FINAL VERSION FOR COPYEDITING (JNCI 11-0690R, AB, 11/17/11)

Genome-Wide Association Study of Classical Hodgkin Lymphoma and Epstein-Barr Virus Status-Defined Subgroups

Kevin Y. Urayama[1†], Ruth F. Jarrett[2†], Henrik Hjalgrim[3†], Arjan Diepstra[4†], Yoichiro Kamatani[5], Amelie Chabrier[1], Valerie Gaborieau[1], Anne Boland[6] , Alexandra Nieters[7], Nikolaus Becker[8], Lenka Foretova[9], Yolanda Benavente[10,11], Marc Maynadié[12] , Anthony Staines[13], Lesley Shield[2], Annette Lake[2], Dorothy Montgomery[2], Malcom Taylor[14], Karin Ekström Smedby[15], Rose-Marie Amini[16], Hans-Olov Adami[17,18], Bengt Glimelius[19], Bjarke Feenstra[3], Ilja M. Nolte[20], Lydia Visser[4], Gustaaf W. van Imhoff[21], Tracy Lightfoot[22], Pierluigi Cocco[23], Lambertus Kiemeney[24], Sita H. Vermeulen[24], Ivana Holcatova[25], Lars Vatten[26], Gary J. Macfarlane[27], Peter Thomson[28], David I. Conway[29], Simone Benhamou[30], Antonio Agudo[31], Claire M. Healy[32], Kim Overvad[33], Anne Tjønneland[34], Beatrice Melin[35] , Federico Canzian[36] , Kay-Tee Khaw[37], Ruth C. Travis[38], Petra H.M. Peeters[39], Carlos A. González[40], José Ramón Quirós[41], María-José Sánchez[42,11] , José María Huerta[43,11], Eva Ardanaz[44,11], Miren Dorronsoro[45,11], Françoise Clavel-Chapelon[46], H. Bas Bueno-de-Mesquita[47,48], Elio Riboli[49], Eve Roman[22], Paolo Boffetta[50,51], Silvia de Sanjosé[10,11], Diana Zelenika[6], Mads Melbye[3], Anke van den Berg[4], Mark Lathrop[5,6], Paul Brennan[1*], James D. McKay[1]

Author affiliations:
1. International Agency for Research on Cancer, Lyon, France
2. MRC University of Glasgow Centre for Virus Research, Garscube Estate, University of Glasgow, Glasgow, United Kingdom
3. Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark
4. Department of Pathology and Medical Biology, University Medical Centre Groningen, University of Groningen, Groningen, Netherlands
5. Fondation Jean Dausset- Centre d'Etude du Polymorphisme Humain (CEPH), Paris, France
6. Centre National de Génotypage, Institut Génomique, Commissariat à l'énergie Atomique, Evry, France
7. Centre of Chronic Immunodeficiency Freiburg, University Medical Centre Freiburg, Freiburg, Germany
8. Division of Cancer Epidemiology, German Cancer Research Center, Hiedelberg, Germany
9. Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic
10. Unit of Infections and Cancer, Cancer Epidemiology Research Programme, Catalan Institute of Oncology (ICO), Barcelona, Spain
11. CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
12. Registry of Hematological Malignancies, EA4184, University of Burgundy, University Hospital of Dijon, Dijon, France
13. School of Nursing, Dublin City University, Glasnevin, Dublin 9, Ireland

14. School of Cancer Sciences, University of Manchester, St Mary's Hospital, Manchester, United Kingdom
15. Department of Medicine Solna, Clinical Epidemiology Unit, Karolinska Institute and Karolinska University Hospital, Stockholm, Sweden
16. Department of Oncology and Pathology, Karolinska University Hospital, Stockholm, Sweden
17. Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden
18. Department of Epidemiology, Harvard School of Public Health, Massachusetts, United States
19. Department of Radiology, Oncology, and Radiation Sciences, Uppsala University, Uppsala, Sweden
20. Unit of Genetic Epidemiology and Bioinformatics, University Medical Centre Groningen, University of Groningen, Groningen, Netherlands
21. Department of Hematology, University Medical Centre Groningen, University of Groningen, Groningen, Netherlands
22. Epidemiology and Genetics Unit, Department of Health Sciences, University of York, York, United Kingdom
23. Department of Public Health, Occupational Health Section, University of Cagliari, Cagliari, Italy
24. Department of Epidemiology and Biostatistics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands
25. Institute of Hygiene and Epidemiology, 1st Faculty of Medicine Charles University, Prague, Czech Republic
26. Norwegian University of Science and Technology, Trondheim, Norway
27. School of Medicine and Dentistry, University of Aberdeen, Aberdeen, United Kingdom
28. University of Newcastle Dental School, Newcastle, United Kingdom
29. University of Glasgow Dental School, Glasgow, Scotland
30. INSERM U946, Paris, France
31. Unit of Nutrition, Environment and Cancer, Catalan Institute of Oncology (ICO), Barcelona, Spain
32. Trinity College School of Dental Science, Dublin, Ireland
33. Department of Epidemiology and Social Medicine, Aarhus University, Aarhus, Denmark
34. Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark
35. Department of Radiation Sciences, Oncology, University of Umeå, Umeå, Sweden
36. Genomic Epidemiology Group, German Cancer Research Center, Hiedelberg, Germany
37. School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom
38. Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, United Kingdom
39. Julius Center for Health Sciences and Primary Care, University Medical Center of Utrecht, Utrecht, The Netherlands
40. Unit of Nutrition, Environment and Cancer (IDIBELL, RETICC DR06-0020), Catalan Institute of Oncology (ICO), Barcelona, Spain
41. Asturian Council for Health and Health Care, Asturias, Spain
42. Andalusian School of Public Health, Granada, Spain
43. Epidemiology Department, Murcia Regional Health Authority, Murcia, Spain
44. Navarra Public Health Institute, Pamplona, Spain
45. Public Health Division of Gipuzkoa, Basque Regional Health Department, San Sebastian, Spain
46. EPIC Group Institut Gustave Roussy, INSERM, U1018, Villejuif, France

47. Centre for Nutrition and Health, National Institute of Public Health and the Environment (RIVM), Bilthoven, The Netherlands
48. Department of Gastroenterology and Hepatology, University Medical Centre Utrecht (UMCU), Utrecht, The Netherlands
49. School of Public Health, Imperial College, London, United Kingdom
50. The Tisch Cancer Institute and Institute for Translational Epidemiology, Mount Sinai School of Medicine, New York, United States
51. International Prevention Research Institute, Lyon, France

[†]These authors contributed equally to this work.

[*]Corresponding Author: Paul Brennan, PhD

International Agency for Research on Cancer

Genetic Epidemiology Group

150 Cours Albert Thomas

69008 Lyon FRANCE

Tel: +33 (0) 4 72 73 83 91

Fax: +33 (0) 4 72 73 83 42

Email: brennan@iarc.fr

Abstract

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 (MHC) region; one resides adjacent to *MICB* (rs2248462: OR = 0.61, 95% CI = 0.53 to 0.69, $P = 1.3 \times 10^{-13}$) and the other at *HLA-DRA* (rs2395185: OR = 0.56, 95% CI = 0.50 to 0.62, $P = 8.3 \times 10^{-25}$) 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, $P = 1.2 \times 10^{-15}$; rs6904029, *HCG9*: OR = 0.46, 95% CI = 0.36 to 0.59, $P = 5.5 \times 10^{-10}$) 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, $P = 6.1 \times 10^{-31}$). The association between rs6903608 and EBV-negative cHL was confined to the nodular sclerosis histologic subtype. Evidence for an association between EBV-negative cHL and rs20541 (5q31, *IL13*: OR = 1.53, 95% CI = 1.32 to 1.76, $P = 5.4 \times 10^{-9}$), 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, $P = 1.5 \times 10^{-4}$) 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 -negative patients.

Introduction

Hodgkin lymphoma (HL) is a malignant condition of B lymphocyte origin and 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 clinicopathological 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 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 -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, rs6903608, 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

Genome-wide Association Study 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 to 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 (15); the International Agency for Research on Cancer Central Europe Study (16); the Pancreatic Cancer Cohort Consortium (17); the Nijmegen Biomedical Study (18); the Nord-Trøndelag Health Study and Tromsø IV Studies (19, 20); 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 (Evry, France, http://www.cng.fr/) as described previously (22). These case patients were matched with control populations genotyped on Illumina BeadChips (Illumina Inc.) from six previous GWAS (described above and Table 1).

We conducted systematic quality control steps separately within case patient samples and each of the six different control sample series. A single nucleotide polymorphism (SNP) was excluded if it was not successfully genotyped in 98% or more of samples, the distribution of genotypes clearly deviated from that expected by Hardy-Weinberg equilibrium ($P < 1 \times 10^{-7}$), or the minor allele frequency was less than 0.01. These exclusions resulted in a total of 502514 SNPs with at least partial genotype data available for both case patients and control subjects.

In addition, quality control metrics were applied to the 1238 cHL case patients resulting in the exclusion of nine samples with less than a 95% overall genotyping success rate, six samples with outlying autosomal heterozygosity rates (>6 SDs from the mean), and three self-reported males and two females with abnormal X-chromosome heterozygosity rates of greater than 10% and less than 20%, respectively. Relatedness was evaluated on the basis of an identity by descent analysis (evaluation of whether loci are identical copies of the same ancestral allele) (23) leading to the exclusion of an additional five samples (ie, first-degree relatives) (Supplementary Table 1, available online).

To maximize the ethnic homogeneity of the case patient and control subject series, we performed a principal components analysis using a subseries of 11029 SNPs present on all Illumina BeadChip panels (Illumina Inc.) that are evenly distributed across the genome in low linkage disequilibrium (LD) ($r^2 < 0.004$) (24) and that passed quality control requirements. The principal components analysis was conducted using EIGENSTRAT of the EIGENSOFT 3.0

software package (Broad Institute, Cambridge, MA) (25). This resulted in further exclusion of 62 outliers (13 patients and 49 control subjects) resulting in a total of 1200 cHL case patients and 6417 control subjects available for the genome-wide analysis (Supplementary Table 1, available online).

Technical Replication

The technical replication stage involved re-genotyping all 1200 cHL patients included in the GWAS and genotyping for the first time 1395 study-specific control subjects for 20 SNPs that reached genome-wide significance in the analysis of total cHL or the subgroup analyses by tumor EBV status. These included 19 SNPs from the MHC region (selected from among greater than 100 genome-wide significant MHC region SNPs based on LD pruning at $r^2 < 0.2$ with an effort to retain SNPs with the lowest *P* values and one SNP at chromosomal region 5q31 (*IL13*, rs20541). Technical replication genotyping was performed using the Sequenom platform (Sequenom Inc., San Diego, CA) at the Centre National de Génotypage (Supplementary Table 2, available online). This stage served two purposes: 1) it allowed for the assessment of genotyping error in the GWAS through re-genotyping select SNPs by an alternate genotyping technology, and 2) it allowed us to confirm the GWAS findings in an analysis of GWAS case patients and study-specific control subjects and aided in the evaluation of bias possibly introduced by the use of generic control subjects in the GWAS.

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 (*ERAP1*), 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 rs2395185 (6p21: *HLA-DRA*), 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___2259921_20 for rs20541, C___16026475_10 for rs2248462, and C___16222465_10 for rs2395185). 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 (United States 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

Hodgkin lymphoma 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 Hodgkin lymphoma 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; eight indicator variables after excluding one country as the reference) 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 1.0×10^{-7} were considered statistically significant at the genome-wide level (21).

Genome-wide analyses were conducted for total cHL and for EBV-positive cHL 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 χ^2test of homogeneity was used to test the null hypothesis that the subtype specific ORs 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).

