12-2020

GENETIC AND CLINICAL DETERMINANTS OF RACIAL/ETHNIC DIFFERENCES IN MULTIPLE MYELOMA SUSCEPTIBILITY AND OUTCOMES FOCUSING ON HISPANICS

Alem Belachew

GENETIC AND CLINICAL DETERMINANTS OF RACIAL/ETHNIC DIFFERENCES IN MULTIPLE MYELOMA SUSCEPTIBILITY AND OUTCOMES FOCUSING ON HISPANICS

by

Alem Abebe Belachew, M.S.

APPROVED:

_____
Michelle A. T. Hildebrandt, Ph.D.
Advisory Professor

_____
Chad D. Huff, Ph.D.

_____
Ann M. Killary, Ph.D.

_____
Robert Z. Orlowski, M.D., Ph.D.

_____
Susan K. Peterson, Ph.D., M.P.H.

APPROVED:

_____
Dean, The University of Texas

MD Anderson Cancer Center UTHealth Graduate School of Biomedical Science

GENETIC AND CLINICAL DETERMINANTS OF RACIAL/ETHNIC DIFFERENCES IN MULTIPLE MYELOMA SUSCEPTIBILITY AND OUTCOMES FOCUSING ON HISPANICS

A

DISSERTATION

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Alem Abebe Belachew, M.S.

Houston, Texas

December 2020

Dedication

To my clan, Lemlem Sissay, Tsegahiwot Abebe, Mulugeta Amha, Yamrot Mulugeta, Bereket Mulugeta, Addis Mulugeta, Samson Mulugeta, and Metasebia Mulugeta. I could not have done it without your unconditional love and support.

To my uncle Tekeba, you instilled in me the value of education at an early age. I hope I made you proud.

To my late father and uncle, Babi and Taye, your influence made me pursue a Ph.D. in cancer research, and you will live forever in my memories.

Acknowledgments

First and foremost, I would like to extend my gratitude to my advisor Dr. Michelle Hildebrandt for her guidance in my scientific, professional, and personal development. From Michelle, I learned scientific rigor, interdisciplinary collaboration, and effective scientific communication. I am thankful for her support and for giving me the freedom to expand my horizon in leadership and career development. I could not have asked for a better mentor.

My thesis advisors, Dr. Chad Huff, Dr. Ann Killary, Dr. Robert Orlowski, and Dr. Susan Peterson, thank you for invaluable insight and for pushing me to thrive for high standards in my work. Our lab collaborator, Dr. Elisabet Manasanch, thank you for helping me capture the clinical data of my study subjects in great detail. I want to extend a special thank you to Dr. Ryan Bohlender, who was instrumental in moving my project forward by helping me with the genetic ancestry analysis.

The people at the MD Anderson Center for Entrepreneurship Advancement and the UT System Venture Mentoring Service, thank you for an unforgettable internship experience that opened my eyes to the business side of science and innovation.

UT Health Student InterCouncil (SIC) and the UT Health student body, thank you for choosing me to be your student leader. Because of SIC, I learned what it means to be a service driven leader.

My friends and family who are too many to name, I am grateful for your presence in my life and for inspiring me to push forward. I appreciate every one of you. My boyfriend, Charlie, thank you for your optimism, support, and for helping me edit this thesis.

The entire GSBS community, thank you for a fun and rewarding Ph.D. experience.

Abstract

GENETIC AND CLINICAL DETERMINANTS OF RACIAL/ETHNIC DIFFERENCES IN MULTIPLE MYELOMA SUSCEPTIBILITY AND OUTCOMES FOCUSING ON HISPANICS

Alem Abebe Belachew, M.S.

Advisory Professor: Michelle A. T. Hildebrandt, Ph.D.

Multiple Myeloma (MM) constitutes 10% of diagnosed hematologic malignancies in the US, with over 12,000 deaths recorded each year. Race/ethnicity is a well-known MM risk factor, where individuals of African descent have over 2- to 3-fold increased risk of incidence compared to those of European descent. Additionally, Hispanics are diagnosed approximately three years younger than white American counterparts, for unknown reasons. Differences in clinical phenotype are also present for MM patients by ancestry, including varying rates of common initiation mutations such as IgH translocations and TP53 mutation between patients of European and African descent. Studies have begun to interrogate the genetic basis for differences in MM susceptibility and other clinical endpoints in populations of European and African lineage. However, there is a gap in our understanding of the genetic etiology of MM susceptibility in Hispanics. Furthermore, MM clinical features have yet to be described in Hispanics, precluding genetic studies of MM clinical outcomes by race/ethnicity.

This study examined the effect of genetic ancestral background on MM susceptibility and clinical endpoints by utilizing the genome-wide genotype dataset and robust medical records of a multi-ethnic patient population seen at MD Anderson Cancer Center. We conducted case-control association analysis in 143 self-identified Hispanic, 211 non-Hispanic black, 262 non-Hispanic white MM cases, and 633 healthy controls. We also

described MM clinical characteristics at diagnosis in Hispanic patients and performed a comparative analysis of clinical phenotypes by self-reported ethnicity and genetic ancestry.

We discovered differential risk in MM susceptibility by genetic ancestry. We also identified unique patterns in Hispanics' baseline clinical phenotype compared to self-reported non-Hispanic black and non-Hispanic white patients. Our study also revealed Hispanics with elevated European ancestry to be at an increased risk of genetic abnormalities associated with poor MM prognosis. Moreover, we identified genetic variants within the Wnt/beta-catenin pathway associated with MM risk that vary by race/ethnicity.

Our findings may be clinically applicable to filling the knowledge gap regarding the genetic contributors of MM susceptibility and outcomes in diverse patient populations and towards eliminating self-report bias of race/ethnicity to better define risk associations and better manage patient outcomes.

Table of Contents

List of Illustrations

List of Tables

Abbreviations

MM – multiple myeloma

MGUS – monoclonal gammopathy of undermined significance

SMM – smoldering myeloma

NHW – non-Hispanic white

NHB – non-Hispanic black

SEER – surveillance, epidemiology, and end results

MRI – magnetic resonance imaging

PET CT – positron emission tomography–computed tomography

B2M – beta-2-microglobulin

ASCT – autologous stem cell transplant

BCL11A – b-cell lymphoma/leukemia 11A

BTRC – beta-transducin repeat containing E3 ubiquitin protein ligase;

CCDC57 – coiled-coil domain containing 57

PCA – principal component analysis

PC – principal component

LD – linkage disequilibrium

CEU – Utah residents with Northern and Western European ancestry

YRI – Yoruba in Ibadan, Nigeria

MXL – Mexican Ancestry in Los Angeles, CA, USA

CI – confidence interval

CSNK1D – casein kinase 1 Delta

DCXR – dicarbonyl and l-xylulose reductase

DKK1 – dickkopf wnt signaling pathway inhibitor 1

DVL – dishevelled

eQTL – expression trait loci

EZH2 – enhancer of zeste homolog 2

IKZF1 – IKAROS family zinc finger 1

IRF4 – interferon regulatory factor 4

LRP6 – low-density lipoprotein receptor-elated protein-6

SPI1 – spi-1 proto-oncogene

MDA-CPSC – MD Anderson Cancer Patient and Survivors Cohort

MOI – model of inheritance

MWG – Myeloma Working Group

OR – odds ratio

HR – hazard ratio

Chapter 1: Introduction

Background

Multiple Myeloma (MM) Epidemiology and Disease Overview

Multiple Myeloma (MM) makes up ~10% of diagnosed hematologic malignancies in the U.S., with over 32,000 new cases and over 12,000 deaths recorded each year(1). The median age of diagnosis is 69, with the elderly (age > 75) making up most cases(2). Although MM accounts for only 1.8% of all cancers, it is an incurable disease with a high symptom burden due to the physical manifestation of the disease, such as bone destruction, which impairs patients' quality of life(3). Improved management strategies and novel effective treatments have dramatically improved MM outcomes over the past two decades(4). Nevertheless, almost all MM patients frequently relapse, and remain on chemotherapy for the rest of their lives(5).

As a blood disorder, MM is characterized by the overproduction of clonal plasma cells and elevated levels of monoclonal immunoglobulin (M-proteins) in the bone marrow. Furthermore, end organ damage events described by hypercalcemia, renal insufficiency, anemia, and lytic bone lesions, collectively known as the CRAB criteria, signify active MM requiring treatment.

The etiology of MM is not fully known. However, a small number of risk factors have been associated with an increased incidence of MM, including older age and high body mass index reported by a meta-analysis showing a risk ratio of 1.12 for overweight individuals, and a 1.21 risk ratio for those that were obese(6). In contrast, some studies reported an inverse relationship between MM incidence and fruit/vegetable consumptions(7, 8). Also, MM is more common in men and individuals of African descent(9), which will be further elaborated in the next section. Interestingly, MM displays geographical differences in incidence that as of yet have unclear reasons.

North American and European countries report the highest incidence rates, followed by populations of Southern and Middle African origins, as well as Caribbean countries. Asian countries and U.S residents of Asian descent report the lowest incidences(10). Other MM risk factors include environmental and occupational hazard through exposure to chemical agents, including pesticides, organic solvents, and hair dyes(11). A group from Sloan Kettering published a strong supporting study for the role of occupational exposures on MM susceptibility, which showed a disproportionate number of firefighters exposed to the 9/11 World Trade Center disaster in 2001 who were at an increased risk of early-onset aggressive MM(12).

Genetics also play a role in MM susceptibility. Individuals with a family history of MM and other hematological malignancies, including 1st degree relatives, were also shown to have a substantial increase of MM incidence(10). Germline genetic studies have identified 23 MM susceptibility loci in individuals of European descent(13). Almost all identified risk loci, including 8q24.21, 6p22.3, 3p22.1, and 7p15.3(14–17), are located in non-coding regions suggesting their influence via gene regulation. Furthermore, two recent meta-analyses comparing patients of African and European ancestry found overlap between 20 of the 23 prior reported risk regions, suggesting shared susceptibility loci across populations(18), with MM risk association variation on p23.3, 17p11.2, 3p22.1, 22q13.1, 7p15.3 regions between African and European ancestry(19). One epigenetic study identified hyperphosphorylation of the antigenic paraprotein target protein, paratarg-7 (pP-7), which is overexpressed in MM, to occur in a substantially higher frequency in patients of African descent, suggesting differential epigenetic events by ancestral lineage leading to MM incidence via antigenic stimulation(20). Nevertheless, much of MM's genetic etiology and hereditability remain

unknown, and MM progression cannot be linked to any unique genetic or environmental event.

Moreover, MM is preceded by asymptomatic conditions, namely monoclonal gammopathy of undermined significance (MGUS) and smoldering myeloma (SMM)(21). MGUS is characterized by $\leq$ 10 % clonal plasma cells infiltration in the bone marrow with $\leq$ 3g/dL of M-protein(22) where SMM presents $\geq$ 3 g/dL serum M-protein and/or 10-60% of plasma cells in the bone marrow(23), but without manifestation of CRAB features. Due to a lack of population-based registries or systematic screening programs for MGUS and SMM, the prevalence rate of these conditions is difficult to estimate. However, approximately 3.2% of Caucasians over 50 are estimated to be living with MGUS based on data from a retrospective study on 28,000+ individuals who underwent routine clinical screening at the Mayo Clinic(24). Interestingly a nation-wide study of over 4 million individuals admitted to 142 Veterans Affairs hospitals found an MGUS prevalence rate 2 to 3-fold higher in African Americans compared to whites(22), for reasons that remain unclear. This striking disparity in MGUS prevalence by race/ethnicity is also a recurring trend in MM susceptibility and will be discussed further in the next section.

The estimated SMM incidence rate is 0.9 cases per 100,000 persons(25) with varying risk of MM progression based on the burden of circulating plasma cell, serum light chain ratio produced by the monoclonal antibodies, mutational events, and suggestions of approaching end-organ damage(26). SMM patients with bone marrow plasma cell > 60%, free light chain ratio of > 100, and more than one focal lesions detected by radiographic imaging are classified as having high-risk SMM and are often treated(27). While not all MGUS and SMM patients will develop MM, the likelihood to progression increases by 1% and 10% per year, respectively(27).

The spectrum of plasma cell expansion between MGUS, SMM, and MM has provided a unique platform for investigating the genomic hierarchy and clonal evolution of these disease stages. MM initiating primary cytogenetic subtypes can be broadly divided into two groups: translocations involving the immunoglobulin heavy chain (IgH) locus, and hyperdiploidy that are often trisomies of odd-numbered chromosomes(28). The most common IgH translocations include t(11;14), t(6;14), t(4;14), t(14;16) and t(14;20); the latter three are associated with poor prognosis(29, 30). In contrast, trisomic tumors are associated with favorable overall survival(31). IgH translocations promote overexpression and dysregulation of oncogenes cyclin D1 (*CCND1*), cyclin D3 (*CCND3)*, fibroblast growth factor receptor 3 (*FGFR3*), MM SET domain (*MMSET*), and transcription factors Maf and MafB(32)— leading to the accumulation of mutations resulting in disease progression. Secondary cytogenetic abnormalities like monosomy 13/del13q have been identified from the onset of the disease, whereas del17p coding *TP53*, 1q gain, and the Ig translocation involving the 8q24 *MYC* oncogene(33, 34) are seen with disease progression. MGUS, SMM, and MM indeed share some of these genetic events. For instance, del(17p), t(4:14),1q gains, t(4;14), t(6;14), t(11;14), t(14;16) and t(14;20), have been found to correlate with increased risk of disease progression from MGUS to SMM. *Cyclin D1*, *FGFR3,* and *MYC* overexpression have also been detected in MGUS and SMM patients(35). Despite these overlaps, differing initiating events promote heterogenous MM with varying molecular subtypes.

MM prognosis and staging described by the revised international staging system in the International Myeloma Working Group is primarily determined by biological markers such as tumor burden (% of plasma cells in the bone marrow), high-risk cytogenetic abnormalities(t(4;14), t(14;16), t(14;20), del(17/17p), elevated serum lactate dehydrogenase, albumin, and beta-2-microglobulin(30). Over time, patients'

outcomes have dramatically improved due to the introduction of autologous stem cell transplant (ASCT) and new treatments such as proteasome inhibitors, immunomodulators, and monoclonal antibodies. With advances in cancer management, the median survival time of patients over the last 20 years has increased from 4 to 8 years (36).

The following sections will discuss the differences in MM susceptibility and clinical endpoints that differ by race/ethnicity and the gaps in our current knowledge of MM development in black and Hispanic patients.

Racial/Ethnic Disparities in MM

Race/ethnicity is a well-known risk factor for MM. According to Surveillance Epidemiology and End Results (SEER) registries, blacks have over 2 to 3-fold increased risk MM when compared to whites and Hispanics. This disparity is thought to be partly due to a higher MGUS prevalence in black Americans compared to white and Mexican Americans(37)**.** This excessive prevalence also continues to increase with progressing age in black cases. Similarly, a study comparing 917 Ghanaians to the predominantly white residents of Olmsted County in Minnesota, revealed a 2-fold increased prevalence of MGUS in the Ghanaian study group(38).

The study was conducted to determine if shared environmental and socioeconomic factors in black Americans contributed to the excessive MGUS prevalence. However, the common genetic ancestry between Ghanaians and black Americans, but differing environmental conditions, support the hypothesis that genetics contribute to the race-related disparity in MM susceptibility in individuals with African ancestry. In contrast, a population-based MM incidence study of Afro-Caribbeans from Curaçaoa in 1993 showed an incidence rate of 3.1 to 100,000 persons(39), considerably lower than the US black incidence rate of 9.5 to 100,000 persons in the same year(40).

Although, race/ethnicity is well known risk factor, environmental and nutritional factors cannot be ruled out. A retrospective study on Afro-Caribbean patients from New York presented a substantial higher incidence rate in females with high BMI than males. It is also worth noting that the rate of MM progression between in those of European and African descendants is constant(41), displaying that the higher prevalence of MGUS does not translate to faster MM progression.

There are notable differences in the patterns of driver genetic abnormalities by race/ethnicity, such as lower rates of IgH translocations in blacks(42–44) compared to whites. Similarly, *TP53* mutations appear more frequently in individuals of European ancestry(45) than those of African ancestry. Also, genome-wide studies have yet to reveal susceptibility loci that uniquely associated with African ancestry and MM risk(18, 19), adding to the gap in knowledge regarding the biological mechanism causing the disproportionately high MM incidence rate in individuals of African descent.

When considering MM outcomes, the death rate in blacks is 2-fold higher than that of whites. This is due, in part, to the excessive MM incidence in black individuals and subsequently, a higher mortality rate. However, a large population-based study, using SEER registries of over 5,700 black and 28,000 white MM cases from 1973-2005, showed a slightly favorable survival in blacks in years 1973-1993, with improved prognosis in whites between 1994-2005, around the time novel MM treatments were introduced clinically(46). The improvements in survival after 1994 were not as significant in blacks as it was in whites, suggesting that treatment access may have affected MM outcomes in black and white cases disproportionately. Moreover, a recent population-based study found that black patients survive longer than whites if both groups have access to similar treatments and autologous stem cell transplant (ASCT)(47). Additionally, black MGUS cases are diagnosed with fewer high-risk IgM MGUS than white MGUS patients(48). This and the lower occurrence rate of mutations like *TP53* (associated with adverse prognosis) in those of African descent suggest a possible biological influence of slightly favorable outcomes in black patients. However, after adjusting for other prognostic covariates, the underlying cause of the minor improved survival in MM patients with African ancestry is unknown.

There is a well-founded highlight on African descendants when studying disparities in MM. However, less emphasis has been given to the fast-growing Hispanic population in the United States. Therefore, a significant portion of this thesis and the next section will focus on MM in the understudied Hispanic population.

MM in Hispanics

According to the Texas Tumor Registry, there are more Hispanics diagnosed with MM than non-Hispanic blacks in the last decade, even though Hispanics reported over half of the age-adjusted incidence rate compared to non-Hispanic black Texans **(Table 1)**. Nonetheless, this fastest-growing minority group in Texas and the United States is overlooked in MM etiology and outcomes research. A PubMed search for “Hispanic AND myeloma” will yield only three SEER based studies on MM outcomes, underscoring the importance of investigating myeloma in this understudied population.

Interestingly, Hispanics are diagnosed at a younger age (65 years), compared to blacks (66 years) and whites (71 years)(47), for unclear reasons.

Other than the early onset of disease, little is known regarding MM development in this group. Therefore, research of MM etiology and clinical endpoints in Hispanics is vital to fill this knowledge gap.

Table 1. Texas Cancer Registry: New MM Cases, 2006 - 2016

	Population at Risk	Cases	Age-adjusted Rate
Hispanic	106,516,575	**3,843**	6.3
Non-Hispanic black	33,616,305	**3,579**	14.3

Tumor Cancer Registry reports the number of MM cases in the Hispanic and non-Hispanic black Texans, adjusted for age

While clinical characteristics and cytogenetic abnormalities dictate outcomes for MM, data on genetic and clinical profiles of disease in Hispanics is limited in the literature. To date there exists no peer-reviewed study characterizing the clinical phenotype of MM in Hispanics. One abstract presented at the American Society of Hematology meeting in 2017 described disease presentation in 100 US Hispanic patients at diagnosis(49), reporting variation in clinical phenotypes, such as younger median age of diagnosis, favorable hemoglobin and creatinine levels, and lower occurrence of t(4:14) and monosomy 13 mutations when compared to white cases. This study provides evidence of distinct MM clinical features in Hispanics that require further investigation (**Figure 1**).

When investigating the survival trends in Hispanics, the 1992-2007 SEER registries reported worse disease specific survival in Hispanics (2.7 years) than white (3.6 years), black (3.8 years), and Asian (4.1 years) patients(50). However, differences in prognosis by ethnicity has narrowed in recent years(51). A study based on a SEER-Medicare dataset reported a comparable and even elevated disease-specific survival in Hispanics (5.4 years) compared to whites (4.5 years), pointing that Hispanics' previous adverse survival may be attributed to external factors such as treatment access. Indeed, treatment use and ASCT have increased among all ethnicities over time, but this increase has been more pronounced among white patients than black and Hispanic patients(47, 52).

From lack of data in the literature, it is unclear how the mentioned differences in clinical phenotypes, overall survival, and early disease onset in Hispanics tie into the racial/ethnic disparity in MM risk and outcomes. To address the tremendous dearth of knowledge regarding MM in Hispanics, this proposal is designed to investigate the influence of genetic ancestry on MM risk and survival utilizing the robust clinical data

from MD Anderson's diverse patient population. This study will be among the first to shed light on the genetic and clinical factors affecting of MM in the Hispanic population and how these factors also influence the differential risk of incidence and outcomes by race/ethnicity.

Characteristic	Hispanic Cohort	Kyle et al, 2003	p value
Gender			
Males (%)	50	59	**0.01**
Females (%)	50	41	
Median age (years)	58	66	**<0.001**
Median BM Plasmacytosis (%)	40	50	**0.003**
Median Cr (mg/dL)	1	1.2	**0.028**
Median Hb (g/dL)	11.4	10.9	**0.02**
Light Chain Subtype			
κ	67%	57%	**<0.001**
λ	31%	34%	
Other	2%	9%	
Lytic Bone Disease at Diagnosis			
Present	71%	66%	0.14
Absent	29%	34%	
Median M Spike at Diagnosis (g/dL)	2.25	0.48	**<0.001**

Characteristic	Hispanic Cohort	Fonseca et al, 2003	p value
Gender			
Males	50	62	**0.003**
Females	50	38	
Median age (years)	58	63	**<0.001**
Median BM Plasmacytosis	40	43	0.57
Median Cr (mg/dL)	1	1.2	**0.028**
Median Hb (g/dL)	11.4	10.7	**<0.001**
Light Chain Subtype			
κ	67%	63%	0.17
λ	31%	33%	
Other	2%	3%	
Lytic Bone Disease at Diagnosis			
Present	71%	61%	**0.013**
Absent	29%	39%	
Cytogenetics			
del17p	9.8%	10.5%	0.87
t(14;16)	3.3%	4.3%	0.80
t(4;14)	5.9%	12%	**0.038**
t(11;14)	11.1%	15.1%	0.26
del13q/monosomy	31.4%	50.1%	**<0.001**

Figure 1. Comparison of Baseline Clinical Characteristics Between Hispanic and White MM Cases

An abstract during the American Society of Hematology meeting presented by Jain and colleagues describe the clinical features of Hispanic patients at diagnosis compared to white patients' baseline clinical characteristics collected by Kyle et al., 2003 and Fonseca et al., 2003. (*Tania Jain, Rafael Fonseca, Ruqin Chen, Raj Patel, Prachi Jani, Veronica Gonzalez De La Calle, Zahara Meghji, James E. Hoffman, Alvaro J. Alencar, Kevin R Kelly, Vivek Roy, Taimur Sher, Asher A. Chanan-Khan, Sikander Ailawadhi; Racial Differences in Disease Characteristics: Understanding Multiple Myeloma in Hispanics.* Blood *2017; 130 (Supplement 1): 864). I have been granted permission by the American Society of Hematology, provided by the Copyright Clearance Center, to re-publish the above figure, for the purpose of my thesis. License ID: 1067178-1*

Study Objective and Approach

This thesis aims to identify links between genetic ancestry and MM susceptibility and outcomes in a diverse study population. Previous studies have identified genetic variations mediating risk, but the full spectrum of genetic factors remain unclear for this complex disease. The striking racial/ethnic disparity in MM susceptibility further alludes to the presence of genetic factors driving these differences. To begin understanding the roles of common germline variants and MM risk, we previously conducted a candidate pathway analysis focusing on variants within the Wnt/beta-catenin pathway. The rationale of this analysis was to assess the impact of genetic variants within pathways previously associated with MM risk. We will build on this discovery association study to investigate the genetic etiology of MM and how that may vary by race/ethnicity.

Moreover, studies have begun to interrogate the genetic basis for differences in survival and other clinical endpoints. However, these studies have been in populations of European descent and thus cannot adequately assess how genetic factors influence MM's outcomes by race/ethnicity. Furthermore, there is a gap in our understanding of the genetic etiology of MM and disease characteristics in Hispanics**.** Together, this underscores a great need for investigation into the genetic mediators of susceptibility and clinical outcomes of MM in multi-ethnic populations. By leveraging the robust medical records of the diverse patient population at MD Anderson Cancer Center, this proposal is designed to investigate the genetic influence on MM risk and outcomes of multi-ethnic subjects, with a special emphasis on Hispanics.

This study may provide a novel understanding of the genetic and clinical MM characteristics in the Hispanic patient population, while also identifying genetic

contributors to racial/ethnic disparities in MM susceptibility and outcomes. Therefore, we tested the central hypothesis that genetic ancestry differentially mediates MM susceptibility and outcomes in populations with varying ancestral backgrounds. Towards this, our approach is as follows:

1. To Identify and confirm MM risk variants within the Wnt/beta-catenin pathway and ascertain if associations vary by race/ethnicity by building on a previous discovery analysis. In a prior study, we conducted a discovery candidate pathway analysis to identify variants associated with MM risk using a patient population from the MD Anderson Cancer Patient and Survivors Cohort (MDA-CPSC), seven variants associated with MM risk in a non-Hispanic white study population were identified.

To replicate the findings and to also establish if these candidate variants differ by race/ethnicity, we conducted a cross-ethnic replication analysis on an additional 731 self-identified non-Hispanic black, Hispanic, and non-Hispanic white MM cases, including 788 race/ethnicity matched controls. Genotyping was performed on the OncoArray platform that includes over 400,000 fixed genetic markers with a GWAS backbone for imputation. We also validated the significant association of the top candidate variants in a non-Hispanic white case-control dataset in collaboration with the Myeloma Working Group of the InterLymph Epidemiology Consortium. We then utilized *in-silico* informatics tools to evaluate the functional significance of variants.

2. To examine the effect of European, African, and Amerindian (Indigenous American) genetic ancestry on MM susceptibility by conducting a case-control association analysis in 143 self-identified Hispanic, 211 non-Hispanic black, and 262 non-Hispanic white MM cases and 633 healthy controls. The inferred genetic ancestry of each study

individual was calculated from a genome-wide genotyping dataset. Then, we analyzed the effect of each genetic ancestry on MM risk in the overall study population and also among self-identified Hispanic individuals.

3. To characterize MM clinical phenotypes in Hispanics and analyze the differences in clinical profiles between race/ethnicities by utilizing the robust electronic medical records (EMR) of MD Anderson Cancer Center. From the EMR we abstracted extensive clinical and follow-up information including patient demographics, history of pre-malignancy, and clinical phenotypes such as subtypes, biomarkers, and diagnostic criteria. Extracted characteristics also included baseline cytogenetic/FISH/karyotype results for somatic mutations in high-risk patients and prognostic indicators such as beta-2-microglobulin (B2M), lactate dehydrogenase (LDH), and albumin levels.

Also included in the abstraction were MM treatment regimens, treatment cycles, dates of response and relapses, death, and last date of follow up. We then conducted a cross-ethnic comparative analysis of abstracted clinical features and survival between self-identified Hispanics and non-Hispanic whites and non-Hispanic black patients.

Furthermore**,** to investigate if differences in clinical phenotypes also vary by genetic ancestry**,** we analyzed the phenotypes by genetic ancestry. Such phenotypes included diagnostic and prognostic blood biomarkers, prior history of pre-malignancy and high-risk cytogenetic mutations (t(4;14), del(17/17p), t(14;16), t(14;20)). We also compared overall survival of patients, adjusting for appropriate prognostic covariates.

This study will take steps towards improving MM risk assessment and cancer management in patients of diverse backgrounds in two ways: (1) filling the knowledge gap regarding the genetic contributors of MM susceptibility and outcomes in Hispanic

patients and (2) utilizing genetic ancestry instead of a self-reported ethnicity to better define risk association in patients.

Chapter 2: Genetic Variants Within the Wnt/beta-catenin Pathway Associated with MM Risk Vary by Race/Ethnicity

This chapter includes data from **Belachew, A. A.,** Wu X., Callender, R. A., Waller R., Orlowski, R. Z., Vachon, C. M., Camp, N. J., Zid, E., Hildebrandt, M. A. T. Genetic variants in the Wnt/beta-catenin signaling pathway as determinants of multiple myeloma risk (submitted). All contributions are from work performed by Belachew A. A.

