ABSTRACT
Several studies have implicated HLA in non-Hodgkin lymphoma (NHL) subtype etiology. However, NHL patients indicated for stem cell transplants are underrepresented in these reports. We therefore evaluated the association between HLA and NHL subtypes among a transplant-indicated population. One thousand three hundred and sixty-six NHL patients HLA-typed and indicated for transplant at the City of Hope National Medical Center (Duarte, CA) were compared to 10,271 prospective donors. Odds ratios and 95% confidence intervals were calculated for HLA haplotype and alleles, adjusted for sex and age. The HLA-A\(^*\)0201\,-\,C\(^*\)0602\,-\,B\(^*\)1302\,-\,DRB1\(^*\)0701\,-\,DQB1\(^*\)0201 haplotype was significantly associated with follicular lymphoma (FL) risk among Caucasians. Several haplotypes were associated with diffuse large B-cell lymphoma (DLBCL) risk among Caucasians, including the previously implicated DLBCL risk loci, HLA-B\(^*\)0801. The HLA-A\(^*\)0101 allele was also observed to be associated with mantle cell lymphoma (MCL) risk. Our results support the association between previously reported susceptibility loci and FL and suggest potentially new DLBCL and MCL risk loci.

ARTICLE HISTORY
Received 13 November 2018
Revised 10 April 2019
Accepted 4 May 2019

KEYWORDS
Epidemiology; lymphoma; HLA; bone marrow transplant

Introduction
The human leukocyte antigens (HLAs) are a class of genes encoded in the major histocompatibility complex (MHC) on chromosome 6 that are critical for human immune response. Among the most polymorphic human genes, HLA alleles and haplotypes are increasingly implicated in the etiologies and outcomes of immune conditions and hematopoietic malignancies, including Hodgkin lymphoma [1,2], non-Hodgkin lymphomas (NHLs) [3–11], and leukemias [12–16]. Recent genome-wide association studies (GWASs) of NHL have now confirmed the role for HLA in four major NHL subtypes diffuse large B-cell lymphoma (DLBCL) [17,18], follicular lymphoma (FL) etiology [19–21], chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL) [22–24], and marginal zone lymphomas (MZLs) [25].

The studies in which associations between HLA loci and hematopoietic malignancies have largely been conducted are epidemiologic and clinical studies, which depend on voluntary participation among the catchment population. As a result, there are well-documented survival biases among these studies where participation is typically higher among patients who have less severe disease [26–28]. Most epidemiologic and association studies of NHL thus under-represent patients for whom bone marrow transplants are conducted and/or indicated. We posited that identifying risk alleles in this population may provide insights into more severe disease etiologies. Our efforts here build upon previously published results, including efforts based on the National Marrow Donor Program (NMDP) which implicated HLA alleles and haplotypes in the risk of chronic lymphocytic leukemia (CLL) in their
transplant-based population [11], and GWAS studies which have evaluated HLA allele associations based on imputed data among epidemiologic and clinical studies [19,24–26]. Although the NMDP effort similarly used HLA typing data to identify loci for a transplant population of CLL, they were unable to evaluate other NHL subtypes due to the lack of subtype information within the NMDP. Here, we link pathology records, HLA typing, and patient data from a single institution’s (City of Hope (COH)) transplant database to evaluate the associations between HLA haplotypes and alleles with NHL subtypes. We note that for the purposes of this evaluation, all patients indicated for a transplant (regardless of whether they had a transplant) were included in this analysis.

Materials and methods

Study population

Our patient population comprised individuals diagnosed with NHL at the COH between January 1995 and October 2015 and indicated for a transplant. To be included in the analysis, patients were identified with the COH pathology database and medical records and then linked to the COH HLA database. Prospective donors who were typed for HLA at COH during the same time period served as the comparison/control group. A total of 1977 patients with histologically confirmed NHL per Interlymph/WHO classification [29,30] and 15,020 donors were identified between January 1995 and October 2015.

NHL definition

NHL subtypes were defined according to the WHO algorithm. Briefly, DLBCL was defined as ICD-O-3 codes 9678, 9679, 9680, and 9684; FL was defined as 9690, 9691, 9695, and 9698; CLL/SLL was defined as 9823 and 9678, 9679, 9680, and 9684; FL was defined as 9690, 9691, 9695, and 9698; mantle cell was defined as 9673 [30]. As expected, NHL patients indicated for transplant had more advanced disease than the general COH population (Supplementary Table 1). Of the patients indicated for transplant and typed for HLA, 30% did not receive a transplant; and of those who did, the majority received an autologous transplant (Supplementary Table 2).

To evaluate the comparability of the COH HLA donor population to that of the general population, we compared HLA allele frequencies (minimum 1% frequency) to previously published frequencies [31] and noted the similarity, including with the NMDP population (Supplementary Table 3).

HLA typing and imputation

HLA typing from 1995 to 2000 in the COH histopathology laboratory comprised a combination of serological, low resolution, and high resolution HLA typing. All HLA typing after May 2000 was conducted with low or high resolution molecular typing.

Briefly, for low to intermediate resolution HLA typing, a reverse sequence-specific oligonucleotide (rSSO) probe hybridization method was utilized. Sequence-specific oligonucleotide (SSO) probes bind to fluorescently coded microspheres to identify alleles encoded by the sample DNA. The target DNA is amplified by polymerase chain reaction (PCR) and then hybridized to probes that are specific to the HLA region. Signal detection is captured on a Luminex® 100 analyzer. Based on the Luminex® xMAP® Technology, the system enables multiplexing of up to 100 analytes in a single microplate well. Each well represents one sample and one HLA locus. Analysis is performed using the HLA Fusion software from One Lambda. The program matches the positive and negative signals from each probe to a reaction pattern that defines a specific HLA allele group and assigns HLA type.

