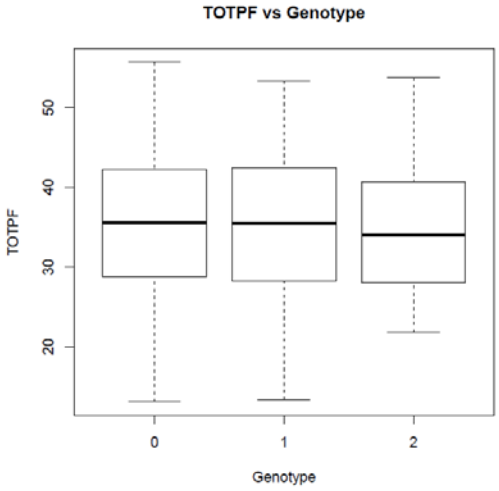

Figure G6: Genotype-TOTPF plot(rs805554)

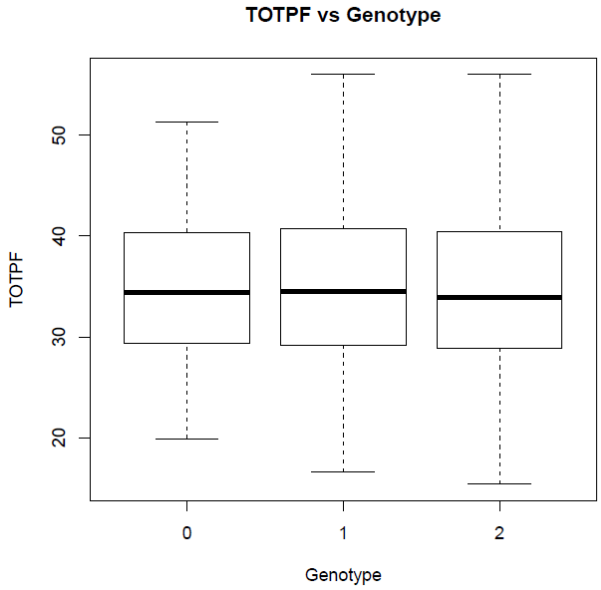

Figure G7: Genotype-TOTPF plot(rs2185594)

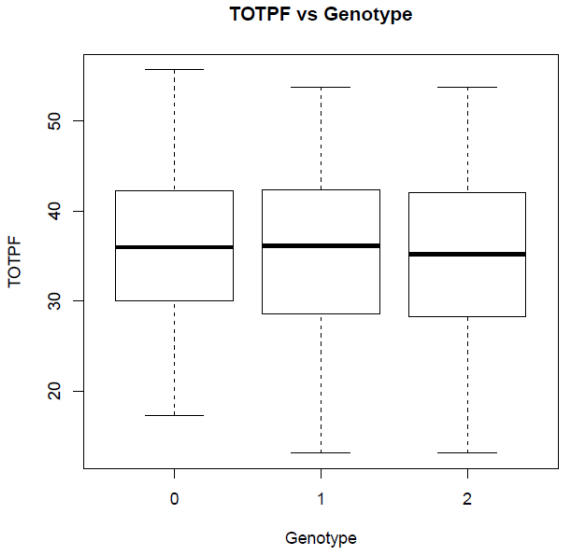

Figure G8: Genotype-TOTPF plot(rs1788781)

APPENDIX H

SNPS WITH SUGGESTIVE P VALUES IN EUROPEAN AMERICANS

Ancestry	Gene	SNP	Location	MAF	Allele	Associated Trait	beta	pvalue
European	*MAGOHB*	rs1026442	Intron1	0.19	A/G	THIMF	-0.0722	2.822×10^{-3}
				0.19	A/G	THMUSD	0.0755	2.582×10^{-3}
		rs10845181	Intron 1	0.49	C/T	THIMF	0.0191	8.433×10^{-3}
		rs11053877	3' UTR	0.19	A/G	THIMF	0.0721	3.062×10^{-3}
				0.19	A/G	THMUSD	-0.0771	2.237×10^{-3}
		rs12811247	Intron 1	0.46	G/T	TOTPF	33.3603	0.714×10^{-3}
	MVK	rs10850436	Intron 1	0.26	A/C	KCTMAX	0.1479	4.848×10^{-3}
	NPC1	rs1808579	intron	0.45	C/T	THIMF	0.0498	7.924×10^{-3}
		rs1805081	exon	0.39	C/T	THIMF	-0.0570	3.174×10^{-3}
	SLC27A4	rs7869223	Intron 1	0.10	C/G	THMUSD	-0.0923	3.41×10^{-3}
		rs7852856	Intron 1	0.10	A/T	THMUSD	-0.0923	3.324×10^{-3}