Introduction and Study Objective

Evidence suggests that the Wnt/beta-catenin pathway is linked to MM susceptibility(53). Additionally, genes within the canonical Wnt pathway have been shown to exhibit changes in expression in the bone micro-environment(54) leading to MM progression (55, 56). Furthermore, the conserved Wnt/beta-catenin pathway, a key player of cellular homeostatic actions(57), is also associated with self-renewal of cancer stem-like cells(58, 59). Hence, Wnt/beta-catenin pathway dysregulation is reported in the tumors of common cancers(60) and downstream effects of this pathway have been intensely studied as potential therapeutic targets(61, 62).

Figure 2 illustrates the canonical Wnt pathway. During the "off" state of the canonical Wnt pathway, the beta-catenin destruction complex composed of ademomatosis polyposis coli (APC)/axin/glycogen synthase kinase-3 beta (GSK-3β) /casein kinase 1(CK1), phosphorylates beta-catenin leading to its ubiquitination and proteasomal degradation after binding to beta transducin repeat containing E3 ubiquitin protein ligase (β-TrCP). This action prevents beta-catenin from traveling from the cytosol to the nucleus, thus activating Wnt-targeted transcription factors.

During the "on" state of this pathway, secreted glycoprotein Wnt ligand binds to the transmembrane G-coupled protein receptor Frizzled (FZD), along with its co-receptor transmembrane low-density lipoprotein receptor-related Protein 5/6 (LRP 5/6), recruiting intracellular protein disheveled (DVL) to the cell membrane and consequently disrupting the beta-catenin destruction complex. This releases beta-catenin to translocate from the cytosol to the nucleus where it trans-activates transcription factors, such as T-cell factor/lymphoid enhancer factor (TCF/LEF), for the transcription of Wnt

target genes responsible for cellular proliferation, polarity, survival, and cell differentiation(63–65).

Germline genetic studies have not yet ascertained the inherited genetic risk of MM development conferred by this pathway. Therefore, we aimed to identify the genetic mediators of MM susceptibility within the Wnt/beta-catenin pathway.
To elucidate these genetic contributors, we performed discovery genotyping non-Hispanic white MM case and control subjects from MD Anderson Cancer Center using variants identified from 26 core genes within the Wnt/beta-catenin pathway. We then verified our findings on an additional population of non-Hispanic white replication cases and controls from the same institute. External validation of replicated findings was conducted using existing genotyping data from the University of Utah and the University of California-San Francisco through the Myeloma Working Group (MWG) of the InterLymph Consortium(66). Given the evidence for racial/ethnic disparities in MM susceptibility, we further examined the association of these genetic variations with MM among non-Hispanic blacks and Hispanics.

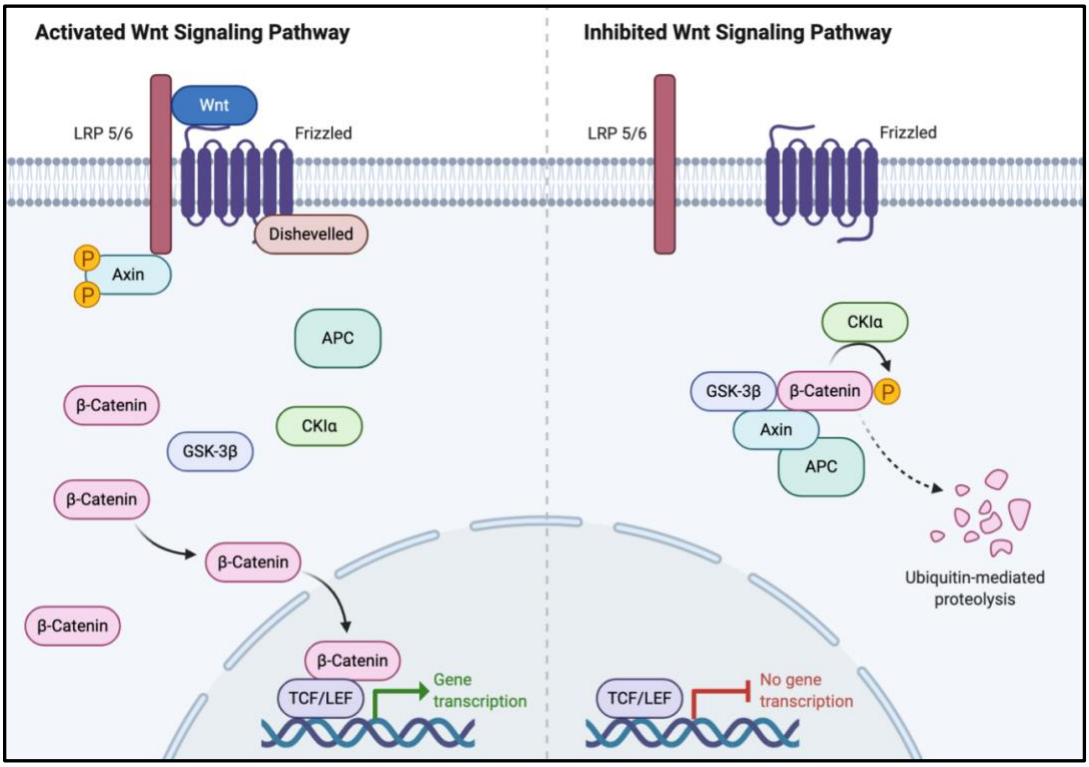

Figure 2. Overview of the Canonical Wnt Signaling Pathway

Canonical Wnt signaling pathway activation signaling cascade after the Wnt ligand binds to FZD and LRP5/6 co-receptors (left) and inactivation in absence of Wnt ligand with continual degradation of beta-catenin through ubiquitin-mediated proteolysis (right). (*Created with BioRender.com)*

Study Design and Methods

Discovery Phase

The overall design of the study is shown in **Figure 3**. For the discovery phase, 269 self-reported non-Hispanic white patients diagnosed with MM were identified from the MD Anderson Cancer Patient and Survivors Cohort (MDA-CPSC)(67), a hospital-based cancer patient cohort at The University of Texas MD Anderson Cancer Center. A total of 272 healthy non-Hispanic white control subjects, with no prior history of cancer, were recruited from Kelsey-Seybold Clinics(68). Cases and controls were matched by age (± 5 years) and sex (**Table 2**). Each subject provided peripheral blood as source of genomic DNA for genotyping conducted on a custom Illumina BeadXpress chip (San Diego, CA), which included 171 variants from 26 core genes of the Wnt/beta-catenin pathway identified from literature search and KEGG(69). Tagging variants ($r^2 > 0.8$) from a 10 kb flanking region upstream and downstream within each core gene from the CEU HapMap population with a minor allele frequency > 5% were identified using Tagger(70). Written informed consent was provided by each patient and the study was approved by the Institutional Review Board of MD Anderson.

Internal Replication Phase

We selected an additional 292 non-Hispanic white MDA-CPSC MM cases from the MDA-CSPC and 331 healthy non-Hispanic white controls(68) matched by age and sex (**Table 2**) for internal replication of the seven variants identified in the discovery phase. Genotyping was performed on the genome-wide Illumina OncoArray followed by imputation to the 1000 Genome Project(71) using the Michigan Imputation Server(72).

Candidate variants were extracted from the dataset using PLINK(73) for replication analysis.

External validation phase

For the external validation phase, the two candidate variants from the internal replication analysis were extracted from existing genome wide association study (GWAS) data generated from 526 non-Hispanic white patients with MM and 878 non-Hispanic white healthy control subjects (**Table 2**) from the University of Utah and University of California-San Francisco(66). Imputation of the external validation phase was performed using the Michigan Imputation Server(72) to the 1000 Genome Project(71). Written informed consent was provided by each patient and the study was approved by the respective Institutional Review Boards.

Cross-Ethnic Internal Replication Phase

Cross-ethnic internal replication of the seven variants from the discovery phase was conducted on self-reported 172 Hispanic and 267 non-Hispanic black MM cases from the MDA-CPSC. Control subjects (180 Hispanic and 277 non-Hispanic black) were selected from Kelsey-Seybold Clinics(68). Genotyping of the 49 Hispanic and 91 non-Hispanic black MM cases, as well as 48 Hispanic and 90 non-Hispanic black control subjects, was performed using the Illumina BeadXpress genotyping chip (San Diego, CA). Genotyping of the remaining MM case/control samples was conducted on the Illumina OncoArray platform. The genotyping data from both platforms were combined and analyzed together. Written informed consent was provided by each patient and the study was approved by the Institutional Review Board of MD Anderson.

Statistical Analysis

For genotype quality control, variants deviating from the Hardy-Weinberg equilibrium in controls and those with call rates < 95% or minor allele frequency (MAF) < 0.01 were omitted from the analysis. Risk of MM for each variant was estimated using odds ratios (ORs) with 95% confidence intervals (95% CI) in multivariable logistic regression adjusting for age and sex. For each variant, analysis was conducted under the dominant, additive, or recessive model of inheritance (MOI) with the model with the lowest P-value reported. Variants associated with MM risk with $P < 0.05$ during the discovery phase underwent bootstrap resampling of 1000 iteration to prioritize candidate variant selection. Variants consistently associated with MM risk with $P < 0.05$ for 80% of the bootstraps were deemed candidates for replication. We also performed a meta-analysis for combined (fixed) effects of discovery-replication (same ethnicity)-validation, as well as discovery-cross ethnic replication study groups. Statistical analysis was conducted using STATA 14 (Stata, College Station, TX).

In-Silico Functional Prediction

The location of variants in the genome was visualized using the UCSC genome browser(74). The regulatory and functional effects of genotyped variants and their proxies ($r^2 > 0.8$) were determined using Haploreg4.1(75), Regulomedb(76), and LDLink(77) by annotating transcription regulators, as well as enhancer and promoter elements in lymphoblastoid cell lines. We also used the open-access expression trait loci (eQTL) browser (https://genenetwork.nl/bloodeqtlbrowser/) to identify eQTL of candidate variants correlating with gene expression.

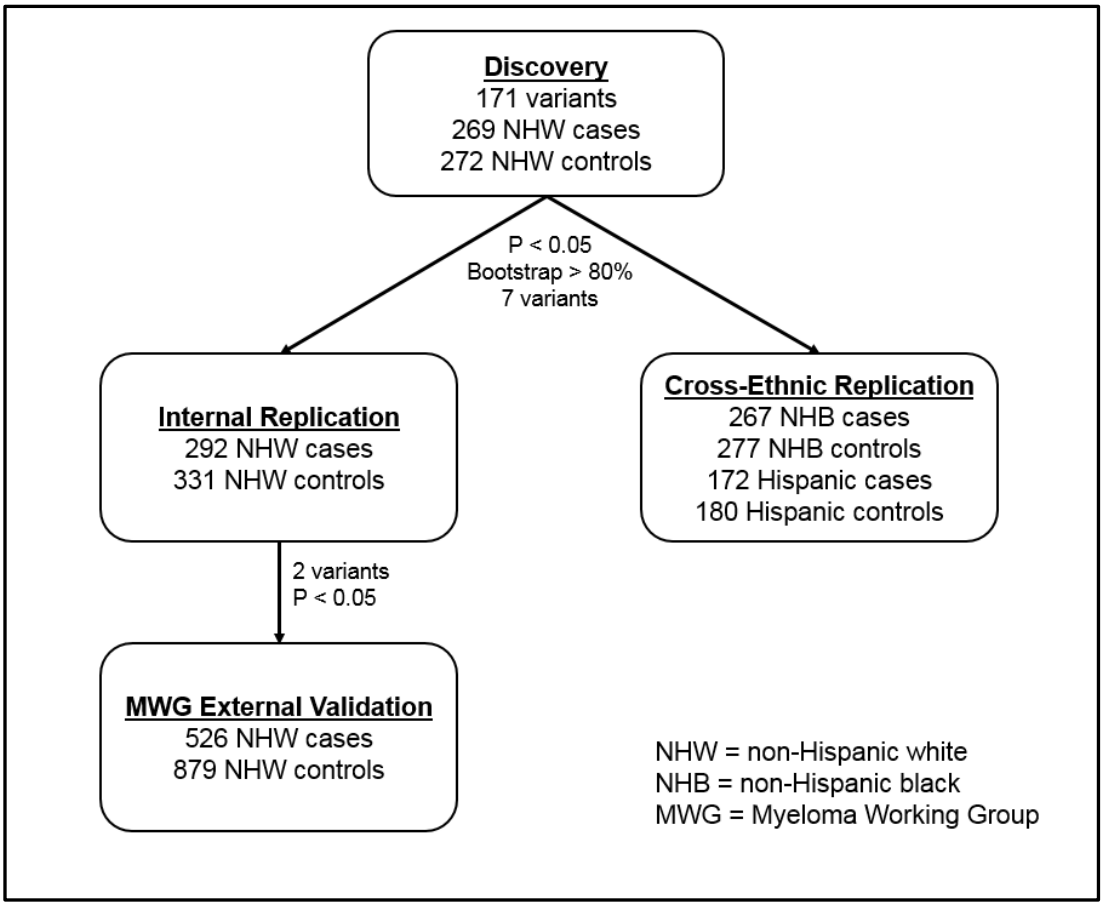

Figure 3. Study Design

The flow chart shows the study design of the discovery, internal replication, and Myeloma Working Group (MWG) external validation in non-Hispanic white (NHW) study populations, along with the cross-ethnic internal replication phase in non-Hispanic black (NHB) and Hispanic populations.

Results

Study Population

Characteristics of each study population are shown in **Table 2**. Males comprised a slight majority of the discovery cases (59.1%) and controls (59.6%), with higher representation in controls from the replication group (68.9%). The mean age of both MM cases and controls in the discovery study group was 60.8 years, and slightly older for the internal validation (62.4 years) and controls (61.6 years). Patients included in the MWG dataset had a median age of 60.0 for cases and 63.7 years for controls. The 267 non-Hispanic black and 172 Hispanic patients with MM had a median age of 57.0 years, slightly younger than the non-Hispanic white patients.

Table 2. Study Populations

	Discovery		Internal Replication		MWG External Validation		Cross-Ethnic Internal Replication			
Population	NHW		NHW		NHW		NHB		Hispanic	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
Total (N%)	269	272	292	331	526	878	267	277	172	180
Gender										
Male	159 (59.1)	162 (59.6)	162 (55.5)	228 (68.9)	333 (63.3)	498 (56.7)	131 (49.1)	123 (44.4)	96 (55.8)	95 (52.8)
Female	110 (40.9)	110 (40.4)	130 (44.5)	103 (31.1)	193 (36.7)	380 (43.3)	136 (50.5)	154 (55.6)	76 (44.2)	85 (47.2)
Age, mean (SD)	60.9 (11.1)	60.83 (10.7)	62.4 (9.7)	61.6 (8.43)	57.8 (10.2)	63.74 (10.7)	57.13 (10.9)	58.89 (9.02)	57.5 (9.65)	59.2 (9.38)

NHW = Non-Hispanic White
NHB = Non-Hispanic Black
MWG = Myeloma Working Group
SD = Standard Deviation

Variants Associated with MM Risk in Non-Hispanic Whites

Of the 171 variants analyzed in the discovery phase, seven were associated with MM risk with P value < 0.05 for over 80% of the bootstrap re-samplings (**Table 3**). These variants were deemed as candidate variants for internal replication in additional cases and controls from MD Anderson (**Figure 3**). Two of the candidate variants in *LRP6*, rs7966410 (OR: 0.57; 95% CI: 0.38-0.88; $P = 9.90 \times 10^{-3}$) and rs7956971 (OR: 0.64; 95% CI: 0.44-0.95; $P = 0.027$) were also associated with reduced MM risk (rs7966410 – OR: 0.65; 95% CI: 0.44-0.97; $P = 0.036$; rs7956971 – OR: 0.69; 95% CI: 0.48-0.99; $P = 0.049$) in the internal replication phase. Likewise, these results externally validated in the MWG dataset (*LRP6*:rs7966410 – OR: 0.57; 95% CI: 0.43-0.76; $P = 1.01 \times 10^{-4}$; *LRP6*:rs7956971 – OR: 0.60; 95% CI: 0.45-0.79; $P = 3.22 \times 10^{-4}$). Meta-analysis across the three phases of this study for these two *LRP6* variants demonstrated 42% and 37% reductions in risk of MM for rs7966410 and rs7956971, respectively (**Figure 4**).

Cross-Ethnic Comparisons of Candidate Variants Associated with MM Risk

Of the seven candidate variants identified in the discovery phase (**Table 3**), two variants (*CSNK1D*:rs9901910 and *BTRC:*rs7916830) replicated when genotyped in our cross-ethnic internal replication of non-Hispanic black and Hispanic populations. Similar to the discovery findings, *CSNK1D*:rs9901910 was associated with > 6-fold increased MM risk in non-Hispanic blacks (OR: 6.42; 95% CI: 2.47-16.7; $P = 3.14 \times 10^{-4}$) and over 4-fold increased risk of MM in Hispanics (OR: 4.31; 95% CI: 1.83-10.1; $P = 8.10 \times 10^{-4}$) (**Table 3**). In addition, *BTRC:*rs7916830 conferred a 24% reduction in risk in the non-Hispanic black population (OR: 0.76; 95% CI: 0.60-0.97; $P = 0.028$) that was similar in

effect to the reduced risk observed in the non-Hispanic white discovery phase. Although not statistically significant ($P = 0.74$), *BTRC:*rs7916830 was associated with a 5% reduced risk in the Hispanic population. Consequently, the remaining of the seven variants did not replicate in our cross-ethnic study population at $P < 0.05$.

Table3. Variants Associated with MM Risk in the Discovery and Cross-Ethnic Internal Replication Set

Gene	Variant	Risk Allele	*MOI	Discovery †OR (95% CI)	Boot-strap	P-value	NHB †OR (95% CI)	P-value	Hispanic †OR (95% CI)	P-value
CSNK1D	rs9901910	C	dom	2.40 (1.67-3.45)	1000	2.43×10^{-6}	**6.42 (2.47-16.7)**	**3.14×10^{-4}**	**4.31 (1.83-10.1)**	**8.10×10^{-4}**
BTRC	rs7916830	C	add	0.63 (0.49-0.82)	999	5.60×10^{-4}	**0.76 (0.60-0.97)**	**0.028**	0.95 (0.71-1.28)	0.74
LRP6	rs7966410	C	dom	0.57 (0.38-0.88)	998	9.90×10^{-3}	0.83 (0.57-1.21)	0.39	0.79 (0.46-1.36)	0.40
LRP6	rs7956971	T	dom	0.64 (0.44-0.95)	881	0.027	1.03 (0.73-1.46)	0.70	0.82 (0.50-1.35)	0.44
APC	rs2431507	G	dom	0.45 (0.27-0.76)	1000	2.77×10^{-4}	1.73 (0.97-3.07)	0.62	0.91 (0.45-1.82)	0.80
BTRC	rs10450405	C	dom	0.52 (0.36-0.75)	1000	4.20×10^{-4}	0.87 (0.56-1.34)	0.53	1.38 (0.90-2.10)	0.13
BTRC	rs7901883	A	dom	0.61 (0.43-0.87)	1000	6.10×10^{-3}	0.95 (0.61-1.47)	0.84	1.36 (0.89-2.08)	0.15

*Mode of inheritance: dominant (dom) or additive (add).
†OR = odds ratio; CI = confidence interval; adjusted for age and sex.
NHB = Non-Hispanic black

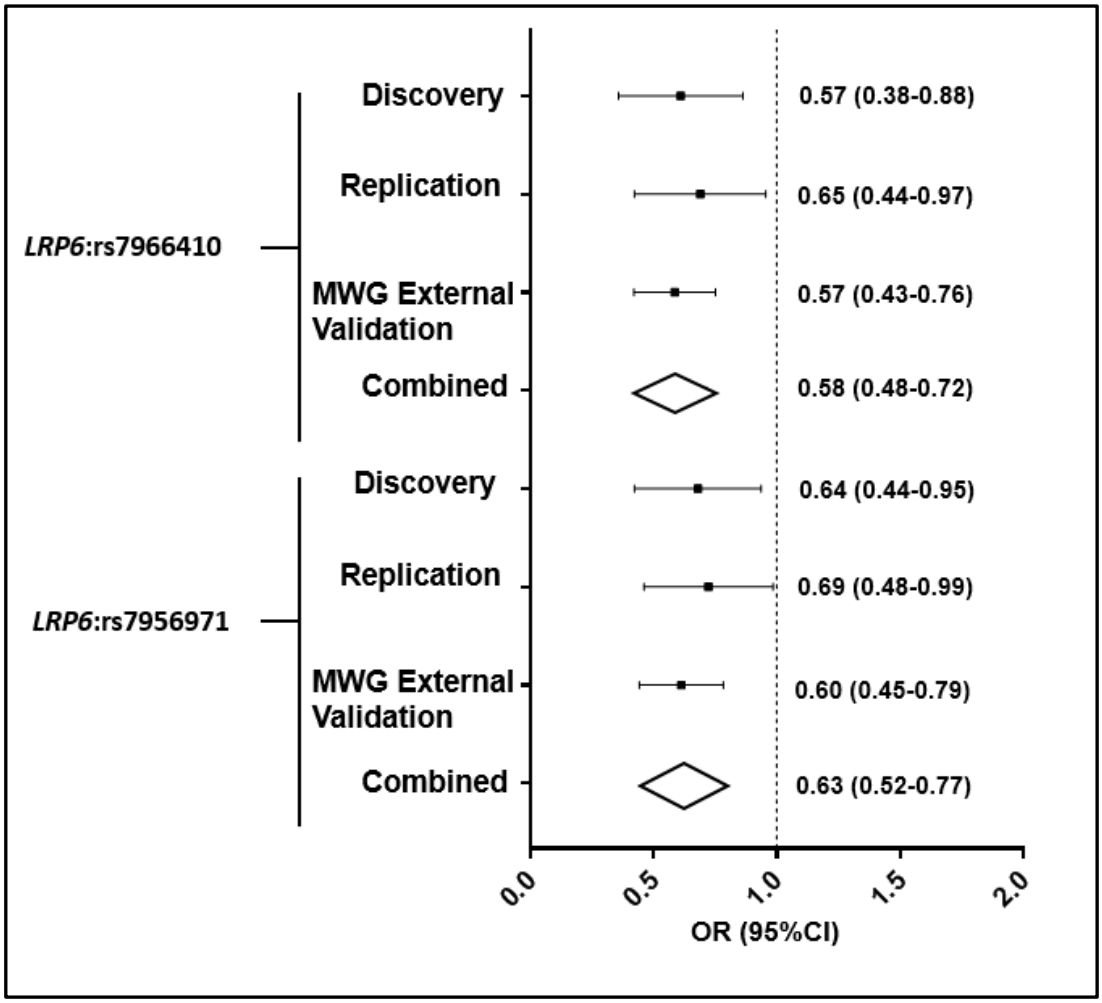

Figure 4. Genetic Variants in LRP6 Associated with Multiple Myeloma Risk in Non-Hispanic Whites

The forest plot shows the estimated odds ratios (OR) with 95% confidence intervals (CI) for *LRP6*:rs7966410 and *LRP6*:rs7956971 risk variants in the non-Hispanic white population during the discovery, internal replication, and external validation, as well as the combined effect across all three phases.

Predicted Biological Function of Candidate Variants Associated with MM Risk

Our analysis identified *LRP6*:rs7966410 as being associated with reduced MM risk in non-Hispanic whites. The functional consequence of this variant is unclear in the hematopoietic lineage based on *in silico* prediction analysis. However, we identified potential causal variants in high LD to the genotyped variant, including rs17819999 (r^2 = 0.97) located in a strong enhancer region of lymphoblastoid cell lines. Other variants in high LD ($r^2 \geq 0.97$), rs11054721, rs2417085, and rs10845496, were located in regions linked to tissue-specific epigenetic changes but were not shown to be linked to the hematopoietic cell lineage.

The other MM susceptibility variant in *LRP6*, rs7956971, was also not predicted to be functional. We identified 13 variants in high LD ($r^2 \geq 0.80$) that were predicted to have functional effects within lymphoblastoid cell lines and other hematopoietic cell lineages. For instance, rs12823243 (r^2 = 0.98), resides in a predicted binding site on *LRP6* for several transcription factors, including IKAROS family zinc finger 1 (*IKZF1*), B-cell lymphoma/leukemia 11A (*BCL11A*), Spi-1 proto-oncogene (*SPI1*), and interferon regulatory factor 4 (*IRF4*), all of which play crucial role in the development of the hematologic lineage. rs7302808 (r^2 = 0.83) is located 285 base pairs 5' upstream of *LRP6* and located within the transcription start site activator region for over 53 tissues, including those of hematologic lineage. Two variants in high LD, rs11054744 (r^2 = 0.93) and rs12366664 (r^2 = 0.97) were also located in regions of *LRP6* associated with weak enhancer histone activity via methylation. Additional variants rs1819871, rs11054731, and, rs4763785 (r^2 = 0.98) were located in transcription-factor binding sites associated with enhancer activity in lymphoblastoid cells lines.

CSNK1D encodes for the delta isoform of the casein kinase 1 involved in maintaining the "off" signal of the pathway. rs9901910 is intronic and located within an enhancer, as well as a genomic region linked to strong histone promoter/enhancer markers via methylation and acetylation activity in lymphoblastoid cell lines. A cis-eQTL for this variant was also reported for Dicarbonyl and L-xylulose Reductase (*DCXR*) in whole blood. This gene is located about 202 kb upstream of *CSNK1D*:rs9901910. A variant with high LD ($r^2 = 1$) with rs9901910, rs4789846, was also predicted to serve as a cis-eQTL for Coiled-Coil Domain Containing 57(*CCDC57*) in lymphoblastoid cells. rs7916830 is located 2.7 kb upstream of the beta-transducin repeat containing E3 ubiquitin protein ligase (*BTRC*) and is linked to gene regulation via polycomb gene repression in lymphoblastoid cell lines.

Discussion

This study identified genetic variation within the Wnt/beta-catenin pathway as contributors to MM susceptibility and explored potential racial/ethnic differences in this risk. We discovered and validated rs7966410 and rs7956971 in *LRP6* associated with reduced MM risk in non-Hispanic white study subjects. We also identified *CSNK1D*:rs9901910 to be associated with a 2- to 6-fold increased MM risk among all three racial/ethnic populations, further clarifying the underlying genetic contributors of MM susceptibility within the Wnt/beta-catenin pathway. An additional candidate variant, *BTRC:*rs7916830, was replicated in the non-Hispanic black population only, suggesting the variability in genetic etiology of MM risk by ancestry.

The validated intronic variants, rs7966410 and rs7956971, are located in the gene for LRP6 encoding the low-density lipoprotein receptor-related protein-6, a transmembrane Wnt binding co-receptor. LRP5/6 in conjunction with the fizzled co-receptor form the signaling complex with Wnt ligands to activate downstream signaling for beta-catenin stabilization and trans-activating Wnt target genes. Specifically related to MM, inhibition of the LRP5/6 co-receptor is reported to reduce tumor burden in MM mouse models(78). Additionally, molecular studies have demonstrated the LRP6 co-receptor to play a direct role in Wnt inhibition activity by sequestering DKK1, an antagonist of the Wnt ligand, leading to the downregulation of the canonical Wnt signaling pathway(79, 80). This event is shown to disrupt the osteoclast/osteoblast homeostasis in the bone marrow, leading to bone destruction in MM patients(81). Some studies proposed anti-DKK1 antibody as a therapeutic agent to improve bone disease(82, 83), one of the four diagnostic criteria of MM(84), pointing *LRP6* to be instrumental in MM risk as a possible regulator of bone homeostasis and bone disease.

Although there were no predicted functional consequences of rs7966410 and rs7956971 in the hematopoietic cell lineage, one possible causal variant, rs12823243, in high LD to rs7956971 pointed to functional effect through a transcription regulatory mechanism involving MYC. Rs12823243 resides on a predicted transcription factor binding site within *LRP6* in lymphoblastoid cell lines for the transcription factors IKZF1 and IRF4, both of which have been proven to play critical roles in MM progression. *IKZF1,* encoding the lymphoid transcription factor IKAROS(85), is normally activated during early lymphocyte differentiation and is a frequently mutated tumor suppressor gene in hematologic malignancies(86, 87). Downregulation of *IKZF1* is shown to lead to the downregulation of the IKZF1 target genes, *IRF4* and *MYC*. Reduced expression of the transcription factor IRF4 is also known to reduce MM cell viability, possibly through the downregulation of MYC(88). Moreover, the MM chemotherapeutic drug lenalidomide has been shown to selectively degrade the IKZF1 transcription factor via E3 ubiquitin ligase activity(85) and induce cell toxicity through reduced expression of IRF4(89). It is important to point out that MYC mutations are reported in 15-20% of MM diagnoses(90), suggesting a potential regulatory mechanism between the causal variant in *LRP6* and transcription factors IKZF1-IRF4 and MYC on MM risk.

