11-12-2013

Genetic Markers, Birth Characteristics, and Childhood Leukemia Risk

Amy Kennedy
Florida International University, akenn001@fiu.edu

DOI: 10.25148/etd.FI13120601

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

GENETIC MARKERS, BIRTH CHARACTERISTICS, AND CHILDHOOD LEUKEMIA RISK

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PUBLIC HEALTH

by

Amy Kennedy

2013

To: Dean Michele Ciccazzo
R.Stempel College of Public Health and Social Work

This dissertation, written by Amy Kennedy, and entitled Genetic Markers, Birth Characteristics, and Childhood Leukemia Risk, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Marianna K. Baum

_____
Wasim Maziak

_____
O. Dale Williams

_____
Mehmet T. Dorak, Major Professor

Date of Defense: November 12, 2013

The dissertation of Amy Kennedy is approved.

_____
Dean Michele Ciccazzo
R.Stempel College of Public Health and Social Work

_____
Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2013

DEDICATION

I dedicate this dissertation to my family, especially my parents, David and Karen, my sister Sara and her growing family, late grandmothers, Ann and Betty, and my amazing friends. Your support and encouragement was always appreciated, and I could not have done this without all of you.

ACKNOWLEDGMENTS

I gratefully acknowledge the support and guidance of my major professor, Dr. Mehmet Dorak, whose encouragement and persistence enabled me to be successful in my work and grow as an individual. I wish to thank my dissertation committee, Dr. Marianna Baum, Dr. Wasim Maziak, and Dr. O. Dale Williams for their time and efforts in guiding my research, and providing valuable feedback and advice. I would also like to thank our collaborators at the Baylor College of Medicine in Houston, Texas for providing the DNA samples that made this research possible.

I would like to acknowledge and thank the Dean's Office for giving me the opportunity to work there, and broadening my view on public health. The assistantship and support has been very rewarding for the past two and a half years. I am also grateful to the University Graduate School for providing me with the Dissertation Evidence Acquisition Fellowship.

Last but certainly not least, I would like to acknowledge my family and friends who have supported me throughout this journey. My parents, David and Karen, have been nothing but positive and supportive throughout, and I greatly appreciate it. I want to thank my sister Sara and brother-in-law Dave for their encouragement, and my nephew Connor for bringing a special light into my life. I cherish the support of my friends living all across the country, who have been encouraging throughout. I would also like to thank my fellow graduate students, former and current, who have provided a supportive and engaging environment over the past five years.

ABSTRACT OF THE DISSERTATION

GENETIC MARKERS, BIRTH CHARACTERISTICS, AND

CHILDHOOD LEUKEMIA RISK

by

Amy Kennedy

Florida International University, 2013

Miami, Florida

Professor Mehmet T. Dorak, Major Professor

The cause for childhood acute lymphoblastic leukemia (ALL) remains unknown, but male gender is a risk factor, and among ethnicities, Hispanics have the highest risk. In this dissertation, we explored correlations among genetic polymorphisms, birth characteristics, and the risk of childhood ALL in a multi-ethnic sample in 161 cases and 231 controls recruited contemporaneously (2007-2012) in Houston, TX. We first examined three lymphoma risk markers, since lymphoma and ALL both stem from lymphoid cells. Of these, rs2395185 showed a risk association in non-Hispanic White males (OR=2.8, P=0.02; $P_{interaction}$=0.03 for gender), but not in Hispanics. We verified previously known risk associations to validate the case-control sample. Mutations of *HFE* (C282Y, H63D) were genotyped to test whether iron-regulatory gene (IRG) variants known to elevate iron levels increase childhood ALL risk. Being positive for either polymorphism yielded only a modestly elevated OR in males, which increased to 2.96 (P=0.01) in the presence of a particular transferrin receptor (*TFRC*) genotype for rs3817672 ($P_{interaction}$=0.04). SNP rs3817672 itself showed an ethnicity-specific association ($P_{interaction}$=0.02 for ethnicity). We then examined additional IRG SNPs

(rs422982, rs855791, rs733655), which showed risk associations in males (ORs=1.52 to 2.60). A polygenic model based on the number of polymorphic alleles in five IRG SNPs revealed a linear increase in risk (OR=2.00 per incremental change; *P*=0.002). Having three or more alleles compared with none was associated with increased risk in males (OR=4.12; *P*=0.004). Significant risk associations with childhood ALL was found with birth length (OR=1.18 per inch, *P*=0.04), high birth weight (>4,000g) (OR=1.93, *P*=0.01), and with gestational age (OR=1.10 per week, *P*=0.04). We observed a negative correlation between *HFE* SNP rs9366637 and gestational age (*P*=0.005), again, stronger in males (*P*=0.001) and interacting with *TFRC* (*P*<0.001; $P_{interaction}$=0.05). Our results showed that (i) ALL risk markers do not show universal associations across ethnicities or between genders, (ii) IRG SNPs modify ALL risk presumably by their effects on iron levels, (iii) a negative correlation between an *HFE* SNP and gestational age exists, which implicates an iron-related mechanism. The results suggest that currently unregulated supplemental iron intake may have implications on childhood ALL development.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

ABBREVIATIONS AND ACRONYMS

Acute lymphoblastic leukemia	ALL
Acute myeloid leukemia	AML
Ancestry-informative marker	AIM
AT rich interactive domain 5B	ARID5B
Baylor College of Medicine Pediatric Hematology/Oncology Clinic	BCM
CCAAT/enhancer binding protein (C/EBP), epsilon	CEBPE
Deoxyribonucleic acid	DNA
Genome-wide association studies	GWAS
Hemochromatosis (gene or protein)	HFE
Hepcidin antimicrobial peptide	HAMP
Hereditary hemochromatosis	HH
Hispanic White	HW
Human leukocyte antigen	HLA
IKAROS family zinc finger 1 (Ikaros)	IKZF1
Iron-regulatory gene	IRG
Non-Hispanic White	NHW
Polymerase chain reaction	PCR
Solute Carrier Family 11 (proton-coupled divalent metal ion transporters), member 2	SLC11A2
Single nucleotide polymorphism	SNP
Transferrin (protein)	TF
Transferrin receptor	TFRC

Transmembrane protease, serine 6	TMPRSS6
Texas Children's Cancer and Hematology Center in Houston, Texas	TXCCC

CHAPTER I

INTRODUCTION

Cancer is the second leading cause of death among children ages 1-14, surpassed only by accidents [1]. Despite making up only a small percentage of the total cancer cases in the United States, approximately 1 in 100 new cancers is a childhood cancer, and the number of childhood cancer cases has increased gradually over the past few decades. Acute lymphoblastic leukemia (ALL) is the most common leukemia among children [2,3]. There are approximately 30-45 children per 1,000,000 affected by this malignancy annually [4], making up 80% of all childhood leukemias [4,5], with a peak incidence occurring between the ages of 2 and 5 [6]. The racial/ethnic variation in the incidence of childhood ALL is well established [4,7]. The highest incidence rates are among Hispanic populations [8-10], followed by non-Hispanic Whites, Asians, and then Blacks [4].

ALL is typically characterized by an excess amount of lymphoblasts. Instead of maturing into functional lymphocytes, an abundance of stem cells develop into lymphoblasts, which are immature, malignant white blood cells [11]. These non-functional leukemia cells accumulate quickly in the blood and bone marrow, reducing the effectiveness of healthy bloods cells in fighting infection, and eventual spread to other organs [12]. The most common subtype of ALL is B-cell precursor (BCP) ALL, accounting for 70% of childhood ALL [3].

Great improvements have been made in the treatment of ALL. In the 1960s, the 5-year survival rate was less than 10% [13]. This rate has continually increased from 77% in 1985-1994 [13] to 87.5% in 2000-2004, and is now estimated to be greater than 90% [14]. The increase in survival rates is due to the combination of better diagnostics through

molecular and microarray testing, and better treatments individualized to patients [15]. Despite the improving prognosis, the incident rate of ALL has increased annually over the past few decades, and the exact cause of the disease remains unknown in more than 90% of the cases [15]. Survivors of childhood leukemia also have an increased risk for other malignancies later in life, and almost two thirds of all pediatric cancer survivors suffer from at least one permanent side effect from their treatment [16], which emphasizes the need for prevention methods.

There are a few well established risk factors for childhood ALL, including race, sex, age peak, maternal age, and certain genetic disorders, including Down Syndrome, [17-22], but the exact cause is unknown in more than 90% of the cases [15,19]. Table I lists the established, suggestive, and controversial risk factors associated with childhood ALL. Birth weight has been recognized consistently as a risk factor for childhood ALL [23-36]. Less than 5% of diagnosed cases are associated solely with genetic defects [19], therefore many researchers are exploring environmental links to the disease. Ionizing radiation can increase the risk for ALL [28], although the relationship is more established with acute myeloid leukemia (AML).

It is likely that increased risk for childhood ALL involves the combination of environmental insults and genetic predisposition [37]. There have been several genome wide association studies (GWAS) conducted to identify risk variants for childhood ALL [38-40]. Table II lists the polymorphisms found to be statistically significant with childhood ALL risk in GWAS.

Candidate gene studies have also yielded many risk associations [38], although many have yet to be replicated. Variants located in the iron-related hemochromatosis

(*HFE*) gene, rs1800562 (C282Y) and rs1799945 (H63D), have been associated with an increase in birth weight and childhood ALL risk, in males only [35,41-44]. These two variants are the major contributors to hereditary hemochromatosis (HH), an inherited disorder causing iron overload, which may progress to cell and tissue damage [45]. Excess iron can cause cell stress and oxidative damage, and has been shown to increase susceptibility to cancer [46-49]. Male carriers of the variants of the *HFE* gene in particular have been shown to have an increased risk for numerous neoplasms [50]. The biological plausibility behind excess iron being a risk factor for childhood ALL is one worth pursuing, for both population health benefits and for scientific advancements.

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin 2013;63:11-30.

2. Mody R, Li S, Dover DC, Sallan S, Leisenring W, Oeffinger KC, Yasui Y, Robison LL, Neglia JP. Twenty-five-year follow-up among survivors of childhood acute lymphoblastic leukemia: a report from the Childhood Cancer Survivor Study. Blood 2008;111:5515-23.

3. Sherborne AL, Hemminki K, Kumar R, Bartram CR, Stanulla M, Schrappe M, Petridou E, Semsei AF, Szalai C, Sinnett D and others. Rationale for an international consortium to study inherited genetic susceptibility to childhood acute lymphoblastic leukemia. Haematologica 2011;96:1049-54.

4. Stiller CA. Epidemiology and genetics of childhood cancer. Oncogene 2004;23:6429-44.

5. Shah A, Coleman MP. Increasing incidence of childhood leukaemia: a controversy re-examined. Br J Cancer 2007;97:1009-12.

6. Swensen AR, Ross JA, Severson RK, Pollock BH, Robison LL. The age peak in childhood acute lymphoblastic leukemia: exploring the potential relationship with socioeconomic status. Cancer 1997;79:2045-51.

7. Stiller CA, Parkin DM. Geographic and ethnic variations in the incidence of childhood cancer. Br Med Bull 1996;52:682-703.

8. Wilkinson JD, Fleming LE, MacKinnon J, Voti L, Wohler-Torres B, Peace S, Trapido E. Lymphoma and lymphoid leukemia incidence in Florida children: ethnic and racial distribution. Cancer 2001;91:1402-8.

9. Wilkinson JD, Gonzalez A, Wohler-Torres B, Fleming LE, MacKinnon J, Trapido E, Button J, Peace S. Cancer incidence among Hispanic children in the United States. Rev Panam Salud Publica 2005;18:5-13.

10. Glazer ER, Perkins CI, Young JL, Jr., Schlag RD, Campleman SL, Wright WE. Cancer among Hispanic children in California, 1988-1994: comparison with non-Hispanic white children. Cancer 1999;86:1070-9.

11. Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A and others. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 2002;1:133-43.

12. Hutter JJ. Childhood leukemia. Pediatr Rev 2010;31:234-41.

13. Ries LAG, Smith MA, Gurney JG, Linet M, Tamra T, Young JL, Bunin GR, editors. Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975-1995. Volume NIH Pub. No. 99-4649. Bethesda, MD: National Cancer Institute, SEER Program 1999.

14. Hunger SP, Lu X, Devidas M, Camitta BM, Gaynon PS, Winick NJ, Reaman GH, Carroll WL. Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. J Clin Oncol 2012;30:1663-9.

15. Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: The impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA Cancer J Clin 2011;61:212-36.

16. Oeffinger KC, Hudson MM. Long-term complications following childhood and adolescent cancer: foundations for providing risk-based health care for survivors. CA Cancer J Clin 2004;54:208-36.

17. Birch JM. Genes and cancer. Arch Dis Child 1999;80:1-3.

18. Pui CH. Childhood leukemias. N Engl J Med 1995;332:1618-30.

19. Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. Lancet 2008;371:1030-43.

20. Belson M, Kingsley B, Holmes A. Risk factors for acute leukemia in children: a review. Environ Health Perspect 2007;115:138-45.

21. Feller M, Adam M, Zwahlen M, Brazzola P, Niggli F, Kuehni C. Family characteristics as risk factors for childhood acute lymphoblastic leukemia: a population-based case-control study. PLoS One 2010;5.

22. Greaves MF. Aetiology of acute leukaemia. Lancet 1997;349:344-9.

23. Robison LL, Codd M, Gunderson P, Neglia JP, Smithson WA, King FL. Birth weight as a risk factor for childhood acute lymphoblastic leukemia. Pediatr Hematol Oncol 1987;4:63-72.

24. Ross JA, Potter JD, Shu XO, Reaman GH, Lampkin B, Robison LL. Evaluating the relationships among maternal reproductive history, birth characteristics, and infant leukemia: a report from the Children's Cancer Group. Ann Epidemiol 1997;7:172-179.

25. Westergaard T, Andersen PK, Pedersen JB, Olsen JH, Frisch M, Sorensen HT, Wohlfahrt J, Melbye M. Birth characteristics, sibling patterns, and acute leukemia risk in childhood: a population-based cohort study. J Natl Cancer Inst 1997;89:939-947.

26. Yeazel MW, Ross JA, Buckley JD, Woods WG, Ruccione K, Robison LL. High birth weight and risk of specific childhood cancers: a report from the Children's Cancer Group. J Pediatr 1997;131:671-7.

27. Okcu MF, Goodman KJ, Carozza SE, Weiss NS, Burau KD, Bleyer WA, Cooper SP. Birth weight, ethnicity, and occurrence of cancer in children: a population-based, incident case-control study in the State of Texas, USA. Cancer Causes Control 2002;13:595-602.

28. Hjalgrim LL, Westergaard T, Rostgaard K, Schmiegelow K, Melbye M, Hjalgrim H, Engels EA. Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. Am J Epidemiol 2003;158:724-35.

29. Milne E, Royle JA, de Klerk NH, Blair E, Bailey H, Cole C, Attia J, Scott RJ, Armstrong BK. Fetal growth and risk of childhood acute lymphoblastic leukemia: results from an Australian case-control study. Am J Epidemiol 2009;170:221-8.

30. Hjalgrim LL, Rostgaard K, Hjalgrim H, Westergaard T, Thomassen H, Forestier E, Gustafsson G, Kristinsson J, Melbye M, Schmiegelow K. Birth weight and risk for childhood leukemia in Denmark, Sweden, Norway, and Iceland. J Natl Cancer Inst 2004;96:1549-56.

31. McLaughlin CC, Baptiste MS, Schymura MJ, Nasca PC, Zdeb MS. Birth weight, maternal weight and childhood leukaemia. Br J Cancer 2006;94:1738-44.

32. Dorak MT, Hammal DM, Pearce MS, McNally RJ, Parker L. Examination of gender effect in birth weight and miscarriage associations with childhood cancer. Cancer Causes Control 2007;18:219-228.

33. Koifman S, Pombo-de-Oliveira MS. High birth weight as an important risk factor for infant leukemia. Br J Cancer 2008;98:664-7.

34. Sprehe MR, Barahmani N, Cao Y, Wang T, Forman MR, Bondy M, Okcu MF. Comparison of birth weight corrected for gestational age and birth weight alone in prediction of development of childhood leukemia and central nervous system tumors. Pediatr Blood Cancer 2010;54:242-9.

35. Dorak MT, Mackay RK, Relton CL, Worwood M, Parker L, Hall AG. Hereditary hemochromatosis gene (HFE) variants are associated with birth weight and childhood leukemia risk Pediatric Blood Cancer 2009;53:1242-8.

36. Caughey RW, Michels KB. Birth weight and childhood leukemia: A meta-analysis and review of the current evidence. Int J Cancer 2009;124:2658-70.

37. Vijayakrishnan J, Houlston R. Candidate gene association studies and risk of childhood acute lymphoblastic leukemia: a systematic review and meta-analysis. Haematologica 2010;95:1405-1414.

38. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, Kinsey SE, Lightfoot T, Roman E, Irving JA and others. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1006-10.

39. Trevino LR, Yang W, French D, Hunger SP, Carroll WL, Devidas M, Willman C, Neale G, Downing J, Raimondi SC and others. Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1001-5.

40. Xu H, Yang W, Perez-Andreu V, Devidas M, Fan Y, Cheng C, Pei D, Scheet P, Burchard EG, Eng C and others. Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. J Natl Cancer Inst 2013;105:733-42.

41. Davis CF, Dorak MT. An extensive analysis of the hereditary hemochromatosis gene HFE and neighboring histone genes: associations with childhood leukemia. Ann Hematol 2010;89:275-84.

42. Dorak MT, Burnett AK, Worwood M. Hemochromatosis gene in leukemia and lymphoma. Leuk Lymphoma 2002;43:467-477.

43. Dorak MT, Sproul AM, Gibson BE, Burnett AK, Worwood M. The C282Y mutation of HFE is another male-specific risk factor for childhood ALL. Blood 1999;94:3957-3958.

44. Dorak MT, Burnett AK, Worwood M. HFE gene mutations in susceptibility to childhood leukemia: HuGE review. Genet Med 2005;7:159-68.

45. Olynyk JK, Trinder D, Ramm GA, Britton RS, Bacon BR. Hereditary hemochromatosis in the post-HFE era. Hepatology 2008;48:991-1001.

46. Torti SV, Torti FM. Iron and cancer: more ore to be mined. Nat Rev Cancer 2013;13:342-55.

47. Huang X. Iron overload and its association with cancer risk in humans: evidence for iron as a carcinogenic metal. Mutat Res 2003;533:153-71.

48. Toyokuni S. Iron-induced carcinogenesis: the role of redox regulation. Free Radic Biol Med 1996;20:553-66.

49. Chitambar CR, Wereley JP. Iron transport in a lymphoid cell line with the hemochromatosis C282Y mutation. Blood 2001;97:2734-2740.

50. Nelson RL, Davis FG, Persky V, Becker E. Risk of neoplastic and other diseases among people with heterozygosity for hereditary hemochromatosis. Cancer 1995;76:875-9.

Table I. Risk factors for childhood acute lymphoblastic leukemia

Established	Suggestive	Controversial
Sex	Maternal reproductive history	Maternal drug use
Age peak	Maternal alcohol use	Exposure to nonionizing radiation
Race	Maternal diet	Parental chemical exposure
Birth weight	Miscarriage history	Household exposure to pesticides
Down syndrome	Advanced maternal age	Parental smoking: prior and during pregnancy
Genetic disorders	Being the first-born or only child	
Ionizing radiation	Post-natal use of chloramphenicol	
Sibling with leukemia, brain tumor or Down syndrome		

Table II. Genetic associations with childhood leukemia risk previously established in genome wide association studies[*]

SNP	PMID[**]	Sample population	Chromosome nucleotide position	Reported gene
rs11978267	19684603	European ancestry	chr7: 50466304	*IKZF1, DDC*
rs11155133	19684603	European ancestry	chr6: 141169825	Intergenic
rs9290663	19684603	European ancestry	chr3: 178429939	*KCNMB2*
rs12621643	19684603	European ancestry	chr2: 223917983	*KCNE4*
rs2191566	19684603	European ancestry	chr19: 44511389	*ZNF230*
rs1879352	19684603	European ancestry	chr18: 2498054	Intergenic
rs2089222	19684603	European ancestry	chr12: 117002658	*KRTHB5*
rs10849033	19684603	European ancestry	chr12: 4425122	*C12orf5*
rs10821936	19684603	European ancestry	chr10: 63723577	*ARID5B*
rs563507	19684603	European ancestry	chr10: 34817988	*PARD3*
rs7554607	19684603	European ancestry	chr1: 237266603	*RYR2*
rs10873876	19684603	European ancestry	chr1: 76772328	*SIAT7C*
rs6428370	19684603	European ancestry	chr1: 196844593	Intergenic
rs1881797	19684603	European ancestry	chr1: 247689532	*OR2C3*
rs4132601	19684604	European ancestry	chr7: 50470604	*IKZF1*
rs2239633	19684604	European ancestry	chr14: 23589057	*CEBPE*
rs7089424	19684604	European ancestry	chr10: 63752159	*ARID5B*
rs6140264	20189245	Korean	chr20: 7376354	*HAO1*
rs9388856, rs9388857, rs1360756	20189245	Korean	chr6: 131148863	*EPB41L2*
rs3776932	20189245	Korean	chr5: 109185988	*MAN2A1*
rs12105972	20189245	Korean	chr2: 76648560	*C2orf3*
rs920590	22076464	European ancestry	chr8: 19651161	*INTS10*
rs7738636	22076464	European ancestry	chr6: 77789808	Intergenic
rs282708	22076464	European ancestry	chr4: 59503726	Intergenic
rs17505102	22076464	European ancestry	chr3: 189401776	*TP63*
rs17505102	22076464	European ancestry	chr3: 189401776	*TP63*
rs207954	22076464	European ancestry	chr15: 92657373	*SLCO3A1*
rs7156960	22076464	European ancestry	chr14: 76703351	*C14orf118*
rs7156960	22076464	European ancestry	chr14: 76703351	*C14orf118*
rs1945213	22076464	European ancestry	chr11: 56175671	*OR8U8*
rs1945213	22076464	European ancestry	chr11: 56175671	*OR8U8*
rs3942852	22076464	European ancestry	chr11: 48115089	*PTPRJ*
rs17837497	23007406	Mixed	chr7: 139702593	*TBXAS1*
rs1496766	23007406	Mixed	Cchr7: 78331465	*MAGI2*
rs17079534	23007406	Mixed	chr3: 39847072	*MYRIP*

Table II continued…

SNP	PMID**	Sample population	Chromosome nucleotide position	Reported gene
rs10170236	23007406	Mixed	chr2: 150457624	*LOC642340*
rs9958208	23007406	Mixed	chr18: 40591084	*RIT2*
rs7578361	23007406	Mixed	chr2: 150397218	Intergenic
rs7142143	23007406	Mixed	chr14: 51403531	*PYGL*
rs41322152	23007406	Mixed	chr10: 72040805	*NPFFR1*
rs6683977	23007406	Mixed	chr1: 66769100	*PDE4B*
rs546784	23007406	Mixed	chr1: 66762466	*PDE4B*
rs6964969	23512250	European, African American, Hispanic	chr7: 50473251	*IKZF1,DDC, GRB10*
rs4982731	23512250	European, African American, Hispanic	chr14: 23585333	*CEBPE*
rs10821936	23512250	European, African American, Hispanic	chr10: 63723577	*ARID5B*

*Hindorff LA, MacArthur J (European Bioinformatics Institute), Morales J (European Bioinformatics Institute), Junkins HA, Hall PN, Klemm AK, and Manolio TA. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies. Accessed September 10, 2013.

