A genome-wide association study of Hodgkin Lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21, and 10p14 (GATA3)

Published in final edited form as:
Nat Genet. 2010 December ; 42(12): 1126–1130. doi:10.1038/ng.696.

A genome-wide association study of Hodgkin Lymphoma identifies new susceptibility loci at 2p16.1 (*REL*), 8q24.21, and 10p14 (*GATA3*)

Victor Enciso-Mora[#1], **Peter Broderick**[#1], **Yussanne Ma**[#1], **Ruth F Jarrett**[2], **Henrik Hjalgrim**[3], **Kari Hemminki**[4], **Anke van den Berg**[5], **Bianca Olver**[1], **Amy Lloyd**[1], **Sara E Dobbins**[1], **Tracy Lightfoot**[6], **Flora E van Leeuwen**[7], **Asta Försti**[4], **Arjan Diepstra**[5], **Annegien Broeks**[8], **Jayaram Vijayakrishnan**[1], **Lesley Shield**[2], **Annette Lake**[2], **Dorothy Montgomery**[2], **Eve Roman**[6], **Andreas Engert**[9], **Elke Pogge von Strandmann**[9], **Katrin S. Reiners**[9], **Ilja M Nolte**[27], **Karin E Smedby**[10], **Hans-Olov Adami**[11,12], **Nicola S Russell**[13], **Bengt Glimelius**[14,15], **Stephen Hamilton-Dutoit**[16], **Marieke de Bruin**[7], **Lars P Ryder**[17], **Daniel Molin**[18], **Karina Meden Sorensen**[19], **Ellen T Chang**[20,21], **Malcolm Taylor**[22], **Rosie Cooke**[23], **Robert Hofstra**[24], **Helga Westers**[24], **Tom van Wezel**[25], **Ronald van Eijk**[25], **Alan Ashworth**[26], **Klaus Rostgaard**[3], **Mads Melbye**[3], **Anthony J Swerdlow**[20], and **Richard S Houlston**[1,¥]

[1]Section of Cancer Genetics, Institute of Cancer Research, Sutton, SM2 5NG, UK [2]Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G61 1QH, UK [3]Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark [4]Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany and Center for Primary Health Care Research, Clinical Research Center, Lund University, Malmö, Sweden [5]Department of Pathology & Medical Biology, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands [6]Epidemiology & Genetics Unit, Department of Health Sciences, University of York, York, YO10 5DD, UK [7]Department of Epidemiology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands [8]Department of Experimental Therapy, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, the

¥Corresponding author: Richard.houlston@icr.ac.uk, Tel: +44 (0) 208 722 4175, Fax: +44 (0) 208 722 4365 .

AUTHOR CONTRIBUTIONS

RSH designed the study and obtained financial support. RSH drafted the manuscript with contributions from PB, VE, YM and SED. YM and VE performed statistical and bioinformatic analyses; PB sample coordination and laboratory analyses; BO, AL and JV performed genotyping. AJS, AA and RC provided samples and data from a study conducted at the ICR. ER initiated ELCCS; TL and ER managed and prepared Epidemiology & Genetics Lymphoma Case-Control Study samples. RFJ designed and conducted studies contributing to the UK replication series and RFJ, LS, AL and DM prepared samples and collated data. FEvL designed the Dutch NKI study and obtained financial support; NSR and MdB, were involved in identification and inclusion of Dutch cases, study design, review board approval and clinical implementation; AB coordinated collection and preparation of NKI samples. AF, KH, AE, EPvS and KR provided samples and data from German cases and controls. AD, IMN and AvdB, collection of samples and data from cHL cases ascertained through Groningen. RH, HW, TvW and RvE, ascertainment and collection of control samples from the Netherlands. HH, MM, KR, LPR, KES, HOA, BG, DM, SHD, KMS, and ETC provided samples and data from the SCALE study in Denmark and Sweden. SHD analyzed samples and provided data from Danish cHL cases. All authors contributed to the final paper. RFJ and HJ contributed equally to the paper and should be considered to have equal positional status in the author list.

