Title Page:

Title of the paper: Birth Characteristics and Childhood Leukemia Risk: Correlations with Genetic Markers

Name and affiliation of each author (in order):
Dr. Amy E. Kennedy - Florida International University (Kennedy AE)
Dr. Kala Y. Kamdar - Baylor College of Medicine (Kamdar KY)
Dr. Philip J. Lupo - Baylor College of Medicine (Lupo PJ)
Dr. M. Fatih Okcu - Baylor College of Medicine (Okcu MF)
Dr. Michael E. Scheurer - Baylor College of Medicine (Scheurer ME)
Dr. Mehmet Tevfik Dorak - Florida International University (Dorak MT)

Corresponding author information:

Dr. Mehmet Tevfik Dorak*, MD, PhD
Florida International University
Robert Stempel College of Public Health & Social Work, Department of Epidemiology
11200 SW 8th Street
AHC II 595-1
Miami, FL 33199
mdorak@fiu.edu
Phone: 305-348-0118
Fax: 305-348-4901

*Present address:
Professor of Health Sciences
School of Health Sciences, HCA EW 208
Liverpool Hope University
Liverpool L16 9JD
United Kingdom
dorakm@hope.ac.uk
Phone: +44 151 291 3236

Other author's information:

Dr. Amy E. Kennedy**, PhD, MPH
Florida International University
Robert Stempel College of Public Health & Social Work, Department of Epidemiology
11200 SW 8th Street
AHC II 595-1
Miami, FL 33199
akenn001@fiu.edu
Phone: 305-348-0118
Fax: 305-348-4901

**Present address:
National Institutes of Health
National Cancer Institute
Epidemiology and Genomics Research Program
Division of Cancer Control and Population Sciences
9609 Medical Center Drive, Rm. 4E114, MSC 9763
Rockville, MD 20850
Phone: (240) 276-6723
Fax: (240) 276-7921

Dr. Kala Y. Kamdar, MD, MS
Baylor College of Medicine
Department of Pediatrics, Section of Hematology-Oncology
One Baylor Plaza
MS-BCM305
Houston, TX 77030
kykamdar@txch.org
Phone: 832-824-4163
Fax: 832-825-4732

Dr. Philip J. Lupo, PhD, MPH
Baylor College of Medicine
Department of Pediatrics, Section of Hematology-Oncology
One Baylor Plaza
MS-BCM305
Houston, TX 77030
Philip.Lupo@bcm.edu
Phone: 713-798-2960
Fax: 713-798-3658

Dr. M. Fatih Okcu, MD, MPH
Baylor College of Medicine
Department of Pediatrics, Section of Hematology-Oncology
One Baylor Plaza
MS-BCM320
Houston, TX 77030
mfokcu@txch.org
Phone: 832-822-4242
Fax: 832-825-4202

Dr. Michael E. Scheurer, PhD, MPH
Baylor College of Medicine
Department of Pediatrics, Section of Hematology-Oncology
One Baylor Plaza
MS-BCM305
Houston, TX 77030

scheurer@bcm.edu
Phone: 713-798-5547
Fax: 713-798-8711

Key Words: Genetic associations, childhood acute lymphoblastic leukemia, genetic predisposition to disease, birth characteristics

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. Examinations of correlations between birth characteristics and leukemia risk markers have been limited to birth weight-related genetic polymorphisms. We integrated information on non-genetic and genetic markers by evaluating the relationship of birth characteristics, genetic markers for childhood acute lymphoblastic leukemia (ALL) susceptibility, and ALL risk together. The multi-ethnic study consisted of cases with childhood ALL (n=161) and healthy controls (n=261). Birth characteristic data were collected through questionnaires, and genotyping was achieved by TaqMan® SNP Genotyping Assays. We observed risk associations for birth weight over 4,000 grams (OR=1.93, 95% CI=1.16 to 3.19), birth length (OR=1.18 per inch, 95% CI=1.01 to 1.38), and with gestational age (OR=1.10 per week, 95% CI=1.00 to 1.21). Only the HFE tag SNP rs9366637 showed an inverse correlation with a birth characteristic, gestational age, with a gene-dosage effect (P=0.005) and in interaction with a transferrin receptor rs3817672 genotype ($P_{interaction}$=0.05). This correlation translated into a strong association for rs9366637 with preterm birth (OR=5.0, 95% CI=1.19 to 20.9). Our study provides evidence for the involvement of prenatal events in the development of childhood ALL. The inverse correlation of 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.