APPENDIX G

SNPS WITH SUGGESTIVE P VALUES IN AFRICAN AMERICANS

Ancestry	Gene	SNP	Location	MAF	Allele	Associated Trait	beta	pvalue
African	*ADCK3*	rs2077734	Intron1	0.13	G/T	KCTMAX	-0.2598	9.976×10^{-3}
	COQ2	rs4693597	Intron 1	0.28	C/T	THMUS	-0.1307	5.402×10^{-3}
	CYP3A5	rs10242455	intron	0.31	A/G	THIMF	-0.0743	4.768×10^{-3}
				0.31	A/G	THMUSD	0.0742	6.564×10^{-3}
				0.31	A/G	TOTPF	-59.2917	1.325×10^{-3}
		rs4646457	intron	0.32	A/C	TOTPF	-51.0726	5.09×10^{-3}
		rs776746	Intron-variant	0.33	C/T	TOTPF	-50.7495	5.395×10^{-3}
	FDPS	rs3020781	intron	0.42	A/G	KCTMAX	0.1855	5.649×10^{-3}
	GOLGA4	rs6800842	intron	0.47	A/G	TOTPF	50.0047	5.149×10^{-3}
		rs11129756	intron	0.44	A/G	TOTPF	54.5200	2.606×10^{-3}
		rs6550470	intron	0.44	A/G	TOTPF	-54.8735	2.438×10^{-3}
		rs7639447	intron	0.42	C/T	TOTPF	18.2568	2.528×10^{-3}
		rs4635655	intron	0.142	A/G	TOTPF	55.3529	2.488×10^{-3}
		rs17266090	Intron	0.47	A/G	TOTPF	-50.7139	4.71×10^{-3}
	JUP	rs9914318	intron	0.25	C/T	TOTLEAN	2.0862	8.07×10^{-3}
	MAGOHB	rs10746384	intron	0.14	A/G	THMUS	0.1763	2.568×10^{-3}
		rs11053877	3’ UTR	0.09	A/G	TOTLEAN	3.2678	3.407×10^{-3}
	NPC1	rs1652344	intron	0.27	C/T	TOTPF	-50.7536	8.776×10^{-3}
		rs1631685	intron	0.27	C/T	TOTPF	51.8288	7.433×10^{-3}
		rs1652348	intron	0.27	C/T	TOTPF	52.7439	6.498×10^{-3}
		rs1788781	intron	0.27	C/T	TOTPF	19.2749	4.256×10^{-3}
	PDSS1	rs1780179	intron	0.44	C/T	THIMF	0.0645	8.747×10^{-3}
		rs1748357	intron	0.44	A/G	THIMF	-0.0645	8.747×10^{-3}
				0.44	A/G	TOTPF	-45.3115	8.285×10^{-3}
		rs1780180	intron	0.44	A/G	THIMF	-0.0646	8.719×10^{-3}
			intron	0.44	A/G	TOTPF	-45.3559	8.223×10^{-3}
	RALGAPB	rs9941745	intron	0.09	C/T	TOTLEAN	-3.3229	4.439×10^{-3}
		rs9941761	intron	0.09	C/T	TOTLEAN	3.3463	4.158×10^{-3}
		rs805554	intron	0.13	A/T	THIMF	0.0983	9.653×10^{-3}
			intron	0.12	A/T	TOTPF	101.7339	1.527×10^{-4}
		rs1115600	Intron 1	0.09	G/T	TOTPF	-98.7616	8.404×10^{-4}
		rs1295992	intron	0.10	A/G	TOTPF	-100.7957	5.133×10^{-4}
		rs1262459	exon	0.06	A/G	TOTPF	-143.6325	0.33×10^{-3}

		rs805559	intron	0.10	A/G	TOTPF	101.4057	5.305×10^{-4}
		rs805552	intron	0.10	C/G	TOTPF	-101.5515	4.658×10^{-4}
		rs1303577	exon	0.26	C/T	THIMF	0.0730	9.794×10^{-3}
		rs1315305	intron	0.26	A/C	THIMF	0.0730	9.794×10^{-3}
		rs1303571	intron	0.26	A/G	THIMF	0.0739	9.38×10^{-3}
		rs11699660	intron	0.26	C/G	THIMF	-0.0734	9.408×10^{-3}
		rs4812025	intron	0.26	A/G	THIMF	-0.0768	6.346×10^{-3}
		rs2185594	intron	0.26	A/G	THIMF	-0.0834	3.142×10^{-3}
		rs2206783	intron	0.26	A/G	THIMF	0.0733	9.216×10^{-3}
	SACM1L	rs2578677	intron	0.09	C/G	KCTMAX	-0.3095	8.576×10^{-3}