CSNK1D:rs9901910 was associated with over 2-to-6-fold increase in MM risk across all populations. This variant is located within an enhancer region of over 22 tissues, including those of hematologic lineage. *CSNK1D* encodes a monomeric serine/threonine kinase and interacts with dishevelled (DVL) within the beta-catenin destructive complex to regulate beta-catenin abundance in the cytoplasm(91). Although *CSNK1D* has not been studied in MM, other members of the highly conserved casein kinase family, CK2 and CK1α, have been shown to consistently sustain activation of well-known oncogenic signaling cascades, PI3K/AKT, JAK/STAT, and NF-κB, in MM

cell lines and mouse models(92, 93). Furthermore, loss of function or inhibition of CK2 and CK1α has led to apoptosis and reduced MM cell survival as a result of impaired phosphorylation of these oncogenic signaling cascades. Hence, it would be useful to study the enhancer/promoter activity on which rs9901910 is located to elucidate if this effect also promotes activity of the mentioned oncogenic signaling cascades in MM.

BTRC:rs7916830 was associated with MM risk in the non-Hispanic white discovery and the non-Hispanic black population. The protective effect associated with this variant is perhaps linked to polycomb gene repression that was predicted by our *in silico* analysis. Polycomb proteins that have gene silencing effect through epigenetic alterations(94) are also shown to undergo post-translational modification via ubiquitin mediated proteasome action(95). One study demonstrated the regulatory role of the E3 ubiquitin protein ligase (β-TrCP) on a critical enzymatic subunit, enhancer of zeste homolog 2 (EZH2), of the Polycom repressor complex which tri-methylates H3K27 to mediate gene repression(96). By averting recognition of β-TrCP and hence degradation, stabilization of EZH2 through gain of function was shown to enhance tri-methylation of the lysine tail of H3K27(me3) and promote B-cell lymphocyte pathogenesis. We also know that enhanced BTRC activity is positively correlated with MM progression(97), although the mode of tumorigenesis is unclear. In-vitro studies are therefore necessary to understand the tumorigenic significances of *BTRC* variants' downstream epigenetic consequences in MM development.

The strength of our study is the three-phase study design comprised of discovery, internal replication, external validation, and cross-ethnic internal replication phases. *In silico* functional prediction point to a biological inference of gene regulatory effect for the identified variants and proxies in the hematologic lineages. Nevertheless, *in vivo* studies are critical to understanding the mechanistic effects of these variants in

MM development. A further strength is the inclusion of study participants from three different institutions and analysis of genetic risk across three different racial/ethnic populations. However, the study is limited in sample size from the non-Hispanic black and Hispanic populations that may hinder the ability to form a definitive conclusion on the significance of identified variants within these subgroups. We acknowledge that the selected variants in this study are from the CEU HapMap population, which is of European descent, and may not accurately tag the underlying genetic structure in the Hispanic and the non-Hispanic black populations. Nevertheless, given the scarcity of genotype data in the Hispanic patient population this study a solid stepping stone for further research to understand genetic variation within a diverse patient population associated with MM risk.

In conclusion, this work identified candidate variants of MM for replication studies that have supporting functional consequences *in silico*. We also identified several variants associated with MM risk that vary by race/ethnicity. Previous studies show the Wnt/beta-catenin pathway as a key player in cancer progression. Our results may provide further insight into the biology of this pathway as well as its role in MM development. This study also serves as a platform for additional studies in understanding the genetic contributors of racial/ethnic disparity in MM susceptibility.

Chapter 3: Genetic Ancestry Mediates MM Susceptibility in Hispanics

Introduction and Study Objective

In Chapter 1, we highlighted the dearth of genetic and clinical studies of MM in Hispanics. We also described the well-established disparity in MM incidence by ancestral background that is yet to be studied in the Hispanic population. This chapter investigates if genetic ancestry mediates MM susceptibility in a multi-ethnic study population, emphasizing this relationship within the Hispanic population.

Admixed human populations, like our Hispanic study subjects, have a non-homogenous genetic inheritance from two or more insulated continental populations. Interbreeding between isolated parental populations create admixed generations through recombinant genetic events that allow their descendants to carry the original parental populations' chromosomal segments. In admixed individuals with a complex disease, chromosomal segments harboring the disease's susceptibility variants will show an excess of genetic ancestry from the parental population that carried the risk allele. Consequently, through admixture mapping, one can identify chromosomal regions that show an excess of ancestry from the high-risk parental population in individuals with the disease. Therefore, we will utilize admixture mapping to understand the relationship between genetic ancestry and MM risk as a complex disease.

Admixture mapping can be performed by (1) local ancestry inference — tracing the parental ancestry of an individual from a particular chromosomal location or by (2) global ancestry inference — estimating the proportion of parental populations of an individual by averaging chromosomal segments of the entire genome of that individual.

We performed global admixture mapping to demonstrate the population structure of a multi-ethnic, case-control population. We then quantified MM risk based

on the proportion of genetic ancestry within our study population. Hispanics are a heterogeneous group comprising European, African, and Amerindian ancestry, providing unique multi-level ancestral reference groups to analyze MM susceptibility by racial/ethnicity.

Study Design and Methodology

Study Population

MM cases were diagnosed between 1981-2019 and selected from the MD Anderson Cancer Patient and Survivors Cohort (MDA-CPSC)(67). Patients were self-identified as Hispanic (N = 143), non-Hispanic black (NHB; N = 211), and non-Hispanic white (NHW; N = 262). Healthy controls (N = 654) were identified from two approaches: (1) NHW and NHB controls were recruited from Kelsey-Seybold Clinics(98) and (2) Hispanic controls were selected from the Mexican American Mano a Mano Cohort(99). Hispanic control individuals were self-reported Mexican descendants who reside in the metropolitan Houston area(99). All controls were frequency matched to cases by age (± 5 years), gender, and self-reported ethnicity (**Table 4**). We also collected our Hispanic cases' geographical origin using their MD Anderson medical records (**Table 5**). Written informed consent was provided by each patient, and the study was approved by the MD Anderson Institutional Review Board.

Genotyping

All case and control subjects provided peripheral blood as a source of genomic DNA for genotyping on the genome-wide Illumina OncoArray platform. Following QC and data cleaning, imputation was conducted to the HRC(71) (mapped to GRCh37/hg19) using the Michigan Imputation Server(72).

Estimating Genetic Ancestry and Population Structure of the Study Population

ADMIXTURE(100) was used to evaluate the population structure and infer the study population's global ancestry. ADMIXTURE estimates ancestry in a model-based

manner from large autosomal SNP genotype datasets with a required pre-defined "K," for the number of assumed ancestries in the dataset. To choose the correct value of pre-defined K with the best predictive accuracy for estimating genetic ancestry, one can conduct cross-validation (CV) analysis for multiple Ks. A good value of K exhibits a low cross-validation error compared to other K values. Therefore, we performed cross-validation analysis on the OncoArray dataset for K1-K10 (CV error: K1: 0.47455, K 2: 0.44339, K3: 0.43855, K4: 0.43841, K5: 0.43771, K6: 0.43775, K7: 0.43784, K8: 0.43799, K9: 0.43873, K10: 0.43842). We found K5 to have the lowest cross-validation error, and thus chose K = 5 to run unstructured ADMIXTURE on the Linkage Disequilibrium (LD) pruned ($r^2 > 0.2$) OncoArray genotype dataset.

Next, we conducted an independent Principle Component Analysis (PCA) on the OncoArray genotype dataset to visualize and confirm the ADMIXTURE findings. Through this step, ancestral outliers of admixed individuals can be identified for removal to minimize confounding. PCA can be a population stratification method by genotyped data and cluster individuals that share the greatest genetic similarities, i.e., genetic ancestry. LD pruned ($r^2 > 0.5$) PCA was performed on the OncoArray genotype using the FlashPCA program(101). PCA and ADMIXTURE outputs were visualization using RStudio(102).

Statistical Analysis

Genotype quality control was performed by filtering call rates < 95% using PLINK. We then calculated the effect for every 10% increase of inferred genetic ancestry on MM risk using logistic regression models (adjusted for age and gender). Odds ratios (95% Cl) for MM risk by genetic ancestry for the overall population and stratified by self-identified NHW, NHB, and Hispanic ethnicity were reported. Moreover,

due to the admixed genetic heterogeneity, Hispanic subjects were further matched by their principal components. Next, the odds ratio (95% CI) of MM risk controlling for age and gender, after removing outliers identified through our PCA strategy were calculated. A p-value of < 0.05 was considered statistically significant. All statistical analysis was conducted using Stata 16 Software(103).

Results

Study Population

For the NHW group, the median age of cases was higher (61.4 years) compared to NHB (57.5 years) and Hispanic (57.4 years) cases. Similar to Hispanics, males made up most of the cases (60.6%) and controls (62.1%) in NHWs. For the NHB group, female cases represented a slightly higher number (50.3%) and controls (50.5%). **Table 4** describes the characteristics of the study population stratified by self-identified race/ethnicity.

Additionally, review of the electronic medical records (EMR) revealed that the majority of MM patients reside in Texas (83.9%), of which almost 22.5% come from Houston metropolitan area (**Table 5**). 4.9% of the patients live in New Mexico, and an additional 4.9% are residents of states other than Texas and New Mexico, including California, Tennessee, and Florida. International patients comprised only 6.3% of study cases, and a third of those came from Mexico. Together, 90.9% of self-reported Hispanic cases have Mexican heritage as many of Texan residents are of Mexican lineage(104), and the rest of the Hispanic patients reside in New Mexico and Mexico.

Table 4. Study Population with Self-Identified Ethnicity (N = 1248)

	NHW (N = 533)		NHB (N = 425)		Hispanic (N = 290)	
	Cases	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
Total	261	272	211	214	143	147
Gender						
Male (%)	158 (60.5)	169 (62.1)	105 (49.7)	106 (49.5)	81 (56.6)	87 (59.2)
Female (%)	103 (39.4)	103 (37.9)	106 (50.3)	108 (50.5)	62 (43.4)	60 (40.8)
Age, mean (SD)	61.4 (9.2)	61.0 (8.5)	57.5 (10.8)	59.9 (8.3)	57.4 (9.6)	59.6 (8.9)

NHW = Non-Hispanic white
NHB = Non-Hispanic black
SD = Standard deviation

Table 5. Self-Identified Hispanic Cases Place of Origin

Place of Origin	Number of Cases (%)	Possible Mexican Ancestry
Texas	120 (83.9)	✔
Houston	27 (22.5)	
New Mexico	7 (4.9)	✔
Los Angeles	2 (1.4)	
Florida	4 (2.8)	
Tennessee	1 (0.7)	
Mexico	3 (2.1)	✔
**Non-US, Not Mexico	6 (4.2)	
Total	**143 (100)**	**130 (90.9)**

****Non-US, Not Mexico**
Costa-Rica, N = 1
Venezuela, N =1
Honduras, N =1
Puerto-Rico, N =1
Colombia, N =1
Ecuador N = 1

Population Structure of Study Subjects

Figure 5 and **Figure 6** demonstrates the study population's genetic population structure with the ancestral distribution of cases and controls stratified by self-reported ethnicity. In **Figure 5**, the y-axis represents the fraction of ancestry for K1-K5, for each of the 1,248 individuals shown in the x-axis. The self-reported NHWs were comprised of primarily K2 (purple) and K4 (orange) and a smaller portion of K3 (yellow), highlighting the genetic diversity of Europe (North-South gradient for example) that is being recapitulated here due to the immigration patterns of Europeans to the US. Because this thesis focused on the admixed minority subjects, we did not further classify K2, K4, and K3 by their geographical origin, but instead will condense these Ks as only European ancestry. Collectively, the average inferred European ancestry (K1, K4, and K2) in NHW cases and controls was (94.1%) and (97.8%), respectively. The highest European ancestral percentage in NHWs (K2) had a mean 53.7% for the cases and 68.1% for the controls. The second common inferred European ancestry (K4) in the NHW subjects reported an average of 22.6% in the cases and 19.5% in the controls.

Furthermore, we observed that African descent (blue, K = 3) mapped primarily with the NHB study group, with an average of 80.3% inferred African ancestry in cases and 80.4% in controls. As expected, self-reported Hispanics were a three-way admixed population between European (K1, K2, K4), Amerindian (K5, green) ancestry, and a small percentage of African K3 heritage. Hispanic subjects comprised of primarily Amerindian (K5) ancestry, averaging 52.7% for controls, and 47.7% for cases. European ancestry made up an average of 46.8% in Hispanic cases and 45.6% in

controls. The average African ancestry was minimal in Hispanic cases (5.3%) and controls (1.7%).

We successfully applied ADMIXTURE to determine the study subjects' population structure as the inferred genetic ancestry mapped to self-reported race/ethnicity. As expected, our self-identified NHW individuals were primarily European, whereas NHBs were primarily of African ancestry. Hispanics were a heterogeneous group consisting primarily of European and Amerindian ancestry. We also found a small sample size of African lineage (1 - 5%), which may preclude analysis in Hispanics by African ancestry. Furthermore, **Figure 5** reveals individuals who harbor a different genetic lineage than their self-reported race/ethnicity—highlighting the advantage of using quantified genetic ancestry instead of self-reporting to investigate genetic contributors of disease or phenotype in admixed populations.

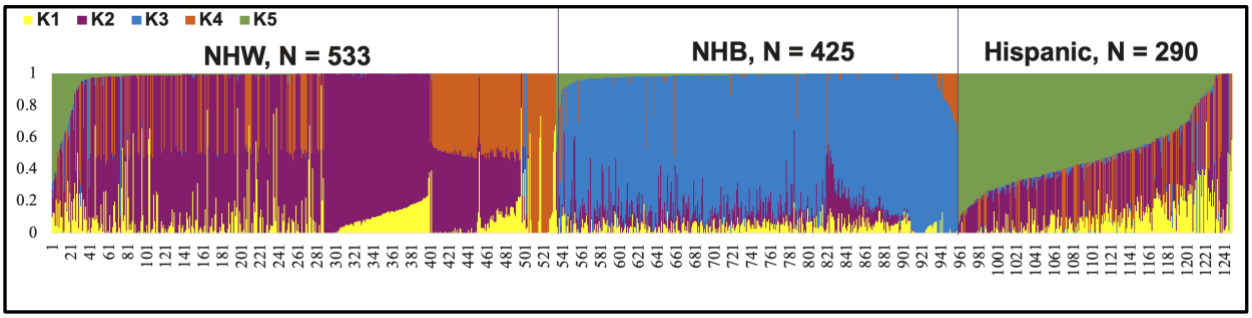

Figure 5. Ancestral Percentages of Individuals in the MM Study Group by Self-Identified Ethnicity

The MM study population (N = 1,248) stratified by self-reported non-Hispanic white (NHW), Non-Hispanic black (NHB), and Hispanic ethnicity map on to the inferred European - K1 (yellow), K2 (purple), K4 (orange), African - K3 (blue), and Amerindian - K5 (green) ancestry.

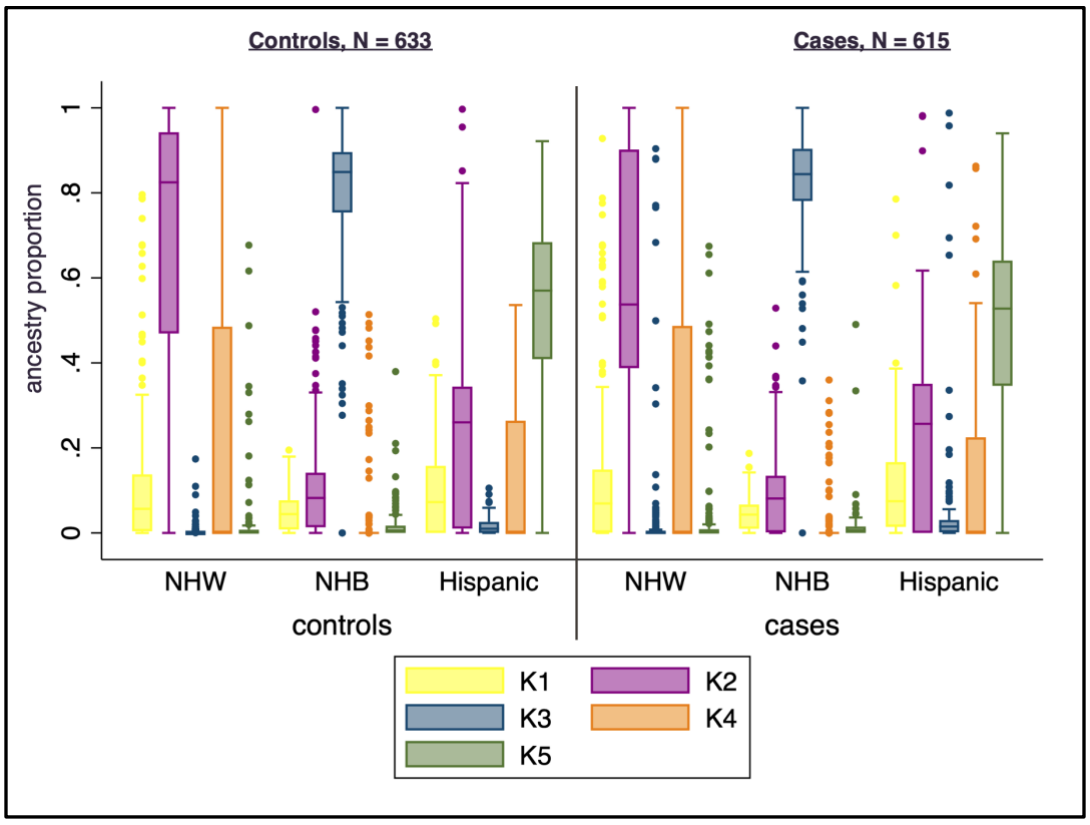

Figure 6. Ancestral Distributions of MM Cases and Controls by Self-Identified Ethnicity

The box plot of MM cases and controls, further stratified by self-reported non-Hispanic white (NHW), Non-Hispanic black (NHB), and Hispanic individuals, illustrates the distribution and median proportion of the inferred genetic ancestry corresponding to European - K1 (yellow), K2 (purple), K4 (orange), African - K3 (blue), and Amerindian - K5 (green) origins.

Association Between Genetic Ancestry and MM risk in the Overall Study Population

We investigated the effect of a 10% increase in European, African, and Amerindian genetic ancestry on MM risk using logistic regression, unadjusted and adjusted for age and gender. For the overall population (**Table 6**), we identified a significant decreased MM risk associated with an increase of the K2 European ancestry. However, K2 is overrepresented in the NHW controls with the median K2 percentage (~ 80%) compared to NHW cases (~ 50%) (**Figure 6**). Therefore, K2 matching between cases and controls is required to conclude that K2 European ancestry is indeed protective of MM risk.

Table 6. MM Risk in the Overall (N = 1248) Study Population by Inferred Ancestry

Unadjusted	OR (95% CI)	P-value
K1 (European)	1.05 (0.95-1.15)	0.34
K2 (European)	0.95 (0.93-0.99)	**0.019**
K3 (African)	1.02 (0.99-1.05)	0.16
K4 (European)	1.03 (0.97-1.09)	0.30
K5 (Amerindian)	0.98 (0.93-1.03)	0.47
Adjusted (age and gender)		
K1 (European)	1.05 (0.95-1.15)	0.33
K2 (European)	0.96 (0.93-0.99)	**0.039**
K3 (African)	1.02 (0.98-1.04)	0.23
K4 (European)	1.03 (0.97-1.09)	0.26
K5 (Amerindian)	0.97 (0.92-1.03)	0.41

OR is the effect size for MM risk for every 10% increase of a given ancestral fraction (K1-K5), unadjusted and adjusted for age and gender

Association Between Genetic Ancestry and MM Risk by Self-Identified Ethnicity

After stratification by self-reported ethnicity, we identified that the K2 European ancestry was significantly associated with a decreased MM risk among NHW

individuals. Although not statistically significant, African ancestry increased MM risk over 2-fold in NHWs. Inferred Amerindian ancestry was also associated with an increase in MM risk in NHW study groups; however, this effect was not significant (**Table 7**)

For the NHBs study subjects (**Table 8**), European ancestry was associated with decreased MM risk, although not statistically significant. Characteristically, a 10% increase in African lineage indicated a significant 15% increase in MM risk, echoing the hypothesis that elevated African genetic ancestry imposed an increase in MM susceptibility. Increasing Amerindian ancestry suggested an increase in MM risk in NHBs; however, with no statistical significance.

When considering self-identified Hispanics (**Table 9**), we identified a significant (P = 0.025) 12% reduced risk of MM for every 10% increase of Amerindian Ancestry, suggesting that Amerindian heritage may be protective of MM risk. Albeit not statistically significant, we identified that MM risk increased by 4 - 6% for increasing European ancestry and by five-fold for elevated African ancestry in Hispanics.

Table 7. MM Risk in NHW (N = 533) Study Population by Ancestry

Unadjusted	OR (95% CI)	P-value
K1 (European)	1.06 (0.95-1.12)	0.27
K2 (European)	0.91 (0.86-0.97)	**0.002**
K3 (African)	2.32 (0.81-6.66)	0.11
K4 (European)	1.04 (0.97-1.12)	0.22
K5 (Amerindian)	1.19 (0.95-1.50)	0.12
Adjusted (age and gender)		
K1 (European)	1.07 (0.95-1.20)	0.33
K2 (European)	0.91 (0.87-0.97)	**0.002**
K3 (African)	2.39 (0.81-7.05)	0.11
K4 (European)	1.04 (0.97-1.12)	0.22
K5 (Amerindian)	1.19 (0.95-1.51)	0.12

OR is the effect size for MM risk for every 10% increase of a given ancestral fraction (K1-K5), unadjusted and adjusted for age and gender

Table 8. MM Risk in NHB (N = 425) Study Population by Ancestry

Unadjusted	OR (95% CI)	P-value
K1 (European)	0.73 (0.38-1.40)	0.35
K2 (European)	0.88 (0.73-1.06)	0.19
K3 (African)	1.13 (0.99-1.30)	0.056
K4 (European)	0.83 (0.60-1.14)	0.30
K5 (Amerindian)	1.00 (0.53-1.87)	0.98
Adjusted (age and gender)		
K1 (European)	0.69 (0.36-1.33)	0.27
K2 (European)	0.86 (0.71-1.05)	0.15
K3 (African)	1.15 (1.00-1.32)	**0.040**
K4 (European)	0.82 (0.59-1.15)	0.26
K5 (Amerindian)	0.92 (0.49-1.72)	0.79

OR is the effect size for MM risk for every 10% increase of a given ancestral fraction (K1-K5), unadjusted and adjusted for age and gender

Table 9. MM Risk in Hispanic (N = 290) Study Population by Ancestry

Unadjusted	OR (95% CI)	P-value
K1 (European)	1.06 (0.85-1.31)	0.59
K2 (European)	0.98 (0.86-1.11)	0.79
K3 (African)	5.09 (0.68-37.7)	0.11
K4 (European)	1.05 (0.92-1.20)	0.45
K5 (Amerindian)	0.88 (0.79-0.98)	**0.024**
Adjusted (age and gender)		
K1 (European)	1.06 (0.85-1.32)	0.57
K2 (European)	0.99 (0.87-1.13)	0.95
K3 (African)	5.35 (0.69-41.2)	0.11
K4 (European)	1.04 (0.91-1.19)	0.50
K5 (Amerindian)	0.88 (0.79-0.98)	**0.025**

OR is the effect size for MM risk for every 10% increase of a given ancestral fraction (K1-K5), unadjusted and adjusted for age and gender

Principle Component Analysis in Self-Reported Hispanics

To match admixed cases and controls by their genetic principal components, we performed a principal component analysis on the LD pruned OncoArray genotype dataset in Hispanics. We include the place of residence for each case (**Table 5**) when visualizing PCA results to discern if the outliers were from international patients.

Figure 7 shows the plot of the principal components (PC1 vs PC2) of the Hispanics study subjects. The population structure of Hispanic cases (orange) and controls (blue) were distributed almost evenly between the European and Amerindian genetic ancestry, with 11 Hispanic outlier cases. When cross-referencing the genetic ancestry of the 11 outliers, we found that those cases exhibited the highest proportion of African lineage. Therefore, PC1 corresponded to African genetic ancestry.

We then removed the 11 outliers and re-ran the PCA analysis on the remaining 279 Hispanics study subjects to match cases and controls. PCA results are illustrated in **Figure 8, Figure 9,** and **Figure 10.** With removed individuals of primarily African descents, **Figure 8** shows an almost perfect match of cases and controls by principal components corresponding to European and Amerindian ancestry. We did, however, identify potential outliers e.g., Texas resident case of PC2 > 0.25.

Therefore, we plotted PC2 vs. PC3 (**Figure 9**) and PC3 vs. PC4 in **Figure 10**, which uncovered primarily international case outliers from Mexico, Honduras, and Puerto Rico. After removing the seven other outliers circled out in **Figure 10**, we have a closely matched population of Hispanic cases and controls by their genetic principal components.

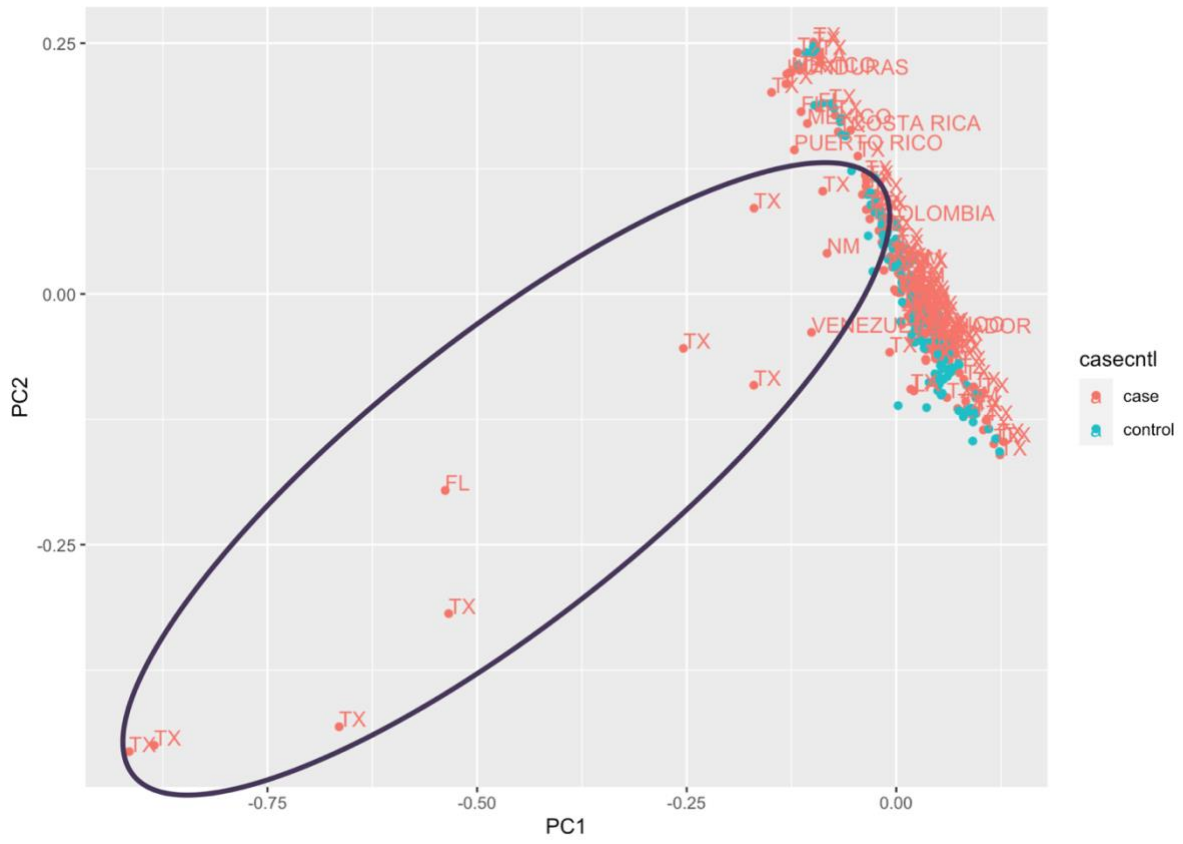

Figure 7. Principal Component Analysis Result for Hispanics (PC1 vs PC2, N = 290)

Population structure of N = 290 Hispanic cases (orange) and controls (blue) is plotted (PC1 vs. PC2**)** using PCA results from flashPCA, with cases' place of residence included.