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 ($\lambda = 1.04$ 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, $P = 8.3 \times 10^{-25}$), 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 ($P_{\text{homogeneity}} < .001$) (Table 2 and Figure 3). Most notable was the variant rs2734986 (OR = 2.45, 95% CI = 2.00 to 3.00, $P = 1.2 \times 10^{-15}$) 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, $P = 5.5 \times 10^{-10}$). 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, $P = 3.2 \times 10^{-27}$), 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 ($P_{homogeneity} < .001$). 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, $P = 6.1 \times 10^{-31}$) (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 all were independently associated with the risk of total cHL and/or EBV subtype-specific cHL (Table 2).

Non-MHC Regions

The genome-wide analysis also indicated a strong signal originating from the chromosomal region 5q31 in proximity to the *IL13* and *IL4* genes (Figure 1). An increased risk of cHL was associated with the minor allele of rs20541 (Table 3 and Figure 4), which appeared predominately in EBV-negative cHL (OR = 1.53, 95% CI = 1.32 to 1.76, $P = 5.4 \times 10^{-9}$). 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 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, $P = 1.5 \times 10^{-4}$) (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, *PTPN22*) also showed borderline statistical significance after a Bonferroni correction (P = 3.7 x 10^{-3}) (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 x 10^{-9}; EBV-negative cHL: OR = 1.47, 95% CI = 1.29 to 1.68, P = 1.1 x 10^{-8}), 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 histologic 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 ($P_{homogeneity}$ = .002).

The EBV-positive cHL associations with rs2734986 and rs6904029 were observed in both MCHL and NSHL subtypes. However, there was some suggestion that the rs2734986 EBV-positive relationship may be stronger in MCHL compared with NSHL ($P_{\text{homogeneity}} = .09$).

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 region supported by five loci independently associated with total cHL or EBV status-specific cHL: rs2248462 (class I region in proximity to *MICB*) and rs2395185 (class II region, *HLA-DRA*) are novel loci associated with cHL, irrespective of EBV status. rs2734986 and rs6904029 (class I region, *HLA-A*) were associated with EBV-positive cHL, and rs6903608 (class II region, *HLA-DRA*) was associated with EBV-negative cHL and was associated predominately with the NSHL histological subtype. We also 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 in 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 between 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 current 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 rs6904029 (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 histologic subtype, indicating 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 non-synonymous SNP in the *IL13* gene (rs20541). rs20541 causes the replacement of a positively charged arginine with a neutral glutamine which 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 1 (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) 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 down-regulation 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 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 UK 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 ($P < 1 \times 10^{-7}$) 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 current study make it unlikely that the reported SNPs are false-positive findings. The patients included in this GWA 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 (*ERAP1*).

In summary, this GWAS identified susceptibility loci that appeared to be common across subtypes of cHL, but also noted 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.

Funding

This work was supported by INCa, France; Spanish Ministry of Health grant CIBERESP (FIS 08-1555 and 06/02/0073 to S.S.); Ministry of Health of the Czech Republic (MZ0 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. Brennan); European Commission 6th Framework Program (FP6-2003-FOOD-2-B to P. Boffetta); La Fondation de France (1999-008471 to M. Maynadie); Compagnia di San Paolo - Programma Oncologia (P.C.); Health Research Board, Ireland (A.S.); Leukaemia Research Fund (08031 and 05045 to R.J.); Kay Kendall Leukaemia Fund (R.J. and M.T); United States National Institutes of Health (R01 CA69269 to M. Melbye), Nordic Cancer Union (16-02-D to H.H.); Plan Danmark; Danish Cancer Research Foundation (41-08 to M. Melbye); 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 06001 to E.R.).

Notes

K. Urayama, R. 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.wtccc.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.

REFERENCES

1. Curado MP, Edwards B, Shin HR, et al. Cancer incidence in five continents. Volume IX. Lyon: IARC Scientific Publications; 2007.

2. Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Fourth Edition. Lyon: IARC Press; 2008.

3. Kuppers R. The biology of Hodgkin's lymphoma. *Nat Rev Cancer.* 2009;9(1):15-27.

4. Proceedings of the IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Epstein-Barr Virus and Kaposi's Sarcoma Herpesvirus/Human Herpesvirus 8. Lyon, France, 17-24 June 1997. *IARC Monogr Eval Carcinog Risks Hum.* 1997;70:1-492.

5. Gulley ML, Eagan PA, Quintanilla-Martinez L, et al. Epstein-Barr virus DNA is abundant and monoclonal in the Reed-Sternberg cells of Hodgkin's disease: association with mixed cellularity subtype and Hispanic American ethnicity. *Blood.* 1994;83(6):1595-1602.

6. 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(11):811-816.

7. Glaser SL, Lin RJ, Stewart SL, et al. Epstein-Barr virus-associated Hodgkin's disease: epidemiologic characteristics in international data. *Int J Cancer.* 1997;70(4):375-382.

8. Diepstra A, Niens M, Vellenga E, et al. Association with HLA class I in Epstein-Barr-virus-positive and with HLA class III in Epstein-Barr-virus-negative Hodgkin's lymphoma. *Lancet.* 2005;365(9478):2216-2224.

9. Hjalgrim H, Rostgaard K, Johnson PC, et al. HLA-A alleles and infectious mononucleosis suggest a critical role for cytotoxic T-cell response in EBV-related Hodgkin lymphoma. *Proc Natl Acad Sci U S A.* 2010;107(14):6400-6405.

10. 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(9):3310-3315.

11. 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(12):1126-1130.

12. Besson H, Brennan P, Becker N, et al. Tobacco smoking, alcohol drinking and Hodgkin's lymphoma: a European multi-centre case-control study (EPILYMPH). *Br J Cancer.* 2006;95(3):378-384.

13. Alexander FE, Lawrence DJ, Freeland J, et al. An epidemiologic study of index and family infectious mononucleosis and adult Hodgkin's disease (HD): evidence for a specific association with EBV+ve HD in young adults. *Int J Cancer.* 2003;107(2):298-302.

14. Smedby KE, Hjalgrim H, Melbye M, et al. Ultraviolet radiation exposure and risk of malignant lymphomas. *J Natl Cancer Inst.* 2005;97(3):199-209.

15. Lagiou P, Georgila C, Minaki P, et al. Alcohol-related cancers and genetic susceptibility in Europe: the ARCAGE project: study samples and data collection. *Eur J Cancer Prev.* 2009;18(1):76-84.

16. Hashibe M, McKay JD, Curado MP, et al. Multiple ADH genes are associated with upper aerodigestive cancers. *Nat Genet.* 2008;40(6):707-709.

17. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet.* 2009;41(9):986-990.

18. Wetzels JF, Kiemeney LA, Swinkels DW, Willems HL, den Heijer M. Age- and gender-specific reference values of estimated GFR in Caucasians: the Nijmegen Biomedical Study. *Kidney Int.* 2007;72(5):632-637.

19. Holmen TL, Barrett-Connor E, Holmen J, Bjermer L. Health problems in teenage daily smokers versus nonsmokers, Norway, 1995-1997: the Nord-Trondelag Health Study. *Am J Epidemiol.* 2000;151(2):148-155.

20. Kronborg J, Jenssen T, Njolstad I, Toft I, Eriksen BO. Metabolic risk factors associated with serum creatinine in a non-diabetic population. *Eur J Epidemiol.* 2007;22(10):707-713.

21. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447(7145):661-678.

22. Purdue MP, Johansson M, Zelenika D, et al. Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. *Nat Genet.* 2011;43(1):60-65.

23. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81(3):559-575.

24. Yu K, Wang Z, Li Q, et al. Population substructure and control selection in genome-wide association studies. *PLoS One.* 2008;3(7):e2551.

25. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006;38(8):904-909.

26. Willett EV, O'Connor S, Smith AG, Roman E. Does smoking or alcohol modify the risk of Epstein-Barr virus-positive or -negative Hodgkin lymphoma? *Epidemiology.* 2007;18(1):130-136.

27. Hindorff LA, Junkins HA, Hall PN, Mehta JP, Manolio TA. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies. (date last accessed, March 29, 2011).

28. Li X, Howard TD, Zheng SL, et al. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. *J Allergy Clin Immunol.* 2010;125(2):328-335 e311.

29. Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med.* 2010;363(13):1211-1221.

30. Nair RP, Duffin KC, Helms C, et al. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nat Genet.* 2009;41(2):199-204.

31. Strange A, Capon F, Spencer CC, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet.* 2010;42(11):985-990.

32. Armstrong AA, Alexander FE, Cartwright R, et al. Epstein-Barr virus and Hodgkin's disease: further evidence for the three disease hypothesis. *Leukemia.* 1998;12(8):1272-1276.

33. Hjalgrim H, Smedby KE, Rostgaard K, et al. Infectious mononucleosis, childhood social environment, and risk of Hodgkin lymphoma. *Cancer Res.* 2007;67(5):2382-2388.

34. Maggio EM, Stekelenburg E, Van den Berg A, Poppema S. TP53 gene mutations in Hodgkin lymphoma are infrequent and not associated with absence of Epstein-Barr virus. *Int J Cancer.* 2001;94(1):60-66.

35. Gulley ML, Glaser SL, Craig FE, et al. Guidelines for interpreting EBER in situ hybridization and LMP1 immunohistochemical tests for detecting Epstein-Barr virus in Hodgkin lymphoma. *Am J Clin Pathol.* 2002;117(2):259-267.

36. Wang D, Sun Y, Stang P, Berlin JA, Wilcox MA, Li Q. Comparison of methods for correcting population stratification in a genome-wide association study of rheumatoid arthritis:

principal-component analysis versus multidimensional scaling. *BMC Proc.* 2009;3(Suppl 7):S109.

37. Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic Reviews in Health Care: Meta-analysis in Context*. London: BMJ Publishing Group; 2001:285-312.

38. SAS Institute. *SAS statistical software version 9*. Cary, NC: SAS Institute; 2004.

39. de Bakker PI, Ferreira MA, Jia X, Neale BM, Raychaudhuri S, Voight BF. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum Mol Genet.* 2008;17(R2):R122-128.

40. Bahram S, Bresnahan M, Geraghty DE, Spies T. A second lineage of mammalian major histocompatibility complex class I genes. *Proc Natl Acad Sci U S A.* 1994;91(14):6259-6263.

41. 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(3):497-505.

42. Oza AM, Tonks S, Lim J, Fleetwood MA, Lister TA, Bodmer JG. A clinical and epidemiological study of human leukocyte antigen-DPB alleles in Hodgkin's disease. *Cancer Res.* 1994;54(19):5101-5105.

43. Taylor GM, Gokhale DA. Increased frequency of HLA-DPB1*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.

44. 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(9):1405-1411.

45. Vladich FD, Brazille SM, Stern D, Peck ML, Ghittoni R, Vercelli D. IL-13 R130Q, a common variant associated with allergy and asthma, enhances effector mechanisms essential for human allergic inflammation. *J Clin Invest.* 2005;115(3):747-754.

46. Graves PE, Kabesch M, Halonen M, et al. A cluster of seven tightly linked polymorphisms in the IL-13 gene is associated with total serum IgE levels in three populations of white children. *J Allergy Clin Immunol.* 2000;105(3):506-513.

47. He JQ, Chan-Yeung M, Becker AB, et al. Genetic variants of the IL13 and IL4 genes and atopic diseases in at-risk children. *Genes Immun.* 2003;4(5):385-389.

48. Skinnider BF, Elia AJ, Gascoyne RD, et al. Interleukin 13 and interleukin 13 receptor are frequently expressed by Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. *Blood.* 2001;97(1):250-255.