For high resolution HLA typing, two different methods were used. The first method utilized was the Sanger sequence-based typing method (SBT, SBT Resolver, Conexio Genomics, Fremantle, Australia) whereby the target sequence is amplified, followed by treatment with ExoSAP-IT® to remove unincorporated primers and deoxynucleotide triphosphates (dNTPs). The amplicon is then used as template for direct automated fluorescent DNA sequencing using customized sequencing primers and the BigDye® Terminator sequencing chemistry (Applied Biosystems™, Life Technologies™, Foster City, CA). The extension products are purified according to the ethanol precipitation method and denatured using HI-DITM formamide (Applied Biosystems™, Life Technologies™, Foster City, CA) before separation and detection on an automated fluorescent DNA sequencer. Results are imported and analyzed using Assign™ sequence analysis software from Conexio Genomics (Fremantle, Australia). The second method used to obtain high resolution HLA typing was next generation sequencing (NGS) from Scisco Genetics (Seattle, WA). The method uses the Illumina MiSeq system, which is based on a sequencing-by-synthesis approach utilizing fluorescently labeled reversible terminator nucleotides. After assay specific amplification, samples are tagged with unique indexes and pooled together and applied to the MiSeq instrument, where they are amplified as individual clusters and sequenced using universal sequencing primers.
The result is several million reads that can be analyzed using the GeMS HLA software (Scisco Genetics, Seattle, WA) to report unambiguous HLA allele types for several individual samples simultaneously.

To maximize four-digit HLA data and haplotype information for analysis, race and HLA typing data were used to impute high resolution alleles and haplotypes when possible. Briefly, given a minimum of low resolution, 2 digit A-B-DR alleles and race, the possible list of high resolution A-B-C-DR-DQ haplotypes is enumerated though an expectation-maximization algorithm [32]. The probability of a given haplotype is calculated based on registry level data [33–35]. We generated five unambiguous imputed HLA datasets based on these probabilities for analysis.

**Final analytic population**

Of the 1977 NHL patients indicated for transplant and 15,020 controls; age, race, and A-B-DR HLA genotyping were available for 1366 patients and 10,271 donors, and who comprised the final analytic population. Of NHL patients with imputed HLA haplotypes, there were 354 DLBCLs, 263 FLs, and 173 mantle cell lymphomas (MCLs).

**Statistical analysis**

We conducted haplotype- and allele-level HLA analyses for four NHL subtypes (DLBCL, FL, CLL/SLL, and MCL). We calculated the risk (odds ratios) with each NHL subtype in logistic regression models, adjusting for age and sex [36]. Geometric mean odds ratios were calculated after adjustment with the false discovery rate (FDR) method at a threshold of 0.05 [37]. Given the high linkage disequilibrium (LD) present in the HLA region and co-dominant fashion in which HLA alleles are expressed, results are presented by haplotype. Due to the high number of haplotypes, we restricted our analysis to haplotypes with at least a frequency of one percent among donors. We further present results for the individual HLA loci affiliated with the HLA haplotypes to determine whether any allele was the dominant association. We also evaluated the association between HLA homozygosity for HLA class I (A, B, C) and Class II (DR, DB, DQ) alleles, whereby ORs and 95% CIs were calculated as estimates of NHL risk with heterozygotes as the referent category, adjusted for sex, age, study, and race. Estimates for each additional number of homozygous loci were calculated as p-trend. For all results, results stratified by race are presented. All analyses were conducted in SAS 9.4 (SAS Institute, Cary, NC).

**Results**

Of the 1366 NHL patients included in this analysis, there were a total of 938 Caucasian patients, 237 Hispanic patients, 138 Asian patients, and 54 African American patients (Table 1). The most common NHL subtype was DLBCL (26%), followed by FL (19%), mantle cell (13%), and CLL/SLL (11%). Overall, there were more male patients (63.5%) than female patients. Among the donors, 6801 were identified as Caucasian, 2076 as Hispanic, 1029 as Asian, and 365 as African American.

We present below, results for HLA associations with DLBCL, FL, and MCL. As COH participants are in the

|                  | Caucasian | Hispanic | Asian | African American |
|------------------|-----------|----------|-------|------------------|
| **B-cell NHL**   |           |          |       |                  |
| N                | 780       | 229      | 191   | 124              |
| Age (median)     | 49        | 50       | 47    | 51               |
| **DLBCL**        |           |          |       |                  |
| N                | 252       | 68       | 49    | 12               |
| Age (median)     | 50        | 47       | 46    | 53               |
| **FL**           |           |          |       |                  |
| N                | 191       | 49       | 49    | 25               |
| Age (median)     | 50        | 47       | 46    | 57               |
| **CLL/SLL**      |           |          |       |                  |
| N                | 124       | 12       | 12    | 20               |
| Age (median)     | 51        | 53       | 53    | 56               |
| **Mantle**       |           |          |       |                  |
| N                | 120       | 25       | 25    | 20               |
| Age (median)     | 54        | 57       | 57    | 56               |
| **PTCL**         |           |          |       |                  |
| N                | 116       | 18       | 18    | 20               |
| Age (median)     | 46        | 48       | 48    | 57               |
| **Other T-cell** |           |          |       |                  |
| N                | 158       | 64       | 38    | 16               |
| Age (median)     | 48        | 36       | 38    | 29               |
| **Gender**       |           |          |       |                  |
| N                | 780       | 229      | 191   | 124              |
| %                | 63.9%     | 64.8%    | 64.8% | 51%              |
| **Male**         |           |          |       |                  |
| N                | 599       | 153      | 78    | 58               |
| %                | 63.9%     | 64.8%    | 57.4% | 35.2%            |
| **Female**       |           |          |       |                  |
| N                | 339       | 88       | 58    | 58               |
| %                | 36.1%     | 35.2%    | 42.6% | 48.8%            |
| **Donors**       |           |          |       |                  |
| N                | 6801      | 2076     | 1029  | 365              |
| %                |           |          |       |                  |

