NIH Public Access
Author Manuscript
Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2012 January 12.

Published in final edited form as:
Cancer Epidemiol Biomarkers Prev. 2010 September ; 19(9): 2152–2163. doi: 10.1158/1055-9965.EPI-10-0389.

Genetic Polymorphisms in Adaptive Immunity Genes and Childhood Acute Lymphoblastic Leukemia

Jeffrey S. Chang[1,2,3], **Joseph L. Wiemels**[1], **Anand P. Chokkalingam**[2], **Catherine Metayer**[2], **Lisa F. Barcellos**[2], **Helen M. Hansen**[4], **Melinda C. Aldrich**[5], **Neela Guha**[2,6], **Kevin Y. Urayama**[2,6], **Ghislaine Scélo**[6], **Janet Green**[2], **Suzanne L. May**[2], **Vincent A. Kiley**[7], **John K. Wiencke**[4], and **Patricia A. Buffler**[2]

[1] Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California, USA

[2] School of Public Health, University of California, Berkeley, Berkeley, California, USA

[3] National Institute of Cancer Research, National Health Research Institutes, Tainan, Taiwan

[4] Department of Neurological Surgery, University of California, San Francisco, San Francisco, California, USA

[5] Department of Medicine, University of California, San Francisco, San Francisco, California, USA

[6] International Agency for Research on Cancer, Lyon, France

[7] Pediatric Oncology, Kaiser Permanente Roseville, Roseville, California, USA

Abstract

Background—Childhood acute lymphoblastic leukemia (ALL) has been hypothesized to have an infection and immune-related etiology. The lack of immune priming in early childhood may result in abnormal immune responses to infections later in life and increase ALL risk.

Methods—The current analyses examined the association between childhood ALL and 208 single nucleotide polymorphisms (SNPs) of 29 adaptive immune function genes among 377 ALL cases and 448 healthy controls. Single SNPs were analyzed with a log-additive approach using logistic regression models adjusted for sex, age, Hispanic ethnicity, and race. Sliding window haplotype analyses were performed with haplotypes consisting of 2 to 6 SNPs.

Results—Of the 208 SNPs, only rs583911 of *IL12A*, which encodes a critical modulator of T-cell development, remained significant after accounting for multiple testing (odds ratio for each copy of the variant G allele = 1.52, 95% confidence interval: 1.25–1.85, $p = 2.9 \times 10^{-5}$). This increased risk was stronger among first-born children of all ethnicities and among non-Hispanic children with less daycare attendance, consistent with the hypothesis regarding the role of early immune modulation in the development of childhood ALL. Haplotype analyses identified additional regions of *CD28*, *FCGR2, GATA3*, *IL2RA*, *STAT4*, and *STAT6* associated with childhood ALL.

Conclusion—Polymorphisms of genes on the adaptive immunity pathway are associated with childhood ALL risk.

Corresponding author: Jeffrey S. Chang, 1F No 367, Sheng-Li Road, Tainan 70456, Taiwan R.O.C. jeffreychang@nhri.org.tw, Tel: 886-6-208-3422 ext 65260, Fax: 886-6-208-3427.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Impact—Results of this study support an immune-related etiology of childhood ALL. Further confirmation is required to detect functional variants in the significant genomic regions identified in this study, in particular for *IL12A*.

INTRODUCTION

The etiology of childhood acute lymphoblastic leukemia (ALL) is likely to be affected by history of infections and immune development as suggested by Greaves' "delayed infection" hypothesis (1) and Kinlen's "population mixing" hypothesis (2). Although these two hypotheses differ on the existence of specific leukemia-causing agents, both suggest that the lack of immune priming in a child's early development may result in abnormal immune responses to microbial challenges later in life thereby increasing the risk of childhood ALL. Both hypotheses are similar to the "hygiene hypothesis" proposed by Strachan to explain the rising prevalence of allergies in the western population (3). Many studies have been conducted to test these two hypotheses using proxy measures of infection (4–11). An increased risk of childhood ALL has been observed with proxy measures of early childhood infections such as low birth order (4, 5) and low daycare attendance (6–9, 12), although negative findings have also been reported (10, 11). Studies examining associations between reports of specific early childhood infections and childhood ALL have yielded variable findings ranging from inverse association (8, 13, 14), no association (7, 15–17), to positive association (18). These inconsistencies may be due to difficulty in obtaining information regarding asymptomatic infections, recall bias, variable sample sizes, and differences in questionnaire design and data collection. Inverse associations have been reported by the majority of studies of allergies and childhood leukemia (13, 19–25), indicating that immune function may play an important role.

Though the majority of studies support an infection and immune-related etiology of childhood ALL, little is known regarding the underlying role of genetics. The development of immune function is a complex process that involves the interplay between many cell types including Th1, Th2, T regulatory, and Th17 cells (26). Variations in the genes affecting the development and the function of these cell types may affect a child's immune responses and thus his/her risk of childhood ALL. The current analyses examine the association between childhood ALL and 208 polymorphisms of 29 adaptive immune function genes involved in the development and the function of Th1, Th2, T regulatory, and Th17 cells. In addition, analysis was performed to assess the interaction between immune function genes and two proxy measures of early childhood infections (early daycare attendance and birth order) on the risk of childhood ALL.

MATERIALS AND METHODS

Study subjects

The study subjects were recruited by the Northern California Childhood Leukemia Study (NCCLS), a case-control study that began in 1995. Major medical centers in 17 counties in the San Francisco Bay Area were included in the study from 1995 to 1999, and 18 additional counties in the California Central Valley were added in 1999. The eligibility criteria for case and control subjects were: 1) being a resident of the study area; 2) being less than 15 years old at the time of the leukemia diagnosis (reference date for controls); 3) having at least one English or Spanish speaking parent; and 4) having no previous diagnosis of cancer. Cases were identified from four (1995–1999) and later nine hospitals (1999–2008) in the study area. Comparison of case ascertainment in the 35-county study area to the California Cancer Registry data (1997–2003) showed that the NCCLS ascertained 93–96% of children diagnosed with leukemia in the participating hospitals. When considering both participating and non-participating hospitals within the 35 study counties, cases ascertained through the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NCCLS protocol represented 76% of all the diagnosed cases, making the study approximately population-based. Controls were randomly selected from birth certificates through the California Office of Vital Records and individually matched to cases on birth date, gender, maternal race, and Hispanic ethnicity and were shown to be representative of the source population of the cases (27). The current genetic study began with 928 NCCLS subjects recruited during 1995–2002 (464 cases, 464 controls), with 825 subjects retained (377 childhood ALL cases, 448 healthy controls) for this analysis. Reasons for exclusion were insufficient DNA for genotyping from either buccal cytobrush swabs or archived newborn dried bloodspot specimens (n = 21), ineligibility after genotyping (respondent was not a biological parent) (n = 1) or Illumina SNP call rates < 95% (n = 22). AML cases (n = 59) were excluded from the analysis, since ALL was the primary childhood leukemia subtype that has been hypothesized to have an infection and immune-related etiology (1). The study subjects included both Hispanic and non-Hispanic subjects. A child was considered Hispanic if either parent self-reported Hispanic ethnicity (156 cases, 179 controls). The non-Hispanic group (221 cases, 269 controls) consisted of 73.5% Whites, 11.8% Asians, 5.5% Blacks, 0.4% Native Americans, and 8.8% others.