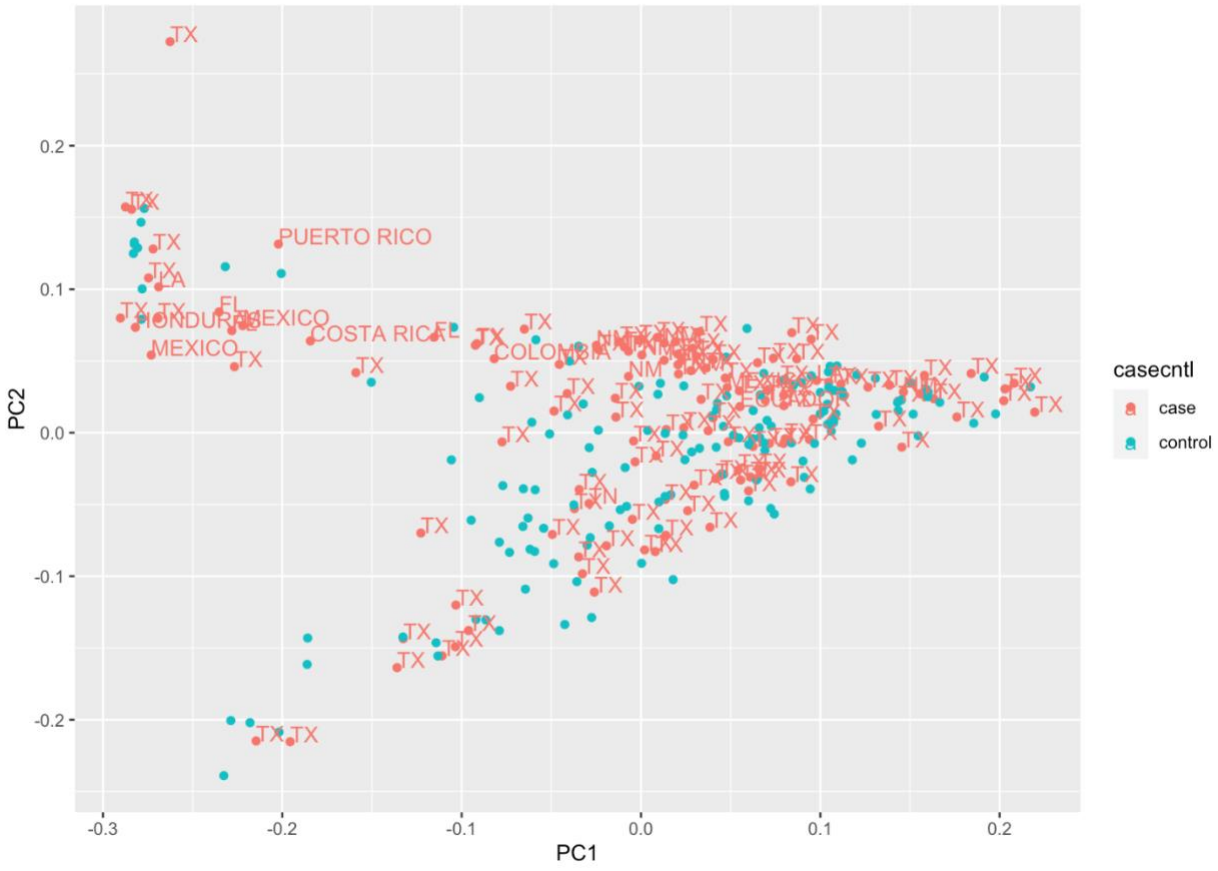

Figure 8. Principal Component Analysis Result for Hispanics (PC1 vs PC2, N = 279)

Population structure of N = 279 Hispanic cases (orange), and controls (blue) is plotted (PC1 vs. PC2**)** using PCA results from flashPCA, with cases' place of residence included.

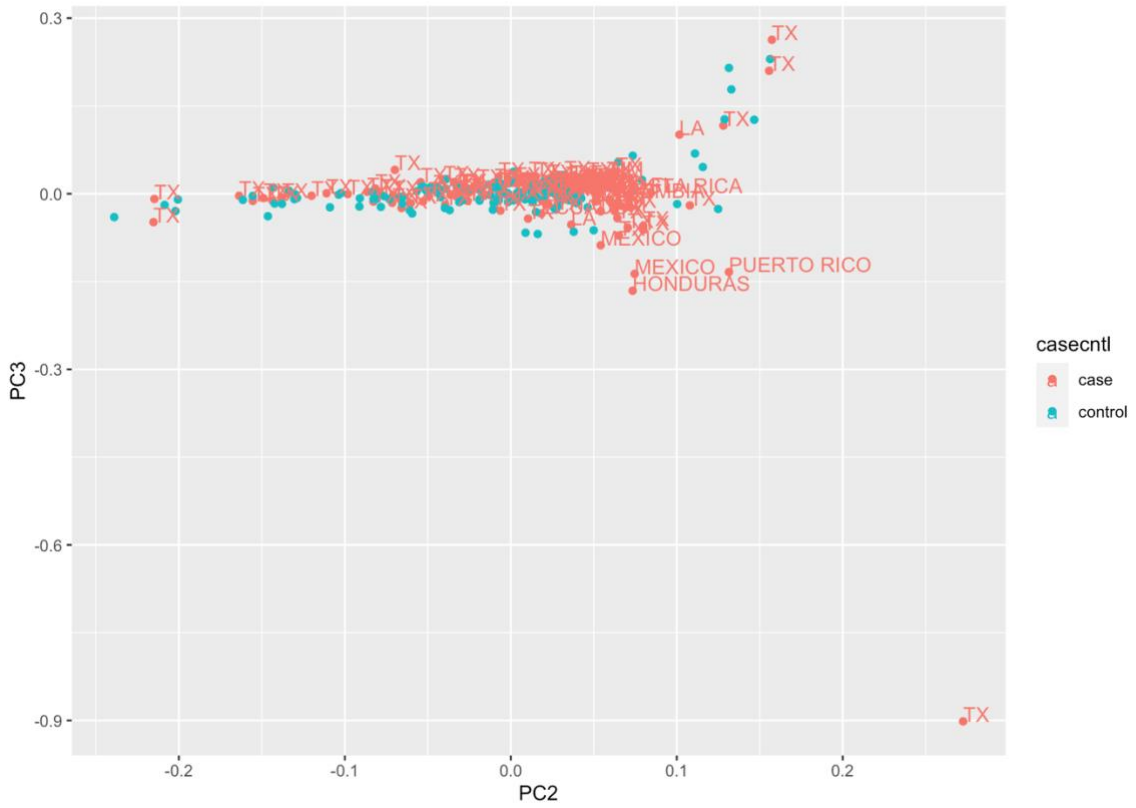

Figure 9. Principal Component Analysis Result for Hispanics (PC2 vs PC3, N = 279)

Population structure of N = 279 Hispanic cases (orange), and controls (blue) is plotted (PC2 vs. PC3**)** using PCA results from flashPCA, with cases' place of residence included

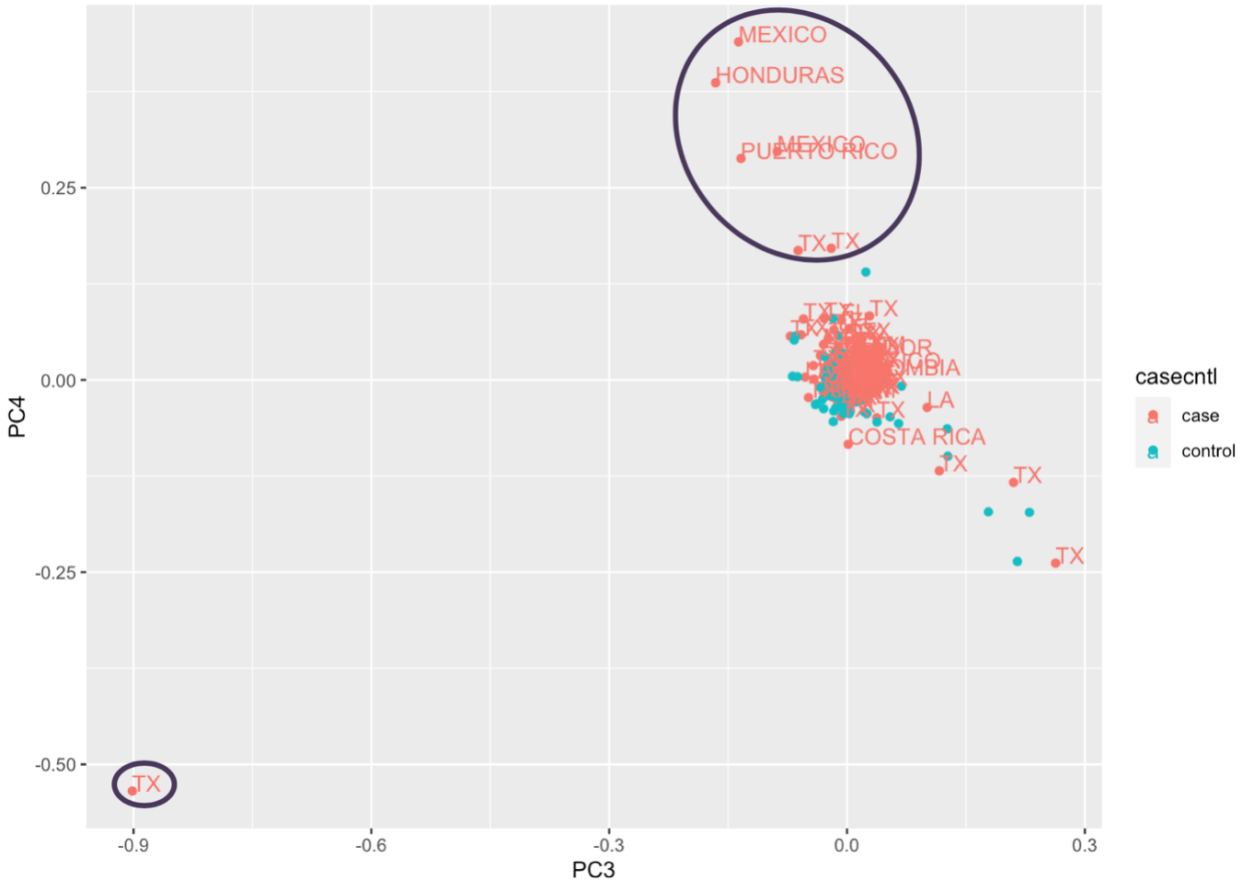

Figure 10. Principal Component Analysis Result for Hispanics (PC3 vs PC4, N = 279)

Population structure of N = 279 Hispanic cases (orange), and controls (blue) is plotted (PC3 vs. PC4**)** using PCA results from flashPCA, with cases' place of residence included.

MM Risk in Self-Identified Hispanics after PCA Matching

The effect size of a 10% increase in ancestry after removing the initial 11 outlier cases is described in **Table 10**. The odds ratio of MM risk after removing the additional seven outlier Hispanic individuals (total, N = 18) is also shown in **Table 11**. Although the protective effect of Amerindian ancestry remained, the statistical significance was attenuated after PCA matching.

Table 10. MM Risk in PC Matched Hispanic (N = 279) Study Population by Ancestry

Unadjusted	OR (95% CI)	P-value
K1 (European)	1.06 (0.85-1.32)	0.55
K2 (European)	1.01 (0.89-1.15)	0.84
K3 (African)	1.11 (0.06-17.9)	0.94
K4 (European)	1.05 (0.92-1.20)	0.46
K5 (Amerindian)	0.92 (0.82-1.03)	0.15
Adjusted (age and gender)		
K1 (European)	1.07 (0.86-1.33)	0.54
K2 (European)	1.02 (0.89-1.16)	0.71
K3 (African)	1.23 (0.076-20.0)	0.88
K4 (European)	1.04 (0.91-1.99)	0.50
K5 (Amerindian)	0.92 (0.82-1.03)	0.15

OR is the effect size for MM risk for every 10% increase of a given ancestral fraction (K1-K5) after removal of 11 outlier Hispanic individuals, unadjusted and adjusted for age and gender

Table 11. MM Risk in PC Matched Hispanic (N = 272) Study Population by Ancestry

Unadjusted	OR (95% CI)	P-value
K1 (European)	0.90 (0.69-1.17)	0.46
K2 (European)	1.04 (0.91-1.18)	0.54
K3 (African)	1.17 (0.06-19.0)	0.90
K4 (European)	0.99 (0.86-1.53)	0.98
K5 (Amerindian)	0.96 (0.86-1.08)	0.57
Adjusted (age and gender)		
K1 (European)	0.91 (0.70-1.19)	0.51
K2 (European)	1.05 (0.92-1.20)	0.42
K3 (African)	1.29 (0.080-21.0)	0.85
K4 (European)	0.99 (0.86-1.14)	0.91
K5 (Amerindian)	0.99 (0.85-1.08)	0.56

OR is the effect size for MM risk for every 10% increase of a given ancestral fraction (K1-K5) after removal of 18 outlier Hispanic individuals, unadjusted and adjusted for age and gender

Discussion

This chapter identified the differential risk of MM by genetic ancestry and elucidated that Amerindian ancestry may have protective associations for MM risk in Hispanics. In our overall study population, we found a trend of increased MM risk by increasing African ancestry and a decrease of MM risk by increasing European and Amerindian ancestry.

The influence of genetic ancestry on MM risk was more pronounced when we stratified our subjects by self-reported race/ethnicity. Consequently, we reported an 8% significantly decreased MM risk by a 10% increase of the predominant K2 European ancestry in NHWs. We also identified a significantly enhanced MM risk by increasing African ancestry in our NHB study subjects (African ancestry; OR: 1.15; 95% CI: 1.00-1.32; $P = 0.040$). The increase of MM risk in African descents is well documented in the literature(105, 106), which aligned with our findings. We also found a 12% decreased MM risk for every 10% increase of Amerindian ancestry ($P = 0.025$), suggesting a protective effect of MM risk for admixed populations that carry a high percentage of Amerindian ancestry as opposed to those of European/Spanish ancestry.

We know that our Hispanic control population is predominately of Mexican descent and therefore have a higher proportion of Amerindian ancestry than most US Hispanic populations. However, we do not have background information for the Hispanic patient population at MD Anderson. Therefore, we controlled for the potential ancestral mismatch of our heterogeneous Hispanic cases and controls to see the protective association using principal component analysis. After PC adjustment by principal genetic components and controlling for ancestral differences between the

Hispanic case and control populations, the significant effect of Amerindian ancestry was attenuated.

Albeit not statically significant, Amerindian consistently remained protective of MM risk after PC matching for a 10% increase of K5 with a wide confidence interval (0.85 to 1.08) (**Table 11**). Additionally, the most recent SEER data shows that the incidence rate of US Hispanics is slightly lower than NHWs, and the incidence rate of American Indians and Alaskan residents are almost half of the general population. Interestingly, one nested case-control study of California Hispanic farmworkers derived from a multi-ethnic cohort study of 130,000 farmers, revealed an elevated risk of incidence in hematological malignancies, i.e., leukemia and non-Hodgkin lymphoma, particularly in female workers exposed to pesticides. However, an increase in MM incidence was not present in this study group(107). Farming is a known MM occupational hazard(108); therefore, it is still plausible that Amerindian ancestry is somewhat protective of MM risk. However, additional validation studies on a larger Hispanic study population is essential to determine if Amerindian ancestry is indeed associated with a reduction of MM risk.

Overall, the cancer incidence rate is lower in Hispanics than the NHW and NHB populations(109). However, infectious based cancers and cancer associated with diabetes and obesity, such as gastrointestinal and liver cancers, are becoming more prevalent in Hispanics(109). Furthermore, the inverse relationship between Amerindian genetic ancestry and risk of incidence has been reported in common cancers like prostate and breast cancer, after adjusting for socioeconomics and lifestyle. One study that confirmed this inverse relationship of breast cancer incidence and Native American heritage in postmenopausal Hispanic women found that those with the highest proportion of Native American ancestry (71–100%) carried risk loci on *IkBKB, mTOR,*

PDK2, PRKAA1, RPS6KA2, and *TSC1*. In contrast, genes NFκB1, PTEN, and RPS6KA2 were associated with Hispanics, with over 70% of European ancestry(110). Also, Latin American women report a high incidence of aggressive HER2+ breast cancer. This association also trends with elevated Amerindian ancestry(111), analogous to our findings that ancestral background is associated with differential cancer risk.

We do not have genetic data on the relationship between MM and Hispanics in the literature. However, one study reported that Hispanic children with Native American ancestry over 10% are at a higher risk of relapse of B-cell acute lymphoblastic leukemia (ALL), a hematological malignancy closely related to MM(112). In parallel, risk loci rs3731217 and rs3824662 partially explain ALL relapse in patients that harbor high Native American genetic lineage(113, 114). Likewise, genome-wide studies or local admixture mapping in a larger Hispanic patient population may identify genetic loci that mediate the protective effect of MM. Nevertheless, extensive research is vital to narrow the gaps in our knowledge on the effects of genetic ancestry in MM susceptibility in these subgroups.

There are limitations to using unstructured ADMIXTURE. To navigate this, we previously projected our inferred ancestry results to three publicly available HapMap3 reference populations(115). The reference populations selected include (1) European ancestry — (CEU: Utah residents with Northern and Western European) (2) African ancestry — (YRI: Yoruba in Ibadan, Nigeria), and (3) Mexican ancestry — (MXL: Mexican ancestry in Los Angeles). Individuals genotyped for the MXL reference population identified themselves as having at least 3 out of 4 grandparents born in Mexico, potentially matching Mexican ancestry to our self-identified Hispanic study group.

We also performed an extensive literature search to identify ancestry informative SNP markers that exhibit substantial allele frequency differences between European, Amerindian, and African reference populations to precisely estimate the population structure for our admixed study subjects. From this, we identified two groups in the literature with a robust AIMs panel(116, 117). One reported 2,120 AIMs derived from reference populations genotype of European origins, as well as Mesoamericans (Maya and Nahua from Mexico), South Americans (Aymara/Quechua from Bolivia and Quechua from Peru), West African (YRI), and East Asians (populations from China and Japan). The other AIMs panel (N = 975) was assembled using the HapMap3 reference European populations (CEU, Utah residents from northern and western European Populations), African populations (YRI, Yoruba in Ibadan, Nigeria and LWK, Luhya in Webuye, Kenya), and Pima Indians in Arizona (PIMA, full heritage Pima or Tohono O'odham individuals or combination of the two tribes)(116).

We then extracted these AIMs from our study population's imputed dataset with high QC cutoff (average $r^2 = 0.99$ and individual call rate > 0.95). Next, we simulated Amerindian reference genotype drawing on genotypes from the 1000 Genomes reference population (CEU, YRI)(118) and by calculating binomial distribution with success probability equal to the allele frequencies of extracted AIM SNPs from Mesoamerican and PIMA source population. Afterward, we ran unstructured ADMIXTURE on the overall study population using the simulated AIMs panel to determine our study group’s population structure and estimated the genetic ancestry of admixed individuals in a precise manner.

Unfortunately, neither of these strategies, i.e., using CEU, YRI, and MXL ancestry as a reference population and using simulated Amerindian AIMs, were feasible. The MXL reference population was admixed within itself and introduced noise

in the study samples' population structure. Moreover, we found differing ancestral estimates from simulated AIMs than the assumed ancestral fraction of our admixed study individuals. For instance, our self-identified Hispanics reported a 32% - 36% Amerindian ancestry and about 60% European ancestry using the simulated AIMs. Past population structure analysis has shown Amerindian lineage (51% - 56%) in those with Mexican lineage followed by European ancestry (40% - 45%) and a small share of African descent (2% - 5%)(119). Given that our Hispanic subjects are primarily Mexican descendants residing in Texas, we postulate that the direct genotype ancestry (K = 5) estimate of 47 - 52% Amerindian followed by 46 - 47% European and 1 - 5% African was a more accurate predictor.

A potential explanation for this inconsistency may be due to our AIMs derived reference populations, i.e., Mesoamericans (Maya and Nahua from Mexico) and PIMA Indians. These populations may have been too specific and isolated as they only captured a smaller subset of Amerindian ancestry in our heterogeneous Hispanic study groups. Another explanation may attribute to the imputation platform, i.e., the 1,000 Genomes Project which is underrepresented in its admixed reference population, potentially skewing our ancestry estimate towards European ancestry.

We acknowledge that the probability of finding specific AIMs in direct genotype platforms is low. That is why we extracted our AIMs from the imputed dataset with 22 million+ SNPs. However, to mitigate this discrepancy, we suggest that future investigators impute direct genotype on platforms such as TOPMed(120) before ancestral estimation. TOPMed contains reference sequencing data of over 100,000 admixed individuals and has also identified risk loci unique to Hispanics that would have otherwise been genome-wide insignificant on the 1,000 Genomes Project.

Overall, the unstructured K = 5 ADMIXTURE approach was the most viable method to ascertain our study subjects' population structure. From this, we found suggestive evidence that Amerindian ancestry may be protective of MM susceptibility with a recommendation for further investigation. We also confirmed the direct relationship between increasing African genetic lineage and increasing MM risk in NHBs with statistical significance.

Chapter 4: Genetic Ancestry Mediates MM Disease Types and Outcomes in Hispanics

Introduction and Study Objective

MM Evaluation and Clinical Presentation

As described in Chapter 1, MM is a heterogeneous disease characterized by the uncontrolled production of plasma cells and the presence of one or more CRAB symptoms. Plasma cells produce excessive immunoglobulin (Ig) heavy chains (G, A, D, E, or M) and one type of light chain (kappa or lambda). IgG is the most common subtype (54%), followed by IgA (21%) and light chain restricted (16%)(121).

For MM diagnosis, the necessary clinical workup includes a bone marrow biopsy to identify excessive CD138+ plasma cells and cytogenetics/FISH analysis. FISH probes detect chromosomal abnormalities such as hyperdiploidy and the presence of high-risk mutations i.e., del17p, t(4;14), t(11;14), and t(14;16). Additional evaluation for suspected MM includes a complete blood count and serum biomarkers, i.e., lactate dehydrogenase, creatinine, beta-2-microglobulin, and albumin. In addition, serum protein electrophoresis is used to quantify M paraproteins and identify Ig subtypes and serum free light chains. Urine studies include an immunofixation and a 24-hour urine protein test to detect Bence Jones(122). Radio-imaging is also essential to detect lytic lesions and compression fractures that cause bone disease(121, 123).

The three staging systems that have been developed for MM include the Durie-Salmon system, the International Staging System (ISS), and the Revised International Staging Symptoms (R-ISS). The ISS and R-ISS staging criteria put forward by the International Myeloma Working Group are the most recent prognosis identifiers. **Figure 11** illustrates the biomarker levels used to determine MM diagnosis and staging.

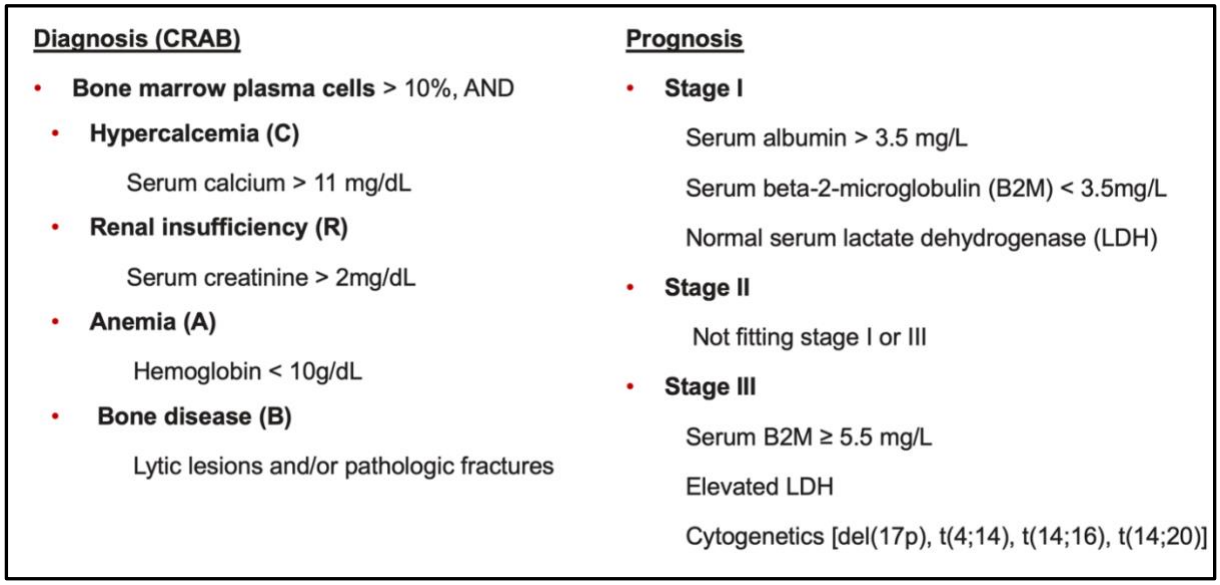

Figure 11. MM Diagnosis and Prognostic Biomarkers

MM diagnosis is based on > 10% of plasma cells in the bone marrow in addition to one or more of the MM CRAB symptoms — diagnosed from serum calcium levels, creatinine, hemoglobin levels, and radio imaging results for detecting bone lesions and pathologic fractures. MM prognosis and stage classification is established by serum beta-2-microglobulin, albumin, and lactate dehydrogenase levels, as well as the detection of high-risk cytogenetic abnormalities.

Study Objective

In Chapter 3, we illustrated that genetic ancestry contributes to MM susceptibility by ethnicity. In this Chapter we analyzed the patterns of clinical phenotypes that characterize disease subtypes that may also drive prognosis in Hispanics. In addition, we utilized the genetic ancestry information from Chapter 3 to ascertain if distinct patterns of clinical profiles in Hispanics also correlate with European, African, or Amerindian genetic ancestry.

We leveraged the robust medical records of the diverse patient population at MD Anderson to determine patterns of MM clinical characteristics and identify differences in clinical features by self-reported race/ethnicity. Next, we used genetic ancestry information to find differences in MM outcomes by ancestral background. MM outcomes studied by genetic ancestry included occurrence rate of somatic mutations/high-risk profiles, MGUS/SMM diagnoses, and survival in the overall study population.

The objective of this Chapter is to identify important clinical features that are enriched in Hispanic patients, while also using this information to uncover differences of MM outcome by race/ethnicity and genetic ancestry. Therefore, we designed a study to ascertain unique MM disease features in Hispanics and compared them with a multi-ethnic MM patient population. Through the increased knowledge of the clinical presentation and course of disease in Hispanic cases this study may help to provide tailored care to this population.

Study Design and Methods

Study Population, Genotype, and Ancestry Analysis

The study populations for this analysis were the N = 615 MM patients described in Chapter 3 and **Table 4.** Genotyping and ancestry inference of the MM study populations are also described in Chapter 3 Study Design and Materials and displayed in **Figure 5** and **6.**

Clinical Data Collection

For clinical data collection from the MD Anderson electronic medical records, we created a comprehensive 9-paged customized abstraction form with N > 100 variable in collaboration with Dr. Elisabet Manasanch (Department of Lymphoma/Myeloma). Information gathered from the abstraction form included demographics, history of MM precursors, MM immunoglobulin (Ig) subtypes, cytogenetics karyotype, FISH data, and lytic lesions detected by X-ray, PET CT, or MRI. Additional data included baseline diagnostic and prognostic indicators from the serum, urine and the bone marrow clinical assays, in addition to serum and urine biomarker levels indicating treatment response. We also abstracted treatment regimen, dates of treatment, clinical indicators of relapse and response, dates of follow up, dates of death.

