Human leukocyte antigens and genetic susceptibility to lymphoma

Karen A McAulay and Ruth F Jarrett

MRC - University of Glasgow Centre for Virus Research, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

Key Words: Hodgkin lymphoma, HLA, EBV, B-cell lymphoma

Word Count: 6971
Short Title: HLA associations with lymphoma

Correspondence: Professor Ruth Jarrett
Ian Botham Building
Garscube Estate
Bearsden Road
Glasgow G61 1QH
Tel: 44 141 330 5775
Fax: 44 141 330 5733
Email: ruth.jarrett@glasgow.ac.uk

Abstract

Familial aggregation, coupled with ethnic variation in incidence, suggests that inherited susceptibility plays a role in the development of lymphoma, and the search for genetic risk factors has highlighted the contribution of the Human Leukocyte Antigen (HLA) complex. In a landmark study published almost fifty years ago, Hodgkin lymphoma was the first disease to be associated with HLA variation. It is now clear that EBV-positive and -negative Hodgkin lymphoma are strongly associated with specific HLA polymorphisms but these differ by EBV status of the tumours. HLA class I alleles are consistently associated with EBV-positive Hodgkin lymphoma while a polymorphism in HLA class II is the strongest predictor of risk of EBV-negative Hodgkin lymphoma. Recent investigations, particularly genome wide association studies, have also revealed associations between HLA and common types of non-Hodgkin lymphoma. Follicular lymphoma is strongly associated with two distinct haplotypes in HLA class II whereas diffuse large B-cell lymphoma is most strongly associated with *HLA-B*08*. Although chronic lymphocytic leukaemia is associated with variation in HLA class II, the strongest signals in genome wide association studies are from non-HLA polymorphisms, suggesting that inherited susceptibility is explained by co-inheritance of multiple low risk variants. Associations between B-cell derived lymphoma and HLA variation suggest that antigen presentation, or lack of, plays an important role in disease pathogenesis but the precise mechanisms have yet to be elucidated.

Introduction

Mature B-cell neoplasms account for 90% of all lymphomas globally and are divided into Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL) (1). HL is one of the most common lymphomas in the developed world and comprises two entities, nodular lymphocyte predominant Hodgkin lymphoma (NLPHL) and classical Hodgkin lymphoma (cHL), which account for <5% and >95% of cases, respectively. The NHLs represent a diverse and heterogeneous group of malignancies. The 2008 World Health Organization (WHO) classification of lymphomas includes more than 60 distinct subtypes of NHL with diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL) and chronic lymphocytic leukaemia (CLL)/small lymphocytic lymphoma (SLL) the most common.

The aetiology of both HL and NHL is still unclear. However, current evidence suggests that environmental exposures, infectious agents, and both immunosuppression and immune activation contribute to the development of these diseases. In addition, as described below, there is compelling evidence for inherited susceptibility to HL and NHL. The search for genetic risk factors has highlighted the importance of the human leukocyte antigen (HLA) region (2, 3). A crucial role for HLA in disease pathogenesis is further suggested by observations that HLA class I and class II expression is frequently down-regulated in tumour tissue (4, 5). Taken together, the data implicate HLA-mediated interactions in the pathogenesis of both HL and NHL. This review aims to summarise the main findings that have emerged from HLA association studies and genome-wide association studies (GWAS) of HL and the three most common forms of NHL, namely DLBCL, FL and CLL/SLL.

Evidence for inherited susceptibility in lymphoma

Inherited susceptibility to HL and NHL is supported by several strands of evidence. First, the incidence of HL and NHL varies significantly between racial and ethnic groups with the Asian population showing one of the lowest incidence rates for both diseases (6-8). Population-based studies in the USA demonstrate that the incidence of HL and CLL is consistently low in Asian groups relative to US Whites, suggesting some genetic protection against disease, given similar environmental exposures (6, 8). Secondly, there is clear evidence of familial aggregation of both HL and NHL. Studies of HL indicate a 3-9 fold increased risk of HL in first-degree relatives of index cases (9-12). Furthermore, a study of monozygotic and dizygotic twins demonstrated that the co-twin of a monozygotic HL case has a 100-fold increased risk of developing the disease (13). The Swedish Family Cancer database estimated the heritability of HL in the white population as 28.4% (14). Individual case-control studies as well as a recent meta-analysis of NHL subtypes also show an increased incidence of DLBCL, FL and CLL in patients with a family history of NHL, with CLL showing the greatest increase of 8-fold (15-21). Familial aggregation of CLL cases has also been documented in over 100 familial studies (reviewed in 22). Collectively, the data support a genetic predisposition to the development of lymphoma.

Human Leukocyte Antigens and Genes

HLA molecules present pathogen and tumour-derived peptides to T-cells thereby initiating the adaptive immune response. The classical HLA class I molecules, HLA-A, -B and -C, are heterodimers of a polymorphic α-polypeptide chain and β2-microglobulin, present endogenous peptides to cytotoxic CD8 positive T-cells (CD8+T-cells), and are expressed by almost all somatic cells. In contrast, the HLA class II molecules HLA-DR, -DQ and -DP are heterodimers of α- and β-polypeptide

chains, present endocytosed extracellular peptides to helper CD4 positive T-cells (CD4+T-cells), and are normally expressed only on antigen presenting cells.

The highly polymorphic HLA genes are contained within the HLA complex, the human equivalent of the major histocompatibility complex, which is located on the short arm of chromosome 6 and divided into three regions. The class I region contains genes for the α-chains of HLA-A, B and C and also the three non-classical HLAs (*HLA-E, -F and –G*), which are differentiated by their limited polymorphism and restricted tissue expression. Similarly, the class II region includes genes for HLA-DR, DQ, and DP as well as HLA-DO and DM, the non-classical class II antigens. *DQA1* and *DQB1* genes encode the α- and β-chains of the HLA-DQ heterodimer and, similarly, *DPA1* and *DPB1* genes encode the respective chains of the DP heterodimer. For DR the situation is more complex as the number of DR genes varies between individual MHC regions. All have a single *DRA* gene and a *DRB1* gene but some haplotypes have an additional, functional *DRB3*, *DRB4*, or *DRB5* gene. For example, haplotypes containing the DRB1 alleles DRB1*04:01, *07:01 and *09:01 also have a DRB4 gene. This results in expression of a second DR molecule with α- and β-chains encoded by the *DRA* and *DRB4* genes, respectively (23). Out of a total of 128 class I and 27 class II genes, 42 and 17 encode proteins, the majority of which have an established role in immunity.

The HLA class I and II regions are separated by the class III region. Although there are no HLA genes within this region, it encodes proteins with immune function, such as tumour necrosis factor (TNF)-α and -β, heat shock proteins and the complement factors C2 and C4.

HLA class I and II molecules have similar conformations including a peptide-binding groove which contains several pockets. The pockets accommodate side chains of amino acids in the bound peptide and some of these, referred to as anchoring pockets, have a critical role in determining peptide-binding specificity (reviewed in 24). Polymorphisms in HLA genes lead to amino acid variation within these pockets, and this in turn affects peptide binding preferences; however, products of different HLA alleles can contain identical anchoring pockets thus enabling grouping of HLA alleles into supertypes on the basis of overlapping peptide specificities (25).

To enable an efficient immune response, HLA molecules must bind a peptide and the T-cell repertoire must include clones that recognise the HLA-peptide combination. Absence of either may render an individual susceptible to a given disease.

HLA association studies: caveats and pitfalls

The analysis of HLA associations with disease is challenging and a number of factors need to be taken into consideration. First, study design and sample size are critical but for reasons of practicality are seldom ideal. In case-control studies, spurious results can arise from poor matching between cases and controls, diagnostic misclassification and disease heterogeneity, biased case selection, and retrospective ascertainment of cases leading to survival bias. Population stratification, resulting from admixture of populations with different ancestry, is a particular problem in genetic association studies. Family-based designs aim to avoid the potential confounding effects of population stratification by using parents or unaffected siblings as controls for the index case. However, family-based studies are often limited by size. Small studies frequently report strong associations that

disappear or become weak when sample size increases. Sample size is also crucial in case-control studies. To have 80% power to detect a significant association ($p<0.05$) with an allele with a carrier frequency of 0.10 with an odds ratio of 2 requires >300 cases; following correction for multiple testing of e.g. 50 alleles, this number increases to >600. Thus, most published studies of HLA associations are under-powered to detect small effects sizes and associations with rare alleles. The necessary larger-sized studies are usually limited by cost and sample availability.

Secondly, it is often assumed that the genotyped locus confers susceptibility/protection from disease development. However, the genotyped locus may not be causal but rather may be in linkage disequilibrium (LD) with the causal variant. This is a particular challenge in HLA association studies as there is extensive LD within the HLA complex and many alleles are inherited in a block. Thus, unadjusted analyses frequently show associations with multiple alleles on the same haplotype; this is less of problem with *DP* genes as there is a recombination hotspot between *DQ* and *DP*. R^2 and D' values are used to describe the extent of LD between two genetic variants. R^2 is a measure of correlation, and high values (>0.8) occur when two alleles are present at equal frequency and inherited together; an r^2 of 1 indicates that only two of the four possible haplotypes are present. For tagging studies, where one marker is used to tag or substitute for another, r^2 is the important measure; however, in population genetics D', a measure of recombination events between two markers which is not dependent on allele frequency, is generally more informative. A high D' value indicates that the rarer allele is always inherited with the other allele or, in the case of single nucleotide polymorphisms (SNPs), with one of the alleles; r-squared is usually lower since most of the time the commoner allele is present without the rarer allele.

Most HLA association studies describe the frequency of a particular allele (or genotype or phenotype) according to case-control status, with unadjusted statistical analyses, i.e. not adjusted for the effects of other alleles. In order to identify potential causal alleles, various statistical methods have been used to adjust for effects of LD; although generally more informative, this can complicate comparisons between studies.