Note: Supplementary information is available on the Nature Genetics website

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests

Europe PMC Funders Author Manuscripts

Netherlands [9]University Hospital of Cologne, Department of Internal Medicine, Cologne, Germany [10]Unit of Clinical Epidemiology, Department of Medicine, Karolinska University Hospital, SE-171 76, Stockholm, Sweden [11]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-171 77, Stockholm, Sweden [12]Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA [13]Department of Radiotherapy, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands [14]Department of Pathology and Oncology, Karolinska Institutet, SE-171 77, Stockholm, Sweden [15]Department of Oncology, Radiology and Clinical Immunology, Uppsala University, SE-751 85, Uppsala, Sweden [16]Institute of Pathology, Aarhus University Hospital, DK-8000, Aarhus, Denmark [17]Department of Clinical Immunology, University Hospital of Copenhagen, Rigshospitalet, DK-2100, Copenhagen, Denmark [18]Department of Oncology, Radiology and Clinical Immunology, Uppsala University, Uppsala, Sweden [19]Department of Clinical Biochemistry, Statens Serum Institut, Copenhagen, Denmark [20]Cancer Prevention Institute of California, Fremont, CA 94538 [21]Division of Epidemiology, Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA 94305, USA [22]Cancer Immunogenetics Group, School of Cancer & Enabling Sciences, University of Manchester, Research Floor, St Mary's Hospital, Manchester, M13 9WL, UK [23]Section of Epidemiology, Institute of Cancer Research, Sutton, SM2 5NG, UK [24]Department of Genetics University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands [25]Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands [26]The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, Fulham Road, London SW3 6JB, UK [27]Unit of Genetic Epidemiology and Bioinformatics, Department of Epidemiology, University Medical Centre Groningen, University of Groningen, the Netherlands

[#] These authors contributed equally to this work.

Abstract

To identify predisposition loci for classical Hodgkin Lymphoma (cHL) we conducted a genome-wide association study of 589 cHL cases and 5,199 controls with validation in 4 independent samples totaling 2,057 cases and 3,416 controls. We identified three new susceptibility loci at 2p16.1 (rs1432295, *REL*; odds ratio [OR]=1.22, $P_{combined}=1.91\times10^{-8}$), 8q24.21 (rs2019960, *PVT1*; OR=1.33, $P_{combined}=1.26\times10^{-13}$) and 10p14 (rs501764, *GATA3*; OR=1.25, $P_{combined}=7.05\times10^{-8}$). Furthermore, we confirmed the role of the MHC in disease etiology by revealing a strong HLA association (rs6903608; OR=1.70, $P_{combined}=2.84\times10^{-50}$). These data provide new insight into the pathogenesis of cHL.

Classical Hodgkin Lymphoma (cHL) is a lymph node cancer of germinal center B-cell origin, characterized by malignant Hodgkin and Reed-Sternberg (HRS) cells mixed with a dominant background population of reactive lymphocytes and other inflammatory cells[1]. cHL is one of the most common tumors in young adults in economically developed countries, with ~1,500 cases being diagnosed each year in the UK, and the disease accounts for ~1 in 3 of all lymphomas[2,3]. While Epstein-Barr virus (EBV) infection may be causally related to a proportion of cases, the etiology of EBV-negative cHL remains largely unknown[4].

Evidence for inherited genetic predisposition to cHL is provided by the 3 to 9-fold increased risk of cHL in first-degree relatives of cHL patients[5]. In the light of a possible viral basis to cHL it is interesting that cHL was the first disease to be associated with the HLA region[6]. Subsequent studies have reported associations between various HLA class I and class II alleles and risk of cHL[7,8]; specifically an association between the HLA-A*01 and A*02 for EBV-positive cHL[9,10]. Genetic variation in HLA is, however, insufficient to account for the observed familial risk of cHL[11]. To date no non-HLA genetic risk factors have been identified and convincingly replicated. Genome-wide linkage studies of cHL families have failed to demonstrate an additional major gene locus for cHL[12]. This coupled with the very high concordance of Hodgkin Lymphoma in monozygotic compared with dizygotic twins[13] is consistent with a genetic model of inherited susceptibility based on the co-inheritance of multiple low-risk variants.