49. Skinnider BF, Mak TW. The role of cytokines in classical Hodgkin lymphoma. *Blood.* 2002;99(12):4283-4297.

50. Ho IC, Tai TS, Pai SY. GATA3 and the T-cell lineage: essential functions before and after T-helper-2-cell differentiation. *Nat Rev Immunol.* 2009;9(2):125-135.

51. Stanelle J, Doring C, Hansmann ML, Kuppers R. Mechanisms of aberrant GATA3 expression in classical Hodgkin lymphoma and its consequences for the cytokine profile of Hodgkin and Reed/Sternberg cells. *Blood.* 2010;116(20):4202-4211.

52. Monroy CM, Cortes AC, Lopez MS, et al. Hodgkin disease risk: role of genetic polymorphisms and gene-gene interactions in inflammation pathway genes. *Mol Carcinog.* 2011;50(1):36-46.

53. Kraft P. Curses--winner's and otherwise--in genetic epidemiology. *Epidemiology.* 2008;19(5):649-651; discussion 657-648.

54. Saveanu L, Carroll O, Lindo V, et al. Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nat Immunol.* 2005;6(7):689-697.

55. Fruci D, Ferracuti S, Limongi MZ, et al. Expression of endoplasmic reticulum aminopeptidases in EBV-B cell lines from healthy donors and in leukemia/lymphoma, carcinoma, and melanoma cell lines. *J Immunol.* 2006;176(8):4869-4879.

56. Landgren O, Engels EA, Pfeiffer RM, et al. Autoimmunity and susceptibility to Hodgkin lymphoma: a population-based case-control study in Scandinavia. *J Natl Cancer Inst.* 2006;98(18):1321-1330.

57. Gelfand JM, Shin DB, Neimann AL, Wang X, Margolis DJ, Troxel AB. The risk of lymphoma in patients with psoriasis. *J Invest Dermatol.* 2006;126(10):2194-2201.

58. Dommasch E, Gelfand JM. Is there truly a risk of lymphoma from biologic therapies? *Dermatol Ther.* 2009;22(5):418-430.

Context and Caveats

Prior knowledge

Hodgkin lymphoma is one of the most common cancers among adolescents and young adults. Epstein-Barr virus (EBV) has been associated with some HL patients who have one of the major forms—classical HL (cHL). Although the relationship between EBV and cHL is not fully understood, it has been hypothesized that genetic factors may be involved.

Study design

To determine if EBV status should be considered when classifying the different subtypes of cHL, a genome-wide association study was done using data from a large series of cHL patients for

whom EBV status was known. The results were validated in an independent replication study and genetic variants associated with cHL, EBV-positive cHL, and EBV-negative cHL were identified.

Contributions

Two variants located in the major histocompatibility complex (MHC) were independently associated with cHL, regardless of EBV status, and were confirmed in the replication series. Two genetic variants within the class I region of the MHC previously identified in reports as having an association with EBV-positive cHL were also observed. One genetic variant in the class II region of the MHC and one variant previously linked to psoriasis and asthma were associated with EBV-negative cHL. These results indicate that although there are some loci associated with cHL, regardless of EBV status, there are also some loci that are heterogeneous in their relationship with cHL EBV subtypes.

Implications

EBV status should be included in the classification of cHL.

Limitations

Although this is the largest GWAS of cHL to date, the statistical power of this study may not have been high enough to detect associations involving rare variants or less common cHL subtypes, such as EBV-positive cHL. Also, systematic bias could have been introduced from differences in genotyping between patients and some control subjects.

Table 1. Subjects included in the genome-wide association and replication analyses examining the relationship between genetic variants and classical Hodgkin lymphoma

Study name, duration	Country	Case patients, No.*				Control subjects, No.	
		EBV+	EBV-	EBV not available	Total	Generic (source) †	Study-specific
GWAS							
Total	NA	265	668	267	1200	6417	1395
EPILYMPH Study, 1998−2003	France, Germany, Spain, Czech Republic, Ireland	8	11	162	181	863 (ARCAGE, PANSCAN, CE)	377
Scotland and Newcastle Lymphoma Group, 1992−2008‡; Young Adult Hodgkin Case−Control Study, 1991−1995	United Kingdom	106	240	45	391	3172 (ARCAGE, PANSCAN, WTCCC)	348
Scandinavian Lymphoma Etiology Study, 1999−2002	Denmark, Sweden	83	211	47	341	578 (PANSCAN, HUNT2/Tromsø)	670
Northern Dutch Hodgkin Lymphoma Study, 1987−2009	The Netherlands	68	206	13	287	1804 (NBS, PANSCAN)	0
Replication							
Total	NA	131	291	141	563	NA	613
EPILYMPH Study, 1998−2003	France, Germany, Spain, Czech Republic, Ireland, Italy	6	11	47	64	NA	396
Scotland and Newcastle	United Kingdom	81	167	34	282	NA	32

Lymphoma Group, 1992−2008‡							
Young Adult Hodgkin Case−Control Study, 1991−1995	United Kingdom	11	43	4	58	NA	10
Epidemiology and Genetics Lymphoma Case−Control Study, 1998−2003	United Kingdom	33	70	56	159	NA	175

*The number of case patients is indicated by Epstein-Barr virus (EBV) status. ARCAGE = Alcohol Related Cancers and Genetic Susceptibility in Europe, used the Illumina HumanHap300 genotyping platform (Illumina Inc., San Diego, CA); CE = National Cancer Institute and International Agency for Research on Cancer Study in Central Europe, used the Illumina HumanHap300 genotyping platform; HUNT2/Tromsø = Nord-Trøndelag Health and Tromsø IV Studies, used the Illumina HumanHap300 genotyping platform; NBS = Nijmegen Biomedical Study, used the Illumina CNV370-Duo genotyping platform (Illumina Inc.); PANSCAN = Pancreatic Cancer Cohort Consortium, used the Illumina HumanHap550 genotyping platform (Illumina Inc.); and WTCCC = Wellcome Trust Case−Control Consortium, used the Illumina Human 1.2M-Duo genotyping platform (Illumina Inc.).

†Generic control subjects were selected on the basis of country of the case patient recruitment from several previously published genome-wide association studies of other cancers.

‡The Scotland and Newcastle Lymphoma Group series is comprised of the Scotland and Newcastle Epidemiological Study of Hodgkin's Disease conducted during 1993−1997 and the "local case series", a prospectively collected series during 1992−2008.

Table 2. Relationship between five single-nucleotide polymorphisms (SNPs) located within the major histocompatibility complex (MHC) region and total classical Hodgkin lymphoma (cHL), Epstein-Barr virus (EBV)-positive cHL, or EBV-negative cHL subgroups

SNP	Position (Region in MHC)	Gene candidate	Frequency*			Single SNP†		Mutually adjusted†	
			Minor allele	Patients	Control subjects	OR (95% CI)	*P*	OR (95% CI)	*P*
Total cHL‡									
rs2248462	31554775 (class I)	*MICB*	A	0.14	0.22	0.61 (0.53 to 0.69)	1.3×10^{-13}	0.71 (0.62 to 0.81)	8.8×10^{-7}
rs2395185	32541145 (class II)	*HLA-DRA*	T	0.21	0.33	0.56 (0.50 to 0.62)	8.3×10^{-25}	0.75 (0.65 to 0.85)	9.6×10^{-6}
EBV+ cHL‡									
rs2734986	29926547 (class I)	*HLA-A*	C	0.34	0.18	2.45 (2.00 to 3.00)	1.2×10^{-15}	1.83 (1.43 to 2.34)	1.5×10^{-6}
rs6904029	30051046 (class I)	*HCG9*	A	0.17	0.30	0.46 (0.36 to 0.59)	5.5×10^{-10}	0.59 (0.45 to 0.76)	7.3×10^{-5}
EBV- cHL‡									
rs6903608	32536263 (class II)	*HLA-DRA*	C	0.46	0.29	2.08 (1.84 to 2.35)	6.1×10^{-31}	1.88 (1.62 to 2.17)	3.1×10^{-17}

* Frequency of the minor allele in case patients and control subjects included in each analysis.

† Odds ratios (ORs) and 95% confidence intervals (CIs) for each SNP in the single SNP analysis were derived using logistic regression assuming a log-additive genetic model of inheritance 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 mutually adjusted analysis included additional adjustment for the effects of all other SNPs in the table. All statistical tests were two-sided.

‡ Indicates the phenotype with which the SNP appears to be predominately associated based on the results of the total cHL and EBV-specific subtype analyses.

Table 3. Summary of results for four previously unreported loci associated with classical Hodgkin lymphoma (cHL) included in the independent replication phase *

SNP (Region)	Position	Gene candidate	Genome-wide analysis†					Independent replication‡		Combined§		P_{hom}§
			Minor allele	Minor allele frequency in patients	Minor allele frequency in control subjects	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	
Total cHL												
rs2248462 (6p21)	31554775	*MICB*	A	0.14	0.22	0.61 (0.53 to 0.69)	1.3×10^{-13}	0.61 (0.46 to 0.83)	1.2×10^{-3}	0.61 (0.54 to 0.69)	6.9×10^{-16}	1.0
rs2395185 (6p21)	32541145	*HLA-DRA*	T	0.21	0.33	0.56 (0.50 to 0.62)	8.3×10^{-25}	0.54 (0.42 to 0.70)	1.6×10^{-6}	0.55 (0.50 to 0.60)	3.8×10^{-31}	.90
rs27524 (5q15)	96127700	*ERAP1*	A	0.40	0.36	1.21 (1.10 to 1.33)	1.5×10^{-4}	1.29 (1.03 to 1.62)	.03	1.22 (1.11 to 1.33)	7.4×10^{-6}	.61
EBV-cHL												
rs20541 (5q31)	132023863	*IL13*	A	0.26	0.18	1.53 (1.32 to 1.76)	5.4×10^{-9}	1.18 (0.83 to 1.67)	.36	1.47 (1.29 to 1.68)	1.1×10^{-8}	.18

* CI = confidence interval, OR = odds ratio, SNP = single nucleotide polymorphism.

† Frequency of the single-nucleotide polymorphism minor allele in case patients and control subjects included in the genome-wide association analysis. 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 (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. All statistical tests were two-sided.

‡ ORs and 95% CIs were derived using logistic regression assuming a log-additive genetic model of inheritance and adjusting for sex and country. All statistical tests were two-sided.

§ Genome-wide association study and replication phase results were combined using inverse variance weighting meta-analysis. Two-sided *P* values (P_{hom}) based on the Cochran Q test statistic were used to evaluate evidence of heterogeneity in risk estimates between the genome-wide association and replication results.

 Indicates the phenotype within which the SNP appears to be predominately associated based on the results of the total cHL and Epstein-Barr virus (EBV)-specific subtype analyses.