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 explore 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.

Materials and 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 Hospitalin 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- and gender matched controls contemporaneously recruited at Texas Children's Hospital general pediatric 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 cancer or other disease. 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 information on 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 tag 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 race/ethnicity as has been used in previous studies of 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 ($P \leq 0.001$). Correlations with genetic markers were assessed using the gene-dosage effect (additive model), which uses all three genotypes coded as 0 (wild-type), 1 (heterozygote) and 2 (homozygote) reflecting the number of polymorphic allele 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). Since there might still be residual confounding after adjustment for self-declared ethnicity/race, we re-examined associations after adjustment for two AIMs.

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 others including mixed race/ethnicity. The controls included 130 males (56%) and 101 females (44%), who had visited the pediatric clinicfor 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 (≥4,000 g) or others (<4,000 g), 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 and gender). 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. When we categorized the birth weight variable into three groups, as customarily done, (<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 1.26 (95% CI = 1.0 to 1.57), and 3.03 (95% CI = 1.19 to 7.74) 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, $P = 0.009$).

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 (≥4,000 g) 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 in which birth characteristics showed the strongest leukemia association. We were able to show a correlation for only one genetic marker and gestational age. *HFE* tag 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). No heterogeneity was found when analyzing the case and control groups separately, but the regression coefficients were larger in the control group (-0.777 (SE: 0.325) in controls vs. -0.497 (SE: 0.345) in cases). Thus, the results were not due to the inclusion of cases in overall analysis. 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* [42, 43], the interaction between *TFRC* rs3817672 and *HFE* rs9366637 was assessed. This analysis revealed a statistically significant interaction ($P_{interaction}$=0.02) in their association with gestational age. 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 (OR = 5.0, 95% CI = 1.19 to 20.9, *P* = 0.03), which was due to five of ten preterm males being positive for the variant allele of rs9366637.

Since rs9366637 is a tag 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, despite one having a

significant FS score (rs9393682), nor was there any published disease associations with any of these SNPs.

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 corresponding to a biological interaction. 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]. 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)[44]. 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, also with increased iron placental transfer.

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, which may negatively impact 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 coronary heart disease 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 [44]. 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 [45]. 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 [42] as well as childhood ALL [11]. Since the *HFE* interaction with *TFRC* allele G is believed to increase placental 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, resulting in an iron-deficient phenotype. 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. Some studies of a childhood disease consist of a control group comprised of full term newborns and include as many cases as possible, regardless of their birth term. When this is the case, the control group could be missing a whole group of children who may be carriers of certain 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 summary, the present study observed associations of birth weight, birth length and gestational age with childhood ALL risk, and the *HFE* variant rs9366637 showed a statistically significant negative correlation with gestational age. The SNP rs9366637, which is not known to be directly 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.* Sep 1957;94(3):289-293.

2. Macmahon B, Levy MA. Prenatal Origin Of Childhood Leukemia. Evidence From Twins. *N Engl J Med.* May 21 1964;270:1082-1085.

3. Fasal E, Jackson EW, Klauber MR. Birth characteristics and leukemia in childhood. *J Natl Cancer Inst.* Sep 1971;47(3):501-509.

4. Ross JA, Swensen AR. Prenatal epidemiology of pediatric tumors. *Curr Oncol Rep.* May 2000;2(3):234-241.

5. Linet MS, Wacholder S, Zahm SH. Interpreting epidemiologic research: lessons from studies of childhood cancer. *Pediatrics.* Jul 2003;112(1 Pt 2):218-232.

6. Mullighan CG, Collins-Underwood JR, Phillips LA, et al. Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. *Nat Genet.* Nov 2009;41(11):1243-1246.

7. Hjalgrim LL, Westergaard T, Rostgaard K, et al. Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. *Am J Epidemiol.* Oct 15 2003;158(8):724-735.