Predicated on this hypothesis we conducted a genome-wide association study (GWAS) of 622 UK cHL cases using Illumina 660w Quad BeadChips. Genotype frequencies were compared with publicly accessible genotype data generated by the UK Wellcome Trust Case-Control Consortium 2 (WTCCC2) study of 2,930 individuals from the 1958 British Birth Cohort (58C)[14] and 2,737 individuals from the UK Blood Service collections (UKBS), that had been genotyped using Illumina Human 1.2M-Duo Custom_v1 Array BeadChips (Online Methods). There was no evidence of systematic bias between these two series (Online Methods; Supplementary Figure 1), which were combined to provide genotype data on 5,667 controls. Data on 521,834 autosomal SNPs common to cases and controls were included in this analysis. After stringent quality control filtering (Online Methods; Supplementary Table 1), we analyzed 504,374 SNPs in 589 cHL cases and 5,199 controls. Principal component analysis (PCA) demonstrated that these cases and controls were genetically well matched (Supplementary Figure 2). We therefore assessed the association between each SNP and cHL risk using the Cochran-Armitage trend test without PCA adjustment. The quantile-quantile (Q-Q) plots of the negative logarithm of genome-wide *P*-values showed a strong deviation from the null distribution (Supplementary Figure 1), which could be ascribed to the strong association observed within the MHC region. After excluding 1,700 SNPs mapping to the major histocompatibility (MHC) region (6p21: 28-33Mb) there was only minimal inflation of test statistics, except at the upper tail of the distribution ($P<10^{-4}$), thereby rendering cryptic population substructure or differential genotype calling between cases and controls unlikely (genomic control inflation factor[15], $\lambda_{gc=}1.04$; Supplementary Figure 1). Using principal components analyses as implemented in Eigenstrat[16], correction for possible population substructure had no influence on findings for subsequently validated loci (Table 1). Furthermore, evidence for loci influencing cHL risk was provided by independent comparison with both 58C and UKBS control series (Supplementary Table 2).

This GWAS revealed multiple associations at chromosome 6, as well as suggestive associations on chromosomes 2, 5, 7, 8, 9, 10, 11 and 19 (Figure 1). To validate these associations we genotyped the HLA class II SNP rs6903608 and 10 SNPs from other regions showing an association, in the UK replication series (524 cases, 1,533 controls) (Online Methods, Supplementary Table 1). In the combined analysis, associations for 6 of

the SNPs were significant at $P_{combined}<1.0\times10^{-4}$ (Supplementary Table 3). These 6 SNPs were successfully genotyped in 3 independent case-control replication series (Online Methods, Supplementary Table 1) - SCALE (482 cases, 590 controls), Germany (498 cases, 655 controls) and Netherlands (553 cases, 638 controls). Combined analysis of all case-control series revealed genome-wide associations (*i.e.*, $P<5.0\times10^{-7}$)[17] at 2p16.1, 6p21, 8q24.21 and 10p14 (Table 1; Supplementary Table 3).

In our GWAS, 42 SNPs mapping to the 4.8Mb interval at 6p21, bordered by the *TRIM27* and *MLN* genes (rs209130, 28,975,779bps and rs1547668, 33,883,424bps respectively) defining the classical MHC region, showed evidence of an association with cHL risk at $P<5.0\times10^{-7}$ (Supplementary Figure 3). The most significant associations were with SNPs mapping to HLA class II; the strongest signal was attained at rs6903608 centromeric to *HLA-DRA* ($P=8.12\times10^{-21}$, 32,536,263bps; Supplementary Figure 3). The association between rs6903608 was consistently seen in each of the replication series, $P_{combined}=2.84\times10^{-50}$ (Table 1, Supplementary Table 3).

The association with rs1432295 ($P_{combined}=1.91\times10^{-8}$, OR=1.22) on 2p16.1 (60,920,170bps) straddles a recombination hotspot between 2 regions of high linkage disequilibrium (LD) (Figure 2; Supplementary Figure 4). The 137Kb region defined by these two LD blocks encompasses the putative transcript *FLJ16341* and *REL* (avian reticuloendotheliosis viral oncogene homolog). *REL* encodes c-Rel, a member of the Rel/NFκB family of transcription factors. Constitutive activity of NFκB transcription factors is a hallmark of cHL[1] and inactivating somatic mutations of the NFκB signaling inhibitors play a major role in cHL pathogenesis[18-20]. Furthermore, studies have shown genomic amplifications of *REL* associated with increased c-Rel expression in cHL [21-23].