Statistical Analysis

To investigate the clinical features of Hispanic MM patients in comparison to NHW and NHB patients, we performed the appropriate chi-square or student's t-test for baseline categorical and continuous biomarkers, respectively. Risk of MM binary outcomes (previous reporting of MGUS/SMM and occurrence of high risk cytogenetic mutations) by genetic ancestry (African (K3), European (K1, K2, K4), or Amerindian

(K5) was assessed using logistic regression with corresponding odds ratios (ORs) and 95% confidence intervals (CIs).

Cox proportional hazards model adjusted for age, sex, and high-risk cytogenetics was used to assess overall survival as a function of race/ethnicity. Additionally, survival hazard ratio (HR with 95% CI) was performed for every 10% increase of genetic ancestry as described in Chapter 3. Recruitment for this study began in 2010, even if those patients had a history of MM prior to 2010. Therefore, to avoid prevalence confounding in our survival analysis, we divided the patients into two groups; those diagnosed prior 2010 and after 2010. Kaplan–Meier survival function and corresponding log-rank tests were used to plot overall survival stratified by race/ethnicity and diagnosis year (prior or after year 2010). Survival time was defined as the duration from the date of diagnosis to the date of death or last follow-up visit. P values < 0.05 were considered statistically significant. Statistical analysis was performed using GraphPad Software (San Diego, California) and Stata software (version 16; StataCorp, College Station, Texas).

Results

Clinical Characteristics

The clinical profile of our study subjects is shown in **Table 12**. Some defining features of Hispanics include a significantly younger age of diagnosis (57.4 years, *P* = 0.003), compared to NHW (61.4), but similar to NHB (57.5 years) patients. Additionally, we found a lower prevalence of MGUS and/or SMM in Hispanics (8.4%) compared to NHW (18.8%) and NHB (16.1%). For Hispanics, IgG comprised the majority (53.8%), followed by IgA and light chain restricted (21.6%) and IgD (1.4%). From this, we observed that the IgA subtype occurs at a slightly higher rate in Hispanics compared to NHW (18.8%) and NHB (13.7%). Hispanics have an intermediate percentage of high-risk cytogenetics abnormalities (11.8%) between NHW (13.0%) and NHW (7.1%). Overall, we observed some variation of clinical phenotype Hispanics when compared to NHW and NHB patients.

Table 12. Characteristics of Study Population by Self-Identified Ethnicity (N = 615)

	Hispanic	NHB	NHW
Total:	143	211	261
Dates of diagnosis:	1998 - 2019	1981 - 2019	2001 - 2019
Gender:			
Male (%)	81 (56.6)	105 (49.7)	158 (60.5)
Female (%)	62 (43.4)	106 (50.3)	103 (39.4)
Median age of diagnosis:			
	57.4 (29.0 - 82.0)	57.5 (28.0 - 87.0)	61.4 (36.0 - 87.0)
Previous case of MGUS/SMM (%):			
	12 (8.4)	34 (16.1)	49 (18.8)
IG Subtype			
IgG (%)	77 (53.8)	144 (68.2)	150 (54.5)
IgA (%)	31 (21.6)	29 (13.7)	49 (18.8)
IgD (%)	2 (1.4)	3 (1.4)	2 (0.8)
κ or λ light chain	31 (21.6)	34 (16.1)	59 (22.6)
N/A	2 (0.7)	1 (0.5)	1 (0.4)
Risk by cytogenetic abnormalities (%):			
High risk (%)	17 (11.8)	15 (7.1)	34 (13.0)
Standard risk (%)	69 (48.3)	108 (51.2)	135 (51.7)
N/A (%)	57 (39.9)	88 (41.7)	92 (35.3)

High risk cytogenetic abnormalities: t(4;14), t(14;16), t(14;20), del(17/17p)
MGUS: Monoclonal gammopathy of unknown significance
SMM: Smoldering multiple myeloma
NHB: Non-Hispanic black
NHW: Non-Hispanic white
N/A = Not available

Clinical Phenotype by Genetic Ancestry

We calculated the relationship between genetic ancestry and clinical features exhibiting differences by self-identified ethnicity in our study subjects and those that have previously indicated differential association by race/ethnicity. Chapter 1 described that MGUS/SMM prevalence and chromosomal abnormalities occur with a varying degree by ancestry.

In the previous section, we discovered a higher frequency of high-risk cytogenetic mutations in self-identified NHWs, followed by Hispanics and NHBs. Some studies have also indicated a greater occurrence of chromosomal abnormalities in individuals of European descent(45, 105). Genetic ancestry results in Chapter 3 indicated that Hispanics were a distribution of inferred European (~40%) and Amerindian (~50%) ancestry. Therefore, to determine if the European ancestry is a driving factor of the higher frequency of the high-risk cytogenetic mutations in Hispanics and NHW patients, we calculated the odds ratio of these mutations in patients with > 40% inferred European ancestry and cases with > 50% Amerindian ancestry.

Consequently, we discovered that Hispanics having > 40% European ancestry showed over 3-fold increased risk of these mutations (P = 0.036). We also found a borderline significant inverse relationship between Amerindian ancestry and high-risk cytogenetic mutations in Hispanics (OR: 0.33; 95% CI: 0.11-1.03; P = 0.055) in this group. When calculating the risk of mentioned genetic abnormalities in the overall study group, patients with > 40% European ancestry also exhibited a 1.70-fold increased risk of these mutations (P = 0.056) with no significant association by Amerindian ancestry (**Table 13**). This suggests that European ancestry was the driving factor for high-risk cytogenetics in the Hispanic, as well as overall study subjects.

We used the same method to determine if the prior case of MGUS/SMM is mediated by European or Amerindian ancestry. However, we did not find any significant association between MGUS/SMM prevalence by European or Amerindian ancestry in Hispanics or the overall study population (**Table 14**)**.**

Table 13. High-Risk Cytogenetic Mutations by European and Amerindian Ancestry

Study Group	OR (95% CI)	P-value
Overall (N = 378)		
European > 40%	1.70 (0.98-2.95)	0.056
Amerindian > 50%	0.68 (0.40-1.76)	0.15
Hispanic (N = 86)		
European > 40%	3.22 (1.06-9.78)	**0.036**
Amerindian > 50%	0.33 (0.11-1.02)	0.055

OR is the risk of occurrence of one or more of the t(4;14), t(14;16), t(14;20), del(17/17p) abnormalities for individuals with > 40% European and Amerindian ancestry in the overall study populations and Hispanics only.

Table 14. Prior Diagnosis of MGUS/SMM by European and Amerindian Ancestry

Study Group	*OR (95% CI)	P-value
Overall (N = 615)		
European > 40%	1.35 (0.85-2.14)	1.96
Amerindian > 50%	0.69 (0.44-1.10)	0.12
Hispanic (N = 413)		
European > 40%	0.81 (0.25-32.56)	0.72
Amerindian > 50%	0.69 (0.21-2.22)	0.54

OR is the risk of a prior diagnosis of MGUS and/or SMM for individuals with > 40% European and Amerindian ancestry for the overall study population and Hispanics only. *OR is adjusted for age

Diagnostic and Prognostic Clinical Biomarkers

We compared the diagnostic and prognostic marker levels of Hispanic patients to NHW and NSW study subjects at diagnosis illustrated in **Figures 12** and **13**. When considering diagnostic blood markers, we observed a significantly lower hemoglobin level in NHB patients (P = 0.019) compared to Hispanics with no significant differences compared to NHWs (**Figure 12A**). A low hemoglobin level < 10.0 g/dL is suggestive of anemia, one of the CRAB symptoms.

A high creatinine level is also an indicator of CRAB manifestation in MM patients through renal failure. We found a significant elevation of median creatinine levels (P = 0.035) NHB patients compared to Hispanics. However, differences in median creatinine levels were not significant different between Hispanics and NHW cases (**Figure 12B**).

The percentage of clonal plasma cells in the bone-marrow and hypercalcemia did not indicate any significant difference by ethnicity (**Figure 12C**, **Figure 12D**). When we examined differences in prognostic markers by ethnicity, we observed significantly elevated beta-2-microglobulin levels in NHB patients compared to Hispanics (**Figure 13A**). We also observed a lower median M spike in Hispanics than NHBs, with no significant difference compared to NHWs (**Figure 13B**). High beta-2-microglobulin levels and M spikes signify tumor burden and adverse prognosis in MM patients, indicating that Hispanics have a favorable survival indicators than NHB patients. We did not find any significant difference in median LDH and albumin levels in Hispanics compared to NHBs and NHWs (**Figure 13C, Figure 13D**).

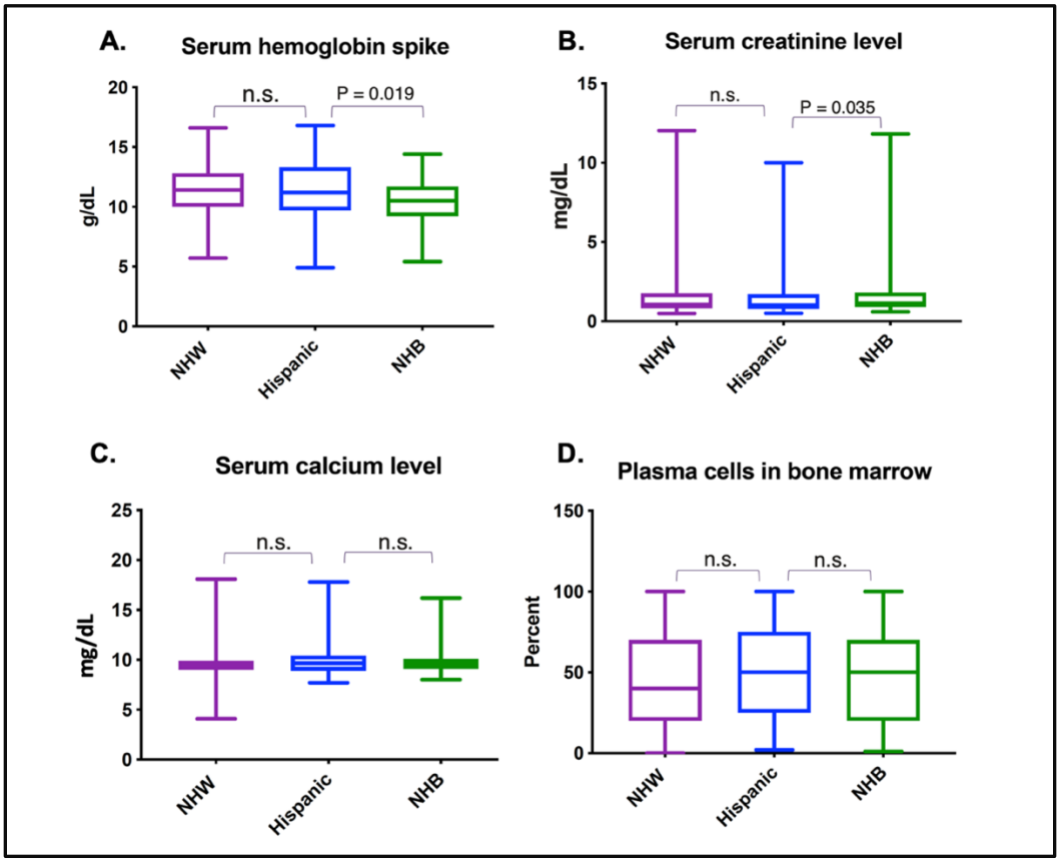

Figure 12. Diagnostic Biomarker Levels in Hispanic, NHW and NHB Patients

The box plots illustrate median levels of diagnostic biomarkers that indicate the CRAB symptoms i.e. (**A.**) serum hemoglobin (**B.**) serum creatinine (**C.**) serum calcium and (**D.**) median percentage of plasma cells in the bone marrow, stratified by NHW (purple), Hispanic (blue) and NHB (green) patients.

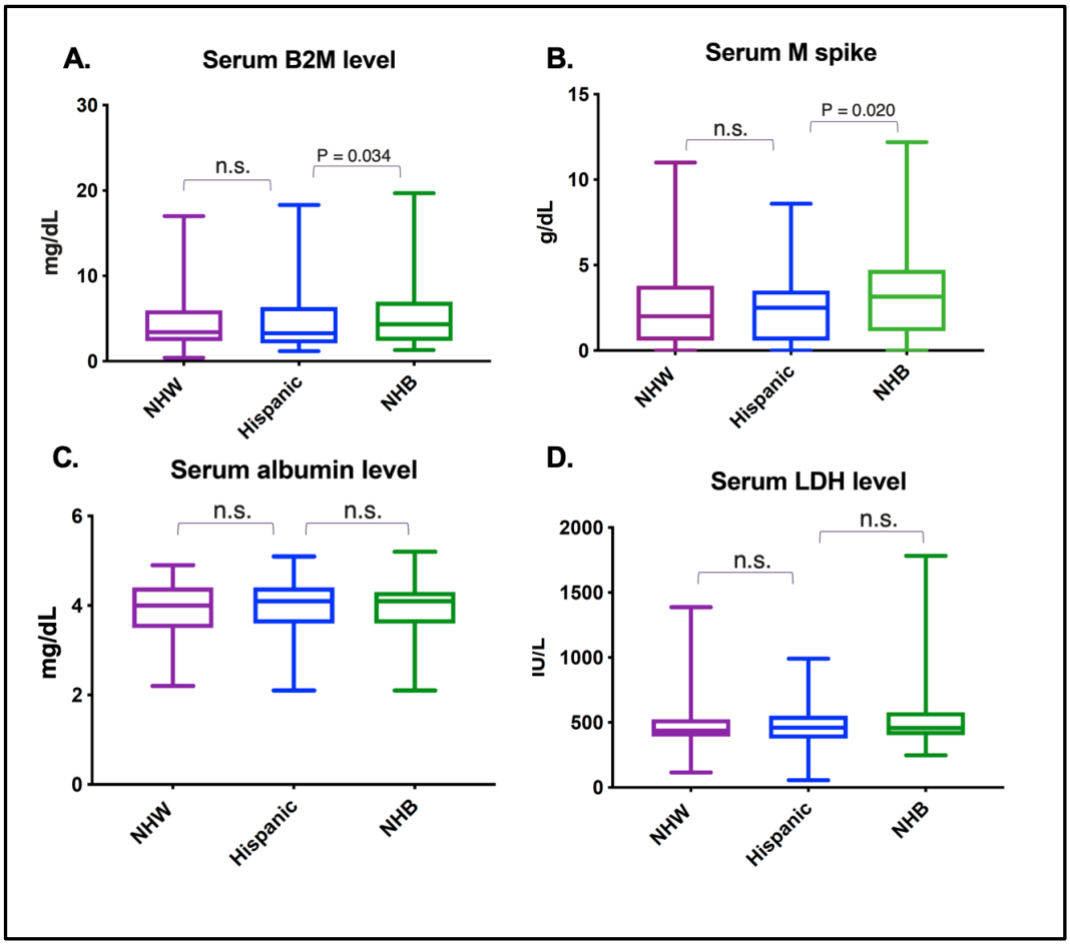

Figure 13. Prognostic Biomarker Levels in Hispanic, NHW and NHB Patients

The box plots illustrate median levels of prognostic biomarkers to dictate staging of MM i.e. (**A.**) serum B2M (beta-2-microglobulin) (**B.**) serum M protein spikes (**C.**) serum albumin and (**D.**) median lactate dehydrogenase (LDH), stratified by NHW (purple), Hispanic (blue) and NHB (green) patients.

Overall Survival by Ethnicity and Genetic Ancestry

We report the overall survival and Kaplan-Meier curve of all patients stratified by ethnicity; but further divided into the incident and prevalent group of cases for those recruited prior and after the year 2010, respectively. As expected, there was a confounding effect from MM prevalence in **Figure 14** with a higher median survival time of 6 - 8 years that do not align with the average median survival time of approximately 5 - 6 years in the general population(47, 124). After controlling for this survival bias by stratifying patients by incident/prevalent cases, the median survival times in Hispanics (6.9 years), NHBs (6.6 years), and NHW (6.5 years) did not significantly differ by race/ethnicity. Consequently, the Cox regression estimates (with NHWs as reference) adjusted for age, gender, and high-risk cytogenetics did not significantly vary by race/ethnicity (**Figure 15**).

Similarly, the effect size for survival for every 10% increase of each European, African, or Amerindian lineage was close to 1.0 (**Table 15**), suggesting that genetic ancestry may not be a factor influencing overall survival in Hispanics or other racial populations. Several iterations of genetic ancestry percentages were tested with no significant associations to overall survival (data not shown).

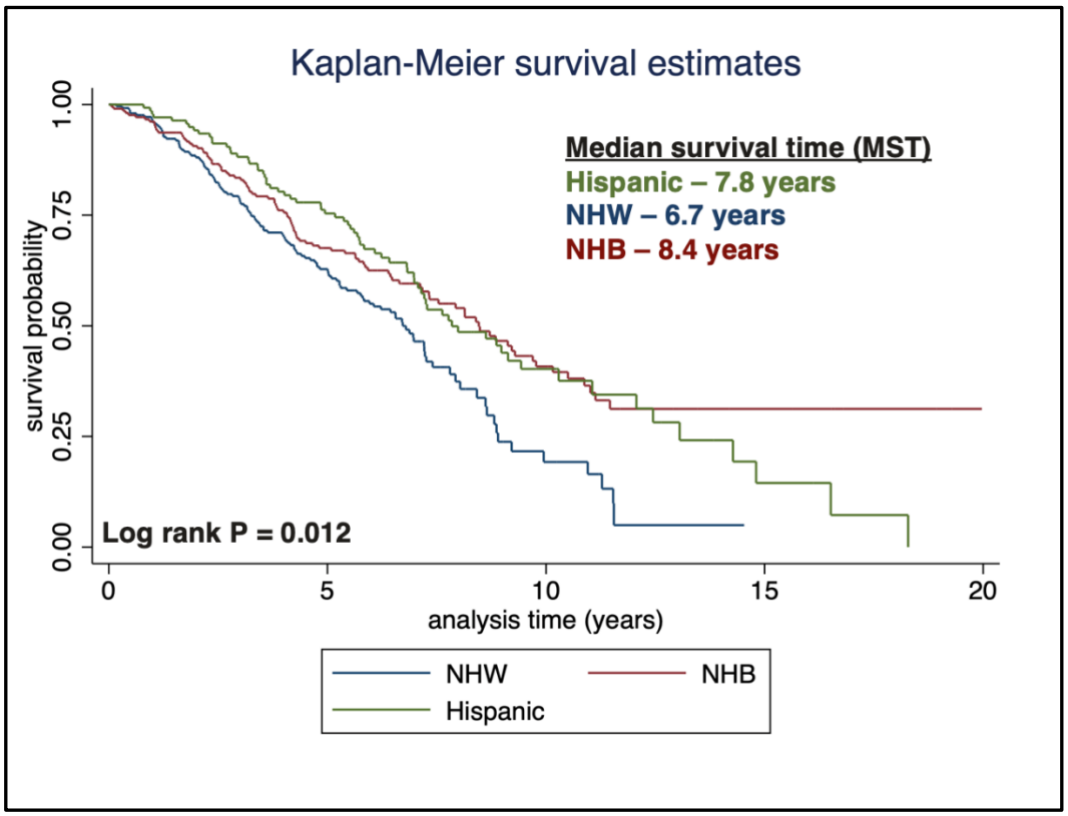

Figure 14. Overall Survival of MM Cases by Ethnicity

Kaplan Meier curves show overall survival of the entire MM patient cohort stratified by ethnicity (NHW, Hispanic, NHB). MST is median survival time.

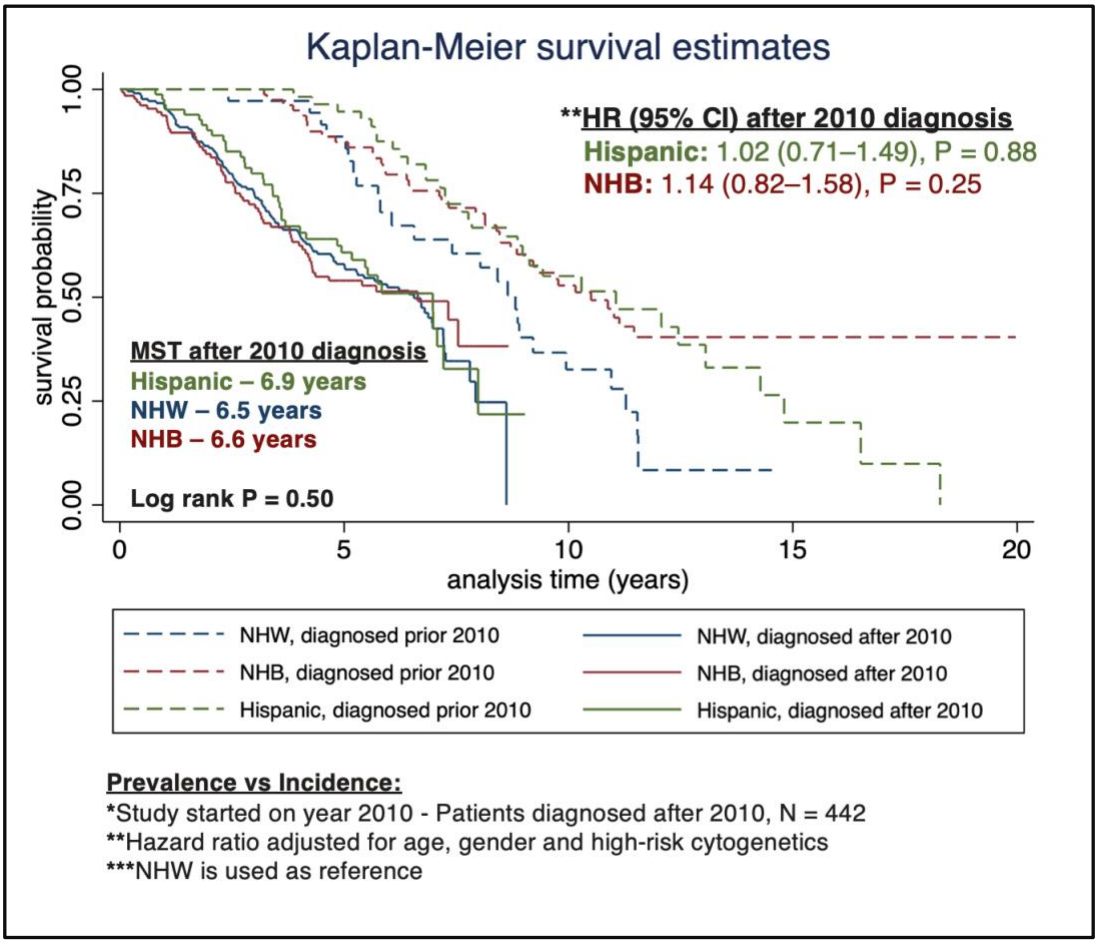

Figure 15. Overall Survival of MM Cases by Ethnicity and Diagnosis Year

Kaplan Meier curves show overall survival of the entire MM patient cohort stratified by ethnicity (NHW, Hispanic, NHB) and further stratified by patients diagnosed before and after year 2010. Solid lines indicate those diagnosed after 2010, and dashed lines correspond to those diagnosed prior to 2010. HR is hazard ratio (95% CI) for patients diagnosed after 2010 adjusted for age, gender, and high-risk cytogenetics. NHW is used as reference for HR. MST is median survival time.

Table 15. Overall Survival by Genetic Ancestry

Ancestry	*HR (95% CI)	P-value
European (K2)	1.00 (0.97-1.03)	0.69
European (K4)	0.99 (0.93-0.86)	0.93
European (K1)	1.00 (0.89-1.12)	0.98
African (K3)	1.02 (0.99-1.05)	0.36
Amerindian (K5)	1.01 (0.94-1.08)	0.74

HR is the effect size for survival for every 10% increase of a given ancestral fraction. *HR was adjusted for age, gender and high-risk cytogenetics

Discussion

This Chapter demonstrated differences in clinical features such as age of diagnosis, Ig subtype, previous diagnosis of MM precursors, and cytogenetic abnormalities in Hispanics compared to NHW and NHB cases. Parallel to past findings, Hispanics had the youngest average age of diagnosis (57.4 years) followed by NHBs (57.5 years) and NHWs (61.2 years).

Moreover, we discovered that mutations characteristically associated with poor prognosis (del 13, del 17p, t(4;14), t(11;14), t(14;16), t(14;20)), occurred at a varying rates in our Hispanic, NHW, and NHB patients. The NHB cases had the lowest frequency of these mutations, followed by Hispanics and NHW patients. We also found that Hispanics with > 40% European ancestry were at increased risk of these mutations compared to those with < 40% European genetic ancestry. Studies have shown some level of association with high-risk cytogenetic abnormalities and European ancestry, but this is the first study that provided a comparative group, Hispanics with a varying fraction of European ancestry, to determine if European ancestry is indeed a contributing factor of high-risk cytogenetic abnormalities.

We currently do not have a comparison for MM cytogenetic data in Hispanics. However, when investigating the chromosomal changes of acute lymphoblastic leukemia (ALL) a B-cell malignancy closely related to MM, the Philadelphia chromosome (Ph+), i.e., t(9;22) translocation mutations associated with favorable prognosis(125), is found less commonly in Hispanics than NHW patients(126, 127). ALL is relatively well studied in Hispanics due to the high incidence rate in Hispanic children(126, 128). Unfortunately, research on mutational changes of B-cell malignancies, including MM, remains limited in the Hispanic population. Here, we

present a meaningful stepping stone for understanding the genetics of MM development in subjects of diverse backgrounds and how that affects their outcomes for personalized and effective disease management.

We also reported that the prior reported diagnosis rate of MGUS/SMM was lower in Hispanics (8.4%) compared to NHW (18.8%) and NHB (16.1%) patients. Literature shows evidence of the differential prevalence of MM premalignancy by ancestral background. Our findings somewhat parallel to Landgren's group results, which described a drop in the prevalence rate of MGUS/SMM after the age of 70 in Mexican Hispanics when compared to NHW and NHB patients(37).

Furthermore, one study reported MGUS prevalence to be 2.4% in Mexico residents compared to the estimated 3% prevalence rate in Caucasians(129), pointing to a potential lower incidence rate of MGUS and SMM in Hispanics. In contrast, another of Landgren's population based study revealed a prevalence rate of MGUS in Mexican Americans aged 10-49 to be almost double that those of their white counterparts, but lower than blacks in the same aged group(130).

The lack of routine screening of MM precursors makes it challenging to ascertain the "true" rate of MGUS and SMM in Hispanics or other ethnic groups. Nevertheless, with additional studies, the rate of MM progression may be better understood for improved observations of patients with the pre-cancerous diagnosis. However, we did not detect an ancestral association between MGUS/SMM prevalence and genetic ancestry in Hispanics.

Moreover, we identified differences in diagnostic markers like hemoglobin level and creatinine levels between Hispanics and NHBs. Compared to Hispanics, NHB patients exhibited a significantly lower median hemoglobin level of 10.1 g/dL compared to Hispanics (11.4g/dL). Similarly, creatinine levels were slightly elevated in NHB cases

compared to Hispanics. Likewise, we found some favorable staging factors in Hispanics such as lower levels of median beta-2-microglobulin (P = 0.034) and serum M spike (P = 0.020) when compared to NHBs. However, these variations did not translate to differences by genetic ancestry (data not shown). Levels of diagnostic and prognostic markers did not show significant variation between self-identified Hispanics and NHW patients (**Figure 12** and **Figure 13**).

To our knowledge, the first baseline clinical characteristics in U.S. Hispanic patient population were described in 2017 abstract at ASH by Tania et al(49) using NHW patients as a comparison group. We take a step further by adding NHB patients in our comparative analysis for inclusive reporting. We found the significant differences in baseline biomarkers were mainly between Hispanics and NHB patients, with Hispanic cases showing favorable levels of diagnostic and prognostic markers. We are limited in our sample size, but we confirm the varied clinical characteristics, such as slightly elevated hemoglobin levels in Hispanics at diagnosis. In contrast, median M spike levels were not significantly different in Hispanics compared to NHW patients but were significantly lower than that of NHBs. These clinical feature variations between ethnicities warrant continued exploration in larger populations and validation in other institutes and public datasets.

Lastly, after adjusting for MM prevalence, age, gender, and poor prognosis cytogenetics, we found no significant differences between overall survival and ethnicity/genetic ancestry. Hispanics have been shown to have adverse MM survival compared to whites. Evidence attributes this disparity to lagged initiation of therapy(47) as well as reduced utilization of ASCT(131) and novel treatments(132) in U.S. minorities. Our study may support this attribution as most of our patients received their therapy from one specialized institute. In parallel, there is exponential growth of MM

experimental treatment advancements, but there is an underrepresentation of minorities in clinical trial participation compared to NHW patients(133). Therefore, advocating for awareness of treatment disparity by race/ethnicity may reduce the poorer survival trends in Hispanics and other minority patients.