**PubMed identifier number

CHAPTER II

HYPOTHESIS AND SPECIFIC AIMS

The purpose of this dissertation project is to explore the correlations among genetic polymorphisms, birth characteristics, and the risk of childhood ALL. It is known that gender is a risk factor for childhood ALL, with males having greater risk, and Hispanics have the highest risk of all ethnicities. We hypothesize that risk markers may not be universal for both genders and across ethnicities. We aim to exploit this differential to obtain etiologic clues for the development of childhood ALL. Another focal point is to test the hypothesis that iron-regulatory gene (IRG) variants known to elevate iron levels increase childhood ALL risk. Iron, which is necessary for cell division and fetal growth, has been shown to increase birth weight and susceptibility to cancer when found in excess. This research will seek correlations between genotypes (IRG variants) and phenotypes (birth characteristics and leukemia) to determine whether genotypes and their combinations associated with birth characteristics also modify the risk for leukemia.

In order to test these hypotheses, the following specific aims will be explored:

Aim 1) to genotype previously reported childhood ALL risk markers to validate the case-control sample used in this project for the first time;

Aim 2) to genotype lymphoma risk markers, since lymphoma and ALL both stem from lymphoid cells, for their correlations with childhood ALL risk and birth characteristics;

Aim 3) to genotype iron-regulatory gene polymorphisms, some known to increase iron levels, for their correlations with childhood ALL risk and birth characteristics

Aim 4) to examine effect modification by race/ethnicity and gender in childhood ALL by stratified statistical analysis

Findings of this study may indicate that genetic risk markers are not universal for both genders and across ethnicities, and may be an important consideration in future genetic association studies. It will also provide evidence for or against the idea that iron excess mediated by genetic variants contribute to childhood ALL risk. If our hypothesis is confirmed, educating the public on lifestyle modifications for preventive measures such as controlled iron supplementation during pregnancy and avoidance from excessive consumption of iron-rich food could reduce the risk of leukemia in children.

CHAPTER III

HEREDITARY HEMOCHROMATOSIS GENE (*HFE*) POLYMORPHISMS AND ASSOCIATIONS WITH DISEASE SUSCEPTIBILITY

Abstract

The implication of excess iron in cancer and disease susceptibility has long been recognized. The increase in oxidative stress due to an overload of the transition metal is believed to be the reason for its harmful effects on the body. The hemochromatosis (*HFE*) gene is known to be involved in iron regulation, and mutations to the gene can have negative effects on certain health conditions. In this review, we discuss the role iron has in the body, and examine what impact *HFE* gene polymorphisms have on disease susceptibility, focusing on childhood leukemia. PubMed was searched using the following key words: *HFE* or hemochromatosis gene with disease susceptibility, polymorphisms, cancer, leukemia, variants, or disease association. The mechanisms behind the disease associations are also explored.

Role of iron

Iron is a necessary element required for oxygen transport, DNA synthesis, and is involved in cell proliferation and fetal growth [1-4]. Iron is capable of donating and accepting an electron, converting from ferrous (Fe^{2+}) to ferric (Fe^{3+}), which makes it useful in metabolic reactions, but also potentially harmful to the body. Iron generates free radicals when abundant and not tightly bound, which can lead to cell damage.

In humans, iron is found in heme proteins and in ferritin, an iron storage protein [5]. Iron is most commonly stored in the liver, spleen, and in bone marrow [6]. Circulating plasma iron is typically bound to one of the two binding sites on the protein

transferrin (TF) in healthy individuals. Cellular uptake of iron is dependent on the number of transferrin receptors (TFRCs) [5]. When cells are in need of iron, there is an increase in transcription of TFRC, and it is found greater numbers, therefore more receptors available for the TF-iron complex to bind and become internalized [5]. The opposite occurs when intracellular iron levels are high, in that transcription of TFRC mRNA is decreased, and therefore there are a reduced number of receptors available [5]. Intracellular iron is used for both ferritin and hemoglobin production, the protein found in red blood cells responsible for carrying oxygen throughout the body.

Health issues may arise when either too little or too much iron is stored and absorbed in the body. Iron deficiency, commonly referred to as iron anemia, has been extensively researched and is a common condition. In individuals with iron anemia, TFRC production is typically increased, while ferritin production is decreased. The lack of iron can lead to a wide range of adverse health outcomes [7]. Iron overload has not been studied as extensively until recently, and is underestimated and underdiagnosed by physicians. The primary reasons behind iron overload include: 1) excessive intake of iron through intestinal absorption, 2) parenteral iron (intravenously or intramuscular), 3) inhaled iron, 4) release of stored iron into blood plasma, 5) the reduction of normal menstrual excretion of iron in premenopausal women, and 6) the lack of transferrin or lactoferrin protein [8]. When found in excess, iron can induce cell stress and oxidative damage due to the increase in free radicals [3,9]. With the depletion of antioxidants, DNA damage can progress cancer development or worsen symptoms of an inherited disease [10]. The role of iron in cancer and disease progression has been widely studied [9], yet the genetics behind iron overload is still not fully understood. The

hemochromatosis gene (*HFE*) encodes for a protein involved in iron regulation, and its mutations may increase the risk of developing numerous cancers.

***HFE* gene**

The *HFE* (high iron FE) gene is located at 6p21.3, the short arm of chromosome 6 at position 21.3, and is surrounded by histone genes at both ends. The gene is approximately 10kb in size, and has at least eleven alternative transcripts consisting of four to seven exons, all which encode a distinct domain of the protein [11]. The gene encodes for a 321 amino acid mature transmembrane protein product that is similar in structure to HLA Class I molecules, with two alpha and beta chains, although it is not involved in antigen presentation [11,12]. It was originally named HLA-H because of the similarities with other HLA genes.

The HFE protein is expressed on cell surfaces in most tissues, but found most in the liver and small intestine [11], where iron absorption, transport, and storage takes place. HFE is also expressed on the apical plasma membrane of the syncytiotrophoblasts, which suggests it is the protein which regulates iron transport from the mother to the fetus [9]. Genome-wide association studies (GWAS) looking at hematological parameters such as serum markers of iron status have found significant associations with SNPs in the *HFE* gene. A list of the traits showing statistical significance with variants in the *HFE* gene is found in Table I.

Although not clearly understood, the *HFE* protein is involved in iron homeostasis regulation, controlling iron uptake by interacting with transferrin receptor (TFRC) and hepcidin (HAMP) [13]. The main function of HFE is to regulate mucosal iron transfer in interaction with TFRC, mainly in the intestinal mucosa (iron absorption) and in the

placenta (materno-fetal iron transfer). Animal studies first discovered high levels of saturated serum transferrin and excessive iron buildup in the liver in *HFE*-deficient mice [14]. The high iron levels in the hepatocytes of *HFE*-deficient mice are similar in histopathology to humans with hereditary hemochromatosis, suggesting that *HFE* is in control of iron homeostasis [9]. Transferrin receptors allow iron intake into the cell when iron concentrations are low by binding with iron-bound transferrin molecules. When iron levels are adequate and pH conditions are neutral, the *HFE* protein binds to the TFRC complex, reducing cellular iron uptake [12]. The hormone hepcidin is a negative regulator of iron absorption that determines how much iron will be stored and released [15]. *HFE* is speculated to regulate hepcidin by keeping production levels high when there is enough iron inside the cell, ensuring that the body will not absorb or store too much iron [15,16].

The interactions between *HFE* and *TFRC* have been observed previously in both adult cancers [17] and in childhood acute lymphoblastic leukemia [18]. In the study conducted by Beckman *et al.,* variants of *HFE* and *TFRC* genes on their own were not associated with breast, colorectal, or multiple myeloma, but their genotype combinations were significantly different between controls and cases [17]. It is speculated that the carcinogenic effect may somehow be amplified with the combination of both gene variants, and be dependent on increased iron uptake [17]. Normal, wild-type HFE protein is able to bind to transferrin receptors, reducing the affinity of TF, therefore controlling iron homeostasis. When *HFE* is mutated with the C282Y mutation, this binding is unable to occur, therefore allowing TF to bind to the receptors, and iron absorption to be uncontrolled. Mutations to the *HFE* gene disrupt iron homeostasis by causing increased

iron levels, which can lead to many negative health conditions, including hereditary hemochromatosis (HH).

HFE variants

Since the discovery of the *HFE* gene in 1996, many single nucleotide polymorphisms (SNPs) have been identified [19,20]. Information describing the most popular SNPs can be found in Table II. The two originally discovered missense mutations, rs1800562 (C282Y) and rs1799945 (H63D), account for most of the disease associations with *HFE*.

C282Y is the result of a point mutation in the coding sequence (exon 4) of the gene, with a guanine-to-adenine transition at nucleotide 845. This alters the production of the amino acid in position 282, changing it from cysteine to tyrosine [21]. This mutation is the most deleterious, as the cysteine residue is highly conserved, and the protein is no longer able to bind properly to TFRC when altered. The transferrin molecule is then able to bind to the TFRC with high affinity, and iron is released [22]. The frequency of C282Y in the United States is estimated to be around 6% in non-Hispanic Whites [23]. The C282Y mutation appeared more than 2000 years ago and is now associated with a particular haplotype in European populations, and is not found or found in low allele frequencies in other populations [24]. Two copies of the C282Y mutation is found in approximately 90% of patients with hereditary hemochromatosis (HH), an autosomal recessive disorder causing iron overload [25].

The second most common HH associated variant is the H63D mutation, which is a result of guanine replacing cytosine at nucleotide 187, causing the amino acid in position 63 to change from aspartate to histidine. This mutation only shows symptoms of

iron overload when the individual also has a copy of the C282Y mutation on the other chromosome, and symptoms are typically mild [25]. The variant protein is still able to bind to TFRC, although it is unable to reduce the affinity of transferrin binding to the receptors [26]. H63D mutations are more frequent than C282Y, with approximately 15% of non-Hispanic Whites being carriers of the mutation in the US, and 3% having two copies [23]. The penetrance of H63D is low, and the variant needs to be coupled with the C282Y variant in order to have an effect on iron levels. However, C282Y and H63D can never exist on the same chromosome.

A third mutation, rs1800730 (S65C), consists of an adenine to thymine substitution at position 193, causing the amino acid serine to change to cysteine at position 65 [27]. Studies have shown that this variant is associated with a mild form of hemochromatosis [28], having a greater effect when found in the presence of other *HFE* variants [29]. S65C is less common than C282Y and H63D, found in less than 5% of the non-Hispanic White population [28].

The SNP rs807212 was identified as tagging for the most common HFE region haplotype [30]. This intergenic SNP appears to tag for the wild-type alleles for all *HFE* variants, meaning the haplotype lacked all disease-associated variants [30]. There was a strong gender effect noticed with this SNP. Males had strong protective association against childhood leukemia, which made biological sense as male carriers of the variant rs807212 were therefore absent of the C282Y mutation, which had previously been associated with childhood leukemia risk [30-32].

The transition substitution in intron 1 creates the *HFE* SNP rs9366637. This SNP has been identified as a tagging SNP as well, but has not been studied extensively. It has

not yet been implicated in iron regulation, however, the haplotype tagged by rs9366637 lacks C282Y and H63D. The minor allele frequency for this SNP varies greatly depending on the race/ethnicity of the population, with the CEU (Western European ancestry) T-allele frequency being 0.06 and the HCB (Han Chinese in Beijing, China) having a MAF of 0.49. The variant (T) allele has been shown to be a marker of increased birth weight in European childhood ALL patients [18]. In a case-control study in a Han Chinese population, the C allele was found to be a significant risk marker for coronary heart disease (CHD) [33]. In a meta-analyses study looking at *HFE* gene variants and CHD in Han Chinese, only an association with rs1799945 variant allele increased risk for CHD, and the association with rs9366637 did not reach overall statistical significance [34]. The SNP was also found to be a marker for height in a GWAS [35].

Disease associations with *HFE* mutations

Hereditary hemochromatosis

Hemochromatosis describes any disorder caused by iron excess and tissue injury [25]. Hereditary hemochromatosis (HH) is the most commonly inherited form of iron overload. This autosomal recessive disorder may lead to severe organ dysfunction over time because of high iron absorption [25]. It is estimated that 90% of those with HH who are of northern European descent are homozygous for the C282Y mutation [36]. It is very common in that particular population, affecting 1 in every 200-300 individuals [37]. The condition is expressed more severely in males rather than females, as women are able to reduce their iron levels through the loss of blood during menstruation and childbirth [38].

Homozygosity for the C282Y variant or heterozygosity of C282Y and H63D on separate chromosomes are the two primary genetic combinations that cause HH [37].

When the C282Y mutations are present, the mutated HFE protein is no longer able to interact with transferrin receptors and hepcidin on at the cell surface [37]. Low hepcidin production is a result of the mutation, causing the cells to think iron levels are low and absorption is necessary. The excessive storage occurs primary in the liver, heart, joints, pancreas, skin, and testes [39]. Homozygosity of H63D on its own will not cause HH, but with a C282Y mutation may [40]. Determining the penetrance of the disorder is difficult to assess [41], as many patients do not show any signs or symptoms, particularly those who are heterozygotes [37]. Even with those that do show symptoms, they take a long time to develop, and clinical indicators are not apparent until at least middle age [31]. The horizontal bar chart depicted in Figure I represents the probability of developing HH based on *HFE* genotype. The highest risk is among C282Y homozygotes, with risk decreasing gradually for C282Y/H63D heterozygotes, H63D homozygotes, and C282Y heterozygotes. Homozygosity for H63D has the same risk as individuals who are not carriers of any *HFE* variants.

Once symptoms of iron excess develop, they can range from mild iron elevation to severe heart and liver disease [37]. Early symptoms include weight loss, lethargy, and stomach pain [39]. The liver is the first organ to typically show signs of HH, with hepatomegaly developing frequently [38]. Approximately 10-15% of patients with hepatocellular carcinoma have HH [38,42]. Other disorders that may stem from HH include diabetes, cardiomyopathy, and hyperpigmentation [38].

Hereditary hemochromatosis can be diagnosed by checking transferrin-iron saturation (TS) and serum ferritin concentrations. Serum transferrin-iron saturation greater than 50% in women and 60% in men is common in individuals with HH [38]. It is

not feasible to do prenatal testing for HH, since the disorder is treatable. The only way to treat HH is by removing iron from the body. The easiest way to do this is through phlebotomy. Individuals with HH should also monitor their iron intake, from both mineral supplements and with food choices [5].

Disease associations

Iron excess has been implicated in many different cancers and diseases, described in Table II. Studies have shown that iron can affect carcinogenesis by either suppressing the hosts' defense cells, acting as a nutrient for growing tumor cells, or by increasing oxidative stress with an increase in free radicals [17]. Free radical generation leads to inflammation and mutagenesis within the body [43]. Hereditary hemochromatosis is a well-established model of iron-induced carcinogenesis [10].

Many HH patients develop cirrhosis of the liver, and are at an increased risk for liver cancer. Primary hepatocellular carcinoma, in particular, is 200 times more common in these patients [42]. The liver is the major site of iron storage, therefore liver disorders are not surprising in HH patients. Increased risk for liver cancer has been seen not only in HH patients, but also in non-HH individuals with iron overload [44,45]. HH patients are also at an increased risk for diabetes [42]. The mechanism is not clearly understood, but it is hypothesized that iron accumulation damages pancreatic beta cells and insulin resistance [46]. Heart conditions such as cardiomyopathy and arrhythmias have been observed in in high numbers of HH patients as well [38]. Studies have shown that reducing iron stores through blood donation reduces the risk for heart disease [43]. Whether or not the C282Y mutation is a risk factor for developing heart disease is still

debated. Factors including smoking and hypertension status in women with heart disease may confound the association [47].

Many studies have explored the association between neurodegenerative disorders association and iron accumulation in the brain [48]. Iron accumulation is consistently observed in Alzheimer disease (AD) and its involvement in neuritic plaques has been well documented [49]. AD is enhanced by oxidative stress, and excess free iron would lead to oxidative stress through generation of free radicals in the neurons. Both C282Y and H63D have been associated with an increased risk for Alzheimer [26,50-52].

HFE variants and childhood leukemia risk

Previous studies have examined the risk association between *HFE* mutations and childhood acute lymphoblastic leukemia (ALL) risk [18,31,32,53,54]. The C282Y mutation was reported to be a risk factor for male children in a case-control study of Welsh and Scottish populations [53]. The minor allele frequencies for C282Y in males was 23.4% in cases and 12.3% in controls for the Wales population, and 34.7% in cases compared to 15.1% in controls in the Scotland group [53]. These populations are quite homogenous; therefore the generalization of the findings is unknown. H63D did not show any association with childhood ALL risk.

A Finnish study [55] is the only other published study that looked at *HFE* variants in childhood leukemia. The study population only included 32 childhood ALL patients, of which 14 were male, and did not find any significant results. Other hematologic malignancies studies on acute myeloid leukemia [56] and Hodgkin disease [32] have found no risk associations with *HFE* variants, and a myelodysplastic syndrome study [57]

revealed a positive association, which was not been replicated in a different population sample [58].

Mechanism

It is speculated that most cancer risk associated with increased iron levels is due to iron accumulation over years, resulting in chronic oxidative stress. The association with childhood leukemia cannot therefore be explained by this same reasoning. Chitambar et al. have suggested that increased intracellular iron levels in lymphoid cells during development may explain the risk [59]. Cell studies have shown that B-lymphoid cell lines that are homozygous for the C282Y variant demonstrate greater iron uptake and increased cell sensitivity to oxidative stress [59]. The increase in oxidative stress due to the high iron levels may increase radiation sensitivity, which has been shown to increase cancer susceptibility [60,61]. Lymphocytes can be radiation sensitized by iron, and Stevens et al. have identified C282Y heterozygotes as risk factors for radiation sensitivity [62,63]. Growing fetuses are sensitive to environmental insults, and those who are carriers of the C282Y variant may be exposed to higher intracellular levels of iron, especially if their mother is a carrier [31]. This hypothesis, of course, must be carried out in functional studies to determine its validity.

Another idea is that the *HFE* gene may play a role with immune function [53]. Mutations in *HFE* causing iron overload may interfere with the immune function of lymphoid cells, and be an underlying reason for the association with childhood leukemia risk. The role of *HFE* with the histone proteins may also be important [30]. The *HFE* gene is flanked by histone coding genes on both sides, and histones are known to be

important in genome biology and potentially with tumor suppression [30]. There may be similarities between the two or interactions between the genes that are unknown.

Conclusion

With well-known biology of its effects, iron excess is likely to have more impact on cancer and neurodegenerative diseases in the Western world. Iron supplementation is currently unregulated and non-personalized, with every male and female given the same suggested amount to intake. Educating the public on lifestyle modifications for those with iron overload such as controlled iron supplementation and avoidance from the unnecessary consumption of iron-rich food, could reduce ill side effects from the disorder. After decades of serological work trying to determine the effects of iron accumulation in the body, genetic epidemiology has made more of an impact in a shorter period of time.

REFERENCES

1. Gambling L, Czopek A, Andersen HS, Holtrop G, Srai SK, Krejpcio Z, McArdle HJ. Fetal iron status regulates maternal iron metabolism during pregnancy in the rat. Am J Physiol Regul Integr Comp Physiol 2009;296:1063-70.

2. Gambling L, Lang C, McArdle HJ. Fetal regulation of iron transport during pregnancy. Am J Clin Nutr 2011;94:1903S-1907S.

3. Srai SK, Bomford A, McArdle HJ. Iron transport across cell membranes: molecular understanding of duodenal and placental iron uptake. Best Pract Res Clin Haematol 2002;15:243-59.

4. Toyokuni S. Iron-induced carcinogenesis: the role of redox regulation. Free Radic Biol Med 1996;20:553-66.

5. Huang X. Iron overload and its association with cancer risk in humans: evidence for iron as a carcinogenic metal. Mutat Res 2003;533:153-71.

6. Puntarulo S. Iron, oxidative stress and human health. Mol Aspects Med 2005;26:299-312.

7. Scholl TO, Reilly T. Anemia, iron and pregnancy outcome. J Nutr 2000;130:443S-447S.

8. Weinberg ED. Cellular iron metabolism in health and disease. Drug Metab Rev 1990;22:531-79.

9. Lieu PT, Heiskala M, Peterson PA, Yang Y. The roles of iron in health and disease. Mol Aspects Med 2001;22:1-87.

10. Fargion S, Valenti L, Fracanzani AL. Hemochromatosis gene (HFE) mutations and cancer risk: expanding the clinical manifestations of hereditary iron overload. Hepatology 2010;51:1119-21.

11. Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A, Dormishian F, Domingo R, Jr., Ellis MC, Fullan A and others. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. Nat Genet 1996;13:399-408.

12. Lebron JA, Bennett MJ, Vaughn DE, Chirino AJ, Snow PM, Mintier GA, Feder JN, Bjorkman PJ. Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor. Cell 1998;93:111-23.

13. Fleming RE, Britton RS, Waheed A, Sly WS, Bacon BR. Pathogenesis of hereditary hemochromatosis. Clin Liver Dis 2004;8:755-73.

14. Zhou XY, Tomatsu S, Fleming RE, Parkkila S, Waheed A, Jiang J, Fei Y, Brunt EM, Ruddy DA, Prass CE and others. HFE gene knockout produces mouse model of hereditary hemochromatosis. Proc Natl Acad Sci USA 1998;95:2492-7.

15. Goswami T, Andrews NC. Hereditary hemochromatosis protein, HFE, interaction with transferrin receptor 2 suggests a molecular mechanism for mammalian iron sensing. J Biol Chem 2006;281:28494-8.

16. Collins JF, Wessling-Resnick M, Knutson MD. Hepcidin regulation of iron transport. J Nutr 2008;138:2284-8.

17. Beckman LE, Van Landeghem GF, Sikstrom C, Wahlin A, Markevarn B, Hallmans G, Lenner P, Athlin L, Stenling R, Beckman L. Interaction between haemochromatosis and transferrin receptor genes in different neoplastic disorders. Carcinogenesis 1999;20:1231-1233.

18. Dorak MT, Mackay RK, Relton CL, Worwood M, Parker L, Hall AG. Hereditary hemochromatosis gene (HFE) variants are associated with birth weight and childhood leukemia risk Pediatric Blood Cancer 2009;53:1242-8.

19. Pointon JJ, Wallace D, Merryweather-Clarke AT, Robson KJ. Uncommon mutations and polymorphisms in the hemochromatosis gene. Genet Test 2000;4:151-61.

20. Toomajian C, Kreitman M. Sequence variation and haplotype structure at the human HFE locus. Genetics 2002;161:1609-23.

21. Beutler E, Gelbart T, West C, Lee P, Adams M, Blackstone R, Pockros P, Kosty M, Venditti CP, Phatak PD and others. Mutation analysis in hereditary hemochromatosis. Blood Cells Mol Dis 1996;22:187-94.

22. Robson KJ, Merryweather-Clarke AT, Cadet E, Viprakasit V, Zaahl MG, Pointon JJ, Weatherall DJ, Rochette J. Recent advances in understanding haemochromatosis: a transition state. J Med Genet 2004;41:721-30.

23. Steinberg KK, Cogswell ME, Chang JC, Caudill SP, McQuillan GM, Bowman BA, Grummer-Strawn LM, Sampson EJ, Khoury MJ, Gallagher ML. Prevalence of C282Y and H63D mutations in the hemochromatosis (HFE) gene in the United States. JAMA 2001;285:2216-22.

24. Rochette J, Pointon JJ, Fisher CA, Perera G, Arambepola M, Arichchi DS, De Silva S, Vandwalle JL, Monti JP, Old JM and others. Multicentric origin of hemochromatosis gene (HFE) mutations. Am J Hum Genet 1999;64:1056-62.