We identified 2 SNPs on 8q24.21 associated with cHL risk, rs2019960 ($P_{combined}=1.26\times10^{-13}$, OR=1.33) and rs2608053 ($P_{combined}=1.16\times10^{-7}$, OR=1.20). rs2608053 mapping at 129,145,014bps localizes to a 56Kb region of LD that encompasses intron 6 of *PVT1* (Figure 2, Supplementary Figure 4). rs2019960 mapping at 129,261,453bps localizes to a 82Kb region of LD telomeric to *PVT1* (Figure 2, Supplementary Figure 4). The effects of rs2019960 and rs2608053 on cHL risk are maintained when adjusted for each other by logistic regression (OR=1.33, 95% CI: 1.23-1.44, $P=1.97\times10^{-13}$; and OR=1.20, 95% CI:1.12-1.28, $P=1.37\times10^{-7}$, respectively). Furthermore, correlation between rs2019960 and rs2608053 is poor ($r^2=0.0$, $D'=0.01$ in HapMap CEU samples, $r^2=0.0$, $D'=0.03$ in our control data) and comparison of haplotype frequencies provided evidence of two haplotypes differing in frequency between cases and controls (Supplementary Table 4). Because rs2019960 or rs2608053 alone cannot fully account for the association between 8q24.21 and cHL, it is possible that a unique variant in LD with and capturing the effects of both SNPs may exist. We did not, however, identify a more significant association in LD with both SNPs through imputation, making it plausible that two independent signals exist at 8q24.21.

PVT1 is frequently involved in translocations occurring in variant Burkitt's lymphoma and murine plasmacytomas[24]. The *PVT1* locus encodes several microRNAs thought to be as important as *MYC* in T-lymphomagenesis and T-cell activation[25]. Co-activation of c-Myc

and *PVT1* has been shown in a variety of human and animal tumors[26-28]. The 128-130Mb genomic interval at 8q24.21 harbors multiple independent loci with different tumor specificities, including chronic lymphocytic leukemia (rs2456449; 128,262,163bps)[29], prostate (rs16901979; 128,194,098bps)[30], breast (rs13281615; 128,424,800bps)[31], colorectal and prostate (rs6983267; 128,482,487bps)[32,33], prostate (rs1447295; 128,554,220bps)[34] and bladder (rs9642880; 128,787,250bps)[35] cancer. The LD blocks defining these loci are distinct from the 8q24.21 cHL association signal (r^2<0.03; Supplementary Table 5). The colorectal cancer SNP rs6983267 shows differential binding of TCF4 to an enhancer element that physically interacts with the *MYC* promoter [36,37]. A similar allele-specific cis-effect either on *MYC* or through *PVT1* impacting on *MYC* expression provides an attractive mechanistic basis for the 8q24.21 association with cHL risk. If the 8q24.21 locus influences risk through differential *MYC* expression, the association is intriguing since c-Myc and Rel/NFκB are the two master transcriptional systems activated in the latency III program of EBV-immortalized B-cells, which are responsible for the phenotype, growth pattern, and biological properties of cells driven into proliferation by EBV[38].

The two SNPs showing an association with cHL mapping to 10p14, rs501764 ($P_{combined}=7.05\times10^{-8}$, OR=1.25) and rs485411 ($P_{combined}=1.29\times10^{-7}$, OR=1.22) are in strong LD ($r^2=0.71$, $D'=0.95$ in HapMap CEU samples, $r^2=0.69$, $D'=1.00$ in our control data) and map to a 40Kb region of LD encompassing the transcription factor and putative tumor suppressor gene, *GATA3* (GATA binding protein 3 isoform 2) (Figure 2, Supplementary Figure 4). The expression of *GATA3* is important in hematopoeitic and lymphoid-cell development, acting as a master transcription factor for differentiation of T_h2 cells[39]. A high proportion of the reactive infiltrate in cHL tumors is composed of T_h2-like cells with T_{reg} phenotype which can influence EBV-positive cHL cell growth, depending on EBV antigenic presentation by MHC molecules[40]. Notably, a key characteristic of HRS cells is the production of cytokines and chemokines driven by *GATA3* expression and other T-cell transcription factors [41]. Evidence for a biological relationship between the 2p16.1, 8q24.21 and 10p14 loci is that members of the Rel-family have differential effects on the *MYC* promoter[42] and *GATA3* is a target for c-Myc[43].

Elucidation of the basis of each of the associations at 2p16.1, 8q24.21 and 10p14 will require fine-mapping and functional analyses. To examine if any directly typed or imputed SNPs annotate a putative transcription factor (TF) binding/enhancer element, we conducted a bioinformatic search of each of the regions of association using Transfac Matrix Database, PReMod and EEL software. At 10p14 an imputed SNP rs369421 provides the best evidence for the association signal ($P=6.20\times10^{-7}$) mapping within module 011553 (Supplementary Table 6, Supplementary Figure 4). Intriguingly, this module includes binding sites for ARID5B and E2F TFs. ARID5B has been previously implicated in development of acute leukemia[44], and loss of PU.1, an E2F TF, has been associated with defective immunoglobulin expression in HRS cells[45].