The study was approved by the Committee for Protection of Human Subjects of the University of California, Berkeley and by the Institutional Review Boards of all collaborating institutions.

Biospecimen collection and DNA processing

Buccal cytobrushes were collected at the time of interview by trained interviewers as the primary DNA source for case and control children. DNA from cytobrush samples was extracted by heating (98–100°C) in the presence of NaOH, followed by neutralization with Tris-HCl buffer, and whole-genome amplification (WGA) using GenomePlex reagents (Sigma Aldrich, St. Luis, MO). Archived newborn blood (ANB) specimens, which are collected at birth on a paper card for each child born in California and archived at −20°C by the California Department of Public Health, were used as a secondary DNA source when buccal cell DNA was insufficient for genotyping (26.6% of subjects). For each child, the NCCLS receives one spot of ANB containing approximately 60 microliters of blood. A small piece of the bloodspot was excised for DNA extraction using the QIAamp DNA mini-extraction kit. Isolated ANB DNA was whole-genome amplified using REPLI-g reagents (Qiagen, Hilden, Germany). WGA products were tested for minimum acceptable amplifiable human DNA content using an ALUq real time PCR method published elsewhere (28). WGA DNA from both buccal cells and ANB specimens and genomic DNA from peripheral blood produced genotypes that were highly concordant when analyzed using multiplexed GoldenGate genotyping (Illumina, San Diego, CA) (28, 29).

SNP selection

We focused on 29 adaptive immunity genes (Th1: *IL12A*, *IL12B*, *IL12RB1*, *IL12RB2*, *PHF11*, *STAT4*; Th2/Allergy: *ADAM33*, *GATA3*, *IL4*, *IL4R*, *MS4A2*, *STAT6*; T regulatory cells: *CD28*, *CD80*, *CTLA4*, *IL2*, *IL2RA*, *IL6*, *IL10*, *STAT5A*, *STAT5B*, *TGFB1*; Th17: *STAT3*; B- cell: *CD40*, *FCGR2A*, *MME*; Others: *NFKB1*, *NFKBIA*, *NFKBIB*) based on assessment of their importance in published literature. SNPs genotyped by the International HapMap Project (30) were selected, including 10 kilobase (kb) regions up- and down-stream to include variants in potential regulatory elements. Haploview (31) was used to construct haplotype blocks according to the block definition by Gabriel *et al.* (32) to select haplotype tagging SNPs. Additional SNPs specific to the Hispanic population were supplemented from the SNP500Cancer database (33). Additionally, SNPs reported previously in the literature for these genes were also included. In total, 244 SNPs in the 29 adaptive immunity genes were selected for genotyping.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Genotyping

The genotyping of 244 SNPs of the 29 adaptive immunity genes was performed on whole genome amplified DNA using a custom Illumina GoldenGate assay panel. SNPs were excluded from statistical analyses if they had a call rate of less than 90% (29 SNPs), had minor allele frequencies less than 5% in both Hispanics and non-Hispanics (6 SNPs), or failed Hardy-Weinberg equilibrium ($p<0.01$) in both Hispanic and non-Hispanic controls (1 SNP), leaving a total of 208 SNPs for analysis. Ninety-five ancestry informative markers (AIMs) were included to account for potential population stratification in our study population but only 80 AIMs passed quality control.

Additional quality control of genotyping was performed by comparing duplicate samples: 1) 59 samples were run in duplicate after processing with the same whole genome amplification method and genotyped on the same plate; these showed a 99.1% concordance of genotype; 2) DNA specimens extracted from both buccal cell and archived newborn dried bloodspots were genotyped for 9 subjects; these showed a 98.9% concordance of genotype; and 3) the Mendelian errors (the child does not have the expected genotype compared to the parents) were estimated in 10 trios from the HapMap Centre d'Etude du Polymorphisme Humain (CEPH) and only 28 pedigree errors in 25 markers were found (overall Mendelian error rate = 0.20%). For subjects that had samples run in duplicate for quality control purpose, one sample from each pair of duplicates was chosen on the basis of SNP call rates. If one of the replicates had a higher call rate than the other (higher number of SNPs successfully genotyped) then its data were retained in the analysis, while the data of the other replicate were discarded.

Proxy Measures of Early Childhood Exposure to Infections

Information on daycare attendance and birth order was collected through an in-person interview with the biological parent (usually the mother) of the child. Detailed information on collection and calculation of daycare attendance was presented previously (7). Briefly, data on daycare and preschool attendance before the date of case diagnosis (reference date for the controls) or before age 6, whichever came first, were collected. For each daycare facility, information was ascertained regarding the duration of attendance (in months), mean hours per week of attendance, and mean number of other children in attendance. These data were used to calculate child-hours at each daycare, which is a composite measure of exposure to other children defined as follows: (number of months attending a daycare) × (mean hours per week at this daycare) × (number of other children at this daycare) × (4.35 weeks per month) (7). The measure for total child-hours of exposure for each child was then determined by summing the child-hours at each daycare attended. To examine specific time windows of exposure to infection early in life, child-hours of daycare attendance prior to 1 year and also 6 months of age were considered in the analysis.

Statistical Analysis

Single SNP analyses were conducted assuming a log-additive model (0, 1, or 2 copies of the variant allele) or a dominant model (genotypes with at least 1 copy of the variant allele vs. homozygous wild type), using unconditional logistic regression. The odds ratio (OR) associated with each SNP was adjusted for age, sex, Hispanic ethnicity (in the analysis with all subjects), and race. Sensitivity analyses with data from the NCCLS (data not shown) showed that the results from conditional logistic regression and those from unconditional logistic regression adjusted for the matching variables were very similar. Tests of heterogeneity were performed to assess the difference in the SNP-disease association by Hispanic ethnicity. If evidence of heterogeneity was present ($p<0.10$) for a SNP, separate analyses by Hispanic ethnicity were performed for that SNP, otherwise all subjects were combined to assess SNP-disease association.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Multifactor dimensionality reduction (MDR) analysis (34) was performed to assess SNP-SNP interactions between the 6 SNPs that were significantly ($p<0.05$) associated with ALL among all subjects. For the current MDR analysis, we allowed for combinations of one to four SNPs. The 10-fold cross-validation was repeated 10 times using 10 different random seeds to reduce the probability of spurious findings due to chance division of the data. P-values were calculated by permutation testing with 1000 permutations. The best combination of SNPs was determined based on cross-validation consistency and testing accuracy.

To assess the influence of adaptive immunity genes on the association between early life exposure to infections and childhood ALL, gene-environment interaction analysis was performed with the one SNP (*IL12A* rs583911) that passed the multiple testing adjustment. The risk associated with the minor variant G allele was evaluated by stratifying on the two proxy measures of early exposure to common infections (total child-hours of daycare attendance and birth order). There were 365 ALL cases and 429 controls retained in the gene-environment interaction analyses after excluding subjects under the age of 1 in order to allow for sufficient exposure to infectious factors prior to the development of childhood leukemia. Product terms for interaction between infectious exposure variables and rs583911 were included in the statistical model and statistical significance was assessed by the log-likelihood ratio test comparing the full model containing the product terms to the sub-model without the product terms. Interaction between rs583911 and daycare attendance was evaluated separately for Hispanics and non-Hispanics because the association between daycare attendance and childhood ALL differed significantly by Hispanic ethnicity ($p<0.05$). Hispanics and non-Hispanics were combined for assessing interaction between rs583911 and birth order since the association between ALL and birth order did not differ by Hispanic ethnicity (p values ranged from 0.35–0.89).