25. Olynyk JK, Trinder D, Ramm GA, Britton RS, Bacon BR. Hereditary hemochromatosis in the post-HFE era. Hepatology 2008;48:991-1001.

26. Robson KJ, Lehmann DJ, Wimhurst VL, Livesey KJ, Combrinck M, Merryweather-Clarke AT, Warden DR, Smith AD. Synergy between the C2 allele of transferrin and the C282Y allele of the haemochromatosis gene (HFE) as risk factors for developing Alzheimer's disease. J Med Genet 2004;41:261-5.

27. Barton JC, Sawada-Hirai R, Rothenberg BE, Acton RT. Two novel missense mutations of the HFE gene (I105T and G93R) and identification of the S65C mutation in Alabama hemochromatosis probands. Blood Cells Mol Dis 1999;25:147-55.

28. Mura C, Raguenes O, Ferec C. HFE mutations analysis in 711 hemochromatosis probands: evidence for S65C implication in mild form of hemochromatosis. Blood 1999;93:2502-5.

29. Holmstrom P, Marmur J, Eggertsen G, Gafvels M, Stal P. Mild iron overload in patients carrying the HFE S65C gene mutation: a retrospective study in patients with suspected iron overload and healthy controls. Gut 2002;51:723-30.

30. Davis CF, Dorak MT. An extensive analysis of the hereditary hemochromatosis gene HFE and neighboring histone genes: associations with childhood leukemia. Ann Hematol 2010;89:275-84.

31. Dorak MT, Burnett AK, Worwood M. HFE gene mutations in susceptibility to childhood leukemia: HuGE review. Genet Med 2005;7:159-68.

32. Dorak MT, Burnett AK, Worwood M. Hemochromatosis gene in leukemia and lymphoma. Leuk Lymphoma 2002;43:467-477.

33. Shi Y, Zhou L, Huang LH, Lian YT, Zhang XM, Guo H, Wu TC, Cheng LX, He MA. Plasma ferritin levels, genetic variations in HFE gene, and coronary heart disease in Chinese: a case-control study. Atherosclerosis 2011;218:386-90.

34. Lian J, Xu L, Huang Y, Le Y, Jiang D, Yang X, Xu W, Huang X, Dong C, Ye M and others. Meta-analyses of HFE variants in coronary heart disease. Gene 2013;527:167-73.

35. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S and others. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 2010;467:832-8.

36. Frazer DM, Anderson GJ. The orchestration of body iron intake: how and where do enterocytes receive their cues? Blood Cells Mol Dis 2003;30:288-97.

37. Lyon E, Frank EL. Hereditary hemochromatosis since discovery of the HFE gene. Clin Chem 2001;47:1147-56.

38. Hanson EH, Imperatore G, Burke W. HFE gene and hereditary hemochromatosis: a HuGE review. Human Genome Epidemiology. Am J Epidemiol 2001;154:193-206.

39. Alexander J, Kowdley KV. HFE-associated hereditary hemochromatosis. Genet Med 2009;11:307-13.

40. Beutler E. The significance of the 187G (H63D) mutation in hemochromatosis. Am J Hum Genet 1997;61:762-4.

41. Rossi E, Jeffrey GP. Clinical penetrance of C282Y homozygous HFE haemochromatosis. Clin Biochem Rev 2004;25:183-90.

42. Niederau C, Fischer R, Purschel A, Stremmel W, Haussinger D, Strohmeyer G. Long-term survival in patients with hereditary hemochromatosis. Gastroenterology 1996;110:1107-19.

43. Worwood M. HFE Mutations as risk factors in disease. Best Pract Res Clin Haematol 2002;15:295-314.

44. Turlin B, Juguet F, Moirand R, Le Quilleuc D, Loreal O, Campion JP, Launois B, Ramee MP, Brissot P, Deugnier Y. Increased liver iron stores in patients with hepatocellular carcinoma developed on a noncirrhotic liver. Hepatology 1995;22:446-50.

45. Mandishona E, MacPhail AP, Gordeuk VR, Kedda MA, Paterson AC, Rouault TA, Kew MC. Dietary iron overload as a risk factor for hepatocellular carcinoma in Black Africans. Hepatology 1998;27:1563-6.

46. Niederau C, Fischer R, Sonnenberg A, Stremmel W, Trampisch HJ, Strohmeyer G. Survival and causes of death in cirrhotic and in noncirrhotic patients with primary hemochromatosis. N Engl J Med 1985;313:1256-62.

47. Roest M, van der Schouw YT, de Valk B, Marx JJ, Tempelman MJ, de Groot PG, Sixma JJ, Banga JD. Heterozygosity for a hereditary hemochromatosis gene is associated with cardiovascular death in women. Circulation 1999;100:1268-73.

48. Thompson KJ, Shoham S, Connor JR. Iron and neurodegenerative disorders. Brain Res Bull 2001;55:155-64.

49. Connor JR, Lee SY. HFE mutations and Alzheimer's disease. J Alzheimers Dis 2006;10:267-76.

50. Pulliam JF, Jennings CD, Kryscio RJ, Davis DG, Wilson D, Montine TJ, Schmitt FA, Markesbery WR. Association of HFE mutations with neurodegeneration and oxidative stress in Alzheimer's disease and correlation with APOE. Am J Med Genet B Neuropsychiatr Genet 2003;119B:48-53.

51. Sampietro M, Caputo L, Casatta A, Meregalli M, Pellagatti A, Tagliabue J, Annoni G, Vergani C. The hemochromatosis gene affects the age of onset of sporadic Alzheimer's disease. Neurobiol Aging 2001;22:563-8.

52. Moalem S, Percy ME, Andrews DF, Kruck TP, Wong S, Dalton AJ, Mehta P, Fedor B, Warren AC. Are hereditary hemochromatosis mutations involved in Alzheimer disease? Am J Med Genet 2000;93:58-66.

53. Dorak MT, Sproul AM, Gibson BE, Burnett AK, Worwood M. The C282Y mutation of HFE is another male-specific risk factor for childhood ALL. Blood 1999;94:3957-3958.

54. Dorak MT. HFE H63D variant and leukemia susceptibility. Leuk Lymphoma 2006;47:2269-70.

55. Hannuksela J, Savolainen ER, Koistinen P, Parkkila S. Prevalence of HFE genotypes, C282Y and H63D, in patients with hematologic disorders. Haematologica 2002;87:131-5.

56. Gimferrer E, Nomdedeu J, Gich I, Barcelo MJ, Baiget M. Prevalence of hemochromatosis related HFE gene mutations in patients with acute myeloid leukemia. Leuk Res 1999;23:597-8.

57. Varkonyi J, Tarkovacs G, Karadi I, Andrikovics H, Varga F, Demeter J, Tordai A. High incidence of hemochromatosis gene mutations in the myelodysplastic syndrome: the Budapest Study on 50 patients. Acta Haematol 2003;109:64-7.

58. Speletas M, Kioumi A, Mandala E, Katodritou E, Papaioannou G, Ritis K, Korantzis I. Prevalence of hemochromatosis gene (HFE) mutations in Greek patients with myelodysplastic syndromes. Acta Haematol 2003;110:53-4.

59. Chitambar CR, Wereley JP. Iron transport in a lymphoid cell line with the hemochromatosis C282Y mutation. Blood 2001;97:2734-2740.

60. Stevens RG, Kalkwarf DR. Iron, radiation, and cancer. Environ Health Perspect 1990;87:291-300.

61. Nelson JM, Stevens RG. Ferritin-iron increases killing of Chinese hamster ovary cells by X-irradiation. Cell Prolif 1992;25:579-85.

62. Stevens RG, Morris JE, Anderson LE. Hemochromatosis heterozygotes may constitute a radiation-sensitive subpopulation. Radiat Res 2000;153:844-7.

63. Stevens RG. Hemochromatosis heterozygotes may constitute a radiation-sensitive subpopulation. Radiat Res 2000;154:725-6.

Table I. GWAS findings for *HFE* polymorphisms[*]

PMID[**]	Trait	SNP
19820699	Iron status biomarkers	**rs1800562**
19084217		
21483845		
23446634	Red blood cell traits	
20927387		
20858683	Glycated hemoglobin levels	
19820697	Hematological parameters	
21785125	Hepcidin levels	
21943158	Cardiovascular disease risk factors	
19862010	Hematocrit Hemoglobin Mean corpuscular volume	
23263863	Hematology traits Mean corpuscular hemoglobin	
20686565	Cholesterol, total LDL cholesterol	
21909115	Diastolic blood pressure Hypertension Systolic blood pressure	**rs1799945**
21208937	Iron levels	
19820698	Hemoglobin	**rs198846**
20139978	Mean corpuscular volume	
21909110	Blood pressure	
19853236	Hematology traits	**rs1408272**
21149283	Iron status biomarkers	
21149283	Iron status biomarkers	**rs17342717**

[*]Hindorff LA, MacArthur J (European Bioinformatics Institute), Morales J (European Bioinformatics Institute), Junkins HA, Hall PN, Klemm AK, and Manolio TA. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies. Accessed September 10, 2013.
[**]PubMed identifier number

Table II. Characteristics of most studied SNPs in the *HFE* gene

				Minor allele and frequency**			
SNP	**Nucleotide position***	**Polymorphism type and location**	**Sequence**	**CEU**	**YRI**	**JPT**	**HCB**
rs1800562	26093141	Coding; nonsynonymous; substitution of cysteine (C) by tyrosine (Y) at amino acid position 282 (C282Y)	ATATACGT(A):(G)CCAGGTGG	(A) 0.053	NA	NA	NA
rs1799945	26091179	Coding; nonsynonymous; substitution of histidine (H) by aspartate (D) at amino acid position 63 (H63D)	TCTATGAT(C):(G)ATGAGAGT	(G) 0.179	(G) 0.00	(G) 0.042	(G) 0.125
rs1800730	26091185	Transversion substitution (S65C)	ATCATGAG(A):(T)GTCGCCGT	(T) 0.033	NA	NA	NA
rs807212	26065621	Transition substitution, intergenic	TTTTACCC(A):(G)GGAGTGGA	(A) 0.358	(A) 0.042	(A) 0.091	(A) 0.089
rs9366637	26089098	Intronic, transition substitution (IVS1)	TTTGCATT(C):(T)TAGTGGGA	(T) 0.066	(T) 0.075	C 0.385	(T) 0.493
rs2794719	26088890	Intron, transversion substitution, intragenic	CAAAGCCC(C):(A)GTGTACCA	(C) 0.397	(C) 0.155	(A) 0.302	(A) 0.256
rs2858996	26094026	Transversion substitution, intron/intragenic	GAGTTTGC(T):(G)TAGCTATC	(T) 0.208	(T) 0.089	(T) 0.066	(T) 0.179
rs2071303	26091336	Transition substitution, intron, intragenic	CTCTCCAC(A):(G)TACCCTTG	(G) 0.341	(G) 0.401	(A) 0.221	(A) 0.367
rs1800708	26093303	Transition substitution, intron, intragenic	GGGTGGGC(C):(T)GAGGGTGG	(C) 0.080	(C) 0.088	(T) 0.319	(T) 0.456
rs1572982	26094367	Intron, transition substitution, intragenic	GCAAGATG(A):(G)TGCCTAGG	(A) 0.456	(G) 0.312	(G) 0.159	(G) 0.296

*Genome Reference Consortium Human Build 37 patch release 10 (GRCh37.p10) used for nucleotide position (http://www.ncbi.nlm.nih.gov/SNP/)
**Population descriptors:
CEU (C): Utah residents with Northern and Western European ancestry from the CEPH collection
YRI (Y): Yoruban in Ibadan, Nigeria
JPT (J): Japanese in Tokyo, Japan
HCB (H): Han Chinese in Beijing, China

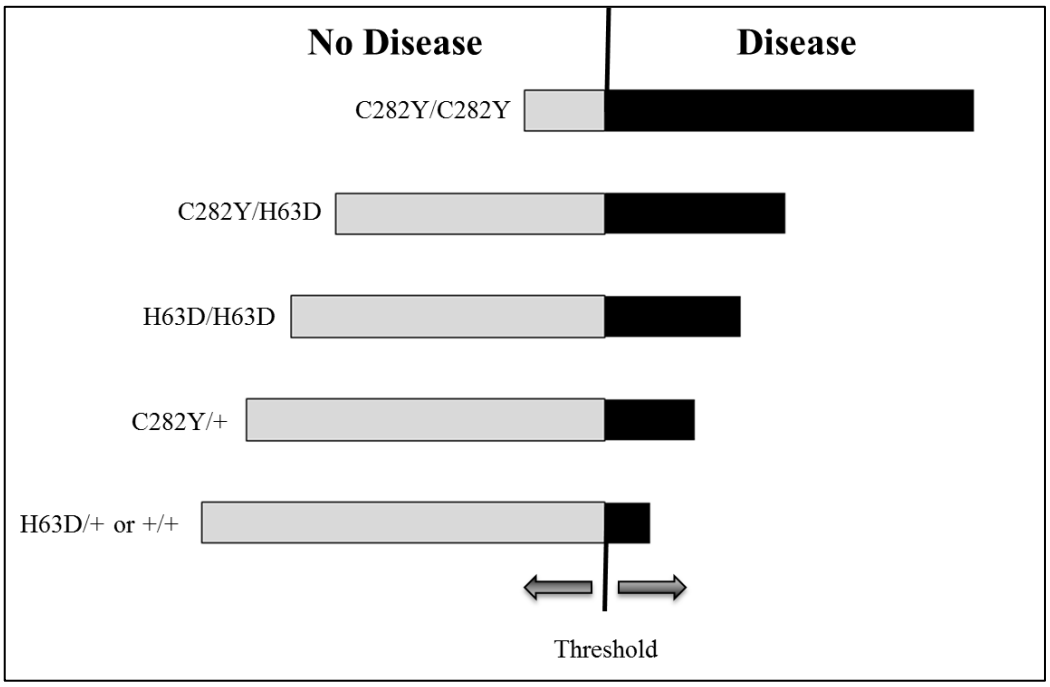


Figure I. Likelihood of developing adverse health effects from hereditary hemochromatosis based on *HFE* genotype

Adapted from: Khoury MJ, Burke W, Thomas EJ, editors. Genetics and Public Health in the 21st Ce Using Genetic Information to Improve Health and Prevent Disease. New York: Oxford Press; 2000.

Table III. Conditions showing increased risk associations with *HFE* variants

Condition	PMID*
Alzheimer disease	12707938, 15060098, 11445256, 10861683, 20029940, 17119292
Amyotrophic lateral sclerosis	17828789
Arthritis	12401309, 17284543, 16468045, 16583477, 16638105
Breast cancer	10383894, 14973098, 20099304, 16216474, 16503999, 23681799
Cardiovascular disease	12401309, 21696736, 20640879, 17389307
Cervical cancer	16414021
Childhood acute lymphoblastic leukemia	19806355, 10627122
Colon Cancer	10383894, 12529348, 20099304, 19291797, 15941956, 23553028, 23281741
Diabetes	8613000, 9654270, 9726605, 9885759, 22908207
Gastric cancer	23389292
Glioma	11591868, 19386095
Hepatic veno-occlusive disease	15834437
Liver cancer	10845668, 10989544, 11500061, 12003382, 15929796, 15017669, 18164971, 23281741, 10918159, 12591066, 1312985, 4058506, 8613000
Neurodegenerative diseases	12401309, 23813494, 22526559, 21349849, 21346098
Ovarian cancer	20669231, 21879820
Prostate cancer	16003728

*PubMed identifier number

CHAPTER IV

GENETIC MARKERS IN A MULTI-ETHNIC SAMPLE FOR CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA RISK

Abstract

Genome-wide association studies (GWAS) have identified multiple risk loci for childhood acute lymphoblastic leukemia (ALL), but mostly in European/White populations, even though Hispanics have a greater risk compared to the European/White population. We re-examined selected SNPs of known associations with childhood ALL and known HLA region lymphoma risk markers in a multi-ethnic population sample from Houston, Texas consisting of 161 incident childhood ALL cases and 231 controls. Significant associations were found in two *ARID5B* variants (rs7089424, OR_{allele}=1.69, P<0.001; and rs10821936, OR_{allele}=1.48, P=0.005), as previously shown in GWAS. Marker rs2395185, which is a protective marker for lymphoma, replicated the previously found strong risk association with childhood ALL in non-Hispanic White males (OR_{allele}=2.79, P= 0.02), but no association was noted in Hispanics. Another HLA region marker rs2647012 showed a statistically significant risk association among Hispanics, but not in non-Hispanics ($P_{interaction}$ = 0.003 for ethnicity). A strong statistically significant protective association was found with rs1048456, a risk marker in follicular lymphoma (OR_{rec}=0.19, P=0.009). Our study validated this new case-control sample by confirming some of the previously discovered genetic markers associated to childhood ALL, and yielded new associations with known lymphoma markers. The association of lymphoma markers were in opposite direction in childhood ALL. Despite positive results, our study did not provide any clues to why Hispanics have a higher susceptibility to childhood

leukemia, suggesting that environmental factors may have a stronger contribution to this differential.

Introduction

Over a third of cancers in children are leukemias [1], with acute lymphoblastic leukemia (ALL) being the most common [2-4]. Approximately 3.8 individuals (less than 14 years of age) per 1,000,000 are diagnosed with childhood ALL per year in the United States [5]. The racial disparity in the incidence of childhood ALL has been well established [6,7]. In the United States, incidence of ALL is greatest in Hispanic children, followed by non-Hispanic Whites, Asians, and then Blacks [7-9].

Research continually tries to determine the etiology of childhood leukemias. While there are a few known risk factors associated with childhood ALL [10-15], more than 90% are of unknown etiology [11]. Increased birth weight is an established risk factor shown to increase risk of childhood ALL [16-33]. Environmental factors are also involved, and may work in conjunction with genetic factors to cause many cancers [10].

Genome-wide association studies (GWAS) [34-37] have identified multiple risk loci showing significant associations with childhood ALL. Most significantly, variants located within the *ARID5B, IKZF1*, and *CEBPE* genes have robust risk associations. Most of these studies, however, have only evaluated risk among those of European ancestry. Two multiethnic studies conducted by Xu *et al*. looked at various risk loci in both African American and Hispanic American populations [8,37], discovering that some markers are universal across races/ethnicities, while others are race/ethnic-specific.

We re-examined five previously discovered single nucleotide polymorphisms (SNPs) of known associations to leukemia (rs7089424, rs10821936, rs10994982,

rs4132601, and rs2239633), along with three HLA region susceptibility markers for lymphomas, since lymphoma and ALL both stem from lymphoid cells. These included rs2395185, a marker of HLA DRB4 lineage [38] which has previously shown risk associations with major leukemias including childhood ALL and other diseases [39-49]; rs10484561, which has been shown to be a strong risk marker in follicular lymphoma (FL) [50]; and rs2647012, which is a protective marker for FL [51]. These SNPs were examined in a multi-ethnic sample (non-Hispanic Whites, Hispanic Whites, and Blacks) from Houston, Texas to assess their association with ALL. We hypothesized that the variation in ethnic/racial susceptibility to childhood ALL has a genetic basis.

Subjects and Methods

Study population

Institutional Review Board (IRB) approval was received at both the Baylor College of Medicine (BCM) and Florida International University prior to the start of the study. The case-control study comprised of 161 incident childhood ALL cases and 231 healthy age-matched controls contemporaneously recruited at the Texas Children's Cancer Center, at BCM in Houston, TX from 2007 to 2012. The children were less than 18 years of age at diagnosis, and exclusion criteria for both cases and controls were refusal to participate in the study and the diagnosis of any other disease or cancer. Subjects or their parents provided informed consent for provision of epidemiological data with a questionnaire and a biological sample. The DNA samples were extracted from saliva or peripheral blood samples at BCM. Race/ethnicity was determined by the responses provided on the questionnaire. Parents were requested to state the race (White, Black/African American, Asian, American Indian/Alaska Native, or Native

Hawaiian/Other Pacific Islander) and ethnicity (Hispanic/Latino or non-Hispanic/Latino) of the child. The questionnaire also collected the race/ethnicity of the parents, which was used to verify the response. Information on clinical subtype of ALL was collected from medical records.

Genotyping

The main features for the SNPs genotyped are shown in Table I. Pre-developed TaqMan® SNP Genotyping Assays (LifeTech, Foster City, CA) were used for all of the SNPs we examined. Genotyping was achieved using the Bio-Rad CFX96 real-time PCR machine (Hercules, CA). The TaqMan assays consist of singleplex reactions carried out in ninety-six well plates. Each plate contained two no template controls (NTCs), a positive control, and random replicate samples. Bio-Rad SsoFast™ Probes Supermix, a 2x reaction buffer which contains the necessary components for running the PCR; Sso7d-fusion polymerase, dNTPs, $MgCI_2$, and stabilizers, was used with the TaqMan Assay. PCR amplifications were performed using the manufacturer's suggestion of 20 μL total volume and with the following PCR thermal cycling conditions: enzyme activation at 95^oC for two minutes, and 49 cycles of denaturation at 95^oC for 5 seconds followed by annealing and extension at 61^oC for 5 seconds. Bio-Rad CFX Manager software (version 3.0) was used for data acquisition and genotype assignment.

Statistical analysis

Statistical analyses were performed using Stata v.11 (StataCorp, College Station, TX). Pearson's X^2, Student's t-test (for means) or median test (for medians) were used to compare characteristics between the cases and controls. Logistic regression methods were used to calculate crude and adjusted odds ratios (OR) and 95% confidence intervals (CIs).

All statistical tests were two-tailed, and the threshold for statistical significance was set at $P \leq 0.05$. The ORs, with 95% CIs, were used as a measure of effect size. Genotype counts were tested for Hardy–Weinberg equilibrium (HWE) in controls for each SNP. By default, we used the additive genetic model to assess associations by Cochrane-Armitage trend test. Due to its previous association being in the recessive model, rs2395185 was analyzed also for the recessive model association. Ethnic- and gender-specific associations were calculated through stratified analyses. None of the SNPs were located in coding regions, therefore we used RegulomeDB (Stanford University, Palo Alto, CA) to calculate a score for regulatory function. To adjust for heterogeneity in our sample, especially in Hispanics, we used two ancestry-informative markers, AIMs, to control for confounding caused by population stratification and to avoid spurious associations [52]. The two AIMs we used, rs285 and rs2891, have been previously used in Hispanic populations to account for the differences in genetic ancestry [53,54]. The risk associations were adjusted by each of the AIMS.

Results

All cases and controls were genotyped for the eight candidate SNPs and two ancestry-informative markers (AIMs). Genotype call rates were greater than 95% for both cases and controls. Table II shows characteristics of the case-control sample. The case samples included 86 males (53%) and 75 females (47%). Out of these cases, 66 identified themselves as non-Hispanic White, 72 as Hispanic White, 17 as Black, and 6 as "other." The group labeled "other" included those identifying themselves as Asian, Native American, or other. The healthy controls included 130 males (56%) and 101 females (44%), who had visited the pediatric clinic at BCM for a non-disease related reason.

Forty-nine were classified as non-Hispanic White, 98 as Hispanic White, and 78 as Black. The distribution of ethnic background was different between cases and controls mainly due to the infrequency of childhood ALL. Because of this difference, results were adjusted for ethnic background or stratified analyses were performed when necessary. There was no significant difference in the means between the case/control groups in regards to birth weight. Birth length and gestational age did not differ between cases and controls, except in specific subgroups. There was a significant difference in gestational age between cases and controls, only in the Hispanic female subgroup. The difference in birth length means was significant, but only in the non-Hispanic White subgroup. Eighty-eight percent of the cases were diagnosed with early precursor B (early pre-B) ALL subtype, and associations did not change in effect size depending on the molecular subtype. There were no significant genotype associations found within the Black subpopulation, possibly due to the small number of cases in the population sample.