BIBLIOGRAPHY

Barany, M. ATPase activity of myosin correlated with speed of muscle shortening. *J Gen Physiol* , 50(Suppl):197–218, 1967.

Bodeman B.O., Orvedahl A., Cheng T., Ram R.R,Ou Y.H., Formstecher E., Maiti M., Hazelett C.C, Wauson E.M, Balakireva M., Camonis J.H, Yeaman C., Levine B., and Whitw M.A. RalB and the exocyst mediate the cellular starvation response by direct activation of autophagosome assembly. *Cell*, 144(2):253-67, 2011.

Brown, P. A., Driscoll, M., Dudaronek, J. M., Hall, D. H., Herndon, L. A. and Listner, K. M. Stochastic and genetic factors influence tissue-specific decline in ageing C. elegans. *Nature* , (volume n):808+, 2002.doi: 10.1038/nature01135.

Chen X.W., Leto D., Chiang S.H., Wang Q., and Saltiel A.R. Activation of RalA is required for insulin-stimulated Glut4 trafficiking to the plasma membrane via the exocyst and the motor protein Myo1c. *Dev. Cell*, 13(3):391-404, 2007.

Chen X.W., Leto D., Xiong T., Cheng A., Decker S. and Saltiel A.R. A Ral GAP complex links PI 3-kinase/Akt signaling to RalA activation in insulin action. *Mol Bio Cell*, 22(1): 141-52, 2011.

Delmonico, M. J., Harris, T. B., Visser, M., Park, S. W., Conroy, M. B., Velasquez-Mieyer, P., Boudreau, R., Manini, T. M., Nevitt, M., Newman, A. B., and Goodpaster, B. H. Longitudinal study of muscle strength, quality, and adipose tissue infiltration. *Am J Clin Nutr.* 90:1579–85, 2009.

Dirks, A.J. and Leeuwenburgh, C. The role of apoptosis in agerelated skeletal muscle atrophy. *Sports Med,* 35:473–483, 2005.

Evans, W. J. What is sarcopenia? *The Journals of Gerontology Series*, 50A:5-8, 1995.

Evans, W. J. Functional and metabolic consequences of sarcopenia. *J Nutr* ,127:998S–1003S, 1997.

Evans, W. J. Skeletal muscle loss: cachexia, sarcopenia, andinactivity. *Am J Clin Nutr*, 91:1123S–1127S, 2010.

Fielding, R. A., Vellas, B., Evans, W.J., Bhasin, S., Morley, J. E., Newman, A. B., vanKan, G.A., Andrieu, S., Bauer, J., Breuille, D., Cederholm, T., Chandler, J., De-Meynard, C., Donini, L., Harris, T., Kannt, A., Guibert, F. K., Onder, G., Papanicolaou, D., Rolland, Y., Rooks, D., Sieber, C., Souhami, E., Verlaan, S., and Zamboni, M. Sarcopenia: An undiagnosed condition in older adults. Current consensus definition: prevalence, etiology, and consequences. International working group on sarcopenia. *J Am Med Dir Assoc*, 12:249-256, 2011.

Fisher, A. L. Of worms and women: sarcopenia and its role in disability and mortality. *J Am Geriatr Soc*, 52:1185–1190, 2004.

Giaccaglia, V., Nicklas, B., Kritchevsky, S., Mychaleckyz, J., Messier, S., Bleecker, E., and Pahor, M. Interaction between angiotensin converting enzyme insertion/deletion genotype and exercisetraining on knee extensor strength in older individuals. *Int JSports Med* 29:40–44, 2008.

Glenn, C. F., Chow, D. K., Gami, M. S., Iser, W. B., Hanselman, K. B., Wolkow, C. A., David, L., Goldberg, I. G., and Cooke, C. A. Behavioral deficits during early stages of aging in *Caenorhabditis elegans* result from locomotory deficits possibly linked to muscle frailty. *Gerontol A Biol Sci Med Sci*, 59: 1251–1260, 2004.

