Background  Accumulating evidence suggests that risk factors for classical Hodgkin lymphoma (cHL) differ by tumor Epstein–Barr virus (EBV) status. This potential etiological heterogeneity is not recognized in current disease classification.

Methods  We conducted a genome-wide association study of 1200 cHL patients and 6417 control subjects, with validation in an independent replication series, to identify common genetic variants associated with total cHL and subtypes defined by tumor EBV status. Multiple logistic regression was used to calculate odds ratios (ORs) and 95% confidence intervals (CIs) assuming a log-additive genetic model for the variants. All statistical tests were two-sided.

Results  Two novel loci associated with total cHL irrespective of EBV status were identified in the major histocompatibility complex region; one resides adjacent to MICB (rs2248462: OR = 0.61, 95% CI = 0.53 to 0.69, P = 1.3 × 10⁻³) and the other at HLA-DRA (rs2395185: OR = 0.56, 95% CI = 0.50 to 0.62, P = 8.3 × 10⁻²⁴) with both results confirmed in an independent replication series. Consistent with previous reports, associations were found between EBV-positive cHL and genetic variants within the class I region (rs2734986, HLA-A: OR = 2.45, 95% CI = 2.00 to 3.00, P = 1.2 × 10⁻¹³; rs6904029, HCG9: OR = 0.46, 95% CI = 0.36 to 0.59, P = 5.5 × 10⁻¹³) and between EBV-negative cHL and rs6903608 within the class II region (rs6903608, HLA-DRA: OR = 2.08, 95% CI = 1.84 to 2.35, P = 6.1 × 10⁻¹⁰). The association between rs6903608 and EBV-negative cHL was confined to the nodular sclerosis histological subtype. Evidence for an association between EBV-negative cHL and rs20541 (5q13, IL13: OR = 1.53, 95% CI = 1.32 to 1.76, P = 5.4 × 10⁻⁰⁴), a variant previously linked to psoriasis and asthma, was observed; however, the evidence for replication was less clear. Notably, one additional psoriasis-associated variant, rs27524 (5q15, ERAP1), showed evidence of an association with cHL in the genome-wide association study (OR = 1.21, 95% CI = 1.10 to 1.33, P = 1.5 × 10⁻¹⁰) and replication series (P = .03).

Conclusion  Overall, these results provide strong evidence that EBV status is an etiologically important classification of cHL and also suggest that some components of the pathological process are common to both EBV-positive and EBV-negative patients.


Hodgkin lymphoma (HL) is a malignant condition of B lymphocyte origin and is among the most common cancers in adolescents and young adults (1). Current classification broadly divides HL into two major forms that are thought to be distinct clinicopathologic entities: classical HL (cHL), which accounts for about 95% of HL, and the less prevalent nodular lymphocyte–predominant HL (2). cHL is characterized by the presence of a small number of Hodgkin and Reed–Sternberg (HRS) tumor cells scattered within an admixture of inflammatory cells and is further subdivided into four histological subtypes consisting of nodular sclerosis HL (NSHL), mixed cellularity HL (MCHL), lymphocyte-depleted HL, and lymphocyte-rich HL (2,3).
Epstein–Barr virus (EBV) has been etiologically associated with a proportion of cHL patients (4,5). EBV latent membrane protein (LMP)-1 and/or EBV-encoded small RNAs (EBERs) have been detected in HRS cells in as many as 40% of cHL patients in economically developed countries (5,6). The presence of clonal EBV genomes in tumor cells and the known transforming potential of LMP-1 and LMP-2 suggest that EBV plays a direct role in the pathogenesis of cHL. The proportion of EBV-positive cHLs varies with sex, age, ethnicity, and regional economic development. Specifically, EBV-positive cHL is more frequently observed in males, children and older adults, non-whites, and in less economically developed regions of the world (7). Compared with the NSHL histological subtype, MCHL is more likely to be EBV positive across all age groups, but particularly for young adults (7). However, because approximately 60% of cHLs are NSHL (vs about 30% MCHL), the majority of EBV-positive cHLs are of this histological subtype.

A small number of B lymphocytes in most healthy adults are persistently infected with EBV. The factors associated with progression to EBV-related cHL are not fully understood, but growing evidence suggests that it is likely to be influenced by genetic factors (8,9). Antiviral immune responses are considered to be essential in the control of EBV-induced lymphomagenesis. Accordingly, a previously reported genetic screening study using 33 microsatellite markers across the major histocompatibility complex (MHC) region, a region densely populated with immunoregulatory genes, identified susceptibility loci within the class I and class III regions that were specifically associated with EBV-positive or EBV-negative cHL (8). Subsequent evidence has emerged that supports a strong involvement of the highly polymorphic HLA-A locus in EBV-positive cHL in which HLA-A*01 confers an approximate twofold increased risk per allele, whereas HLA-A*02 was associated with an approximate 40% reduction in risk (9,10). Also, a recent genome-wide association study (GWAS) showed an association between cHL and an MHC class II variant, rs6901608, which appeared to be largely confined to the EBV-negative subtype (11).

Taken together, these previous reports provide a strong indication that cHL is an etiologically heterogeneous disease that may warrant consideration of EBV status as an addition to the current classification of cHL, which does not include EBV. We therefore conducted a GWAS among a large series of EBV-classified cHL patients and control subjects, with an emphasis on incorporating EBV status in the definition of tumor phenotype, to identify common genetic variants involved in the susceptibility to cHL.