Haplotype analyses were performed using the haplo.stats R package (http://cran.r-project.org/) for each gene separately to capture the information potentially missed by the single SNP analyses by increasing the statistical power to tag causal variants and by accounting for *cis*-interactions between two or more SNPs (35). All subjects were combined for haplotype analysis if none of the SNPs in a gene had evidence of heterogeneity by Hispanic ethnicity; otherwise, separate analyses by Hispanic ethnicity were performed. Haplotype analyses were performed using the sliding window approach (haplotype windows of 2 to 6 SNPs) confined to the region of a single gene. Global p-values were calculated for each haplotype window to evaluate whether the distribution of haplotypes was significantly different between cases and controls. Graphical representations of the sliding window results were constructed using GrASP (36). The most significant p-value (minimal p-value) for each SNP across all haplotype windows was used to determine the most significant genomic region for each gene. Analyses to examine specific haplotypes in significant genomic regions were performed with haplotype trend regression to calculate the OR associated with each copy of a specific haplotype using the most frequent haplotype as the referent group.

Eighty AIMS were used to estimate genetic ancestry (percent of European, Amerindian, and African ancestry) using the methods described by Chakraborty *et al.* (37) and Hanis *et al.* (38). Genetic ancestry was included in statistical models to assess the impact of potential population stratification on the association between adaptive immunity SNPs and childhood ALL.

RESULTS

Cases and controls were comparable in the distribution of sex, age, race, genetic ancestry, and birth order (Table 1). For Hispanics, cases had more total child-hours of day care

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

attendance before the 6 months and 1 year of age. In contrast, cases had less total child-hours of daycare attendance before the 6 months and 1 year of age for non-Hispanics. Among controls, non-Hispanics had more total child-hours of daycare attendance before the 6 months and 1 year of age compared to Hispanics.

The quantile-quantile (Q-Q) plot (Figure 1) compares the distribution of the observed vs. the expected $-\log_{10}$ p-values (log-additive model not adjusted for genetic ancestry using data of combined race/ethnicity) of the 208 SNPs. All of the observed $-\log_{10}$ p-values except the most significant one follow closely the 45 degree angle line expected under the null hypothesis of no association, indicating minimal evidence of an inflated test statistic associated with population stratification or some other systematic bias. In addition, sensitivity analyses including genetic ancestry in the statistical models did not change the ORs by more than 10%, suggesting minimal impact of population substructure, and therefore genetic ancestry was not included in the final analytical models.

Single SNP analyses

Among the 235 single SNP tests performed (Supplementary Table 1) for the 208 SNPs (27 SNPs were analyzed stratified on Hispanic ethnicity due to presence of heterogeneity), 19 had a p-value of less than 0.05 using the log-additive model (Table 2). These 19 SNPs occur in 10 genes involved in the development and function of different immune cells: Th1 (*IL12A*, *STAT4*, *PHF11, IL12B*), Th2/allergy (*GATA3*, *STAT6*), T regulatory cells (*IL10*, *CTLA4*, *IL2RA*), and B-cell (*MME*). However, only rs583911 of *IL12A* (OR for each copy of the minor variant G allele = 1.52, 95% confidence interval (CI): 1.25–1.85, $p = 2.9 \times 10^{-5}$) remained significant after accounting for multiple testing using the Bonferroni correction (significance threshold=$0.05/235=2.1 \times 10^{-4}$). No p-values from the dominant model reached statistical significance after correcting for multiple testing (Supplementary Table 1). Supplementary Table 2 compares the results of the 19 statistically significant SNPs ($p < 0.05$) with and without adjustment for genetic ancestry, and the results were very similar. Single SNP analyses were further performed among three ALL subgroups with a sufficient subject number (N>50): 1) ALL with TEL-AML translocation (N=62); 2) ALL with high hyperdiploidy (N=110); and c-ALL (N= 189). Although statistical power decreased with the smaller sample size in the subgroup analyses, rs583911 of *IL12A* remained statistically significant ($p<0.05$) in all subgroups.

Multifactor dimensionality reduction (MDR)

MDR analysis showed that rs583911 is the best predictor of case-control status with the highest testing accuracy (P = 0.05; Supplementary Table 3). No additional SNPs were able to improve the prediction accuracy, suggesting no evidence of SNP-SNP interactions.

Interaction between the variant G allele of *IL12A* rs583911 and proxies for infectious exposures

The increased ALL risk associated with each copy of the variant G allele of *IL12A* rs583911 appeared stronger among first-born children (OR = 2.14; 95% CI:1.52–3.01) compared to children with older siblings (OR = 1.30; 95% CI:1.00–1.69) (Table 3).

Among non-Hispanics, those children who had less than two thousand child-hours in daycare before the age of six months had an increased risk of ALL associated with each copy of the G allele of *IL12A* rs583911 (OR = 1.68; 95% CI: 1.27–2.22). This was not apparent among children with two thousand child-hours or more of daycare attendance. Similar results were seen for daycare attendance before the age of one using five thousand child-hours as the cutoff.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Among Hispanic children, the risk of ALL associated with each copy of the G allele of *IL12A* rs583911 did not differ significantly by child-hours of daycare attendance.

Haplotype analyses with all subjects combined

Because no evidence of heterogeneity was observed among all SNPs of each gene by Hispanic status, haplotype analyses using a sliding window approach were performed with all subjects combined for 16 of the 29 genes (Supplementary Figure 1). Among these 16 genes, 4 (*CD28*, *CTLA4, FCGR2A*, and *IL12A*) showed significant haplotype associations (global p-value <0.05). However, for *CTLA4* and *IL12A*, haplotype analyses did not contribute additional information because the strength of association either did not improve or weakened with increasing haplotype window size compared to the results of single SNP analyses.

Haplotype analyses of *CD28* and *FCGR2A* identified significant regions in the genes not observed with the single SNP analyses. For *CD28*, a region tagged by rs1879877, rs3181096, rs1181389, and rs3769683 showed the strongest significance (global p = 0.02), and the haplotype CGAA was associated with an increased risk of ALL (OR for each copy of the haplotype = 1.55, 95% CI: 1.03–2.34, p = 0.04; Table 4). For *FCGR2A*, a region tagged by rs10800309 and rs4656308 showed the strongest significance (global p = 0.02), and the haplotype AA was associated with an increased risk of ALL (OR for each copy of the haplotype = 1.46, 95% CI: 1.13–1.90, p = 0.004; Table 4).