Table 4. Results of a genome-wide association study of classical Hodgkin lymphoma (cHL) subtypes defined jointly by tumor Epstein-Barr virus (EBV) status and histology *

SNP (gene candidate) by cHL EBV status	Pooled (NSHL and MCHL)		MCHL (n=154)		NSHL (n=685)		$P_{hom(MC/NS)}$†
	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	
rs2248462 (*MICB)*							
EBV+ cHL	0.58 (0.44 to 0.77)	1.4×10^{-4}	0.56 (0.36 to 0.87)	.01	0.60 (0.42 to 0.85)	4.1×10^{-3}	.81
EBV- cHL	0.58 (0.49 to 0.70)	4.0×10^{-9}	0.57 (0.34 to 0.96)	.04	0.58 (0.48 to 0.71)	2.1×10^{-8}	.95
$P_{hom\ (EBV+/-)}$	.95		.96		.87		
rs2395185 (*HLA-DRA)*							
EBV+ cHL	0.51 (0.40 to 0.65)	4.7×10^{-8}	0.43 (0.29 to 0.64)	3.0×10^{-5}	0.55 (0.41 to 0.74)	6.2×10^{-5}	.33
EBV- cHL	0.58 (0.50 to 0.67)	3.0×10^{-13}	0.68 (0.45 to 1.01)	.05	0.56 (0.48 to 0.66)	5.0×10^{-13}	.38
$P_{hom\ (EBV+/-)}$	.41		.11		.92		
rs27524 (*ERAP1)*							
EBV+ cHL	1.30 (1.07 to 1.58)	.009	1.19 (0.88 to 1.63)	.26	1.35 (1.06 to 1.72)	.016	.53
EBV- cHL	1.23 (1.08 to 1.40)	.001	1.03 (0.71 to 1.48)	.88	1.26 (1.10 to 1.44)	7.0×10^{-4}	.31
$P_{hom\ (EBV+/-)}$	.65		.56		.63		
rs2734986 (*HLA-A)*							
EBV+ cHL	2.60 (2.06 to 3.27)	7.2×10^{-16}	3.35 (2.34 to 4.81)	5.5×10^{-11}	2.25 (1.68 to 3.00)	4.7×10^{-8}	.09
EBV- cHL	0.95 (0.80 to 1.14)	.601	1.13 (0.71 to 1.81)	.62	0.93 (0.77 to 1.12)	.456	.45
$P_{hom\ (EBV+/-)}$	<.001		<.001		<.001		
rs6904029 (*HCG9)*							
EBV+ cHL	0.46 (0.36 to 0.60)	4.4×10^{-9}	0.41 (0.27 to 0.62)	3.0×10^{-5}	0.50 (0.37 to 0.69)	1.9×10^{-5}	.45
EBV- cHL	1.05 (0.92 to 1.20)	.466	0.99 (0.68 to 1.45)	.98	1.06 (0.92 to 1.22)	.403	.74
$P_{hom\ (EBV+/-)}$	<.001		.002		<.001		

rs6903608 (*HLA-DRA)*							
EBV+ cHL	1.05 (0.85 to 1.30)	.652	0.92 (0.65 to 1.30)	.62	1.17 (0.90 to 1.52)	.244	.28
EBV- cHL	2.16 (1.90 to 2.46)	1.1×10^{-31}	1.23 (0.85 to 1.80)	.28	2.33 (2.04 to 2.67)	2.6×10^{-34}	.002
$P_{\text{hom (EBV+/-)}}$	<.001		.27		<.001		
rs20541 (*IL13)*							
EBV+ cHL	1.12 (0.88 to 1.43)	.366	0.84 (0.55 to 1.27)	.401	1.28 (0.96 to 1.72)	.094	.11
EBV- cHL	1.54 (1.33 to 1.78)	1.0×10^{-8}	1.36 (0.90 to 2.07)	.145	1.55 (1.33 to 1.81)	2.1×10^{-8}	.56
$P_{\text{hom (EBV+/-)}}$	.03		.11		.26		

* Analysis included only patients with mixed cellularity Hodgkin lymphoma (MCHL) and nodular sclerosis Hodgkin lymphoma (NSHL) that have Epstein-Barr virus (EBV) status available (MCHL: 89 EBV-positive and 65 EBV-negative patients; NSHL: 146 EBV-positive and 539 EBV-negative patients) and 5653 control subjects. Odds ratios (ORs) and 95% confidence intervals (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. $P_{\text{hom (EBV+/-)}}$ comparing the ORs between analyses of EBV-positive and EBV-negative cHL are from two-sided χ^2 tests of homogeneity. All statistical tests were two-sided.

† $P_{\text{hom (MC/NS)}}$ comparing the ORs between analyses of MCHL and NSHL are from two-sided χ^2 tests of homogeneity.

Figure Legends

Figure 1. Genome-wide association analysis of 502 514 single-nucleotide polymorphisms (SNPs) in 1200 classical Hodgkin lymphoma patients and 6417 control subjects. Multiple logistic regression analysis 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* value) for each SNP is plotted against their chromosomal position. All statistical tests were two-sided. The horizontal dotted line represents the genome-wide significance threshold level ($P < 1 \times 10^{-7}$). The regions in which SNPs associated with classical Hodgkin lymphoma are located include 5q31 in the vicinity of the *IL13* and *IL4* genes, and 6p21 within the major histocompatibility complex (MHC) region.

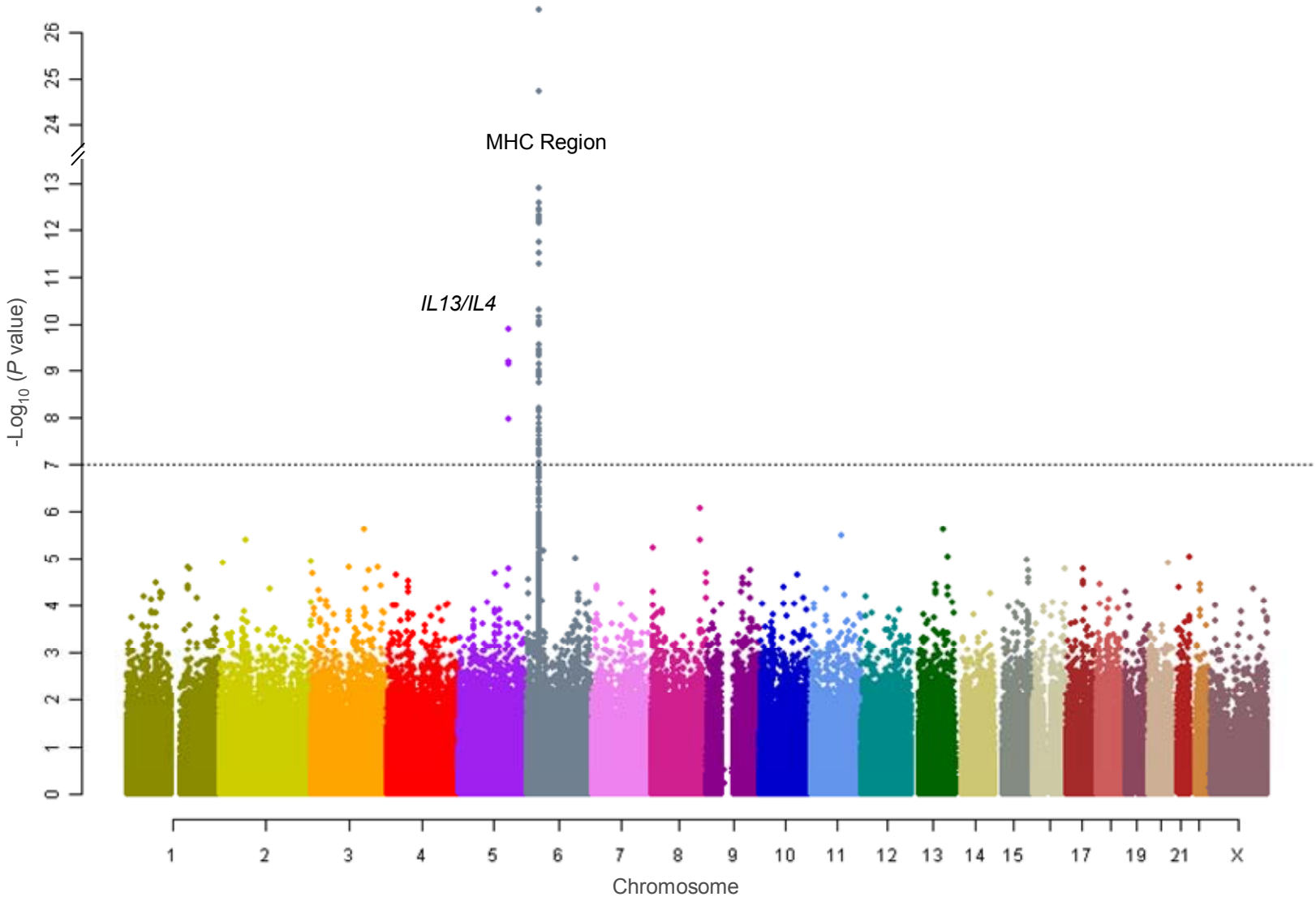
Figure 2. Investigation of the association between genetic variants and total classical Hodgkin lymphoma (cHL), Epstein-Barr virus (EBV)-positive cHL and EBV-negative cHL within an approximately 6.5 Mb region of the extended major histocompatibility complex located at 6p21. Overlayed plots of the $-\log_{10}$(*P* values) for 1973 single-nucleotide polymorphisms (SNPs) by their chromosomal position resulting from three separate analyses of all patients who have classical Hodgkin lymphoma (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),

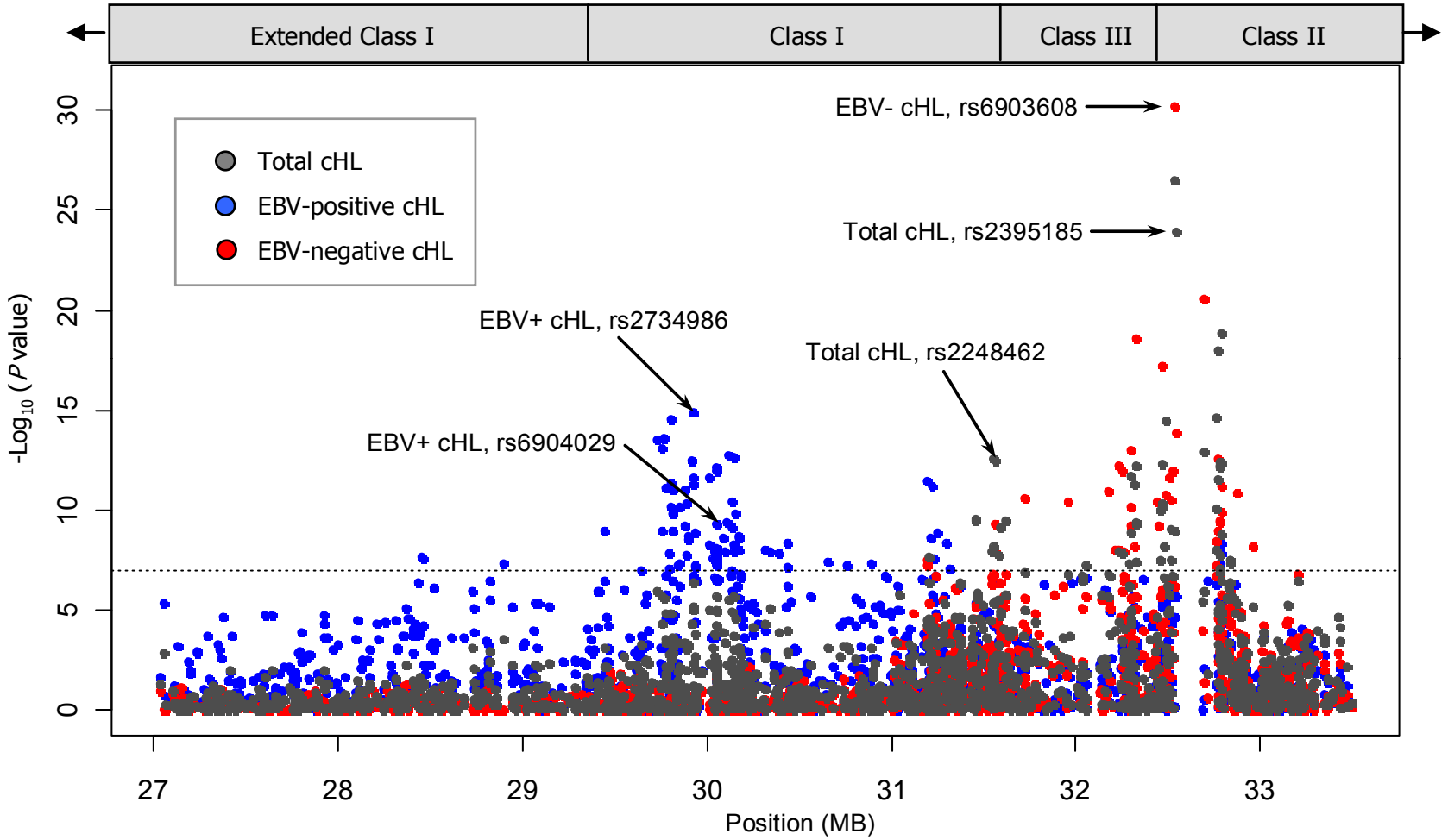
and eight principal components analysis eigenvectors. All statistical tests were two-sided. **Arrows** indicate cHL-associated regions indexed by five SNPs.