A hallmark of cHL epidemiology is the bimodal age specific incidence and it has been argued that the disease in young adults and older adults are etiologically different; in particular there is a low prevalence of EBV in younger cHL patients[46]. We assessed the

relationship between cHL and EBV-status, age and sex at the 6p21, 2p16.1, 8q24.21 and 10p14 loci (defined by rs6903608, rs1432295, rs2019960, rs2608053, and rs501764 genotypes) by case-only analysis using data from SCALE, UK and Netherlands replication series (1,100 cases; Supplementary Table 7). Associations at all loci were not influenced by sex after adjustment for age and EBV-status. The rs501764 association with cHL was not related to age or EBV-status (Supplementary Table 7). The HLA class II association at 6p21 was primarily driven by EBV-negative cHL after adjustment for age and sex ($P_{adjusted}=1.63\times10^{-11}$). Similarly, rs1432295 (2p16.1) risk alleles were significantly enriched in EBV-negative cHL ($P_{adjusted}=0.01$). At 8q24.21, while rs2608053 was associated with EBV-negative cHL ($P_{adjusted}=0.01$), rs2019960 showed a relationship with early-onset cHL, independent of EBV-status or sex ($P_{adjusted}=0.002$) (Supplementary Table 7). These phenotypic differences provide further support for two independent cHL risk loci at 8q24.21.

To explore whether any of the associations at 2p16.1, 8q24.21 and 10p14 reflect cis-acting regulatory effects on a nearby gene we searched for genotype-expression correlations in 90 EBV-transformed lymphoblastoid cell lines using previously described data[47,48]. We did not find any significant relationship between SNP genotype and gene expression, after adjustment for multiple testing (Supplementary Figure 5). This does not preclude the possibility that the causal variants at these disease loci have subtle effects on expression as the dynamic range of transcripts, such as *MYC*, is small. Furthermore, it is likely that only a cumulative long-term imbalance in expression in target genes will influence cHL development and expression differences may only be relevant to a specific subpopulation of B-cells, which may not be well modelled by EBV-transformed lymphocytes.

While the HLA association with cHL is a very strong genetic effect, the identification of risk variants at 2p16.1, 8q24.21 and 10q14 implicates important roles for networks involving *MYC*, *GATA3* and the NFκB pathway in cHL disease etiology. In the combined dataset there was some evidence for interactions between HLA (rs6903608) and 2p16.1 (rs1432295; $P=0.05$) and between 8q24.21 (rs2608053) and 10p14 (rs501764 and rs485411; $P=0.01$), albeit non-significant after correction for multiple testing (Supplementary Table 8). Further studies are needed to investigate possible interactions between these susceptibility loci and their interplay with EBV infection. Finally, the modest size of our study makes it likely that further risk variants for cHL can be identified through additional studies.

ONLINE METHODS

Patients and samples

Genome-wide association study—We analyzed constitutional DNA of 622 cHL patients (International Classification of Diseases [ICD] 10 codes C81.0-3) ascertained through the Royal Marsden Hospitals NHS Trust Family History study, during 2004-2008 (n=104, 63 male; mean age of diagnosis [AOD]=38, SD=16) and an ongoing national study of cHL in females (n=518, mean AOD=23, SD=6) conducted by the Institute of Cancer Research (ICR). 146 cases had been diagnosed with breast cancer subsequent to cHL diagnosis. All cases British residents and self-reported to be of European Ancestry.

For controls we used publicly accessible data generated by the UK Wellcome Trust Case-Control Consortium 2 (WTCCC2) study on 5,667 individuals from two sources: 2,930 individuals from the British 1958 Birth Cohort (58C; also known as the National Child Development Study) which includes all births in England, Wales and Scotland[14]; and 2,737 UK Blood Services Controls (UKBS) aged 18-69, sex- and geographically matched to reproduce the distribution of samples within 58C.