GWAS risk markers

Results including genotype frequencies and odds ratios are described in Tables III-V. Analyses yielded significant associations with some of the genetic markers similar to previous reports. In total, three *ARID5B* SNPs were examined for associations. Two of the *ARID5B* SNPs, rs7089424 and rs10821936, showed expected risk associations, while rs10994982 did not show an association (Table III). The SNP rs7089424 had an overall odds ratio per allele (OR_{allele}) of 1.69 (P>0.001). The association showed a somewhat stronger risk in the non-Hispanic subgroup (OR_{allele}=2.11, P=0.01), compared with Hispanics (OR_{allele}=1.61, P=0.02). Similarly, rs10821936 had an overall OR_{allele}=1.48

(*P*=0.05). Adjustments of the analyses for race/ethnicity did not change the observed results. Adjustment for AIMs did not appreciably alter the results either.

There were no significant associations found for *IKZF1* rs4132601 or for *CEBPE* rs2239633. Hardy-Weinberg equilibrium was violated in controls for rs4132601, which could lead to spurious results, and was therefore excluded.

HLA region lymphoma risk markers

The SNP located in the HLA-DR region, rs2395185, showed a weak, non-significant risk overall (Table IV). This SNP is an exclusive marker for the HLA-DRB4 (DR53) lineage [38]. Since this lineage was shown to be a risk marker for childhood ALL, but only in males [46], we examined rs2395185 association in males. The non-Hispanic White male group had an OR_{allele} of 2.79 (*P*=0.016). Figure I depicts the childhood ALL risk in rs2395185 subgroup analysis, focusing on males and non-Hispanic White males, specifically. The OR reached 6.21 (95% CI=0.70-54.96) for homoyzgosity for the variant allele, which corresponds to the original association [46]. The known protective marker for follicular lymphoma, rs2647012, showed a statistically significant association in Hispanics OR_{allele} =2.21 (*P*=0.007), but not in non-Hispanics ($P_{interaction}$ = 0.003 for ethnicity), as shown in Figure II. The significance remained after adjusting for both rs23951885 and rs10484561. The variant rs10484561 was shown to be a strong protective marker in this study, opposite of what was found in follicular lymphoma [50]. Using the recessive model, the variant allele showed a strong association, OR_{rec}=0.19 (*P*=0.009). The two AIMs used, rs285 and rs2891, did not show any association as expected (Table V).

Discussion

Despite being well established, the racial disparity in the incidence of childhood ALL is not always addressed in genetic association studies. Most GWAS, until recently, have identified risk loci using only European-origin populations. Variant polymorphisms located within the *ARID5B*, *IKZF1*, and *CEBPE* genes have reported strong risk associations in multiple GWAS [34-37,55-57]. Our study provides some confirmation of previously discovered genetic markers associated with childhood ALL, which also validated our case-control set for further exploration.

Of the three *ARID5B* SNPs, rs7089424 and rs10821936 showed significant risk associations, and rs10994982 did not show any associations. The *ARID5B* gene is involved in transcriptional regulation with embryonic development [55]. Overexpression of the gene in particular acute leukemias have led some to speculate that variations within the gene may affect B-lineage development, and increase susceptibility to B-lineage leukemia [35]. The marker rs7089424 was associated with a stronger risk of leukemia in the non-Hispanic subgroup compared with Hispanics, replicating recent findings from case-control studies using Hispanic populations [8,37].

In the recent multi-ethnic GWAS by Xu *et al*., rs10821936 was found to be a significant risk marker across all ethnicities. Xu *et al*. noted that the risk allele frequencies for rs10821936 increased in order by race incident rates: Black/African Americans, non-Hispanic/European American, and Hispanic Americans [37]. Our results showed a similar trend with an increasing risk allele frequency (RAF) in cases of Blacks, non-Hispanic Whites and Hispanics. The multi-ethnic GWAS reported that rs10821936 was highly correlated with Native American genetic ancestry [37], substantiating their

previous observations that Native American ancestry correlates with higher risk of relapse in Hispanics, and leading to speculation that this may be a factor with the increased risk for Hispanic children who have a high proportion of Native American ancestry [58].

The SNP rs4132601, located in the Ikaros family zinc finger 1 (*IKZF1*) gene, is associated with increased risk of childhood ALL in multiple studies [34,36,56,57,59]. The Ikaros proteins are known to be involved with lymphocyte development and differentiation [34], and deletions are frequent and associated with unfavorable prognosis in B-cell precursor ALL [34, 60]. One study found this variant to be a significant risk marker amongst non-Hispanic Whites, but not in Hispanics, despite similar allele frequencies [57]. Chokkalingam *et al*. hypothesized that that this marker's association may be due to linkage disequilibrium with a functional variant, and because of admixture in Hispanic populations the linkage disequilibrium may vary [57]. Our study was unable to examine this SNP, due to Hardy-Weinberg disequilibrium found in controls, even after stratification for race/ethnicity.

The present study confirmed some, but not all previous findings of GWAS. With the *ARID5B* risk SNPs, there was heterogeneity even between the first two GWAS reports [34,35]. The modest sample size we had also reduced the statistical power of our study. Nevertheless, confirmed results validated the present case-control sample for further genetic association studies.

The relevance of lymphoma-associated polymorphisms in childhood ALL was assessed by genotyping rs2395185, rs10484561, and rs2647012. The SNP near the HLA-DRA gene, rs2395185, is a marker for the HLA-DRB4 (DR53) lineage [38]. The HLA-

DRB4 lineage or its marker SNP have been previously shown as a risk marker in lung cancer [39], asthma [40], rheumatoid arthritis [49], type I diabetes [48], adult acute myeloblastic leukemia [61], chronic myeloid leukemia [62], chronic lymphoid leukemia [63-65] and in childhood ALL (males only) [46,66,67], and as a protective marker for non-Hodgkin lymphoma [41], and ulcerative colitis [42-44,47]. The DRB4/DR53 lineage has been shown previously to have a risk association with childhood ALL, with male specificity, within a European sample via HLA typing [46]. The first GWAS association of rs2395185 was with ulcerative colitis [42,47]. The variant allele, T, was later found to be a protective marker in a GWAS examining risk factors for classical Hodgkin lymphoma [41], and most recently a risk marker factor in Asian females for lung cancer [39]. Our results replicated the strong male specificity of the risk for childhood ALL, specifically in non-Hispanic White males, with no association in Hispanics. The DRB4 lineage has unique features, such as lower expression levels of HLA-DR molecules, poor interaction with CD4, disrupted intracellular transport, and possibly contains extra amount of DNA in the DR/DQ region which may contribute to this risk association in childhood ALL [45]. One important finding of the present study is that the risk modifiers of lymphoma showed opposite associations in childhood ALL.

The variant rs2647012, a marker for DRB3/DRB5 lineages and a protective marker for lymphoma, showed a statistically significant risk association among Hispanics, but not in non-Hispanics ($P_{\text{interaction}} = 0.003$ for ethnicity). Our study appears to be the first looking at a Hispanic population with this SNP, and further studies are warranted to determine if this inverse relationship of risk exists in non-Hispanic and

Hispanic Whites. Variant rs10484561 showed signs of being a protective marker for childhood ALL, opposite to the follicular lymphoma findings [50].

It is now customary that the additive model is used to assess statistical associations of SNPs. While the additive model has sufficient power to detect associations in most situations, there are certain scenarios that it may not show statistical significance when in fact, there is an association. An association conforming to the strictly recessive model when allele frequency is low is one example [68-70]. For this reason, and also because the original rs2395185 association was a recessive model association, we also assessed this genetic model in HLA region associations. This approach consistently yielded larger effect sizes for HLA region SNPs, especially for rs2395185 as in previous studies. We are in favor of routine use of the recessive model analysis in exploration of associations in the HLA region.

There are limitations of our study. With childhood ALL being a rare disease, the sample size for our study was small, resulting in limited statistical power. The issue of self-reported ethnicity may also be of concern. This method is common with population-based association studies, and residual confounding is often suspected. Even though it has been described that self-reported ethnicity may be reliable [71], the heterogeneity within the Hispanic population is still a concern. A recent study conducted in a Spanish population was unable to replicate original risk associations found in Hispanic Americans, demonstrating the large heterogeneity in this high risk group [37,72].

To adjust for heterogeneity in our sample, especially in Hispanics, we used two ancestry-informative markers, AIMs, to control for confounding caused by population stratification and to avoid spurious associations [52]. The AIMs have widely different

allele frequencies in major human continental groups. The two AIMs we used, rs285 and rs2891, have been previously used in Hispanic populations to account for the differences in genetic ancestry [53,54]. Adjusting the risk associations by AIMs did not alter the results.

Our study did have a well-defined phenotype, with molecular ALL types determined. The use of a multi-ethnic sample population was a strength, especially for the ethnic disparity that persists with childhood ALL. Another strength is the use of multiple genetic models, where appropriate, to determine associations that may remain undetectable by the exclusive use of the additive model association. The replication of known leukemia markers validated our sample set for further studies. This pilot study is part of an ongoing effort at BCM. Recruitment for the second phase is continuing together with clinical follow-up.

In summary, we validated a new multi-ethnic case-control set and also examined some new markers with their association with childhood ALL. The examination of lymphoma risk markers yielded associations in opposite directions for childhood ALL, but also confirmed a previously identified childhood ALL risk marker. Two HLA region associations were ethnicity-specific. Still, our study did not provide clues as to why Hispanics have a higher susceptibility to childhood ALL, suggesting that environmental factors may play a stronger role in this differential. Studies with information on environmental exposures may help explain how gene-environment interactions contribute to childhood ALL susceptibility and its variation among different populations.

REFERENCES

1. Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: The impact of eliminating socioeconomic and racial disparities on premature cancer deaths. CA Cancer J Clin 2011;61:212-36.

2. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. Nat Rev Cancer 2006;6:193-203.

3. Mody R, Li S, Dover DC, Sallan S, Leisenring W, Oeffinger KC, Yasui Y, Robison LL, Neglia JP. Twenty-five-year follow-up among survivors of childhood acute lymphoblastic leukemia: a report from the Childhood Cancer Survivor Study. Blood 2008;111:5515-23.

4. Sherborne AL, Hemminki K, Kumar R, Bartram CR, Stanulla M, Schrappe M, Petridou E, Semsei AF, Szalai C, Sinnett D and others. Rationale for an international consortium to study inherited genetic susceptibility to childhood acute lymphoblastic leukemia. Haematologica 2011;96:1049-54.

5. Howlader N NA, Krapcho M, Garshell J, Neyman N, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2010. Bethesda, MD: National Cancer Institute; 2013.

6. McNeil DE, Cote TR, Clegg L, Mauer A. SEER update of incidence and trends in pediatric malignancies: acute lymphoblastic leukemia. Med Pediatr Oncol 2002;39:554-7; discussion 552-3.

7. Matasar MJ, Ritchie EK, Consedine N, Magai C, Neugut AI. Incidence rates of the major leukemia subtypes among US Hispanics, Blacks, and non-Hispanic Whites. Leuk Lymphoma 2006;47:2365-70.

8. Xu H, Cheng C, Devidas M, Pei D, Fan Y, Yang W, Neale G, Scheet P, Burchard EG, Torgerson DG and others. ARID5B genetic polymorphisms contribute to racial disparities in the incidence and treatment outcome of childhood acute lymphoblastic leukemia. J Clin Oncol 2012;30:751-7.

9. Swinney RM, Beuten J, Collier AB, 3rd, Chen TT, Winick NJ, Pollock BH, Tomlinson GE. Polymorphisms in CYP1A1 and ethnic-specific susceptibility to acute lymphoblastic leukemia in children. Cancer Epidemiol Biomarkers Prev 2011;20:1537-42.

10. Birch JM. Genes and cancer. Arch Dis Child 1999;80:1-3.

11. Pui CH. Childhood leukemias. N Engl J Med 1995;332:1618-30.

12. Belson M, Kingsley B, Holmes A. Risk factors for acute leukemia in children: a review. Environ Health Perspect 2007;115:138-45.

13. Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. Lancet 2008;371:1030-43.

14. Feller M, Adam M, Zwahlen M, Brazzola P, Niggli F, Kuehni C. Family characteristics as risk factors for childhood acute lymphoblastic leukemia: a population-based case-control study. PLoS One 2010;5.

15. Greaves MF. Aetiology of acute leukaemia. Lancet 1997;349:344-9.

16. Robison LL, Codd M, Gunderson P, Neglia JP, Smithson WA, King FL. Birth weight as a risk factor for childhood acute lymphoblastic leukemia. Pediatr Hematol Oncol 1987;4:63-72.

17. Kaye SA, Robison LL, Smithson WA, Gunderson P, King FL, Neglia JP. Maternal reproductive history and birth characteristics in childhood acute lymphoblastic leukemia. Cancer 1991;68:1351-1355.

18. Petridou E, Trichopoulos D, Kalapothaki V, Pourtsidis A, Kogevinas M, Kalmanti M, Koliouskas D, Kosmidis H, Panagiotou JP, Piperopoulou F and others. The risk profile of childhood leukaemia in Greece: a nationwide case-control study. Br J Cancer 1997;76:1241-7.

19. Ross JA, Potter JD, Shu XO, Reaman GH, Lampkin B, Robison LL. Evaluating the relationships among maternal reproductive history, birth characteristics, and infant leukemia: a report from the Children's Cancer Group. Ann Epidemiol 1997;7:172-179.

20. Westergaard T, Andersen PK, Pedersen JB, Olsen JH, Frisch M, Sorensen HT, Wohlfahrt J, Melbye M. Birth characteristics, sibling patterns, and acute leukemia risk in childhood: a population-based cohort study. J Natl Cancer Inst 1997;89:939-947.

21. Yeazel MW, Ross JA, Buckley JD, Woods WG, Ruccione K, Robison LL. High birth weight and risk of specific childhood cancers: a report from the Children's Cancer Group. J Pediatr 1997;131:671-7.

22. Murray L, McCarron P, Bailie K, Middleton R, Davey Smith G, Dempsey S, McCarthy A, Gavin A. Association of early life factors and acute lymphoblastic leukaemia in childhood: historical cohort study. Br J Cancer 2002;86:356-61.

23. Okcu MF, Goodman KJ, Carozza SE, Weiss NS, Burau KD, Bleyer WA, Cooper SP. Birth weight, ethnicity, and occurrence of cancer in children: a population-based, incident case-control study in the State of Texas, USA. Cancer Causes Control 2002;13:595-602.

24. Hjalgrim LL, Westergaard T, Rostgaard K, Schmiegelow K, Melbye M, Hjalgrim H, Engels EA. Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. Am J Epidemiol 2003;158:724-35.

25. Milne E, Royle JA, de Klerk NH, Blair E, Bailey H, Cole C, Attia J, Scott RJ, Armstrong BK. Fetal growth and risk of childhood acute lymphoblastic leukemia: results from an Australian case-control study. Am J Epidemiol 2009;170:221-8.

26. Hjalgrim LL, Rostgaard K, Hjalgrim H, Westergaard T, Thomassen H, Forestier E, Gustafsson G, Kristinsson J, Melbye M, Schmiegelow K. Birth weight and risk for childhood leukemia in Denmark, Sweden, Norway, and Iceland. J Natl Cancer Inst 2004;96:1549-56.

27. McLaughlin CC, Baptiste MS, Schymura MJ, Nasca PC, Zdeb MS. Birth weight, maternal weight and childhood leukaemia. Br J Cancer 2006;94:1738-44.

28. Dorak MT, Hammal DM, Pearce MS, McNally RJ, Parker L. Examination of gender effect in birth weight and miscarriage associations with childhood cancer. Cancer Causes Control 2007;18:219-228.

29. Koifman S, Pombo-de-Oliveira MS. High birth weight as an important risk factor for infant leukemia. Br J Cancer 2008;98:664-7.

30. Caughey RW, Michels KB. Birth weight and childhood leukemia: A meta-analysis and review of the current evidence. Int J Cancer 2009;124:2658-70.

31. Dorak MT, Mackay RK, Relton CL, Worwood M, Parker L, Hall AG. Hereditary hemochromatosis gene (HFE) variants are associated with birth weight and childhood leukemia risk Pediatric Blood Cancer 2009;53:1242-8.

32. Sprehe MR, Barahmani N, Cao Y, Wang T, Forman MR, Bondy M, Okcu MF. Comparison of birth weight corrected for gestational age and birth weight alone in prediction of development of childhood leukemia and central nervous system tumors. Pediatr Blood Cancer 2010;54:242-9.

33. Oksuzyan S, Crespi CM, Cockburn M, Mezei G, Kheifets L. Birth weight and other perinatal characteristics and childhood leukemia in California. Cancer Epidemiol 2012;36:e359-65.

34. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, Kinsey SE, Lightfoot T, Roman E, Irving JA and others. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1006-10.

35. Trevino LR, Yang W, French D, Hunger SP, Carroll WL, Devidas M, Willman C, Neale G, Downing J, Raimondi SC and others. Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1001-5.

36. Orsi L, Rudant J, Bonaventure A, Goujon-Bellec S, Corda E, Evans TJ, Petit A, Bertrand Y, Nelken B, Robert A and others. Genetic polymorphisms and childhood acute lymphoblastic leukemia: GWAS of the ESCALE study (SFCE). Leukemia 2012;26:2561-4.

37. Xu H, Yang W, Perez-Andreu V, Devidas M, Fan Y, Cheng C, Pei D, Scheet P, Burchard EG, Eng C and others. Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. J Natl Cancer Inst 2013;105:733-42.

38. Kennedy AE, Singh SK, Dorak MT. Re: genome-wide association study of classical hodgkin lymphoma and epstein-barr virus status-defined subgroups. J Natl Cancer Inst 2012;104:884-5.

39. Lan Q, Hsiung CA, Matsuo K, Hong YC, Seow A, Wang Z, Hosgood HD, 3rd, Chen K, Wang JC, Chatterjee N and others. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. Nat Genet 2012;44:1330-5.

40. Li X, Ampleford EJ, Howard TD, Moore WC, Torgerson DG, Li H, Busse WW, Castro M, Erzurum SC, Israel E and others. Genome-wide association studies of asthma indicate opposite immunopathogenesis direction from autoimmune diseases. J Allergy Clin Immunol 2012;130:861-8 e7.

41. Urayama KY, Jarrett RF, Hjalgrim H, Diepstra A, Kamatani Y, Chabrier A, Gaborieau V, Boland A, Nieters A, Becker N and others. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. J Natl Cancer Inst 2012;104:240-53.

42. Asano K, Matsushita T, Umeno J, Hosono N, Takahashi A, Kawaguchi T, Matsumoto T, Matsui T, Kakuta Y, Kinouchi Y and others. A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. Nat Genet 2009;41:1325-9.

43. Latiano A, Palmieri O, Latiano T, Corritore G, Bossa F, Martino G, Biscaglia G, Scimeca D, Valvano MR, Pastore M and others. Investigation of multiple susceptibility loci for inflammatory bowel disease in an Italian cohort of patients. PLoS One 2011;6:e22688.

44. Juyal G, Prasad P, Senapati S, Midha V, Sood A, Amre D, Juyal RC, Bk T. An investigation of genome-wide studies reported susceptibility loci for ulcerative colitis shows limited replication in north Indians. PLoS One 2011;6:e16565.

45. Dorak MT, Oguz FS, Yalman N, Diler AS, Kalayoglu S, Anak S, Sargin D, Carin M. A male-specific increase in the HLA-DRB4 (DR53) frequency in high-risk and relapsed childhood ALL. Leuk Res 2002;26:651-6.

46. Dorak MT, Lawson T, Machulla HK, Darke C, Mills KI, Burnett AK. Unravelling an HLA-DR association in childhood acute lymphoblastic leukemia. Blood 1999;94:694-700.

47. Silverberg MS, Cho JH, Rioux JD, McGovern DP, Wu J, Annese V, Achkar JP, Goyette P, Scott R, Xu W and others. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. Nat Genet 2009;41:216-20.

48. Nakanishi K, Shima Y. Capture of type 1 diabetes-susceptible HLA DR-DQ haplotypes in Japanese subjects using a tag single nucleotide polymorphism. Diabetes Care 2010;33:162-4.

49. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burtt NP, Blumenstiel B, DeFelice M and others. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007;39:1477-82.

50. Conde L, Halperin E, Akers NK, Brown KM, Smedby KE, Rothman N, Nieters A, Slager SL, Brooks-Wilson A, Agana L and others. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nat Genet 2010;42:661-4.

51. Smedby KE, Foo JN, Skibola CF, Darabi H, Conde L, Hjalgrim H, Kumar V, Chang ET, Rothman N, Cerhan JR and others. GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. PLoS Genet 2011;7:e1001378.

52. Enoch MA, Shen PH, Xu K, Hodgkinson C, Goldman D. Using ancestry-informative markers to define populations and detect population stratification. J Psychopharmacol 2006;20:19-26.

53. Ziv E, John EM, Choudhry S, Kho J, Lorizio W, Perez-Stable EJ, Burchard EG. Genetic ancestry and risk factors for breast cancer among Latinas in the San Francisco Bay Area. Cancer Epidemiol Biomarkers Prev 2006;15:1878-85.

54. Lee YL, Teitelbaum S, Wolff MS, Wetmur JG, Chen J. Comparing genetic ancestry and self-reported race/ethnicity in a multiethnic population in New York City. J Genet 2010;89:417-23.

55. Healy J, Richer C, Bourgey M, Kritikou EA, Sinnett D. Replication analysis confirms the association of ARID5B with childhood B-cell acute lymphoblastic leukemia. Haematologica 2010;95:1608-11.

56. Prasad RB, Hosking FJ, Vijayakrishnan J, Papaemmanuil E, Kohler R, Greaves MF, Sheridan E, Gast A, Kinsey SE, Lightfoot T and others. Verification of the susceptibility loci on 7p12.2, 10q21.2 and 14q11.2 in precursor B-cell acute lymphoblastic leukemia of childhood. Blood 2010;In Press.

57. Chokkalingam AP, Hsu LI, Metayer C, Hansen HM, Month SR, Barcellos LF, Wiemels JL, Buffler PA. Genetic variants in ARID5B and CEBPE are childhood ALL susceptibility loci in Hispanics. Cancer Causes Control 2013.

58. Yang JJ, Cheng C, Devidas M, Cao X, Fan Y, Campana D, Yang W, Neale G, Cox NJ, Scheet P and others. Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. Nat Genet 2011;43:237-41.

59. Lautner-Csorba O, Gezsi A, Semsei AF, Antal P, Erdelyi DJ, Schermann G, Kutszegi N, Csordas K, Hegyi M, Kovacs G and others. Candidate gene association study in pediatric acute lymphoblastic leukemia evaluated by Bayesian network based Bayesian multilevel analysis of relevance. BMC Med Genomics 2012;5:42.

60. Mullighan CG, Downing JR. Genome-wide profiling of genetic alterations in acute lymphoblastic leukemia: recent insights and future directions. Leukemia 2009;23:1209-18.

61. Seremetis S, Cuttner J, Winchester R. Definition of a possible genetic basis for susceptibility to acute myelogenous leukemia associated with the presence of a polymorphic Ia epitope. J Clin Invest 1985;76:1391-7.