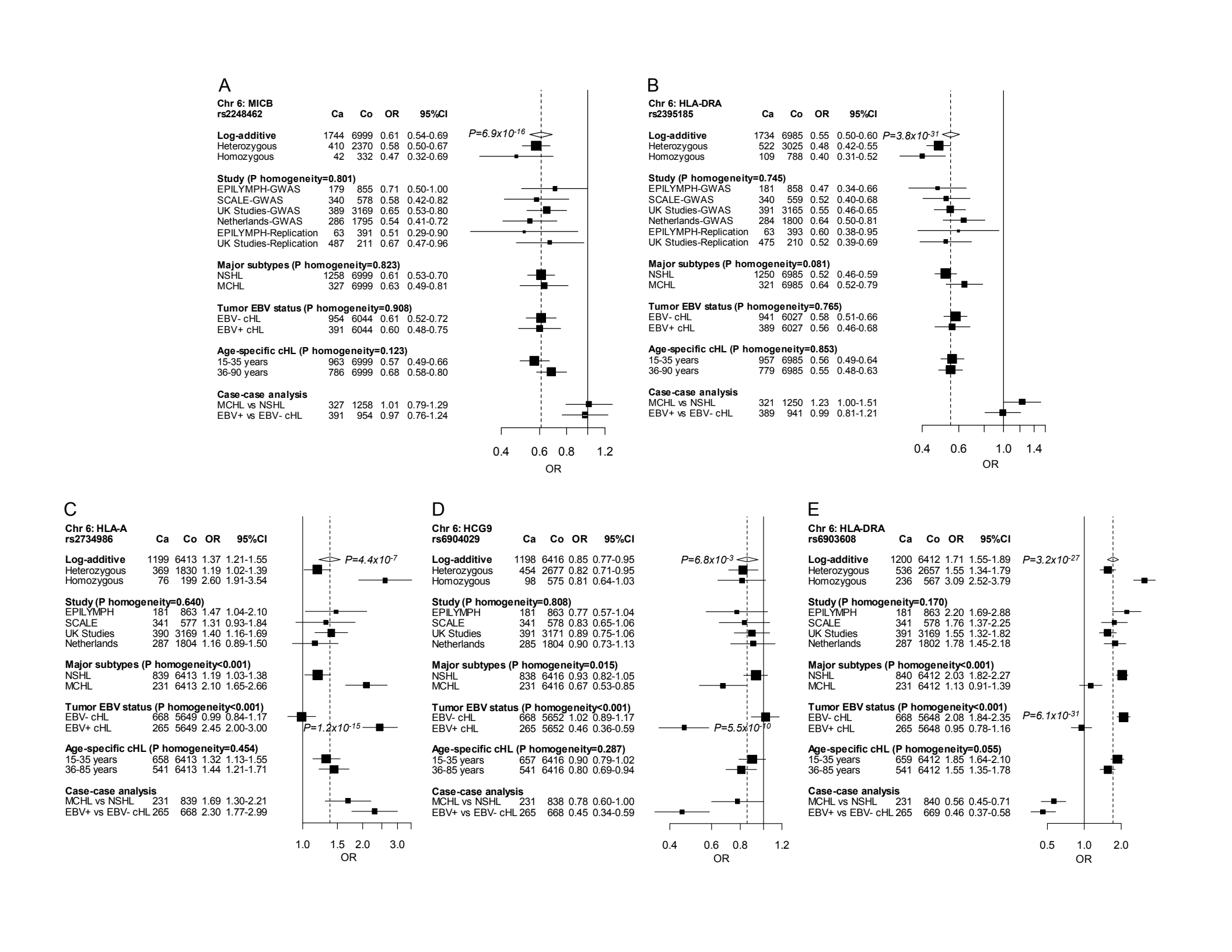
Figure 3. Stratified and subgroup analyses of five single-nucleotide polymorphisms (SNPs) within the major histocompatibility complex region which were independently associated with classical Hodgkin lymphoma (cHL). Results are shown for **A**) rs2248462 (class I region, *MICB*), **B**) rs2395185 (class II region, *HLA-DRA*), **C**) rs2734986 (class I region, *HLA-A*), and **D**) rs6904029 (class I region, *HLA-A*), and **E**) rs6903608 (class II region, *HLA-DRA*). 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 genome-wide association study (GWAS) included the EPILYMPH study (EPILYMH-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 EPILYMPH study (EPILYMPH-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).

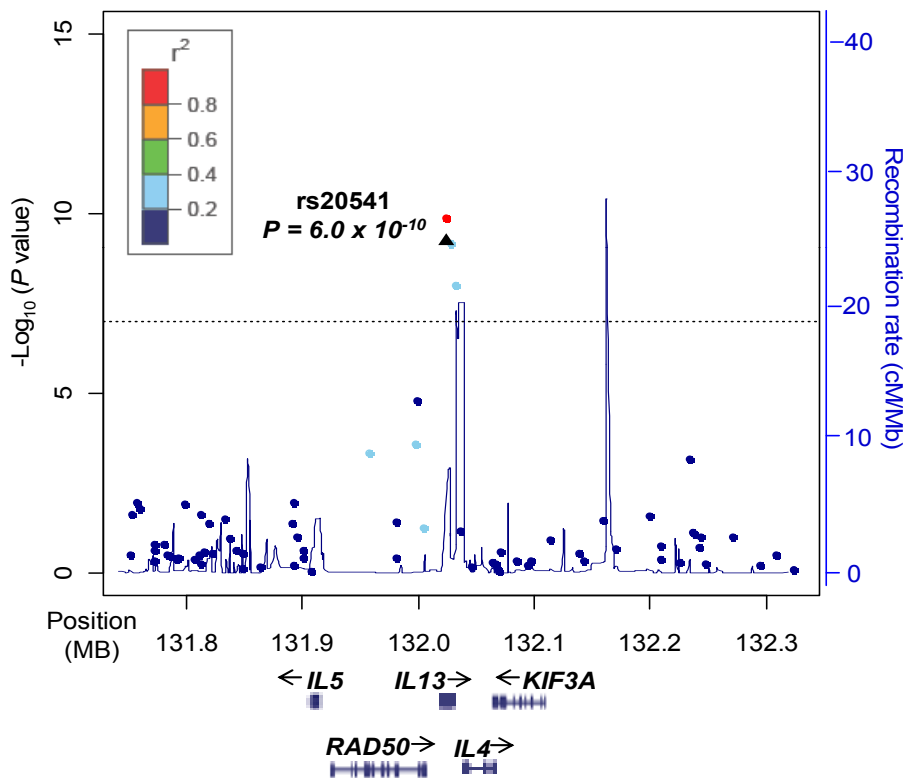
$P_{homogeneity}$ was on the basis of the Cochran's Q test statistic and was used to evaluate between-study heterogeneity in results. Associations between the SNPs and cHL subgroups (including histologic subtype, EBV status, and age) were performed, and *P* values ($P_{homogeneity}$) for the χ^2test of homogeneity are presented to indicate differences in the OR between subgroup analyses. For rs2248462 (**A**) and rs 2395185 (**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.

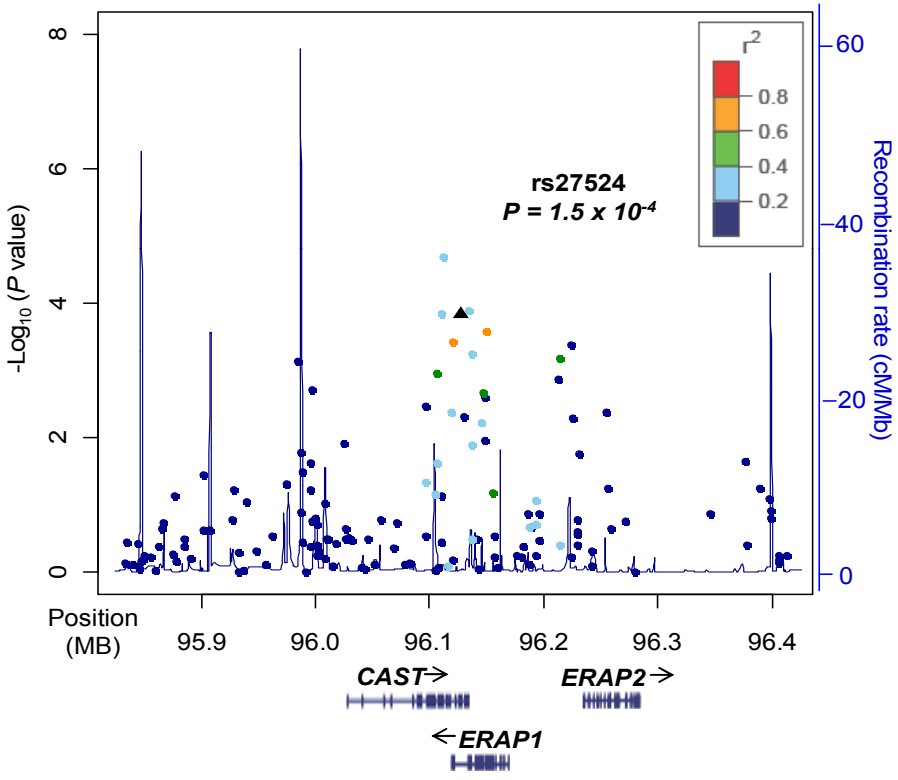
Figure 4. Genome-wide association study (GWAS) results in total classical Hodgkin lymphoma (cHL) for single-nucleotide polymorphisms (SNPs) within a 300 kilobase flanking region of rs20541at chromosome 5q31 (left) and rs27524 at chromosome 5q15 (right). Multiple logistic regression was performed assuming 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* 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 color 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.