**Table 1.** Distribution of hematopoietic malignancies indicated for bone marrow transplant at the City of Hope 1995–2015 and potential donors (controls without cancer) for whom high resolution HLA haplotyping was imputed.
National Marrow and Donor Program, data from the majority of our CLL patients were included in a previous publication on HLA and CLL associations based on data at the NMDP [11]. We therefore present our institution-specific results for CLL in Supplemental Tables 4 and 5. For all NHL subtypes, results for Caucasians are shown in the main tables and those for nonwhite race groups, when sample size permits, are presented in Supplemental Tables 6–9.

**DLBCL.** Among Caucasians, the most common haplotype in donors was the AH 8.1 haplotype (HLA-A*0101–C*0701~B*0801~DRB1*0301~DQB1*0201) at 7.2% (Supplemental Table 3), but was not statistically significantly associated with DLBCL risk (OR 1.19, 95% CI 0.66–2.15) (Table 2). We identified three haplotypes, HLA-A*0301~C*0401~B*3501~DRB1*0101~DQB1*0501 (OR 2.70, 95% CI 1.21–6.01), HLA-A*0201~C*0701~B*0801~DRB1*0301~DQB1*0201 (OR 3.27, 95% CI 1.21–8.62). For non-Caucasians, Table 4 showed the most common HLA haplotype in the majority of our CLL patients were included in a previous publication on HLA and CLL associations based on data at the NMDP [11]. We therefore present our institution-specific results for CLL in Supplemental Tables 4 and 5. For all NHL subtypes, results for Caucasians are shown in the main tables and those for nonwhite race groups, when sample size permits, are presented in Supplemental Tables 6–9.

### Table 2. HLA haplotype association with B-cell NHL subtypes (DLBCL, FL, and MCL), in Caucasian patients and donors (analyses adjusted for sex and age, associations significant after multiple testing correcting in bold italics).

| HLA Region | Patient (%) | Donor (%) | OR     | 95% CI      | p value |
|------------|-------------|-----------|--------|-------------|---------|
| **DLBCL**  |             |           |        |             |         |
| A*01:01~C*07:01~B*08:01~DRB1*03:01~DQB1*02:01 | 7.2% | 7.2% | 1.14 | (0.67–1.93) | .63     |
| A*03:01~C*07:02~B*07:02~DRB1*15:01~DQB1*06:02 | 4.1% | 4.3% | 1.01 | (0.53–2.05) | .86     |
| A*29:02~C*16:01~B*44:03~DRB1*07:01~DQB1*02:01 | 3.2% | 2.4% | 1.48 | (0.67–3.29) | .36     |
| A*30:01~C*06:02~B*13:02~DRB1*07:01~DQB1*02:01 | 2.1% | 1.6% | 1.38 | (0.64–2.51) | .29     |
| A*01:03~C*04:01~B*35:01~DRB1*01:01~DQB1*05:01 | 3.7% | 1.5% | 2.70 | (1.21–6.01) | .04     |
| A*02:01~C*07:01~B*08:01~DRB1*03:01~DQB1*02:01 | 3.1% | 1.1% | 3.27 | (1.46–7.32) | .016    |
| A*26:01~C*12:03~B*38:01~DRB1*04:02~DQB1*03:02 | 3.0% | 1.1% | 3.06 | (1.34–6.97) | .01     |

| **FL**     |             |           |        |             |         |
|------------|-------------|-----------|--------|-------------|---------|
| A*01:01~C*07:01~B*08:01~DRB1*03:01~DQB1*02:01 | 7.2% | 7.2% | 1.13 | (0.64–1.97) | .68     |
| A*29:02~C*16:01~B*44:03~DRB1*07:01~DQB1*02:01 | 4.4% | 2.4% | 2.01 | (0.96–4.22) | .1      |
| A*23:01~C*04:01~B*44:03~DRB1*07:01~DQB1*02:01 | 3.0% | 1.1% | 2.69 | (0.95–7.58) | .17     |
| A*02:01~C*06:02~B*13:02~DRB1*07:01~DQB1*02:01 | 2.8% | 1.1% | 2.70 | (1.09–6.66) | .04     |

(continued)
Table 2. Continued.

| Loci          | Donor (%) | Patient (%) | OR     | 95% CI   | p value |
|---------------|-----------|-------------|--------|----------|---------|
| DR1*03:01    | 16.8%     | 17.9%       | 1.02   | (0.67–1.54) | .9      |
| DR1*07:01    | 41.3%     | 22.9%       | 2.41   | (1.75–3.31) | <.0001  |
| DQB1*02:01   | 45.3%     | 35.0%       | 1.63   | (1.16–2.3)  | .005    |

MCL

| Loci          | Donor (%) | Patient (%) | OR     | 95% CI   | p value |
|---------------|-----------|-------------|--------|----------|---------|
| A*01:01–C*07:01–B*08:01–DRB1*03:01–DQB1*02:01 | 8.2% | 7.2% | 1.42 | (0.72–2.80) | .37 |
| A*03:01–C*07:02–B*07:02–DRB1*15:01–DQB1*06:02 | 4.8% | 4.3% | 1.10 | (0.43–2.79) | .7   |
| A*02:01–C*03:04–B*15:01–DRB1*04:01–DQB1*03:02 | 2.1% | 2.0% | 1.20 | (0.30–4.81) | .8   |
| A*26:01–C*12:03–B*38:01–DRB1*04:02–DQB1*03:02 | 3.0% | 1.1% | 3.13 | (1.00–9.83) | .06  |

DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; MCL: mantle cell lymphoma.