8. Caughey RW, Michels KB. Birth weight and childhood leukemia: A meta-analysis and review of the current evidence. *Int J Cancer.* Dec 18 2009;124(11):2658-2670.

9. Milne E, Greenop KR, Metayer C, et al. Fetal growth and childhood acute lymphoblastic leukemia: Findings from the childhood leukemia international consortium. *Int J Cancer.* Jun 10 2013.

10. Ou SX, Han D, Severson RK, et al. Birth characteristics, maternal reproductive history, hormone use during pregnancy, and risk of childhood acute lymphocytic leukemia by immunophenotype (United States). *Cancer Causes Control.* Feb 2002;13(1):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(7):1242-1248.

12. Callan AC, Milne E. Involvement of the IGF system in fetal growth and childhood cancer: an overview of potential mechanisms. *Cancer Causes Control.* Dec 2009;20(10):1783-1798.

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 1991;68(6):1351-1355.

14. Savitz DA, Ananth CV. Birth characteristics of childhood cancer cases, controls, and their siblings. *Pediatr Hematol Oncol.* 1994 1994;11(6):587-599.

15. Westergaard T, Andersen PK, Pedersen JB, et al. Birth characteristics, sibling patterns, and acute leukemia risk in childhood: a population-based cohort study. *J Natl Cancer Inst.* 1997 1997;89(13):939-947.

16. Reynolds P, Von Behren J, Elkin EP. Birth characteristics and leukemia in young children. *Am J Epidemiol.* Apr 1 2002;155(7):603-613.

17. Ma X, Metayer C, Does MB, Buffler PA. Maternal pregnancy loss, birth characteristics, and childhood leukemia (United States). *Cancer Causes Control.* Nov 2005;16(9):1075-1083.

18. Podvin D, Kuehn CM, Mueller BA, Williams M. Maternal and birth characteristics in relation to childhood leukaemia. *Paediatr Perinat Epidemiol.* Jul 2006;20(4):312-322.

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.* Dec 2012;36(6):e359-365.

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.* Jul 1 2005;115(4):611-617.

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.* Aug 2013;37(8):936-942.

23. Greaves MF, Wiemels J. Origins of chromosome translocations in childhood leukaemia. *Nat Rev Cancer.* Sep 2003;3(9):639-649.

24. McHale CM, Wiemels JL, Zhang L, et al. Prenatal origin of childhood acute myeloid leukemias harboring chromosomal rearrangements t(15;17) and inv(16). *Blood.* Jun 1 2003;101(11):4640-4641.

25. Gruhn B, Taub JW, Ge Y, et al. Prenatal origin of childhood acute lymphoblastic leukemia, association with birth weight and hyperdiploidy. *Leukemia.* Sep 2008;22(9):1692-1697.

26. Wiemels JL, Cazzaniga G, Daniotti M, et al. Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet.* Oct 30 1999;354(9189):1499-1503.

27. Barker DJ, Gluckman PD, Godfrey KM, Harding JE, Owens JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. *Lancet.* Apr 10 1993;341(8850):938-941.

28. Gluckman PD, Hanson MA. Living with the past: evolution, development, and patterns of disease. *Science.* Sep 17 2004;305(5691):1733-1736.

29. Walker CL, Ho SM. Developmental reprogramming of cancer susceptibility. *Nat Rev Cancer.* Jul 2012;12(7):479-486.

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.* May 2010;16(5):338-346.

31. Larsson HE, Lynch K, Lernmark B, et al. Diabetes-associated HLA genotypes affect birthweight in the general population. *Diabetologia.* Aug 2005;48(8):1484-1491.

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.* Feb 2004;31(1):21-26.

33. Stene LC, Magnus P, Ronningen KS, Joner G. Diabetes-associated HLA-DQ genes and birth weight. *Diabetes.* Dec 2001;50(12):2879-2882.

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.* Jun 6 2012;104(11):884-885.

35. Dorak MT, Chalmers EA, Gaffney D, et al. Human major histocompatibility complex contains several leukemia susceptibility genes. *Leuk Lymphoma.* Jan 1994;12(3-4):211-222.

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.* Jan 17 1996;65(2):134-139.