Haplotype analyses stratified by Hispanic ethnicity

Due to the presence of significant heterogeneity for at least one SNP in the gene by Hispanic ethnicity, haplotype analyses using a sliding window approach were performed separately for Hispanic and Non-Hispanic subjects for 13 of the 29 genes (Supplementary Figure 2). Eleven of the 13 genes had significant haplotype regions among either the Hispanics or non-Hispanics; however only 4 (*GATA3*, *IL2RA*, *STAT4*, and *STAT6*) of the 11 genes had common regions shared by Hispanics and non-Hispanics in their association with childhood ALL.

For *GATA3*, a region tagged by rs4143094, rs3781093, and rs3802604 was significantly associated with childhood ALL (global p=0.03 for non-Hispanics and 0.04 for Hispanics) in both Hispanics and non-Hispanics (Table 5). The CAG haplotype was associated with a reduced childhood ALL risk compared to the most common haplotype among both non-Hispanics and Hispanics (although only borderline significant among non-Hispanics). However, the CGG haplotype was positively associated with childhood ALL only among non-Hispanics.

For *IL2RA*, a region tagged by rs6602398, rs942201, rs791587, and rs706778 was significantly associated with childhood ALL (global p=0.02 for non-Hispanics and 0.004 for Hispanics) in both Hispanics and non-Hispanics (Table 5). The ACAG haplotype was associated with an increased risk of childhood ALL compared to the most common haplotype among both non-Hispanics and Hispanics, but Hispanics had two additional common (>5%) haplotypes (CAGA and CCGA) that were associated with an increased risk of childhood ALL compared to the most common haplotype.

For *STAT4*, even though a region tagged by rs17769459, rs4853546, and rs1031509 was found to be associated with childhood ALL among both Hispanics and non-Hispanics, these associations were mainly driven by the rare (<5%) haplotypes (Table 5).

For *STAT6*, a region tagged by rs4559, rs1059513, and rs324015 was associated with childhood ALL among both non-Hispanics and Hispanics. Particularly, the GAG haplotype

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

was associated with an increased risk of childhood ALL, although more significantly for Hispanics than non-Hispanics (Table 5). In addition, the AGG haplotype was associated with an increased risk of childhood ALL among non-Hispanics but not among Hispanics.

DISCUSSION

Of the 208 SNPs analyzed in the study, 19 SNPs belonging to 10 genes (*IL12A*, *STAT4*, *IL12B*, *GATA3*, *PHF11*, *STAT6*, *IL10*, *CTLA4*, *IL2RA*, and *MME*) showed a significant ($p<0.05$) association with childhood ALL. However, only rs583911 of *IL12A* remained significant after accounting for multiple testing (OR for each copy of variant allele = 1.52, 95% confidence interval: 1.25–1.85, $p = 2.9 \times 10^{-5}$). The increased risk associated with *IL12A* rs583911 G allele was stronger among first-born children of all ethnicities and among children with less child-hours at daycare in non-Hispanics. In addition to single SNP analyses, haplotype analyses further identified regions of *CD28*, *FCGR2A, GATA3*, *IL2RA*, *STAT4*, and *STAT6* that may be associated with childhood ALL risk.

Previous studies have shown that newborns have Th2-skewed immune profiles (39–41). Furthermore, during the normal course of immune development, a shift from Th2-dominant to Th1-dominant immune profiles occurs with increasing age (42). It is thought that the major driving force for this immune shift is the production of IL12 by innate immune cells (e.g. dendritic cells) after exposure to microbial challenges (43). The IL12 protein is a heterodimer that consists of two subunits, IL12A (p35) and IL12B (p40) (44). In the current study, the most significant result was observed with rs583911 of *IL12A*, a finding that is strikingly robust against multiple testing. Functional impact of the *IL12A* SNPs has yet to be characterized. A recent study by Pistiner *et al.* showed that rs2243123 of *IL12A*, a SNP in intron 2 that is 739 base pairs away from rs583911, is associated with immune sensitization to cockroach antigen (45); this lends further support that the region around rs583911 may be important in either the function or the expression of IL12A and may be a promising candidate region to perform fine mapping and functional studies to determine causal variants. The current study also observed three SNPs in *IL12B* (rs3181224, rs1368439, and rs11574790) that are associated with childhood ALL risk only among non-Hispanics, although not statistically significant after multiple testing adjustment. Rs3181224 is located near the 3′ end of *IL12B* and rs1368439 is located in the 3′ untranslated region (UTR), and may tag polymorphisms in regulatory elements located in this region. Another SNP in the 3′ UTR of *IL12B*, rs3212227, has been associated with psoriasis, a chronic T-cell-mediated inflammatory disease of the skin (46); however, we observed a null association between this SNP and childhood ALL.

The current infectious hypotheses of childhood leukemia propose that childhood leukemia may result from abnormal responses to microbial challenges due to lack of immune priming during early childhood (1, 2). Consistent with this hypothesis are the findings showing a reduced risk of childhood ALL associated with higher birth order (4, 5) and early daycare attendance (6–9, 12). While our data showed an overall increase in the risk of ALL associated with the variant G allele of *IL12A* rs583911, this association was stronger among first-born children who likely have less exposures to infections early in life compared to those with older siblings. These observations further support that childhood ALL may result from inadequate immune modulation early in life due to decreased microbial exposures.

Similar gene-environment interaction was observed between *IL12A* rs583911 and early daycare attendance on the risk of childhood ALL among non-Hispanics. An increased ALL risk associated with the variant G allele of *IL12A* rs583911 was observed among those children who had fewer child-hours of daycare attendance. However, a similar modifying effect of daycare attendance was not observed in Hispanics, consistent with published data

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

from the NCCLS reporting ethnic differences in the association between daycare attendance and childhood ALL (7). In this paper, Ma *et al.* suggested that daycare attendance may not be the primary source of early life exposure to infections for Hispanic children since it was observed that fewer Hispanic children started daycare before the age of one and Hispanic children tended to live with more other children in the same household compared to non-Hispanic white children (7). Among the subjects included in the current analysis, fewer Hispanic control children than non-Hispanic control children attended daycare before the age of one (12.8% vs 31.6%), and among those who did attend daycare, the mean child-hours of daycare was significantly lower among Hispanics compared to non-Hispanics (3.9 vs. 7.3 thousand child-hours, p=0.002).

As stated previously, IL12 is a key cytokine for the normal switching, triggered by exposure to microbial challenges, of a Th2-dominant to a Th1-dominant immune profile in early childhood (42, 43). Though the functional of impact of rs583911 or the actual causal SNP linked to rs583911 has yet to be elucidated, it seems that the increased ALL risk associated with genetic factors, possibly due to a decreased IL12 function, could be mitigated by increasing infectious exposures early in life.

In addition to the association between rs583911 of *IL12A* and childhood ALL there were 18 other SNPs that had p<0.05, though none of them remained significant after correcting for multiple testing. Haplotype analyses further identified regions of *CD28*, *FCGR2A, GATA3*, *IL2RA*, *STAT4*, and *STAT6* that may be associated with risk of childhood ALL. Though these findings may not have reached statistical significance after accounting for multiple testing, it is nevertheless worthwhile to take note of their potential importance in the context of the adaptive immunity pathway. Therefore, the discussion for these results is presented in the Supplementary Discussion for completeness.