62. Oguz FS, Kalayoglu S, Diler AS, Tozkir H, Sargin D, Carin M, Dorak MT. HLA system affects the age-at-onset in chronic myeloid leukemia. Am J Hematol 2003;73:256-62.

63. Machulla HK, Muller LP, Schaaf A, Kujat G, Schonermarck U, Langner J. Association of chronic lymphocytic leukemia with specific alleles of the HLA-DR4:DR53:DQ8 haplotype in German patients. Int J Cancer 2001;92:203-7.

64. Dyer PA, Ridway JC, Flanagan NG. HLA-A,B and DR antigens in chronic lymphocytic leukaemia. Dis Markers 1986;4:231-7.

65. Dorak MT, Machulla HK, Hentschel M, Mills KI, Langner J, Burnett AK. Influence of the major histocompatibility complex on age at onset of chronic lymphoid leukaemia. Int J Cancer 1996;65:134-9.

66. Dorak MT, Owen G, Galbraith I, Henderson N, Webb D, Mills KI, Darke C, Burnett AK. Nature of HLA-associated predisposition to childhood acute lymphoblastic leukemia. Leukemia 1995;9:875-8.

67. Dorak MT, Chalmers EA, Gaffney D, Wilson DW, Galbraith I, Henderson N, Worwood M, Mills KI, Burnett AK. Human major histocompatibility complex contains several leukemia susceptibility genes. Leuk Lymphoma 1994;12:211-22.

68. Zheng G, Joo J, Lin JP, Stylianou M, Waclawiw MA, Geller NL. Robust ranks of true associations in genome-wide case-control association studies. BMC Proc 2007;1 Suppl 1:S165.

69. Lettre G, Lange C, Hirschhorn JN. Genetic model testing and statistical power in population-based association studies of quantitative traits. Genet Epidemiol 2007;31:358-62.

70. Freidlin B, Zheng G, Li Z, Gastwirth JL. Trend tests for case-control studies of genetic markers: power, sample size and robustness. Hum Hered 2002;53:146-52.

71. Mez JB, Cole JW, Howard TD, Macclellan LR, Stine OC, O'Connell JR, Wozniak MA, Stern BJ, Sorkin JD, Mitchell BD and others. Evaluation of self-reported ethnicity in a case-control population: the stroke prevention in young women study. BMC Res Notes 2009;2:260.

72. Lopez-Lopez E, Gutierrez-Camino A, Martin-Guerrero I, Garcia-Orad A. Re: Novel Susceptibility Variants at 10p12.31-12.2 for Childhood Acute Lymphoblastic Leukemia in Ethnically Diverse Populations. J Natl Cancer Inst 2013.

Table I. Main features of SNPs analyzed

Gene	SNP	Chromosome nucleotide position*	Inclusion criteria	PMID**	Minor allele and frequency†	Location	RegulomeDB score††
ARID5B	rs7089424	chr10: 63752159	GWAS identified risk loci for childhood ALL	19684604, 22660188, 20042726	(G) 0.314	Intronic	3a
ARID5B	rs10821936	chr10: 63723577	GWAS identified risk loci for childhood ALL	19684603, 20054350, 22660188, 23512250, 22291082	(C) 0.318	Intronic	5
ARID5B	rs10994982	chr10: 63710104	GWAS identified risk loci for childhood ALL	19684603, 22660188	(A) 0.457	Intronic	NA
IKZF1	rs4132601	chr7: 50470604	GWAS identified risk loci for childhood ALL	19684604, 22660188, 20054350	(G) 0.306	3'-UTR	5
CEBPE	rs2239633	chr14: 23589057	GWAS identified risk loci for childhood ALL	19684604, 22660188	(A) 0.466	5'-upstream	4
HLA-DR region	rs2395185	chr6: 32433167	HLA-DRB4/DR53 lineage; ALL risk marker (in European males); Hodgkin lymphoma risk marker	10397736, 12008082, 22286212, 7909466	(T) 0.423	Intronic	6
HLA-DQB1 region	rs2647012	chr6: 32664458	HLA-DRB3/DRB5 lineage, protective marker for non-Hodgkin (follicular) lymphoma	21533074, 22911334, 23455380	(T) 0.381	Intergenic	6
HLA-DQA1 region	rs10484561	chr6: 32665420	HLA-DR1/DR10 lineage, risk marker in follicular lymphoma	20639881, 21533074, 23025665	(G) 0.084	Intergenic	6

*Genome Reference Consortium Human Build 37 patch release 10 (GRCh37.p10) used for nucleotide position (http://www.ncbi.nlm.nih.gov/SNP/)

**PubMed identifier number

†Minor allele frequencies are from a reference Caucasian population (U.S. residents of northern and western European ancestry) genotyped in HapMap project

††RegulomeDB scores range from 1 (most functional) to 5 (least functional) (6=other). Not all SNPs have a RegulomeDB score (http://regulome.stanford.edu/)

Table II. Characteristics of cases and controls

	Cases n=161	**Controls** n=231	***P* value**
Ethnic background			
Non-Hispanic White	66	49	
Hispanic White	72	98	**<0.001**
Black	17	78	
Other*	6	6	
Gender			
Male	86	130	0.58
Female	75	101	
Ratio	1.15	1.29	
Birth weight (*grams*)			
Mean (SD)	3349.3 (584)	3263.3 (684)	0.23
Median (IQR)	3400 (760)	3311.5 (850)	0.48
Gestational age *(weeks)* (SD)	38.8 (2.2)	38.2 (2.8)	**0.04**
Birth length *(cm)* (SD)	51.4 (3.4)	50.1 (5.05)	**0.024**

*Other includes Asian, Native American, or other

Table III. SNP associations previously shown as ALL risk markers in GWAS (overall*)

			Case Subjects			Control Subjects			Minor Allele Frequency			
Gene	**SNP**	**Minor Allele**	**A/A**	**A/B**	**B/B**	**A/A**	**A/B**	**B/B**	**Case**	**Control**	**OR_{allele}**(95% CI)**	***P* value**
ARID5B	**rs7089424**	G	37	80	43	113	68	46	0.52	0.35	**1.69 (1.28-2.24)**	**<0.001**
ARID5B	**rs10821936**	C	39	72	44	104	69	50	0.52	0.38	**1.48 (1.12-1.95)**	**0.005**
ARID5B	**rs10994982**	A	55	78	25	90	78	43	0.41	0.39	1.00 (0.75-1.34)	0.992
IKZF1	**rs4132601**	G	76	68	16	141	56	29	0.31	0.25	1.19 (0.88-1.60)	0.269†
CEBPE	**rs2239633**	A	68	68	23	118	74	36	0.36	0.32	1.02 (0.76-1.36)	0.881

*ORs adjusted for self-reported ethnicity and race

**OR per allele (OR_{allele}) for the additive model

†Hardy-Weinberg disequilibrium in controls (overall and after stratification) for race/ethnicity

Table IV. Association of HLA region lymphoma susceptibility markers (overall*)

			Case Subjects			Control Subjects			Minor allele frequency					
SNP	**Gene**	**Minor allele**	**A/A**	**A/B**	**B/B**	**A/A**	**A/B**	**B/B**	**Case**	**Control**	**OR_{allele}**** (95% CI)	***P* value**	**OR_{rec}**† (95% CI)	***P* value**
rs2395185	HLA-DR region	T	69	68	22	120	72	25	0.35	0.28	1.27 (0.94-1.71)	0.127	1.1 (0.65-2.26)††	0.553
rs2647012	HLA-DQB1 region	A	78	64	16	116	47	30	0.3	0.28	1.09 (0.80-1.47)	0.595	0.57 (0.30-1.12)	0.103
rs10484561	HLA-DQA1 region	G	128	23	3	149	21	18	0.09	0.15	0.70 (0.46-1.06)	0.094	**0.19 (0.05-0.66)**	**0.009**

*ORs adjusted for self-reported ethnicity and race

**OR per allele (OR_{allele}) for the additive model;

†OR recessive (OR_{rec}) for the variant homozygous genotype

††OR_{allele}=1.88 (*P*=0.003) in males; OR_{allele}=2.79 (*P*=0.016) in non-Hispanic White males

Table V. Ancestry-informative marker SNPs*

				Case Subjects			Control Subjects			Minor allele frequency			
SNP	**Gene**	**Chromosome nucleotide position****	**Minor allele**	**A/A**	**A/B**	**B/B**	**A/A**	**A/B**	**B/B**	**Case**	**Control**	**OR_{allele} (95% CI)**	***P* value**
rs285	*LPL*	chr8: 19815189	C	43	68	48	92	61	53	0.52	0.41	1.26 (0.95-1.65)	0.103
rs2891	*ITGAE*	chr17: 3705526	C	56	66	37	114	72	42	0.44	0.34	1.14 (0.86-1.52)	0.362

*ORs adjusted for self-reported ethnicity and race

**Genome Reference Consortium Human Build 37 patch release 10 (GRCh37.p10) used for nucleotide position (http://www.ncbi.nlm.nih.gov/SNP/)

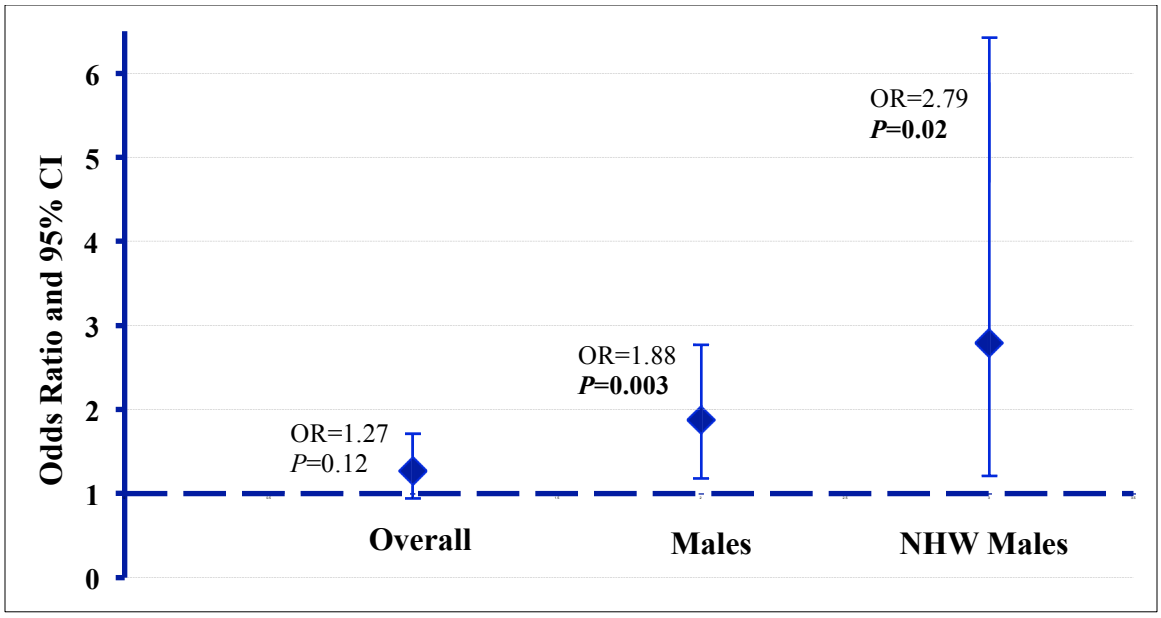


Figure I: Childhood ALL risk in rs2395185 subgroup analysis

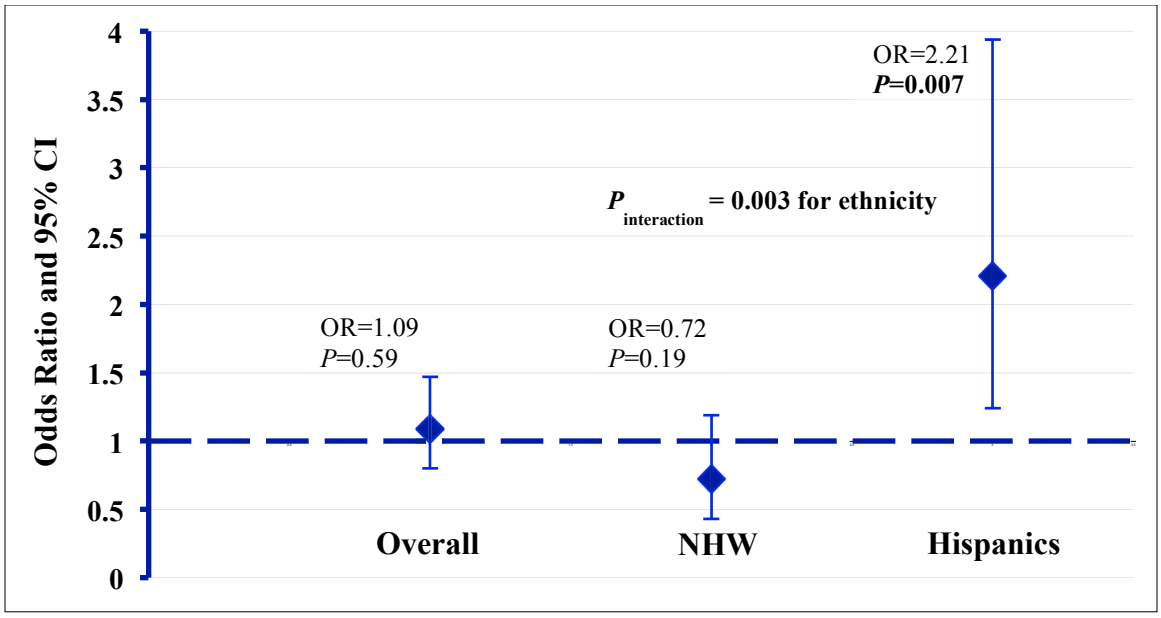


Figure II: Childhood ALL risk in rs2647012 subgroup analysis

CHAPTER V

EXAMINATION OF *HFE* ASSOCIATIONS WITH CHILDHOOD LEUKEMIA RISK AND EXTENSION TO OTHER IRON REGULATORY GENES

Abstract

Hereditary hemochromatosis (*HFE*) variants that correlate with body iron levels also show associations with cancer risk, including childhood acute lymphoblastic leukemia (ALL). Two previous studies reported *HFE* associations with different gender effects, but this association has not been evaluated in a US-based population. Using a multi-ethnic sample of cases (n=161) and controls (n=231) from Houston, TX, we examined two *HFE* variants (rs1800562/C282Y and rs1799945/H63D), one transferrin receptor gene (*TFRC*) variant (rs3817672/S142G) and three additional iron-regulatory gene (IRG) variants (*SLC11A2* rs422982; *TMPRSS6* rs855791 and rs733655) for their associations with childhood ALL. Being positive for either of the two *HFE* variants (C282Y or H63D) yielded only a modestly elevated odds ratio (OR) for childhood ALL risk in males (1.40, 95% CI = 0.83 to 2.35), which increased to 2.96 (95% CI = 1.29 to 6.80; $P = 0.01$) in the presence of a particular *TFRC* genotype for rs3817672 ($P_{interaction}$= 0.04). These findings were similar to the results reported in previous childhood ALL studies. The *TFRC* genotype also showed an ethnicity-specific association, with increased risk observed in non-Hispanic Whites (OR = 2.54, 95% CI = 1.05 to 6.12; $P = 0.04$; $P_{interaction}$ with ethnicity = 0.02). Further support to the hypothesis that elevated iron levels contribute to leukemia risk came from the three additional IRG SNPs. They all showed individual risk associations with childhood ALL in males (OR = 1.52 to 2.60). A polygenic model based on the number of variant alleles in these five IRG SNPs revealed

a linear increase in risk among males with the increasing number of variants possessed (OR = 2.00 per incremental change, 95% CI = 1.29 to 3.12; *P* = 0.002). Having three or more variants in five IRGs was associated with increased risk (OR = 4.12; *P* = 0.004) compared with having none, while corresponding OR in females was 1.22 (*P* = 0.71). Our results replicated previous *HFE* risk associations with childhood ALL in a US population and also demonstrated novel associations for IRG SNPs, thereby strengthening the hypothesis that iron excess mediated by genetic variants contribute to childhood ALL risk.

Introduction

The hereditary hemochromatosis gene, *HFE,* has shown multiple associations with cancer susceptibility [1-7], including risk childhood acute lymphoblastic leukemia (ALL) [8], which has been replicated in one other study [9]. In multiple cancers [1], including childhood ALL [9], the association of *HFE* variants with cancer risk gets stronger in interaction with a polymorphism in the transferrin receptor gene (*TFRC)*. Since HFE and TFRC proteins biologically interact in iron transfer across membranes such as intestinal mucosa and placenta, this interaction supports the notion that the involvement of *HFE* variants in cancer risk modification is mediated via their effect on body iron levels [10]. Body iron levels have long been known to be positively correlated with cancer risk as several cohort studies have shown [11-14], and iron's carcinogenic effect has been well documented [15]. Thus, *HFE* associations with cancer have strong biological plausibility.

Recent genome-wide association studies (GWAS) have identified the *HFE* variant C282Y as a major determinant of body iron levels [10]. The mediation of iron

homeostasis by genetic variants extends beyond the *HFE* gene, with the strongest association being *TMPRSS6* rs855791 [16]. We reasoned that if *HFE* associations are due to their effect on iron levels, other iron regulatory gene (IRG) polymorphisms should show similar associations with childhood ALL risk. We therefore aimed to extend our study beyond *HFE* variants to other IRG polymorphisms. To further test our hypothesis that *HFE* variants modify the risk for childhood ALL via their effects on iron levels, we also included the *TFRC* polymorphism that is known to interact with *HFE* variants in the genotyping scheme to test whether this interaction raises the risk of childhood ALL. The other IRG do not interact with *TFRC* biologically, so we did not predict any other interaction. To test our hypothesis, we used a new case-control set, which was first validated by replicating known childhood ALL associations.

Subjects & Methods

Subjects

Institutional Review Boards of Baylor College of Medicine and Florida International University approved the study protocol. The case-control sample was from Houston, TX, consisting of 161 incident cases with childhood ALL diagnosed at Texas Children's Hospital from 2007 to 2012, and 231 age-matched healthy controls contemporaneously recruited. The children were less than 18 years of age at diagnosis, and exclusion criteria for both cases and controls were refusal to participate in the study and the diagnosis of any other disease or cancer. Subjects and their parents were approached to obtain informed consent for provision of epidemiological data with a questionnaire and a biological sample. The DNA samples were extracted from saliva or peripheral blood samples at BCM. The sample was multiethnic to allow us to examine

effect modification of childhood ALL risk by race and ethnicity (Table I). Race/ethnicity was determined by the responses provided on the questionnaire by the children's parents. Our main interest was the contrast between non-Hispanic Whites (NHW) and Hispanic Whites (HW), since childhood ALL is very rare in Blacks, and we had a very small number of Blacks in the case group.

SNP selection

We included two *HFE* variants known to influence body iron levels commonly known as C282Y (rs1800562) and H63D (rs1799945), as well as the *TFRC* variant S142G (rs3817672), which is known to interact with *HFE* variants in previously reported cancer associations. As other IRG variants, we included the GWAS-identified iron-related SNP *TMPRSS6* rs855791 [16], as well as two additional SNPs we selected also from the *TMPRSS6* gene (rs733655) and the *SLC11A2* gene (rs422982) involved in the non-transferrin receptor-related iron transfer across membranes. These two SNPs were selected as the promoter region haplotype tagging SNPs for these two genes. The selected SNP from *TMPRSS6* is 32kb away and not in linkage disequilibrium with the GWAS-identified marker in the same gene according to the HapMap project European population data ($r^2 = 0.286$). Two more SNPs were included as ancestry-informative markers (AIMs) to adjust for the racial/ethnic heterogeneity in the multiethnic sample to supplement the self-reported race/ethnicity data. The two SNPs were rs285 and rs2891, which were identified as AIMs in previous studies due to their largely different allele frequencies in major ancestral human populations [17,18]. Characteristics of each SNP are given in Table II.

Genotyping

TaqMan allelic discrimination assay was the choice of method for genotyping. All SNPs were genotyped by TaqMan assays purchased from Life Technologies (Foster City, CA) on the Bio-Rad CFX96 real-time PCR instrument (Bio-Rad, Hercules, CA). The assay ID of each assay is given in Table II.

Statistical analysis

Genetic associations (both crude and adjusted) were evaluated by logistic regression using Stata v.11 (StataCorp, College Station, TX). Statistical interactions were also analyzed by logistic regression. All statistical tests were two-tailed and threshold for statistical significance was set at $P \leq 0.05$. All genetic associations, except the *TFRC* locus, were assessed by using the dominant genetic model which corresponds to variant allele positivity and coded as 1 for heterozygote and variant allele homozygote genotypes, and 0 for the common allele homozygosity (referent). Due to the low frequency of their variant alleles, the two *HFE* SNPs were pooled together by creating a new variable for the number of cumulative variant alleles at both SNPs (0 for no variant allele, 1 for variant allele at either SNP, 2 for heterozygosity at both SNPs or variant allele homozygosity at either SNP (compound heterozygosity)). To be consistent with the previous studies, *TFRC* SNP was analyzed in recessive model (by coding variant allele homozygosity as 1 and the other genotypes as 0). A similar approach was used for a polygenic risk model using the total number of variant alleles at all five IRGs (0 for no variant allele at any SNP, 1 for one or two variant alleles at any SNP, 2 for three or more variant alleles at any of the five SNPs). All statistical associations in the overall group were adjusted for the race/ethnicity variable which had four categories (NHW, HW,

Blacks, and others). The efficiency of adjustment for race/ethnicity was double-checked by adjustment for each of the AIMs separately. Before proceeding to the statistical analysis of genetic associations, Hardy-Weinberg disequilibrium was ruled out in controls as a test for genotyping error.

Results

HFE C282Y and H63D frequencies in the sample population

As expected, the *HFE* C282Y mutation was more common in the NHW subjects: variant allele frequencies were 0.113, 0.030 and 0.011 in NHWs, HWs, and Blacks, respectively. The H63D variant positivity also had similar variation across race groups with frequencies of 0.254, 0.151 and 0.032 in NHWs, HWs, and Blacks, respectively. Only two cases and two controls (all NHWs) were compound heterozygotes for *HFE* variants C282Y and H63D.

Univariable genetic markers analyses

Genotype frequencies for each SNP were in Hardy-Weinberg equilibrium in the control group when analyzed for each race/ethnicity group. All associations reported below for the whole group were adjusted for self-reported race/ethnicity. Replacing the race/ethnicity variable by either AIM did not appreciably alter the results. As shown in Table III, neither C282Y nor H63D showed an overall association with childhood ALL risk. The *TFRC* SNP, which was included in the study to assess its interaction with *HFE* SNPs did not show any individual association in the overall group. The two IRG variants selected as haplotype tagging SNPs for the respective promoter regions showed statistically significant associations in the overall analysis, but the association by the GWAS-identified marker for the iron levels did not reach statistical significance in the

overall analysis (Table III). This SNP only showed an association in males, as described below.

Gender- and race/ethnicity-specific analyses and statistical interactions

The differences in risk between genders were not significantly different, as judged by the statistical interaction test, but for the two *TMPRSS6* SNPs, males had statistically significant risk associations (Table IV). Despite yielding greater ORs for males, the male-specific *HFE* associations did not reach statistical significance in individual analysis. Results from pooling of the two *HFE* variants in one variable (as described in the methods) are shown in Figure I. The bars depict the risk genotype frequencies in the case and control groups, and ORs for childhood ALL risk are provided. The risk did not yield a statistically significant association in the overall group (OR = 1.46, P=0.17)(Graph A in Figure I), but the pooled variable was associated with a significant risk in males (OR = 2.09, P = 0.04)(Graph B in Figure I), and even stronger when interacting with the *TFRC* variant (OR = 4.92, P = 0.002)(Graph D in Figure I). Graph C depicts the frequencies and OR for females cases/controls, and Graph E shows the frequencies and OR for males with the wild type allele for *TFRC*.