Bold-italics indicate alleles significant after false discovery rate correction.

46–73.2), and HLA-A*2601–C*1203–B*3801–DRB1*0402–DQB1*0302 (OR 3.06, 95% CI 1.34–6.97), that appeared to confer an increased risk for DLBCL. Evaluation of the individual alleles which comprise those haplotypes did not yield any significant associations with DLBCL risk. However, we observed two individual HLA alleles which were significantly associated with increased DLBCL risk after FDR correction. HLA-A*0501 (OR 1.99, 95% CI 1.21–3.36) and HLA-C*0602 (OR 1.72, 95% CI 1.24–2.40). Both loci are part of the HLA-A*3001–C*0602–B*1302–DRB1*0701–DQB1*0201 haplotype, but that haplotype was not significantly associated with DLBCL (OR 1.36, 95% CI 0.46–4.10).

The AH 8.1 haplotype was only seen in 3% of Hispanic donors and 0.5% of Asian donors (Supplemental Table 3) and not associated with DLBCL risk (Supplemental Tables 6 and 8). The most common haplotype in Hispanic donors was HLA-A*6802–C*0802–B*1402–DRB1*0102–DQB1*0501 (4%), but was not statistically associated with DLBCL (OR 2.29, 95% CI 0.65–8.06). The most common haplotype in Asian donors was HLA-A*3303–C*0302–B*5801–DRB1*0301–DQB1*0201 (9%) and also not associated with DLBCL (OR 1.48, 95% CI 0.18–12.48, Supplemental Table 8). No other haplotypes or individual-level alleles were associated with DLBCL risk after correction for FDR among our Hispanic or Asian populations.

FL. Among Caucasians, the HLA-A*0201–C*0602–B*1302–DRB1*0701–DQB1*0201 haplotype was associated with increased FL risk (OR 2.70, 95% CI 1.09–6.66) (Table 2). We also observed a few haplotypes that were borderline significant, including HLA-A*2902–C*1601–B*4403–DRB1*0701–DQB1*0201 (OR 2.01, 95% CI 0.96–4.22) and HLA-A*2301–C*0401–B*4403–DRB1*0701–DQB1*0201 (OR 2.69, 95% CI 0.95–7.58).

Evaluation of the individual HLA alleles which comprise these haplotypes identified significant risk associations with HLA-DRB1*0701 (OR 2.41, 95% CI 1.75–3.31) and HLA-DQB1*0201 (OR 1.63, 95% CI 1.16–2.30), both of which remained statistically significant after FDR correction (Table 2). Among Hispanics, we observed an increased risk for FL with the haplotype HLA-A*0301–C*0401–B*3501–DRB1*0101–DQB1*0501 (OR = 8.71, 95% CI 2.06–36.88), with significant associations present for the HLA-DRB1*0101 (OR 4.83, 95% CI 2.39–9.78) and HLA-DQB1*0501 (OR 2.65, 95% CI 1.35–5.79) alleles (Supplemental Table 7). HLA associations for FL among Asians could not be evaluated as we only had 18 Asian FL patients with full HLA haplotype information.
Table 3. Effect of homozygosity at the three HLA class I loci -A, -B, and -C and three HLA class II loci -DRB1, -DQB1, and -DPB1 on susceptibility to three B-cell NHL subtypes (DLBCL, FL, and MCL) in Caucasian patients and donors (analyses adjusted for sex and age).

| Class I locus | Donors (n = 6801) | DLBCL (n = 229) | FL (n = 191) | MCL (n = 120) |
|---------------|------------------|-----------------|-------------|--------------|
| **HLA-A**     |                  |                 |             |              |
| Heterozygote  | 84.5%            | 82.5%           | 1.00 (REF)  |             |
| Homozygote    | 15.5%            | 17.5%           | 1.18 (0.83–1.68) | 1.36% 0.87 (0.57–1.33) | 18.3% 1.26 (0.78–2.04) |
| **HLA-B**     |                  |                 |             |              |
| Heterozygote  | 93.7%            | 92.1%           | 1.00 (REF)  |             |
| Homozygote    | 6.3%             | 7.9%            | 1.27 (0.78–2.10) | 9.4% 1.54 (0.93–2.53) | 9.2% 1.43 (0.75–2.73) |
| **HLA-C**     |                  |                 |             |              |
| Heterozygote  | 89.8%            | 87.3%           | 1.00 (REF)  |             |
| Heterozygote  | 10.2%            | 12.7%           | 1.28 (0.86–1.92) | 12.6% 1.26 (0.81–1.95) | 12.5% 1.23 (0.70–2.15) |
| **Total # of homozygote loci** | | | | |
| 0              | 75.5%            | 72.5%           | 1.00 (REF)  |             |
| 1              | 18.6%            | 20.5%           | 1.17 (0.84–1.63) | 16.2% 0.89 (0.60–1.31) | 18.3% 1.06 (0.65–1.71) |
| 2+             | 5.9%             | 7.0%            | 1.26 (0.75–2.15) | 8.4% 1.43 (0.84–2.43) | 9.2% 1.62 (0.84–3.11) |