The results of these analyses need to be interpreted in the context of several limitations. We attempted to be inclusive in gene selection, but we were not able to include all of the known adaptive immune function genes. In addition, the statistical power to detect an OR < 1.5 may be insufficient in our study especially when the minor allele frequency is 0.10 or less. For example, based on 377 ALL cases and 448 controls, type I error (alpha) = 0.05, and a minor allele frequency of 0.10, the statistical power to detect an OR of 1.4 is 0.59. Nevertheless, this makes the finding associated with rs583911 of *IL12A* all the more impressive. Besides an insufficient statistical power for detecting a main effect OR < 1.5 when the minor allele frequency is low, the current study may have a even more limited statistical power for haplotype and interaction analyses (both gene-environment and gene-gene via MDR analysis).

The interpretation of results from complex genetic analyses when multiple comparisons are inherent is always an issue of concern and this constraint exists with the current analysis. However, our most significant result associated with rs583911 of *IL12A* remained highly significant even after a stringent multiple testing correction using the Bonferroni method.

Results of this genetic study complement those of studies that used non-genetic measures and surrogates of infection such as daycare attendance, birth order, vaccination history, allergies, and parental reports of infections for investigating the infection and immune-related etiology of childhood leukemia. In addition, unlike interview-based studies, genetic studies do not suffer from recall errors or recall biases.

Ancestry informative markers (AIMs) were included in the genotyping to assess potential population stratification especially among the Hispanic population. It was reassuring that the results of the analysis with AIMs indicated that population stratification is minimal in our study most likely because of the matched design. Another strength of the current analysis is

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the haplotype approach to provide a more comprehensive assessment of variations within candidate genes than has been possible using single SNP analyses. These analyses indicate regions within several genes that would be suitable candidates for replication in other study populations, as well as further studies to identify potential causal variants.

In summary, the current analysis identified associations between polymorphisms of several adaptive immunity genes and childhood ALL. Although only one single-SNP association (rs583911 of *IL12A*) was statistically robust, these findings provide important support for a role of the adaptive immunity pathway in childhood ALL through immune modulation in early childhood development. Additional support was provided by the results showing that the risk associated with *IL12A* rs583911 G allele was stronger among children with fewer opportunities for infectious exposures (among first-born children of all ethnicities and non-Hispanic children with less daycare attendance). Further confirmation is needed to determine functional variants in the significant genomic regions identified by this study in particular for *IL12A* which encodes a critical modulator of T-cell development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by research awards from the National Institute of Environmental Health Sciences (PS42ES04705 and R01ES09137) and the Children with Leukaemia Foundation. Dr. Jeffrey S. Chang also received support from the National Cancer Institute (R25 CA112355). This research could not have been conducted without the strong support from our clinical collaborators and participating hospitals which include: University of California Davis Medical Center (Dr. Jonathan Ducore), University of California San Francisco (Dr. Mignon Loh and Dr. Katherine Matthay), Children's Hospital of Central California (Dr. Vonda Crouse), Lucile Packard Children's Hospital (Dr. Gary Dahl), Children's Hospital Oakland (Dr. James Feusner), Kaiser Permanente Roseville (Dr. Kent Jolly and Dr. Vincent Kiley), Kaiser Permanente Santa Clara (Dr. Alan Wong and Dr. Carolyn Russo), Kaiser Permanente San Francisco (Dr. Kenneth Leung) and Kaiser Permanente Oakland (Dr. Daniel Kronish and Dr. Stacy Month). Finally, we acknowledge the entire Northern California Childhood Leukemia Study staff and the UCB Survey Research Center for their effort and dedication.

References

1. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. Nat Rev Cancer. 2006; 6:193–203. [PubMed: 16467884]
2. Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. Br J Cancer. 1995; 71:1–5. [PubMed: 7819022]
3. Strachan DP. Hay fever, hygiene, and household size. Bmj. 1989; 299:1259–60. [PubMed: 2513902]
4. Dockerty JD, Draper G, Vincent T, Rowan SD, Bunch KJ. Case-control study of parental age, parity and socioeconomic level in relation to childhood cancers. Int J Epidemiol. 2001; 30:1428–37. [PubMed: 11821358]
5. 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; 89:939–47. [PubMed: 9214673]
6. Ma X, Buffler PA, Selvin S, et al. Daycare attendance and risk of childhood acute lymphoblastic leukaemia. Br J Cancer. 2002; 86:1419–24. [PubMed: 11986774]
7. Ma X, Buffler PA, Wiemels JL, et al. Ethnic difference in daycare attendance, early infections, and risk of childhood acute lymphoblastic leukemia. Cancer Epidemiol Biomarkers Prev. 2005; 14:1928–34. [PubMed: 16103439]
8. Perrillat F, Clavel J, Auclerc MF, et al. Day-care, early common infections and childhood acute leukaemia: a multicentre French case-control study. Br J Cancer. 2002; 86:1064–9. [PubMed: 11953850]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