Guralnik, J. M., Ferrucci, L., Simonsick, E. M., Salive, M. E., and Wallace, R. B. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med,* 332:556-61, 1995.

Kashyap, L., Perera, S., and Fisher, A. L. Identification of novel genes involved in sarcopenia through RNAi screening in caenorhabditis elegans. *J Gerontol A Biol Sci Med Sci*, 67A:56–65, 2012. doi:10.1093/gerona/glr072.

Klei, L., Luca, D., Devlin, B., and Roeder, K. Pleiotropy and principal components of heritability combine to increase power for association analysis. *Genet. Epidemiol*, 32: 9–19, 2008. doi: 10.1002/gepi.20257.

Morley, J. E., Baumgartner, R. N., Roubenoff, R., Mayer, J., and Nair, K. S. Sarcopenia. *J Lab Clin Med*, 137:231-43, 2001.

Pruim, R. J., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., Gliedt, T. P., Boehnke, M., Abecasis, G. R., and Willer, C. J. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics, 26:2336–2337, 2010. doi: 10.1093/bioinformatics/btq419.

Rantanen, T. Muscle strength, disability and morality. Scand J Med Sci Sports, 13:3-8, 2003.

Rantanen, T., Guralnik, J. M., Ferrucci, L., Penninx, B. W. J. H., Leveille, S., Sipilä, S., and Fried, L. P. Coimpairments as predictors of severe walking disability in older women. *J Am Geriatr Soc,* 49:21–27, 2001.

Rekeneire, N. D., Visser, M., Peila, R., Nevitt, M. C., Cauley, J. A., Tylavsky, F. A., Simonsick, E. M., and Harris, T. B. Is a fall just a fall: correlates of falling in healthy older persons. The health, aging and body composition study. *J Am Geriatr Soc,* 51:841–846, 2003.

Roignant J.Y. and Treisman J.E. Exon junction complex subunits are required to splice Drosophila MAP kinase, a large heterochromatic gene. *Cell*, 143(2):238-50, 2010.

Roubenoff, R. Sarcopenia: effects on body composition and function. *Journal of Gerontology: Medical Sciences,* 58A:1012–1017, 2003.

Segre, A. V., DIAGRAM Consortium, MAGIC Investigators, Groop, L., Mootha, V. K., Daly, M. J., Altshuler, D., and et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet* 6(8):1-19, 2010.

Silver D.L., Watkins-Chow D.E., Schreck K.C, Pierfecile T.J., Larson D.M, Burnetti A.J., Liaw H.J., Myung K., Walsh C.A., Galano N., and Pavan W.J. The exon junction complex component Magoh controls brain size by regulating neural stem cell division. *Nat. Neuroscience*, 13(5):551-8, 2010.

Sternfeld, B., Ngo, L., Satariano, W. A., and Tager, I. B. Associations of body composition with physical performance and self-reported functional limitation in elderly men and women. *Am J Epidemiol,* 156:110–21, 2002.

Strotmeyer ES, de Rekeneire N, Schwartz AV, Resnick HE, Goodpaster BH, Faulkner KA, Shorr RI, Vinik AI, Harris TB, Newman AB; Health ABC Study. Sensory and motor peripheral nerve function and lower-extremity quadriceps strength: the health, aging and body composition study.*J Am Geriatr Soc*. 2009 Nov; 57(11):2004-10, 2009

Tan, L.-J., Liu, S.-L., Lei, S. -F., Papasian, C. J., and Deng, H.-W. Molecular genetic studies of gene identification for sarcopenia. *Hum Genet,* 131:1–31, 2012.

Vandervoort, A. A. Aging of the human neuromuscular system. *Muscle Nerve,* 25: 17–25, 2002.

Visser, M., Kritchevsky, S. B., Goodpaster, B. H., Newman, A. B., Nevitt, M., Stamm, E., and Harris, T. B. Leg muscle mass and composition in relation to lower extremity performance in men and women aged 70 to 79: the health, aging and body composition study. *J Am Geriatr Soc,* 50:897–904, 2002.

Volpi, E., Nazemi, R., and Fujita, S. Muscle tissue changes with aging. *Curr Opin Clin Nutr Metab Care*, 7:405–410, 2004.

Wang, K., Li, M., and Bucan, M. Pathway-based approaches for analysis of genomewide association studies. *Am. J. Hum. Genet,* 81:1278–1283, 2007.
doi: 10.1086/522374.