| Class II locus | Donors (n = 6801) | DLBCL (n = 229) | FL (n = 191) | MCL (n = 120) |
|---------------|------------------|-----------------|-------------|--------------|
| **HLA-DRB1**  |                  |                 |             |              |
| Heterozygote  | 91.9%            | 85.9%           | 1.00 (REF)  |             |
| Heterozygote  | 8.1%             | 14.1%           | 1.89 (1.28–2.79) | 8.1% 2.06 (1.37–3.09) | 6.7% 0.80 (0.38–1.66) |
| **HLA-DQB1**  |                  |                 |             |              |
| Heterozygote  | 86.0%            | 76.1%           | 1.00 (REF)  |             |
| Heterozygote  | 14.0%            | 23.9%           | 1.46 (1.04–2.06) | 14.0% 1.50 (1.04–2.16) | 15.0% 1.09 (0.65–1.83) |
| **HLA-DPB1**  |                  |                 |             |              |
| Heterozygote  | 75.8%            | 74.8%           | 1.00 (REF)  |             |
| Heterozygote  | 24.2%            | 25.2%           | 1.03 (0.65–1.62) | 19.7% 0.76 (0.42–1.37) | 19.5% 0.71 (0.31–1.61) |
| **Total # of homozygote loci** | | | | |
| 0              | 65.1%            | 62.6%           | 1.00 (REF)  |             |
| 1              | 26.3%            | 26.2%           | 1.06 (0.67–1.68) | 15.5% 0.53 (0.27–1.02) | 26.8% 0.97 (0.47–2.03) |
| 2+             | 8.6%             | 11.2%           | 1.60 (0.84–3.07) | 11.3% 1.30 (0.61–2.79) | 2.4% 0.33 (0.04–2.48) |

*p-trend* 0.39 0.5 0.18

*Restricted to those with DPB molecular typing: 4594 Donors, 107 DLBCL, 74 FL, and 41 MCL.

**MCL.** The HLA-A*0201~C*0304~B*1501~DRB1*0401~DQB1*0302 haplotype was borderline significantly associated with increased MCL risk (OR 3.13, 95% CI 1.00–9.83) among Caucasians, but no individual alleles were significantly associated with MCL (Table 2). HLA-A*0101 was associated with MCL (OR 1.60, 95% CI 1.05–2.43); remaining significant after FDR correction, but the affiliated haplotype was not associated with MCL. Despite the relatively small number of Asians in our study population, we observed a strong association between the HLA-A*2402~C*0702~B*0702~DRB1*0101~DQB1*0501 haplotype and MCL (OR 11.7, 95% CI 2.25–60.9, Supplemental Table 9); the HLA-B*0702 (OR 8.41, 95% CI 2.20–32.2) and HLA DRB1*0101 (OR 11.12, 95% CI 2.50–49.4) loci both remained statistically significant after FDR correction. MCL risk among Hispanics could not be calculated due to the small number of patients in our study population.

**HLA class I and II homozygosity.** Among our Caucasian patients, we did not observe statistically significant associations between overall HLA class I homozygosity and NHL subtype risk, though the associations were trending toward increased risk with increasing homozygosity (Table 3). We observed increased DLBCL and FL risk with homozygosity at HLA class II loci, including HLA-DRB1 (DLBCL OR 1.89, 95% CI 1.28–2.79; FL OR 2.06, 95% CI 1.37–3.09) and HLA-DQB1 (DLBCL OR 1.46, 95% CI 1.04–2.06; FL OR 1.50, 95% CI 1.04–2.16). Among FL, we further observed an association between increasing number of homozygous HLA class II loci and increased FL risk among Caucasians (*p-trend* = 0.3).

**Discussion**

Our analysis of HLA variation and NHL risk among a transplant-indicated patient population for NHL confirmed previously identified risk alleles and identified some potentially novel alleles that warrant follow-up and replication. Specifically, statistically significant associations, even after FDR correction, were observed between HLA-DRB1*0701 and FL among our Caucasian patients, which was previously reported in GWAS efforts [21]. Associations between HLA-DRB1*0101 and HLA-DQB1*0501 and FL were also evident in our Hispanic patients [10,38]. For DLBCL, significant association between the HLA-A*0201~C*0701~B*0801~DRB1*0301~DQB1*0201 haplotype was observed, which notably shares the same previously GWAS-implicated HLA-B*0801~DRB1*0301 region that is also in high LD with the previously reported risk allele, TNF-308A, as well as the HLA-DR3-DQ2 serotype that is associated with several other autoimmune conditions [39]. Individually, however, associations between GWAS-confirmed loci, HLA-B*0801 and HLA-DRB1*0301,
and the ancestral haplotype 8.1 (HLA-A*0101~C*0701~B*0801~DRB1*0301~DQB1*0201) were null in our population [5,17].