37. Dorak MT, Lawson T, Machulla HK, Darke C, Mills KI, Burnett AK. Unravelling an HLA-DR association in childhood acute lymphoblastic leukemia. *Blood.* Jul 15 1999;94(2):694-700.

38. Dorak MT, Oguz FS, Yalman N, et al. A male-specific increase in the HLA-DRB4 (DR53) frequency in high-risk and relapsed childhood ALL. *Leuk Res.* Jul 2002;26(7):651-656.

39. Oguz FS, Kalayoglu S, Diler AS, et al. HLA system affects the age-at-onset in chronic myeloid leukemia. *Am J Hematol.* Aug 2003;73(4):256-262.

40. Ziv E, John EM, Choudhry S, et al. Genetic ancestry and risk factors for breast cancer among Latinas in the San Francisco Bay Area. *Cancer Epidemiol Biomarkers Prev.* Oct 2006;15(10):1878-1885.

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.* Dec 2010;89(4):417-423.

42. Beckman LE, Van Landeghem GF, Sikstrom C, et al. Interaction between haemochromatosis and transferrin receptor genes in different neoplastic disorders. *Carcinogenesis.* 7/1999 1999;20(7):1231-1233.

43. Benyamin B, McRae AF, Zhu G, et al. Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. *Am J Hum Genet.* Jan 2009;84(1):60-65.

44. Shi Y, Zhou L, Huang LH, et al. Plasma ferritin levels, genetic variations in HFE gene, and coronary heart disease in Chinese: a case-control study. *Atherosclerosis.* Oct 2011;218(2):386-390.

45. Singhal A, Lucas A. Early origins of cardiovascular disease: is there a unifying hypothesis? *Lancet.* May 15 2004;363(9421):1642-1645.

46. Dorak MT. HFE H63D variant and leukemia susceptibility. *Leuk Lymphoma.* 2006;47(11):2269-2270.

47. Allen LH. Anemia and iron deficiency: effects on pregnancy outcome. *Am J Clin Nutr.* May 2000;71(5 Suppl):1280S-1284S.

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.* Jan 1 1997;145(1):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	Missense (C282Y)
HFE	rs1799945	chr6: 26091179	(G) 0.179	Missense (H63D)
HFE	rs9366637	chr6: 26089098	(T) 0.064	Intronic (intron 1)
TFRC	rs3817672	chr3: 195800811	(G) 0.383	Missense (G142S)
SLC11A1	rs422982	chr12: 51406354	(A) 0.246	Intronic (intron 1)
TMPRSS6	rs733655	chr22: 37495051	(C) 0.221	Intronic (intron 2)
TMPRSS6	rs855791	chr22: 37462936	(T) 0.412	Missense (V736A)
ARID5B	rs7089424	chr10: 63752159	(G) 0.314	Intronic (intron 3)
ARID5B	rs10821936	chr10: 63723577	(C) 0.318	Intronic (intron 3)
ARID5B	rs10994982	chr10: 63710104	(A) 0.457	Intronic (intron 3)
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	Intronic (intron 6)
ITGAE	rs2891	chr17: 3705526	(G) 0.496	5'-upstream

*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 II. Characteristics of cases and controls

	Cases n=161	**Controls** n=231	***P* value**
Ethnic background			
Non-Hispanic White	66	49	**<0.001**
Hispanic White	72	98	
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)* (SEM)	51.4 (3.4)	50.1 (5.05)	**0.024**

*Asian, Native American, or other

Table III. *HFE* rs9366637 and SNPs tagged

Proxy	Distance	r^{2*}	Chr.	Nucleotide number	SNP Location	FS score**	RegulomeDB***
rs9366637	0	1.00	6	26089098	Intronic	0.208	4
rs2050947	19019	1.00	6	26070079	Intergenic	NA	6
rs9295682	19604	1.00	6	26069494	Intronic	NA	NA
rs9379826	23058	0.89	6	26112156	Intergenic	NA	6
rs9393682	32048	0.89	6	26165029	Intergenic	0.815	4

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

** FS (functional significance) scores are provided by F-SNP and range from 0 (no tools predict a deleterious effect) to 1 (all tools predict a deleterious effect),
with FS scores ≥0.5 considered significant (http://compbio.cs.queensu.ca/F-SNP)