**Methods**

**GWAS Samples**

Genome-wide genotyping was performed on cHL case samples from five European-based HL studies: the EPILYMPH Study, the Scotland and Newcastle Lymphoma Group, the Young Adult Hodgkin’s Disease Case–Control Study, the Scandinavian Lymphoma Etiology Study, and the Northern Dutch Hodgkin Lymphoma Study (Table 1). Briefly, EPILYMPH is a multicenter case–control study of lymphoma conducted between January 1998 and June 2003 that included patients from centers in six European countries and followed a common core protocol (12). The Scotland and Newcastle Lymphoma Group includes the Scotland and Newcastle Epidemiological Study of Hodgkin’s Disease, which is a case–control study conducted in Scotland and Northern England between January 1993 and July 1997 (6), and the “local case series,” a prospectively collected series of patients within those regions between January 1992 and March 2008 (9). The Young Adult Hodgkin’s Disease Case–Control Study is a case–control study of HL in adolescents and young adults (aged 16–24 years) conducted in the United Kingdom in the regions of Yorkshire, Wessex, Southwest Family Health Service Authority areas, and parts of Cumbria and Lancashire between October 1991 and May 1995 (13). The Scandinavian Lymphoma Etiology Study is a case–control study of lymphomas conducted in Denmark and Sweden between January 1999 and August 2002 (14), and the Northern Dutch Hodgkin Lymphoma Study comprises a population-based case patient series of HL diagnosed in Northern Netherlands.
between January 1987 and April 2009 (8). DNA samples of sufficient quality and quantity for genome-wide genotyping were available for 1238 cHL patients aged 15–80 years (median age = 33 years), as well as 1395 study-specific control subjects representing nine European countries, including the Czech Republic, Denmark, France, Germany, Ireland, the Netherlands, Spain, Sweden, and the United Kingdom. The study protocol was approved by the institutional review boards, regional scientific ethics committees, and/or data protection agencies of all collaborating centers or institutions, and written informed consent was obtained for all participating subjects.

A large series of generic control subjects were selected on the basis of the country of recruitment of case patients from several previously published GWAS of other cancers conducted among populations of European ancestry (Table 1). These studies included the Alcohol Related Cancers and Genetic Susceptibility in Europe Study (15); the International Agency for Research on Cancer Central Europe Study (16); the Pancreatic Cancer Cohort Consortium (17); the Nijmegen Biomedical Study (18); the Scandinavian Lymphoma Etiology Study, 1999–2002 (19); the Northern Dutch Hodgkin Lymphoma Study, 1987–2009 (20); the Scotland and Newcastle Lymphoma Group, 1992–2008‡; the Young Adult Hodgkin Case–Control Study, 1991–1995 (21); and the Wellcome Trust Case–Control Consortium (21). From these six studies, genome-wide data from a total of 6466 control subjects were assembled.

### Genome-Wide Genotyping and Quality Control

Genome-wide genotyping of cHL case patient samples was performed using the Illumina Infinium Human660-Quad BeadChip (Illumina Inc, San Diego, CA) at the Centre National de Génotypage.
Independent Replication

Further evaluation of the independence of the associations between cHL and the 20 technical replication SNPs identified five SNPs within the MHC region and one at 5q31 (rs20541) that best explain the associations with each locus. Three of the six SNPs have not been previously reported in the cHL literature. Thus, they were studied in the independent replication series comprising a different sample of 563 case patients and 613 control subjects from EPILYMPH (12), the Scotland and Newcastle Lymphoma Group (6,9), the Young Adult Hodgkin’s Disease Case–Control Study (13), and the Epidemiology and Genetics Lymphoma Case–Control Study conducted between January 1998 and July 2003 (http://www.elccs.info) (26) (Table 1). The rs20541 SNP has been associated with psoriasis and asthma in previous GWAS (27–31). One additional psoriasis SNP that showed a putative association in the cHL GWAS, rs27524 (ERAPI), was selected for independent replication on the basis of the hypothesis that there may be overlap in the biological mechanisms of the two immune-related diseases. In total, four SNPs were included in the independent replication stage.

Genotyping was performed using TaqMan Pre-Designed SNP Genotyping Assays (Applied Biosystems, Carlsbad, CA) in 384-well plates for variants rs27524, rs20541, rs2248462 (6p21: MICB), and rs2381585 (6p21: HLA-DR4), including 3.6% duplicate samples. A specific TaqMan Pre-Designed SNP Genotyping Assay Mix (containing probes and primers) was used for each SNP (Applied Biosystems assay-on-demand order code: C\_3056837\_10 for rs27524, C\_1259921\_20 for rs20541, C\_1624657\_10 for rs2248462, and C\_1622465\_10 for rs2381585). Assays were run on the GeneAmp PCR System 9700 thermal cycler (Applied Biosystems) for 15 minutes at 95°C followed by 30 cycles at 95°C for 15 seconds and a 1-minute incubation at 60°C. Subsequently, the endpoint fluorescence was read with an ABI Prism 7900HT Sequence Detection System, and genotypes were “called” using the SDS 2.1 software (Applied Biosystems). The performance of the assays was validated by re-genotyping CEU HapMap samples (US residents with northern and western European ancestry) and comparing the results to HapMap genotypes (http://hapmap.ncbi.nlm.nih.gov). All assays were found to be functioning robustly. Within the study samples, duplicates genotyping concordance was greater than 99%.

Tumor EBV Status Classification

HL EBV status was ascertained through in situ hybridization for EBERs and/or through immunohistochemical staining for EBV LMP-1 protein on formalin-fixed paraffin-embedded tumor samples, with commercially available reagents and protocols that are used in diagnostic settings (32–34). In most patients, in situ hybridization was performed to detect EBER1-2 RNAs, which are associated with latent EBV infection, by using an EBER-specific peptide nucleic acid probe according to the manufacturer’s instructions (DAKO, Glostrup, Denmark). An appropriate positive control (known EBV-positive HL case) was included.