MHC Region
IL13/IL4
-Log10 (P value)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
24
25
26
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
17
19
21
X
Chromosome


Extended Class I
Class I
Class III
Class II
Total cHL
EBV-positive cHL
EBV-negative cHL
EBV- cHL, rs6903608
Total cHL, rs2395185
EBV+ cHL, rs2734986
Total cHL, rs2248462
EBV+ cHL, rs6904029
-Log10 (P value)
0
5
10
15
20
25
30
27
28
29
30
31
32
33
Position (MB)


A
Chr 6: MICB
rs2248462 Ca Co OR 95%CI
Log-additive 1744 6999 0.61 0.54-0.69
Heterozygous 410 2370 0.58 0.50-0.67
Homozygous 42 332 0.47 0.32-0.69
Study (P homogeneity=0.801)
EPILYMPH-GWAS 179 855 0.71 0.50-1.00
SCALE-GWAS 340 578 0.58 0.42-0.82
UK Studies-GWAS 389 3169 0.65 0.53-0.80
Netherlands-GWAS 286 1795 0.54 0.41-0.72
EPILYMPH-Replication 63 391 0.51 0.29-0.90
UK Studies-Replication 487 211 0.67 0.47-0.96
Major subtypes (P homogeneity=0.823)
NSHL 1258 6999 0.61 0.53-0.70
MCHL 327 6999 0.63 0.49-0.81
Tumor EBV status (P homogeneity=0.908)
EBV- cHL 954 6044 0.61 0.52-0.72
EBV+ cHL 391 6044 0.60 0.48-0.75
Age-specific cHL (P homogeneity=0.123)
15-35 years 963 6999 0.57 0.49-0.66
36-90 years 786 6999 0.68 0.58-0.80
Case-case analysis
MCHL vs NSHL 327 1258 1.01 0.79-1.29
EBV+ vs EBV- cHL 391 954 0.97 0.76-1.24
P=6.9x10-16
0.4 0.6 0.8 1.2
OR
B
Chr 6: HLA-DRA
rs2395185 Ca Co OR 95%CI
Log-additive 1734 6985 0.55 0.50-0.60
Heterozygous 522 3025 0.48 0.42-0.55
Homozygous 109 788 0.40 0.31-0.52
Study (P homogeneity=0.745)
EPILYMPH-GWAS 181 858 0.47 0.34-0.66
SCALE-GWAS 340 559 0.52 0.40-0.68
UK Studies-GWAS 391 3165 0.55 0.46-0.65
Netherlands-GWAS 284 1800 0.64 0.50-0.81
EPILYMPH-Replication 63 393 0.60 0.38-0.95
UK Studies-Replication 475 210 0.52 0.39-0.69
Major subtypes (P homogeneity=0.081)
NSHL 1250 6985 0.52 0.46-0.59
MCHL 321 6985 0.64 0.52-0.79
Tumor EBV status (P homogeneity=0.765)
EBV- cHL 941 6027 0.58 0.51-0.66
EBV+ cHL 389 6027 0.56 0.46-0.68
Age-specific cHL (P homogeneity=0.853)
15-35 years 957 6985 0.56 0.49-0.64
36-90 years 779 6985 0.55 0.48-0.63
Case-case analysis
MCHL vs NSHL 321 1250 1.23 1.00-1.51
EBV+ vs EBV- cHL 389 941 0.99 0.81-1.21
P=3.8x10-31
0.4 0.6 1.0 1.4
OR
C
Chr 6: HLA-A
rs2734986 Ca Co OR 95%CI
Log-additive 1199 6413 1.37 1.21-1.55
Heterozygous 369 1830 1.19 1.02-1.39
Homozygous 76 199 2.60 1.91-3.54
Study (P homogeneity=0.640)
EPILYMPH 181 863 1.47 1.04-2.10
SCALE 341 577 1.31 0.93-1.84
UK Studies 390 3169 1.40 1.16-1.69
Netherlands 287 1804 1.16 0.89-1.50
Major subtypes (P homogeneity<0.001)
NSHL 839 6413 1.19 1.03-1.38
MCHL 231 6413 2.10 1.65-2.66
Tumor EBV status (P homogeneity<0.001)
EBV- cHL 668 5649 0.99 0.84-1.17
EBV+ cHL 265 5649 2.45 2.00-3.00
Age-specific cHL (P homogeneity=0.454)
15-35 years 658 6413 1.32 1.13-1.55
36-85 years 541 6413 1.44 1.21-1.71
Case-case analysis
MCHL vs NSHL 231 839 1.69 1.30-2.21
EBV+ vs EBV- cHL 265 668 2.30 1.77-2.99
P=4.4x10-7
P=1.2x10-15
1.0 1.5 2.0 3.0
OR
D
Chr 6: HCG9
rs6904029 Ca Co OR 95%CI
Log-additive 1198 6416 0.85 0.77-0.95
Heterozygous 454 2677 0.82 0.71-0.95
Homozygous 98 575 0.81 0.64-1.03
Study (P homogeneity=0.808)
EPILYMPH 181 863 0.77 0.57-1.04
SCALE 341 578 0.83 0.65-1.06
UK Studies 391 3171 0.89 0.75-1.06
Netherlands 285 1804 0.90 0.73-1.13
Major subtypes (P homogeneity=0.015)
NSHL 838 6416 0.93 0.82-1.05
MCHL 231 6416 0.67 0.53-0.85
Tumor EBV status (P homogeneity<0.001)
EBV- cHL 668 5652 1.02 0.89-1.17
EBV+ cHL 265 5652 0.46 0.36-0.59
Age-specific cHL (P homogeneity=0.287)
15-35 years 657 6416 0.90 0.79-1.02
36-85 years 541 6416 0.80 0.69-0.94
Case-case analysis
MCHL vs NSHL 231 838 0.78 0.60-1.00
EBV+ vs EBV- cHL 265 668 0.45 0.34-0.59
P=6.8x10-3
P=5.5x10-10
0.4 0.6 0.8 1.2
OR
E
Chr 6: HLA-DRA
rs6903608 Ca Co OR 95%CI
Log-additive 1200 6412 1.71 1.55-1.89
Heterozygous 536 2657 1.55 1.34-1.79
Homozygous 236 567 3.09 2.52-3.79
Study (P homogeneity=0.170)
EPILYMPH 181 863 2.20 1.69-2.88
SCALE 341 578 1.76 1.37-2.25
UK Studies 391 3169 1.55 1.32-1.82
Netherlands 287 1802 1.78 1.45-2.18
Major subtypes (P homogeneity<0.001)
NSHL 840 6412 2.03 1.82-2.27
MCHL 231 6412 1.13 0.91-1.39
Tumor EBV status (P homogeneity<0.001)
EBV- cHL 668 5648 2.08 1.84-2.35
EBV+ cHL 265 5648 0.95 0.78-1.16
Age-specific cHL (P homogeneity=0.055)
15-35 years 659 6412 1.85 1.64-2.10
36-85 years 541 6412 1.55 1.35-1.78
Case-case analysis
MCHL vs NSHL 231 840 0.56 0.45-0.71
EBV+ vs EBV- cHL 265 669 0.46 0.37-0.58
P=3.2x10-27
P=6.1x10-31
0.5 1.0 2.0
OR


r²
0.8
0.6
0.4
0.2
rs20541
P = 6.0 x 10-10
-Log10 (P value)
Recombination rate (cM/Mb)
Position (MB)
131.8
131.9
132.0
132.1
132.2
132.3
IL5
IL13
KIF3A
RAD50
IL4


r²
0.8
0.6
0.4
0.2
rs27524
P = 1.5 x 10-4
-Log10 (P value)
Recombination rate (cM/Mb)
Position (MB)
95.9
96.0
96.1
96.2
96.3
96.4
CAST
ERAP2
ERAP1

JNCI 11-0690R
Supplementary Materials

Supplementary Table 1. Exclusions performed in the genome-wide association study (GWAS) patients and independent replication control subjects

Reason for exclusion	GWAS case patients, No. (%)	Independent replication series, No. (%)	
		Case patients	Control subjects
Total samples genotyped	1238 (100)	599 (100)	633 (100)
Completely failed	6 (0.48)	10 (1.67)	6 (0.95)
Genotype rate	3 (0.24)	23 (3.84)	13 (2.05)
Autosomal heterozygosity rate	6 (0.48)	NA	NA
Sex discrepancy	5 (0.40)	3 (0.50)	1 (0.16)
Relatedness (first-degree relatives)	5 (0.40)	NA	NA
Ethnic heterogeneity based on PCA*	13 (1.05)	NA	NA
Total included in analysis	1200 (96.9)	563 (93.9)	613 (96.8)

* Principal components analysis (PCA) was conducted using a subseries of 11029 single-nucleotide polymorphisms present on all Illumina BeadChip panels (Illumina Inc., San Diego, CA) that are evenly distributed across the genome in low linkage disequilibrium. NA = not applicable.

Supplementary Table 2. Comparison between genome-wide association study (GWAS) genotypes and technical replication genotyping in case patients and study-specific control subjects for select single-nucleotide polymorphism (SNPs)

Chromosome	SNP	Allele		Case patients		Control subjects	
				Frequency*	Genotype Concordance†	Frequency*	
		A1*	A2	Sequenom	No. of case patients (%)	Generic	Study-specific
5	rs20541	A	G	0.24	1210 (100)	0.19	0.20
6	rs9393777‡	C	T	NA	NA	0.14	NA
6	rs3799499	T	G	0.28	1210 (100)	0.25	0.23
6	rs2523399	C	T	0.50	1151 (100)	0.46	0.43
6	rs2734986‡	C	T	NA	NA	0.17	NA
6	rs6904029	A	G	0.27	1207 (99.5)	0.30	0.31
6	rs1150741	G	C	0.34	NA	NA	0.30
6	rs1245371‡	G	A	NA	NA	0.28	NA
6	rs1049623	C	T	0.40	1205 (99.9)	0.39	0.38
6	rs3094211	G	A	0.30	1209 (100)	0.26	0.26
6	rs3094204	G	A	0.45	1197 (100)	0.49	0.52
6	rs746647	G	A	0.35	1208 (100)	0.31	0.31
6	rs3130542	A	G	0.24	1205 (100)	0.20	0.21
6	rs2248462	A	G	0.14	1204 (100)	0.22	0.20
6	rs206015	A	G	0.07	1208 (100)	0.11	0.12
6	rs443198	G	A	0.32	1172 (99.0)	0.36	0.37
6	rs3134931	C	T	0.22	1177 (99.8)	0.30	0.32
6	rs411326	T	C	0.19	1201 (100)	0.26	0.25
6	rs2395174	G	T	0.31	1207 (100)	0.30	0.29
6	rs6903608	C	T	0.41	1210 (100)	0.31	0.30
6	rs2395185	T	G	0.22	1204 (100)	0.32	0.32
6	rs7775228	C	T	0.08	1210 (100)	0.12	0.12
6	rs9275572	A	G	0.54	1187 (100)	0.44	0.43

* Frequency of the minor allele (A1). NA = not applicable.

† Percent concordance among subjects for which genotypes were available from both the GWAS and technical replication Sequenom genotyping (Sequenom Inc., San Diego, CA).

‡ These SNPs were not successfully genotyped in the technical replication stage.