Replication series—The UK-replication series comprised 524 cHL cases (ICD10 C81.0-3; 290 male, mean AOD=38, SD=16) ascertained from the Scotland and Newcastle Epidemiological Study of Hodgkin Disease (SNEHD), the Young adult Hodgkin Case-Control Study (YHCCS) and the Epidemiology & Genetics Lymphoma Case-Control Study (ELCCS; www.elccs.info). Full details of SNEHD, YHCCS and ELCCS studies provided previously[49-51]. Briefly, SNEHD involved ascertainment of incident cases from Scotland and Northern England during 1993-1997. YHCCS was based on newly diagnosed patients aged 16-24 from Northern England during 1991-1995. ELCCS comprised patients residing in the north of England aged 16-69, with newly diagnosed, non-HIV-related HL, during 1998-2003. UK population controls obtained from SNEHD and YHCCS (n=495, 268 male, mean age 41, SD=17) and ongoing epidemiological studies of cancer conducted at the ICR (n=1,038, 524 male, mean age 60, SD=9)[52].

The Scandinavian Lymphoma Etiology (SCALE) study has been described previously[53,54]. Briefly, SCALE is a population-based case-control study of HL and non-Hodgkin lymphomas conducted in Denmark and Sweden during 1999-2002. The study population encompassed Danish and Swedish speaking residents aged 18-74 with no history of HIV infection, solid organ transplantation or previous hematopoietic malignancy in Denmark from June 1, 2000 to August 30, 2002, and in Sweden from October 1, 1999, to April 15, 2002. Participants recruited in a Danish regional pilot phase starting November 1, 1999, were also included, as were prevalent cases of HL diagnosed since January 1, 1999 in both countries. In total, 586 patients diagnosed with cHL according to the WHO classification in the study period and 3,187 controls representing 91% and 71% of eligible cases and controls, respectively, participated in the study, which included telephone interview and blood sampling. For the present investigation, DNA from 482 cases (82% of all SCALE cHL cases, 282 male, mean AOD=40, SD=16) and from 255 Danish controls was extracted from dried filter paper blood spots with Extract-N-AmpT as per manufacturer's instructions (Sigma-Aldrich, St. Louis, MO, USA) and subjected to whole genome amplification with AmpliQ Genomic Amplifier Kit (Ampliqon, Denmark)[55]. In addition, germline DNA extracted from buffy coat for 335 Danish SCALE controls (randomly selected from 590 controls) was also included. (Mean age for combined SCALE controls 59, SD=13).

The Netherlands replication series comprised: (i) 281 cHL patients (149 male, mean AOD=36, SD=15) collected from the north of the Netherlands diagnosed during 1997-2000 as part of an ascertainment by the University Medical Centre Groningen; (ii) 272 cHL cases, 97 diagnosed with breast cancer subsequent to cHL (mean AOD=24, SD=6). These patients were selected in the framework of an ongoing case-control study of risk factors for breast cancer after HL conducted by the Netherlands Cancer Institute, Amsterdam, within a larger cohort study of women treated for cHL before age 60, during 1965-1995 and who survived

at least 5 years. Patient selection, methods of data/blood collection and DNA isolation described previously[56-58]. Samples from healthy blood donors, aged 19-69, ascertained through medical centers in Groningen (mean age=52, SD=11) and Leiden (mean age=47, SD=12), served as controls.

The German replication series comprised 498 cHL patients ascertained by the German Hodgkin Study Group during 1998-2007 (292 male, mean AOD=34, SD=12). Controls were 655 healthy blood donors from Mannheim, located 200km from Cologne (381 male, mean age=36, SD=13).

EBV status of tumors—EBV status of cHL tumors was determined by immunohistochemical staining for EBV latent membrane antigen (LMP)-1 and/or EBV EBER *in situ* hybridization using sections of paraffin-embedded material[53,59].

Ethics—Collection of blood samples and clinico-pathological information from subjects was undertaken with informed consent and relevant ethical review board approval in accordance with the tenets of the Declaration of Helsinki.

Genotyping—DNA extracted from samples using conventional methodologies and quantified using PicoGreen (Invitrogen, Carlsbad, USA).

Genotyping of cHL cases in the GWAS conducted using Illumina Infinium HD Human660-Quad BeadChips according to manufacturer's protocols (Illumina, San Diego, USA). DNA samples with GenCall scores<0.25 at any locus considered "no-calls". A SNP was considered failed if fewer <95% of DNA samples generated a genotype at the locus. Cluster plots manually inspected for all SNPs considered for replication.