Immunohistochemical staining for EBV LMP-1 protein was typically used to supplement the in situ hybridization analyses, but in a small proportion of samples (<10%), EBV LMP-1 staining was the only test for EBV. Following antigen retrieval, paraffin
sections were incubated with the antibody cocktail CS1-4 (DAKO). Reactivity was detected using standard EnVision (DAKO), VECTASTAIN ABC (Vector Laboratories, Peterborough, UK), or Ventana kit (Ventana, Tuscon, AZ) immunohistochemistry. EBV-positive patients were defined as those for whom their HRS cells expressed EBERs or LMP-1 protein (35).

**Statistical Analysis**

The association between each genetic variant and the disease phenotype was estimated by the odds ratio (OR) per allele and 95% confidence intervals (CIs) using multiple logistic regression, assuming a log-additive genetic model with sex (male or female) and country of recruitment (France, Germany, Spain, Czech Republic, Ireland, United Kingdom, Denmark, Sweden, and the Netherlands; up to eight indicator variables after excluding one country as the reference, depending on the analysis) included in the model as covariates. To adjust for any additional population stratification not captured by adjustment for country, eight principal components analysis eigenvectors, which were suggested to be informative based on the Tracy–Widom statistic ($P < .05$), were included in the regression model (36). To evaluate the independence of the associations between cHL and multiple SNPs, additional model variables representing the variants in question were included together in a single logistic regression model. Results of the genome-wide and independent replication analyses were combined with meta-analyses applying the inverse variance weighting method (37). Between-study heterogeneity in risk estimate was evaluated using the Cochran $Q$ test statistic. Differences for which $P$ values were less than $1.0 \times 10^{-7}$ were considered statistically significant at the genome-wide level (21).

Genome-wide analyses were conducted for total cHL and for EBV-positive and EBV-negative cHL as the phenotypes of interest. For SNPs that were identified as being associated with cHL, analyses were also conducted on subtypes that were defined using both histology (NSHL or MCHL) and tumor EBV status. The $\chi^2$ test of homogeneity was used to test the null hypothesis that the subtype-specific odds ratios were equal. Genome-wide analyses were not stratified by sex to minimize the total number of statistical tests performed in light of no compelling previously published evidence for strong sex-specific differences in genetic associations. The study population included only whites of European origin; thus, no analyses by racial/ethnic groups were performed. All statistical tests were two-sided. Analyses were conducted using PLINK (23) and SAS software version 9 (SAS, Cary, NC) (38).

![Figure 1](https://academic.oup.com/jnci/article-abstract/104/3/240/972531)
Results

The GWAS analysis included 1200 cHL case patients, of which tumor EBV data were available for 933 (77.8%) of 1200 patients. Of the EBV-classified patients, 265 (28.4%) of 933 were EBV-positive (Table 1). The control subject series consisted of 6417 control subjects with existing genome-wide SNP data representing the countries from which the patients were recruited (Table 1 and Supplementary Table 3, available online). The overall analysis of cHL and 502 514 SNPs adjusting for sex, country, and eight principal components analysis eigenvectors showed minimal evidence of genomic inflation (λe = 1.04 when adjusted to a sample size of 1000 case patients and 1000 control subjects) (39) (Supplementary Figure 1, available online).

MHC Region

Strong associations were observed between cHL and SNPs within the MHC region (chromosomal region 6p21) (Figure 1). Because of the availability of EBV status for the large majority of the cHL patients, three groups of risk-associated variants were distinguishable: those relevant to total cHL, EBV-positive cHL, and EBV-negative cHL (Figure 2 and Supplementary Figure 2, available online). Two loci showed strong evidence of association with total cHL, with no evidence of heterogeneity in effect by tumor EBV status or histological subtype (Table 2 and Figure 3). One of these, indexed by SNP rs2248462, located in the class I region nearing the border with the class III region at the MICB gene, was associated with a reduced risk of cHL (OR = 0.61, 95% CI = 0.53 to 0.69, \(P = 1.3 \times 10^{-13}\)). A reduced risk of cHL was also independently associated with rs2395185 (OR = 0.56, 95% CI = 0.50 to 0.62, \(P = 8.3 \times 10^{-13}\)), a class II SNP located at HLA-DRA.

Multiple SNPs in the class I region were associated with EBV-positive cHL with only limited evidence of an association with EBV-negative cHL (\(P_{\text{heterogeneity}} < .001\)) (Table 2 and Figure 3). Most notable was the variant rs2734986 (OR = 2.45, 95% CI = 2.00 to 3.00, \(P = 1.2 \times 10^{-13}\)) and another class I SNP in weak LD (\(r^2 < 0.10\), rs6904029, which showed an independent association with EBV-positive cHL (OR = 0.46, 95% CI = 0.36 to 0.59, \(P = 5.5 \times 10^{-11}\)). Previously genotyped HLA-A data were available for the majority of patients (257 EBV-positive and 642 EBV-negative cHL), and strong LD was observed for rs2734986 (\(r^2 = 0.98\)) and rs6904029 (\(r^2 = 0.88\)) with the previously documented HLA-A*01 and A*02 allelic groups, respectively (9).

Also within the MHC in the class II region, a marked cHL association was found for rs6903608 (OR = 1.71, 95% CI = 1.55 to 1.89, \(P = 3.2 \times 10^{-13}\)), a SNP previously reported (11) (Table 2, Figure 2, and Supplementary Figure 2, available online). Strong evidence of heterogeneity in the risk estimate by tumor EBV status for rs6903608 was observed (\(P_{\text{heterogeneity}} < .001\)). SNP rs6903608 showed no association with EBV-positive cHL (OR = 0.95, 95% CI = 0.78 to 1.16, \(P = .63\)), but a markedly increased risk of EBV-negative cHL was observed (OR = 2.08, 95% CI = 1.84 to 2.35, \(P = 6.1 \times 10^{-11}\)) (Figure 3). An analysis including the five currently reported MHC region SNPs (rs2248462, rs2395185, rs2734986, rs6904029, and rs6903608) in a single logistic regression model showed that all were independently associated with the risk of total cHL and/or EBV subtype–specific cHL (Table 2).