There was no ethnicity-specific association of *HFE* variants. Ethnicity/race-specific analyses revealed a strong risk association (OR = 2.54, CI: 1.05-6.12, P value = 0.04) of the *TFRC* rs3817672 allele A homozygote genotype in NHWs, while the OR was less than 1.0 (non-significant) in Hispanics. The interaction of ethnicity and this *TFRC* genotype was statistically significant ($P_{interaction}$ = 0.02 for ethnicity). Thus, the only truly ethnicity-specific association with childhood ALL was the NHW-specific association of

TFRC. The strong association of the *TMPRSS6* promoter region tagging SNP rs733655 was equally strong in both major ethnicities examined in this study.

We examined the *TFRC* and *HFE* gene-gene interactions previously observed in multiple cancers [1], including childhood ALL. In the overall sample, there was no interaction (data not shown). Our main group of interest was males because of the previous findings in childhood ALL, and also because of generally higher ORs in males in univariable analysis of *HFE* variants. Since the small sample size would not allow a reliable assessment of interactions for rare *HFE* variants, especially C282Y, we used the *HFE* pooled variant variable for this analysis. In the two groups of males with and without the *TFRC* homozygous genotypes, the ORs were 0.59 (CI = 0.24 to 1.45) and 2.96 (CI = 1.29 to 6.80), yielding a statistically significant interaction ($P_{interaction}$ = 0.04). We also examined interactions of non-*HFE* SNPs with *TFRC*. This analysis did not reveal any interaction (*P* = 0.33, 0.46, and 0.96).

Polygenic risk model

We constructed a polygenic risk variable consisting of five IRG SNPs as described in the Methods section. Analysis using this variable showed that for stepwise increase in the number of variant alleles, there was a linear increase in childhood ALL risk in the overall group (OR = 1.63, 95% CI = 1.18 to 2.26, *P* = 0.003)(Figure II), and in males (OR = 2.00, 95% CI = 1.29 to 3.12, *P* = 0.002)(Figure II), but not in females (OR = 1.26, 95% CI = 0.77 to 2.08, *P* = 0.36) in stratified analyses. In ethnicity-specific analysis, the association remained statistically significant in NHW (OR = 2.19, 95% CI = 1.18 to 4.06, *P* = 0.01), but not in Hispanics (OR = 1.42, 95% CI = 0.89 to 2.27, *P* = 0.14).

Discussion

We examined previously reported *HFE* associations and interactions with *TFRC* with the risk of childhood ALL. In this multi-ethnic sample, we observed associations similar to previously reported ones with the same gender effect, and extended the observations to other iron regulatory gene polymorphisms to provide further support for our hypothesis that *HFE* and *TFRC* association in childhood ALL is due to their effect on iron homeostasis. The only statistically significant gender-specific associations with IRG variants and childhood ALL risk were in males, and we also noted a novel ethnicity-specific association with the *TFRC* variant.

Due to the limited sample size, we pooled the two *HFE* variants to increase statistical power to detect their associations with childhood ALL. The ORs for the pooled variables were always in the risk direction for individual SNPs, and were statistically significant in males. When the interaction with the *TFRC* genotype and the gender effect was considered, a more robust statistically significant association was found, as in a previous study [9]. The same interaction was also observed in multiple cancers without a gender effect [1,2]. While interaction analysis is usually seen as a challenge in terms of statistical power, as happened in the present study, the increase in the effect size may compensate for the loss of statistical power due to comparison of smaller subsets of the sample. Like any statistical association, our results should be considered cautiously. However, similarities with previous observations provide sufficient credibility to the cumulative results, which now strongly suggest that iron excess, whether environmentally- or genetically-induced increases the risk for cancer in general, and in particular childhood ALL. As previously postulated [9,19], the mechanism of the

childhood ALL risk association with *HFE* variants known to elevate body iron levels may include increased materno-fetal iron transport through placenta. This process is mediated by *HFE* and *TFRC* [20,21], and these genetically-mediated alterations in fetal iron homeostasis may also have implications on the developmental origins of health and disease [22].

Another novel finding of the present study was the risk associations of previously unexamined IRG SNPs (rs422982 and rs733655) with childhood ALL. Together with the association of rs855791, a GWAS identified marker for iron levels, which reached statistical significance only in males, these new findings lend support to our hypothesis that iron homeostasis related risk modification in childhood ALL extend beyond *HFE*/*TFRC* polymorphisms. We do not yet know whether the SNPs selected by us and used in any association study for the first time have any correlation with body iron levels. Their locations in crucial IRGs suggest that they will be somewhat involved in some aspect of gene function and subsequently in iron homeostasis, but only functional studies can confirm their roles. The lack of statistical interaction between these additional SNPs and the *TFRC* SNP suggested the specificity of the *HFE* and *TFRC* interaction. We were not surprised by the lack of interaction between *SLC11A2* / *TMPRSS6* and *TFRC* since they do not interact biologically.

The present study highlights the benefits of explorations of effect modification by gender or ethnicity. Although such explorations are usually reserved for well powered studies, when backed up by previous observations or strong biological hypotheses, stratified analyses complemented by statistical interaction analyses are powerful approaches to unravel otherwise masked associations. It is only natural that in a

multigenic disorder like cancer, effect modification, which can be seen as multicausality, will be operational. Researchers usually shy away from analysis of effect modification or statistical interaction to avoid performing multiple comparisons and subsequent chance findings, but there are ways to rule out chance findings by additional replication studies.

Besides the limitations already mentioned such as sample size and statistical power, our study has another limitation, which has stemmed from one of its strengths. Examination of these associations in a multi-ethnic cohort has benefits, but also brings about heterogeneity, which should be accounted for during analysis. We had self-reported race/ethnicity data, but adjustment of the results by these data may still leave some residual confounding. We also used two AIMs to make sure that the heterogeneity in the population will not result in spurious findings. The current practice in well-resourced GWAS studies is to use thousands of AIMs to adjust genetic ancestry, which is particularly crucial when the sample includes recently admixed populations such as African Americans or Hispanics. We could not do that, but could include a couple of AIMs to control for population heterogeneity. Another limitation of using a multi-ethnic sample was the constraints it adds on checking genotyping error. We followed the usual safeguards of genotyping error avoidance at the experimental phase, and checked for errors at the analysis phase by using Hardy-Weinberg equilibrium testing. This test, however, has to be done in each race/ethnicity subgroup separately. This practice further reduces the statistical power of this test and may have caused inefficiency of genotyping error checking. Genotyping errors have the potential to cause both false positives and false negatives. Our results basically replicated previously observed associations, and

there is no overwhelming reason to consider genotyping errors as an alternative explanation.

By replicating the *HFE* and *TFRC* interaction in childhood ALL risk association and revealing new associations with IRGs, our study provided further evidence for the hypothesis on the iron connection in childhood ALL susceptibility. These findings have far reaching implications beyond childhood leukemia in the cancer field. We hope that our results will stimulate interest in secondary analyses of existing GWAS data on multiple cancers to explore the pathways involved in iron homeostasis. We also report novel associations with gender or ethnicity specificity. These associations can also be replicated using already existing datasets. Since no childhood ALL GWAS dataset is currently available through public databases, we could not try replication, but we hope to be able to do so in the future as databases become available for additional analysis.

REFERENCES:

1. Beckman LE, Van Landeghem GF, Sikstrom C, Wahlin A, Markevarn B, Hallmans G, Lenner P, Athlin L, Stenling R, Beckman L. Interaction between haemochromatosis and transferrin receptor genes in different neoplastic disorders. Carcinogenesis 1999;20:1231-1233.

2. Beckman LE, Hagerstrand I, Stenling R, Van Landeghem GF, Beckman L. Interaction between haemochromatosis and transferrin receptor genes in hepatocellular carcinoma. Oncology 2000;59:317-22.

3. Shaheen NJ, Silverman LM, Keku T, Lawrence LB, Rohlfs EM, Martin CF, Galanko J, Sandler RS. Association between hemochromatosis (HFE) gene mutation carrier status and the risk of colon cancer. J Natl Cancer Inst 2003;95:154-9.

4. Kallianpur AR, Hall LD, Yadav M, Christman BW, Dittus RS, Haines JL, Parl FF, Summar ML. Increased prevalence of the HFE C282Y hemochromatosis allele in women with breast cancer. Cancer Epidemiol Biomarkers Prev 2004;13:205-12.

5. Dorak MT, Burnett AK, Worwood M. HFE gene mutations in susceptibility to childhood leukemia: HuGE review. Genet Med 2005;7:159-68.

6. Dorak MT. HFE H63D variant and leukemia susceptibility. Leuk Lymphoma 2006;47:2269-70.

7. Osborne NJ, Gurrin LC, Allen KJ, Constantine CC, Delatycki MB, McLaren CE, Gertig DM, Anderson GJ, Southey MC, Olynyk JK and others. HFE C282Y homozygotes are at increased risk of breast and colorectal cancer. Hepatology 2010;51:1311-8.

8. Dorak MT, Sproul AM, Gibson BE, Burnett AK, Worwood M. The C282Y mutation of HFE is another male-specific risk factor for childhood ALL. Blood 1999;94:3957-3958.

9. Dorak MT, Mackay RK, Relton CL, Worwood M, Parker L, Hall AG. Hereditary hemochromatosis gene (HFE) variants are associated with birth weight and childhood leukemia risk Pediatric Blood Cancer 2009;53:1242-8.

10. Benyamin B, McRae AF, Zhu G, Gordon S, Henders AK, Palotie A, Peltonen L, Martin NG, Montgomery GW, Whitfield JB and others. Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. Am J Hum Genet 2009;84:60-5.

11. Stevens RG, Jones DY, Micozzi MS, Taylor PR. Body iron stores and the risk of cancer. N Engl J Med 1988;319:1047-1052.

12. Knekt P, Reunanen A, Takkunen H, Aromaa A, Heliovaara M, Hakulinen T. Body iron stores and risk of cancer. Int J Cancer 1994;56:379-82.

13. Stevens RG, Graubard BI, Micozzi MS, Neriishi K, Blumberg BS. Moderate elevation of body iron level and increased risk of cancer occurrence and death. Int J Cancer 1994;56:364-369.

14. Zacharski LR, Chow BK, Howes PS, Shamayeva G, Baron JA, Dalman RL, Malenka DJ, Ozaki CK, Lavori PW. Decreased cancer risk after iron reduction in patients with peripheral arterial disease: results from a randomized trial. J Natl Cancer Inst 2008;100:996-1002.

15. Toyokuni S. Role of iron in carcinogenesis: cancer as a ferrotoxic disease. Cancer Sci 2009;100:9-16.

16. Benyamin B, Ferreira MA, Willemsen G, Gordon S, Middelberg RP, McEvoy BP, Hottenga JJ, Henders AK, Campbell MJ, Wallace L and others. Common variants in TMPRSS6 are associated with iron status and erythrocyte volume. Nat Genet 2009;41:1173-5.

17. Choudhry S, Coyle NE, Tang H, Salari K, Lind D, Clark SL, Tsai HJ, Naqvi M, Phong A, Ung N and others. Population stratification confounds genetic association studies among Latinos. Hum Genet 2006;118:652-64.

18. Tsai HJ, Kho JY, Shaikh N, Choudhry S, Naqvi M, Navarro D, Matallana H, Castro R, Lilly CM, Watson HG and others. Admixture-matched case-control study: a practical approach for genetic association studies in admixed populations. Hum Genet 2006;118:626-39.

19. Dorak MT, Hammal DM, Pearce MS, McNally RJ, Parker L. Examination of gender effect in birth weight and miscarriage associations with childhood cancer. Cancer Causes Control 2007;18:219-228.

20. Parkkila S, Waheed A, Britton RS, Bacon BR, Zhou XY, Tomatsu S, Fleming RE, Sly WS. Association of the transferrin receptor in human placenta with HFE, the protein defective in hereditary hemochromatosis. Proc Natl Acad Sci U S A 1997;94:13198-13202.

21. Balesaria S, Hanif R, Salama MF, Raja K, Bayele HK, McArdle H, Srai SK. Fetal iron levels are regulated by maternal and fetal Hfe genotype and dietary iron. Haematologica 2012;97:661-9.

22. Gluckman PD, Hanson MA, Beedle AS. Early life events and their consequences for later disease: a life history and evolutionary perspective. Am J Hum Biol 2007;19:1-19.

Table I. Characteristics of cases and controls

	Cases n=161	**Controls** n=231	***P* value**
Ethnic background			
Non-Hispanic White	66	49	
Hispanic White	72	98	**<0.001**
Black	17	78	
Other*	6	6	
Gender			
Male	86	130	0.58
Female	75	101	
Ratio	1.15	1.29	

*Asian, Native American, or other

Table II. Main features of SNPs analyzed

Gene	SNP	Chromosome nucleotide position*	Minor allele and frequency†	SNP Type	Assay ID
HFE	rs1800562	chr6: 26093141	(A) 0.053	Transition substitution, missense mutation	C___1085595_10
HFE	rs1799945	chr6: 26091179	(G) 0.179	Missense mutation, transversion substitution	C___1085600_10
TFRC	rs3817672	chr3: 195800811	(G) 0.383	Transition substitution, missense mutation	C___3259537_10
SLC11A1	rs422982	chr12: 51406354	(A) 0.246	Transversion substitution, intragenic	C____570333_10
TMPRSS6	rs733655	chr22: 37495051	(C) 0.221	Transition substitution, intragenic	C___3289858_1_
TMPRSS6	rs855791	chr22: 37462936	(T) 0.412	Transition substitution, missense mutation	C___3289902_10
LPL	rs285	chr8: 19815189	(T) 0.500	Transition substitution, intragenic	C__12104266_10
ITGAE	rs2891	chr17: 3705526	(G) 0.496	Transition substitution, intragenic	C___3211308_20

*Genome Reference Consortium Human Build 37 patch release 10 (GRCh37.p10) used for nucleotide position (http://www.ncbi.nlm.nih.gov/SNP/)

†Minor allele frequencies are from a reference Caucasian population (U.S. residents of northern and western European ancestry) genotyped in HapMap project

Table III. Univariable analyses of associations with childhood ALL risk[*]

SNP	OR (95% CI)	*P* value
***HFE* rs1800562**	1.37 (0.52 to 3.60)	0.52
***HFE* rs1799945**	1.33 (0.74 to 2.38)	0.35
***TFRC* rs3817672**	0.8 (0.52 to 1.23)	0.31
***SLC11A2* rs422982**	1.55 (1.01 to 2.37)	**0.04**
***TMPRSS6* rs733655**	2.06 (1.33 to 3.20)	**0.001**
***TMPRSS6* rs855791**	1.41 (0.91 to 2.18)	0.12

[*]Adjusted for self-reported ethnicity and race (non-Hispanic White, Hispanic White, Blacks and others)

Table IV. *HFE* and non-*HFE* associations with childhood ALL in gender and race/ethnicity groups (ORs and 95% CIs)

	Females	Males	Non-Hispanic Whites	Hispanic Whites
***HFE* rs1800562**	0.4	3.41	1.78	0.91
	$P_{interaction} = 0.09$		$P_{interaction} = 0.55$	
***HFE* rs1799945**	1.01	1.51	1.26	0.87
	$P_{interaction} = 0.78$		$P_{interaction} = 0.56$	
***TFRC* rs3817672**	1.66	0.7	2.54 (1.05 to 6.12) ***P* = 0.04**	0.69
	$P_{interaction} = 0.21$		$P_{interaction} =$ **0.02**	
***SLC11A2* rs422982**	1.58	1.52	1.47	1.91 (0.99 to 3.68)
	$P_{interaction} = 1.00$		$P_{interaction} = 0.61$	
***TMPRSS6* rs733655**	1.56	**2.6** (1.44 to 4.70) ***P* = 0.002**	**2.35** (1.07 to 5.16) ***P* = 0.03**	**2.52** (1.26 to 5.04) ***P* = 0.009**
	$P_{interaction} = 0.19$		$P_{interaction} = 0.90$	
***TMPRSS6* rs855791**	1.12	**1.91** (1.04 to 3.51) ***P* = 0.04**	1.71	1.22
	$P_{interaction} = 0.40$		$P_{interaction} = 0.52$	

*Adjusted for self-reported ethnicity and race (non-Hispanic White, Hispanic White, Blacks and others).
P values for individual analyses are only shown when <0.05

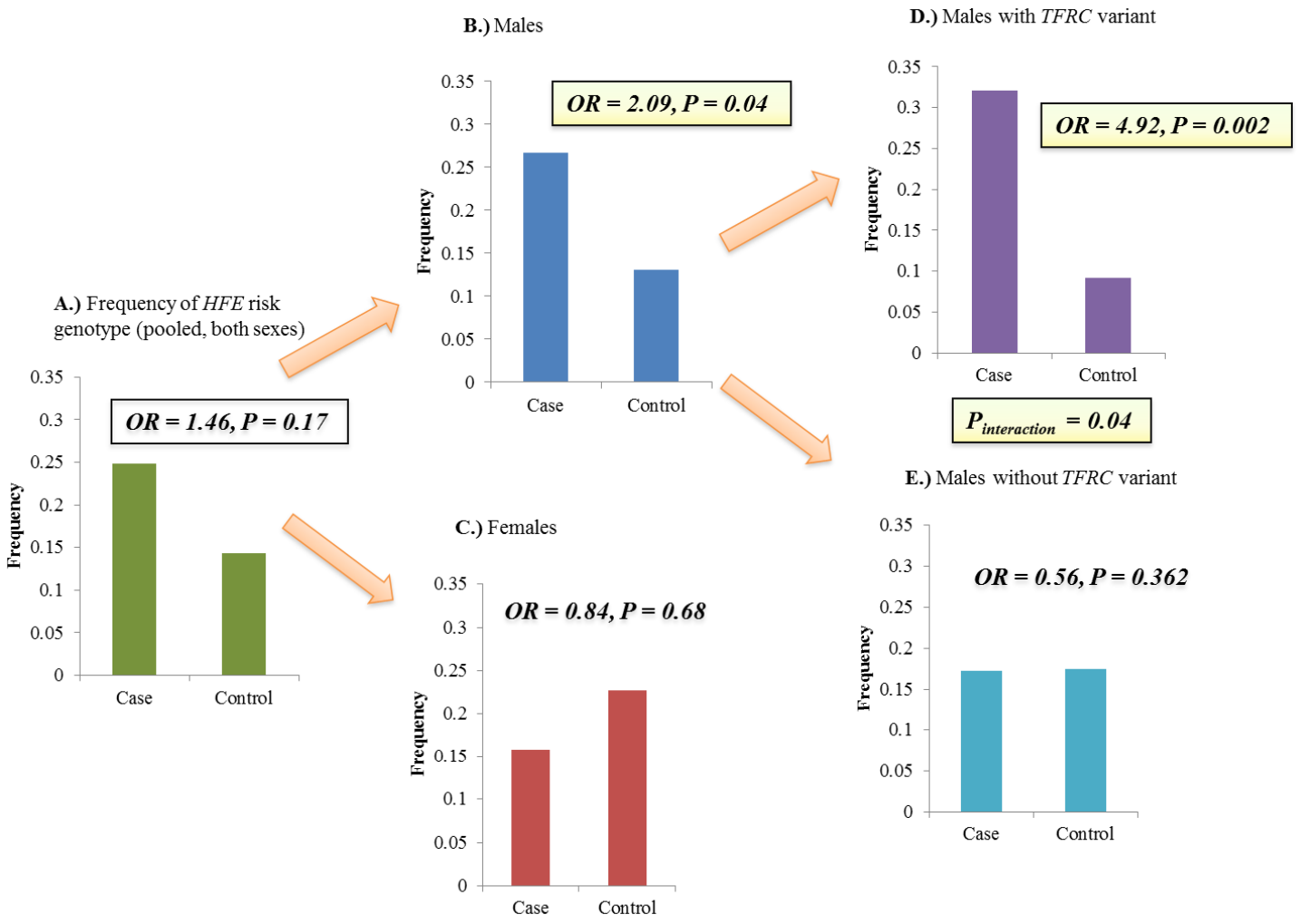
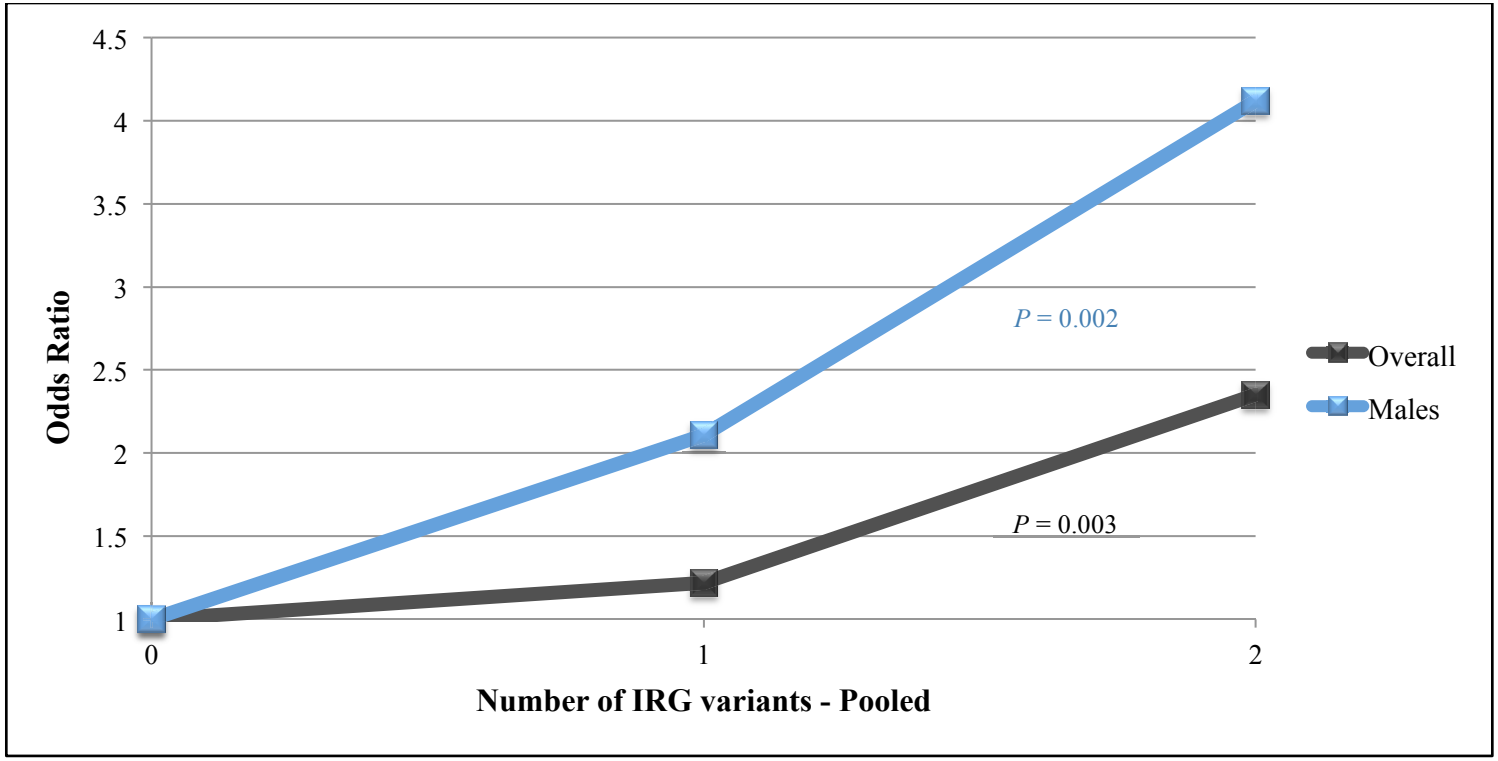


Figure I. Risk genotype frequencies with pooled *HFE* association in childhood ALL in case and control groups, by gender and *TFRC* genotype group



0 = no variants, 1 = one or two variants, 2 = three+ variants

Figure II. Polygenic risk variable consisting of five IRG SNPs and childhood ALL risk

CHAPTER VI

BIRTH CHARACTERISTICS AND CHILDHOOD LEUKEMIA RISK: CORRELATIONS WITH GENETIC MARKERS

Abstract

Birth characteristics such as birth order, birth weight, birth defects, and Down syndrome showed some of the first risk associations with childhood leukemia. Besides these non-genetic factors, recent genetic association studies have also identified markers for susceptibility. Examinations of correlations between birth characteristics and leukemia risk markers have been limited to birth weight-related genetic polymorphisms. We aimed to revisit associations of birth characteristics with childhood acute lymphoblastic leukemia (ALL) as well as their correlations with selected ALL risk markers to integrate information on non-genetic and genetic markers. The study sample was multi-ethnic/racial consisting of cases with childhood (≤18yr) ALL (n=161) and healthy controls (n=261) recruited contemporaneously between 2007 and 2012. Non-genetic data on birth weight, birth length, and gestational age were collected through administration of questionnaires, and genetic markers were genotyped using TaqMan allelic discrimination assays. We observed risk associations for having a birth weight of over 4,000 grams compared with being less than 4,000 grams (OR = 1.93, 95% CI = 1.16 to 3.19, $P = 0.01$), birth length (OR = 1.18 per inch, 95% CI = 1.01 to 1.38, $P = 0.04$), and with gestational age (OR = 1.10 per week, 95% CI = 1.00 to 1.21, $P = 0.04$). None of the associations interacted with gender or ethnicity. When we examined correlations between these birth characteristics and genetic markers, only the *HFE* tagging SNP rs9366637 showed an inverse correlation with gestational age with a gene-dosage effect

($P = 0.005$) and in interaction with a transferrin receptor (*TFRC*) rs3817672 genotype ($P_{interaction} = 0.05$). This correlation was translated into a strong association for this SNP with preterm birth (gestational age ≤ 37 weeks) with five of ten subjects born preterm being positive for the variant allele of rs9366637 (OR = 5.0, 95% CI = 1.19 to 20.9, $P = 0.03$). The strong interaction with *TFRC* suggests that rs9366637 may be involved in the modification of *HFE*-mediated placental iron transfer. Our study provides evidence for the involvement of prenatal events in the development of childhood ALL. The inverse correlation of *HFE* rs9366637 with gestational age has implications on the design of *HFE* association studies in birth weight and childhood conditions using full-term newborns as controls. If this statistical association is confirmed and its biological mechanisms include iron homeostasis changes, as suggested by our findings, the use of iron supplementation during pregnancy may need to be more targeted.