We used data on controls from the 1958 Birth Cohort (58C) and National Blood Service (UKBS) which had been generated by the WTCCC. Genotyping of both sets of controls was conducted using Illumina Human 1.2M-Duo Custom_v1 Array chips. SNP calling performed using Illuminus Software. Full details of genotyping, SNP calling and QC reported previously (www.wtccc.org.uk). Concordant with previous findings[17] comparison of the two control series showed little evidence for systematic bias (inflation factor λ=1.022; Supplementary Figure 1).

Validation and replication of associations were performed using competitive allele-specific PCR KASPar chemistry (KBiosciences Ltd, Hertfordshire, UK). Primers and probes used available on request. Samples having SNP call rates <90% excluded from analysis. To ensure quality of genotyping in all assays, at least two negative controls and 1-2% duplicates (showing a concordance >99.99%) were genotyped. We performed cross-platform validation and sequenced a random series of 96 samples to exclude technical artifact confirm genotyping accuracy (concordance>99.9%).

Statistical and bioinformatic analysis—We applied pre-determined quality-control metrics to the GWAS data. We restricted analyses to samples for whom >95% of SNPs were successfully genotyped, eliminating 12 cases. We computed identity-by-state (IBS) probabilities for all pairs to search for duplicates and closely related individuals amongst

cases and controls (defined as IBS ≥0.80, thereby excluding first-degree relatives). For all identical pairs the sample with highest call rate was retained, eliminating 2 cases. To identify individuals with possible non-Western European ancestry, we merged our case and control data with HapMapII samples (60 western European [CEU], 60 Nigerian [YRI], 90 Japanese [JPT] and 90 Han Chinese [CHB]). For each pair of individuals we calculated genome-wide IBS distances on markers shared between HapMap and our SNP panel, used as dissimilarity measures upon which to perform principal component analysis. The first two principal components for each individual were plotted; any individual not present in the main CEU cluster (*i.e.* 5% furthest from cluster centroids) was excluded. We removed 30 cases with non-CEU ancestry (some of which had poor call rates) and 1 WTCCC2 control which was a duplicate case. We excluded SNPs with minor allele frequency <1%, and call rate <95% (cases or controls) and those showing departure from Hardy-Weinberg equilibrium ($P<10^{-5}$) in controls. For replication and validation analysis call rates were >95% per 384-well plate for each SNP; cluster plots visually examined by two researchers.

Main analyses were undertaken using R(v2.6), Stata10 (State College, Texas, US) and PLINK(v1.06). Association between each SNP and cHL risk was assessed by the Cochran-Armitage trend test. The adequacy of case-control matching and possibility of differential genotyping of cases and controls were formally evaluated using quantile-quantile (Q-Q) plots of test statistics. The inflation factor λ was based on the 90% least significant SNPs[15]. We adjusted for possible population substructure using Eigenstrat software[16]. Odds ratios (ORs) and associated 95% confidence intervals (CIs) were calculated by unconditional logistic regression. Meta-analysis was conducted using standard methods[60]. Cochran's Q statistic and I^2 statistic were calculated to test for heterogeneity and quantify total variation due to heterogeneity; large heterogeneity typically defined as $I^2 \geq 75\%$[61]. We conducted a pooled analysis incorporating Eigenstrat-adjusted *P*-values from the GWAS using the weighted Z-method implemented in the program METAL. We examined each SNP for dose response by comparing 1-d.f. and 2-d.f. logistic regression models, adjusting for stage using a likelihood ratio test, and examined the combined effects of multiple SNPs by evaluating the effect of adding an interaction term on the model by using a likelihood ratio test and adjusting for stage. Associations by sex, age and EBV-status were examined by logistic regression in case-only analyses.

Prediction of the untyped SNPs was carried out using IMPUTEv2, based on HapMapIII/Release27 (Feb2009, NCBI B36, dbSNP26) and 1000genomes. Imputed data were analysed using SNPTESTv2 to account for uncertainties in SNP prediction. LD-metrics between HapMap SNPs were based on HapMapIII/Release27, viewed using Haploview(v4.2) and plotted using SNAP. LD-blocks defined on the basis of HapMap recombination rate (cM/Mb) as defined using Oxford recombination hotspots[62] and on the basis of distribution of confidence intervals defined by Gabriel *et al*. [63]

To annotate potential regulatory sequences within disease loci we implemented *in silico* searches using Transfac Database(v7.29)[64], PReMod10[65] and EEL[66].