Supplementary Table 3. Characteristics of classical Hodgkin lymphoma case patients and control subjects

Characteristics	Genome-wide association study			Independent replication	
	Case patients, No. (%)	Generic control subjects, No. (%)	Study-specific control subjects, No. (%)	Case patients, No. (%)	Control subjects, No. (%)
Total	1200 (100)	6417 (100)	1395 (100)	563 (100)	613 (100)
Country					
Czech Republic	54 (4.50)	543 (8.46)	106 (7.60)	2 (0.36)	21 (3.43)
Denmark	189 (15.8)	81 (1.26)	525 (37.6)	0 (0)	0 (0)
France	13 (1.08)	157 (2.45)	21 (1.51)	16 (2.84)	49 (7.99)
Germany	86 (7.17)	52 (0.81)	187 (13.4)	6 (1.07)	86 (14.03)
Ireland	3 (0.25)	12 (0.19)	5 (0.36)	15 (2.66)	26 (4.24)
Italy	0 (0)	0 (0)	0 (0)	3 (0.53)	20 (3.26)
Netherlands	287 (23.9)	1804 (28.1)	0 (0)	0 (0)	0 (0)
Norway	0 (0)	389 (6.06)	0 (0)	0 (0)	0 (0)
Spain	25 (2.08)	99 (1.54)	58 (4.16)	22 (3.91)	194 (31.7)
Sweden	152 (12.7)	108 (1.68)	145 (10.4)	0 (0)	0 (0)
United Kingdom	391 (32.6)	3172 (49.4)	348 (25.0)	499 (88.6)	217 (35.4)
Gender					
Male	637 (53.1)	3489 (54.4)	795 (57.0)	321 (57.0)	334 (54.5)
Female	563 (46.9)	2928 (45.6)	600 (43.0)	242 (43.0)	279 (45.5)
Age, y					
15−24	295 (24.6)	6 (0.09)	135 (9.68)	153 (27.2)	57 (9.30)
25−34	332 (27.7)	28 (0.44)	181 (13.0)	147 (26.1)	119 (19.4)
35−44	210 (17.5)	217 (3.38)	189 (13.6)	94 (16.7)	123 (20.1)
45−54	137 (11.4)	3526 (55.0)	200 (14.3)	63 (11.2)	122 (19.9)
55−64	126 (10.5)	1159 (18.1)	306 (21.9)	68 (12.1)	101 (16.5)
65−74	86 (7.17)	1175 (18.3)	362 (26.0)	27 (4.80)	76 (12.4)
75−84	12 (1.00)	289 (4.50)	19 (1.36)	9 (1.60)	14 (2.28)
85−94	0 (0)	11 (0.17)	2 (0.14)	1 (0.18)	1 (0.16)
missing	2	6	1	1	0

Supplementary Table 4. Summary of classical Hodgkin lymphoma (cHL) results for single-nucleotide polymorphisms (SNPs) recently reported in genome-wide association studies (GWAS) of psoriasis and asthma *

Chromosome	SNP	Position	Gene	Total cHL		EBV+ cHL		EBV- cHL		Previous GWAS phenotype†
				OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	
5q31	rs20541	132023863	*IL13*	1.43 (1.28 to 1.60)	6.0×10^{-10}	1.07 (0.85 to 1.35)	.56	1.53 (1.32 to 1.76)	5.4×10^{-9}	Asthma, psoriasis, serum IgE
5q15	rs27524	96127700	*ERAP1*	1.21 (1.10 to 1.33)	1.5×10^{-4}	1.27 (1.05 to 1.53)	.01	1.24 (1.10 to 1.40)	5.5×10^{-4}	Psoriasis
1p13	rs2476601	114179091	*PTPN22*	1.24 (1.07 to 1.44)	3.7×10^{-3}	1.23 (0.92 to 1.64)	.16	1.26 (1.05 to 1.51)	.02	Psoriasis
2p16	rs702873	60935046	*REL*	1.12 (1.02 to 1.24)	.02	1.06 (0.88 to 1.27)	.54	1.15 (1.02 to 1.30)	.02	Psoriasis
19p13	rs12720356	10330975	*TYK2*	1.18 (1.01 to 1.39)	.04	1.30 (0.96 to 1.78)	.09	1.19 (0.97 to 1.45)	.10	Psoriasis
1p31	rs2201841	67466790	*IL23R*	1.10 (0.99 to 1.21)	.08	0.95 (0.78 to 1.16)	.62	1.15 (1.01 to 1.31)	.03	Psoriasis
1p36	rs4649203	24392507	*IL28RA*	0.94 (0.85 to 1.05)	.27	1.04 (0.85 to 1.28)	.71	0.88 (0.76 to 1.01)	.06	Psoriasis
5q33	rs2546890	158692478	*IL12B*	0.95 (0.87 to 1.05)	.31	0.92 (0.77 to 1.11)	.40	0.94 (0.84 to 1.07)	.35	Psoriasis
7q36	rs916514	154099909	*DPP6*	1.07 (0.92 to 1.25)	.40	1.01 (0.75 to 1.37)	.93	1.09 (0.89 to 1.32)	.40	Psoriasis
12q13	rs2066808	55024240	*STAT2*	1.06 (0.88 to 1.27)	.54	1.04 (0.73 to 1.49)	.82	1.04 (0.82 to 1.32)	.74	Psoriasis
11q22	rs1939015	102081585	*MMP27*	1.04 (0.91 to 1.18)	.58	1.00 (0.77 to 1.30)	.99	1.08 (0.92 to 1.27)	.36	Psoriasis
5q33	rs953861	158705160	*IL12B*	1.03 (0.91 to 1.17)	.63	1.05 (0.83 to 1.33)	.70	1.05 (0.90 to 1.22)	.55	Psoriasis
16p12	rs1859308	27305499	*IL21R*	1.03 (0.90 to 1.18)	.67	1.23 (0.95 to 1.58)	.11	0.97 (0.81 to 1.16)	.74	Serum IgE
15q22	rs744910	65233839	*SMAD3*	0.98 (0.89 to 1.08)	.70	1.07 (0.89 to 1.29)	.44	0.93 (0.83 to 1.05)	.27	Asthma
6p21	rs240993	111780407	*REV3L*	1.02 (0.92 to 1.14)	.70	1.00 (0.81 to 1.24)	1.0	1.04 (0.90 to 1.19)	.62	Psoriasis
9p24	rs1342326	6180076	*IL33*	0.98 (0.86 to 1.11)	.73	1.07 (0.84 to 1.36)	.61	0.93 (0.79 to 1.10)	.42	Asthma
17q12	rs2305480	35315722	*GSDML*	1.02 (0.93 to 1.12)	.73	0.92 (0.76 to 1.10)	.35	1.02 (0.91 to 1.16)	.70	Asthma
12q13	rs167769	55790042	*STAT6*	0.98 (0.89 to 1.08)	.74	1.05 (0.88 to 1.27)	.58	0.96 (0.84 to 1.08)	.47	Serum IgE
6p21	rs458017	111802784	*REV3L*	0.97 (0.80 to 1.19)	.80	1.08 (0.74 to 1.57)	.71	0.97 (0.75 to 1.25)	.80	Psoriasis
6p21	rs10484554	31382534	*HLA-C*	1.01 (0.87 to 1.16)	.92	1.35 (1.05 to 1.74)	.02	0.90 (0.75 to 1.10)	.30	Psoriasis
22q12	rs2284033	35863980	*IL2RB*	1.00 (0.91 to 1.10)	.99	1.03 (0.85 to 1.24)	.77	0.98 (0.87 to 1.11)	.78	Asthma

*Odds ratio (OR) and 95% confidence interval (CI) were derived using logistic regression assuming a log-additive genetic model of inheritance and adjusting for sex (male or female), country (eight indicator variables after excluding one country as the reference), and eight principal components analysis eigenvectors. All statistical tests were two-sided. EBV = Epstein-Barr virus. †Indicates the phenotype that the SNP was associated with in recent genome-wide association studies of psoriasis (30, 31) and asthma/immunoglobulin E (IgE) (28, 29).

Supplementary Table 5. Summary of genome-wide association study (GWAS) and technical replication stage results for the five independently associated major histocompatibility complex (MHC) region single-nucleotide polymorphisms (SNPs) and rs20541 (5q31, *IL13*), excluding the Netherlands subjects *

SNP (region)	Gene candidate	GWAS†		Technical replication†		P_{hom}‡
		OR (95% CI)	*P*	OR (95% CI)	*P*	
Total cHL						
rs2248462 (6p21)	*MICB*	0.64 (0.55 to 0.75)	1.8×10^{-8}	0.65 (0.55 to 0.77)	5.0×10^{-7}	.89
rs2395185 (6p21)	*HLA-DRA*	0.54 (0.47 to 0.61)	2.2×10^{-20}	0.57 (0.49 to 0.66)	1.3×10^{-14}	.59
EBV+ cHL						
rs2734986 (6p21) §	*HLA-A*	2.31 (1.78 to 2.99)	2.5×10^{-10}	NA	NA	NA
rs6904029 (6p21)	*HCG9*	0.51 (0.39 to 0.68)	2.6×10^{-6}	0.50 (0.38 to 0.66)	1.1×10^{-6}	.92
EBV- cHL						
rs6903608 (6p21)	*HLA-DRA*	2.13 (1.82 to 2.48)	2.7×10^{-21}	1.98 (1.67 to 2.35)	5.4×10^{-15}	.54
rs20541 (5q31)	*IL13*	1.48 (1.23 to 1.77)	2.4×10^{-5}	1.40 (1.15 to 1.70)	8.7×10^{-4}	.68

* Study-specific control subjects for the Northern Dutch Hodgkin Lymphoma Study were not available for the technical replication stage. To make results directly comparable, the technical replication analysis excluded the Netherlands case patients, and the GWAS analysis excluded the Netherlands case patients and control subjects. cHL = classical Hodgkin lymphoma, CI = confidence interval, EBV = Epstein-Barr virus, NA = not applicable, OR = odds ratio.

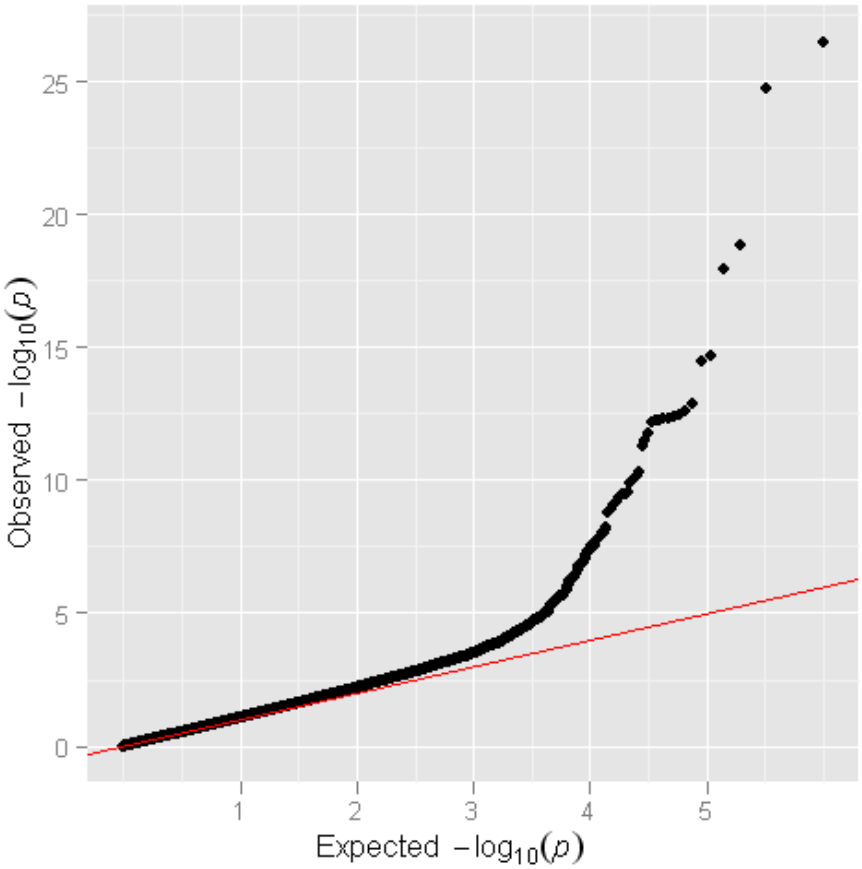
† The GWAS and technical replication analyses included 883 case patients (194 EBV-positive and 448 EBV-negative cHL). The GWAS and technical replication analysis included 4613 generic and 1395 study-specific control subjects, respectively. ORs and 95% CIs were derived using logistic regression assuming a log-additive genetic model of inheritance and adjusting for sex (male or female)

and country (seven indicator variables after excluding one country as the reference). The genome-wide association analysis additionally included eight principal components analysis eigenvectors. All statistical tests were two-sided.

‡ Two-sided *P* value for the χ^2 test of homogeneity (P_{hom}) comparing the risk estimates derived from the GWAS and technical replication analyses.