In conclusion, we presented unique disease characteristics in self-reported Hispanics and genetic ancestry, which may provide meaningful and timely information for a systematic evaluation of MM disease in the fastest sub-population in the United States.

Chapter 5: Findings Summary, Discussion, and Suggestions for Further Research

Findings Summary and Discussion

This study identified genetic and clinical contributors of MM susceptibility and outcomes and addressed the mediating factors of racial/ethnic disparities of this disease in a diverse study population. Our study also addressed gaps in our knowledge of MM disease profile in Hispanics.

Chapter 2 aimed to identify the genetic mediators of MM susceptibility within the Wnt/beta-catenin pathway. We found seven variants associated with MM risk in non-Hispanic whites in the discovery population, of which LRP6:rs7966410 and LRP6:rs7956971 remained protective of MM risk in the internal and external populations. Rs7966410 and rs7956971 also tagged causal variants with potential regulatory effects in known genes associated with MM development, such as *DKK1* and *Myc*.

Furthermore, by performing cross-ethnic comparisons of candidate variants associated with MM risk, we identified two variants, *CSNK1D*:rs9901910 and *BTRC*:rs7916830, that replicated in the non-Hispanic black and Hispanic patient populations. *CSNK1D*:rs9901910 was found to be a consistent risk locus among non-Hispanics white (OR: 2.40; 95% CI: 1.67-3.45; $P = 2.43 \times 10^{-6}$), non-Hispanic black (OR: 6.42; 95% CI: 2.47-16.7; $P = 3.14 \times 10^{-4}$), and Hispanic (OR: 4.31; 95% CI: 1.83-10.1; $P = 8.10 \times 10^{-4}$) patients. *BTRC*:rs7916830 was associated with a 37% and 21% reduced risk of MM in the non-Hispanic white and non-Hispanic black populations, respectively, indicating differences in MM genetic etiology by race/ethnicity.

The biological inference of candidate variants through gene regulations are described in detail in Chapter 2 using *in silico* tools. However, additional studies, such

as the downstream cellular effects of identified variants in MM cells, is needed to evaluate the functional consequences. For instance, transcriptional regulation of essential genes modulated by the Wnt/beta-catenin pathway, such as Myc and cyclin D1 via candidate causal variants on *LRP6*, may be good candidates for exploring biological mechanism.

We are limited with the small sample size of Hispanic and NHB subjects to form definitive conclusions on our findings' significance. Therefore, additional analysis on a larger patient population is needed for validation of our results. However, MM genetic association scans are held primarily in European descents, and to our knowledge, there are no MM germline genetic studies in Hispanics(134). Our study provides a much-needed contribution to genetic research that encapsulates diverse patient populations in identifying MM risk loci.

In Chapter 3, we estimated the proportions of genetic ancestry in our study populations. Using these estimations, we found, for the first time, suggestive evidence of a 12% protective effect of Amerindian genetic ancestry in MM susceptibility. However, this discovery requires external validation on a larger Hispanic study population. Moreover, our findings revealed a significant increase of MM risk for every 10% increase of African ancestry among our NHB study groups, confirming the relationship between excessive MM incidence and genetic ancestry in individuals of African descent.

The average age diagnosis in our Hispanic study subjects (57.4 years) compared to NHBs (57.5 years) and NHWs (61.0 years), parallel past studies showing early disease onset in this subgroup(135-137). Prior cases of MM precursors were also the lowest in Hispanics, compared to NHW and NHB patients. Moreover, high-risk cytogenetic abnormalities were more common in NHWs and Hispanics than NHBs.

Furthermore, Hispanic patients with European ancestry of > 40% had a 3-fold increased risk of carrying high-risk cytogenetic abnormalities than those with < 40% of European ancestry.

Our study population did indicate some protective diagnostic and prognostic biomarker levels of hemoglobin, creatinine, beta-2-microglobulin, and serum M spikes in Hispanics, compared to NHBs; but no significant differences between Hispanics and NHW patients. Moreover, we did not find significant differences in prognosis by self-reported ethnicity or genetic lineage, suggesting that the reported adverse survival in Hispanics may not be influenced by biology but instead by treatment access(47).

When investigating genetic heritage and disease/phenotype risk, it is crucial not to lump Hispanic/Latino populations as one admixed group due to their highly diverse genetic, continental, and regional backgrounds. The Hispanic/Latino lineage in the present-day Americas (including the U.S.), the Caribbean, and Mexico is a varying combination of European, African, Amerindian, and some East Asian ancestry depending on the history of population mixture between Indigenous Americans (Amerindians), Africans brought to the Americas and the Caribbean through the transatlantic slave trade, as well as European and East Asian settlers(138, 139).

Several studies have presented the differential ancestral population structure of Hispanics using AIMs from reference populations that harbor genetic loci with varying allele frequencies due to geographic isolation(71, 140–142). For instance, Wang et al., using reference population from Africa, Europe, and East Asia, found admixed Puerto Rican individuals to have a high European ancestral proportion of over 70%, compared to Mexicans and Peruvians, that display lower European genetic ancestry of 44% and 46%, respectively. Also, Mexicans exhibit the influence of East Asian origin (32%)

compared to Peruvians (51%), which had Chinese-origin populations initially settled in the coastal valleys of Peru(119).

Similarly, Salzano et al. illustrates the predominantly African influence in the Caribbean Latino population, i.e., Jamaicans (78% - 82%), and Haitians (96%). In contrast we find more European/Spanish influence in countries like Cuba (73% - 86%) and Puerto Rico (60% - 76%) and a higher prevalence of Amerindian ancestry in Guatemala (53%) and Mexico (51% - 56%). Mexican lineage also harbors some European ancestry (40% - 45%) and a small share of African ancestral influence (2% - 5%)(119), demonstrating the vast genetic diversity of European, Amerindian, and African ancestry within Hispanics and the Latin American microcosm.

Additionally, there is variability in the Hispanic/Latino lineage by regions in the United States. Bryc and colleagues published a paper demonstrating in great detail the admixture trends of Hispanics and Europeans that display high variability in ancestral percentages based on recent migration patterns within the United States(139). They reported the highest percentage of Native American/Amerindian ancestry in self-reported Latinos from Southwest states, especially those bordering Mexico, mirroring the Amerindian legacy in the area and the recent immigration trends through the Southwest border. Interestingly, they found a high percentage of African ancestry (20%) in self-identified Latinos living in southern states like Louisiana, Georgia, and North Carolina, and also states further north like New York and Pennsylvania. The study also highlighted the prevalence of European ancestry in self-reported Hispanics residing in states like Florida, Kentucky, and Tennessee(139).

These differences in population structure across the United States is also reflected in the concentration of Hispanic or Latino population in different regions of the U.S. According to the 2010 U.S. Census Briefs, over half of the Mexican origin

population reside in California and Texas alone. Salvadorians make up most of the Hispanic/Latino population in Maryland and the District of Colombia(104). Furthermore, Hispanics with countries of origin from the Dominican Republic and Puerto Ricans were more likely to reside in the Northeast, whereas Cubans were more likely to live in the South. "More than three-quarters of the Cuban population (77 percent) resided in the South, more than three-quarters of Dominicans (78 percent) resided in the Northeast, and more than half of the Puerto Rican population (53 percent) lived in the Northeast"(104). This may explain the higher proportion of European ancestry in states like Florida through residents of Cuban origins and Northeastern states' African influence through concentration of self-reported Puerto-Ricans and Dominicans in the region.

Moreover, there is genetic diversity within the Mexican population alone. The Monero-Estrada group at USCF showed divergence in ancestry in indigenous Mexicans as well as Mexican-Americans in Los Angeles (MXL), stating, "Some groups [indigenous population in Mexico] were as differentiated as Europeans are from East Asians"(143). This underscores the nuances of the Hispanic/Latino ancestral diversity through geographic origin regions of residence and to take into consideration the vast heterogeneity of these groups when studying complex diseases.

In addition, cancer incidence rates also differ substantially across Latinos by residency and national origin. Cuban Latinos and Puerto Ricans residing in the U.S. report a higher incidence rate of colorectal and lung cancer than those living in their respective countries of origin and also compared to Mexican Americans(109), suggesting consideration of environmental factors when investigating the incidence rate of MM and other cancers by genetic ancestry. Therefore, it is essential to understand the fine-scale population structure and cultural as well as environmental background of

the Hispanic/Latino study subjects when conducting epidemiologic or biomedical studies.

Study Limitations and Suggestions for Further Investigation

In addition to the small sample size, this study has several limitations. Our patient population was collected from a specialized institute, echoed by the above-average 6.4 - 7 years median survival rates of our subjects. Although there have been continuous improvements in survival throughout the years, the median survival time in the general population is about five years. Therefore, validation is necessary for other multi-ethnic study populations in a larger and non-specialized center for generalizability findings.

As mentioned above, the Hispanic population is a highly diverse group with varying genetics by regions. Therefore, the inclusion of Hispanic/Latino residents in multiple states of the US and collaborative efforts with hospitals in Latin and Central American countries will provide a strong understanding of the genetic and clinical mediators of MM development in Hispanics. In addition, it would also be interesting to study differences in MM incidence and outcomes in Hispanics/Latinos with varying degree of African genetic ancestry.

Similarly, genome-wide association studies are useful to identify the genetic etiology of MM individuals of elevated Amerindian ancestry. Local admixture mapping studies have successfully identified multiple independent risk variants on 8q24 and found commonly in men of African descent to explain the high incidence rate of prostate cancer in black Americans(144, 145). Likewise, conducting local admixture

and fine-mapping studies in MM cases of Amerindian heritage may help identify chromosomal segments associated with reduced risk of MM in Hispanics.

Moreover, active and systematic recruitment of diverse group of subjects in MM investigative studies through community outreach and clinical trials would be very beneficial to narrow the gaps in our understanding of highly heterogenous MM and also the racial/ethnic disparities that exist in MM development.

Final Remarks

In conclusion, we performed the first inclusive, multi-ethnic comparison of MM disease characteristics and presented unreported clinical and genetic features of MM in Hispanics. Our study is applicable for clinical research addressing the racial/ethnic disparity associated with MM and providing a better understanding of disease in the Hispanic cancer populations.

Bibliography

1. Siegel, R. L., K. D. Miller, and A. Jemal. 2019. Cancer statistics, 2019. *CA. Cancer J. Clin.* 69(1): 7-34.

2. Kazandjian, D. 2016. Multiple myeloma epidemiology and survival: A unique malignancy. *Semin. Oncol.* 43(6): 676-681.

3. Kvam, A. K., and A. Waage. 2015. Health-related quality of life in patients with multiple myeloma-does it matter? *Haematologica.* 100(6): 704-705.

4. Howlader, N., A.M. Noone, M. Krapcho, D. Miller, A. Brest, M. Yu, J. Ruhl, Z . Tatalovich, A. Mariotto, D.R. Lewis, H.S. Chen, E.J. Feuer, and K.A. Cronin. 2019. SEER Cancer Statistics Review. 1975-2016. https://seer.cancer.gov/csr/1975_2016/

5. João, C., C. Costa, I. Coelho, M. J. Vergueiro, M. Ferreira, and M. G. da Silva. 2014. Long-term survival in multiple myeloma. *Clin. Case Reports.* 2(5): 173-179.

6. Wallin, A., and S. C. Larsson. 2011. Body mass index and risk of multiple myeloma: A meta-analysis of prospective studies. *Eur. J. Cancer*. 47(11): 1606-15

7. Brown, L. M., G. Gridley, L. M. Pottern, D. Baris, C. A. Swanson, D. T. Silverman, R. B. Hayes, R. S. Greenberg, G. M. Swanson, J. B. Schoenberg, A. G. Schwartz, and J. F. Fraumeni. 2001. Diet and nutrition as risk factors for multiple myeloma among blacks and whites in the United States. *Cancer Causes Control.* 12(2): 117-25.

8. Vlajinac, H. D., T. D. Pekmezović, B. J. Adanja, J. M. Marinković, M. S. Kanazir, N. D. Suvajdžć, and M. D. Čolović. 2003. Case-control study of multiple myeloma with special reference to diet as risk factor. *Neoplasma*. 50(1): 79-83.

9. Sergentanis, T. N., F. Zagouri, G. Tsilimidos, A. Tsagianni, M. Tseliou, M. A. Dimopoulos, and T. Psaltopoulou. 2015. Risk Factors for Multiple Myeloma: A Systematic Review of Meta-Analyses. *Clin. Lymphoma, Myeloma Leuk.* 15(10): 563-

77.e1-3.

10. Becker, N. 2011. Epidemiology of multiple myeloma. *Recent Results Cancer Res.* .

11. Alexander, D. D., P. J. Mink, H. O. Adami, P. Cole, J. S. Mandel, M. M. Oken, and D. Trichopoulos. 2007. Multiple myeloma: A review of the epidemiologic literature. *Int. J. Cancer*. 12: 40-61.

12. Landgren, O., R. Zeig-Owens, O. Giricz, D. Goldfarb, K. Murata, K. Thoren, L. Ramanathan, M. Hultcrantz, A. Dogan, G. Nwankwo, U. Steidl, K. Pradhan, C. B. Hall, H. W. Cohen, N. Jaber, T. Schwartz, L. Crowley, M. Crane, S. Irby, M. P. Webber, A. Verma, and D. J. Prezant. 2018. Multiple myeloma and its precursor disease among firefighters exposed to theworld trade center disaster. *JAMA Oncol.* 4(6):821-827

13. Went, M., A. Sud, A. Försti, B. M. Halvarsson, N. Weinhold, S. Kimber, M. van Duin, G. Thorleifsson, A. Holroyd, D. C. Johnson, N. Li, G. Orlando, P. J. Law, M. Ali, B. Chen, J. S. Mitchell, D. F. Gudbjartsson, R. Kuiper, O. W. Stephens, U. Bertsch, P. Broderick, C. Campo, O. R. Bandapalli, H. Einsele, W. A. Gregory, U. Gullberg, J. Hillengass, P. Hoffmann, G. H. Jackson, K. H. Jöckel, E. Johnsson, S. Y. Kristinsson, U. H. Mellqvist, H. Nahi, D. Easton, P. Pharoah, A. Dunning, J. Peto, F. Canzian, A. Swerdlow, R. A. Eeles, Zs. S. Kote-Jarai, K. Muir, N. Pashayan, J. Nickel, M. M. Nöthen, T. Rafnar, F. M. Ross, M. I. da Silva Filho, H. Thomsen, I. Turesson, A. Vangsted, N. F. Andersen, A. Waage, B. A. Walker, A. K. Wihlborg, A. Broyl, F. E. Davies, U. Thorsteinsdottir, C. Langer, M. Hansson, H. Goldschmidt, M. Kaiser, P. Sonneveld, K. Stefansson, G. J. Morgan, K. Hemminki, B. Nilsson, R. S. Houlston, B. E. Henderson, C. A. Haiman, S. Benlloch, F. R. Schumacher, A. A. Al Olama, S. I. Berndt, D. V. Conti, F. Wiklund, S. Chanock, V. L. Stevens, C. M. Tangen, J. Batra, J. Clements, H. Gronberg, J. Schleutker, D. Albanes, S. Weinstein, A. Wolk, C. West, L. Mucci, G. Cancel-Tassin, S. Koutros, K. D. Sorensen, E. M. Grindedal, D. E. Neal, F.

C. Hamdy, J. L. Donovan, R. C. Travis, R. J. Hamilton, S. A. Ingles, B. Rosenstein, Y. J. Lu, G. G. Giles, A. S. Kibel, A. Vega, M. Kogevinas, K. L. Penney, J. Y. Park, J. L. Stanford, C. Cybulski, B. G. Nordestgaard, H. Brenner, C. Maier, J. Kim, E. M. John, M. R. Teixeira, S. L. Neuhausen, K. De Ruyck, A. Razack, L. F. Newcomb, D. Lessel, R. Kaneva, N. Usmani, F. Claessens, P. A. Townsend, M. G. Dominguez, M. J. Roobol, F. Menegaux, K. T. Khaw, L. Cannon-Albright, H. Pandha, and S. N. Thibodeau. 2018. Identification of multiple risk loci and regulatory mechanisms influencing susceptibility to multiple myeloma. *Nat. Commun.* 9(1): 3707.

14. Mitchell, J. S., N. Li, N. Weinhold, A. Försti, M. Ali, M. van Duin, G. Thorleifsson, D. C. Johnson, B. Chen, B.-M. Halvarsson, D. F. Gudbjartsson, R. Kuiper, O. W. Stephens, U. Bertsch, P. Broderick, C. Campo, H. Einsele, W. A. Gregory, U. Gullberg, M. Henrion, J. Hillengass, P. Hoffmann, G. H. Jackson, E. Johnsson, M. Jöud, S. Y. Kristinsson, S. Lenhoff, O. Lenive, U.-H. Mellqvist, G. Migliorini, H. Nahi, S. Nelander, J. Nickel, M. M. Nöthen, T. Rafnar, F. M. Ross, M. I. da Silva Filho, B. Swaminathan, H. Thomsen, I. Turesson, A. Vangsted, U. Vogel, A. Waage, B. A. Walker, A.-K. Wihlborg, A. Broyl, F. E. Davies, U. Thorsteinsdottir, C. Langer, M. Hansson, M. Kaiser, P. Sonneveld, K. Stefansson, G. J. Morgan, H. Goldschmidt, K. Hemminki, B. Nilsson, and R. S. Houlston. 2016. Genome-wide association study identifies multiple susceptibility loci for multiple myeloma. *Nat. Commun.* 7: 12050.

15. Martino, A., D. Campa, K. Jamroziak, R. M. Reis, J. Sainz, G. Buda, R. García-Sanz, F. Lesueur, H. Marques, V. Moreno, M. Jurado, R. Ríos, Z. Szemraj-Rogucka, J. Szemraj, A. Tjønneland, K. Overvad, A. J. Vangsted, U. Vogel, G. Mikala, K. Kádár, G. Szombath, J. Varkonyi, E. Orciuolo, C. Dumontet, F. Gemignani, A. M. Rossi, S. Landi, M. Petrini, R. S. Houlston, K. Hemminki, and F. Canzian. 2012. Impact of polymorphic variation at 7p15.3, 3p22.1 and 2p23.3 loci on risk of multiple myeloma. *Br. J.*

Haematol. 158: 805-809.

16. Broderick, P., D. Chubb, D. C. Johnson, N. Weinhold, A. Försti, A. Lloyd, B. Olver, Y. P. Ma, S. E. Dobbins, B. A. Walker, F. E. Davies, W. A. Gregory, J. A. Child, F. M. Ross, G. H. Jackson, K. Neben, A. Jauch, P. Hoffmann, T. W. Mühleisen, M. M. Nöthen, S. Moebus, I. P. Tomlinson, H. Goldschmidt, K. Hemminki, G. J. Morgan, and R. S. Houlston. 2011. Common variation at 3p22.1 and 7p15.3 influences multiple myeloma risk. *Nat. Genet.* 44: 58-61.

17. Johnson, D. C., N. Weinhold, J. S. Mitchell, B. Chen, M. Kaiser, D. B. Begum, J. Hillengass, U. Bertsch, W. A. Gregory, D. Cairns, G. H. Jackson, A. Försti, J. Nickel, P. Hoffmann, M. M. Nöethen, O. W. Stephens, B. Barlogie, F. E. Davis, K. Hemminki, H. Goldschmidt, R. S. Houlston, and G. J. Morgan. 2016. Genome-wide association study identifies variation at 6q25.1 associated with survival in multiple myeloma. *Nat. Commun.* 7: 10290.

18. Du, Z., N. Weinhold, G. C. Song, K. A. Rand, D. J. Van Den Berg, A. E. Hwang, X. Sheng, V. Hom, S. Ailawadhi, A. K. Nooka, S. Singhal, K. Pawlish, E. S. Peters, C. Bock, A. Mohrbacher, A. Stram, S. I. Berndt, W. J. Blot, G. Casey, V. L. Stevens, R. Kittles, P. J. Goodman, W. R. Diver, A. Hennis, B. Nemesure, E. A. Klein, B. A. Rybicki, J. L. Stanford, J. S. Witte, L. Signorello, E. M. John, L. Bernstein, A. M. Stroup, O. W. Stephens, M. Zangari, F. Van Rhee, A. Olshan, W. Zheng, J. J. Hu, R. Ziegler, S. J. Nyante, S. A. Ingles, M. F. Press, J. D. Carpten, S. J. Chanock, J. Mehta, G. A. Colditz, J. Wolf, T. G. Martin, M. Tomasson, M. A. Fiala, H. Terebelo, N. Janakiraman, L. Kolonel, K. C. Anderson, L. Le Marchand, D. Auclair, B. C. H. Chiu, E. Ziv, D. Stram, R. Vij, L. Bernal-Mizrachi, G. J. Morgan, J. A. Zonder, C. A. Huff, S. Lonial, R. Z. Orlowski, D. V. Conti, C. A. Haiman, and W. Cozen. 2020. A meta-analysis of genome-wide association studies of multiple myeloma among men and women of African ancestry.

Blood Adv. 4(1): 181-190.

19. Rand, K. A., C. Song, E. Dean, D. J. Serie, K. Curtin, X. Sheng, D. Hu, C. A. Huff, L. Bernal-Mizrachi, H. M. Tomasson, S. Ailawadhi, S. Singhal, K. Pawlish, E. S. Peters, C. H. Bock, A. Stram, D. J. Van Den Berg, C. K. Edlund, D. V. Conti, T. Zimmerman, A. E. Hwang, S. Huntsman, J. Graff, A. Nooka, Y. Kong, S. L. Pregja, S. I. Berndt, W. J. Blot, J. Carpten, G. Casey, L. Chu, W. R. Diver, V. L. Stevens, M. R. Lieber, P. J. Goodman, A. J. M. Hennis, A. W. Hsing, J. Mehta, R. A. Kittles, S. Kolb, E. A. Klein, C. Leske, A. B. Murphy, B. Nemesure, C. Neslund-Dudas, S. S. Strom, R. Vij, B. A. Rybicki, J. L. Stanford, L. B. Signorello, J. S. Witte, C. B. Ambrosone, P. Bhatti, E. M. John, L. Bernstein, W. Zheng, A. F. Olshan, J. J. Hu, R. G. Ziegler, S. J. Nyante, E. V. Bandera, B. M. Birmann, S. A. Ingles, M. F. Press, D. Atanackovic, M. J. Glenn, L. A. Cannon-Albright, B. Jones, G. Tricot, T. G. Martin, S. K. Kumar, J. L. Wolf, S. L. Deming Halverson, N. Rothman, A. R. Brooks-Wilson, S. V. Rajkumar, L. N. Kolonel, S. J. Chanock, S. L. Slager, R. K. Severson, N. Janakiraman, H. R. Terebelo, E. E. Brown, A. J. De Roos, A. F. Mohrbacher, G. A. Colditz, G. G. Giles, J. J. Spinelli, B. C. Chiu, N. C. Munshi, K. C. Anderson, J. Levy, J. A. Zonder, R. Z. Orlowski, S. Lonial, N. J. Camp, C. M. Vachon, E. Ziv, D. O. Stram, D. J. Hazelett, C. A. Haiman, and W. Cozen. 2016. A meta-analysis of multiple myeloma risk regions in African and European ancestry populations identifies putatively functional loci. *Cancer Epidemiol. Biomarkers Prev.* 25(12): 1609-1618.

20. Morgan, G. J., D. C. Johnson, N. Weinhold, H. Goldschmidt, O. Landgren, H. T. Lynch, K. Hemminki, and R. S. Houlston. 2014. Inherited genetic susceptibility to multiple myeloma. *Leukemia*. 4(4): 291-297.

21. Kumar, S. K., V. Rajkumar, R. A. Kyle, M. Van Duin, P. Sonneveld, M. V. Mateos, F. Gay, and K. C. Anderson. 2017. Multiple myeloma. *Nat. Rev. Dis. Prim.* 3:17046.

22. Landgren, O., G. Gridley, I. Turesson, N. E. Caporaso, L. R. Goldin, D. Baris, T. R. Fears, R. N. Hoover, and M. S. Linet. 2006. Risk of monoclonal gammopathy of undetermined significance (MGUS) and subsequent multiple myeloma among African American and white veterans in the United States. *Blood* 107: 904-906.

23. Kyle, R. A., E. D. Remstein, T. M. Therneau, A. Dispenzieri, P. J. Kurtin, J. M. Hodnefield, D. R. Larson, M. F. Plevak, D. F. Jelinek, R. Fonseca, L. J. Melton, and S. V. Rajkumar. 2007. Clinical Course and Prognosis of Smoldering (Asymptomatic) Multiple Myeloma. *N. Engl. J. Med.* 356: 2582–2590.

24. Kyle, R. A., T. M. Therneau, S. V. Rajkumar, D. R. Larson, M. F. Plevak, J. R. Offord, A. Dispenzieri, J. A. Katzmann, and L. J. Melton. 2006. Prevalence of Monoclonal Gammopathy of Undetermined Significance. *N. Engl. J. Med.* 354: 1362–1369.

25. Ravindran, A., A. C. Bartley, S. J. Holton, W. I. Gonsalves, P. Kapoor, M. A. Siddiqui, S. K. Hashmi, A. L. Marshall, A. A. Ashrani, A. DIspenzieri, R. A. Kyle, S. V. Rajkumar, and R. S. Go. 2016. Prevalence, incidence and survival of smoldering multiple myeloma in the United States. *Blood Cancer J.* 6(10): e486.

26. Muchtar, E., S. K. Kumar, H. Magen, and M. A. Gertz. 2018. Diagnosis and management of smoldering multiple myeloma: the razor's edge between clonality and cancer. *Leuk. Lymphoma.* 59(2): 288-299.

27. Rajkumar, S. V., O. Landgren, and M. V. Mateos. 2015. Smoldering multiple myeloma. *Blood*. 125(20): 3069-3075.

28. Rajan, A. M., and S. V. Rajkumar. 2015. Interpretation of cytogenetic results in multiple myeloma for clinical practice. *Blood Cancer J.* 5(10): e365.

29. Neben, K., A. Jauch, T. Hielscher, J. Hillengass, N. Lehners, A. Seckinger, M. Granzow, M. S. Raab, A. D. Ho, H. Goldschmidt, and D. Hose. 2013. Progression in

smoldering myeloma is independently determined by the chromosomal abnormalities del(17p), t(4;14), gain 1q, hyperdiploidy, and tumor load. *J. Clin. Oncol.* 31(34): 4325-32.

30. Palumbo, A., H. Avet-Loiseau, S. Oliva, H. M. Lokhorst, H. Goldschmidt, L. Rosinol, P. Richardson, S. Caltagirone, J. J. Lahuerta, T. Facon, S. Bringhen, F. Gay, M. Attal, R. Passera, A. Spencer, M. Offidani, S. Kumar, P. Musto, S. Lonial, M. T. Petrucci, R. Z. Orlowski, E. Zamagni, G. Morgan, M. A. Dimopoulos, B. G. M. Durie, K. C. Anderson, P. Sonneveld, J. S. Miguel, M. Cavo, S. V. Rajkumar, and P. Moreau. 2015. Revised international staging system for multiple myeloma: A report from international myeloma working group. *J. Clin. Oncol.* 33: 2863-2869.

31. Kumar, S., R. Fonseca, R. P. Ketterling, A. Dispenzieri, M. Q. Lacy, M. A. Gertz, S. R. Hayman, F. K. Buadi, D. Dingli, R. A. Knudson, A. Greenberg, S. J. Russell, S. R. Zeldenrust, J. A. Lust, R. A. Kyle, L. Bergsagel, and S. V. Rajkumar. 2012. Trisomies in multiple myeloma: Impact on survival in patients with high-risk cytogenetics. *Blood* .

32. Morgan, G. J., B. A. Walker, and F. E. Davies. 2012. The genetic architecture of multiple myeloma. *Nat. Rev. Cancer*. 12(5): 335-48

33. Agarwal, A., and I. M. Ghobrial. 2013. Monoclonal gammopathy of undetermined significance and smoldering multiple myeloma: A review of the current understanding of epidemiology, biology, risk stratification, and management of myeloma precursor disease. *Clin. Cancer Res.* 19(5): 985-994.