The other two HLA haplotypes we observed with increased DLBCL risk in our Caucasian population (HLA-A*2601~C*1203~B*3801~DRB1*0402~DQB1*0302 and HLA-A*0301~C*0401~B*3501~DRB1*0101~DQB1*0501) have not previously been reported and require replication. We note that the HLA-A*2601~B*3801~DRB1*0402 haplotype is more commonly seen in Ashkenazi Jewish populations (2.3–6.7% frequency) [40,41] compared to other non-Hispanic European populations (<0.5% frequency) [31] and it is possible that this association may reflect a bias in our case population. We also observed increased DLBCL risk with the HLA-B*1302 and HLA-C*0602 alleles, even after FDR correction, the latter which is implicated in increased risk of CLL [11] and psoriasis [42] though the two alleles are in high LD (D’=0.98). HLA haplotype and allele frequencies varied widely among Hispanics and Asians, compared to Caucasians; the associations we observed among Caucasians were not replicated among either Hispanic or Asian subgroups, though our sample size was admittedly limited.

In our Caucasian population, we replicated previously reported GWAS association for FL (HLA-DRB1*0701) [21] and further implicate the HLA-A*0201~C*0602~B*1302~DRB1*0701~DQB1*0201 haplotype as a possible FL risk haplotype. In addition, we observed increased risk with the HLA-DQB1*0201 allele which was in high LD with HLA-DRB1*0701 (D’=0.76). The replication of GWAS results is not entirely unexpected, as FL is in general a more indolent disease for which minimal survival bias would be expected among population-based studies. Of the NHL subtypes evaluated, the difference between our clinical subset of transplant-indicated FL patients and that of previously published population studies would be the least dissimilar [26]. These associations were not, however, replicated in the other race groups. Among Hispanics, we observed a very different haplotype association between HLA-A*0301~C*0401~B*3501~DRB1*0101~DQB1*0501 and FL. Although we cannot exclude the possibility that HLA associations differ by race/ethnicity, we also cannot exclude the possibility that these associations are false positive results due to the relatively small sample size available for analysis in non-Caucasian race groups.

In MCLs, we found no significant HLA haplotype association, although we did observe an association with the HLA-A*0101 allele, an association previously reported for both FL [6] and EBV + HL [1]. HLA-A*0101 had been implicated as a risk factor in type 1 diabetes but recent studies have attributed this to the risk associated with the AH8.1 haplotype, and specifically to its linkage with the HLA-DRB1*0301 allele [43]. Although we observed no association between the AH8.1 haplotype and MCL, the HLA-DRB1*0301 allele was significantly associated with increased MCL risk but did not remain significant after FDR correction. We observed different associations among Asians in our study population, specifically associations between the HLA-A*2402~C*0702~B*0702~DRB1*0101~DQB1*0501 haplotype and MCL. Similar to the observations observed for DLBCL and FL, our lack of replication of observed results in Caucasians coupled with new distinct significant results among other race groups, suggests either that our results among other race groups may be due to chance or a larger implication that distinct HLA associations exist for different race/ethnic groups. We also note that MCL patients may also undergo transplant as part of consolidation therapy and as such may not necessarily be a higher risk population compared to our DLBCL or FL patients. At time of transplant, 51% of the MCL patients were in first complete response, compared to only 10% of DLBCL and 7% of FL patients.

Our evaluation of HLA homozygosity was not entirely consistent to previous reports, which were largely derived from population-based studies [10,44]. We did not observe a statistically significant overall increase in disease risk with increasing homozygosity at the HLA class I loci. However, we did observe associations with HLA class II homozygosity, specifically between HLA-DRB1 and HLA-DQB1 homozygosity and both FL and DLBCL risk [44]. Our data thus appear to support the notion that class II heterozygosity provides greater ability for the host to enhance tumor surveillance in its role in antigen presentation [45,46] and in clearing related infections [47,48]. At present, it remains unclear what the underlying biology is, should one exist, for the observed associations between the two specific HLA class II zygosity and FL and DLBCL.

Strengths of our study include a relatively large sample size and availability of subtype information for NHL. Our HLA typing was largely ascertained through genotyping, in contrast to GWAS-based analyses which are imputation-based. The available HLA typing information was critical in permitting us to impute HLA haplotype information for data analysis. In addition, our comparison group of donors was drawn from among individuals typed as possible matches for our patients. As donors include some family members, it is plausible that the comparison donor population might
be slightly more likely to have shared alleles, which would likely attenuate our results. However, we did not see large differences in allele frequencies between our donors and the NMDP population. Study strengths also include the unique study population of transplant-indicated NHLs, which differ from those reflected in published HLA-association literature. In particular, our patients are diagnosed at higher stage, as would be expected of those undergoing transplantation. Study limitations include the exclusion of patients and donors for whom we did not have age, sex, and race. We also note that unlike GWASs conducted to date which rely on genetic characterization of race/ethnicity for stratified analysis, we relied on self-reported race/ethnicity from our patient population. We were thus unable to examine genetic admixture in our population, and as such allele and haplotype frequencies may be affected, especially among Hispanics [49]. Further studies that are able to incorporate genetic characterization of race/ethnicity are thus needed to confirm our associations. In addition, while our sample size is relatively large by subtype, the large number of HLA haplotypes and alleles limits our statistical power. We also note that a number of patients that underwent autologous transplant had insufficient HLA typing for inclusion in our analysis; this is largely due to historical data at COH for which only low resolution A and B typing and anti-HLA antibody testing were available among those patients. Finally, even though COH performs the most transplants for NHL in the nation [50], our results only represent a subset of patients seen at a single institution and will require replication in larger populations, particularly for minority populations.

To our knowledge, this and the previous publication on CLL [11] are among the first evaluations of HLA association in NHL subtypes specifically among transplant-indicated populations. Similar to the report of CLL, some GWAS associations reported among the general population were replicated in this unique transplant-indicated population, but our results further suggest new associations among more severe disease which require replication in other similar study populations. Complementary analyses of these and other risk factors (genetic and non-genetic) for more severe disease warranting transplant would shed light on whether there are unique risk factors or risk factor combinations which indeed result in more aggressive disease.