Figure 2. Investigation of the association between genetic variants and total classical Hodgkin lymphoma (cHL), Epstein–Barr virus (EBV)–positive cHL, and EBV-negative cHL within an approximately 6.5 Mb region of the extended major histocompatibility complex located at 6p21. Overlaid plots of the \(-\log_10(P\text{-values})\) for 1973 single-nucleotide polymorphisms (SNPs) by their chromosomal position resulting from three separate analyses of all patients who have cHL (total cHL, in gray), patients who have EBV-positive cHL (in blue), and patients who have EBV-negative cHL (in red) are shown. Multiple logistic regression analysis was performed assuming a log-additive genetic model and adjusting for sex (male or female), country (up to eight indicator variables after excluding one country as the reference, depending on the analysis: France, Germany, Spain, Czech Republic, Ireland, United Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors. All statistical tests were two-sided. Arrows indicate cHL-associated regions indexed by five SNPs.
Non-MHC Regions

The genome-wide analysis also indicated a strong signal originating from the chromosomal region 5q31 in proximity to the \textit{IL13} and \textit{IL4} genes (Figure 1). An increased risk of cHL was associated with the minor allele of rs20541 (Table 3 and Figure 4), which appeared predominantly in EBV-negative cHL (OR = 1.53, 95% CI = 1.32 to 1.76, \(P = 5.4 \times 10^{-7}\)). The rs20541 variant has been shown to be involved in the genetic susceptibility to psoriasis and asthma in recently published GWAS (27–31). As the overlap suggests that there may be potential similarities in biological mechanism between cHL and these other immune-related diseases, we placed emphasis on SNPs associated with psoriasis and asthma identified by previous GWAS (Supplementary Table 4, available online). One additional SNP, rs27524, was associated with total cHL (OR = 1.21, 95% CI = 1.10 to 1.33, \(P = 1.5 \times 10^{-8}\)) (Table 3 and Figure 4) after applying a Bonferroni correction for the 20 psoriasis- and asthma-associated SNPs (\(P < .003\)) that were considered. The association between total cHL and SNP rs2476601 (chromosomal region 1p13, \textit{PTPN22}) also showed borderline statistical significance after a Bonferroni correction (\(P = 3.7 \times 10^{-7}\)) (Supplementary Table 4, available online).

Technical and Independent Replication

Subsequent to the GWAS analysis, the key associations between SNPs and cHL were investigated by technical (genotype validation and analysis using study-specific control subjects) and independent (replication of findings in an independent series) replication. In the technical replication, all SNPs had a genotype concordance between the Illumina (Illumina Inc) and Sequenom (Sequenom Inc) platforms of 99% or greater (Supplementary Table 2, available online). Furthermore, the associations were consistent between the analysis using generic and study-specific control subjects (Supplementary Table 5, available online). For the independent replication conducted in the additional series of 563 cHL patients and 613 control subjects for four previously unreported SNPs (rs2248462, rs2395185, rs20541, and rs27524), consistent evidence of associations between rs2248462, rs2395185, and rs27524 and cHL were found with the exception of rs20541 (EBV-negative cHL: OR = 1.18, 95% CI = 0.83 to 1.67, \(P = .361\)) (Table 3). Nevertheless, the combined GWAS and replication results for rs20541 were statistically significant at the genome-wide level (total cHL: OR = 1.38, 95% CI = 1.24 to 1.54, \(P = 1.8 \times 10^{-8}\); EBV-negative cHL: OR = 1.47, 95% CI = 1.29 to 1.68, \(P = 1.1 \times 10^{-6}\)), but additional studies are needed to confirm this finding (Supplementary Figure 3, available online).

Associations by Tumor EBV Status in MCHL and NSHL

An evaluation of the relationship between cHL and the SNPs by tumor EBV status within the separate MCHL and NSHL histological subtypes was performed. The exclusively EBV-negative cHL association with rs6903608 was largely restricted to the NSHL subtype (Table 4) with evidence of heterogeneity between EBV-negative NSHL and MCHL (\(P_{\text{heterogeneity}} = .002\)). The EBV-positive cHL associations with rs27524 and rs6904029 were observed in both MCHL and NSHL subtypes. However, there was some suggestion that the rs27524 and EBV-positive relationship may be stronger in MCHL compared with NSHL (\(P_{\text{heterogeneity}} = .09\)).
Figure 3. Stratified and subgroup analyses of five single-nucleotide polymorphisms (SNPs) within the major histocompatibility complex region, which were independently associated with classical Hodgkin lymphoma (cHL). Results are shown for (A) rs2248462 (class I region, MICA), (B) rs2395185 (class II region, HLA-DRA), (C) rs2734986 (class III region, HLA-DQA1), (D) rs6904029 (class IV region, HLA-DQB1), and (E) rs6903608 (class V region, HLA-DRB1). Odds ratios (ORs, represented by boxes with the area of each box inversely proportional to the variance of the estimate) and 95% confidence intervals (CIs, error bars) were derived using multiple logistic regression assuming a log-additive genetic model and adjusting for sex (male or female), country (up to eight indicator variables after excluding one country as the reference, depending on the analysis: France, Germany, Spain, Czech Republic, Ireland, United Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors for the genome-wide association study (GWAS) analyses only. The dashed vertical line represents the OR of the SNP in the analysis of total cHL among all subjects, and the width of the diamond is the corresponding 95% CI. In the analysis stratified by study, the GWAS included the EPILYPH study (EPILYPH-GWAS), Scandinavian Lymphoma Etiology Study (SCALE-GWAS), the Scotland and Newcastle Lymphoma Group, and Young Adult Hodgkin Case–Control Study analyzed together (referred to as the UK Studies-GWAS), and the Northern Dutch Hodgkin Lymphoma Study (Netherlands-GWAS). The independent replication was done with data from the EPILYPH study (EPILYPH - Replication), and the Scotland and Newcastle Lymphoma Group, Young Adult Hodgkin Case–Control Study, and Epidemiology and Genetics Lymphoma Case–Control Study analyzed together (referred to as the UK Studies-Replication). \( \chi^2 \) test was the basis for the Cochran Q test statistic and was used to evaluate between-study heterogeneity in results. Associations between the SNPs and cHL subgroups (including histological subtype, EBV status, and age) were performed, and \( P \) values (\( \chi^2 \) test) were presented to indicate differences in the OR between subgroup analyses. For rs2248462 (A) and rs2395185 (B), results were derived from a combined analysis of the GWAS and independent replication phase results using an inverse variance weighting meta-analysis. All statistical tests were two-sided. Ca = case patients; Chr = chromosome; Co = control subjects; MCHL = mixed cellularity Hodgkin lymphoma; NSHL = nodular sclerosis Hodgkin lymphoma.