34. Bergsagel, P. L., W. M. Kuehl, F. Zhan, J. Sawyer, B. Barlogie, and J. Shaughnessy. 2005. Cyclin D dysregulation: An early and unifying pathogenic event in multiple myeloma. *Blood. 106(1): 296-303*.

35. Fonseca, R., R. J. Bailey, G. J. Ahmann, S. Vincent Rajkumar, J. D. Hoyer, J. A. Lust, R. A. Kyle, M. A. Gertz, P. R. Greipp, and G. W. Dewald. 2002. Genomic

abnormalities in monoclonal gammopathy of undetermined significance. *Blood*. 100(4):1417-24.

36. Kumar, S. K., A. Dispenzieri, M. Q. Lacy, M. A. Gertz, F. K. Buadi, S. Pandey, P. Kapoor, D. Dingli, S. R. Hayman, N. Leung, J. Lust, A. McCurdy, S. J. Russell, S. R. Zeldenrust, R. A. Kyle, and S. V. Rajkumar. 2014. Continued improvement in survival in multiple myeloma: Changes in early mortality and outcomes in older patients. *Leukemia*. 28(5): 1122-1128.

37. Landgren, O., B. I. Graubard, J. A. Katzmann, R. A. Kyle, I. Ahmadizadeh, R. Clark, S. K. Kumar, A. Dispenzieri, A. J. Greenberg, T. M. Therneau, L. J. Melton, N. Caporaso, N. Korde, M. Roschewski, R. Costello, G. M. McQuillan, and S. V. Rajkumar. 2014. Racial disparities in the prevalence of monoclonal gammopathies: A population-based study of 12 482 persons from the national health and nutritional examination survey. *Leukemia*. 28(7): 1537-42.

38. Landgren, O., J. A. Katzmann, A. W. Hsing, R. M. Pfeiffer, R. A. Kyle, E. D. Yeboah, R. B. Biritwum, Y. Tettey, A. A. Adjei, D. R. Larson, A. Dispenzieri, L. J. Melton, L. R. Goldin, M. L. McMaster, N. E. Caporaso, and S. V. Rajkumar. 2007. Prevalence of Monoclonal Gammopathy of Undetermined Significance Among Men in Ghana. *Mayo Clin. Proc.* 82: 1468-1473.

39. Nossent, J. C., C. N. Winkel, and J. C. J. van Leeuwen. 1993. Multiple myeloma in the Afro-Caribbean population of Curaçao. *Neth. J. Med.* 43(5-6): 210-4.

40. Howlader, N., A.M. Noone, M. Krapcho, J.Garshell, N. Neyman, S.F. Altekruse, C.L. Kosary, M. Yu, J. Ruhl, Z. Tatalovich, H. Cho, A. Mariotto, D.R. Lewis, H.S. Chen, E.J. Feuer, K.A. Cronin. 2013. SEER Cancer Statistics Review 1975-2010. https://seer.cancer.gov/archive/csr/1975_2010/

41. Landgren, O., G. Gridley, I. Turesson, N. E. Caporaso, L. R. Goldin, D. Baris, T. R. Fears, R. N. Hoover, and M. S. Linet. 2006. Risk of monoclonal gammopathy of undetermined significance (MGUS) and subsequent multiple myeloma among African American and white veterans in the United States. *Blood*. 107(3): 904-906.

42. Kazandjian, D., E. Hill, M. Hultcrantz, E. H. Rustad, V. Yellapantula, T. Akhlaghi, N. Korde, S. Mailankody, A. Dew, E. Papaemmanuil, I. Maric, M. Kwok, and O. Landgren. 2019. Molecular underpinnings of clinical disparity patterns in African American vs. Caucasian American multiple myeloma patients. *Blood Cancer J. 9(2): 15*.

43. Baughn, L. B., K. Pearce, D. Larson, M. Y. Polley, E. Elhaik, M. Baird, C. Colby, J. Benson, Z. Li, Y. Asmann, T. Therneau, J. R. Cerhan, C. M. Vachon, A. K. Stewart, P. L. Bergsagel, A. Dispenzieri, S. Kumar, and S. V. Rajkumar. 2018. Differences in genomic abnormalities among African individuals with monoclonal gammopathies using calculated ancestry. *Blood Cancer J.* 8(10): 96.

44. Baker, A., E. Braggio, S. Jacobus, S. Jung, D. Larson, T. Therneau, A. Dispenzieri, S. A. Van Wier, G. Ahmann, J. Levy, L. Perkins, S. Kim, K. Henderson, D. Vesole, S. Vincent Rajkumar, D. F. Jelinek, J. Carpten, and R. Fonseca. 2013. Uncovering the biology of multiple myeloma among African Americans: A comprehensive genomics approach. *Blood* 121: 3147-3152.

45. Manojlovic, Z., A. Christofferson, W. S. Liang, J. Aldrich, M. Washington, S. Wong, D. Rohrer, S. Jewell, R. A. Kittles, M. Derome, D. Auclair, D. W. Craig, J. Keats, and J. D. Carpten. 2017. Comprehensive molecular profiling of 718 Multiple Myelomas reveals significant differences in mutation frequencies between African and European descent cases. *PLoS Genet.* 13(11): e1007087.

46. Waxman, A. J., P. J. Mink, S. S. Devesa, W. F. Anderson, B. M. Weiss, S. Y.

Kristinsson, K. A. McGlynn, and O. Landgren. 2010. Racial disparities in incidence and outcome in multiple myeloma: A population-based study. *Blood.* 116(25):5501-6.

47. Ailawadhi, S., K. Parikh, S. Abouzaid, Z. Zhou, W. Tang, Z. Clancy, C. Cheung, Z. Y. Zhou, and J. Xie. 2019. Racial disparities in treatment patterns and outcomes among patients with multiple myeloma: A SEER-Medicare analysis. *Blood Adv.* 3(20): 2986–2994.

48. Smiths C.J, S. Ambs, O. Landgren. 2018. Biological determinants of health disparities in multiple myeloma. *Blood Cancer J.* 8(9): 85.

49. Jain, T., R. Fonseca, R. Chen, R. Patel, P. Jani, V. Gonzalez De La Calle, Z. Meghji, J. E. Hoffman, A. J. Alencar, K. R. Kelly, V. Roy, T. Sher, A. A. Chanan-Khan, S. Ailawadhi. 2017. Racial Differences in Disease Characteristics: Understanding Multiple Myeloma in Hispanics. Racial differences in disease characteristics: Understanding multiple myeloma in hispanics. *Blood.* 130 (Supplement 1): 864

50. Ailawadhi, S., I. T. Aldoss, D. Yang, P. Razavi, W. Cozen, T. Sher, and A. Chanan-Khan. 2012. Outcome disparities in multiple myeloma: A SEER-based comparative analysis of ethnic subgroups. *Br. J. Haematol.* 158: 91-98.

51. Pulte, D., M. T. Redaniel, H. Brenner, L. Jansen, and M. Jeffreys. 2014. Recent improvement in survival of patients with multiple myeloma: Variation by ethnicity. *Leuk. Lymphoma.* 55(5): 1083-9.

52. Ailawadhi, S., R. D. Frank, P. Advani, A. Swaika, M. Temkit, R. Menghani, M. Sharma, Z. Meghji, S. Paulus, N. Khera, S. K. Hashmi, A. Paulus, T. S. Kakar, D. O. Hodge, D. T. Colibaseanu, M. R. Vizzini, V. Roy, G. Colon-Otero, and A. A. Chanan-Khan. 2017. Racial disparity in utilization of therapeutic modalities among multiple myeloma patients: a SEER-medicare analysis. *Cancer Med.* 6(12): 2876-2885.

53. Fulciniti M., C. D. R. 2013. Role of Wnt Signaling Pathways in Multiple Myeloma

Pathogenesis. In *Advances in Biology and Therapy of Multiple Myeloma* Biomedical and Life Sciences. 85-95.

54. Giuliani, N., F. Morandi, S. Tagliaferri, M. Lazzaretti, G. Donofrio, S. Bonomini, R. Sala, M. Mangoni, and V. Rizzoli. 2007. Production of Wnt inhibitors by myeloma cells: Potential effects on canonical Wnt pathway in the bone microenvironment. *Cancer Res.* 67: 7665-7674.

55. Edwards, C. M., J. R. Edwards, S. T. Lwin, J. Esparza, B. O. Oyajobi, B. McCluskey, S. Munoz, B. Grubbs, and G. R. Mundy. 2008. Increasing wnt signaling in the bone marrow microenvironment inhibits the development of myeloma bone disease and reduces tumor burden in bone in vivo. *Blood* 111: 2833-2842.

56. Derksen, P. W. B., E. Tjin, H. P. Meijer, M. D. Klok, H. D. MacGillavry, M. H. J. van Oers, H. M. Lokhorst, A. C. Bloem, H. Clevers, R. Nusse, R. van der Neut, M. Spaargaren, and S. T. Pals. 2004. Illegitimate WNT signaling promotes proliferation of multiple myeloma cells. *Proc. Natl. Acad. Sci. U. S. A.* 101: 6122-7.

57. Kikuchi, A., H. Yamamoto, A. Sato, and S. Matsumoto. 2011. New Insights into the Mechanism of Wnt Signaling Pathway Activation. *Int. Rev. Cell Mol. Biol.* 291: 21-71.

58. Reya, T., A. W. Duncan, L. Ailles, J. Domen, D. C. Scherer, K. Willert, L. Hintz, R. Nusse, and I. L. Weissman. 2003. A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* 423: 409-414.

59. Reya, T., and H. Clevers. 2005. Wnt signalling in stem cells and cancer. *Nature* 434: 843-850.

60. Rosenbluh, J., X. Wang, and W. C. Hahn. 2014. Genomic insights into WNT/β-catenin signaling. *Trends Pharmacol. Sci.* 35(2): 103-9.

61. Anastas, J. N., and R. T. Moon. 2013. WNT signalling pathways as therapeutic targets in cancer. *Nat. Rev. Cancer.* 13(1): 11-26.

62. King, T. D., M. J. Suto, and Y. Li. 2012. The wnt/β-catenin signaling pathway: A potential therapeutic target in the treatment of triple negative breast cancer. *J. Cell. Biochem.* 113(1): 13-8.

63. MacDonald, B. T., K. Tamai, and X. He. 2009. Wnt/β-Catenin Signaling: Components, Mechanisms, and Diseases. *Dev. Cell* 17: 9-26.

64. Moon, R. T., A. D. Kohn, G. V. De Ferrari, and A. Kaykas. 2004. WNT and β-catenin signalling: diseases and therapies. *Nat. Rev. Genet.* 5: 691-701.

65. Komiya, Y., and R. Habas. 2008. Wnt signal transduction pathways. *Organogenesis* 4: 68-75.

66. Ziv, E., E. Dean, D. Hu, A. Martino, D. Serie, K. Curtin, D. Campa, B. Aftab, P. Bracci, G. Buda, Y. Zhao, J. Caswell-Jin, R. Diasio, C. Dumontet, M. Dudziński, L. Fejerman, A. Greenberg, S. Huntsman, K. Jamroziak, A. Jurczyszyn, S. Kumar, D. Atanackovic, M. Glenn, L. A. Cannon-Albright, B. Jones, A. Lee, H. Marques, T. Martin, J. Martinez-Lopez, V. Rajkumar, J. Sainz, A. J. Vangsted, M. Wątek, J. Wolf, S. Slager, N. J. Camp, F. Canzian, and C. Vachon. 2015. Genome-wide association study identifies variants at 16p13 associated with survival in multiple myeloma patients. *Nat. Commun.* 6: 7539.

67. Wu, X., M. A. Hildebrandt, Y. Ye, W. H. Chow, J. Gu, S. Cunningham, H. Zhao, E. T. Hawk, E. Wagar, A. Rodriguez, and S. R. Hamilton. 2016. Cohort Profile: The MD Anderson Cancer Patients and Survivors Cohort (MDA-CPSC). *Int. J. Epidemiol.* 45(3): 713-713f.

68. Wu, X., Y. Ye, L. A. Kiemeney, P. Sulem, T. Rafnar, G. Matullo, D. Seminara, T. Yoshida, N. Saeki, A. S. Andrew, C. P. Dinney, B. Czerniak, Z. Zhang, A. E. Kiltie, D. T. Bishop, P. Vineis, S. Porru, F. Buntinx, E. Kellen, M. P. Zeegers, R. Kumar, P. Rudnai, E. Gurzau, K. Koppova, J. I. Mayordomo, M. Sanchez, B. Saez, A. Lindblom, P. de

Verdier, G. Steineck, G. B. Mills, A. Schned, S. Guarrera, S. Polidoro, S.-C. Chang, J. Lin, D. W. Chang, K. S. Hale, T. Majewski, H. B. Grossman, S. Thorlacius, U. Thorsteinsdottir, K. K. H. Aben, J. A. Witjes, K. Stefansson, C. I. Amos, M. R. Karagas, and J. Gu. 2009. Genetic variation in the prostate stem cell antigen gene PSCA confers susceptibility to urinary bladder cancer. *Nat. Genet.* 41: 991-995.

69. Kanehisa, M., and S. Goto. 2000. KEGG : Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acid Res.* 28(1): 27-30.

70. de Bakker, P. I. W., R. Yelensky, I. Pe'er, S. B. Gabriel, M. J. Daly, and D. Altshuler. 2005. Efficiency and power in genetic association studies. *Nat. Genet.* 37: 1217-1223.

71. Auton, A., G. R. Abecasis, D. M. Altshuler, R. M. Durbin, D. R. Bentley, A. Chakravarti, A. G. Clark, P. Donnelly, E. E. Eichler, P. Flicek, S. B. Gabriel, R. A. Gibbs, E. D. Green, M. E. Hurles, B. M. Knoppers, J. O. Korbel, E. S. Lander, C. Lee, H. Lehrach, E. R. Mardis, G. T. Marth, G. A. McVean, D. A. Nickerson, J. P. Schmidt, S. T. Sherry, J. Wang, R. K. Wilson, E. Boerwinkle, H. Doddapaneni, Y. Han, V. Korchina, C. Kovar, S. Lee, D. Muzny, J. G. Reid, Y. Zhu, Y. Chang, Q. Feng, X. Fang, X. Guo, M. Jian, H. Jiang, X. Jin, T. Lan, G. Li, J. Li, Y. Li, S. Liu, X. Liu, Y. Lu, X. Ma, M. Tang, B. Wang, G. Wang, H. Wu, R. Wu, X. Xu, Y. Yin, D. Zhang, W. Zhang, J. Zhao, M. Zhao, X. Zheng, N. Gupta, N. Gharani, L. H. Toji, N. P. Gerry, A. M. Resch, J. Barker, L. Clarke, L. Gil, S. E. Hunt, G. Kelman, E. Kulesha, R. Leinonen, W. M. McLaren, R. Radhakrishnan, A. Roa, D. Smirnov, R. E. Smith, I. Streeter, A. Thormann, I. Toneva, B. Vaughan, X. Zheng-Bradley, R. Grocock, S. Humphray, T. James, Z. Kingsbury, R. Sudbrak, M. W. Albrecht, V. S. Amstislavskiy, T. A. Borodina, M. Lienhard, F. Mertes, M. Sultan, B. Timmermann, M. L. Yaspo, L. Fulton, V. Ananiev, Z. Belaia, D. Beloslyudtsev, N. Bouk, C. Chen, D. Church, R. Cohen, C. Cook, J.

Garner, T. Hefferon, M. Kimelman, C. Liu, J. Lopez, P. Meric, C. O'Sullivan, Y. Ostapchuk, L. Phan, S. Ponomarov, V. Schneider, E. Shekhtman, K. Sirotkin, D. Slotta, H. Zhang, S. Balasubramaniam, J. Burton, P. Danecek, T. M. Keane, A. Kolb-Kokocinski, S. McCarthy, J. Stalker, M. Quail, C. J. Davies, J. Gollub, T. Webster, B. Wong, Y. Zhan, C. L. Campbell, Y. Kong, A. Marcketta, F. Yu, L. Antunes, M. Bainbridge, A. Sabo, Z. Huang, L. J. M. Coin, L. Fang, Q. Li, Z. Li, H. Lin, B. Liu, R. Luo, H. Shao, Y. Xie, C. Ye, C. Yu, F. Zhang, H. Zheng, H. Zhu, C. Alkan, E. Dal, F. Kahveci, E. P. Garrison, D. Kural, W. P. Lee, W. F. Leong, M. Stromberg, A. N. Ward, J. Wu, M. Zhang, M. J. Daly, M. A. DePristo, R. E. Handsaker, E. Banks, G. Bhatia, G. Del Angel, G. Genovese, H. Li, S. Kashin, S. A. McCarroll, J. C. Nemesh, R. E. Poplin, S. C. Yoon, J. Lihm, V. Makarov, S. Gottipati, A. Keinan, J. L. Rodriguez-Flores, T. Rausch, M. H. Fritz, A. M. Stütz, K. Beal, A. Datta, J. Herrero, G. R. S. Ritchie, D. Zerbino, P. C. Sabeti, I. Shlyakhter, S. F. Schaffner, J. Vitti, D. N. Cooper, E. V. Ball, P. D. Stenson, B. Barnes, M. Bauer, R. K. Cheetham, A. Cox, M. Eberle, S. Kahn, L. Murray, J. Peden, R. Shaw, E. E. Kenny, M. A. Batzer, M. K. Konkel, J. A. Walker, D. G. MacArthur, M. Lek, R. Herwig, L. Ding, D. C. Koboldt, D. Larson, K. Ye, S. Gravel, A. Swaroop, E. Chew, T. Lappalainen, Y. Erlich, M. Gymrek, T. F. Willems, J. T. Simpson, M. D. Shriver, J. A. Rosenfeld, C. D. Bustamante, S. B. Montgomery, F. M. De La Vega, J. K. Byrnes, A. W. Carroll, M. K. DeGorter, P. Lacroute, B. K. Maples, A. R. Martin, A. Moreno-Estrada, S. S. Shringarpure, F. Zakharia, E. Halperin, Y. Baran, E. Cerveira, J. Hwang, A. Malhotra, D. Plewczynski, K. Radew, M. Romanovitch, C. Zhang, F. C. L. Hyland, D. W. Craig, A. Christoforides, N. Homer, T. Izatt, A. A. Kurdoglu, S. A. Sinari, K. Squire, C. Xiao, J. Sebat, D. Antaki, M. Gujral, A. Noor, K. Ye, E. G. Burchard, R. D. Hernandez, C. R. Gignoux, D. Haussler, S. J. Katzman, W. J. Kent, B. Howie, A. Ruiz-Linares, E. T. Dermitzakis, S. E. Devine, H. M. Kang, J. M.

Kidd, T. Blackwell, S. Caron, W. Chen, S. Emery, L. Fritsche, C. Fuchsberger, G. Jun, B. Li, R. Lyons, C. Scheller, C. Sidore, S. Song, E. Sliwerska, D. Taliun, A. Tan, R. Welch, M. K. Wing, X. Zhan, P. Awadalla, A. Hodgkinson, Y. Li, X. Shi, A. Quitadamo, G. Lunter, J. L. Marchini, S. Myers, C. Churchhouse, O. Delaneau, A. Gupta-Hinch, W. Kretzschmar, Z. Iqbal, I. Mathieson, A. Menelaou, A. Rimmer, D. K. Xifara, T. K. Oleksyk, Y. Fu, X. Liu, M. Xiong, L. Jorde, D. Witherspoon, J. Xing, B. L. Browning, S. R. Browning, F. Hormozdiari, P. H. Sudmant, E. Khurana, C. Tyler-Smith, C. A. Albers, Q. Ayub, Y. Chen, V. Colonna, L. Jostins, K. Walter, Y. Xue, M. B. Gerstein, A. Abyzov, S. Balasubramanian, J. Chen, D. Clarke, Y. Fu, A. O. Harmanci, M. Jin, D. Lee, J. Liu, X. J. Mu, J. Zhang, Y. Zhang, C. Hartl, K. Shakir, J. Degenhardt, S. Meiers, B. Raeder, F. P. Casale, O. Stegle, E. W. Lameijer, I. Hall, V. Bafna, J. Michaelson, E. J. Gardner, R. E. Mills, G. Dayama, K. Chen, X. Fan, Z. Chong, T. Chen, M. J. Chaisson, J. Huddleston, M. Malig, B. J. Nelson, N. F. Parrish, B. Blackburne, S. J. Lindsay, Z. Ning, Y. Zhang, H. Lam, C. Sisu, D. Challis, U. S. Evani, J. Lu, U. Nagaswamy, J. Yu, W. Li, L. Habegger, H. Yu, F. Cunningham, I. Dunham, K. Lage, J. B. Jespersen, H. Horn, D. Kim, R. Desalle, A. Narechania, M. A. W. Sayres, F. L. Mendez, G. D. Poznik, P. A. Underhill, D. Mittelman, R. Banerjee, M. Cerezo, T. W. Fitzgerald, S. Louzada, A. Massaia, F. Yang, D. Kalra, W. Hale, X. Dan, K. C. Barnes, C. Beiswanger, H. Cai, H. Cao, B. Henn, D. Jones, J. S. Kaye, A. Kent, A. Kerasidou, R. Mathias, P. N. Ossorio, M. Parker, C. N. Rotimi, C. D. Royal, K. Sandoval, Y. Su, Z. Tian, S. Tishkoff, M. Via, Y. Wang, H. Yang, L. Yang, J. Zhu, W. Bodmer, G. Bedoya, Z. Cai, Y. Gao, J. Chu, L. Peltonen, A. Garcia-Montero, A. Orfao, J. Dutil, J. C. Martinez-Cruzado, R. A. Mathias, A. Hennis, H. Watson, C. McKenzie, F. Qadri, R. LaRocque, X. Deng, D. Asogun, O. Folarin, C. Happi, O. Omoniwa, M. Stremlau, R. Tariyal, M. Jallow, F. S. Joof, T. Corrah, K. Rockett, D. Kwiatkowski, J. Kooner, T. T. Hien, S. J. Dunstan, N. ThuyHang,

R. Fonnie, R. Garry, L. Kanneh, L. Moses, J. Schieffelin, D. S. Grant, C. Gallo, G. Poletti, D. Saleheen, A. Rasheed, L. D. Brooks, A. L. Felsenfeld, J. E. McEwen, Y. Vaydylevich, A. Duncanson, M. Dunn, and J. A. Schloss. 2015. A global reference for human genetic variation. *Nature.* 526(7571): 68-74.

72. Das, S., L. Forer, S. Schönherr, C. Sidore, A. E. Locke, A. Kwong, S. I. Vrieze, E. Y. Chew, S. Levy, M. McGue, D. Schlessinger, D. Stambolian, P. R. Loh, W. G. Iacono, A. Swaroop, L. J. Scott, F. Cucca, F. Kronenberg, M. Boehnke, G. R. Abecasis, and C. Fuchsberger. 2016. Next-generation genotype imputation service and methods. *Nat. Genet.* 48(10): 1284-1287.

73. Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81(3): 559-575.

74. James Kent, W., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and D. Haussler. 2002. The human genome browser at UCSC. *Genome Res.* .

75. Ward, L. D., and M. Kellis. 2016. HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* 44: D877-D881.

76. Boyle, A. P., E. L. Hong, M. Hariharan, Y. Cheng, M. A. Schaub, M. Kasowski, K. J. Karczewski, J. Park, B. C. Hitz, S. Weng, J. M. Cherry, and M. Snyder. 2012. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* 22(9): 1790-7.

77. Machiela, M. J., and S. J. Chanock. 2015. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics.* 31(21): 3555-3557.

78. Bo Mi Park, Eun Jin Kim, Hee Jin Nam, Dongdong Zhang, Chu Hyun Bae, Myeongmo Kang, Heeyoun Kim, Weontae Lee, Bjarne Bogen, and S.-K. L. author. 2017. Cyclized Oligopeptide Targeting LRP5/6-DKK1 Interaction Reduces the Growth of Tumor Burden in a Multiple Myeloma Mouse Model. *Yonsei Med. J.* 58: 505–513.

79. Tian, E., F. Zhan, R. Walker, E. Rasmussen, Y. Ma, B. Barlogie, and J. D. Shaughnessy. 2003. The Role of the Wnt-Signaling Antagonist DKK1 in the Development of Osteolytic Lesions in Multiple Myeloma. *N. Engl. J. Med.* 349: 2483-2494.

80. Kristensen, I. B., J. H. Christensen, M. B. Lyng, M. B. M??ller, L. Pedersen, L. M. Rasmussen, H. J. Ditzel, and N. Abildgaard. 2014. Expression of osteoblast and osteoclast regulatory genes in the bone marrow microenvironment in multiple myeloma: Only up-regulation of Wnt inhibitors SFRP3 and DKK1 is associated with lytic bone disease. *Leuk. Lymphoma* 55: 911-919.

81. Fowler, J. A., C. M. Edwards, and P. I. Croucher. 2011. Tumor-host cell interactions in the bone disease of myeloma. *Bone* 48: 121-128.

82. Fulciniti, M., P. Tassone, T. Hideshima, S. Vallet, P. Nanjappa, S. A. Ettenberg, Z. Shen, N. Patel, Y. T. Tai, D. Chauhan, C. Mitsiades, R. Prabhala, N. Raje, K. C. Anderson, D. R. Stover, and N. C. Munshi. 2009. Anti-DKK1 mAb (BHQ880) as a potential therapeutic agent for multiple myeloma. *Blood.* 114(2): 371-9.

83. Pozzi, S., M. Fulciniti, H. Yan, S. Vallet, H. Eda, K. Patel, L. Santo, D. Cirstea, T. Hideshima, L. Schirtzinge, S. Kuhstoss, K. C. Anderson, N. Munshi, D. Scadden, H. M. Kronenberg, and N. Raje. 2013. In vivo and in vitro effects of a novel anti-Dkk1 neutralizing antibody in multiple myeloma. *Bone.* 53(2): 487-96.

84. Edwards, C. M., J. Zhuang, and G. R. Mundy. 2008. The pathogenesis of the bone disease of multiple myeloma. *Bone.* 7: 33-42.

85. Krönke, J., N. D. Udeshi, A. Narla, P. Grauman, S. N. Hurst, M. McConkey, T. Svinkina, D. Heckl, E. Comer, X. Li, C. Ciarlo, E. Hartman, N. Munshi, M. Schenone, S. L. Schreiber, S. A. Carr, and B. L. Ebert. 2014. Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science.* 343(6168): 301-305.

86. Marke, R., F. N. Van Leeuwen, and B. Scheijen. 2018. The many faces of IKZF1 in B-cell precursor acute lymphoblastic leukemia. *Haematologica*. 103(4): 565-574.

87. Mullighan, C. G., C. B. Miller, I. Radtke, L. A. Phillips, J. Dalton, J. Ma, D. White, T. P. Hughes, M. M. Le Beau, C. H. Pui, M. V. Relling, S. A. Shurtleff, and J. R. Downing. 2008. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature.* 453(7191): 110-4.

88. Shaffer, A. L., N. C. T. Emre, L. Lamy, V. N. Ngo, G. Wright, W. Xiao, J. Powell, S. Dave, X. Yu, H. Zhao, Y. Zeng, B. Chen, J. Epstein, and L. M. Staudt. 2008. IRF4 addiction in multiple myeloma. *Nature.* 454(7201): 226-31.

89. Yang, Y., A. L. Shaffer, N. C. T. Emre, M. Ceribelli, M. Zhang, G. Wright, W. Xiao, J. Powell, J. Platig, H. Kohlhammer, R. M. Young, H. Zhao, Y. Yang, W. Xu, J. J. Buggy, S. Balasubramanian, L. A. Mathews, P. Shinn, R. Guha, M. Ferrer, C. Thomas, T. A. Waldmann, and L. M. Staudt. 2012. Exploiting Synthetic Lethality for the Therapy of ABC Diffuse Large B Cell Lymphoma. *Cancer Cell.* 21(6): 723-737.

90. Holien, T., and A. Sundan. 2015. Oncogene addiction to c-MYC in myeloma cells. *Oncotarget.* 3(8): 739–740.

91. Wang, X., E. L. Goode, Z. S. Fredericksen, R. A. Vierkant, V. S. Pankratz, W. Liu-Mares, D. N. Rider, C. M. Vachon, J. R. Cerhan, J. E. Olson, and F. J. Couch. 2008. Association of genetic variation in genes implicated in the β-catenin destruction complex with risk of breast cancer. *Cancer Epidemiol. Biomarkers Prev.* 17: 2101-

2108.