Introduction

Birth order, birth weight, birth defects, and Down syndrome were among the first risk factors shown to be associated with childhood leukemia risk [1-3]. Subsequent studies have established Down syndrome, increased birth weight, and several congenital disorders as consistent risk factors for childhood acute lymphoblastic leukemia (ALL) [4,5]. More limited data are available on the suggestive associations of being first-born, maternal age and miscarriages in maternal reproductive history with childhood ALL [5]. Of these, only the molecular mechanism of Down syndrome association has been established [6]. Birth weight shows a consistent positive association with childhood ALL that has been confirmed by meta-analyses [7-9]. Several biological mechanisms for the

birth weight association have been speculated [10-12], but none have been established firmly.

Although a number of studies have examined the associations of birth characteristics with childhood ALL [10,13-20] and adult cancer risk [21], few examined the correlations between genetic risk markers and birth characteristics, mainly birth weight [22]. Those studies that did examine genetic risk factors used polymorphisms previously identified in genome-wide association studies (GWAS) and mostly yielded negative results. The importance of the examination of birth characteristics in childhood ALL research stems from recent evidence suggesting a prenatal origin for childhood ALL development [23-26], and more broadly the concept of developmental origins of health and disease [27-29].

Our aim was to examine associations of birth characteristics with childhood ALL as well as their correlations with selected risk markers for childhood ALL to integrate information on non-genetic and genetic markers. We had data on birth weight, birth length, and gestational age, and we included a representative set of risk markers previously identified in GWAS, HLA region markers due to the association of HLA region polymorphisms with birth weight [30-33], iron regulatory gene *HFE* polymorphisms previously shown to correlate with birth weight [11] and additional iron-regulatory genes for their potential associations with birth characteristics. Since our sample was multi-ethnic/racial, we also explored heterogeneity among non-Hispanic Whites and Hispanics in any association observed.

Subjects & Methods

Subjects

Institutional Review Board (IRB) approval was received at the Baylor College of Medicine (BCM) and Florida International University prior to the start of the research project. Subjects were recruited at Texas Children's Hospital in Houston, TX. The study was originally designed as a leukemia association study. The case-control group consisted of 161 incident childhood ALL cases and 231 healthy age-matched controls contemporaneously recruited at Texas Children's Hospital general pediatrics clinics. The children were less than 18 years of age at diagnosis, and exclusion criteria for both groups included refusal to participate in the study and the diagnosis of any other disease or cancer. Subjects and their parents were approached to acquire informed consent and epidemiological data with a questionnaire and a biological sample, thus the gestational age data was not record-based. DNA samples were obtained from either saliva or peripheral blood samples at Texas Children's Hospital. Parents were asked to provide race (White, Black/African American, Asian, American Indian/Alaska Native, or Native Hawaiian/Other Pacific Islander) and ethnicity (Hispanic/Latino or non-Hispanic/Latino) of the child.

Genotyping

Genotyping was completed on the Bio-Rad CFX96 real-time PCR machine (Hercules, CA) using Life Technologies TaqMan® single nucleotide polymorphism (SNP) genotyping assays (LifeTech, Foster City, CA). The singleplex reactions were carried out in 96-well plates and used Bio-Rad's SsoFast™ Probes Supermix as the reaction buffer. PCR amplifications were performed using the manufacturer's suggestion

of total volume/well and PCR thermal cycling conditions. SNPs of interest are listed in Table I together with their features. SNPs included were: established ALL risk markers in previous GWAS, *HFE* SNPs known to correlate with body iron levels, another IRG SNP that shows a very high correlation with serum iron parameters in GWAS, and two additional IRG SNPs selected by us as haplotype tagging SNPs for the promoter regions of two IRGs. We also included three HLA region SNPs known to modify lymphoma risk, one being the marker for the *HLA-DRB4* lineage [34], which is also associated with risk for major leukemia types, as previously shown in candidate gene studies [35-39]. Two additional SNPs (rs285 and rs2891) were included as ancestry informative markers (AIMs) for statistical adjustment of ethnicity and race as has been used in Hispanic populations to account for the differences in genetic ancestry [40,41]. Bio-Rad CFX Manager software (version 3.0) was used for data acquisition and genotype assignment.

Statistical analysis

Statistical analyses were performed using Stata v.11 (StataCorp, College Station, TX, USA). Logistic regression was used to explore the associations of birth characteristics (birth weight, birth length, gestational age) with childhood ALL risk. Linear regression was used to assess correlation of genetic markers with continuous birth characteristic variables. The threshold for statistical significance was set at $P \leq 0.05$, and 95% confidence intervals (CI) of odds ratios (OR) were computed. Pearson's X^2, Student's t-test (for means) or median test (for medians) were used to compare characteristics between the case and control groups. Genotype counts were tested for Hardy–Weinberg equilibrium (HWE) in controls for each SNP. Correlations with genetic markers were assessed using the dominant model in which polymorphic allele carrying

genotypes are coded as 1 and the wildtype (common allele homozygote) genotype is coded as 0. When a SNP association was found, gene-dosage effect was explored by using the additive genetic model which uses all three genotypes coded as 0, 1 and 2 reflecting the number of polymorphic alleles present in the genotype. For most analyses, cases and controls were pooled to increase statistical power, but only after ruling out heterogeneity. As the sample was heterogeneous in its ethnic/racial composition, and genetic markers show variation in these groups, all results are adjusted for self-declared ethnic/racial groups (coded as non-Hispanic Whites, Hispanic Whites, Blacks and others). To rule out residual confounding, two SNPs were also used for adjustments, individually, to check whether the results were changed after adjustment for genetic ancestry.

Results

Characteristics of the case-control sample are shown in Table II. The case samples included 86 males (53%) and 75 females (47%). Out of the cases, 66 identified themselves as non-Hispanic White (NHW), 72 as Hispanic White (HW), 17 as Black, and 6 as "other." The "other" group included those identifying themselves as Asian, Native American, or other. The controls included 130 males (56%) and 101 females (44%), who had visited the pediatric clinic for a non-disease related reason. Forty-nine were classified as NHW, 98 as HW, and 78 as Black. The distribution of race and ethnic background was different between cases and controls, mainly due to the infrequency of childhood ALL in Blacks. Because of this difference, results were adjusted for racial/ethnic background or stratified analyses were performed when necessary. The differences in birth characteristics between cases and controls are detailed below.

Birth weight

There was no association between childhood ALL and birth weight when the birth weight variable was retained as a continuous variable. Since most studies reported a risk association with birth weight of 4,000 grams or higher, we categorized the variable as high birth weight (≥4,000g) - others (<4,000g), and re-examined the association. There was a statistically significant association with high birth weight in the overall group (OR = 1.93, 95% CI = 1.16 to 3.19, $P = 0.01$, adjusted for self-declared ethnicity/race). Although there was no statistical interaction with gender or ethnicity, among the subgroups, the association reached statistical significance in males and in Hispanics. The OR was as high as 3.18 (95% CI = 1.16 to 8.73, $P = 0.02$) in Hispanic males. The overall association with high birth weight did not change when adjusted for each AIM instead of self-declared ethnicity/race. The association with birth weight could not be attributed to gestational age heterogeneity, and adjustment for gestational age or restriction of the analysis to subjects born at or later than 38 weeks of gestational age did not result in a substantial change in the OR of the association. By categorizing the birth weight variable into three groups (<2500g; 2500-3999g; ≥4000g), the overall association and association in males remained statistically significant (data not shown).

Birth length

As a continuous variable, birth length showed an overall association with childhood ALL risk (OR = 1.18 per inch, 95% CI = 1.01 to 1.38, $P = 0.04$). Among the subgroups, this association was statistically significant in females and in non-Hispanic Whites, but there was no statistical interaction with gender or ethnicity. In non-Hispanic White females, the OR was 1.81 (95% CI = 1.03 to 3.18, $P = 0.04$). Restricting the

analysis to subjects born at gestational age 39 week or later (full-term) did not weaken the results. On the contrary, the OR in the whole group became 1.25, and 3.03 in non-Hispanic females, both remaining statistically significant.

Gestational age

In the overall group, there was a risk association with gestational age (OR = 1.10 per week, 95% CI = 1.00 to 1.21, $P = 0.04$, adjusted for ethnicity/race). This association did not interact with gender or ethnicity/race although it was stronger in females and in Hispanics with the OR reaching 1.43 (95% CI = 1.09 to 1.87, $P = 0.009$). Categorization of the gestational age variable into three groups (<38 wk; 38wk; >38wk) resulted in the loss of statistical significance for the association.

Correlation of birth characteristics with genetic markers

Having found associations between birth characteristics and childhood ALL risk, we sought correlations between the birth characteristics and SNPs included in this study. For these analyses, the birth weight variable was used as high birth weight (≥4,000g) versus others; and birth length and gestational age variables were used as continuous variables. Correlations were sought in the whole group as well as in the subgroup that showed the strongest leukemia association. We were able to show a correlation for only one genetic marker and gestational age. *HFE* tagging SNP rs9366637 showed a significant negative correlation with a gene-dosage effect (P=0.005) with gestational age. Males had a stronger negative correlation (P=0.001) with this variant. There was no association in females (P=0.98), and the interaction with gender reached statistical significance ($P_{interaction}$=0.02 for gender). No heterogeneity was found when analyzing the case and control groups separately. In fact, the regression coefficients were larger in the

control group. Replacing the self-declared ethnicity/race variable with either AIM for adjustment did not alter the results. The average gestational age showed a stepwise change with increasing number of rs9366637 variant alleles, with the mean gestational age decreasing from 38.6 weeks (homozygous wild-type) to 36.7 weeks (homozygous variant) (Figure I, Graph A).

Because of the known biological relationship between *HFE* and *TFRC*, rs3817672 was genotyped, and the interaction between rs3817672 and rs9366637 variant alleles was assessed. There was a significant statistical interaction between them ($P_{interaction}$=0.02). Overall, the negative association between rs9366637 and decreasing gestational age increased in statistical significance (P=0.001), and the male subgroup again showed the strongest negative correlation (P<0.001) in subjects with the particular *TFRC* genotype. Figure I depicts the mean gestational age by *HFE* rs9366637 and *TFRC* rs3817672 genotypes. Bar graph A in Figure I shows the gestational age means for just the *HFE* variant on its own, in the overall group. Graphs B and C depict the gestational age means for the *HFE* variant depending on the TFRC genotype (AA in Graph B and AG/GG combined in Graph C). Lastly, bar graphs D and E show the mean gestational age for males (D) and females (E) who had the *HFE* and *TRFC* genotypes that jointly showed the strongest correlation with gestational age ($P_{interaction}$ = 0.001).

We also examined the association of the *HFE* SNP rs9366637 with preterm birth defined as gestational age less than 37 weeks. Despite very small number of preterm births in the overall sample (n=65), there was an association between this SNP and preterm birth in non-Hispanic White males (P = 0.03), which was due to five of ten preterm males being positive for the variant allele of rs9366637.

Discussion

This study observed a birth weight association with childhood ALL risk [7-9], and novel associations with birth length and gestational age. While birth length association may be due to similar mechanisms as the birth weight association, the gestational age association is novel. The birth weight association is well known in childhood ALL, and fetal growth rate is probably even more crucial [9], but we were not able to assess fetal growth rate due to sample size limitations. Another novel finding was the inverse correlation between an *HFE* SNP and gestational age. This association showed gender-specificity, gene-dosage effect and a statistical interaction with another genotype that has biological basis. This is a novel finding worth pursuing in future research, as it may lead to development of a marker with some clinical utility.

The *HFE* SNP rs9366637 (IVS1) has not been studied as extensively as the other *HFE* variants rs1800562 (C282Y) or rs1799945 (H63D), and has not been directly implicated in iron regulation. It has shown a weak correlation with birth weight [11] and a strong association with adult height [42]. In a case-control study in a Han Chinese sample, the variant allele for rs9366637 was also found to be a significant risk marker for coronary heart disease (CHD) [43]. The haplotype tagged by this SNP is always devoid of the two variants C282Y and H63D that are associated with increased iron levels and presumably increased iron placental transfer. Since rs9366637 is a tagging SNP, we explored polymorphisms tagged by rs9366637 to see if any were functional and/or previously assessed (Table III). We were not able to attribute any functionality to any of the tagged SNPs for which rs9366637 is a proxy, nor was there any published disease associations with any of these SNPs.

The inverse correlation found between rs9366637 and gestational age, and its interaction with *TFRC* brings speculation on whether this SNP is also involved in iron homeostasis. There is no known effect with this SNP on *HFE* function, but being positive for this SNP would assure the lack of H63D or C282Y on the same chromosome negatively affecting placental iron transfer. Thus, if the association of rs9366637 is due to its modification of *HFE* function, it is more likely to be not facilitating iron transfer into the developing fetus.

The risk association of rs9366627 and CHD in a Han Chinese population where the minor allele frequency is much higher than in European populations may be consistent with its association with preterm birth [43]. Babies born before term are also low birth weight and at high risk for cardiovascular diseases later in life [27]. This issue has been recently revisited and nutrition of preterm babies with special formula and accelerated growth in very early life have been considered in the pathogenesis of later development of disease in preterm babies [44]. It is important to note that rs9366637 is also associated with body height in adults [42], and may predispose babies to accelerated growth. Although our findings are purely statistical correlations, future studies should explore potential implications of these findings.

We included the interaction analysis of the *TFRC* variant rs3817672 with *HFE* variant rs9366637 to see if a gene-gene interaction existed. The presence of the *TFRC* "A" allele was associated with increased risk for preterm birth in rs9366637 variant-positive individuals, specifically males. Interactions between *HFE* and *TFRC* allele "G" have also been noticed in various cancers, including multiple myeloma, breast, and colorectal cancer [45] as well as childhood ALL [11]. Since the *HFE* interaction with

TFRC allele G is believed to increase iron transport leading to increased birth weight [11] and cancer risk [46], it is important to note that the interaction of rs9366637 was with *TFRC* allele A in its association with shorter gestational age. This contrast suggests that the inverse correlation with gestational age, if causal, is likely to be due to insufficient iron transfer to the fetus. Iron deficiency is indeed a known risk factor for preterm birth [47].

There were important limitations with our study, including sample size. We were able to pool the cases and controls (n=392) to gain statistical power, since cases and controls did not show statistical heterogeneity for the correlation with the *HFE* SNP. Gestational age was provided by the mother of the child enrolled in the study through a questionnaire. Previous studies, however, have shown a high correlation between birth weight, gestational ages, and other birth characteristics provided by the mother of a patient and medical chart recordings [17, 48]. Non-differential measurement error may have occurred, causing towards-the-null bias, if not all gestational ages were correctly recalled. Missing data was another limitation in our study, with 88% of the total group having gestational age data.

Results from this study raise an important issue in study designs for future childhood studies. Most studies of a childhood disease consist of a control group comprised of full term babies and include as many cases as possible, regardless of their birth term. By doing this, the control group could be missing a whole group of children who may be carriers of certain variant genetic markers. Variants that may be associated with non-term births would be eliminated from the control group. The outcome of this mistake could lead to spurious results.

In conclusion, the *HFE* variant rs9366637 has shown a statistically significant negative correlation with gestational age. The SNP rs9366637, which is not known to be involved with iron homeostasis, showed a positive interaction with a *TRFC* genotype and yielded a stronger association with shorter gestational age, specifically in males. This correlation may reveal a connection between altered placental iron transfer and the risk of preterm birth. Since our study did not observe other correlations between a number of genetic markers and birth characteristics investigated, this may be another area that needs to be addressed in larger and more comprehensive studies, ideally using a GWAS design. If our results are validated in different populations, the findings will have implications on iron supplementation strategies during pregnancy.

REFERENCES

1. Krivit W, Good RA. Simultaneous occurrence of mongolism and leukemia; report of a nationwide survey. AMA J Dis Child 1957;94:289-93.

2. Macmahon B, Levy MA. Prenatal Origin Of Childhood Leukemia. Evidence From Twins. N Engl J Med 1964;270:1082-5.

3. Fasal E, Jackson EW, Klauber MR. Birth characteristics and leukemia in childhood. J Natl Cancer Inst 1971;47:501-9.

4. Ross JA, Swensen AR. Prenatal epidemiology of pediatric tumors. Curr Oncol Rep 2000;2:234-41.

5. Linet MS, Wacholder S, Zahm SH. Interpreting epidemiologic research: lessons from studies of childhood cancer. Pediatrics 2003;112:218-32.

6. Mullighan CG, Collins-Underwood JR, Phillips LA, Loudin MG, Liu W, Zhang J, Ma J, Coustan-Smith E, Harvey RC, Willman CL and others. Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. Nat Genet 2009;41:1243-6.

7. Hjalgrim LL, Westergaard T, Rostgaard K, Schmiegelow K, Melbye M, Hjalgrim H, Engels EA. Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. Am J Epidemiol 2003;158:724-35.

8. Caughey RW, Michels KB. Birth weight and childhood leukemia: A meta-analysis and review of the current evidence. Int J Cancer 2009;124:2658-70.

9. Milne E, Greenop KR, Metayer C, Schuz J, Petridou E, Pombo-de-Oliveira MS, Infante-Rivard C, Roman E, Dockerty JD, Spector LG and others. Fetal growth and childhood acute lymphoblastic leukemia: Findings from the childhood leukemia international consortium. Int J Cancer 2013.

10. Ou SX, Han D, Severson RK, Chen Z, Neglia JP, Reaman GH, Buckley JD, Robison LL. Birth characteristics, maternal reproductive history, hormone use during pregnancy, and risk of childhood acute lymphocytic leukemia by immunophenotype (United States). Cancer Causes Control 2002;13:15-25.

11. Dorak MT, Mackay RK, Relton CL, Worwood M, Parker L, Hall AG. Hereditary hemochromatosis gene (HFE) variants are associated with birth weight and childhood leukemia risk Pediatric Blood Cancer 2009;53:1242-8.

12. Callan AC, Milne E. Involvement of the IGF system in fetal growth and childhood cancer: an overview of potential mechanisms. Cancer Causes Control 2009;20:1783-98.

13. Kaye SA, Robison LL, Smithson WA, Gunderson P, King FL, Neglia JP. Maternal reproductive history and birth characteristics in childhood acute lymphoblastic leukemia. Cancer 1991;68:1351-1355.

14. Savitz DA, Ananth CV. Birth characteristics of childhood cancer cases, controls, and their siblings. Pediatr Hematol Oncol 1994;11:587-599.

15. Westergaard T, Andersen PK, Pedersen JB, Olsen JH, Frisch M, Sorensen HT, Wohlfahrt J, Melbye M. Birth characteristics, sibling patterns, and acute leukemia risk in childhood: a population-based cohort study. J Natl Cancer Inst 1997;89:939-947.

16. Reynolds P, Von Behren J, Elkin EP. Birth characteristics and leukemia in young children. Am J Epidemiol 2002;155:603-13.

17. Ma X, Metayer C, Does MB, Buffler PA. Maternal pregnancy loss, birth characteristics, and childhood leukemia (United States). Cancer Causes Control 2005;16:1075-83.

18. Podvin D, Kuehn CM, Mueller BA, Williams M. Maternal and birth characteristics in relation to childhood leukaemia. Paediatr Perinat Epidemiol 2006;20:312-22.

19. Johnson KJ, Soler JT, Puumala SE, Ross JA, Spector LG. Parental and infant characteristics and childhood leukemia in Minnesota. BMC Pediatr 2008;8:7.

20. Oksuzyan S, Crespi CM, Cockburn M, Mezei G, Kheifets L. Birth weight and other perinatal characteristics and childhood leukemia in California. Cancer Epidemiol 2012;36:e359-65.

21. McCormack VA, dos Santos Silva I, Koupil I, Leon DA, Lithell HO. Birth characteristics and adult cancer incidence: Swedish cohort of over 11,000 men and women. Int J Cancer 2005;115:611-7.

22. Linabery AM, Blommer CN, Spector LG, Davies SM, Robison LL, Ross JA. ARID5B and IKZF1 variants, selected demographic factors, and childhood acute lymphoblastic leukemia: a report from the Children's Oncology Group. Leuk Res 2013;37:936-42.

23. Greaves MF, Wiemels J. Origins of chromosome translocations in childhood leukaemia. Nat Rev Cancer 2003;3:639-49.

24. McHale CM, Wiemels JL, Zhang L, Ma X, Buffler PA, Feusner J, Matthay K, Dahl G, Smith MT. Prenatal origin of childhood acute myeloid leukemias harboring chromosomal rearrangements t(15;17) and inv(16). Blood 2003;101:4640-1.