Thirdly, the methodology used to type HLA alleles has changed dramatically over the past 40 years. Traditionally HLA antigens were defined using serological techniques. Whilst serology performed adequately for typing families, it was unsatisfactory for typing unrelated donors particularly once the extent of HLA polymorphism became known. The availability of nucleotide sequence data for HLA genes allowed the rapid development of PCR-based typing methods. These can be grouped into two categories: (i) those which generate a product containing internal polymorphisms, which are then identified using a second technique, e.g. PCR-sequence specific oligonucleotide probing, PCR-restriction fragment length polymorphism; (ii) those in which the polymorphism is identified as part of the PCR process, e.g. PCR-sequence specific primer. These techniques have proven reliable, robust and accurate and are currently the most widely used methods. In recent research investigations, HLA alleles have also been imputed from SNPs identified in GWAS; as more data are becoming available, this method is becoming increasingly reliable. Ultimately, sequencing is required to unequivocally identify specific alleles.

Changes in typing methodology coupled with an improved understanding of the HLA complex have led to changes in HLA nomenclature, which can complicate comparisons across studies. Serotypes, defined by serological typing, are largely separated into class I and II groups and numbered sequentially, e.g. HLA-A1, A2 etc.

Broad specificities are sometimes further divided into ‘split specificities’. With the introduction of molecular typing the naming system was altered and allele designations now include up to four fields separated by colons. The first field contains an asterisk to denote molecular typing and defines the allele of the gene, which often but not always corresponds to the serotype; the second field defines the subtype of the allele. Early HLA association studies generally reported only the first field of the descriptor with more recent, higher resolution studies reporting the first two fields. In describing reported HLA associations we have attempted to adhere to the nomenclature used in the individual studies.

Hodgkin lymphoma

A distinguishing feature of HL, shared by both NLPHL and cHL, is the paucity of malignant cells within the tumour mass, which is largely composed of an inflammatory infiltrate (26). Until 2001, HL was classified as a single disease, Hodgkin’s disease, with four histological subtypes. Following the acceptance that Hodgkin’s disease is a B-cell derived lymphoma and that lymphocyte predominance Hodgkin’s disease has many distinctive features, the disease was renamed HL and separated into two main entities, NLPHL and cHL (27). Based upon the infiltrate composition and architecture and the morphology of the Hodgkin and Reed-Sternberg (HRS) cells, cHL is now further subdivided into four histological subtypes. The nodular sclerosis (NSHL) subtype is the most common, accounting for up to 80% of cases, whilst the mixed cellularity (MCHL) subtype accounts for around 15%.

In the late 1980s it was recognized that a proportion of cHL cases, in developed countries, are associated with the EBV (EBV+cHL) and can be distinguished from EBV-negative (EBV-cHL) cases by the presence of the virus within HRS cells (28,

29). EBV association rates differ by cHL subtype, age, gender and geography reflecting the complexity of the disease (7, 30, 31). Many HLA association studies were performed before the WHO classification system was introduced and before it was appreciated that EBV was aetiologically associated with a subgroup of cases. Since most MCHL, but only a minority of NSHL, cases are EBV-positive, MCHL histology and EBV-positivity are frequently used as proxies for each other; although this can help interpretation, it is important to note that at least half of all EBV+cHL cases are of NSHL subtype (our unpublished results).

Early HLA association studies in Hodgkin lymphoma

HL was the first disease for which any HLA association was described (32). Using 45 sera from multiparous women, Amiel *et al* (1967) reported an increase in frequency of the 'Payne-Bodmer 4c' antigen from 23% in healthy controls to 51% in HL cases from France. A similar analysis of Dutch cases failed to confirm these results (33). However, the results were corroborated in 1970 in Australia by Forbes *et al* (1970) who demonstrated an increased frequency of the 4c antigen and further showed that this resulted from an increased frequency of W5, one of the two specificities recognised by the 4c antibody (34). Forbes et al (1970) also showed that W5 was less common in females with NSHL and more common in females with MCHL. The 4c specificities were subsequently shown to include HLA-B5, B15, B18 and Bw35 (W5) antigens (35).

Subsequent investigations using serological typing methods demonstrated an increased frequency of HLA-B5, B8, B18 and HLA-A1 subtypes in HL patients (Table 1) (36-44). The strongest association was shown for HLA-A1 with a 6-fold increase in risk of developing cHL (42). Despite the well-documented linkage between HLA-

A1 and B8 antigens, HLA-B8 did not always show a similar increase (39, 43). Kissmeyer-Nielson et al (1975) also demonstrated an increased HLA-A1 and A8 prevalence in MCHL cases compared to NSHL and NLPHL cases (39), suggesting subtype-specific HLA associations. Larger, case-control studies further corroborated the association with HLA-A1 and identified a lower prevalence of HLA-A3 and A11 alleles in HL cases compared to healthy controls, suggesting a potential protective effect (Table 1) (35, 36, 45).

Concurrent family-based studies provided some support for these findings. Marshall et al (1977) investigated seven cases of HL within one family and found no HLA haplotype or allele associated with HL development, although HLA-B8 was detected more often in first-degree relatives of cases (46). Further studies demonstrated an increased prevalence of HLA-B18, B35, and B37 in cases within families (47, 48). Using a shared versus non-shared HLA approach for sibling pairs, a greater than expected HLA identity between siblings with HL was documented (43); however, previously reported allele associations were not detected in this study. Similar haplotype concordance was observed between HL-affected siblings in a study of 16 families where half the families displayed the HLA-A*01-B*08 haplotype (49).

Investigation of associations with class II alleles lagged behind those of class I alleles due to a lack of appropriate reagents. The earliest observations arose from family studies. In a study of four siblings with HL, HLA typing revealed that all carried the HLA-DR5 antigen (HLA-DR*11 and DR*12 gene products) and two were also HLA-A1-positive (50). With no obvious differences in EBV antibody titres, chromosomal abnormalities or exposure histories, it was concluded that genetic factors were more important than environmental factors in disease pathogenesis.

Interestingly, some of the earlier studies performed at a time with much higher mortality from HL also observed a survival advantage in association with particular HLA class I types (41, 42, 44, 45, 51). An increased frequency of HLA-A28 and A8 was observed in HL patients with greater than 15 years and 5 years survival, respectively (41, 44, 45). Although HLA-A1 was associated with increased disease risk, it was also associated with improved survival and treatment response (42). Patients positive for HLA-Aw19 (HLA-A29, 30, 31, 32, 33 gene products) and B5 antigens also had a poorer prognosis (51).

HLA and HL in the era of molecular typing

Following the introduction of molecular typing methods, attention shifted to HLA class II. Case-control studies demonstrated an increased frequency of HLA-DPB1*03:01 in white patients, with a relative risk of 1.95 ($p < 0.01$) in the largest study (Table 2) (52-57). The HLA-DPB1*04:01 allele was significantly reduced risk in Asian patients, while there was a non-significant reduction in DPB1*02:01 in white patients (52, 53). Taylor et al (1999) analysed cases by histological subtype and gender and demonstrated an increased relative risk for females carrying HLA-DPB1*03:01 or DPB1*10:01 alleles and a reduced risk for females with NSHL who carried DPB1*02:01 or DPB1*11:01 alleles (58). In one of the first studies to stratify cases by EBV status, the HLA-DPB1*03:01 allele was detected more frequently in patients with EBV+cHL, but differences by EBV status were not statistically significant (54).

Familial studies highlighted the importance of the HLA-DRB1*15:01-DQA1*01:02-DQB1*06:02 haplotype in relation to HL, in particular NSHL, with transmission of the haplotype 73% of the time to an affected offspring but only 44% of the time to an

unaffected offspring (59). Analysis of the individual genes in the haplotype revealed the DRB1*15:01 and DQB1*06:02 were associated with risk of cHL overall whilst DQA1*01:02 was associated only with NSHL.

A much more detailed, but as yet incomplete, understanding of HLA associations with cHL has come from recent HLA typing studies coupled with analysis of microsatellites and SNPs in the HLA region. Typing of tumour EBV status has also revealed that associations with EBV+cHL and EBV-cHL are largely distinct. A landmark study by Diepstra et al (2005) analysed microsatellite markers spanning the HLA class I, II and III regions in EBV-stratified HL cases and controls. Two consecutive microsatellite markers in HLA class I were associated with EBV+cHL. These markers are in LD with HLA-A*01 and HLA-A*02 alleles and subsequent studies provided evidence that HLA-A*01:01 is associated with an increased and A*02:01 with a decreased risk of EBV+cHL (60, 61). The effects of these two alleles are independent and result in A*01:01 and A*02:01 homozygotes having an almost 10-fold difference in odds of developing EBV+cHL (62). The only GWAS to include cases typed by EBV status in the discovery analysis (63), identified independent associations between two HLA class I SNPs (rs2734986 and rs6904029) and EBV+cHL; the effects of these SNPs could be accounted for by the effects of A*01 and A*02 alleles (63). Thus, several lines of evidence suggest associations between EBV+cHL and these alleles, at least in white populations (Table 1). In a study of Chinese cHL cases, Huang et al (2012) reported that the frequency of A*02:07, an allele more common in the Chinese population than in Whites, was increased in patients with EBV+cHL but decreased in patients with EBV-cHL relative to controls (64).

Subsequent case-control studies have identified further associations specific to EBV+cHL. In unadjusted analyses, Huang et al (2012) detected a significant increase in B37, an allele in LD with A1 (65). Johnson et al (2015) confirmed an association with B*37:01, which was significant after adjusting for the effects of other HLA alleles, thus demonstrating that this association is independent of A*01:01 (66). In addition, Huang et al (2012) reported an increased disease risk in association with the rarer DR10 allele. Although an increased frequency of DR10 was detected by Johnson et al (2015), case-control differences were not significant and a larger sample size will be required to validate this association. In allele selection regression modelling, Johnson et al (2015) detected additional associations with DRB1*15:01 and DPB1*01:01. Both alleles were associated with a decreased disease risk, suggesting that these alleles could have protective effects; interestingly, this association between DRB1*15:01 and EBV+cHL is in the opposite direction from that described below for EBV-cHL.