Discussion

The large size of the current GWAS including 1200 cHL patients and more than 6000 control subjects, together with EBV and histology information available for the majority of case patients, allowed us to evaluate genome-wide associations by tumor EBV status and further refine the specificity of previously reported susceptibility loci. Overall, we confirmed the strong role of the MHC (SCALE-GWAS), the Scotland and Newcastle Lymphoma Group, and Young Adult Hodgkin Case–Control Study analyzed together (referred to as the UK Studies-GWAS), and the Northern Dutch Hodgkin Lymphoma Study (Netherlands-GWAS). The independent replication was done with data from the EPILYPH study (EPILYPH - Replication), and the Scotland and Newcastle Lymphoma Group, Young Adult Hodgkin Case–Control Study, and Epidemiology and Genetics Lymphoma Case–Control Study analyzed together (referred to as the UK Studies-Replication). \( \chi^2 \) test was the basis for the Cochran Q test statistic and was used to evaluate between-study heterogeneity in results. Associations between the SNPs and cHL subgroups (including histological subtype, EBV status, and age) were performed, and \( P \) values (\( \chi^2 \) test) were presented to indicate differences in the OR between subgroup analyses. For rs2248462 (A) and rs2395185 (B), results were derived from a combined analysis of the GWAS and independent replication phase results using an inverse variance weighting meta-analysis. All statistical tests were two-sided. Ca = case patients; Chr = chromosome; Co = control subjects; MCHL = mixed cellularity Hodgkin lymphoma; NSHL = nodular sclerosis Hodgkin lymphoma.
showed evidence of an association between cHL and novel loci in close proximity to key immunoregulatory genes at chromosome 5q31 (rs20541, IL13) and 5q15 (rs27524, ERAP1), both of which have also been previously linked to psoriasis.

rs2248462 and rs2395185, both located within the MHC region nearly 1 Mb apart ($r^2 < 0.1$), are independently associated with total cHL with little evidence of heterogeneity in risk estimate by EBV status and histology. The rs2248462 SNP is located in the class I region adjacent to the highly polymorphic MHC class I–related chain (MICB) gene, which encodes a heavily glycosylated protein that is a ligand for the NKG2D receptor. The MICB ligand, structurally resembling the MHC class I glycoproteins, is expressed on epithelial cells in response to cellular stress and activates the cytolytic response of T lymphocytes and natural killer cells that are capable of eliminating tumor and virally infected cells (40). A previous study examining the MHC region reported an association with EBV-negative cHL and a microsatellite marker that possibly tagged the same susceptibility locus (8). It involved a broadly associated region centered on the D6S273 microsatellite marker that is located about 100 kb from rs2248462. The second cHL-associated SNP reported in this study, rs2395185, is in the 3′ untranslated region of HLA-DRA in close proximity to the EBV-negative cHL SNP, rs6903608 ($r^2 = 0.20$). However, in contrast to rs6903608, rs2395185 was associated with EBV-positive and EBV-negative cHL, providing evidence that these two SNPs contribute independently to cHL risk.

A cluster of SNPs within the class I region was associated specifically with EBV-positive cHL. The strongest signal was observed for rs2734986 ($P = 1.2 \times 10^{-15}$), located in the 3′ untranslated region of HLA-G and in close proximity to the HLA-A locus. Further analysis showed that rs4694029 (located approximately 124 kb downstream in HCG9) was independently associated with EBV-positive cHL. Previous evidence based on a subset of the subjects included in this GWAS showed an increased and decreased risk of EBV-positive cHL associated with HLA-A*01 and A*02, respectively (9,10). An evaluation among the patients with HLA-A data available showed HLA-A*01 to be in strong LD with rs2734986 ($r^2 = 0.98$) but not with rs6904029 ($r^2 = 0.10$). Similarly, HLA-A*02 was in LD with rs6904029 ($r^2 = 0.88$) but not with rs2734986 ($r^2 = 0.10$).

The analysis by EBV status showed the class II SNP (11), rs6903608, to be associated only with EBV-negative cHL, whereas no association was observed with EBV-positive cHL. This SNP resides in the 3′ untranslated region of the HLA-DRA locus and is in relatively close proximity to the other HLA class II loci (HLA-DR, HLA-DQ, and HLA-DP), which have been implicated in earlier studies of HL (41–44). In our study, we identified an additional level of heterogeneity in the risk estimate of this variant by histological subtype, indicating that associations were predominantly between rs6903608 and EBV-negative NSHL, rather than EBV-negative MCHL. Although cHL histology and EBV status are related, they are far from synonymous. To our knowledge, this is the first study to show clear evidence of a histology-specific association between genetic loci and cHL risk within an EBV-defined series of cHL patients.