Acknowledgements

This work was supported by City of Hope Tim Nesvig Lymphoma Foundation.

References

[1] Johnson PCD, McAulay KA, Montgomery D, et al. Modeling HLA associations with EBV-positive and -negative Hodgkin lymphoma suggests distinct mechanisms in disease pathogenesis. Int J Cancer. 2015;137:1066–1075.
[2] McAulay KA, Jarrett RF. Human leukocyte antigens and genetic susceptibility to lymphoma. Tissue Antigens. 2015;86:98–113.
[3] Wang SS, Lu Y, Rothman N, et al. Variation in effects of non-Hodgkin lymphoma risk factors according to the human leukocyte antigen (HLA)-DRB1*01:01 allele and ancestral haplotype 8.1. PLoS ONE. 2011;6:e26949.
[4] Riemersma SA, Jordanova ES, Haasnot GW, et al. The relationship between HLA class II polymorphisms and somatic deletions in testicular B cell lymphomas of Dutch patients. Hum Immunol. 2006;67:303–310.
[5] Abdou AM, Gao X, Cozen W, et al. Human leukocyte antigen (HLA) A1-B8-DR3 (8.1) haplotype, tumor necrosis factor (TNF) G-308A, and risk of non-Hodgkin lymphoma. Leukemia. 2010;24:1055–1058.
[6] Lu Y, Abdou AM, Cerhan JR, et al. Human leukocyte antigen class I and II alleles and overall survival in diffuse large B-cell lymphoma and follicular lymphoma. Sci World J. 2011;11:2062–2070.
[7] Akers NK, Curry JD, Conde L, et al. Association of HLA-DQB1 alleles with risk of follicular lymphoma. Leuk Lymphoma. 2011;52:53–58.
[8] Skibola CF, Bracci PM, Nieters A, et al. Tumor necrosis factor (TNF) and lymphotoxin-alpha (LTA) polymorphisms and risk of non-Hodgkin lymphoma in the InterLymph Consortium. Am J Epidemiol. 2010;171:267–276.
[9] Rothman N, Skibola CF, Wang SS, et al. Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. Lancet Oncol. 2006;7:27–38.
[10] Wang SS, Abdou AM, Morton LM, et al. Human leukocyte antigen class I and II alleles in non-Hodgkin lymphoma etiology. Blood. 2010;115:4820–4823.
[11] Gragert L, Fingerson S, Albrecht M, et al. Fine-mapping of HLA associations with chronic lymphocytic leukemia in US populations. Blood. 2014;124:2657–2665.
[12] El Ansary MM, Mohammed LA, Hassan TH, et al. Human leukocyte antigen-DRB1 polymorphism in childhood acute lymphoblastic leukemia. Mol Clin Oncol. 2015;3:425–429.
[13] Speedy HE, Di Bernardo MC, Sava GP, et al. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2014;46:56–60.

Potential conflict of interest: Disclosure forms provided by the authors are available with the full text of this article online at http:\/\<doi:10.1080/10428194.2019.1617858>.

ORCID

Charlie Zhong http://orcid.org/0000-0001-6973-887X
Joo Song http://orcid.org/0000-0003-3497-2513
Leanne Goldstein http://orcid.org/0000-0002-7721-6358
[14] Taylor GM, Hussain A, Verhave V, et al. Strong association of the HLA-DP6 supertype with childhood leukemia is due to a single allele, DPB1*0601. Leukemia. 2009;23:863–869.

[15] Orooji E, Tavakkol Afshari J, Badlee Z, et al. Association between HLA-DQ81 gene and patients with acute lymphoblastic leukemia (ALL). Int J Hematol. 2012;95:551–555.

[16] Naugler C, Liwski R. Human leukocyte antigen class I alleles and the risk of chronic myelogenous leukemia: a meta-analysis. Leuk Lymphoma. 2010;51:1288–1292.

[17] Cerhan JR, Berndt SI, Vijai J, et al. Genome-wide association study identifies multiple susceptibility loci for diffuse large B cell lymphoma. Nat Genet. 2014;46:1233–1238.

[18] Basso BA, Cerhan JR, Au WY, et al. Genetic susceptibility to diffuse large B-cell lymphoma in a pooled study of three Eastern Asian populations. Eur J Haematol. 2015;95:442–448.

[19] Smedby KE, Foo JN, Skibola CF, et al. GWAS of follicular lymphoma reveals allelic heterogeneity at 6p21.32 and suggests shared genetic susceptibility with diffuse large B-cell lymphoma. PLoS Genet. 2011;7:e1001378.

[20] Conde L, Halperin E, Akers NK, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nat Genet. 2010;42:661–664.

[21] Skibola CF, Berndt SI, Vijai J, et al. Genome-wide association study identifies five susceptibility loci for follicular lymphoma outside the HLA region. Am J Hum Genet. 2014;95:462–471.

[22] Berndt SI, Skibola CF, Joseph V, et al. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. Nat Genet. 2013;45:868–876.

[23] Slager SL, Goldin LR, Strom SS, et al. Genetic susceptibility variants for chronic lymphocytic leukemia. Cancer Epidemiol Biomarkers Prev. 2010;19:113.

[24] Slager SL, Camp NJ, Conde L, et al. Common variants within 6p21.31 locus are associated with chronic lymphocytic leukemia and, potentially, other non-Hodgkin lymphoma subtypes. Br J Haematol. 2012;159:572–576.