Analyses of associations with EBV-cHL are more difficult to interpret, largely due to issues with LD. Although only a small number of studies have stratified cases by EBV status, EBV-cHL cases constitute the majority of cHL cases in industrialised countries, are mainly NSHL, and usually occur in the young adult age group. Studies of NSHL and of young adult cHL are therefore likely to reflect associations with EBV-cHL (and vice versa). Taken together, such studies provide compelling evidence that DRB1*15:01, DQA1*01:02 and DQB1*06:02, three alleles that are in LD, are associated with an increased risk of EBV-cHL (55, 59, 65, 66). More recent studies have shown an association between DRB1*07 and decreased risk of EBV-cHL (65, 66). However, as described below, it is not clear whether these alleles are

biologically important or whether they are simply tagging another gene(s) or regulatory sequence that is critical in disease pathogenesis.

GWAS have consistently shown that the HLA class II SNP rs6903608, located ~15kb centromeric to DRA, is the variant most strongly associated with cHL (Table 3) (63, 67-70). The association is specific for EBV-cHL and within this patient subgroup there is a significant association with NSHL (63). Increased disease risk is associated with the 'C' allele, which is in LD with DRB1*15:01 (r^2 = 0.46, D' = 1), DQA1*02:01 and DQB1*06:02 (r^2 = 0.37, D' = 0.91). In a GWAS of NSHL, Cozen et al (2011) identified a 5-variant haplotype, incorporating rs6903608, which was superior to rs6903608 alone at predicting disease status (68). The haplotype containing the alleles associated with decreased disease risk, including the T variant at rs6903608, was associated with a 60% decreased risk of NSHL; all individuals carrying this haplotype were positive for DRB1*07:01. Although the above results suggested that rs6903608, and other class II SNPs, were simply tagging class II alleles, this is not borne out by logistic regression modelling.

Moutsianas et al (2011) imputed classical HLA alleles from SNP data in a study which largely included young adult cHL cases. In unconditional logistic regression analysis, DRB1*15:01 and DQB1*06:02 were associated with an increased and DRB1*07:01, DQA1*02:01 and DQB1*03:03 with a decreased disease risk, consistent with previous studies (71). Subsequent stepwise logistic regression with inclusion of selected SNPs and imputed HLA alleles revealed that most of the class II variation could be explained by the SNP rs6903608. The *DPB1* SNP rs2281389 and DQA1*02:01 allele contributed independent additional signals. DPB1 alleles were not imputed in this study and it was suggested that rs2281389 could be tagging the DPB1*03:01 allele, previously associated with increased disease risk. Since

DQA1*02:01 is in almost complete LD with DRB1*07:01 it was not possible to reliably determine which allele was contributing the additional signal. Johnson et al (2015) used a Bayesian variable selection method to analyse typed HLA alleles and three SNPs, including rs6903608, in a study with stratification of cases by EBV status (66). In the analysis of EBV-cHL without inclusion of the SNPs, DQB1*06:02 and DRB1*07:01 were associated an increased and decreased disease risk, respectively; however, following inclusion of the SNPs in the model, rs6903608 was the variant most strongly associated with disease risk. DQB1*03:01, DRB1*03:01 and B*15:01 contributed independent smaller effects (66). Taken together, these data suggest that rs6903608 is tagging a gene or regulatory sequence that is strongly associated with risk of EBV-cHL. The most consistent associations with *DRB1*, *DQA1* and *DQB1* alleles described above are likely to result from LD with this variant.

HLA associations with cHL as a whole largely reflect associations with EBV-stratified subgroups (65, 66). Urayama et al (2012) identified two HLA SNPs that were independently associated with 'total cHL' with no heterogeneity of effect by EBV status – the class I SNP rs2248462 at the *MICB* gene and the class II SNP rs2395185 at *HLA-DRA* (63). These SNPs were included in the variable selection modelling described by Johnson et al (2015), but neither was selected in the resultant models suggesting that their effects are explained by HLA alleles.

Biological implications of HLA Associations with EBV+cHL

The biological basis for the relationship between HLA alleles and disease risk is unknown. However, in the case of EBV+cHL, this is premised on the ability of HLA class I alleles to present EBV antigens to CD8+T-cells and to induce an effective

cytotoxic response. Interestingly, no EBV-specific T-cell responses restricted by the EBV+cHL risk allele HLA-A*01 have been identified to date (72). This suggests that HLA-A*01 is unable to elicit a cytotoxic response to effectively control EBV-infected B-cells, or HRS cells, although it can efficiently present peptides from other viruses (e.g. CMV and Influenza A (73)). On the other hand, HLA-A*02:01 alleles effectively present a number of EBV-derived peptides (74), suggesting that HLA-A*02-positive individuals may control EBV infection more effectively.

Given the breadth of EBV-specific responses (74), it is surprising that the lack of a single allele has such a large effect upon risk of EBV+cHL, perhaps suggesting that the critical responses are directed against a small number of antigens, such as the EBV latency II antigens expressed by HRS cells. An alternative hypothesis is that HLA-A*01 educes inhibitory effects that influence responses restricted through other HLA alleles. The elevated risk of disease seen in HLA-A*01:01/A*02:01 heterozygotes compared to HLA-A*02:01 homozygotes (62) provides some support for the latter model.

Similarly, the biological mechanism underlying HLA class II allele associations may be due to differing capacities in antigen presentation to CD4+T-cells. CD4+T-cells are primarily known for their helper roles; however, a cytotoxic role and the ability to be poly-functional, i.e. simultaneously produce multiple cytokines and effect cytotoxic function, have also been demonstrated, particularly in response to viral infection (75-77). The specificity, breadth, poly-functional capacity and timing of the CD4+T-cell response may all be important for disease development. Indeed, decreased EBNA-1 specific CD4+T-cell responses have previously been documented for EBV+cHL cases (78). Variation in HLA alleles might therefore result in modulation of CD8+T-cell activation and the recruitment and functionality of the CD4+T-cell population. It

is possible that different peptide-HLA combinations modulate the composition of the cHL associated infiltrate leading to the observed heterogeneity of cHL, survival of HRS cells and prevention of tumour suppression.

HLA Associations with Non-Hodgkin Lymphoma

NHL encompasses a heterogeneous group of cancers, 85‑90% of which arise from B-cells. It is the sixth most common cancer in the U.S. and the fifth most common in the UK. Together, DLBCL and FL, account for approximately 60% of all NHL cases. CLL/SLL is the third most common accounting for around 12% of cases (79). Although immune suppression and autoimmune disorders have been identified as risk factors for NHL, host characteristics also clearly play a role in this disease and its various subtypes (80).

Similar to cHL, HLA is implicated in NHL susceptibility with early studies showing down-regulation of the expression of HLA class II antigens (81). Several small studies have reported links between HLA class II alleles and NHL with conflicting results that may be attributable to small sample size, the combined analysis of various NHL subtypes, and/or analysis of different ethnic groups (82-84). In contrast, recent analyses of individual subtypes provide more consistent results.

Diffuse Large B-Cell Lymphoma

DLBCL itself encompasses a biologically and clinically diverse set of diseases. Although the WHO classification system defines more than a dozen subtypes, the majority of cases are classified as DLBCL, not otherwise specified (NOS). Gene expression profiling has divided DLBCL-NOS into germinal-centre B-cell like (GCB) and activated peripheral B-cell like (ABC) with 15% of cases remaining unclassifiable

(85). These groups have different molecular aberrations and clinical outcomes and may have different aetiologies.

Candidate gene studies and GWAS have identified associations between risk of DLBCL and immune-regulatory genes (86). A large, collaborative study from the International Lymphoma Epidemiology Consortium provided evidence that the *TNF promoter* polymorphism (TNF G-308A), which is thought to increase TNF-α expression resulting in inflammation, is strongly associated with DLBCL in white populations (87). This SNP is located within the HLA class III region, between the class I gene *HLA-B* and the class II gene *HLA-DR*. The common 8.1 ancestral haplotype includes the TNF-308A allele (HLA-A1-B8-TNF-308A-DR3-DQ2) suggesting that TNF associations may be confounded by linkage with HLA alleles. However, Abdou et al (2010) demonstrated that the TNF-A allele was associated with increased risk of DLBCL even in the absence of the 8.1 haplotype (88). Their results also suggested that the B*08 allele was independently associated with disease risk. In a recent GWAS meta-analysis, including 3,857 cases and 7,666 controls of European ancestry, 19 SNPs reached genome-wide significance and 134 were suggestive of significance (89). Of these 153 SNPs, 123 mapped to the HLA region. The strongest HLA signal was at rs2523607, located near *HLA-B* (Table 4). Following imputation of classical HLA alleles, only the HLA-B*08 allele and rs2523607 variant achieved genome-wide significance ($p < 5 \times 10^{-8}$). These two markers are in LD ($r^2 = 0.91$) and after adjustment for the effect of HLA-B*08, the association at rs2523607 was attenuated; consistent with earlier observations, these data suggest that HLA-B*08 is the HLA allele most strongly associated with DLBCL. Smaller GWAS identified weak associations with the class I SNP rs6457327 and class II SNP rs10484561, two SNPs that are strongly associated with FL (see below)

(90, 91) and the association at rs10484561 was corroborated in the above GWAS meta-analysis ($p = 1.5 \times 10^{-4}$). Two genotyping studies of DLBCL patients from China and Western Siberia reported evidence of an association at rs2647012, a further SNP associated with FL (91-93) Smaller HLA typing studies have documented associations with: DRB1*01 and B*51 (risk) and DRB1*15 and C*03 (protection) in European populations (94); and B*51, DRB1*09 and DQB1*03 and the B*51-DRB1*09-DQB1*03 haplotype (risk) in Korean patients (82).