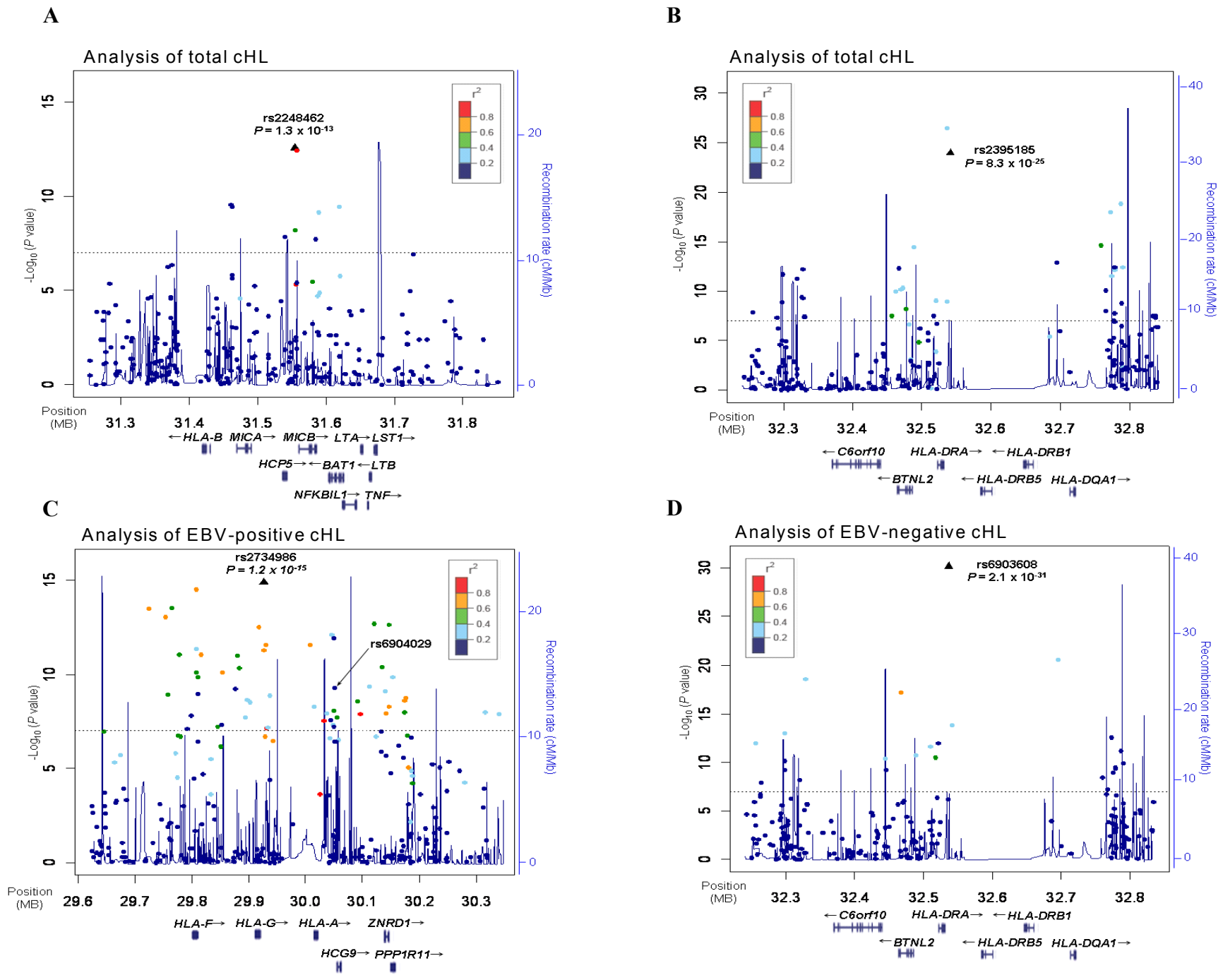
§ This SNP was not successfully genotyped in the technical replication stage.

Supplementary Figures

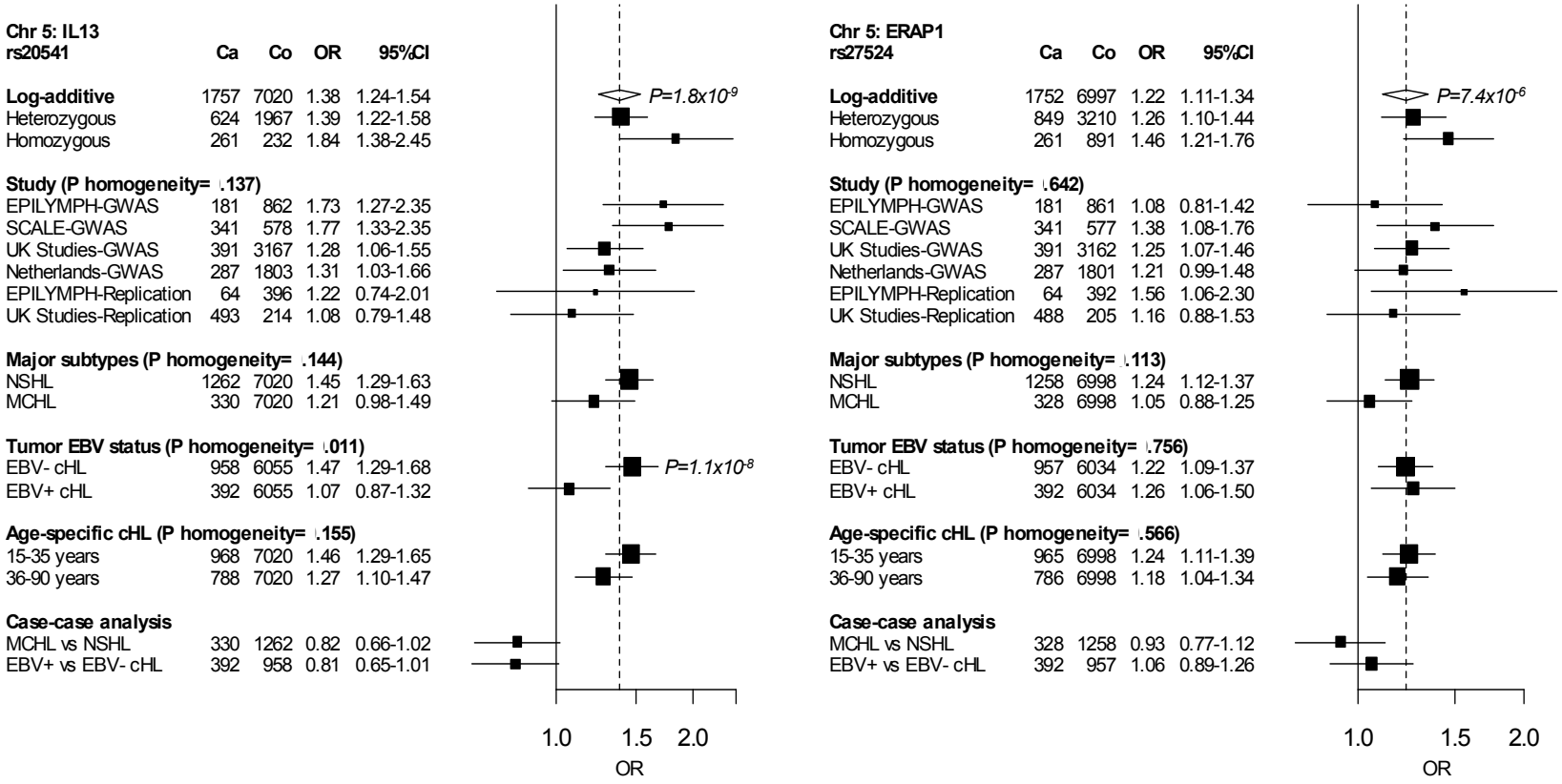


Supplementary Figure 1. Quantile-quantile plot of the expected vs observed -$\log_{10}$(*P* value) distribution in the genome-wide association analysis of total classical Hodgkin lymphoma. Association results were derived by multiple logistic regression assuming a log-additive genetic model and adjusting for sex (male or female), country (France, Germany, Spain, Czech Republic, Ireland, United

Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors. The **red line** represents the plot where the observed distribution of the $-\log_{10}$ (*P* value) is same as the expected distribution given the number of SNPs tested. The genomic inflation factor adjusted to the sample size of 1000 case patients and 1000 control subjects, using the method of de Bakker et al. (39), was equal to 1.04. All statistical tests were two-sided.


A
Analysis of total cHL
rs2248462
P = 1.3 x 10-13
r2
0.8
0.6
0.4
0.2
-Log10 (P value)
Recombination rate (cM/Mb)
Position (MB)
31.3
31.4
31.5
31.6
31.7
31.8
HLA-B
MICA
MICB
LTA
LST1
HCP5
BAT1
LTB
NFKBIL1
TNF
B
Analysis of total cHL
rs2395185
P = 8.3 x 10-25
32.3
32.4
32.5
32.6
32.7
32.8
C6orf10
HLA-DRA
HLA-DRB1
BTNL2
HLA-DRB5
HLA-DQA1
C
Analysis of EBV-positive cHL
rs2734986
P = 1.2 x 10-15
rs6904029
29.6
29.7
29.8
29.9
30.0
30.1
30.2
30.3
HLA-F
HLA-G
HLA-A
ZNRD1
HCG9
PPP1R11
D
Analysis of EBV-negative cHL
rs6903608
P = 2.1 x 10-31
32.3
32.4
32.5
32.6
32.7
32.8
C6orf10
HLA-DRA
HLA-DRB1
BTNL2
HLA-DRB5
HLA-DQA1

Supplementary Figure 2. Plots of genome-wide associations between classical Hodgkin lymphoma (cHL) and single-nucleotide polymorphisms (SNPs) located within a 300 kb region flanking the index SNP (indicated by a black triangle) of the major histocompatibility complex region. The relationship between total cHL and two novel loci, **A**) rs2248462 (class I region, *MICB*) and **B**) rs2395185 (class II region, *HLA-DRA*), was investigated. **C)** The relationship between two loci in the *HLA-A* class I region indexed by rs2734986 and rs6904029 and Epstein Bar virus (EBV)-positive cHL was determined. **D)** The relationship between class II region locus, rs6903608 (*HLA-DRA*) and EBV-negative cHL was investigated. Multiple logistic regression was performed assuming a log-additive genetic model and adjusting for sex (male or female), country (France, Germany, Spain, Czech Republic, Ireland, United Kingdom, Denmark, Sweden, and the Netherlands), and eight principal components analysis eigenvectors. The $-\log_{10}$ (*P* value) for each SNP are plotted against their chromosomal position. All statistical tests were two-sided. The colors of the dots indicate the degree of linkage disequilibrium (based on r^2) in relation to the 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.



Supplementary Figure 3. Stratified and subgroup analyses of two single-nucleotide polymorphisms (SNPs) located outside the major histocompatibility complex region. The relationship between **A)** rs20541 (*IL13*) and **B)** rs27524 (*ERAP1*), with classical Hodgkin lymphoma (cHL) was investigated. 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. The results are on the basis of a combined analysis of genome-wide association and independent replication phase results using inverse variance weighting meta-analysis. In the analysis stratified by study, the GWAS included the EPILYMPH study (EPILYMH-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). Results by study for the independent replication included the EPILYMPH study (EPILYMPH-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). $P_{\text{homogeneity}}$ was on the basis of the Cochran Q test statistic and was used to evaluate between-study heterogeneity in results. Associations between the SNPs and cHL subgroups (including histologic subtype, EBV status, and age) were performed, and $P_{\text{homogeneity}}$ for the χ^2test of homogeneity indicates the differences in the OR between subgroup analyses. All statistical tests were two-sided. Ca = cases, Chr = chromosome, Co = control subjects, EBV = Epstein-Barr virus, MCHL = mixed cellularity Hodgkin lymphoma, NSHL = nodular sclerosis Hodgkin lymphoma.