[25] Vijai J, Wang Z, Berndt SI, et al. A genome-wide association study of marginal zone lymphoma shows association to the HLA region. Nat Commun. 2015;6:5751.

[26] Cerhan JR, Frederiksen ZS, Wang AH, et al. Design and validity of a clinic-based case-control study on the molecular epidemiology of lymphoma. Int J Mol Epidemiol Genet. 2011;2:95–113.

[27] Holly EA, Gautam M, Bracci PM. Comparison of interviewed and non-interviewed non-Hodgkin’s lymphoma (NHL) patients in the San Francisco Bay Area. Ann Epidemiol. 2002;12:419–425.

[28] Hu ZH, Connett JE, Yuan JM, et al. Role of survivor bias in pancreatic cancer case-control studies. Ann Epidemiol. 2016;26:50–56.

[29] Morton LM, Turner JJ, Cerhan JR, et al. Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). Blood. 2007;110:695–708.

[30] Turner JJ, Morton LM, Linet MS, et al. InterLymph hierarchical classification of lymphoid neoplasms for epidemiologic research based on the WHO classification (2008): update and future directions. Blood. 2010;116:e90–e98.

[31] González-Galarza Faviel F, Takeshita Louise YC, Santos Eduardo JM, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. Nucleic Acids Res. 2015;43:D784–D788.

[32] Gourraud P-A, Lamiraux P, El-Kadhi N, et al. Inferred HLA haplotype information for donors from hematopoietic stem cells donor registries. Hum Immunol. 2005;66:563–570.

[33] Madbouly A, Gragert L, Freeman J, et al. Validation of statistical imputation of allele-level multilocus phased genotypes from ambiguous HLA assignments. Tissue Antigens. 2014;84:285–292.

[34] Pauñic V, Gragert L, Madbouly A, et al. Measuring ambiguity in HLA typing methods. PLoS One. 2012;7:e43585.

[35] Gragert L, Madbouly A, Freeman J, et al. Six-locus high resolution HLA haplotype frequencies derived from mixed-resolution DNA typing for the entire US donor registry. Hum Immunol. 2013;74:1313–1320.

[36] Klein SL, Jedlicka A, Pekosz A. The Xs and Y of immune responses to viral vaccines. Lancet Infect Dis. 2010;10:338–349.

[37] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B (Methodol). 1995;57:289–300.

[38] Skibola CF, Akers NK, Conde L, et al. Multi-locus HLA class I and II allele and haplotype associations with follicular lymphoma. Tissue Antigens. 2012;79:279–286.

[39] Smigoc Schweiger D, Mendez A, Kunilo Jamnik S, et al. High-risk genotypes HLA-DR3-DQ2/DR4-DQ2 and DR3-DQ2/DR4-DQ8 in co-occurrence of type 1 diabetes and celiac disease. Autoimmunity. 2016;49:240–247.

[40] Klitz W, Gragert L, Maiers M, et al. Genetic differentiation of Jewish populations. Tissue Antigens. 2010;76:442–458.

[41] Manor S, Halagan M, Shriki N, et al. High-resolution HLA A~B~DRB1 haplotype frequencies from the Ezer Mizion Bone Marrow Donor Registry in Israel. Hum Immunol. 2016;77:1114–1119.

[42] Guðjónsson JE, Valdimarsson H, Kárason A, et al. HLA-Cw6-positive and HLA-Cw6-negative patients with psoriasis vulgaris have distinct clinical features. J Invest Dermatol. 2002;118:362–365.

[43] Noble JA. Immunogenetics of type 1 diabetes: a comprehensive review. J Autoimmun. 2015;64:101–112.

[44] Wang SS, Carrington M, Berndt SI, et al. HLA class I and II diversity contributes to the etiologic heterogeneity of non-Hodgkin lymphoma subtypes. Cancer Res. 2018;78:4086–4096.

[45] Wang SS, Hartge P. Etiology and epidemiology of non-Hodgkin lymphoma. In: Armitage JO, Mauch PM, Harris NL, et al., editors. Non-Hodgkin lymphomas. 2nd ed. Philadelphia (PA): Lippincott, Williams and Wilkins; 2010. p. 64–82.

[46] Blackwell JM, Jamieson SE, Burgner D. HLA and infectious diseases. Clin Microbiol Rev. 2009;22:370–385.

[47] Thio LC, Thomas DL, Karacki P, et al. Comprehensive analysis of class I and class II HLA antigens and
chronic hepatitis B virus infection. J Virol. 2003;77:12083–12087.

[48] De La Concha EG, Fernandez-Arquero M, Martinez A, et al. HLA class II homozygosity confers susceptibility to common variable immunodeficiency (CVID). Clin Exp Immunol. 1999;116:516–520.

[49] Arrieta-Bolaños E, Madrigal JA, Shaw BE. Human leukocyte antigen profiles of Latin American populations: differential admixture and its potential impact on hematopoietic stem cell transplantation. Bone Marrow Res. 2012;2012:13.

[50] Center for International Blood and Marrow Transplant, a contractor for the C.W. Bill Young Cell Transplantation Program operated through the U.S. Department of Health and Human Services, Health Resources and Services Administration, Healthcare Systems Bureau. U.S. Transplant Data by Center Report, NHL – Non-Hodgkin Lymphoma, Number of Transplants Reported in 2016. U.S. Transplant Data by Center Report. Last Updated: May 17, 2017; [cited 9 March 2018]. Available from: https://bloodcell.transplant.hrsa.gov/research/transplant_data/us_tx_data/data_by_center/center.aspx