Additionally, HLA class I and class II alleles have been associated with clinical outcome. Initial reports suggested that HLA-DR2 (DRB1*15 and *16) negative patients had a poorer progression free survival (PFS) and overall survival (OS) (95). Later studies demonstrated associations between B*18, B*44, Cw*07 and DRB1*04 and shorter PFS and OS (94, 96). HLA-B*18 and B*44 share overlapping peptide binding specificity and form part of the B44 supertype. When all alleles comprising this supertype (B*18, B*37, B*40, B*41, B*44, B*45, B*50) were analysed as a group, the supertype also showed poorer PFS and OS (94).

Follicular Lymphoma

FL is the second most common subtype of NHL, comprising up to 30% of all NHL cases worldwide. Although FL has a relatively long median survival, 8–10 years, more than 20% of FLs transform to an aggressive lymphoma with a poor clinical outcome. A defining feature of FL is the presence of the t(14:18) translocation, which results in dysregulation of bcl-2 expression. Akin to cHL and DLBCL, a hereditary component is proposed based upon ethnic variation (rarer in Asian populations) and an increased risk of FL in first-degree relatives (16, 17).

The first GWAS to include samples from FL patients identified an association with SNPs within the HLA class I region. Disease risk was associated with rs6457327 and neighboring SNPs in a 26 kb block of sequence that overlaps the *STG* gene *(C6orf15)* and is near psoriasis susceptibility region 1 *(PSORS1)* (90). This region is close to *HLA-C* and a later analysis showed that the 'A' allele at rs6457327 is associated with C*07:01 and B*07:02 (D' = 0.93 and 1.0, respectively); however, there is no evidence that these alleles account for the effect of the SNP. Subsequent studies confirmed association with rs6457327 (97, 98); in addition, the risk allele 'A' and associated genotypes 'AA' and 'AC' were associated with an increased risk of, and, shorter time to, transformation as well as inferior OS (91, 97, 98).

In a follow-up GWAS, Conde et al (2010) reported further associations in a region including the *HLA-DR* and *HLA-DQ* genes (Table 4) (99). The SNPs with the most significant p-values were rs10484561 and rs7755224, two SNPs in complete LD telomeric to the *HLA-DQB1* gene. Imputation of surrounding SNPs revealed a further four associated SNPs in a 100 kb region of LD including *HLA-DQB1* and *HLA-DQA1*, although rs10484561 remained the strongest signal. One of these four SNPs, rs6457614, is a tag SNP for HLA-DQB1*05:01, and follow up analyses revealed an association between FL and the extended haplotype DRB1*01:01-DQA1*01:01-DQB1*05:01 (OR = 2.07). Subsequent HLA typing studies confirmed associations with DQB1*05:01 and DRB1*01:01, as well as an inverse association with DQB1*06 (96; 97). A further GWAS with a larger number of cases in the discovery analysis corroborated earlier findings and demonstrated more associations (91, 100). The strongest signal was at rs2647012, and this was independent of the association at rs10484561 located 960 base pairs in a centromeric direction.

Subsequent fine-mapping, haplotype and coalescence analyses suggested that these two SNPs are associated with two distinct haplotypes that have opposite effects on risk of FL (91). HLA typing of class I and II alleles by next generation sequencing identified a decreased FL risk in association with DRB1*15 and DQB1*06 alleles and showed that carriers of the DRB1*15:01-DQA1*02:01-DQB1*06:02 haplotype were all positive for the 'A' allele at rs2647012 (but not vice versa) (101). This study also extended HLA analyses to DPB1 alleles and identified a protective effect of DPB1*03:01. A GWAS meta-analysis identified a further SNP association at rs311722, located downstream of the *DPB1* genes, which was explained by the DPB1*03:01 association. The current model of FL therefore suggests that the class I SNP rs6457327, the haplotypes HLA-DRB1*01:01-DQA1*01:01-DQB1*05:01 and DRB1*15:01-DQA1*02:01-DQB1*06:02, and DPB1*03:01 are all independently associated with risk of FL.

To determine whether specific coding variants within HLA genes contribute to the above associations identified by GWAS and HLA allele typing, Foo et al (2013) imputed SNPs, classical HLA alleles and coding variants across the HLA region (102). They identified a hexa-allelic amino acid polymorphism at position 13 of the HLA-DR beta chain that showed a stronger association with FL than any other variant or SNP that was tested. From a possible six amino acids, two were classified as high risk (Tyr and Phe), two as low risk (Ser and Arg), and two as moderate risk (His and Gly). There was a 4.2-fold difference in risk between subjects carrying two alleles encoding high-risk amino acids and those carrying two alleles encoding low-risk amino acids. Further analysis revealed that the risk allele 'C' from rs10484561 tags haplotypes carrying the allele encoding the high risk Phe, whilst the protective allele 'T' from rs2647012 tags haplotypes carrying the allele encoding the low risk

Ser residue. The data suggest that the DRB1 polymorphism at position 13 is the primary driver of the above DR and DQ associations, thus advocating a HLA-DR antigen-driven mechanism in the pathogenesis of FL.

HLA alleles have also been investigated in relation to outcome. HLA-A*01 and the ancestral 8.1 haplotype (A*01-B*08-DR*03) have been associated with poorer OS and HLA-Bw4 and DRB1*13 with improved OS (96).

Chronic Lymphocytic Leukaemia /Small Lymphocytic Lymphoma

CLL is an indolent malignancy resulting from the accumulation of slowly proliferating CD5-positive neoplastic B-cells, which accounts for about a quarter of all leukaemias in Western countries. Small lymphocytic lymphoma has the same immunophenotype and is essentially the same disease but the malignant cells are in the lymph nodes and spleen rather than the blood. Around half of all CLLs have evidence of somatic hypermutation in their immunoglobulin heavy chain variable region genes (IGHVs), and IGHV mutation status is used to divide CLLs into two subgroups. Cases with mutated IGHV have a much better clinical outcome.

Early case-control and family-based HLA typing studies did not identify consistent associations with HLA alleles, most probably because of small case numbers. However, later studies provide support for some of the earlier findings, and the work of Dorak et al (1996) focused attention on *DRB4* and associated haplotypes (103). Machulla et al (2001) compared the frequency of HLA class I and II alleles in 101 cases and 157 controls using molecular and serological typing methods (104). PCR-based typing revealed a striking increase in DRB4*01:03 alleles among cases (RR = 2.74, p = 0.0025); higher frequencies of DRB1*04:01, DQB1*03:02 and DPB1*03:01 and a lower frequency of DQB1*02:02 alleles were also detected. None of these

differences remained significant after correction for multiple testing. The haplotype DRB1*04:01-DRB4*01:03 was also increased and CLL-specific LD was observed between these alleles. The association with DQB1*03:02 was thought to result from LD with DRB1*04:01-DRB4*01:03 on the ancestral haplotype HLA-Cw3-B62-DR4-DQ8, which is frequent in white populations. There was no evidence of LD between DPB1*03:01 and the other risk alleles.

In by far the largest study to date, Gragert et al (2014) compared HLA genotypes from 3491 US white, 397 African-American and 90 Hispanic CLL patients with 50,000 controls from each of these populations (105). Cases were ascertained through the National Bone Marrow Program and this is likely to have resulted in a bias towards patients with more severe disease and younger age. A large number of allele and haplotype associations were identified, some of which had been previously reported. In Whites, 12 alleles were decreased and 16 increased in frequency with DQB1*05:04 (OR = 5.62) and DRB1*04:03 (OR = 0.67) having the highest and lowest ORs, respectively; despite the size of the study, effect sizes of many of the associations were modest with upper or lower bounds of the 95% confidence intervals close to one. The DRB4*01:01 allele and haplotype DRB4*01:01-DRB1*07:01-DQB1*03:03 were strongly associated with increased disease risk in all study populations, i.e. Whites, African Americans and Hispanics. Both DRB4*01:01 and this extended haplotype are much rarer in African Americans than whites (allele frequency 18.2% versus 29.7%, haplotype frequency 0.3% versus 3.4%), and it is therefore possible that this contributes to the lower incidence of CLL in African Americans. C*04:01, an allele found at relatively high frequency in African Americans, was associated with decreased disease risk but only in the African American population. Associations were confirmed with A*02:01 (OR = 1.2), which

had been identified following imputation of HLA alleles in a GWAS (106), and with the haplotype A*02:01-B*15:01-DRB1*04:01 (OR = 1.41). A further haplotype association with A*01:01-C*07:01-B*08:01-DRB1*03:01-DQB1*02:01 was also confirmed at higher resolution (OR = 0.83). In addition, homozygosity at all three class I alleles was shown to significantly increase risk of CLL (105). *DPB1* associations were not investigated.

GWAS have identified few associations between HLA SNPs and CLL. Di Bernardo et al (2013) reported an association with the HLA class I SNP rs6904029, which is also associated with EBV+cHL, and this led to the identification of the association with A*02:01 described above (Table 4) (106). Following analysis of familial CLL cases included in a larger GWAS, Slager et al (2011) identified association with five SNPs in a region of HLA class II that includes the DQA1 and DRB5 genes (Figure 1). Following conditional analyses of the five SNPs, only rs674313 retained significance thus suggesting that these SNPs are tagging the same locus (107). Imputation of HLA alleles in this region was not performed. In contrast to the findings for cHL and FL described above, the strongest signals in GWAS were from non-HLA SNPs. At least 30 common risk variants have been identified thus far, with several in proximity to genes involved in apoptosis or telomere function (108, 109). These data are consistent with the idea that co-inheritance of multiple low risk variants is likely to explain the heritability in CLL/SLL.