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

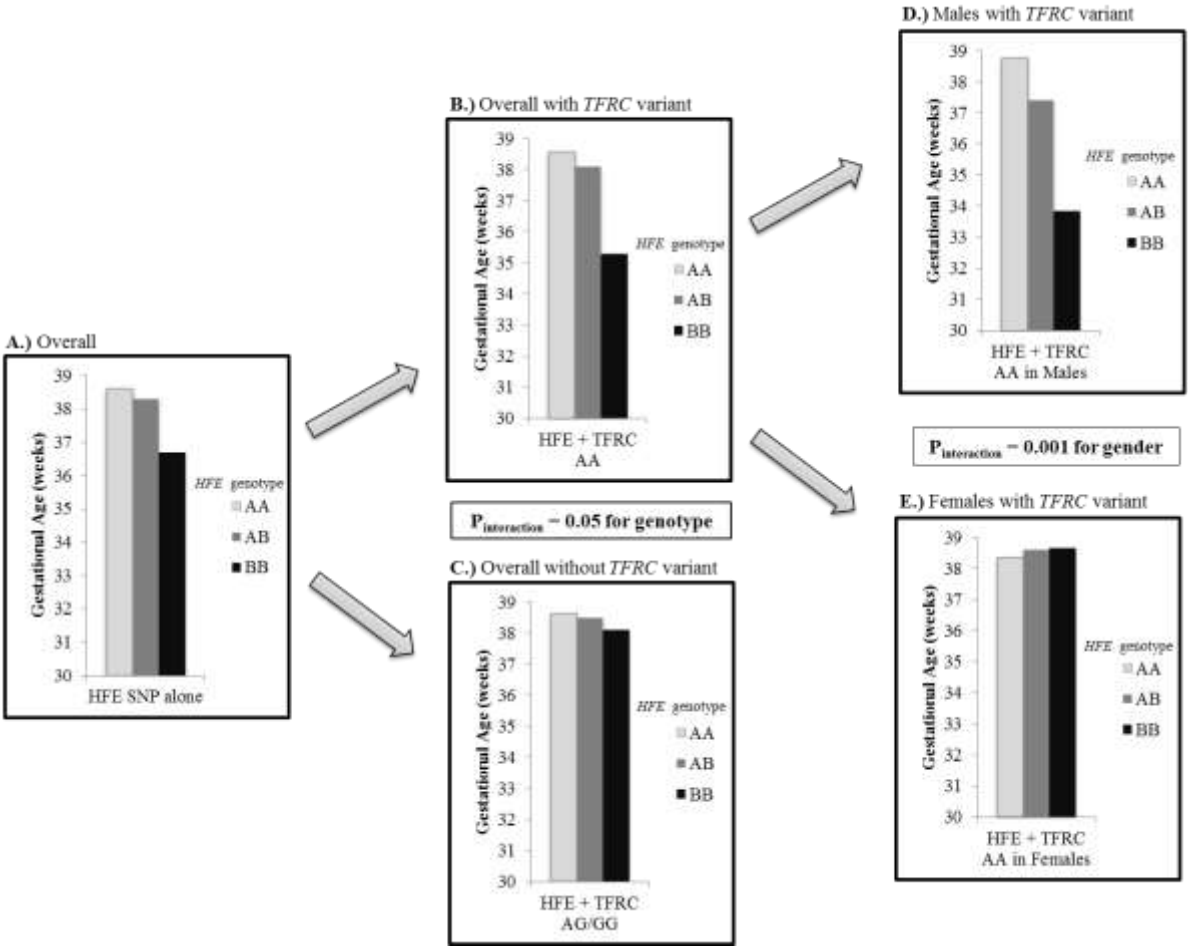

Figure I. Gene-gene and gene-gene-gender interactions in gestational age association of HFE rs9366637

A.) Gestational age means in the overall group, depending on the *HFE* variant

B.) Gestational age means for the *HFE* variant and AA *TFRC* genotype

C.) Gestational age means for the *HFE* variant and the AG/GG *TFRC* genotypes

D.) Gestational age means for males who had the *HFE* and *TRFC* genotypes

E.) Gestational age means for females who had the *HFE* and *TRFC* genotype