In this GWAS, additional genome-wide significant associations between genetic variants and cHL were identified outside of the
MHC region within a cytokine gene cluster at chromosomal region 5q31, including a nonsynonymous SNP in the IL13 gene (rs20541). rs20541 causes the replacement of a positively charged arginine with a neutral glutamine that has been shown to result in increased activity of IL13, a type 2 helper T cell cytokine (45). Association studies have also attributed this SNP to increased serum immunoglobulin E levels (46), atopic conditions (47), asthma (28), and psoriasis (30,31). In cHL, IL13 protein expression has been specifically detected in HRS cells (48) and rarely within the reactive infiltrate of cHL tumors. Evidence from gene expression studies increasingly suggests that this cytokine may function as an autocrine growth factor for the HRS cells in cHL (49). Involvement of IL13 in cHL is also consistent with a recent GWAS reporting an association between cHL and the GATA3 gene (11). GATA3 acts as a master transcription factor for key cytokines, including IL13, in lymphoid cell development (50), and concordantly aberrant expression of these two genes has been detected in HRS cells derived from cHL tissue (51). Further support for an involvement of cytokines in cHL has been demonstrated in a recent candidate gene study reporting evidence of epistatic interactions within the inflammatory response pathway (52).

Although the overall association between rs20541 and cHL (GWAS and replication combined) was statistically significant at the genome-wide level, weak evidence for this relationship was observed in the independent replication series alone. The inconsistency may be partially attributed to insufficient statistical power in the replication series, particularly if the GWAS estimate is overestimated as a consequence of the stringent criteria for genome-wide statistical significance (53). Notably, in the genome-wide analysis, rs20541 appeared to be predominately associated with EBV-negative cHL. The replication dataset included a total of 563 cHL patients, of which only half were classified as EBV-negative, which may have further influenced statistical power. Although strong evidence in the GWAS was found for an association between rs20541 and cHL, and a biological role for rs20541 in cHL is plausible, this reported association requires further confirmation.

The overlap of the cHL association with rs20541 found in our study with recent GWAS of psoriasis (31) and asthma (28) led us to consider other asthma and psoriasis SNPs as possible candidate loci associated with cHL. In particular, rs27524, which maps to the endoplasmic reticulum aminopeptidase I (ERAP1, chromosome 5q15) gene locus, has been strongly associated with psoriasis (31) and has shown evidence of an association with cHL in our genome-wide analysis. A consistent association was observed in the independent replication study and yielded strong evidence in the GWAS and replication combined analysis. ERAP1 is an aminopeptidase involved in the trimming of N-terminal amino acid residues of precursor polypeptides in preparation for loading and antigen presentation by the MHC class I molecule. Aberrant activity of ERAP1 (and ERAP2, a gene located in close proximity) expression resulting in impaired antigen processing plays a key role in regulating the surface expression of HLA class I molecules (54). Fruci et al. (55) reported consistently high ERAP1 expression levels in lymphoblastoid cell lines established from healthy donors, but markedly varied expression levels correlated with surface HLA class I expression in tumor cell lines. Thus, it has been hypothesized that downregulation of HLA class I expression and/or deranged presentation of certain tumor antigens caused by abnormal expression of ERAP1 may contribute to the ability of tumor cells to evade immune surveillance (55).

The nature of a possible relationship of cHL with psoriasis and potentially with other autoimmune diseases is unclear and requires

**Figure 4.** Genome-wide association study results in total classical Hodgkin lymphoma for single-nucleotide polymorphisms (SNPs) within a 300-kb flanking region of rs20541 at chromosome 5q31 (left) and rs27524 at chromosome 5q15 (right). Multiple logistic regression was performed assuming a log-additive genetic model and adjusting for sex (male or female), country (eight indicator variables after excluding one country as the reference: France, Germany, Spain, Czech Republic, Ireland, United Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors. The \(- \log_{10}(P\text{ value})\) for each SNP are plotted against their chromosomal position. All statistical tests were two-sided. The black triangle indicates the reported associated SNP and the colors of the dots represent the degree of linkage disequilibrium (based on \(r^2\)) in relation to that index SNP. Recombination rates (cM/Mb) overlay the plots and are based on HapMap phase I and II data (http://hapmap.ncbi.nlm.nih.gov). cM/Mb = centiMorgans/megabase.
Table 4. Results of a genome-wide association analysis of classical Hodgkin lymphoma (cHL) subtypes defined jointly by tumor Epstein–Barr virus (EBV) status and histology*  