9. Gilham C, Peto J, Simpson J, et al. Day care in infancy and risk of childhood acute lymphoblastic leukaemia: findings from UK case-control study. Bmj. 2005; 330:1294. [PubMed: 15849205]
10. Rosenbaum PF, Buck GM, Brecher ML. Early child-care and preschool experiences and the risk of childhood acute lymphoblastic leukemia. Am J Epidemiol. 2000; 152:1136–44. [PubMed: 11130619]
11. Neglia JP, Linet MS, Shu XO, et al. Patterns of infection and day care utilization and risk of childhood acute lymphoblastic leukaemia. Br J Cancer. 2000; 82:234–40. [PubMed: 10638995]
12. Urayama KY, Buffler PA, Gallagher ER, Ayoob JM, Ma X. A meta-analysis of the association between day-care attendance and childhood acute lymphoblastic leukaemia. Int J Epidemiol. 2010; 39:718–32. [PubMed: 20110276]
13. Jourdan-Da Silva N, Perel Y, Mechinaud F, et al. Infectious diseases in the first year of life, perinatal characteristics and childhood acute leukaemia. Br J Cancer. 2004; 90:139–45. [PubMed: 14710221]
14. McKinney PA, Juszczak E, Findlay E, Smith K, Thomson CS. Pre- and perinatal risk factors for childhood leukaemia and other malignancies: a Scottish case control study. Br J Cancer. 1999; 80:1844–51. [PubMed: 10468308]
15. MacArthur AC, McBride ML, Spinelli JJ, Tamaro S, Gallagher RP, Theriault GP. Risk of childhood leukemia associated with vaccination, infection, and medication use in childhood: the Cross-Canada Childhood Leukemia Study. Am J Epidemiol. 2008; 167:598–606. [PubMed: 18079130]
16. Naumburg E, Bellocco R, Cnattingius S, Jonzon A, Ekbom A. Perinatal exposure to infection and risk of childhood leukemia. Med Pediatr Oncol. 2002; 38:391–7. [PubMed: 11984799]
17. Cardwell CR, McKinney PA, Patterson CC, Murray LJ. Infections in early life and childhood leukaemia risk: a UK case-control study of general practitioner records. Br J Cancer. 2008; 99:1529–33. [PubMed: 18827817]
18. Roman E, Simpson J, Ansell P, et al. Childhood acute lymphoblastic leukemia and infections in the first year of life: a report from the United Kingdom Childhood Cancer Study. Am J Epidemiol. 2007; 165:496–504. [PubMed: 17182983]
19. Nishi M, Miyake H. A case-control study of non-T cell acute lymphoblastic leukaemia of children in Hokkaido, Japan. J Epidemiol Community Health. 1989; 43:352–5. [PubMed: 2614325]
20. Petridou E, Trichopoulos D, Kalapothaki V, et al. The risk profile of childhood leukaemia in Greece: a nationwide case-control study. Br J Cancer. 1997; 76:1241–7. [PubMed: 9365177]
21. Schuz J, Kaletsch U, Meinert R, Kaatsch P, Michaelis J. Association of childhood leukaemia with factors related to the immune system. Br J Cancer. 1999; 80:585–90. [PubMed: 10408870]
22. Wen W, Shu XO, Linet MS, et al. Allergic disorders and the risk of childhood acute lymphoblastic leukemia (United States). Cancer Causes Control. 2000; 11:303–7. [PubMed: 10843442]
23. Schuz J, Morgan G, Bohler E, Kaatsch P, Michaelis J. Atopic disease and childhood acute lymphoblastic leukemia. Int J Cancer. 2003; 105:255–60. [PubMed: 12673688]
24. Rosenbaum PF, Buck GM, Brecher ML. Allergy and infectious disease histories and the risk of childhood acute lymphoblastic leukaemia. Paediatr Perinat Epidemiol. 2005; 19:152–64. [PubMed: 15787890]
25. Hughes AM, Lightfoot T, Simpson J, et al. Allergy and risk of childhood leukaemia: Results from the UKCCS. Int J Cancer. 2007
26. Romagnani S. Regulation of the T cell response. Clin Exp Allergy. 2006; 36:1357–66. [PubMed: 17083345]
27. Ma X, Buffler PA, Layefsky M, Does MB, Reynolds P. Control selection strategies in case-control studies of childhood diseases. Am J Epidemiol. 2004; 159:915–21. [PubMed: 15128601]
28. Hansen HM, Wiemels JL, Wrensch M, Wiencke JK. DNA quantification of whole genome amplified samples for genotyping on a multiplexed bead array platform. Cancer Epidemiol Biomarkers Prev. 2007; 16:1686–90. [PubMed: 17684147]
29. Paynter RA, Skibola DR, Skibola CF, Buffler PA, Wiemels JL, Smith MT. Accuracy of multiplexed Illumina platform-based single-nucleotide polymorphism genotyping compared between genomic and whole genome amplified DNA collected from multiple sources. Cancer Epidemiol Biomarkers Prev. 2006; 15:2533–6. [PubMed: 17164381]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

30. The International HapMap Consortium. The International HapMap Project. Nature. 2003; 426:789–96. [PubMed: 14685227]
31. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21:263–5. [PubMed: 15297300]
32. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science. 2002; 296:2225–9. [PubMed: 12029063]
33. Packer BR, Yeager M, Burdett L, et al. SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. Nucleic Acids Res. 2006; 34:D617–21. [PubMed: 16381944]
34. Ritchie MD, Hahn LW, Roodi N, et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. American journal of human genetics. 2001; 69:138–47. [PubMed: 11404819]
35. Liu N, Zhang K, Zhao H. Haplotype-association analysis. Adv Genet. 2008; 60:335–405. [PubMed: 18358327]
36. Mathias RA, Gao P, Goldstein JL, et al. A graphical assessment of p-values from sliding window haplotype tests of association to identify asthma susceptibility loci on chromosome 11q. BMC Genet. 2006; 7:38. [PubMed: 16774684]
37. Chakraborty R, Weiss KM. Frequencies of complex diseases in hybrid populations. Am J Phys Anthropol. 1986; 70:489–503. [PubMed: 3766715]
38. Hanis CL, Chakraborty R, Ferrell RE, Schull WJ. Individual admixture estimates: disease associations and individual risk of diabetes and gallbladder disease among Mexican-Americans in Starr County, Texas. Am J Phys Anthropol. 1986; 70:433–41. [PubMed: 3766713]
39. Prescott SL, Macaubas C, Holt BJ, et al. Transplacental priming of the human immune system to environmental allergens: universal skewing of initial T cell responses toward the Th2 cytokine profile. J Immunol. 1998; 160:4730–7. [PubMed: 9590218]
40. Delespesse G, Yang LP, Ohshima Y, et al. Maturation of human neonatal CD4+ and CD8+ T lymphocytes into Th1/Th2 effectors. Vaccine. 1998; 16:1415–9. [PubMed: 9711781]
41. Ribeiro-do-Couto LM, Boeije LC, Kroon JS, et al. High IL-13 production by human neonatal T cells: neonate immune system regulator? Eur J Immunol. 2001; 31:3394–402. [PubMed: 11745358]
42. Prescott SL, Macaubas C, Smallacombe T, et al. Reciprocal age-related patterns of allergen-specific T-cell immunity in normal vs. atopic infants. Clin Exp Allergy. 1998; 28(Suppl 5):39–44. discussion 50-1. [PubMed: 9988446]
43. Romagnani S. The increased prevalence of allergy and the hygiene hypothesis: missing immune deviation, reduced immune suppression, or both? Immunology. 2004; 112:352–63. [PubMed: 15196202]
44. Trinchieri G. Interleukin-12 and the regulation of innate resistance and adaptive immunity. Nat Rev Immunol. 2003; 3:133–46. [PubMed: 12563297]
45. Pistiner M, Hunninghake GM, Soto-Quiros ME, et al. Polymorphisms in IL12A and cockroach allergy in children with asthma. Clin Mol Allergy. 2008; 6:6. [PubMed: 18671862]
46. Cargill M, Schrodi SJ, Chang M, et al. A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. Am J Hum Genet. 2007; 80:273–90. [PubMed: 17236132]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

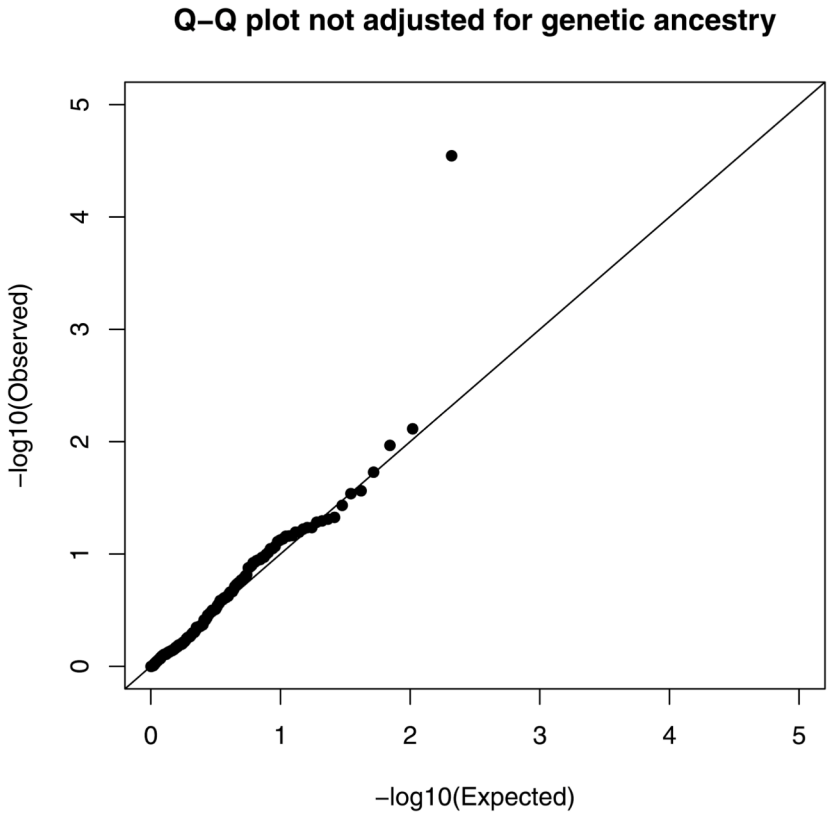

Figure 1.
Quantile-quantile (Q-Q) plot comparing the distribution of the observed vs. the expected $-\log_{10}$ p-values (log-additive model, not adjusted for genetic ancestry using data of combined race/ethnicity) of the 208 adaptive immunity single nucleotide polymorphisms

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2012 January 12.

Table 1

Demographic characteristics, genetic ancestry, birth order, and day care attendance of Hispanic and Non-Hispanic subjects, the Northern California Childhood Leukemia Study, 1995–2002

	Hispanic			Non-Hispanic		
	Case N (%)	**Control N (%)**	**P-value***	**Case N (%)**	**Control N (%)**	**P-value***
Sex						
Male	77 (49.4)	87 (48.6)	0.89	123 (55.7)	150 (55.8)	0.98
Female	79 (50.6)	92 (51.4)		98 (44.3)	119 (44.2)	
Age						
Mean age, years (SE)*	5.3 (0.3)	5.4 (0.3)	0.89	5.5 (0.2)	5.5 (0.2)	0.95
Race						
White	52 (33.3)	57 (31.8)	0.77	162 (73.3)	198 (73.6)	0.99
Black	2 (1.3)	2 (1.1)		14 (5.9)	13 (5.2)	
Native Americans	6 (3.9)	7 (3.9)		1 (0.5)	1 (0.4)	
Asian	0 (0)	2 (1.1)		33 (11.3)	25 (12.3)	
Others	96 (61.5)	111 (62.0)		23 (9.1)	20 (8.6)	
Genetic ancestry						
European, % (SE)†	51.8 (1.4)	52.7 (1.3)	0.60	80.4 (1.8)	82.7 (1.6)	0.33
African, % (SE)†	7.2 (0.7)	7.7 (0.6)	0.60	7.3 (1.2)	6.6 (1.1)	0.67
Amerindian, % (SE)†	41.0 (1.3)	39.5 (1.2)	0.41	12.3 (1.4)	10.7 (1.2)	0.36
Birth order						
First born	57 (38.3)	59 (34.9)	0.54	90 (43.3)	104 (41.1)	0.64
Having at least one older sibling	92 (61.7)	110 (65.1)		118 (56.7)	149 (58.9)	
Daycare attendance						
Thousand child hours before 6 months of age (SE)†	0.36 (0.11)	0.15 (0.05)	0.09	0.41 (0.10)	0.76 (0.13)	0.03
Thousand child hours before 1 year of age (SE)†	1.19 (0.30)	0.50 (0.13)	0.03	1.44 (0.25)	2.29 (0.32)	0.04

*P-values were generated using chi-squared tests for categorical variables (sex and race) and t-tests for continuous variables (age and genetic ancestry).

†SE= standard error

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 2

Single SNP analysis of adaptive immunity genes with p-value <0.05, the Northern California Childhood Leukemia Study, 1995–2002.

Rs number	Gene	Alleles†	Chromosomal location	Location	Tag or literature SNP	Race/Ethnicity‡	Cases No. HMa/Ht/HMi§	Controls No. HMa/Ht/HMi§	Log additive OR‖ (95% CI)	P-value
Th1										
rs583911	IL12A	A/**G**	3q25.33-q26	Intron 2	Tag	All	94/164/118	147/218/83	1.52 (1.25–1.85)	2.85×10^{-5}
rs2243154	IL12A	**A**/G			Tag	All	297/67/5	385/59/1	1.59 (1.11–2.27)	1.08×10^{-2}
rs17769459	STAT4	**A**/G	2q32.2-q32.3	Intron 3	Tag	NH	212/9/0	236/29/4	0.30 (0.14–0.63)	1.37×10^{-3}
rs16833215	STAT4	A/**G**		Intron 14	Tag	NH	81/107/32	125/116/27	1.38 (1.06–1.81)	1.85×10^{-2}
rs2247119	PHF11	A/**G**	13q14.3	Intron 2	Tag	NH	116/91/12	118/121/30	0.69 (0.52–0.92)	1.21×10^{-2}
rs1925742	PHF11	A/**G**			Tag	NH	73/117/31	118/117/34	1.33 (1.01–1.74)	4.02×10^{-2}
rs3181224	IL12B *	A/**G**	5q31.1-q33.1	3′ near gene	Tag	NH	165/55/1	227/37/3	1.73 (1.12–2.65)	1.27×10^{-2}
rs11574790	IL12B *	**A**/G		Intron 6	Tag	NH	165/55/1	226/39/4	1.55 (1.03–2.35)	3.75×10^{-2}
rs1368439	IL12B	A/**C**		3′ UTR	Tag	NH	169/51/1	189/71/9	0.68 (0.46–0.98)	4.07×10^{-2}
Th2/Allergies										
rs3781093	GATA3	A/**G**	10p15	Intron 3	Tag	NH	131/80/9	196/62/9	1.62 (1.17–2.26)	4.08×10^{-3}
rs1059513	STAT6	A/**G**	12q13	3′ UTR	Literature	NH	169/46/6	225/42/1	1.68 (1.11–2.54)	1.40×10^{-2}
Treg										
rs1554286	IL10 *	**A**/G	1q31-q32	Intron 3	Tag	All	167/173/35	242/168/36	1.36 (1.08–1.70)	7.67×10^{-3}
rs3024505	IL10	**A**/G		3′ near gene	Tag	Hispanic	107/44/2	144/34/1	1.73 (1.07–2.80)	2.62×10^{-2}
rs3024490	IL10 *	**A**/C		Intron 1	Tag	All	150/177/44	210/191/45	1.24 (1.00–1.54)	4.91×10^{-2}
rs926169	CTLA4	**A**/C	2q33		Tag	All	140/175/62	136/218/93	0.80 (0.66–0.98)	2.74×10^{-2}
rs733618	CTLA4	A/**G**		5′ near gene	Tag	All	316/55/5	347/93/8	0.70 (0.50–0.96)	2.90×10^{-2}
rs706778	IL2RA	**A**/G	10p15-p14	Intron 1	Tag	NH	82/102/37	69/149/50	0.75 (0.57–0.98)	3.35×10^{-2}
rs942201	IL2RA	**A**/C		Intron 1	Tag	All	243/114/18	311/122/12	1.29 (1.00–1.67)	4.71×10^{-2}
B-cell										
rs1385477	MME	**A**/C	3q25.1-q25.2		Tag	NH	116/90/15	120/118/31	0.74 (0.56–0.98)	3.78×10^{-2}

*Markers in linkage disequilibrium ($r^2 > 0.8$): IL12B: rs3181224 - rs11574790, r^2=0.95 for Non-Hispanics; IL10: rs1554286 - rs3024490, r^2=0.83 for all subjects

†Bolded letter indicates the minor allele

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

‡NH=Non-Hispanic

§HMa=homozygous with 2 major alleles; Ht=heterozygous; HMi=homozygous with 2 minor alleles

//Odds ratios were adjusted for age, sex, Hispanic status, and race using unconditional logistic regression.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 3

Interaction between birth order or daycare attendance and *IL12A* rs583911 on the risk of childhood ALL, the Northern California Childhood Leukemia Study, 1995–2002.

	All ethnicities					
	Case No. AA/AG/GG	**Control No. AA/AG/GG**	**OR for every one copy of G allele of rs583911***			
Birth order						
First born	29/66/52	57/80/25	2.14 (1.52–3.01)			
Having at least one older sibling	58/88/63	81/124/53	1.30 (1.00–1.69)			
		Interaction p-value = 0.01				
	Non-Hispanics			**Hispanics**		
	Case No. AA/AG/GG	Control No. AA/AG/GG	OR for every one copy of G allele of rs583911 *	Case No. AA/AG/GG	Control No. AA/AG/GG	OR for every one copy of G allele of rs583911 *
Day care attendance before 6 months of age (thousand child hours)†						
< 2	49/86/61	83/102/41	1.68 (1.27–2.22)	27/64/50	45/88/33	1.69 (1.21–2.37)
≥ 2	10/4/2	9/19/3	0.44 (0.15–1.31)	2/6/3	3/1/2	2.04 (0.33–12.44)
		Interaction p-value = 0.02			Interaction p-value = 0.94	
Day care attendance before 1 year of age (thousand child hours)†						
< 5	48/83/58	80/97/39	1.64 (1.24–2.18)	26/63/48	44/87/33	1.68 (1.18–2.34)
≥ 5	11/7/5	12/24/5	0.84 (0.37–1.90)	3/7/5	4/2/2	6.00 (0.90–40.15)
		Interaction p-value = 0.11			Interaction p-value = 0.66	

*Odds ratios were adjusted for age, sex, Hispanic status (for birth order analysis), race, and annual household income using unconditional logistic regression.

†Cutoff was selected based on the median thousand hours of day care attendance among controls who had day care attendance.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Haplotype analyses of *CD28* and *FCGR2A* for combined Race/Ethnicity, the Northern California Childhood Leukemia Study, 1995–2002

Genes/Haplotypes	Control %	Case %	OR (95% CI)*	P-value
CD28				
rs1879877, rs3181096, rs1181389, rs3769683				
1. CAAG	36	37	Reference	
2. CGAG	30	27	0.89 (0.70–1.13)	0.33
3. AGGG	15	16	1.09 (0.81–1.46)	0.58
4. AGAA	9	9	0.98 (0.67–1.44)	0.92
5. CGAA	5	9	1.55 (1.03–2.34)	0.04
6. Rare haplotypes†	4	2	0.47 (0.25–0.88)	0.01
			Global P-value: 0.02	
FCGR2A				
rs10800309, rs4656308				
1. GA	61	57	Reference	
2. AG	23	21	0.98 (0.78–1.25)	0.90
3. AA	16	21	1.46 (1.13–1.90)	0.004
			Global P-value: 0.02	

*ORs were adjusted for age, sex, Hispanic status, and race using unconditional logistic regression

†Haplotypes with frequencies less than 5% were grouped together for analysis

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 5

Haplotype analyses of *GATA3*, *IL2RA*, *STAT4*, and *STAT6* by Hispanic Status, the Northern California Childhood Leukemia Study, 1995–2002

Genes/Haplotypes	Non-Hispanics				Hispanics			
	Control %	Case %	OR (95% CI)*	P-value	Control %	Case %	OR (95% CI)*	P-value
GATA3								
rs4143094, rs3781093, and rs3802604								
1. CAA	52	51	Reference		38	43	Reference	
2. AAG	16	15	0.95 (0.65–1.39)	0.78	13	14	0.92 (0.57–1.49)	0.73
3. CGG	14	21	1.51 (1.05–2.18)	0.03	38	38	0.89 (0.63–1.26)	0.52
4. AAA	9	7	0.79 (0.46–1.35)	0.39	†	†	†	
5. CAG	8	4	0.55 (0.30–1.03)	0.06	6	2	0.32 (0.13–0.79)	0.01
6. Rare haplotypes‡	1	2	1.51 (0.34–6.75)	0.59	5	2	0.36 (0.14–0.92)	0.03
			Global P-value: 0.03				Global P-value: 0.04	
IL2RA								
rs6602398, rs942201, rs791587, and rs706778								
1. CCGG	25	21	Reference		14	6	Reference	
2. ACAG	20	26	1.51 (1.02–2.24)	0.04	24	25	2.35 (1.19–4.62)	0.01
3. CCAA	16	9	0.64 (0.39–1.04)	0.07	8	6	1.84 (0.80–4.24)	0.15
4. CAGA	15	19	1.32 (0.87–1.98)	0.19	10	12	2.74 (1.26–6.00)	0.01
5. CCAG	9	11	1.40 (0.81–2.43)	0.22	11	10	1.95 (0.86–4.42)	0.11
6. CCGA	9	8	1.04 (0.55–1.97)	0.90	30	32	2.33 (1.19–4.55)	0.01
7. Rare haplotypes‡	7	5	0.82 (0.44–1.53)	0.53	2	9	9.15 (2.82–29.61)	0.0003
			Global P-value: 0.02				Global P-value: 0.004	
STAT4								
rs17769459, rs4853546, and rs1031509								
1. GGC	59	62	Reference		71	64	Reference	
2. GAA	28	29	0.99 (0.74–1.33)	0.96	25	25	1.12 (0.77–1.62)	0.55
3. GAC	6	7	1.13 (0.66–1.92)	0.66	†	†	†	
4. Rare haplotypes‡	7	2	0.30 (0.14–0.64)	0.002	5	11	2.65 (1.41–4.99)	0.003
			Global P-value: 0.008				Global P-value: 0.008	
STAT6								

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Genes/Haplotypes	Non-Hispanics				Hispanics			
	Control %	Case %	OR (95% CI)*	P-value	Control %	Case %	OR (95% CI)*	P-value
rs4559, rs1059513, and rs324015								
1. AAG	56	47	Reference		51	50	Reference	
2. GAA	26	27	1.23 (0.91–1.65)	0.18	31	27	0.88 (0.61–1.26)	0.48
3. GAG	11	13	1.49 (0.99–2.25)	0.06	9	16	1.82 (1.12–2.96)	0.02
4. AGG	8	13	1.92 (1.24–2.97)	0.003	8	7	0.83 (0.45–1.55)	0.56
			Global P-value: 0.02				Global P-value: 0.04	

*ORs were adjusted for age, sex, Hispanic status, and race using unconditional logistic regression.

†included in the rare haplotypes group

‡Haplotypes less with frequencies less than 5% were grouped together for analysis.