Summary and Conclusions

Recent GWAS and HLA typing studies provide consistent and reliable evidence linking specific HLA polymorphisms with subtypes of human lymphoma. Clustering of SNPs associated with cHL, FL and CLL within the HLA class II region is perhaps

one of the most striking observations to emerge from these studies. However, associations within this region, which encompasses the *HLA-DRA*, *DRB1* and *DQB1* genes, are distinct for each of the three diseases (Figure 1). Stratification of cHL cases by EBV status further demonstrates that the association with the *DRA* SNP rs6903608 is particular to EBV-cHL. In contrast, EBV+cHL is strongly associated with the class I allele HLA-A1*01, an allele first linked to HL over forty years ago. Several alleles are associated with multiple lymphoma subtypes. Associations with DPB1*03:01 and DRB1*15:01 are described for FL and cHL; however, associations with FL are in the opposite direction from those reported for total cHL and EBV-cHL although the DRB1*15:01 association is in the same direction to that recently reported for EBV+cHL. Likewise, the HLA-A*02:01 allele is associated with increased risk of CLL but reduced risk of cHL. In contrast, the class II SNPs rs10484561 and rs2647012 are associated with increased risk of both FL and DLBCL, albeit with less significance for DLBCL, possibly implying that the DLBCL association is confined to a subset of DLBCLs with features more closely related to FL. Although the evidence for specific HLA associations with lymphoma subtypes is compelling, understanding the biology underlying these complex associations awaits further study.

Acknowledgements

We thank Dr. James MacKay, International Agency for Research on Cancer, for helpful advice.

Figure Legends

Figure 1. Location of lymphoma-associated SNPs on chromosome 6

Schematic of chromosome 6 showing relative positions of SNPs associated with risk of classical Hodgkin lymphoma (cHL, blue), diffuse large B cell lymphoma (DLBCL, grey), follicular lymphoma (FL, red) and chronic lymphocytic leukaemia (CLL, purple). SNPs associated with more than one disease are highlighted in green (FL and DLBCL) and brown (CLL and cHL).

Reference List

(1) Jaffe ES, Harris NL, Stein H, Campo E, Pileri SA, Swerdlow S. Introduction and overview of the classification of the lymphoid neoplasms. In Swerdlow SH, Campo E, Harris NL, et al., eds. WHO classification of tumours of haematopoietic and lymphoid tissues. 4th ed. Lyon, International Agency for Research on Cancer, 2008: 158-166.

(2) Hors J, Steinberg G, Andrieu J, et al. HLA genotypes in familial Hodgkin's disease. Excess of HLA identical affected sibs. *European Journal of Cancer* 1980: **16**: 809-815.

(3) Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. *Am J Hum Genet* 1987: **40**: 1-14.

(4) Yamamoto W, Nakamura N, Tomita N, et al. Human leukocyte antigen-DR expression on flow cytometry and tumor-associated macrophages in diffuse large B-cell lymphoma treated by rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone therapy: retrospective cohort study. *Leuk Lymphoma* 2014: **55**: 2721-2727.

(5) Oudejans JJ, Jiwa NM, Kummer JA, et al. Analysis of major histocompatibility complex class I expression on Reed- Sternberg cells in relation to the cytotoxic T-cell response in Epstein-Barr virus-positive and -negative Hodgkin's disease. *Blood* 1996: **87**: 3844-3851.

(6) Gale RP, Cozen W, Goodman MT, Wang FF, Bernstein L. Decreased chronic lymphocytic leukemia incidence in Asians in Los Angeles County. *Leukemia Research* 2000: **24**: 665-669.

(7) Glaser S, Lin R, Stewart S, et al. Epstein-Barr virus-associated Hodgkin's disease: epidemiologic characteristics in international data. *Int J Cancer* 1997: **70**: 375-382.

(8) Glaser SL, Hsu JL. Hodgkin's disease in Asians: incidence patterns and risk factors in population-based data. *Leukemia Research* 2002: **26**: 261-269.

(9) Kerzin-Storrar L, Faed MJ, MacGillivray JB, Smith PG. Incidence of familial Hodgkin's disease. *Br J Cancer* 1983: **47**: 707-712.

(10) Razis DV, Diamond HD, Craver LF. Familial Hodgkin's disease: its significance and implications. *Annals of Internal Medicine* 1959: **51**: 933-971.

(11) Grufferman S, Cole P, Smith PG, Lukes RJ. Hodgkin's disease in siblings. *New England Journal of Medicine* 1977: **296**: 248-250.

(12) Haim N, Cohen Y, Robinson E. Malignant lymphoma in first-degree blood relatives. *Cancer* 1982: **49**: 2197-2200.

(13) Mack TM, Cozen W, Shibata DK, et al. Concordance for Hodgkin's disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease. *New England Journal of Medicine* 1995: **332**: 413-419.

(14) Shugart YY, Hemminki K, Vaittinen P, Kingman A, Dong C. A genetic study of Hodgkin's lymphoma: an estimate of heritability and anticipation based on the familial cancer database in Sweden. *Human Genetics* 2000: **106**: 553-556.

(15) Chang ET, Smedby KE, Hjalgrim H, et al. Family history of hematopoietic malignancy and risk of lymphoma. *Journal of the National Cancer Institute* 2005: **97**: 1466-1474.

(16) Altieri A, Bermejo JL, Hemminki K. Familial risk for non-Hodgkin lymphoma and other lymphoproliferative malignancies by histopathologic subtype: the Swedish Family-Cancer Database. *Blood* 2005: **106**: 668-672.

(17) Wang SS, Slager SL, Brennan P, et al. Family history of hematopoietic malignancies and risk of non-Hodgkin lymphoma (NHL): a pooled analysis of 10 211 cases and 11 905 controls from the International Lymphoma Epidemiology Consortium (InterLymph). *Blood* 2007: **109**: 3479-3488.
(18) Goldin LR, Bjorkholm M, Kristinsson SY, Turesson I, Landgren O. Elevated risk of chronic lymphocytic leukemia and other indolent non-Hodgkin's lymphomas among relatives of patients with chronic lymphocytic leukemia. *Haematologica* 2009: **94**: 647-653.
(19) Cerhan JR, Kricker A, Paltiel O, et al. Medical history, lifestyle, family history, and occupational risk factors for diffuse large B-cell lymphoma: the interlymph non-Hodgkin lymphoma subtypes project. *JNCI Monographs* 2014: **2014**: 15-25.
(20) Linet MS, Vajdic CM, Morton LM, et al. Medical history, lifestyle, family history, and occupational risk factors for follicular lymphoma: the interlymph non-Hodgkin lymphoma subtypes project. *JNCI Monographs* 2014: **2014**: 26-40.
(21) Slager SL, Benavente Y, Blair A, et al. Medical history, lifestyle, family history, and occupational risk factors for chronic lymphocytic leukemia/small lymphocytic lymphoma : the interlymph non-Hodgkin lymphoma subtypes project. *JNCI Monographs* 2014: **2014**: 41-51.
(22) Yuille MR, Matutes E, Marossy A, Hilditch B, Catovsky D, Houlston RS. Familial chronic lymphocytic leukaemia: a survey and review of published studies. *Br J Haematol* 2000: **109**: 794-799.
(23) Marsh SGE, Parham P, Barber LD. The HLA FactsBook. Academic Press, London, UK, 2000.
(24) Klein J, Sato A. The HLA System. *New England Journal of Medicine* 2000: **343**: 702-709.
(25) Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. *BMC Immunol* 2008: **9**: 1-15.
(26) Aldinucci D, Gloghini A, Pinto A, De Filippi R, Carbone A. The classical Hodgkin's lymphoma microenvironment and its role in promoting tumour growth and immune escape. *The Journal of Pathology* 2010: **221**: 248-263.
(27) World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. Lyon, International Agency for Research on Cancer, 2001.
(28) Weiss L, Movahed L, Warnke R, Sklar J. Detection of Epstein-Barr viral genomes in Reed-Sternberg cells of Hodgkin's disease. *New England Journal of Medicine* 1989: **320**: 502-506.
(29) Wu TC, Mann RB, Charache P, et al. Detection of EBV gene expression in Reed-Sternberg cells of Hodgkin's disease. *Int J Cancer* 1990: **46**: 801-804.
(30) Jarrett RF, Krajewski AS, Angus B, et al. The Scotland and Newcastle epidemiological study of Hodgkin's disease: impact of histopathological review and EBV status on incidence estimates. *J Clin Pathol* 2003: **56**: 811-816.
(31) Jarrett RF, Gallagher A, Jones DB, et al. Detection of Epstein-Barr virus genomes in Hodgkin's disease: relation to age. *J Clin Pathol* 1991: **44**: 844-848.
(32) Amiel J. Study of the leukocyte phenotypes in Hodgkin's disease. In Terasaki P, ed. Histocompatibility Testing.Copenhagen, 1967: 79-81.
(33) van Rood JJ, van Leeuwen A, Schippers A, Balner H. Human Histocompatibility Antigens in Normal and Neoplastic Tissues. *Cancer Res* 1968: **28**: 1415-1422.
(34) Forbes J, Morris P. Leukocyte antigens in Hodgkin's disease. *The Lancet* 1970: **296**: 849-851.
(35) Harris R, Lawler SD, Oliver R. The HLA system in acute leukaemia and Hodgkin's disease. *Br Med Bull* 1978: **34**: 301-304.
(36) Falk JA, Osoba D. The association of the human histocompatibility system with Hodgkin's disease. *International Journal of Immunogenetics* 1974: **1**: 53-61.
(37) Daussett J, Hors J. Some contributions of the HL-A complex to the genetics of human diseases. *Transplant Reviews* 1975: **22**: 44-74.