<table>
<thead>
<tr>
<th>SNP (gene candidate) by cHL EBV status</th>
<th>Pooled (NSHL and MCHL)</th>
<th>MCHL (n = 154)</th>
<th>NSHL (n = 685)</th>
<th>(P_{\text{MCHL} / \text{NSHL}})†</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2248462 (MCIB)</td>
<td>EBV-positive cHL</td>
<td>OR (95% CI)</td>
<td>(P)</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td></td>
<td>EBV-negative cHL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs2351818 (HLA-DRA)</td>
<td>EBV-positive cHL</td>
<td>0.51 (0.40 to 0.65)</td>
<td>4.7 (\times) 10^{-6}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EBV-negative cHL</td>
<td>0.58 (0.50 to 0.67)</td>
<td>3.0 (\times) 10^{-13}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs27524 (ERAPI)</td>
<td>EBV-positive cHL</td>
<td>1.30 (1.07 to 1.58)</td>
<td>0.009</td>
<td>1.19 (0.88 to 1.63)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EBV-negative cHL</td>
<td>1.23 (1.08 to 1.40)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2734986 (HLA-A)</td>
<td>EBV-positive cHL</td>
<td>2.60 (2.06 to 3.27)</td>
<td>7.2 (\times) 10^{-10}</td>
<td>3.35 (2.34 to 4.81)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EBV-negative cHL</td>
<td>0.95 (0.80 to 1.14)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6904029 (HCG9)</td>
<td>EBV-positive cHL</td>
<td>0.46 (0.36 to 0.60)</td>
<td>4.4 (\times) 10^{-9}</td>
<td>0.41 (0.27 to 0.62)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EBV-negative cHL</td>
<td>1.05 (0.92 to 1.20)</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6903608 (HLA-DRA)</td>
<td>EBV-positive cHL</td>
<td>1.05 (0.85 to 1.30)</td>
<td>0.65</td>
<td>0.92 (0.65 to 1.30)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EBV-negative cHL</td>
<td>2.16 (1.90 to 2.46)</td>
<td>1.1 (\times) 10^{-4}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs20541 (IL13)</td>
<td>EBV-positive cHL</td>
<td>1.12 (1.08 to 1.43)</td>
<td>0.37</td>
<td>0.84 (0.55 to 1.27)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EBV-negative cHL</td>
<td>1.54 (1.33 to 1.78)</td>
<td>1.0 (\times) 10^{-8}</td>
</tr>
</tbody>
</table>

* Analysis included only patients with MCHL and NSHL that have EBV status available (MCHL: 89 EBV-positive and 65 EBV-negative patients; NSHL: 146 EBV-positive and 563 NSHL control subjects). ORs and 95% CIs were derived using logistic regression assuming a log-additive genetic model of inheritance and adjusting for sex (male or female), country (four indicator variables after excluding one country as the reference: Spain, United Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors. \(P_{\text{MCHL} / \text{NSHL}}\) comparing the ORs between analyses of EBV-positive and EBV-negative cHL are from two-sided \(\chi^2\) tests of homogeneity. All statistical tests were two-sided. CI = confidence interval; MCHL = mixed cellularity Hodgkin lymphoma; NSHL = nodular sclerosis Hodgkin lymphoma; OR = odds ratio.  
† \(P_{\text{MCHL} / \text{NSHL}}\) comparing the ORs between analyses of MCHL and NSHL are from two-sided \(\chi^2\) tests of homogeneity.

**Further investigation.** Previous epidemiological studies have reported associations between cHL risk and a personal and family history of certain autoimmune diseases, namely rheumatoid arthritis and systemic lupus erythematosus, although the reports have been less consistent for psoriasis (56). A large retrospective cohort study conducted in the United Kingdom reported a statistically significant increased risk of HL in psoriasis patients (57). However, there is uncertainty about whether the putative relationship may be attributed to the adverse effects of psoriasis treatment or whether there is some pathophysiological feature common to both diseases (57, 58). Nevertheless, the overlap in genetic risk factors indicated in our study provides support for possible similarities in molecular pathways involved in the pathogenesis of these immune-related diseases.

Our study has several limitations. Although this is the largest GWAS of cHL to date and the first to evaluate EBV status-specific heterogeneity, there may have been insufficient statistical power to detect certain associations within the analysis of the less common cHL subtypes (eg, EBV-positive cHL) and for rarer SNPs and those associated with smaller risk estimates. Assuming a SNP allele frequency of approximately 0.20, the current GWAS sample size had adequate power (>70%) to detect statistically significant (\(P < 1 \times 10^{-8}\)) associations for risk estimates as low as 1.4, 1.9, and 1.5 in the analyses of total cHL, EBV-positive cHL, and EBV-negative cHL, respectively. Presence of systematic bias because of differences in genotyping between patients and the generic control subjects is a possibility. However, consistency with previous literature and successful validation of genotyping together with replication of novel SNP loci in this study make it unlikely that the reported SNPs are false-positive findings. The patients included in this GWAS analysis, with the exception of those from EPILYMPH, have contributed to an independent replication series for a previous GWAS of cHL (11). In consideration of this overlap, we were careful in the interpretation of our findings and pursued only novel loci in the independent replication stage, which included two additional MHC region loci associated irrespective of EBV status, and SNPs at chromosome 5q31 (IL13) and 5q15 (ERAPI).

In summary, this GWAS identified susceptibility loci that showed markedly heterogeneous associations with cHL by...
tumor EBV status. Furthermore, we provided evidence for an overlap in associated loci between cHL and psoriasis. This may indicate some common biological mechanism between the two diseases that can further aid our understanding of cHL pathogenesis.

References

43. Taylor GM, Gokhale DA. Increased frequency of HLA-DRB1*0301 in Hodgkin's disease suggests that susceptibility is HVR-sequence and subtype associated: response to Dorak, Mills, Poynton and Burnett. Leukemia. 1996;10(11):1850–1851.

### Funding

This work was supported by L’Institut National du Cancer, France; Spanish Ministry of Health grant CIBERESP (FIS 08-1555 and 06/02/0073 to S.D.S.); Ministry of Health of the Czech Republic (MZO MOU2005 to L.F.); German José Carreras Leukemia Foundation (DJCLS_R04/08 to A.N.); Federal Office for Radiation Protection (StSch4261 and StSch4420 to N.B.); European Commission 5th Framework Program Quality of Life (QLK4-CT-2000-00422 to P.Br.); European Commission 6th Framework Program (FP6-2003-FOOD-2-B to P.Br.); La Fondation de France (1999-008471 to M.Ma.); Compagnia di San Paolo - Programma Oncologia (P.C.); Health Research Board, Ireland (A.S.); Leukaemia Research Fund (08031 and 05045 to R.F.J.); Kay Kendall Leukaemia Fund (R.F.J. and M.T.); United States National Institutes of Health (R01 CA96290 to M.Me.); Nordic Cancer Union (16-02-D to H.H.); Plan Danmark; Danish Cancer Research Foundation (41-08 to M.Me.); Lundbeck Foundation (R19-A2364 to H.H.); Danish Cancer Society (DP 08-155 to H.H.); Swedish Cancer Society (2009/1084 to K.E.S.); Dutch Cancer Society (KWF grants RUG 2000-2315 and RUG 2010-4860 to A.v.d.B.); the Netherlands Organisation of Scientific Research (NWO-MW grant 920-03-136 to A.D.); Leukaemia & Lymphoma Research (00/73 and 0601 to E.Ro.).