92. Karen L Burger, Mark B. Meads, Ariosto Siqueira Silva, Paula Oliveira, Tuan Nguyen, Allison Distler, Maria Coelho Siqueira Silva, Timothy Jacobson, Jamie Teer, Hongyue Dai, William Roush, J. C. and K. H. S. 2017. A CK1δ/CK1ε-to-β-Catenin Circuit Is a Therapeutic Vulnerability for Primary and Drug Resistant Multiple Myeloma. *Blood* 130: 591.

93. Manni, S., M. Carrino, M. Manzoni, K. Gianesin, S. C. Nunes, M. Costacurta, L. Q. Tubi, P. Macaccaro, E. Taiana, A. Cabrelle, G. Barila, A. Martines, R. Zambello, L. Bonaldi, L. Trentin, A. Neri, G. Semenzato, and F. Piazza. 2017. Inactivation of CK1alpha in multiple myeloma empowers drug cytotoxicity by affecting AKT and beta-catenin survival signaling pathways. *Oncotarget.* 8(9):14604-14619.

94. Golbabapour, S., N. A. Majid, P. Hassandarvish, M. Hajrezaie, M. A. Abdulla, and A. H. A. Hadi. 2013. Gene silencing and polycomb group proteins: An overview of their structure, mechanisms and phylogenetics. *Omi. A J. Integr. Biol.* 17(6): 283–29.

95. Zoabi, M., R. Sadeh, P. de Bie, V. E. Marquez, and A. Ciechanover. 2011. Corrigendum to “PRAJA1 is a ubiquitin ligase for the polycomb repressive complex 2 proteins”. *Biochem. Biophys. Res. Commun.* 410(3): 705.

96. Sahasrabuddhe AA, Chen X, Chung F, Velusamy T, Lim MS, E.-J. K. 2015. Oncogenic Y641 mutations in EZH2 prevent Jak2/β-TrCP-mediated degradation. *Oncogene* 34: 445-54.

97. Sharma R, Williams PJ, Gupta A, McCluskey B, Bhaskaran S, Muñoz S, O. B. 2015. A dominant-negative F-box deleted mutant of E3 ubiquitin ligase, β-TrCP1/FWD1, markedly reduces myeloma cell growth and survival in mice. *Oncotarget* 6: 21589-602.

98. Wu, X., Y. Ye, L. A. Kiemeney, P. Sulem, T. Rafnar, G. Matullo, D. Seminara, T.

Yoshida, N. Saeki, A. S. Andrew, C. P. Dinney, B. Czerniak, Z. F. Zhang, A. E. Kiltie, D. T. Bishop, P. Vineis, S. Porru, F. Buntinx, E. Kellen, M. P. Zeegers, R. Kumar, P. Rudnai, E. Gurzau, K. Koppova, J. I. Mayordomo, M. Sanchez, B. Saez, A. Lindblom, P. De Verdier, G. Steineck, G. B. Mills, A. Schned, S. Guarrera, S. Polidoro, S. C. Chang, J. Lin, D. W. Chang, K. S. Hale, T. Majewski, H. B. Grossman, S. Thorlacius, U. Thorsteinsdottir, K. K. H. Aben, J. A. Witjes, K. Stefansson, C. I. Amos, M. R. Karagas, and J. Gu. 2009. Genetic variation in the prostate stem cell antigen gene PSCA confers susceptibility to urinary bladder cancer. *Nat. Genet.* 41: 991-995.

99. Chow, W. H., M. Chrisman, C. R. Daniel, Y. Ye, H. Gomez, Q. Dong, C. E. Anderson, S. Chang, S. Strom, H. Zhao, and X. Wu. 2017. Cohort profile: The Mexican American Mano a Mano Cohort. *Int. J. Epidemiol.* 46(2): e3.

100. Alexander, D. H., and K. Lange. 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics.* 12: 246.

101. Abraham, G., Y. Qiu, and M. Inouye. 2017. FlashPCA2: principal component analysis of Biobank-scale genotype datasets. *Bioinformatics.* 33(17): 2776-2778.

102. Booth, D. S., H. Szmidt-Middleton, N. King, M. J. Westbrook, S. L. Young, A. Kuo, M. Abedin, J. Chapman, S. R. Fairclough, U. Hellsten, Y. Isogai, I. Letunic, M. Marr, D. Pincus, N. Putnam, A. Rokas, K. J. Wright, R. Zuzow, W. Dirks, M. Good, D. Goodstein, D. Lemons, W. W. Li, J. B. Lyons, A. Morris, S. A. Nichols, D. J. Richter, A. Salamov, J. Sequencing, P. Bork, W. A. Lim, G. Manning, W. T. Miller, W. McGinnis, H. Shapiro, R. Tjian, I. V. Grigoriev, D. Rokhsar, M. J. Dayel, R. A. Alegado, S. R. Fairclough, T. C. Levin, S. A. Nichols, K. L. McDonald, N. King, A. J. Greaney, L. A. Wetzel, N. King, D. J. Richter, P. Fozouni, M. B. Eisen, N. King, B. S. C. Leadbeater, M. Abedin, N. King, T. Brunet, N. King, E. Szathmáry, J. M. Smith, S. R. Fairclough, M. J. Dayel, N. King, L. A. Wetzel, T. C. Levin, R. E. Hulett, D. Chan, G. A. King, R.

Aldayafleh, D. S. Booth, M. A. Sigg, N. King, I. Ruiz-Trillo, A. J. Roger, G. Burger, M. W. Gray, B. F. Lang, K. L. McDonald, A. Apprill, S. McNally, R. Parsons, L. Weber, R Core Team, A. E. Parada, D. M. Needham, J. A. Fuhrman, W. W. Li, A. Godzik, and Rstudio Team. 2018. RStudio: Integrated Development for R.

103. Stata Press. 2019. Stata Statistical Software: Release 16. *StataCorp LLC*.

104. Ennis, S. R., M. Rios-Vargas, and N. G. Albert. 2011. *The Hispanic Population: 2010 Census Briefs*.

105. Baker, A., E. Braggio, S. Jacobus, S. Jung, D. Larson, T. Therneau, A. Dispenzieri, S. A. Van Wier, G. Ahmann, J. Levy, L. Perkins, S. Kim, K. Henderson, D. Vesole, S. Vincent Rajkumar, D. F. Jelinek, J. Carpten, and R. Fonseca. 2013. Uncovering the biology of multiple myeloma among African Americans: A comprehensive genomics approach. *Blood.* 121(16): 3147-52.

106. Smith, C. J., S. Ambs, and O. Landgren. 2018. Biological determinants of health disparities in multiple myeloma. *Blood Cancer J.* 8(9): 85.

107. Mills, P. K., R. Yang, and D. Riordan. 2005. Lymphohematopoietic cancers in the United Farm Workers of America (UFW), 1988-2001. *Cancer Causes Control.* 16(7):823-30.

108. Khuder, S. A., and A. B. Mutgi. 1997. Meta-analyses of multiple myeloma and farming. *Am. J. Ind. Med.* 2008. 3: 27.

109. Stern, M. C., L. Fejerman, R. Das, V. W. Setiawan, M. R. Cruz-Correa, E. J. Perez-Stable, and J. C. Figueiredo. 2016. Variability in Cancer Risk and Outcomes Within US Latinos by National Origin and Genetic Ancestry. *Curr. Epidemiol. Reports* .

110. Slattery, M. L., E. M. John, G. Torres-Mejia, A. Lundgreen, J. S. Herrick, K. B. Baumgartner, L. M. Hines, M. C. Stern, and R. K. Wolff. 2012. Genetic variation in genes involved in hormones, inflammation and energetic factors and breast cancer risk

in an admixed population. *Carcinogenesis.* 33(8): 1512-1521.

111. Marker, K. M., V. A. Zavala, T. Vidaurre, P. C. Lott, J. N. Vásquez, S. Casavilca-Zambrano, M. Calderón, J. E. Abugattas, H. L. Gómez, H. A. Fuentes, R. L. Picoaga, J. M. Cotrina, S. P. Neciosup, C. A. Castañeda, Z. Morante, F. Valencia, J. Torres, M. Echeverry, M. E. Bohórquez, G. Polanco-Echeverry, A. P. Estrada-Florez, S. J. Serrano-Gómez, J. A. Carmona-Valencia, I. Alvarado-Cabrero, M. C. Sanabria-Salas, A. Velez, J. Donado, S. Song, D. Cherry, L. I. Tamayo, S. Huntsman, D. Hu, R. Ruiz-Cordero, R. Balassanian, E. Ziv, J. Zabaleta, L. Carvajal-Carmona, and L. Fejerman. 2020. Human Epidermal Growth Factor Receptor 2–Positive Breast Cancer Is Associated with Indigenous American Ancestry in Latin American Women. *Cancer Res.* 80(9): 1893-1901.

112. Yang, J. J., C. Cheng, M. Devidas, X. Cao, Y. Fan, D. Campana, W. Yang, G. Neale, N. J. Cox, P. Scheet, M. J. Borowitz, N. J. Winick, P. L. Martin, C. L. Willman, W. P. Bowman, B. M. Camitta, A. Carroll, G. H. Reaman, W. L. Carroll, M. Loh, S. P. Hunger, C.-H. Pui, W. E. Evans, and M. V Relling. 2011. Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. *Nat. Genet.* 43: 237-41.

113. Perez-Andreu, V., K. G. Roberts, R. C. Harvey, W. Yang, C. Cheng, D. Pei, H. Xu, J. Gastier-Foster, E. Shuyu, J. Y. S. Lim, I. M. Chen, Y. Fan, M. Devidas, M. J. Borowitz, C. Smith, G. Neale, E. G. Burchard, D. G. Torgerson, F. A. Klussmann, C. R. N. Villagran, N. J. Winick, B. M. Camitta, E. Raetz, B. Wood, F. Yue, W. L. Carroll, E. Larsen, W. P. Bowman, M. L. Loh, M. Dean, D. Bhojwani, C. H. Pui, W. E. Evans, M. V. Relling, S. P. Hunger, C. L. Willman, C. G. Mullighan, and J. J. Yang. 2013. Inherited GATA3 variants are associated with Ph-like childhood acute lymphoblastic leukemia and risk of relapse. *Nat. Genet.* 45(12): 1494-8.

114. Walsh, K. M., A. P. Chokkalingam, L. I. Hsu, C. Metayer, A. J. De Smith, D. I. Jacobs, G. V. Dahl, M. L. Loh, I. V. Smirnov, K. Bartley, X. Ma, J. K. Wiencke, L. F. Barcellos, J. L. Wiemels, and P. A. Buffler. 2013. Associations between genome-wide Native American ancestry, known risk alleles and B-cell ALL risk in Hispanic children. *Leukemia.* 27(12): 2416-9.

115. Altshuler, D. M., R. A. Gibbs, L. Peltonen, S. F. Schaffner, F. Yu, E. Dermitzakis, P. E. Bonnen, P. I. W. De Bakker, P. Deloukas, S. B. Gabriel, R. Gwilliam, S. Hunt, M. Inouye, X. Jia, Aarno Palotie, M. Parkin, P. Whittaker, K. Chang, A. Hawes, L. R. Lewis, Y. Ren, D. Wheeler, D. M. Muzny, C. Barnes, K. Darvishi, M. Hurles, J. M. Korn, K. Kristiansson, C. Lee, S. A. McCarroll, J. Nemesh, A. Keinan, S. B. Montgomery, Samuela Pollack, A. L. Price, N. Soranzo, C. Gonzaga-Jauregui, V. Anttila, W. Brodeur, M. J. Daly, S. Leslie, G. McVean, L. Moutsianas, H. Nguyen, Q. Zhang, M. J. R. Ghori, R. McGinnis, W. McLaren, S. Pollack, F. Takeuchi, S. R. Grossman, I. Shlyakhter, E. B. Hostetter, P. C. Sabeti, C. A. Adebamowo, M. W. Foster, D. R. Gordon, J. Licinio, M. C. Manca, P. A. Marshall, I. Matsuda, D. Ngare, V. O. Wang, D. Reddy, C. N. Rotimi, C. D. Royal, R. R. Sharp, C. Zeng, D. B. Brooks, and J. E. McEwen. 2010. Integrating common and rare genetic variation in diverse human populations. *Nature.* 467(7311): 52-8.

116. Williams, R. C., R. C. Elston, P. Kumar, W. C. Knowler, H. E. Abboud, S. Adler, D. W. Bowden, J. Divers, B. I. Freedman, R. P. Igo, E. Ipp, S. K. Iyengar, P. L. Kimmel, M. Klag, O. Kohn, C. D. Langefeld, D. J. Leehey, R. G. Nelson, S. B. Nicholas, M. V. Pahl, R. S. Parekh, J. I. Rotter, J. R. Schelling, J. R. Sedor, V. O. Shah, M. W. Smith, K. D. Taylor, F. Thameem, D. Thornley-Brown, C. A. Winkler, X. Guo, P. Zager, R. L. Hanson, FIND Research Group, S.K. Iyengar, R. C. Elston, K. A. B. Goddard, J. M. Olson, S. Ialacci, J. Fondran, A. Horvath, R. Igo, G. Jun, K. Kramp, J. Molineros, S. R.

E. Quade, J. Schelling, A. Pickens, L. Humbert, L. Getz-Fradley, S. Adler, E. Ipp, M. Pahl, M. F. Seldin, S. Snyder, J. Tayek, E. Hernandez, J. LaPage, C. Garcia, J. Gonzalez, M. Aguilar, R. Parekh, L. Kao, L. Meoni, T. Whitehead, J. Chester, W. C. Knowler, R. L. Hanson, R. G. Nelson, J. Wolford, L. Jones, R. Juan, R. Lovelace, C. Luethe, L. M. Phillips, J. Sewemaenewa, I. Sili, B. Waseta, M. F. Saad, S. B. Nicholas, Y. D. I. Chen, X. Guo, J. Rotter, K. Taylor, M. Budgett, F. Hariri, P. Zager, V. Shah, M. Scavini, A. Bobelu, H. Abboud, N. Arar, R. Duggirala, B. S. Kasinath, F. Thameem, M. Stern, B. I. Freedman, D. W. Bowden, C. D. Langefeld, S. C. Satko, S. S. Rich, S. Warren, S. Viverette, G. Brooks, R. Young, M. Spainhour, C. Winkler, M. W. Smith, M. Thompson, R. Hanson, B. Kessing, D. J. Leehey, G. Barone, D. Thornley-Brown, C. Jefferson, O. F. Kohn, C. S. Brown, J. P. Briggs, P. L. Kimmel, R. Rasooly, D. Warnock, L. Cardon, R. Chakraborty, G. M. Dunston, T. Hostetter, S. J. O'Brien, J. Rioux, and R. Spielman. 2016. Selecting SNPs informative for African, American Indian and European Ancestry: Application to the Family Investigation of Nephropathy and Diabetes (FIND). *BMC Genomics.* 17: 325.

117. Mao, X., A. W. Bigham, R. Mei, G. Gutierrez, K. M. Weiss, T. D. Brutsaert, F. Leon-Velarde, L. G. Moore, E. Vargas, P. M. McKeigue, M. D. Shriver, and E. J. Parra. 2007. A genomewide admixture mapping panel for hispanic/latino populations. *Am. J. Hum. Genet.* 80(6): 1171-8.

118. Clarke, L., X. Zheng-Bradley, R. Smith, E. Kulesha, C. Xiao, I. Toneva, B. Vaughan, D. Preuss, R. Leinonen, M. Shumway, S. Sherry, and P. Flicek. 2012. The 1000 Genomes Pproject: Data management and community access. *Nat. Methods* .

119. Salzano, F. M., and M. Sans. 2014. Interethnic admixture and the evolution of Latin American populations. *Genet. Mol. Biol.* 37(1 Suppl): 151-170.

120. Kowalski, M. H., H. Qian, Z. Hou, J. D. Rosen, A. L. Tapia, Y. Shan, D. Jain, M.

Argos, D. K. Arnett, C. Avery, K. C. Barnes, L. C. Becker, S. A. Bien, J. C. Bis, J. Blangero, E. Boerwinkle, D. W. Bowden, S. Buyske, J. Cai, M. H. Cho, S. H. Choi, H. Choquet, L. Adrienne Cupples, M. Cushman, M. Daya, P. S. de Vries, P. T. Ellinor, N. Faraday, M. Fornage, S. Gabriel, S. K. Ganesh, M. Graff, N. Gupta, J. He, S. R. Heckbert, B. Hidalgo, C. J. Hodonsky, M. R. Irvin, A. D. Johnson, E. Jorgenson, R. Kaplan, S. L. R. Kardia, T. N. Kelly, C. Kooperberg, J. A. Lasky-Su, R. J. F. Loos, S. A. Lubitz, R. A. Mathias, C. P. McHugh, C. Montgomery, J. Y. Moon, A. C. Morrison, N. D. Palmer, N. Pankratz, G. J. Papanicolaou, J. M. Peralta, P. A. Peyser, S. S. Rich, J. I. Rotter, E. K. Silverman, J. A. Smith, N. L. Smith, K. D. Taylor, T. A. Thornton, H. K. Tiwari, R. P. Tracy, T. Wang, S. T. Weiss, L. C. Weng, K. L. Wiggins, J. G. Wilson, L. R. Yanek, S. Zöllner, K. E. North, P. L. Auer, L. M. Raffield, A. P. Reiner, and Y. Li. 2019. Use of >100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium whole genome sequences improves imputation quality and detection of rare variant associations in admixed African and Hispanic/Latino populations. *PLoS Genet.* 15: 1-25.

121. Brigle, K., and B. Rogers. 2017. Pathobiology and Diagnosis of Multiple Myeloma. *Semin. Oncol. Nurs.* 33(3): 225-236.

122. Cioli, D., and C. Baglioni. 1968. Catabolic origin of a Bence Jones protein fragment. *J. Exp. Med.* 128 (3): 517-532.

123. Rajkumar, S. V., M. A. Dimopoulos, A. Palumbo, J. Blade, G. Merlini, M. V. Mateos, S. Kumar, J. Hillengass, E. Kastritis, P. Richardson, O. Landgren, B. Paiva, A. Dispenzieri, B. Weiss, X. LeLeu, S. Zweegman, S. Lonial, L. Rosinol, E. Zamagni, S. Jagannath, O. Sezer, S. Y. Kristinsson, J. Caers, S. Z. Usmani, J. J. Lahuerta, H. E. Johnsen, M. Beksac, M. Cavo, H. Goldschmidt, E. Terpos, R. A. Kyle, K. C. Anderson, B. G. M. Durie, and J. F. S. Miguel. 2014. International Myeloma Working Group

updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol.* 15: e538-e548.

124. Nandakumar, B., M. Binder, A. Dispenzieri, P. Kapoor, F. Buadi, M. A. Gertz, M. Lacy, D. Dingli, L. Hwa, N. Leung, S. R. Hayman, W. I. Gonsalves, T. Kourelis, E. Muchtar, R. M. Warsame, R. S. Go, M. A. Hobbs, R. A. Kyle, S. V. Rajkumar, and S. Kumar. 2019. Continued improvement in survival in multiple myeloma (MM) including high-risk patients. *J. Clin. Oncol.* 28(5): 1122-1128.

125. Igwe, I. J., D. Yang, A. Merchant, N. Merin, G. Yaghmour, K. Kelly, and G. Ramsingh. 2017. The presence of Philadelphia chromosome does not confer poor prognosis in adult pre-B acute lymphoblastic leukaemia in the tyrosine kinase inhibitor era – a surveillance, epidemiology, and end results database analysis. *Br. J. Haematol.* 179(4): 618-626.

126. Aldrich, M. C., L. Zhang, J. L. Wiemels, X. Ma, M. L. Loh, C. Metayer, S. Selvin, J. Feusner, M. T. Smith, and P. A. Buffler. 2006. Cytogenetics of Hispanic and White children with acute lymphoblastic leukemia in California. *Cancer Epidemiol. Biomarkers Prev.* 15(3): 578-81.

127. Quiroz, E., R. Nelson, V. A. Pullarkat, I. Aldoss, T. Slavin, E. M. Rego, G. Marcucci, and D. Douer. 2018. Clinical, Cytogenetic and Immunophenotype Distribution of Adult Acute Lymphoblastic Leukemia Among Latinos — a Report from a Large Single Institution Cohort in Southern California. *Blood.* 132 (Supplement 1): 2816.

128. Walsh, K. M., A. J. de Smith, T. C. Welch, I. Smirnov, M. J. Cunningham, X. Ma, A. P. Chokkalingam, G. V. Dahl, W. Roberts, L. F. Barcellos, P. A. Buffler, C. Metayer, and J. L. Wiemels. 2014. Genomic ancestry and somatic alterations correlate with age at diagnosis in Hispanic children with B-cell acute lymphoblastic leukemia. *Am. J. Hematol.* 89(7): 721-5.

129. Colunga-Pedraza, P. R., G. B. Gomez-Cruz, J. E. Colunga-Pedraza, and G. J.

Ruiz-Argüelles. 2018. Geographic Hematology: Some Observations in Mexico. *Acta Haematol.* 140: 114-120.

130. Landgren, O., B. I. Graubard, S. Kumar, R. A. Kyle, J. A. Katzmann, K. Murata, R. Costello, A. Dispenzieri, N. Caporaso, S. Mailankody, N. Korde, M. Hultcrantz, T. M. Therneau, D. R. Larson, J. R. Cerhan, and S. V. Rajkumar. 2017. Prevalence of myeloma precursor state monoclonal gammopathy of undetermined significance in 12372 individuals 10-49 years old: a population-based study from the National Health and Nutrition Examination Survey. *Blood Cancer J.* 7(10): e618.

131. Schriber, J. R., P. N. Hari, K. W. Ahn, M. Fei, L. J. Costa, M. A. Kharfan-Dabaja, M. Angel-Diaz, R. P. Gale, S. Ganguly, S. K. Girnius, S. Hashmi, A. Pawarode, D. H. Vesole, P. H. Wiernik, B. M. Wirk, D. I. Marks, T. Nishihori, R. F. Olsson, S. Z. Usmani, T. M. Mark, Y. L. Nieto, and A. D'Souza. 2017. Hispanics have the lowest stem cell transplant utilization rate for autologous hematopoietic cell transplantation for multiple myeloma in the United States: A CIBMTR report. *Cancer.* 123(16): 3141-3149.

132. Duma, N., T. Azam, I. Bin Riaz, M. Gonzalez-Velez, S. Ailawadhi, and R. Go. 2018. Representation of Minorities and Elderly Patients in Multiple Myeloma Clinical Trials. *Oncologist.* 23(9): 1076-1078.

133. Costa, L. J., P. N. Hari, and S. K. Kumar. 2016. Differences between unselected patients and participants in multiple myeloma clinical trials in US: a threat to external validity. *Leuk. Lymphoma.* 57(12): 2827-2832.

134. Park, S. L., I. Cheng, and C. A. Haiman. 2018. Genome-wide association studies of cancer in diverse populations. *Cancer Epidemiol. Biomarkers Prev.* 27(4): 405–417.

135. Patel, M. I., Y. Ma, B. S. Mitchell, and K. F. Rhoads. 2015. Age and genetics: How do prognostic factors at diagnosis explain disparities in acute myeloid leukemia? *Am. J. Clin. Oncol. Cancer Clin. Trials* 38(2): 159-64.

136. Costa, L. J., I. K. Brill, J. Omel, K. Godby, S. K. Kumar, and E. E. Brown. 2017. Recent trends in multiple myeloma incidence and survival by age, race, and ethnicity in the United States. *Blood Adv.* 1(4): 282–287.

137. Darbinyan, K., A. Shastri, A. Budhathoki, D. Helbig, R. Snyder, K. Pradhan, J. Saleh-Esa, N. S. Kornblum, A. F. Binder, S. Goel, M. Janakiram, O. Derman, K. Gritsman, U. Steidl, I. Braunschweig, A. Verma, and I. Mantzaris. 2017. Hispanic ethnicity is associated with younger age at presentation but worse survival in acute myeloid leukemia. *Blood Adv.* 1(24): 2120-2123.

138. Price, A. L., N. Patterson, F. Yu, D. R. Cox, A. Waliszewska, G. J. McDonald, A. Tandon, C. Schirmer, J. Neubauer, G. Bedoya, C. Duque, A. Villegas, M. C. Bortolini, F. M. Salzano, C. Gallo, G. Mazzotti, M. Tello-Ruiz, L. Riba, C. A. Aguilar-Salinas, S. Canizales-Quinteros, M. Menjivar, W. Klitz, B. Henderson, C. A. Haiman, C. Winkler, T. Tusie-Luna, A. Ruiz-Linares, and D. Reich. 2007. A genomewide admixture map for latino populations. *Am. J. Hum. Genet.* 80(6): 1024-1036.

139. Bryc, K., E. Y. Durand, J. M. Macpherson, D. Reich, and J. L. Mountain. 2015. The genetic ancestry of african americans, latinos, and european Americans across the United States. *Am. J. Hum. Genet.* 96(1): 37-53.

140. Rosenberg, N. A., L. M. Li, R. Ward, and J. K. Pritchard. 2003. Informativeness of Genetic Markers for Inference of Ancestry. *Am. J. Hum. Genet.* 73(6): 1402-22.

141. Wang, L. J., C. W. Zhang, S. C. Su, H. I. H. Chen, Y. C. Chiu, Z. Lai, H. Bouamar, A. G. Ramirez, F. G. Cigarroa, L. Z. Sun, and Y. Chen. 2019. An ancestry informative marker panel design for individual ancestry estimation of Hispanic population using whole exome sequencing data. *BMC Genomics.* 20(Suppl 12): 1007.

142. Montinaro, F., G. B. J. Busby, V. L. Pascali, S. Myers, G. Hellenthal, and C. Capelli. 2015. Unravelling the hidden ancestry of American admixed populations. *Nat.*

Commun. 6: 6596.

143. Moreno-Estrada, A., C. R. Gignoux, J. C. Fernández-López, F. Zakharia, M. Sikora, A. V. Contreras, V. Acuña-Alonzo, K. Sandoval, C. Eng, S. Romero-Hidalgo, P. Ortiz-Tello, V. Robles, E. E. Kenny, I. Nuño-Arana, R. Barquera-Lozano, G. Macín-Pérez, J. Granados-Arriola, S. Huntsman, J. M. Galanter, M. Via, J. G. Ford, R. Chapela, W. Rodriguez-Cintron, J. R. Rodríguez-Santana, I. Romieu, J. J. Sienra-Monge, B. Del Rio Navarro, S. J. London, A. Ruiz-Linares, R. Garcia-Herrera, K. Estrada, A. Hidalgo-Miranda, G. Jimenez-Sanchez, A. Carnevale, X. Soberón, S. Canizales-Quinteros, H. Rangel-Villalobos, I. Silva-Zolezzi, E. G. Burchard, and C. D. Bustamante. 2014. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science.* 344(6189): 1280-1285.

144. Freedman, M. L., C. A. Haiman, N. Patterson, G. J. McDonald, A. Tandon, A. Waliszewska, K. Penney, R. G. Steen, K. Ardlie, E. M. John, I. Oakley-Girvan, A. S. Whittemore, K. A. Cooney, S. A. Ingles, D. Altshuler, B. E. Henderson, and D. Reich. 2006. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. *Proc. Natl. Acad. Sci. U. S. A.* 103(38): 14068-73.

145. Salinas, C. A., E. Kwon, C. S. Carlson, J. S. Koopmeiners, Z. Feng, D. M. Karyadi, E. A. Ostrander, and J. L. Stanford. 2008. Multiple independent genetic variants in the 8q24 region are associated with prostate cancer risk. *Cancer Epidemiol. Biomarkers Prev.* 17(5): 1203-13.

Vita

Alem Abebe Belachew was born in Addis Ababa, Ethiopia, the daughter of Lemlem Sissay and Abebe Belachew. After completing her High School work at Nativity Girls School, she moved to Vienna, Austria. The following year, she entered California State University Northridge (CSUN) for her undergraduate education. She received the degree of Bachelor of Science with a major in Biochemistry from CSUN in May 2011. For the next four years, she worked as a Research Trainee and Research Technologist at the Mayo Clinic while pursuing a Masters in Science in Biochemistry and Molecular Biology from Mayo Graduate School. In August of 2015, she entered The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences.