

HLA-DP genetic variation, proxies for early life immune modulation and childhood acute lymphoblastic leukemia risk

Kevin Y. Urayama,^{1,2} Anand P. Chokkalingam,¹ Catherine Metayer,¹ Xiaomei Ma,³ Steve Selvin,¹ Lisa F. Barcellos,⁴ Joseph L. Wiemels,⁵ John K. Wiencke,⁵ Malcolm Taylor,⁶ Paul Brennan,⁷ Gary V. Dahl,⁸ Priscilla Moonsamy,⁹ Henry A. Erlich,⁹ Elizabeth Trachtenberg,¹⁰ and Patricia A. Buffler¹

¹School of Public Health, University of California, Berkeley, CA; ²Center for Clinical Epidemiology, St Luke's Life Science Institute, Tokyo, Japan; ³Yale University School of Medicine, New Haven, CT; ⁴Genetic Epidemiology and Genomics Laboratory, University of California, Berkeley, CA; ⁵Laboratory for Molecular and Neuroepidemiology, University of California, San Francisco, CA; ⁶Cancer Immunogenetics Group, University of Manchester, St Mary's Hospital, Manchester, United Kingdom; ⁷Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France; ⁸Division of Pediatric Hematology/Oncology/BMT, Stanford University School of Medicine, Stanford, CA; ⁹Roche Molecular Systems Inc, Pleasanton, CA; and ¹⁰Center for Genetics, Children's Hospital Oakland Research Institute, CA

The human leukocyte antigen (HLA) genes are candidate genetic susceptibility loci for childhood acute lymphoblastic leukemia (ALL). We examined the effect of *HLA-DP* genetic variation on risk and evaluated its potential interaction with 4 proxies for early immune modulation, including measures of infectious exposures in infancy (presence of older siblings, daycare attendance, ear infections) and breastfeeding. A total of 585 ALL cases and 848 controls were genotyped at the *HLA-DPA1* and *DPB1* loci. Because

of potential heterogeneity in effect by race/ethnicity, we included only non-Hispanic white (47%) and Hispanic (53%) children and considered these 2 groups separately in the analysis. Logistic regression analyses showed an increased risk of ALL associated with *HLA-DPB1*01:01* (odds ratio [OR] = 1.43, 95% CI, 1.01-2.04) with no heterogeneity by Hispanic ethnicity ($P = .969$). Analyses of *DPB1* super-types showed a marked childhood ALL association with DP1, particularly for high-hyperdiploid ALL (OR = 1.83; 95% CI,

1.20-2.78). Evidence of interaction was found between DP1 and older sibling ($P = .036$), and between DP1 and breastfeeding ($P = .094$), with both showing statistically significant DP1 associations within the lower exposure categories only. These findings support an immune mechanism in the etiology of childhood ALL involving the *HLA-DPB1* gene in the context of an insufficiently modulated immune system. (*Blood*. 2012;120(15):3039-3047)

Introduction

Evidence from a growing number of studies indicates that exposure to common infections early in life may play a role in the etiology of childhood leukemia.¹ However, unlike leukemia in some animals² and HTLV-associated adult T-cell leukemia,³ studies have not been able to show a direct pathologic involvement of specific infection(s) for leukemia in children.⁴ It has been suggested that for childhood leukemia, particularly for precursor B-cell acute lymphoblastic leukemia (ALL), a delay in exposure to common childhood infections leaves the immune network undermodulated, and subsequent exposure to infections results in an unfavorable immune response; specifically, a proliferative advantage to preleukemia cells.⁵ This "delayed infection" hypothesis suggests that exposure to infections or other immune-modulatory factors early in life may result in a protective effect against childhood ALL. Evidence from epidemiologic studies using proxies for early life immune modulation support this hypothesis, such as childhood social contacts increasing the likelihood of infectious transmissions and breastfeeding, but provide limited insight on mechanistic aspects. Host genetic variation within immune response genes may contribute to risk of childhood ALL and may operate in concert with early life exposures that are involved in immune modulation.

Within the MHC region on chromosome 6 reside a large number of genes involved in immunologic processes, including the highly

polymorphic human leukocyte antigen (HLA) genes. The cell surface glycoproteins encoded by the classical HLA genes of the class II region, *HLA-DP*, *-DQ*, and *-DR*, are expressed exclusively on antigen-presenting cells (eg, macrophages, B lymphocytes, and dendritic cells) and primarily bind and present endocytosed extracellular proteins to CD4⁺ T lymphocytes to initiate a targeted immune response.^{6,7} The known polymorphisms are largely localized to exon 2, the region of the HLA class II gene that determines characteristics of the peptide-binding groove of the HLA molecule and the resulting diversity of antigens that are presented to the immune system.⁸ Because of their association with the peptide-binding groove, the HLA polymorphic regions have functional relevance and significance for disease susceptibility; > 50 autoimmune and infectious disease associations with HLA alleles have been identified.^{7,9} There are many examples of HLA associations with neoplasms, including Hodgkin lymphoma,^{10,11} nasopharyngeal carcinoma,¹² gastric cancer,¹³ invasive cervical carcinoma,¹⁴ and childhood leukemia.¹⁵⁻¹⁸ Specifically, in the United Kingdom Childhood Cancer Study (UKCCS), Taylor et al in 2002 reported a significantly increased risk of childhood ALL associated with presence of the *DPB1*02:01* allele, as well as with other rarer alleles that share antigen binding site pocket profiles with the *DPB1*02:01* allotype.¹⁸ Further examination of these data as

Submitted January 16, 2012; accepted July 30, 2012. Prepublished online as *Blood* First Edition paper, August 24, 2012; DOI 10.1182/blood-2012-01-404723.

The publication costs of this article were defrayed in part by page charge

payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2012 by The American Society of Hematology

supertypes (a classification based on overlapping peptide binding specificity) showed that childhood ALL risk increased in the presence of DP2 and DP8 supertypes and decreased in the presence of the DP1 supertype.^{19,20}

In the current analysis, we examined the associations of *HLA-DPA1* and *DPB1* genetic variation, both independently and jointly with 4 separate proxies indicative of early life immune modulation, including markers for infectious exposures in infancy (presence of older siblings, daycare attendance by age 6 months, and reported ear infections in infancy) and breastfeeding. We focused on non-Hispanic white and Hispanic children enrolled in the Northern California Childhood Leukemia Study (NCCLS), which together compose the majority of the study population. We previously reported an ethnic difference in risk for birth order and daycare attendance^{21,22} and, thus, consider these 2 groups separately in the current analysis. This study is unique in the inclusion of Hispanic children, an assessment of potential contribution of the *HLA-DPA1* locus, and the evaluation of gene-environment interaction.

Methods

Study participants

The NCCLS is an ongoing case-control study designed to investigate the etiology of pediatric leukemias. Beginning in 1995, newly diagnosed childhood leukemia cases were ascertained at the time of diagnosis from major pediatric hospitals located in a 17-county San Francisco Bay Area study region, which was expanded in 1999 to 35 counties in Northern and Central California. Comparison with the California Cancer Registry (1997-2003) showed that the NCCLS case ascertainment protocol has captured ~95% of children diagnosed with leukemia in the participating study hospitals. When considering both participating and nonparticipating hospitals within the study region, cases ascertained through the NCCLS represent ~76% of all diagnosed cases. For each eligible case, statewide birth records maintained by the Center for Health Statistics of the California Department of Public Health were used to generate a list of randomly selected controls that matched the case on child's date of birth, sex, Hispanic status (a biologic parent who is Hispanic), and maternal race. Information obtained through the birth certificates and commercially available searching tools were used to trace and enroll 1 or 2 matched controls for each case.

Cases and controls were considered eligible if they were under 15 years of age at date of diagnosis for cases (or corresponding reference date for controls), resided in the study region at the date of diagnosis, had a parent or guardian who spoke either English or Spanish, and had no prior history of malignancy. Approximately 85% of eligible cases consented to participate. Among all eligible controls contacted, 86% consented to participate.²³ The overall participation for the control subjects was 59% (the number of enrolled controls divided by the total number of control searches, excluding the known and presumed ineligible). A detailed description of control selection in the NCCLS is reported elsewhere.^{23,24} A previous evaluation showed that participating controls in the NCCLS are representative of the sampled population with respect to parental age, parental education, and mother's reproductive history.²⁴

The current analysis included non-Hispanic white and Hispanic childhood ALL case and control subjects recruited between 1995 and 2008 (study phases 1-3), the 2 largest racial/ethnic groups that together compose ~85% of enrolled subjects. The other race/ethnicity groups were not considered because of the small number of subjects. Children were classified as Hispanic if at least 1 biologic parent self-identified as Hispanic. The non-Hispanic white group was composed of children with both biologic parents self-identifying as non-Hispanic white. Children younger than 1 year at diagnosis/reference date were excluded because of growing evidence that these leukemias may be etiologically distinct compared with

leukemia diagnosed at later ages.²⁵ This resulted in 669 ALL cases and 977 controls 1 year of age and older who were non-Hispanic white or Hispanic. Among these, only subjects with a DNA sample available at the start of this current project were included resulting in 590 cases (88.2%) and 854 controls (87.4%).

The study protocol was approved by the Institutional Review Boards of University of California, Berkeley and all collaborating institutions, and written informed consent was obtained from all participating subjects. This study was conducted in accordance with the Declaration of Helsinki.

HLA-DPA1/DPB1 genotyping

DNA specimens from buccal cytobrushes were obtained from case and control children either at the hospital (for some cases) or during the in-home personal interview and were extracted using an automated DNA extraction system (AutoGen) and FlexiGene reagents from QIAGEN. Whole genome amplification of buccal cell DNA was performed using GenomePlex reagents (Rubicon Genomics) according to the manufacturer's protocol. Whole genome amplification products were cleaned with a Montage PCR9 filter plate (Millipore). When buccal cytobrush DNA was inadequate or not available (26.6% of subjects), DNA was isolated from dried bloodspots collected at birth and archived at -20°C by the Genetic Diseases Screening Program of the California Department of Public Health. After extraction using the QIAamp DNA Mini Kit (QIAGEN), these DNA samples were whole-genome amplified using REPLI-g reagents (QIAGEN). Regardless of source, DNA specimens were quantified using human-specific Alu-PCR to confirm a minimum level of amplifiable human DNA.²⁶

HLA-DPA1 and *DPB1* genotyping was conducted on 1444 unique samples, in addition to 10 sets of Center d'Etude du Polymorphisme Humain family trios and duplicates for 10% of study samples for quality control. We used immobilized sequence-specific oligonucleotide probe strips that were designed to type both *HLA-DPA1* and *DPB1* loci on the same strip (Roche Molecular Systems). There were 21 probes for *DPA1* and 48 probes for *DPB1* on these strips. Exon 2 of the *DPA1* and *DPB1* loci were amplified in separate PCR reactions using biotinylated primer pairs specific for each locus. Subsequent to amplification, 35 μ L of each PCR product was heat denatured at 95°C for 5 minutes and allowed to hybridize to the immobilized sequence-specific oligonucleotide probe strip membranes using an automated hybridization/wash instrument, the SLT Pro-fiblot (Tecan Systems). The developed probe patterns on the strips were then scanned, and the scanned images were analyzed using a computer pattern matching program, StripScan Version 5.7.8 (Roche Molecular Systems) with the ImMunoGeneTics/HLA reference alignment Version HLADB-2.10.0-July 2005 (<http://www.ebi.ac.uk/imgt/hla/>) to determine the genotypes. A resource cataloging all common and well-documented HLA alleles formulated by the American Society for Histocompatibility and Immunogenetics was used to help resolve ambiguities to the lowest value common allele in the genotype determination process.²⁷

No Mendelian errors based on the *HLA-DPA1* and *DPB1* genotypes were observed for the Center d'Etude du Polymorphisme Humain family trios, and the genotypes between duplicate samples were 100% concordant. Of the 1444 unique study samples genotyped, 11 samples (0.8%) failed to provide an interpretable strip probe hybridization pattern for genotype determination. This resulted in a total of 585 ALL cases (265 non-Hispanic white and 320 Hispanic) and 848 controls (408 non-Hispanic white and 440 Hispanic) available for the analysis. For a large subset of the ALL cases (87%), data on hyperdiploidy and *TEL-AML1* chromosomal translocation were available as described in detail previously.²⁸ Subtypes of the cases included 291 common ALL (cALL, defined as CD10⁺ and CD19⁺ ALL age 2-5 years), 172 high-hyperdiploid (51-67 chromosomes), 89 positive for the *TEL-AML1* chromosomal translocation, and 70 ALL with normal karyotypes.

Data collection

The proxies for early life immune-modulatory exposures considered in this current analysis have been described in detail in previous NCCLS publications.²² Data on child's social contacts inside and outside the home (birth order and daycare attendance), common infections during the first

year of life, and breastfeeding were obtained using a standardized questionnaire administered in-person with the parents/guardians of each child.

The child's birth order was determined based on a reproductive history of the biologic mother. Information on the child's social contacts outside the home was obtained through a history of daycare and preschool attendance before the reference date (date of diagnosis for cases and corresponding date for matched controls) or before the age of 6 years, whichever occurred first. For each daycare and/or preschool the child attended, information on age at attendance, duration of attendance, hours per week of attendance, and number of other children present was obtained. The calculation of a quantitative summary measure, called total child-hours of exposure, has been previously described in detail.²¹ Briefly, child-hours at each daycare facility were calculated as follows: (number of months attending the daycare) × (mean hours per week at this daycare) × (number of other children at this daycare) × (4.35 weeks per month). For each child, the child-hours in each daycare setting were summed to obtain the total child-hours of exposure.

Respondents were asked for a history of common infectious illnesses the child had during the first year of life, including severe diarrhea/vomiting, ear infection, persistent cough, mouth and eye infection, influenza, and unspecified "other infection." Questions were asked with an emphasis on the timing of exposure, specifically during the age of < 3 months, 3-5 months, and/or 6-12 months.

HLA-DPB1 supertype classification

Although each HLA class II allele encodes a distinct molecule, it has been shown that many of these molecules have overlapping structural and functional features through shared peptide-binding pockets and thus can be clustered into distinct functional groups referred to as superotypes. Twenty-three of the 28 observed *DPB1* alleles can be clustered into 1 of 6 superotypes described previously by Taylor et al.¹⁹ The 6 superotypes were defined by 3 dimorphic amino acid residues in pockets 6, 4, and 1 at positions β 11, β 69, and β 84, respectively, which have been shown to play significant roles in the binding characteristics of the peptide-binding groove.^{19,29,30} Three of the superotypes have a positively charged lysine (K) at β 69, whereas the other 3 have a negatively charged glutamic acid (E). This is accompanied by either a glycine (G) or leucine (L) at β 11 and a G or an aspartic acid (D) at β 84. The 6 superotypes have been referred to as DP1 (GKD), DP2 (GEG), DP3 (LKD), DP4 (GKG), DP6 (LED), and DP8 (GED).

Statistical analysis

Python for Population Genetics Version 0.7.0 (Pypop) software was used to test *HLA-DPA1* and *DPB1* genotype frequencies for deviation from Hardy-Weinberg equilibrium in controls, to estimate *HLA-DPA1-DPB1* haplotype frequencies from genotype data of unrelated persons using the expectation-maximization algorithm, and to evaluate the patterns of linkage disequilibrium between the 2 loci.³¹ Comparison of *HLA-DPA1* and *DPB1* allele frequencies between cases and controls was performed using the Fisher exact test. A logistic regression model adjusting for child's age and sex was used to calculate odds ratios (ORs) and 95% CIs associated with carrier status of specific *HLA-DPA1* and *DPB1* alleles (referred to as phenotype) and superotypes. Analyses were conducted among non-Hispanic white and Hispanic children separately; and if no evidence of heterogeneity between the 2 groups based on the χ^2 test of homogeneity ($P > .05$) was present, a pooled analysis was performed with adjustment for race/ethnicity (non-Hispanic white, Hispanic white, and Hispanic other).

The multivariable evaluation of the 4 proxy exposure measures (older sibling, daycare child-hours, ear infection, and breastfeeding) have been reported in previous NCCLS publications,²² and those results served as the basis for the focus on the specific variables chosen for evaluation in the joint analysis. The joint effect between the DP1 supertype and each of the 4 proxy measures was evaluated separately using logistic regression models that included an interaction term representing the product of the supertype (presence or absence) and exposure measure, and additionally adjusting for the other 3 proxy measures, child's age and sex, maternal age, maternal

education, annual household income, and study phase. The risk estimate for the interaction term is defined as the ratio of the joint effect of the exposure variable and DP supertype (OR_{GE}) and the product of the individual effects [$OR_{interaction} = OR_{GE} / (OR_{GE} \times OR_{G-E})$]. An $OR_{interaction}$ significantly different from 1.0 is an indication of multiplicative interaction. P values for the interaction term of < .10 were considered statistically significant.

Results

For both non-Hispanic white and Hispanic children, compared with controls, cases tended to have less annual household income and mothers with less formal education (Table 1). Case-control difference for maternal age was observed only in non-Hispanic white children with mothers of cases appearing younger.

A total of 6 *DPA1* alleles were observed in non-Hispanic white and Hispanic subjects. *DPA1**01:03 was the most common with a frequency of about 0.80 (Table 2). In contrast to *DPA1*, *DPB1* was considerably more polymorphic. A total of 28 different alleles were observed in the non-Hispanic white and Hispanic populations combined (Table 2). However, only 5 alleles were present at 5% or higher, including *DPB1**01:01, *02:01, *03:01, *04:01, and *04:02. Observed genotype frequencies for *DPA1* and *DPB1* did not differ from what was expected based on Hardy-Weinberg proportions in non-Hispanic white and Hispanic controls.

Strong linkage disequilibrium was present between the 2 loci ($D' = 0.92$, $W_n = 0.73$, $P < .001$), with only 4 common haplotypes (> 5%) observed (*DPA1**01:03-*DPB1**02:01, *DPA1**01:03-*DPB1**03:01, *DPA1**01:03-*DPB1**04:01, *DPA1**01:03-*DPB1**04:02). Cases and controls had similar frequencies of common *DPA1-DPB1* haplotypes.

In the logistic regression analysis evaluating *DPA1* and *DPB1* phenotypes and childhood ALL risk, no evidence of heterogeneity in effect between non-Hispanic white and Hispanic children was observed (Table 3). Similarly elevated risk estimates were observed for *DPB1**01:01 carriers in both groups, and the pooled analysis showed a statistically significant increased risk of childhood ALL associated with carriers of the allele (OR = 1.43; 95% CI, 1.01-2.04; Table 3).

Previous work has shown more than 90% of the *DPB1* alleles can be clustered into 6 superotypes based on similarities in functional (peptide binding) properties.¹⁹ The most common superotype, DP4, was present in a large majority of non-Hispanic white (79.2%) and Hispanic (82.5%) control children. The DP1 superotype, which includes *DPB1* alleles *01:01, *05:01, and *50:01, showed elevated ORs and a statistically significant increased risk associated with the high-hyperdiploid ALL subtype (OR = 1.83; 95% CI, 1.20-2.78; $P = .005$; Table 4).

For the evaluation of interaction, we focused on the associated DP1 supertype and 4 proxies for early life immune modulation, including breastfeeding, history of ear infections during the first year of life, and birth order and daycare attendance as indicators for the likelihood of infectious transmissions. This analysis was conducted within a subset of the subjects, including those with complete data for all variables (non-Hispanic whites: 231 cases and 364 controls; Hispanics: 276 cases and 398 controls). Within this subset of subjects, similar to previous NCCLS reports,^{21,22} having an older sibling (OR = 0.57; 95% CI, 0.40-0.82) and greater daycare child-hours (≥ 2000 vs < 2000: OR = 0.46; 95% CI, 0.25-0.85) was associated with a statistically significant reduced risk of childhood ALL only in non-Hispanic whites and not Hispanic children (Table 5). In contrast, the risk estimate for history of ear infection before 6 months of age was the same in both

Table 1. Selected sociodemographic characteristics of ALL cases and controls in non-Hispanic white and Hispanic children, NCCLS, 1995-2008

	Non-Hispanic white					Hispanic				
	Cases		Controls		<i>P</i> *	Cases		Controls		<i>P</i> *
	n	%	n	%		n	%	n	%	
Total	265		408			320		440		
Child's sex										
Male	152	57.4	243	59.6	.571	168	52.5	229	52.1	.941
Female	113	42.6	165	40.4		152	47.5	211	48.0	
Child's age										
1 y	19	7.2	35	8.6	.470	27	8.4	43	9.8	.706
2-5 y	156	58.9	215	52.7		183	57.2	233	53.0	
6-10 y	58	21.9	100	24.5		75	23.4	112	25.5	
11-14 y	32	12.1	58	14.2		35	10.9	52	11.8	
Maternal age										
< 20 y	17	6.4	8	2.0	.003	44	13.8	56	12.7	.062
20-24 y	33	12.5	41	10.1		107	33.5	117	26.6	
25-29 y	66	24.9	115	28.2		74	23.2	138	31.4	
30-34 y	102	38.5	137	33.6		57	17.9	88	20.0	
≥ 35 y	47	17.7	107	26.2		37	11.6	41	9.3	
Missing	0		0			1		0		
Mean (SD)	30.1	(5.7)	31.5	(5.6)	.002	26.5	(6.2)	26.9	(5.7)	.287
Maternal education										
≤ High school	63	23.9	70	17.2	.088	219	68.4	271	61.7	.162
Some college	86	32.6	136	33.3		70	21.9	116	26.4	
≥ College	115	43.6	202	49.5		31	9.7	52	11.9	
Missing	1		0			0		1		
Annual income, household										
< \$15 000	16	6.0	16	3.9	< .001	85	26.6	74	16.8	.001
\$15 000-\$29 000	25	9.4	17	4.2		84	26.3	100	22.7	
\$30 000-\$44 000	33	12.5	31	7.6		59	18.4	87	19.8	
\$45 000-\$59 000	43	16.2	50	12.3		44	13.8	65	14.8	
\$60 000-\$74 000	32	12.1	52	12.8		18	5.6	42	9.6	
≥ \$75 000	116	43.8	242	59.3		30	9.4	72	16.4	

**P* values were calculated using the Pearson χ^2 test for categorical variables and Student *t* test for continuous variables, not including those missing.

race/ethnicity groups and the pooled analysis showed a statistically significant reduced risk of childhood ALL (vs no ear infection in the first year: OR = 0.51; 95% CI, 0.31-0.83). Breastfeeding did not show a statistically significant association with childhood ALL risk in either group (Table 5). Based on these main effect results, interaction analyses focused on non-Hispanic whites for the evaluation of older sibling and daycare child-hours and both non-Hispanic white and Hispanic children combined for ear infection and breastfeeding.

The risk of childhood ALL associated with DP1 supertype differed between children without an older sibling (OR = 2.68; 95% CI, 1.22-5.91) and with an older sibling (OR = 0.77; 95% CI, 0.38-1.57) and showed statistically significant evidence of interaction (OR_{interaction} = 0.33; 95% CI, 0.12-0.93; *P* = .036) in non-Hispanic white children (Figure 1). A consistent finding was observed for breastfeeding (OR_{interaction} = 0.47; 95% CI, 0.20-1.14; *P* = .094) where the DP1 supertype was associated with an increased risk of ALL among children who did not breastfeed (OR = 3.04; 95% CI, 1.26-7.30), whereas the DP1 supertype was not associated with risk among children who did breastfeed (OR = 1.18; 95% CI, 0.83-1.69). In contrast, daycare child-hours and ear infections in infancy showed little evidence of multiplicative interaction with the DP1 supertype in the risk of childhood ALL (Figure 1), but these analyses had limited statistical power. However, the point estimates were in the direction consistent with the results observed for the older sibling and breastfeeding proxy measures.

Discussion

Exposure to common infections and the role of immune-related processes have both emerged as strong candidate risk factors for childhood ALL.⁵ In this current study, we examined the *HLA-DP* association in a large non-Hispanic white and Hispanic population of childhood ALL cases and controls in California. The DP1 supertype was associated with childhood ALL; this effect was most prominent in the high-hyperdiploid ALL subtype. Furthermore, we observed a statistically significant interaction between the DP1 supertype and 2 separate measures used as proxies for early life immune-modulatory exposures, having an older sibling and breastfeeding, in the risk of childhood ALL.

In a previous study by Taylor et al, DPB1*02:01 and rarer alleles that share functional properties with the DPB1*02:01 allotypic protein were associated with risk of cALL.¹⁸ In addition, DPB1*04:02 was associated with an increased risk, whereas DPB1*01:01 showed a reduced risk. Further investigation of *HLA-DPB1* alleles as superotypes indicated a reduced risk associated with DP1 (GKD),¹⁹ which appeared to be driven by the DPB1*01:01 allele.²⁰ This effect was most marked for *TEL-AML1* and high hyperdiploid-positive ALL.²⁰ A DP1 association in childhood ALL, particularly in high-hyperdiploid ALL, is also present in the NCCLS; however, the risk estimate was in the opposite direction (ie, increased risk) from that of the UKCCS

Table 2. Frequency of HLA-DPA1 and DPB1 alleles in ALL cases and controls in non-Hispanic white and Hispanic children, NCCLS, 1995-2008

	Non-Hispanic white					Hispanic				
	Cases		Controls		P	Cases		Controls		P
	n*	%	n*	%		n*	%	n*	%	
DPA1 alleles										
01:03	427	80.6	669	82.0	.679	492	77.1	688	78.5	.510
02:01	81	15.3	111	13.6	.390	94	14.7	136	15.5	.666
02:02	18	3.4	30	3.7	.883	44	6.9	46	5.3	.217
Others†	4	0.7	6	0.7	NA	8	1.3	6	0.7	NA
Total chromosomes	530		816			638		876		
DPB1 alleles										
01:01	36	6.8	37	4.5	.104	35	5.5	36	4.1	.266
02:01	70	13.2	116	14.2	.633	77	12.0	97	11.0	.622
03:01	57	10.8	84	10.3	.789	37	5.8	57	6.5	.595
04:01	206	38.9	336	41.2	.468	155	24.2	233	26.5	.338
04:02	64	12.1	96	11.8	.864	214	33.4	290	33.0	.954
05:01	10	1.9	18	2.2	.845	20	3.1	23	2.6	.638
10:01	11	2.1	14	1.7	.681	6	0.9	14	1.6	.363
11:01	10	1.9	21	2.6	.456	11	1.7	17	1.9	.850
13:01	12	2.3	15	1.8	.696	11	1.7	19	2.2	.593
14:01	6	1.1	5	0.6	.354	20	3.1	25	2.8	.883
17:01	10	1.9	12	1.5	.666	20	3.1	24	2.7	.751
Others†	38	7.1	62	7.6	NA	34	5.3	45	5.1	NA
Total chromosomes	530		816			640		880		

NA indicates not applicable.

*Number and percent of chromosomes typing for each allele.

†Group of rare HLA-DP alleles < 0.02 frequency in both non-Hispanic white and Hispanic children. Alleles not shown include: DPA1, 01:04, 03:01, and 04:01; and DPB1, 02:02, 102:01, 06:01, 09:01, 15:01, 16:01, 18:01, 19:01, 20:01, 23:01, 26:01, 29:01, 34:01, 40:01, 51:01, 54:01, and 60:01.

study,²⁰ and the other previously reported DPB1 associations were not observed.

In interpreting this discrepancy, one observation to note is the marked differences in allele frequencies for DPB1*01:01 and DPB1*02:01 between the UKCCS controls and NCCLS non-Hispanic white controls. The DPB1*01:01 allele appears to be less common in the NCCLS non-Hispanic white controls (4.5%)

compared with the UKCCS controls (8.0%) but is similar to those reported in other United States white populations (4.5%-6.0%).³²⁻³⁴ Higher frequencies for DPB1*01:01 (more consistent with the UKCCS observations) have been reported in studies of whites residing in Europe, including populations of Germany (8.0%), France (6.6%), and Great Britain (7.5%).³⁵⁻³⁷ Similarly, the allele frequency of DPB1*02:01, one of the alleles driving the DP2

Table 3. Association between HLA-DPA1 and DPB1 phenotypes and risk of childhood ALL, NCCLS, 1995-2008

	Non-Hispanic white					Hispanic					P _{hom}	
	Cases		Controls		OR (95% CI)†	Cases		Controls		OR (95% CI)†		Pooled OR (95% CI)†
	n*	%	n*	%		n*	%	n*	%			
DPA1 phenotypes‡												
01:03	251	94.7	391	95.8	0.76 (0.37-1.57)	301	94.4	418	95.4	0.80 (0.41-1.54)	0.78 (0.48-1.27)	.918
02:01	73	27.5	101	24.8	1.19 (0.83-1.69)	86	27.0	125	28.5	0.93 (0.67-1.28)	1.04 (0.82-1.31)	.315
02:02	18	6.8	29	7.1	0.94 (0.51-1.72)	44	13.8	44	10.1	1.43 (0.92-2.24)	1.24 (0.86-1.77)	.275
n	265		408			319		438				
DPB1 phenotypes‡												
01:01	31	11.7	35	8.6	1.44 (0.86-2.41)	35	10.9	35	8.0	1.42 (0.87-2.33)	1.43 (1.01-2.04)	.969
02:01	64	24.2	108	26.5	0.88 (0.62-1.26)	72	22.5	92	20.9	1.10 (0.78-1.56)	0.99 (0.77-1.27)	.378
03:01	52	19.6	79	19.4	1.01 (0.68-1.49)	34	10.6	56	12.7	0.81 (0.52-1.28)	0.92 (0.68-1.23)	.469
04:01	161	60.8	269	65.9	0.80 (0.58-1.10)	133	41.6	193	43.9	0.91 (0.68-1.22)	0.84 (0.68-1.05)	.560
04:02	58	21.9	89	21.8	1.00 (0.69-1.45)	174	54.4	242	55.0	0.97 (0.73-1.30)	1.00 (0.79-1.26)	.895
05:01	10	3.8	17	4.2	0.89 (0.40-1.98)	20	6.3	23	5.2	1.21 (0.65-2.25)	1.08 (0.66-1.75)	.552
11:01	10	3.8	21	5.1	0.74 (0.34-1.59)	11	3.4	17	3.9	0.89 (0.41-1.92)	0.81 (0.47-1.40)	.740
14:01	6	2.3	5	1.2	1.90 (0.57-6.34)	20	6.3	24	5.5	1.16 (0.63-2.13)	1.27 (0.73-2.18)	.474
17:01	10	3.8	12	2.9	1.33 (0.56-3.13)	20	6.3	24	5.5	1.16 (0.63-2.14)	1.18 (0.72-1.95)	.800
n	265		408			320		440				

*Number and percent of subjects typing for each allele.

†ORs and 95% CIs were calculated using logistic regression adjusting for child's age and sex. The pooled analysis additionally adjusted for race/ethnicity, and the P_{hom} value for the χ^2 test of homogeneity in risk estimates between non-Hispanic whites and Hispanics are shown.

‡HLA-DPA1 and DPB1 alleles < 0.05 phenotype frequency not shown. Includes DPA1, 01:04, 03:01, and 04:01; and DPB1, 02:02, 102:01, 06:01, 09:01, 10:01, 13:01, 15:01, 16:01, 18:01, 26:01, 29:01, 34:01, 40:01, 51:01, 54:01, and 60:01.

Table 4. Association between HLA-DPB1 supertypes (carrier/noncarrier) and risk of childhood ALL overall and by subtype, NCCLS, 1995-2008

Supertype†	% carriers of DPB1 supertype						OR (95% CI)*				
	Total ALL	cALL	Hyperdiploid	TEL-AML1	Normal‡	Controls	Total ALL	cALL	Hyperdiploid	TEL-AML1	Normal karyotype‡
DP2 (GEG)	24.8	24.1	22.1	24.7	22.9	24.5	1.03 (0.80-1.31)	0.97 (0.70-1.34)	0.88 (0.59-1.31)	0.99 (0.59-1.66)	0.93 (0.52-1.66)
DP4 (GKG)	79.2	79.4	77.9	76.4	74.3	80.9	0.88 (0.68-1.15)	0.91 (0.64-1.29)	0.82 (0.55-1.23)	0.78 (0.46-1.32)	0.67 (0.38-1.18)
DP6 (LED)	14.2	14.8	15.1	15.7	22.9	15.0	0.94 (0.70-1.27)	1.17 (0.79-1.74)	1.03 (0.65-1.64)	1.13 (0.61-2.08)	1.68 (0.93-3.04)
DP3 (LKD)	19.5	17.9	19.2	20.2	21.4	20.2	0.96 (0.73-1.25)	0.80 (0.56-1.15)	0.94 (0.62-1.42)	0.92 (0.53-1.60)	1.11 (0.61-2.01)
DP8 (GED)	3.4	3.1	2.9	2.3	2.9	2.2	1.52 (0.80-2.88)	1.09 (0.47-2.51)	1.21 (0.44-3.33)	0.90 (0.20-3.99)	1.29 (0.29-5.66)
DP1 (GKD)	16.2	16.2	20.9	19.1	14.3	12.7	1.33 (0.99-1.79)	1.33 (0.90-1.96)	1.83 (1.20-2.78)	1.64 (0.92-2.91)	1.14 (0.26-2.29)
n subjects	585	291	172	89	70	848					

cALL indicates common ALL; E, glutamic acid; G, glycine; L, leucine; D, aspartic acid; and K, lysine.

*ORs and 95% CIs were calculated using logistic regression adjusting for child's age, sex, and race/ethnicity. Importance of adjustment by other supertypes was evaluated based on the criteria of a > 10% change in OR of the supertype in question. This adjustment did not remarkably influence the risk estimate for all supertypes.

†Supertypes based on peptide-binding pocket residue sequences: pocket 6, position 11; pocket 4, position 69; and pocket 1, position 84.

‡Normal cytogenetics results with a minimum of 20 metaphases counted.

association observed in the UKCCS, appears to be lower among UKCCS controls (6.0%) versus other studies conducted in United States and European whites (10.0%-15.6%).³²⁻³⁷ Variation in DPB1 allele frequencies between populations is an indication of the potential diversity in underlying genetic structure and that the populations may have evolved under different selective pressures. In adaptive immunity, the HLA class II molecule selectively binds to the antigen and together with the T-cell receptor, forms a specialized complex that activates a cascade of intercellular signals that are specifically designed to elicit a productive immune response.³⁸ The involvement of the HLA molecule as a functional unit in disease susceptibility is therefore dependent on the presence of the immunodominant antigenic peptide.^{17,39} Accordingly, heterogeneity by geographic region in the prevalence of certain infectious

agents or other relevant exposures may potentially contribute to the variability observed.

Studies have shown DPB1*01:01, the principal allele contributing to the DP1 supertype, to be associated with an increased risk of other disease outcomes as well. Earlier case-control studies have found increased risks of both dermatitis herpetiformis and celiac disease associated with DPB1*01:01 in white populations, including those residing in the United Kingdom.^{40,41} More recently, a large study found that, along with other HLA class II alleles, DPB1*01:01 is associated with an increased susceptibility to sarcoidosis in a United States black population.⁴²

The evaluation of interaction between DP1 supertypes and proxy measures in the risk of childhood ALL showed that the increased risk associated with DP1 is only found in children

Table 5. Multivariable analysis of DP1 supertype, proxy measures for early life immune modulation (older sibling, daycare attendance, ear infection in infancy, and breastfeeding), and risk of childhood ALL, NCCLS, 1995-2008

	Non-Hispanic white			Hispanic			Pooled
	Cases (%)	Controls (%)	OR (95% CI)*	Cases (%)	Controls (%)	OR (95% CI)*	OR (95% CI)*
Total†	231	364		276	398		
Genetic factor							
DP1 supertype							
No copies	191 (82.7)	317 (87.1)	1.00	227 (82.3)	345 (86.7)	1.00	1.00
At least 1 copy	40 (17.3)	47 (12.9)	1.32 (0.82-2.15)	49 (17.8)	53 (13.3)	1.50 (0.97-2.33)	1.41 (1.02-1.94)
Proxy exposure measure							
Older sibling							
No	113 (48.9)	134 (36.8)	1.00	109 (39.5)	143 (35.9)	1.00	1.00
Yes	118 (51.1)	230 (63.2)	0.57 (0.40-0.82)	167 (60.5)	255 (64.1)	0.85 (0.59-1.22)	0.71 (0.55-0.92)
Daycare child-hours by age 6 months‡							
< 2000	212 (91.8)	313 (86.0)	1.00	265 (96.0)	381 (95.7)	1.00	1.00
≥ 2000	19 (8.2)	51 (14.0)	0.46 (0.25-0.85)	11 (4.0)	17 (4.3)	1.19 (0.53-2.69)	0.67 (0.41-1.07)
Ear infection during infancy							
Not in infancy	136 (58.9)	198 (54.4)	1.00	175 (63.4)	232 (58.3)	1.00	1.00
At age < 6 mo only	8 (3.5)	23 (6.3)	0.48 (0.20-1.15)	18 (6.5)	41 (10.3)	0.52 (0.28-0.95)	0.51 (0.31-0.83)
At age 6-11 mo only	42 (18.2)	81 (22.3)	0.83 (0.52-1.30)	42 (15.2)	76 (19.1)	0.67 (0.43-1.05)	0.76 (0.56-1.04)
Both time periods	45 (19.5)	62 (17.0)	1.36 (0.83-2.22)	41 (14.9)	49 (12.3)	1.11 (0.68-1.78)	1.20 (0.85-1.68)
Breastfeeding							
No	31 (13.4)	37 (10.2)	1.00	49 (17.8)	60 (15.1)	1.00	1.00
Yes	200 (86.6)	327 (89.8)	0.97 (0.55-1.69)	227 (82.3)	338 (84.9)	0.89 (0.57-1.38)	0.88 (0.62-1.23)

*ORs and 95% CIs were calculated using a single additive logistic regression model, including all variables listed in the table and additionally adjusting for age, sex, maternal age, maternal education, and annual household income. The pooled analysis was additionally adjusted for race/ethnicity (non-Hispanic white, Hispanic white, and Hispanic other).

†Total numbers of cases and controls with complete data for DP1 supertype, all proxy exposure measures indicated in the table, and covariates, including age, sex, maternal age, maternal education, and annual household income.

‡The child-hours in each daycare setting were summed to obtain the total child-hours for each child. Child-hours = (number of months attending the daycare) × (mean hours per week at this daycare) × (number of other children at this daycare) × (4.35 weeks per month).

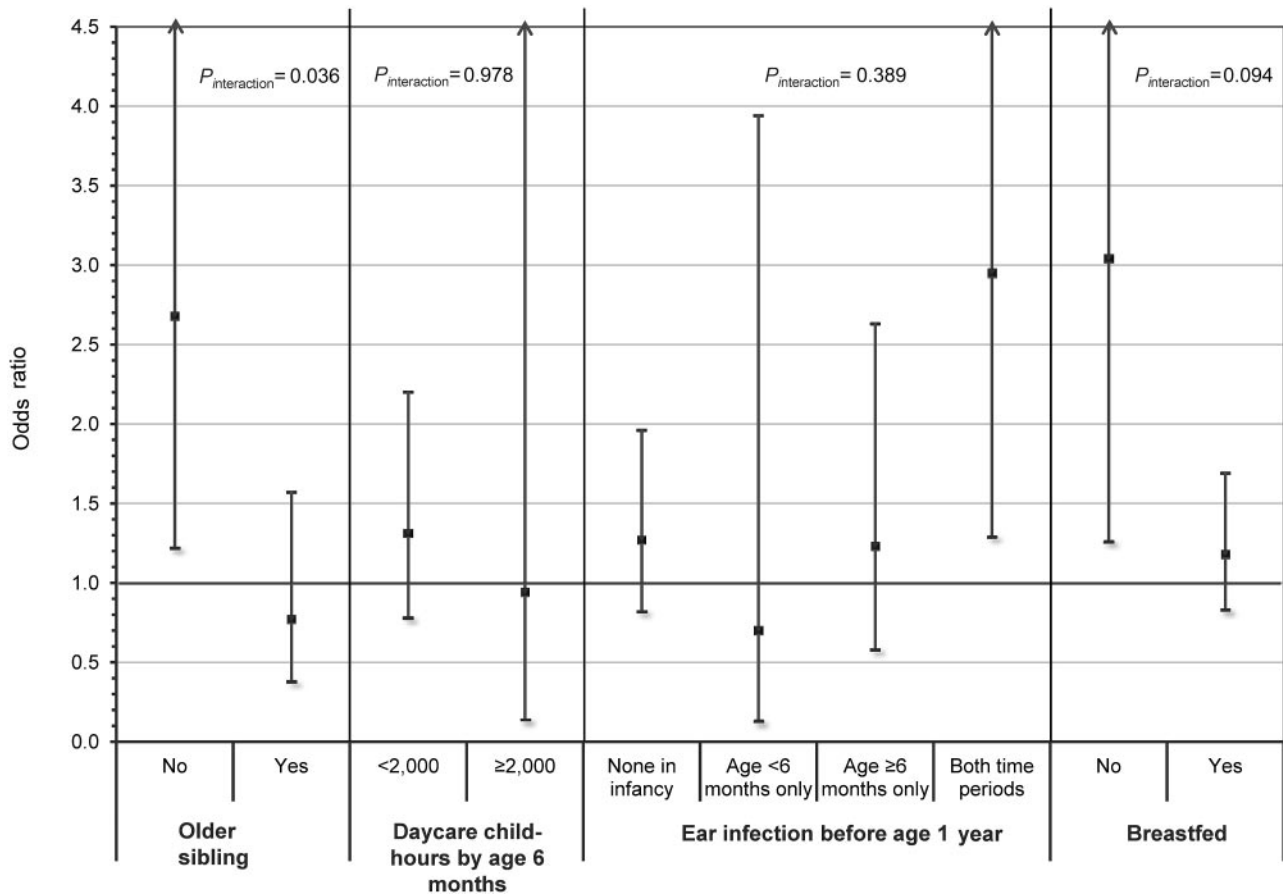


Figure 1. Summary plot showing the risk of childhood ALL associated with carriers of the DP1 supertype stratified by levels of select proxy exposure measures for early life immune modulation. ORs and 95% CIs represent the risk of childhood ALL associated with carriers of the DP1 supertype and were calculated using multivariable logistic regression adjusting for age, sex, maternal age, maternal education, annual household income, phase of study enrolled, and all other proxy measures presented. Two-way interactions between DP1 supertype carrier status and each of the 4 proxy measures separately were evaluated using a similar model, which additionally included an interaction term representing the product of the DP1 supertype and proxy exposure measure. Results for the 2 social contact variables, older sibling and daycare child-hours, were based on an analysis conducted in non-Hispanic white children only. The other 2 measures, ear infection and breastfeeding, were analyzed among all subjects because no evidence of heterogeneity was observed by race/ethnicity. Two-sided *P* values (*P*_{interaction}) of < .10 were considered evidence of a significantly different effect of the DP1 supertype on ALL risk between levels of the proxy exposure measure.

without an older sibling. A complementary finding was also seen for interaction with breastfeeding, where an increased risk of the DP1 supertype was found only among non-breastfed children. The dichotomous variable indicating presence or absence of an older sibling was used as a proxy measure of exposure to common infections early in life, based on the assumption that contact with older siblings would expose the index child to infections at a very young age. In the context of the “delayed infection” hypothesis, this exposure would, in turn, contribute to immune modulation, a critical component of normal immunologic development.^{5,43} The immune-modulatory effects of breastfeeding are also well documented.⁴⁴ Findings of the current analysis, demonstrated by interactions with 2 separate proxy measures, suggest that T-cell recognition of the DP1-peptide complex in the context of an insufficiently modulated immune system may trigger an unfavorable immune response that contributes to leukemogenesis.

Several lines of evidence underscore the importance of immunomodulatory exposures early in life and the effects of exposures on the subsequent development of certain health outcomes, particularly allergic disorders and autoimmune diseases.⁴⁵ The neonatal immune system is significantly down-regulated and proceeds along a series of postnatal developmental stages, most of which are mediated by environmental exposures during the first year of life.

Among the notable characteristics of the neonatal immune system are the presence of high amounts of regulatory T cells and the down-regulated CD4⁺ T helper 1 (Th1) and T helper 2 (Th2) activities that display a strong skewing toward Th2 responses.⁴⁶ A critical element of immune maturation is to increase the functional capacity of Th1, which is necessary for effective cell-mediated immunity, antitumor defense, and progression toward a normal immune balance. Various endogenous conditions and environmental exposures, particularly microbial agents as well as breastfeeding, have been shown to promote Th1 capacity and contribute to the process of Th1/Th2 balance during immune maturation.⁴⁶ Capacity for T regulatory cell activity is also stimulated by early life microbial exposures, which have the ability to enhance the regulatory networks that are responsible for down-regulating immune responses after infection is controlled.⁴⁷ Taken together, extensive research indicates that proper immune modulation, largely through exposures to infections early in life, may have a significant impact on later life responses to immunologic stimuli.

As described in previous NCCLS publications,^{21,22} the main effect of birth order and daycare attendance on childhood ALL risk was observed only in non-Hispanic whites, and not in Hispanic children. Observed differences in family structure and daycare

attendance patterns between these 2 populations suggested that this differential effect may be the result of heterogeneity in the degree to which these factors serve as appropriate proxies for early life exposure to common infections. Accordingly, evidence of interaction using these surrogate measures was not found in Hispanic children, but the main effect of DP1 does appear to be consistent in both non-Hispanic white and Hispanic children. From a social contacts perspective (indirect measures of infectious exposure), a more refined measure of exposure to infections among Hispanics may be the total number of people living in the household at the time of child's birth, including nonsibling children and adults. However, such measures are not available in the NCCLS. Nevertheless, support for a role of infections in the Hispanic population also exists. History of ear infections by the age of 6 months, a direct measure of exposure, was found to be associated with a reduced risk in both Hispanics and non-Hispanic whites.²² With respect to interaction with DP1 in Hispanics, an effect was found when considering history of breastfeeding as the immune-modulatory exposure. The lack of strong evidence for DP1 interactions with daycare child-hours and ear infection by age 6 months may be the result of limited statistical power. Stratified analyses did show, however, risk estimates pointing in the directions consistent with DP1 interaction effects found for older sibling and breastfeeding.

Our study had sufficient power to detect moderate to strong *HLA-DP* associations with ALL, but associations within certain subgroup analyses (eg, molecular subtypes) and of rare alleles and supertypes may not have been detectable with the current sample size. Assuming a phenotype frequency of ~ 0.15 (frequency of carriers of the allele or supertype) and a type I error rate of 0.05, the current analysis had adequate power (> 70%) to detect associations for risk estimates as low as 1.4. Assuming a causal relationship, the functionally driven classification of HLA alleles into supertypes would be expected to improve chances of detecting associations. An increased probability of false-positive findings resulting from multiple testing is a possibility. However, we focused on a previously reported candidate locus, and we approached this unique analysis of gene-environment interactions within the context of a strong a priori hypothesis stemming from an infection-related premise that has been studied over the past 2 decades.¹ Although multiallelic, this investigation included primarily 1 genetic locus (*HLA-DPA1* and *DPB1* are in strong linkage disequilibrium) and the proxy exposure measures were carefully selected based on findings from a previous publication.²² Additional confidence in the findings of the current analysis comes from observing a consistent main effect for the *DP1* supertype in both non-Hispanic white and Hispanic children, and consistent evidence of interaction of DP1 observed with 2 different proxy measures related to immune modulation.

The effect of population stratification is probably minimal in the NCCLS because of the careful and detailed account of race and ethnicity obtained from the subjects and statistical adjustment. In addition, the extent of genetic admixture was assessed using a series of 80 ancestry informative markers for a subset of the cases and controls included in this study.⁴⁸ Estimates of genetic ancestry (percent of European, Amerindian, and African ancestry) were determined with these ancestry informative markers, and comparison of these estimates between cases and controls showed no significant differences (data not shown).⁴⁸

Potential selection biases resulting in systematic differences between cases and controls should be considered. In the NCCLS, population-based controls are selected from the statewide birth

registry among all children born within the study region. A previous methodologic evaluation has shown that controls enrolled in the NCCLS are comparable with "ideal" controls who could have been enrolled under the optimal circumstances (ie, no difficulty in tracing, no refusal of participation).²⁴ In the current analysis, cases and controls appeared to differ in maternal age in non-Hispanic whites, and annual household income in both race/ethnicities. The extent to which these differences are a result of systematic biases is unknown, but the influence of these and other factors was evaluated and addressed accordingly in the multivariable analyses. However, uncontrolled or residual confounding on risk estimates is a possibility.

Although there are numerous studies supporting the importance of early life immune modulation in childhood ALL based on proxy measures, unresolved inconsistencies still exist. The current study demonstrates that concurrent evaluation of host genetic susceptibility factors, together with early life exposures, may be important. Our findings support the conclusion that *HLA-DPB1* genetic variation combined with an insufficiently modulated immune system may trigger an adverse immune response that contributes to childhood ALL risk.

Acknowledgments

The authors thank the clinical collaborators and participating hospitals: University of California Davis, University of California San Francisco, Children's Hospital of Central California, Lucile Packard Children's Hospital, Children's Hospital and Research Center Oakland, Kaiser Permanente Roseville, Kaiser Permanente Santa Clara, Kaiser Permanente San Francisco, and Kaiser Permanente Oakland. The authors also thank the HLA Laboratory staff at the Children's Hospital Research Center Oakland and staff at Roche Molecular Systems for their assistance with the genotyping assays, the Northern California Childhood Leukemia Study staff, the Survey Research Center, and the participating children and their families for their important contributions to this study.

This work was supported by the United States National Institute of Environmental Health Sciences (R01ES09137 and P42ES0470518), the National Cancer Institute (R03CA125823), and the United Kingdom Children with Cancer Foundation (2005/027, 2005/028, 2006/051, and 2006/052).

Authorship

Contribution: K.Y.U., A.P.C., C.M., L.F.B., E.T., and P.A.B. conceived and designed the study; K.Y.U., A.P.C., C.M., X.M., J.L.W., J.K.W., G.V.D., P.M., H.A.E., E.T., and P.A.B. assisted in assembling the data; K.Y.U., A.P.C., C.M., S.S., L.F.B., X.M., J.L.W., J.K.W., M.T., P.B., E.T., and P.A.B. analyzed and interpreted the data; and all authors critically reviewed and edited the manuscript for intellectual content and gave final approval of the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Kevin Y. Urayama, University of California, Berkeley, School of Public Health, 1995 University Ave, Suite 460, Berkeley, CA 94704; e-mail: kurayama@luke.or.jp.

References

- McNally RJ, Eden TO. An infectious aetiology for childhood acute leukaemia: a review of the evidence. *Br J Haematol*. 2004;127(3):243-263.
- Penrose M. Cat leukemia. *Br Med J*. 1970;1(5698):755.
- Boxus M, Willems L. Mechanisms of HTLV-1 persistence and transformation. *Br J Cancer*. 2009;101(9):1497-1501.
- O'Connor SM, Boneva RS. Infectious etiologies of childhood leukemia: plausibility and challenges to proof. *Environ Health Perspect*. 2007;115(1):146-150.
- Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer*. 2006;6(3):193-203.
- Marsh SG, Parham P, Barber LD. *The HLA Facts Book*. London, United Kingdom: Academic Press; 2000.
- Shiina T, Inoko H, Kulski JK. An update of the HLA genomic region, locus information and disease associations: 2004. *Tissue Antigens*. 2004;64(6):631-649.
- Robinson J, Waller MJ, Parham P, et al. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res*. 2003;31(1):311-314.
- Thorsby E. HLA associated diseases. *Hum Immunol*. 1997;53(1):1-11.
- Diepstra A, Niens M, Vellenga E, et al. Association with HLA class I in Epstein-Barr-virus-positive and with HLA class III in Epstein-Barr-virus-negative Hodgkin's lymphoma. *Lancet*. 2005;365(9478):2216-2224.
- Urayama KY, Jarrett RF, Hjalgrim H, et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. *J Natl Cancer Inst*. 2012;104(3):240-253.
- Ooi EE, Ren EC, Chan SH. Association between microsatellites within the human MHC and nasopharyngeal carcinoma. *Int J Cancer*. 1997;74(2):229-232.
- Azuma T, Ito S, Sato F, et al. The role of the HLA-DQA1 gene in resistance to atrophic gastritis and gastric adenocarcinoma induced by *Helicobacter pylori* infection. *Cancer*. 1998;82(6):1013-1018.
- Arends MJ, Buckley CH, Wells M. Aetiology, pathogenesis, and pathology of cervical neoplasia. *J Clin Pathol*. 1998;51(2):96-103.
- Dorak MT, Lawson T, Machulla HK, Darke C, Mills KI, Burnett AK. Unravelling an HLA-DR association in childhood acute lymphoblastic leukemia. *Blood*. 1999;94(2):694-700.
- Dorak MT, Ogun FS, Yalman N, et al. A male-specific increase in the HLA-DRB4 (DR53) frequency in high-risk and relapsed childhood ALL. *Leuk Res*. 2002;26(7):651-656.
- Klitz W, Gragert L, Trachtenberg E. Spectrum of HLA associations: the case of medically refractory pediatric acute lymphoblastic leukemia. *Immunogenetics*. 2012;64(6):409-419.
- Taylor GM, Dearden S, Ravetto P, et al. Genetic susceptibility to childhood common acute lymphoblastic leukaemia is associated with polymorphic peptide-binding pocket profiles in HLA-DPB1*0201. *Hum Mol Genet*. 2002;11(14):1585-1597.
- Taylor GM, Hussain A, Lightfoot TJ, Birch JM, Eden TO, Greaves MF. HLA-associated susceptibility to childhood B-cell precursor ALL: definition and role of HLA-DPB1 supertypes. *Br J Cancer*. 2008;98(6):1125-1131.
- Taylor M, Harrison C, Eden T, et al. HLA-DPB1 supertype-associated protection from childhood leukaemia: relationship to leukaemia karyotype and implications for prevention. *Cancer Immunol Immunother*. 2008;57(1):53-61.
- 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(8):1928-1934.
- Urayama KY, Ma X, Selvin S, et al. Early life exposure to infections and risk of childhood acute lymphoblastic leukemia. *Int J Cancer*. 2011;128(7):1632-1643.
- Bartley K, Metayer C, Selvin S, Ducore J, Buffler P. Diagnostic X-rays and risk of childhood leukemia. *Int J Epidemiol*. 2010;39(6):1628-1637.
- 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(10):915-921.
- Alexander FE, Patheal SL, Biondi A, et al. Transplacental chemical exposure and risk of infant leukemia with MLL gene fusion. *Cancer Res*. 2001;61(6):2542-2546.
- 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(8):1686-1690.
- Cano P, Klitz W, Mack SJ, et al. Common and well-documented HLA alleles: report of the Ad-Hoc Committee of the American Society for Histocompatibility and Immunogenetics. *Hum Immunol*. 2007;68(5):392-417.
- Aldrich MC, Zhang L, Wiemels JL, et al. Cytogenetics of Hispanic and White children with acute lymphoblastic leukemia in California. *Cancer Epidemiol Biomarkers Prev*. 2006;15(3):578-581.
- Castelli FA, Buhot C, Sanson A, et al. HLA-DP4, the most frequent HLA II molecule, defines a new supertype of peptide-binding specificity. *J Immunol*. 2002;169(12):6928-6934.
- Diaz G, Amicosante M, Jaraquemada D, et al. Functional analysis of HLA-DP polymorphism: a crucial role for DPbeta residues 9, 11, 35, 55, 56, 69 and 84-87 in T cell allorecognition and peptide binding. *Int Immunol*. 2003;15(5):565-576.
- Lancaster AK, Single RM, Solberg OD, Nelson MP, Thomson G. PyPop update: a software pipeline for large-scale multilocus population genomics. *Tissue Antigens*. 2007;69(Suppl 1):192-197.
- Begovich AB, Helmuth RC, Oksenberg JR, et al. HLA-DP beta and susceptibility to multiple sclerosis: an analysis of Caucasoid and Japanese patient populations. *Hum Immunol*. 1990;28(4):365-372.
- Moonsamy PV, Klitz W, Tilanus MG, Begovich AB. Genetic variability and linkage disequilibrium within the DP region in the CEPH families. *Hum Immunol*. 1997;58(2):112-121.
- Noble JA, Valdes AM, Thomson G, Erlich HA. The HLA class II locus DPB1 can influence susceptibility to type 1 diabetes. *Diabetes*. 2000;49(1):121-125.
- al-Daccak R, Wang FQ, Theophile D, Lethielleux P, Colombani J, Loiseau P. Gene polymorphism of HLA-DPB1 and DPA1 loci in Caucasoid population: frequencies and DPB1-DPA1 associations. *Hum Immunol*. 1991;31(4):277-285.
- Eiermann TH, Fakler J, Muller CR, Ballas M, Goldmann SF. HLA-DPB1 oligonucleotide typing of a southwest German Caucasian population. *Tissue Antigens*. 1991;38(5):193-198.
- Sage DA, Evans PR, Howell WM. HLA DPA1-DPB1 linkage disequilibrium in the British Caucasian population. *Tissue Antigens*. 1994;44(5):335-338.
- Madden DR. The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol*. 1995;13:587-622.
- Yewdell JW, Bennink JR. Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu Rev Immunol*. 1999;17:51-88.
- Hall MA, Lanchbury JS, Ciclitira PJ. HLA class II region genes and susceptibility to dermatitis herpetiformis: DPB1 and TAP2 associations are secondary to those of the DQ subregion. *Eur J Immunogenet*. 1996;23(4):285-296.
- Rosenberg WM, Wordsworth BP, Jewell DP, Bell JI. A locus telomeric to HLA-DPB1 encodes susceptibility to coeliac disease. *Immunogenetics*. 1989;30(4):307-310.
- Rossmann MD, Thompson B, Frederick M, et al. HLA-DRB1*1101: a significant risk factor for sarcoidosis in blacks and whites. *Am J Hum Genet*. 2003;73(4):720-735.
- M'Rabet L, Vos AP, Boehm G, Garssen J. Breastfeeding and its role in early development of the immune system in infants: consequences for health later in life. *J Nutr*. 2008;138(9):1782S-1790S.
- Labbok MH, Clark D, Goldman AS. Breastfeeding: maintaining an irreplaceable immunological resource. *Nat Rev Immunol*. 2004;4(7):565-572.
- Wills-Karp M, Santeliz J, Karp CL. The germless theory of allergic disease: revisiting the hygiene hypothesis. *Nat Rev Immunol*. 2001;1(1):69-75.
- Dietert RR, Piepenbrink MS. The managed immune system: protecting the womb to delay the tomb. *Hum Exp Toxicol*. 2008;27(2):129-134.
- Wing K, Fehervari Z, Sakaguchi S. Emerging possibilities in the development and function of regulatory T cells. *Int Immunol*. 2006;18(7):991-1000.
- Chokkalingam AP, Aldrich MC, Bartley K, et al. Matching on race and ethnicity in case-control studies as a means of control for population stratification. *Epidemiology*. 2011;1:101.