### Notes

K. Y. Urayama, R. F. Jarrett, H. Hjalgrim, and A. Diepstra contributed equally to this work. This study used data generated by the Wellcome Trust Case–Control Consortium (http://www.wtcc.org.uk). The sponsors had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the article; and the decision to submit the article for publication.

### Affiliations of authors:

Genetics Section, International Agency for Research on Cancer, Lyon, France (KU, AG, VG, PB, JDM); MRC University of Glasgow Centre for Virus Research, University of Glasgow, Glasgow, UK (RJF, LS, AL, DM); Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark (HH, BF, MMc); Department of Pathology and Medical Biology (AD, LV), Upptäckta, Sverige (FH A); Department of Immunology and Department of Hematology (GWvU), University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands; Fondation Jean Dausset, Centre d’Étude du Polymorphisme Humain (CEPH), Paris, France (YK, ML); Centre National de Génotypage, Institut Génomique, Commissariat à l’énergie Atomique, Evry, France (AB, DZ, ML); Centre of Chronic Immunodeficiency, University Medical Centre Freiburg, Freiburg, Germany (AN); Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany (NI); Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic (LF); Unit of Infections and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of Oncology (ICO), Barcelona, Spain (YB, SDS); CIBER Epidemiologia y Salud Pública (CIBERESP), Madrid, Spain (YB, M-JS, JMH, EA, MD, SDS); Registry of Hematological Malignancies, EA4184, University of Burgundy, University Hospital of Dijon, Dijon, France (MVa); School of Nursing, Dublin City University, Dublin, Ireland (AS); Cancer Immunogenetics Group, School of Cancer Sciences, University of Manchester, St Mary’s Hospital, Manchester, UK (MT); Department of Medicine-Solina, Clinical Epidemiology Unit, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden (KES); Department of Immunology, Genetics, and Pathology, Uppsala University, Uppsala, Sweden (KG); Department of Medical and Biostatistics, Karolinska Institutet, Stockholm, Sweden (HOA); Department of Epidemiology, Harvard School of Public Health, Boston, MA (HOA); Department of Radiology, Oncology, and Radiation Science, Uppsala University, Uppsala, Sweden (BG); Epidemiology and Genetics Unit, Department of Health Sciences, University of York, York, UK (TL, ERO); Department of Public Health, Occupational Health Section, University of Cagliari, Cagliari, Italy (PC); Department of Epidemiology and Biostatistics, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands (LK, SHV); Institute of Hygiene and Epidemiology, 1st Faculty of Medicine, Charles University, Prague, Czech Republic (IH); Department of Public Health, Norwegian University of Science and Technology, Trondheim, Norway (LVa); Epidemiology Group, School of Medicine and Dentistry, University of Aberdeen, Aberdeen, UK (GJM); Oral & Maxillofacial Surgery, School of Dental Sciences, Newcastle University, Newcastle, UK (PT); Dental School, College of Medical, Veterinary & Life Sciences, University of Glasgow, Glasgow, UK (DIC); Unit of Genetic Variability and Human Diseases, INSERM U946, Paris, France (SB); Unit of Nutrition, Environment and Cancer (IDIBELL), RETICC DR06-0020, Catalan Institute of Oncology (ICO), Barcelona, Spain (AA, CAG); Division of Oral and Maxillofacial Surgery, Oral Medicine and Oral Pathology, Trinity College School of Dental Science, Dublin, Ireland (CMH); Department of Epidemiology and Social Medicine, Aarhus University, Aarhus, Denmark (KO); Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark (AT); Department of Radiation Sciences, Oncology, University of Umeå, Umeå, Sweden (BM); Genomic Epidemiology Group, German Cancer Research Center, Heidelberg, Germany (FC); School of Clinical Medicine, University of Cambridge, Cambridge, UK (K-TK); Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK (RCT); Julius
Center for Health Sciences and Primary Care, University Medical Center of Utrecht, Utrecht, the Netherlands (PHMP); Public Health Directorate, Asturian Council for Health and Health Care, Asturias, Spain (JRO); Andalusian School of Public Health, Granada, Spain (M-JS); Department of Epidemiology, Murcia Regional Health Authority, Murcia, Spain (JMH); Health Department, Navarra Public Health Institute, Pamplona, Spain (EA); Public Health Division of Gipuzkoa, Basque Regional Health Department, San Sebastian, Spain (MD); EPIC Group Institut Gustave Roussy, INSERM U1018, Villejuif, France (FC-C); Centre for Nutrition and Health, National Institute of Public Health and the Environment (RIVM), Bilthoven, the Netherlands (HBB-d-M); Department of Gastroenterology and Hepatology, University Medical Centre Utrecht (UMCU), Utrecht, the Netherlands (HBB-d-M); School of Public Health, Imperial College, London, UK (ERi); The Tisch Cancer Institute and Institute for Translational Epidemiology, Mount Sinai School of Medicine, New York City, New York (PBo); International Prevention Research Institute, Lyon, France (PBo).