(38) Svejgaard A, Platz P, Ryder LP, Nelsen L, Thomsen M. HL-A and disease associations - a survey. *Transplant Reviews* 1975: **22**: 3-43.
(39) Kissmeyer-Nielson F, Kjerbye K, Lamm L. HL-A in Hodgkin's disease. III. A prospective study. *Transplant Reviews* 1975: **22**: 168-174.
(40) Hansen J, Young C, Whitsett C, et al. HLA and MLC typing in patients with Hodgkin's disease. *Progress in Clinical and Biological Research* 1977: **16**: 217-227.
(41) Hornmark-Stenstam B, Landberg T, Low B. HLA antigens in Hodgkin's disease of very long survival. *Acta Radiologica: Oncology, Radiation, Physics, Biology* 1978: **17**: 283-288.
(42) Hafez M, el-Wehedy G, el-Kholy N, el-Shawaf I, Mahmoud M, Gamil T. The value of HLA phenotypes in the prognosis of Hodgkin's lymphoma. *Int J Cancer* 1985: **36**: 19-22.
(43) Hors J, Daussett J. HLA and susceptability to Hodgkin's disease. *Immunology Reviews* 1983: **70**: 167-182.
(44) Falk J, Osoba D. HL-A antigens and survival in Hodgkin's disease. *The Lancet* 1971: **298**: 1118-1121.
(45) Bjorkholm M, Holm G, Johansson B, Melstedt H, Moler E. A prospective study of HL-A antigen phenotypes and lymphocyte abnormalities in Hodgkin's disease. *Tissue Antigens* 1975: **6**: 247-256.
(46) Marshall W, Barnard J, Buehier S, Crumley J, Larsen B. HLA in familial Hodgkin's disease. Results and a new hypothesis. *Int J Cancer* 1977: **19**: 450-455.
(47) Greene M, McKeen E, Li F, Blatner W, Fraumen JJ. HLA antigens in familial Hodgkin's disease. *Int J Cancer* 1979: **23**: 777-780.
(48) Conte R, Lauria F, Zuccheli P. HLA in familial Hodgkin's disease. *Journal of Immunogenetics* 1983: **10**: 251-255.
(49) Berberich FR, Berberich MS, King MC, Engleman EG, Grumet FC. Hodgkin's disease susceptibility: Linkage to the HLA locus demonstrated by a new concordance method. *Human Immunology* 1983: **6**: 207-217.
(50) Robertson S, Lowman J, Grufferman S, et al. Familial Hodgkin's disease. A clinical and laboratory investigation. *Cancer* 1987: **59**: 1314-1319.
(51) Osoba D, Falk JA, Sousan P, Ciampi A, Till J. The prognostic value of HLA phenotypes in Hodgkin's disease. *Cancer* 1980: **46**: 1825-1832.
(52) Oza AM, Tonks S, Lim J, et al. A clinical and epidemiological study of human leukocyte antigen-DPB alleles in Hodgkin's disease. *Cancer Res* 1994: **54**: 5101-5105.
(53) Tonks S, Oza AM, Lister TA, Bodmer JG. Association of HLA-DPB with Hodgkin's disease. *The Lancet* 1992: **340**: 968-969.
(54) Alexander FE, Jarrett RF, Cartwright RA, et al. Epstein-Barr virus and HLA-DPB1-*0301 in young adult Hodgkin's disease: evidence for inherited susceptibility to Epstein-Barr virus in cases that are EBV+ve. *Cancer Epidemiol Biomarkers Prev* 2001: **10**: 705-709.
(55) Klitz W, Aldrich CL, Fildes N, Horning SJ, Begovich AB. Localization of predisposition to Hodgkin disease in the HLA class II region. *Am J Hum Genet* 1994: **54**: 497-505.
(56) Bodmer JG, Tonks S, Oza AM, Lister TA, Bodmer W. HLA-DP based resistance to Hodgkin's disease. *Lancet* 1989: **1**: 1455-1456.
(57) Taylor GM, Gokhale DA, Crowther D, et al. Increased frequency of HLA-DPB1*0301 in Hodgkin's disease suggests that susceptibility is HVR-sequence and subtype-associated. *Leukemia* 1996: **10**: 854-859.
(58) Taylor GM, Gokhale DA, Crowther D, et al. Further investigation of the role of HLA-DPB1 in adult Hodgkin's disease (HD) suggests an influence on susceptibility to different HD subtypes. *Br J Cancer* 1999: **80**: 1405-1411.
(59) Harty LC, Lin AY, Goldstein AM, et al. HLA-DR, HLA-DQ, and TAP genes in familial Hodgkin disease. *Blood* 2002: **99**: 690-693.

(60) Niens M, Jarrett RF, Hepkema B, et al. HLA-A*02 is associated with a reduced risk and HLA-A*01 with an increased risk of developing EBV+ Hodgkin lymphoma. *Blood* 2007: **110**: 3310-3315.
(61) 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. *The Lancet* 2005: **365**: 2216-2224.
(62) Hjalgrim H, Rostgaard K, Johnson PCD, et al. HLA-A alleles and infectious mononucleosis suggest a critical role for cytotoxic T-cell response in EBV-related Hodgkin lymphoma. *PNAS* 2010: **107**: 6400-6405.
(63) Urayama KY, Jarrett RF, Hjalgrim H, et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein–Barr virus status–defined subgroups. *Journal of the National Cancer Institute* 2012: **104**: 240-253.
(64) Huang X, Hepkema B, Nolte I, et al. HLA-A*02:07 Is a protective allele for EBV negative and a susceptibility allele for EBV positive classical Hodgkin lymphoma in China. *PLoS ONE* 2012: **7**: e31865.
(65) Huang X, Kushekhar K, Nolte I, et al. HLA associations in classical Hodgkin lymphoma: EBV status matters. *PLoS ONE* 2012: **7**: e39986.
(66) 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: n/a.
(67) Enciso-Mora V, Broderick P, Ma Y, et al. A genome-wide association study of Hodgkin's lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21 and 10p14 (GATA3). *Nat Genet* 2010: **42**: 1126-1130.
(68) Cozen W, Li D, Best T, et al. A genome-wide meta-analysis of nodular sclerosing Hodgkin lymphoma identifies risk loci at 6p21.32. *Blood* 2011: **119**: 469-475.
(69) Cozen W, Timofeeva MN, Li D, et al. A meta-analysis of Hodgkin lymphoma reveals 19p13.3 TCF3 as a novel susceptibility locus. *Nat Commun* 2014: **5**.
(70) Frampton M, da Silva Filho MI, Broderick P, et al. Variation at 3p24.1 and 6q23.3 influences the risk of Hodgkin's lymphoma. *Nat Commun* 2013: **4**.
(71) Moutsianas L, Enciso-Mora V, Ma YP, et al. Multiple Hodgkin lymphoma-associated loci within the HLA region at chromosome 6p21.3. *Blood* 2011.
(72) Brennan RM, Burrows SR. A mechanism for the HLA-A*01-associated risk for EBV+ Hodgkin lymphoma and infectious mononucleosis. *Blood* 2008: **112**: 2589-2590.
(73) Currier JR, Kuta EG, Turk E, et al. A panel of MHC class I restricted viral peptides for use as a quality control for vaccine trial ELISPOT assays. *Journal of Immunological Methods* 2002: **260**: 157-172.
(74) Taylor GS, Long HM, Brooks JM, Rickinson AB, Hislop AD. The Immunology of Epstein-Barr Virus–Induced Disease. *Ann Rev Immunol* 2014: **33**: 787-821.
(75) Zaunders JJ, Dyer WB, Wang B, et al. Identification of circulating antigen-specific CD4+ T lymphocytes with a CCR5+, cytotoxic phenotype in an HIV-1 long-term nonprogressor and in CMV infection. *Blood* 2004: **103**: 2238-2247.
(76) van Leeuwen EMM, Remmerswaal EBM, Vossen MTM, et al. Emergence of a CD4+CD28– granzyme B+, cytomegalovirus-specific T cell subset after recovery of primary cytomegalovirus infection. *J Immunol* 2004: **173**: 1834-1841.
(77) Makedonas G, Betts M. Polyfunctional analysis of human T cell responses: importance in vaccine immunogenicity and natural infection. *Springer Semin Immun* 2006: **28**: 209-219.
(78) Heller KN, Arrey F, Steinherz P, et al. Patients with Epstein Barr virus-positive lymphomas have decreased CD4+ T-cell responses to the viral nuclear antigen 1. *Int J Cancer* 2008: **123**: 2824-2831.
(79) Jaffe ES. The 2008 WHO classification of lymphomas: implications for clinical practice and translational research. *ASH Education Program Book* 2009: **2009**: 523-531.

(80) Morton LM, Slager SL, Cerhan JR, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes: the interLymph non-Hodgkin lymphoma subtypes project. *JNCI Monographs* 2014: **2014**: 130-144.
(81) Guy K, Krajewski AS, Dewar AE. Expression of MHC class II antigens in human B-cell leukaemia and non-Hodgkin's lymphoma. *Br J Cancer* 1986: **53**: 161-173.
(82) Choi HB, Roh SY, Choi EJ, et al. Association of HLA alleles with non-Hodgkin's lymphoma in Korean population. *Int J Hematol* 2008: **87**: 203-209.
(83) Nathalang O, Tatsumi N, Hino M, et al. HLA class II polymorphism in Thai patients with non-Hodgkin's lymphoma. *European Journal of Immunogenetics* 1999: **26**: 389-392.
(84) Al-Tonbary Y, Abdel-Razek N, Zaghloul H, Metwaly S, El-Deek B, El-Shawel R. HLA class II polymorphism in Egyptian children with lymphomas. *Hematology* 2004: **9**: 139-145.
(85) Sehn LH, Gascoyne RD. Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. *Blood* 2014: **125**: 22-32.
(86) Cerhan JR, Ansell SM, Fredericksen ZS, et al. Genetic variation in 1253 immune and inflammation genes and risk of non-Hodgkin lymphoma. *Blood* 2007: **110**: 4455-4463.
(87) 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. *The Lancet Oncology* 2006: **7**: 27-38.
(88) 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.
(89) 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.
(90) Skibola CF, Bracci PM, Halperin E, et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. *Nat Genet* 2009: **41**: 873-875.
(91) 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.
(92) Weiner AS, Berezina OV, Ovchinnikov VS, et al. Role of polymorphic loci in HLA-region rs2647012 and rs805288 in the development of non-Hodgkin's malignant lymphomas in western Siberia. *Bull Exp Biol Med* 2014: **157**: 180-183.
(93) Qiao Y, Zhou Y, Wu C, et al. Risk of genome-wide association study-identified genetic variants for non-Hodgkin lymphoma in a Chinese population. *Carcinogenesis* 2013: **34**: 1516-1519.
(94) Alcoceba M, Sebastian E, Marin L, et al. HLA specificities are related to development and prognosis of diffuse large B-cell lymphoma. *Blood* 2013: **122**: 1448-1454.
(95) Juszczynski P, Kalinka E, Bienvenu J, et al. Human leukocyte antigens class II and tumor necrosis factor genetic polymorphisms are independent predictors of non-Hodgkin lymphoma outcome. *Blood* 2002: **100**: 3037-3040.
(96) 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. *ScientificWorldJournal* 2011: **11**: 2062-2070.
(97) Wrench D, Leighton P, Skibola CF, et al. SNP rs6457327 in the HLA region on chromosome 6p is predictive of the transformation of follicular lymphoma. *Blood* 2011: **117**: 3147-3150.
(98) Berglund M, Enblad G, Thunberg U. SNP rs6457327 is a predictor for overall survival in follicular lymphoma as well as survival after transformation. *Blood* 2011: **118**: 4489.
(99) 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.
(100) Skibola CF, Conde L, Foo JN, et al. A meta-analysis of genome-wide association studies of follicular lymphoma. *BMC Genomics* 2012: **13**: 516.

(101) 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.
(102) Foo J-N, Smedby K-E, Akers N-K, et al. Coding variants at hexa-allelic amino acid 13 of HLA-DRB1 explain independent SNP associations with follicular lymphoma risk. *Am J Hum Genet* 2013: **93**: 167-172.
(103) Dorak M, Machulla HKG, Hentschel M, Mills K, Langner J, Burnett A. Influence of the major histocompatibility complex on age at onset of chronic lymphoid leukaemia. *Int J Cancer* 1996: **65**: 134-139.
(104) Machulla HKG, Muller LP, Schaaf A, Kujat G, Schonermarck U, Langner J. Association of chronic lymphocytic leukemia with specific alleles of the HLA-DR4:DR53:DQ8 haplotype in German patients. *Int J Cancer* 2001: **92**: 203-207.
(105) Gragert L, Fingerson S, Albrecht M, Maiers M, Kalaycio M, Hill BT. Fine-mapping of HLA associations with chronic lymphocytic leukemia in US populations. *Blood* 2014: **124**: 2657-2665.
(106) Di Bernardo MC, Broderick P, Harris S, et al. Risk of developing chronic lymphocytic leukemia is influenced by HLA-A class I variation. *Leukemia* 2013: **27**: 255-258.
(107) Slager SL, Rabe KG, Achenbach SJ, et al. Genome-wide association study identifies a novel susceptibility locus at 6p21.3 among familial CLL. *Blood* 2011: **117**: 1911-1916.
(108) 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.
(109) 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.

Table 1. HLA class I associations with HL

Study	Ethnicity or Country	Study Design	Study Subjects	Typing Method	Main Findings		
					HL subgroup association	HLA Allele	Direction of risk
Falk et al (1971)	Canada	Case-control	122 cases 122 controls	Serological	HL	**A*01** **B*05** **B*08** A*03	**Increased** **Increased** **Increased** Decreased
Falk et al (1974)	Canada	Case-control	115 cases 200 controls	Serological	HL (<1 year post diagnosis) HL (<1 year post diagnosis) HL (>5 years post diagnosis)	A*03 A*11 **B*08**	Decreased Decreased **Increased**
Svejgaard et al (1975)[†]	Various	Meta-analysis of 17 studies	1500 cases 5400 controls	Serological	HL	**A*01** **A*09** **A*10** **A*11** **B*05** **B*07** **B*08** **B*10** **B*18**	**Increased** **Increased** **Increased** **Decreased** **Increased** **Decreased** **Increased** **Decreased** **Increased**
Kissmeyer-Nielson et al (1975)	Denmark	Case-control	201 cases 562 controls	Serological	HL (MCHL, males >30 years)	**A*01** B*08	**Increased** Increased
Bjorkman et al (1975)	Sweden	Case-control	65 cases 100 controls	Serological	HL (>40 years) HL (<40 years)	**B*08** **B*12**	**Increased** **Increased**
Hansen et al (1977)	United States	Case-	137 cases	Serological	HL	A*01	Increased

		control	855 controls			A*33	Decreased
Marshall et al (1977)	Canada	Family-based	7 cases 1277 relatives	Serological	HL	B*18	Increased
Hornmark-Stenstam et al (1978)	Sweden	Case-control	10 long term survivors 30 new cases 1263 controls	Serological	HL (long term survivors) HL, NSHL	**A*28** **B*18**	**Increased** **Increased**
Greene et al (1979)	United States	Family-based	411 unrelated cases 629 controls 13 families with 21 cases	Serological	HL	B*35 B*37	Increased Increased
Osoba et al (1980)	Canada	Case series	79 cases	Serological	HL (>40 years, advanced disease, LDHL, MCHL)	Aw19 (A*29,30,31,32,33,34)	Increased
Conte et al (1983)	Italy	Family-based	4 cases 15 relatives	Serological	NSHL	B*18	Increased
Hafez et al (1985)	Africa	Case-control	52 cases 234 controls	Serological	HL	**A*01**	**Increased**
Niens et al (2007)	Netherlands	Case-control	101 cases 59 controls	Molecular	EBV+cHL	**A*01** **A*02**	**Increased** **Decreased**
Hjalgrim et al (2010)	Scandinavia Northern United Kingdom	Case series	278 EBV+cases 656 EBV-cases	Molecular	EBV+cHL	**A*01** **A*02**	**Increased** **Decreased**
Huang et al (2012)	Netherlands	Case-control	338 cases 7754 controls	Molecular Serological	**EBV+cHL** **cHL**	**A*01** **A*02** **B*37** **B*05**	**Increased** **Decreased** **Increased** **Increased**
Huang et al (2012)	China	Case-control	161 cases 119 controls	Molecular	**EBV+cHL** EBV-cHL	**A*02:07** **A*02:07**	**Increased** **Decreased**
Johnson et al (2015)*	Northern United Kingdom	Case-control	503 cases 347 controls	Molecular	EBV+cHL EBV-cHL	**A*01** **B*37** **B*15**	**Increased** **Increased** **Increased**

HLA class I associations by study for HL or HL subgroup. HL, Hodgkin lymphoma; cHL, classical Hodgkin lymphoma; EBV, Epstein-Barr virus; NSHL, nodular sclerosis Hodgkin lymphoma; MCHL, mixed cellularity Hodgkin lymphoma; LDHL, lymphocyte depleted Hodgkin lymphoma

Results highlighted in bold are significant, $p<0.05$

[†] includes subjects from Falk et al (1971) and Falk et al (1974)

*adjusted for effects of class II alleles; these cases were also included in Hjalgrim et al (2010) and Niens et al (2007)

Table 2. HLA class II associations with HL

Study	Ethnicity or Country	Study Design	Study Subjects	Typing Method	Main Findings		
					HL subgroup association	HLA Allele	Direction of risk
Robertson et al (1987)	United States	Family-based	1 family 4 cases	Serological	HL	DR5 (DR*11, DR*12)	Increased
Bodmer et al (1989)	United Kingdom	Case-control	86 cases 91 controls	Molecular	HL	DPw2 (DPB1*02:01)	Decreased
Tonks et al (1992)	White	Case-control	741 cases 686 controls	Molecular	HL	DPB1*02:01 **DPB1*03:01** **DPB1*04:01**	Decreased **Increased** **Decreased**
	Asian					**DPB1*04:01**	**Decreased**
Oza et al (1994)[†]	White	Case-control	741 cases 686 controls	Molecular	HL	**DPB1*03:01**	**Increased**
	Asian					**DPB1*04:01**	**Decreased**
Klitz et al (1994)	White United States	Case-control	196 cases 306 controls	Molecular	NSHL	**DRB1*15:01** **DQB1*06:02** **DRB1*11:04** **DQB1*03:03** **DPB1*03:01** **DPB1*13:01** **DPB1*11:01**	**Increased** **Increased** **Increased** **Decreased** **Increased** **Increased** **Decreased**
Taylor et al (1996)	White United Kingdom	Case-control	118 cases 92 controls	Molecular	HL	DPB1*03:01	Increased
Taylor et al (1999)~	White United Kingdom	Case-control	147 cases 183 controls	Molecular	HL	DPB1*03:01 DPB1*02:01 DPB1*10:01 DPB1*1101	Increased Decreased Increased Decreased
Alexander et al (2001)	United Kingdom	Case series	19 EBV+cases 84 EBV-cases	Molecular	EBV+cHL	DPB1*03:01	Increased

Harty et al (2002)	United States	Family-based	13 families 28 cases 69 unaffected	Molecular	HL	**DRB1*15:01** **DQB1*06:02**	**Increased** **Increased**
					NSHL	DQA1*01:02	Increased
Huang et al (2012)	Netherlands	Case-control	338 cases 7754 controls	Molecular Serological	cHL	**DR7**	**Decreased**‡
					EBV-cHL	DR5 **DR2**	Increased **Increased**‡
					EBV+cHL	**DR10**	**Increased**‡
Johnson et al (2015)*	Northern United Kingdom	Case-control	503 cases 347 controls	Molecular	EBV-cHL	**DRB1*03:01** **DQB1*03:03**	**Increased** **Increased**
					EBV+cHL	**DRB1*15:01** **DPB1*01:01**	**Decreased** **Decreased**

HLA class II associations by study for HL or HL subtype. HL, Hodgkin lymphoma; cHL, classical Hodgkin lymphoma; EBV, Epstein-Barr virus; NSHL, nodular sclerosis Hodgkin lymphoma

Results highlighted in bold are significant, $p<0.05$, ‡ $p<0.001$

† includes subjects from Tonks et al (1992)

~ includes subjects from Taylor et al (1996)

*adjusted for effects of other HLA alleles and selected SNPs

Table 3. HLA associations with HL identified in GWAS

Study	Country	Study Subjects in Discovery Set	Significant Findings					Candidate Gene
			Subgroup association	SNP	Minor Allele	Odds Ratio	p-value	
Enciso-Mora et al (2011) [†]	United Kingdom	589 cases 5199 controls	cHL	rs6903608*[α]	C	1.81	8.12×10^{-23}	HLA-DRA
Cozen et al (2012)	United States European origin	393 cases 3315 controls	NSHL	rs2858870*	G	0.4	1.69×10^{-8}	HLA-DRA
			NSHL	rs6903608*	C	1.6	3.52×10^{-10}	HLA-DRA
			NSHL	rs9268542	G	1.6	5.35×10^{-10}	HLA-DRA
			NSHL	rs9268528	G	1.6	1.19×10^{-9}	HLA-DRA
			NSHL	rs204999	G	0.5	1.44×10^{-9}	PRRT1
Urayama et al (2012)	Various European	1200 cases 6417 controls	cHL	rs2248462*[#]	A	0.61	1.3×10^{-13}	MICB
			cHL	rs2395185*[#]	T	0.56	8.3×10^{-23}	HLA-DRA
			EBV+cHL	rs2734986[#]	C	2.45	1.2×10^{-15}	HLA-A
			EBV+cHL	rs6904029[#]	A	0.46	5.5×10^{-10}	HcG9/HLA-A
			EBV-cHL	rs6903608[#]	C	1.71	3.2×10^{-27}	HLA-DRA
Frampton et al (2013) [†§]	United Kingdom Germany	1465 cases 6417 controls	cHL	rs6903608	C	1.62	5.36×10^{-27}	HLA-DRA
			cHL	rs2395185	T	0.66	4.44×10^{-16}	HLA-DRA

cHL, classical Hodgkin lymphoma; EBV, Epstein-Barr virus; NSHL, nodular sclerosis Hodgkin lymphoma

[†] replication group includes a subset of subjects from Urayama et al (2012)

[§] includes the subjects from Enciso-Mora et al (2011)

*significant in independent replication

[#] significant in technical replication

[α] significant association with EBV-cHL in replication set

Variants rs9268542 and rs9268528 are in LD ($r^2 = 1$), and there is some degree of LD between these variants and rs2395185 ($r^2 = 0.38$). SNPs identified by Urayama et al (2012) were independently associated with cHL.

Table 4. HLA associations with DLBCL, FL and CLL identified in GWAS

Lymphoma type	Study	Country	Study Subjects in Discovery Set	SNP	Minor allele	Odds Ratio	p-value	Candidate Gene/Haplotype
DLBCL	Skibola et al (2009)	United States European origin	380 cases 1412 controls	rs6457327	A	0.79	2.7×10^{-2}	C6orf15 (STG)/ PSORS1
	Smedby et al (2011)	Denmark	379 cases 791 controls	rs10484561	C	1.36	1×10^{-7}	HLA-DQB1
	Cerhan et al (2014)	European	2661 cases 6221 controls	rs2523607	A	1.45	7.1×10^{-10}	HLA-B*08:01
FL	Skibola et al (2009)	United States European origin	278 cases 1412 controls	rs6457327	A	0.55	1.1×10^{-5}	C6orf15 (STG)/ PSORS1
	Conde et al (2010)	United States	213 cases 750 controls	rs10484561 rs7755224	G G	1.81 1.8	5.88×10^{-5} 2.6×10^{-5}	HLA-DQB1
				rs6457614 rs4947332 rs1794265	G T T	1.75 - -	7.06×10^{-5} - -	HLA-DQB1*05:01 HLA-DRB1*01:01 HLA-DQA1*01:01
	Smedby et al (2011)	United States, Canada, Denmark, Australia	1592 cases 5220 controls	rs2647012	T	0.6	1×10^{-7}	DRB1*15:01-DQA1*01:02-DQB1*06:02
	Skibola et al (2012)*	Denmark, United States	592 cases 1541 controls	rs3117222	A	0.66	1.45×10^{-7}	HLA-DPB1*03:01
CLL	Slager et al (2011)	United States	407 cases 296 controls	rs615672 rs674313	C T	1.95 2.4	6.42×10^{-5} 1.12×10^{-6}	HLA-DRB5

				rs502771	C	1.93	1.07×10^{-4}	
				rs9272219	T	2.32	1.65×10^{-6}	HLA-DQA1
				rs9272535	A	2.33	1.33×10^{-6}	
	Di Bernardo et al (2013)	United Kingdom	502 cases 2697 controls	rs6904029	A	1.32	1.38×10^{-4}	HcG9/HLA-A

SNP associations arranged by lymphoma type. DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; CLL, chronic lymphocytic leukaemia.

*meta-analysis includes Conde et al (2010) and Smedby et al (2011)

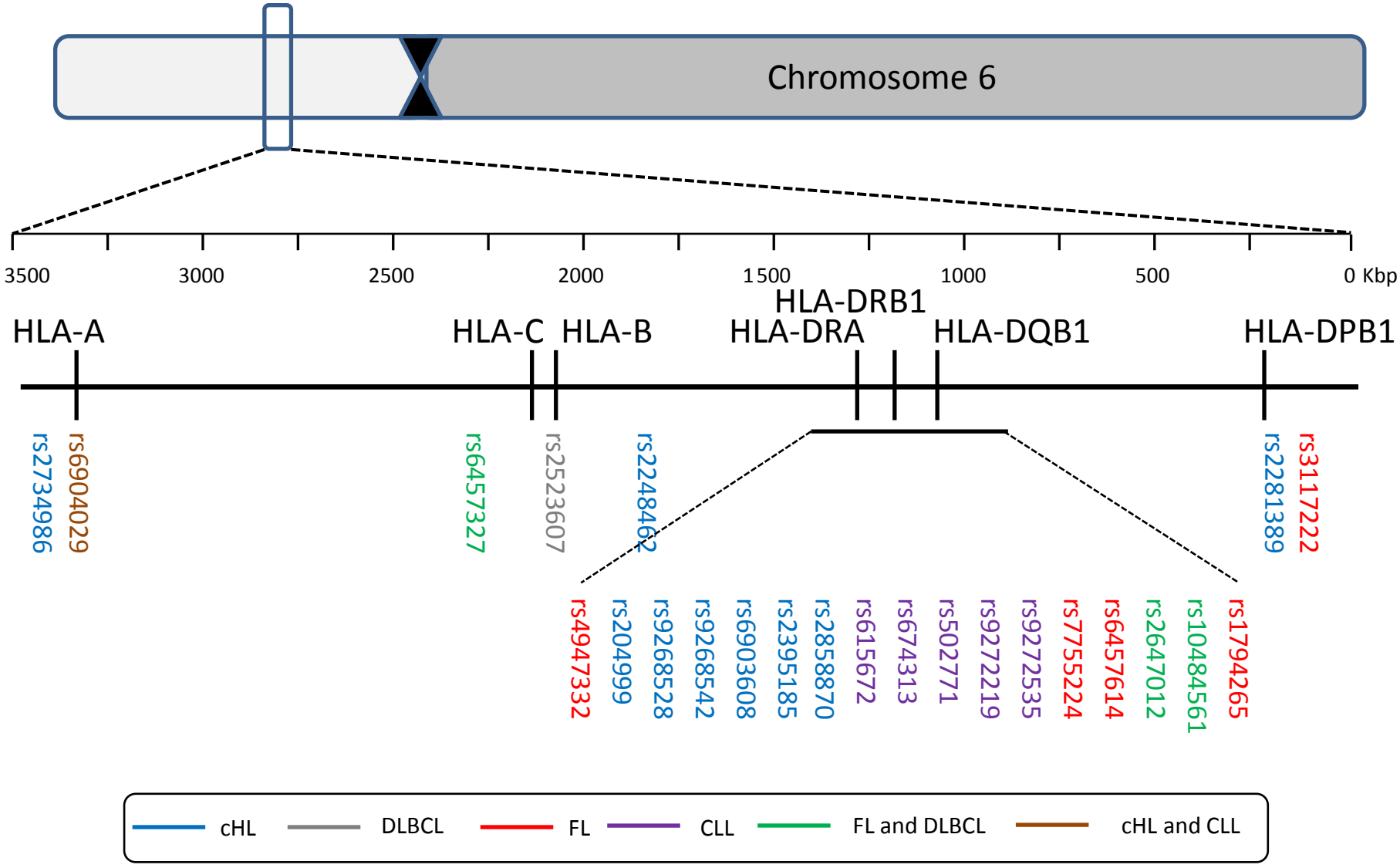

Figure 1. Position of lymphoma-associated SNPs on chromosome 6