25. Gruhn B, Taub JW, Ge Y, Beck JF, Zell R, Hafer R, Hermann FH, Debatin KM, Steinbach D. Prenatal origin of childhood acute lymphoblastic leukemia, association with birth weight and hyperdiploidy. Leukemia 2008;22:1692-7.

26. Wiemels JL, Cazzaniga G, Daniotti M, Eden OB, Addison GM, Masera G, Saha V, Biondi A, Greaves MF. Prenatal origin of acute lymphoblastic leukaemia in children. Lancet 1999;354:1499-503.

27. Barker DJ, Gluckman PD, Godfrey KM, Harding JE, Owens JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. Lancet 1993;341:938-41.

28. Gluckman PD, Hanson MA. Living with the past: evolution, development, and patterns of disease. Science 2004;305:1733-6.

29. Walker CL, Ho SM. Developmental reprogramming of cancer susceptibility. Nat Rev Cancer 2012;12:479-86.

30. Shin S, Yoon JH, Lee HR, Hwang SM, Roh EY. Association of HLA-A, -B and -DRB1 genotype with birthweight and CD34+ cell content: analysis of Korean newborns and their cord blood. Mol Hum Reprod 2010;16:338-46.

31. Larsson HE, Lynch K, Lernmark B, Nilsson A, Hansson G, Almgren P, Lernmark A, Ivarsson SA. Diabetes-associated HLA genotypes affect birthweight in the general population. Diabetologia 2005;48:1484-91.

32. Aroviita P, Partanen J, Sistonen P, Teramo K, Kekomaki R. High birth weight is associated with human leukocyte antigen (HLA) DRB1*13 in full-term infants. Eur J Immunogenet 2004;31:21-6.

33. Stene LC, Magnus P, Ronningen KS, Joner G. Diabetes-associated HLA-DQ genes and birth weight. Diabetes 2001;50:2879-82.

34. Kennedy AE, Singh SK, Dorak MT. Re: genome-wide association study of classical hodgkin lymphoma and epstein-barr virus status-defined subgroups. J Natl Cancer Inst 2012;104:884-5.

35. Dorak MT, Chalmers EA, Gaffney D, Wilson DW, Galbraith I, Henderson N, Worwood M, Mills KI, Burnett AK. Human major histocompatibility complex contains several leukemia susceptibility genes. Leuk Lymphoma 1994;12:211-22.

36. Dorak MT, Machulla HK, Hentschel M, Mills KI, Langner J, Burnett AK. Influence of the major histocompatibility complex on age at onset of chronic lymphoid leukaemia. Int J Cancer 1996;65:134-9.

37. Dorak MT, Lawson T, Machulla HK, Darke C, Mills KI, Burnett AK. Unravelling an HLA-DR association in childhood acute lymphoblastic leukemia. Blood 1999;94:694-700.

38. Dorak MT, Oguz FS, Yalman N, Diler AS, Kalayoglu S, Anak S, Sargin D, Carin M. A male-specific increase in the HLA-DRB4 (DR53) frequency in high-risk and relapsed childhood ALL. Leuk Res 2002;26:651-6.

39. Oguz FS, Kalayoglu S, Diler AS, Tozkir H, Sargin D, Carin M, Dorak MT. HLA system affects the age-at-onset in chronic myeloid leukemia. Am J Hematol 2003;73:256-62.

40. Ziv E, John EM, Choudhry S, Kho J, Lorizio W, Perez-Stable EJ, Burchard EG. Genetic ancestry and risk factors for breast cancer among Latinas in the San Francisco Bay Area. Cancer Epidemiol Biomarkers Prev 2006;15:1878-85.

41. Lee YL, Teitelbaum S, Wolff MS, Wetmur JG, Chen J. Comparing genetic ancestry and self-reported race/ethnicity in a multiethnic population in New York City. J Genet 2010;89:417-23.

42. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S and others. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 2010;467:832-8.

43. Shi Y, Zhou L, Huang LH, Lian YT, Zhang XM, Guo H, Wu TC, Cheng LX, He MA. Plasma ferritin levels, genetic variations in HFE gene, and coronary heart disease in Chinese: a case-control study. Atherosclerosis 2011;218:386-90.

44. Singhal A, Lucas A. Early origins of cardiovascular disease: is there a unifying hypothesis? Lancet 2004;363:1642-5.

45. Beckman LE, Van Landeghem GF, Sikstrom C, Wahlin A, Markevarn B, Hallmans G, Lenner P, Athlin L, Stenling R, Beckman L. Interaction between haemochromatosis and transferrin receptor genes in different neoplastic disorders. Carcinogenesis 1999;20:1231-1233.

46. Dorak MT. HFE H63D variant and leukemia susceptibility. Leuk Lymphoma 2006;47:2269-70.

47. Allen LH. Anemia and iron deficiency: effects on pregnancy outcome. Am J Clin Nutr 2000;71:1280S-4S.

48. Olson JE, Shu XO, Ross JA, Pendergrass T, Robison LL. Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group. Am J Epidemiol 1997;145:58-67.

Table I. Main features of SNPs analyzed

Gene	SNP	Chromosome nucleotide position*	Minor allele and frequency**	SNP Type
HFE	rs1800562	chr6: 26093141	(A) 0.053	Transition substitution, mis-sense mutation
HFE	rs1799945	chr6: 26091179	(G) 0.179	Mis-sense mutation, transversion substitution
HFE	rs9366637	chr6: 26089098	(T) 0.064	Intronic; transition substitution
TFRC	rs3817672	chr3: 195800811	(G) 0.383	Transition substitution, mis-sense mutation
SLC11A1	rs422982	chr12: 51406354	(A) 0.246	Transversion substitution, intragenic
TMPRSS6	rs733655	chr22: 37495051	(C) 0.221	Transition substitution, intragenic
TMPRSS6	rs855791	chr22: 37462936	(T) 0.412	Transition substitution, mis-sense mutation
ARID5B	rs7089424	chr10: 63752159	(G) 0.314	Intronic
ARID5B	rs10821936	chr10: 63723577	(C) 0.318	Intronic
ARID5B	rs10994982	chr10: 63710104	(A) 0.457	Intronic
IKZF1	rs4132601	chr7: 50470604	(G) 0.306	3'-UTR
CEBPE	rs2239633	chr14: 23589057	(A) 0.466	5'-upstream
HLA-DR region	rs2395185	chr6: 32433167	(T) 0.423	Intronic
HLA-DQB1 region	rs2647012	chr6: 32664458	(T) 0.381	Intergenic
HLA-DQA1 region	rs10484561	chr6: 32665420	(G) 0.084	Intergenic
LPL	rs285	chr8: 19815189	(T) 0.500	Transition substitution, intragenic
ITGAE	rs2891	chr17: 3705526	(G) 0.496	Transition substitution, intragenic

*Genome Reference Consortium Human Build 37 patch release 10 (GRCh37.p10) used for nucleotide position (http://www.ncbi.nlm.nih.gov/SNP/)

**Minor allele frequencies are from a reference Caucasian population (U.S. residents of northern and western European ancestry) genotyped in HapMap project

†† IVS1: intervening sequence 1 SNP

Table II. Characteristics of cases and controls

	Cases n=161	**Controls** n=231	***P* value**
Ethnic background Non-Hispanic White Hispanic White Black Other*	66 72 17 6	49 98 78 6	**<0.001**
Gender Male Female Ratio	86 75 1.15	130 101 1.29	0.58
Birth weight (*grams*) Mean (SD) Median (IQR)	 3349.3 (584) 3400 (760)	 3263.3 (684) 3311.5 (850)	 0.23 0.48
Gestational age *(weeks)* (SD)	38.8 (2.2)	38.2 (2.8)	**0.04**
Birth length *(cm)* (SD)	51.4 (3.4)	50.1 (5.05)	**0.024**

*Asian, Native American, or other

Table III. *HFE* rs9366637 and its tagging SNPs

Proxy	Distance	r^{2*}	Chr.	Nucleotide number	SNP Location
rs9366637	0	1.000	6	26089098	Intronic
rs2050947	19019	1.000	6	26070079	Intergenic
rs9295682	19604	1.000	6	26069494	Intronic
rs9379826	23058	0.892	6	26112156	Intergenic
rs9393682	32048	0.892	6	26165029	Intergenic

*An indication of the correlation with rs9366637; scale of 0-1 with 1 denoting maximum correlation

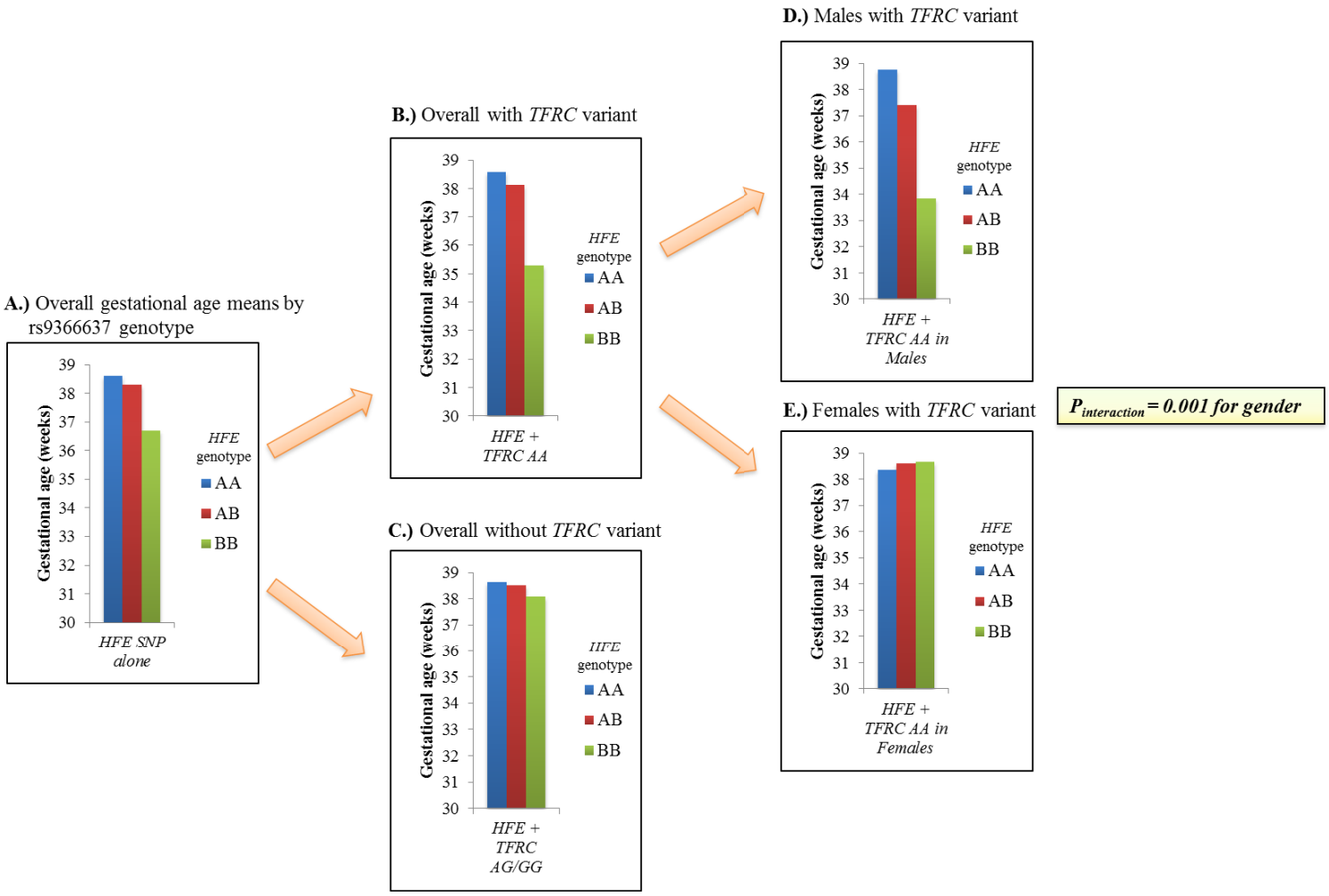


Figure I. Gene-gene and gene-gene-gender interactions in gestational age association of *HFE* rs9366637

CHAPTER VII
CONCLUSIONS AND FUTURE RESEARCH

Conclusions

This dissertation research was conducted to explore the correlations among genetic polymorphisms, birth characteristics, and the risk of childhood ALL. The replication of GWAS childhood leukemia markers [1,2] validated our case-control group and added credibility to the results of our present project. The genotyping of HLA region lymphoma risk markers not only confirmed a previous male-only strong risk association from rs2395185 [3], but also identified new risk markers for childhood ALL. All significant associations with lymphoma risk markers were in opposite risk direction compared to the lymphoma association studies. HLA region marker rs2647012 was found to be a significant risk only in Hispanics, and rs1048456, a risk marker in follicular lymphoma showed a statistically significant protective association.

Our hypothesis that iron-regulatory gene (IRG) variants known to elevate iron levels increase childhood ALL risk was driven by previous research, which showed a male-specific childhood ALL risk [4-6], and the biological implications showing iron excess associated with other cancers [7-11]. We examined the effect modification by gender and race/ethnicity in childhood ALL susceptibility, and also looked into gene-gene interactions to test our hypotheses.

The main objective of re-examining the association with *HFE* variants and childhood ALL susceptibility was accomplished. Our results not only replicated previous *HFE* risk associations with childhood ALL [4,5], but by extending similar findings to other IRG SNPs, strengthened the hypothesis that iron excess mediated by genetic

variants contributes to childhood ALL risk. It was previously known that *HFE* variants that correlate with body iron levels show associations with cancer risk, including childhood ALL [4,5], and also show a correlation with birth weight. Our results confirmed this leukemia risk, with replication of the male-only association [4,5].

Birth characteristics such as birth weight, height, and gestational age were examined to see if correlations with childhood ALL existed. Birth length showed an overall risk association with childhood ALL, with females and non-Hispanic Whites showing significance among the subgroups. There was no statistical interaction with gender or ethnicity, however. The correlation between birth weight and childhood ALL [12,13] was evident when comparing those having a birth weight of over 4,000 grams with those less than 4,000 grams. No individual SNP showed a correlation with birth length or birth weight. A novel significant risk association with gestational age overall did not interact with gender or ethnicity/race, although it was stronger in females and in Hispanics.

Correlations between *HFE* SNPs and gestational age were also examined. A significant negative correlation was found with the *HFE* variant rs9366637, more robustly among males. The inverse correlation of rs9366637 with gestational age has implications on the design of *HFE* association studies in birth weight and childhood conditions using newborns as controls.

Future Research

As with any epidemiological study, and especially important with genetic association studies, there is a need for replication and follow up studies to ensure the findings are valid [14-16]. Our conclusions are based on statistical correlations, therefore properly executed genetic and functional replication studies are necessary. Proper study design addressing the sample size issue for a rare disease like childhood ALL is crucial in follow up studies.

Our conclusions require independent replications done properly; meaning the methods of analysis should be consistent. A recent publication failed to replicate the *HFE* variant risk association with childhood ALL [17]. The researchers did not take into account gender effect and gene-gene interactions that are stated in the parent study, therefore the results are not a true replication failure of the original study.

The only way to confirm our hypotheses would be to examine pre-diagnostic iron levels, making cohort studies a necessity. Even as unrealistic and unfeasible as it may seem because childhood ALL is such a rare disease, ideally a cohort study would be best for this research. Any replication should be done in an international collaboration format as well, so that the sample size is increased, therefore statistical power for more robust and credible results would also increase. Both of these criteria are achievable because of the Childhood Leukemia International Consortium (CLIC), a giant international consortium that pools data from independent studies from 12 countries across the world [18].

Our research is significant in terms of both population health and in scientific advancements. Results may indicate a possible biological mechanism of the role iron

plays in the cancer association, leading to an impact on both child and maternal health. Educating the public on lifestyle modifications for preventive measures, such as controlled iron supplementation during pregnancy, unregulated access to iron-containing multivitamins, and avoidance from the unnecessary consumption of iron-rich food, could reduce the risk of leukemia in children.

REFERENCES

1. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, Kinsey SE, Lightfoot T, Roman E, Irving JA and others. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1006-10.

2. Trevino LR, Yang W, French D, Hunger SP, Carroll WL, Devidas M, Willman C, Neale G, Downing J, Raimondi SC and others. Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet 2009;41:1001-5.

3. Dorak MT, Lawson T, Machulla HK, Darke C, Mills KI, Burnett AK. Unravelling an HLA-DR association in childhood acute lymphoblastic leukemia. Blood 1999;94:694-700.

4. Dorak MT, Sproul AM, Gibson BE, Burnett AK, Worwood M. The C282Y mutation of HFE is another male-specific risk factor for childhood ALL. Blood 1999;94:3957-3958.

5. Dorak MT, Mackay RK, Relton CL, Worwood M, Parker L, Hall AG. Hereditary hemochromatosis gene (HFE) variants are associated with birth weight and childhood leukemia risk Pediatric Blood Cancer 2009;53:1242-8.

6. Dorak MT, Burnett AK, Worwood M. Hemochromatosis gene in leukemia and lymphoma. Leuk Lymphoma 2002;43:467-477.

7. Toyokuni S. Role of iron in carcinogenesis: cancer as a ferrotoxic disease. Cancer Sci 2009;100:9-16.

8. Zacharski LR, Ornstein DL, Woloshin S, Schwartz LM. Association of age, sex, and race with body iron stores in adults: analysis of NHANES III data. Am Heart J 2000;140:98-104.

9. Knekt P, Reunanen A, Takkunen H, Aromaa A, Heliovaara M, Hakulinen T. Body iron stores and risk of cancer. Int J Cancer 1994;56:379-82.

10. Stevens RG, Jones DY, Micozzi MS, Taylor PR. Body iron stores and the risk of cancer. N Engl J Med 1988;319:1047-1052.

11. Stevens RG, Graubard BI, Micozzi MS, Neriishi K, Blumberg BS. Moderate elevation of body iron level and increased risk of cancer occurrence and death. Int J Cancer 1994;56:364-369.

12. Dorak MT, Hammal DM, Pearce MS, McNally RJ, Parker L. Examination of gender effect in birth weight and miscarriage associations with childhood cancer. Cancer Causes Control 2007;18:219-228.

13. Samuelsen SO, Bakketeig LS, Tretli S, Johannesen TB, Magnus P. Birth weight and childhood cancer. Epidemiology 2009;20:484-7.

14. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. Nat Genet 2001;29:306-9.

15. Moonesinghe R, Khoury MJ, Janssens AC. Most published research findings are false-but a little replication goes a long way. PLoS Med 2007;4:e28.

16. Vieland VJ. The replication requirement. Nat Genet 2001;29:244-5.

17. Rodriguez-Lopez R, Donoso M, Fernandez-Cavada M, Gonzalez LM, Margallo A, Corral C, Gallego M, Garcia de Caceres MT, Herrera T, Gonzalez C and others. Diagnostic utility of HFE variants in Spanish patients: association with HLA alleles and role in susceptibility to acute lymphoblastic leukemia. Gene 2013;514:31-5.

18. Metayer C, Milne E, Clavel J, Infante-Rivard C, Petridou E, Taylor M, Schuz J, Spector LG, Dockerty JD, Magnani C and others. The Childhood Leukemia International Consortium. Cancer Epidemiol 2013;37:336-47.

VITA

AMY KENNEDY

Born, Melrose, Massachusetts

2008-present	Ph.D. Candidate in Public Health (Epidemiology) Florida International University Miami, Florida
2011-Present	Graduate Assistant Robert Stempel College of Public Health and Social Work Dean's Office Florida International University Miami, Florida
2008-2011	Graduate Research Assistant Robert Stempel College of Public Health and Social Work Department of Environmental and Occupational Health Florida International University Miami, Florida
2007-2008	M.P.H., Environmental and Occupational Health Florida International University Miami, Florida
2002-2006	B.S., Biology University of Miami Coral Gables, Florida

PUBLICATIONS AND PRESENTATIONS (ABRIDGED LIST)

Kennedy AE, Singh SK, Villalba K, Dorak MT. Analysis of HLA region polymorphisms associated with cancer. Oral presentation at the 39th Annual Meeting of the American Society for Histocompatibility and Immunogenetics Society. Chicago, IL, Nov 17-21, 2013. Hum Immunol 2013;74(Suppl 1):S35.

Singh SK, Ben Taleb Z, Kennedy AE, Villalba K, Dorak MT. Further exploration of HLA region associations with lung cancer risk. Poster presentation at the 39th Annual Meeting of the American Society for Histocompatibility and Immunogenetics Society. Chicago, IL, Nov 17-21, 2013. Hum Immunol 2013;74(Suppl 1):S106.

Villalba, K, Kennedy AE, Singh, S, Dorak MT. Investigation of Molecular Mechanism of Diverse Disease Associations with HLA-DRB4 Lineage. 15th Annual Florida

International University Biomedical & Comparative Immunology Symposium. March 8, 2013, Miami, FL.

Singh, S, Kennedy AE, Dorak MT. DDX3X Gene Polymorphisms and Genetic Susceptibility to Childhood Leukemia. 15th Annual Florida International University Biomedical & Comparative Immunology Symposium. March 8, 2013, Miami, FL.

Kennedy AE, Singh SK, Scheurer ME, Okcu MF, Dorak MT. Association of HLA-linked lymphoma risk markers with childhood lymphoblastic leukemia. Accepted for poster presentation at the 38th Annual Meeting of the American Society for Histocompatibility and Immunogenetics. San Juan, PR, October 8-12, 2012.

Zeledon A, Kennedy AE, Dorak MT. Genetic Risk Associations in Childhood Leukemia Differ by Ethnicity. 3rd Annual Florida International University Summer Research Internship Mini-Symposium. August 10, 2012, Miami, FL.

Kennedy, AE., Singh, S., Dorak M.T. (2012) Re: Genome-Wide Association Study of Classical Hodgkin Lymphoma and Epstein–Barr Virus Status–Defined Subgroups. J Natl Cancer Inst,104: 884-885.

Kennedy, A., Singh, S., Samikkannu, M., Dorak, MT. Correlations of Complex Disease-associated HLA Region SNPs with HLA Alleles. Accepted for Poster Presentation, 37th Annual American Society for Histocompatibility and Immunogenetics Meeting, New Orleans, LA, October 17-21, 2011.

Alvarez K, Singh SK, Kennedy AE, Dorak MT. Lymphoid malignancies and HLA complex associations. 2nd Annual Florida International University Summer Research Internship Biomedical Mini-Symposium. August 12, 2011, Miami, FL.

Torres M, Singh SK, Kennedy AE, Dorak MT. Detection of C4A and C4B deletion by quantitative real-time PCR. Florida International University & University of Puerto Rico at Arecibo Summer Research Program Biomedical Mini-Symposium. August 3, 2011, Miami, FL.

Felty, Q., Yoo, C., & Kennedy, A. (2010) Gene expression profile of endothelial cells exposed to estrogenic environmental compounds: implications to pulmonary vascular lesions. Life Sciences, 86(25-26), 919-927.

Hotz, G., Kennedy, A., Lutfi, K., & Cohn, S. M. (2009) Preventing pediatric pedestrian injuries. The Journal of Trauma, 66(5), 1492-1499.

Hotz, G., de Marcilla, A. G., Lutfi, K., Kennedy, A., Castellon, P., & Duncan, R. (2009) The WalkSafe Program: developing and evaluating the educational component. The Journal of Trauma, 66(3 Suppl), S3-9.