Relationship between SNP genotypes and expression levels—To examine for a relationship between SNP genotype and expression levels of *GATA3*, *REL*, and *MYC* in

lymphocytes we made use of publicly available expression data generated from analysis of 90 Caucasian derived Epstein-Barr virus–transformed lymphoblastoid cell lines using Sentrix Human-6 Expression BeadChips (Illumina, San Diego, USA)[47,48]. Online recovery of data performed using WGAViewer(v1.25). Differences in distribution of mRNA expression levels between SNP genotypes were compared using a Wilcoxon-type trend test[67].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

Leukemia Lymphoma Research (UK) and Cancer Research UK (C1298/A8362 supported by the Bobby Moore Fund) provided principal funding for the study. We acknowledge NHS funding to the NIHR Biomedical Research Centre. This study made use of control genotyping data generated by the Wellcome Trust Case-Control Consortium. We acknowledge use of genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113 and 085475. At the Institute of Cancer Research sample and data acquisition was supported by Breakthrough Breast Cancer and the European Union and we acknowledge NHS funding to the NIHR Biomedical Research Centre. We are grateful to the patients and their clinicians who participated in this collection (see Supplementary Note). Work at the LRF Virus Centre was funded by Leukaemia and Lymphoma Research. Sample and data acquisition for the UK replication series was also supported by the Kay Kendall Leukaemia Fund. ELCCS was funded by Leukaemia & Lymphoma Research. Grant support to the German Study Group was through Deutsche Krebshilfe and the EU, HEALTH-F4-2007-200767. The SCALE study is supported by the Lundbeck Foundation R19 A2364; Danish Cancer Research Foundation grant 41-08 and Danish Cancer Society grant DP 08155. At the Department of Pathology & Medical Biology, University of Groningen, sample and data acquisition was supported by two grants from the Dutch Cancer Society (RUG 200-2315 and RUG 2009-4313). The Dutch NKI study was supported by the Dutch Cancer Society (Grants No. NKI 98-1833, NKI 04-3068, NKI 08-3994) and the EU6th project GRR (Project no 012926). We thank Ausra Kesminiene for coordinating the EU-GRR project, Michael Schaapveld and Anja Eggermond for data management, Linde Braaf and Izabela Mikolajewska for lab assistance. We are indebted to the patients and physicians who participated in this collection (see Supplementary Note).

URLs

The R suite can be found at http://www.r-project.org/

Detailed information on the tag SNP panel can be found at http://www.illumina.com/

dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/

HapMap: http://www.hapmap.org/

1000Genomes: http://www.1000genomes.org/

1958 Birth Cohort: http://www.cls.ioe.ac.uk/studies.asp?section=000100020003

KBioscience: http://kbioscience.co.uk/

WGAViewer: http://www.genome.duke.edu/centers/pg2/downloads/wgaviewer.php

SNAP http://www.broadinstitute.org/mpg/snap/

Europe PMC Funders Author Manuscripts

IMPUTE: https://mathgen.stats.ox.ac.uk/impute/impute.html

SNPTEST: http://www.stats.ox.ac.uk/~marchini/software/gwas/snptest.html

EEL: http://www.cs.helsinki.fi/research/algodan/EEL/

PReMod: http://genomequebec.mcgill.ca/PReMod/welcome.do

Transfac Matrix Database: http://www.biobase-international.com/pages/index.php?id=transfac

JASPAR2 database: http://jaspar.cgb.ki.se/

EIGENSTRAT: http://genepath.med.harvard.edu/~reich/Software.htm

Wellcome Trust Case Control Consortium: www.wtccc.org.uk

METAL: www.sph.umich.edu/csg/abecasis/metal

REFERENCES

1. Kuppers R. The biology of Hodgkin's lymphoma. Nat Rev Cancer. 2009; 9:15–27. [PubMed: 19078975]
2. Swerdlow AJ. Epidemiology of Hodgkin's disease and non-Hodgkin's lymphoma. Eur J Nucl Med Mol Imaging. 2003; 30(Suppl 1):S3–12. [PubMed: 12734687]
3. Smith A, et al. The Haematological Malignancy Research Network (HMRN): a new information strategy for population based epidemiology and health service research. Br J Haematol. 148:739–53. [PubMed: 19958356]
4. Kutok JL, Wang F. Spectrum of Epstein-Barr virus-associated diseases. Annu Rev Pathol. 2006; 1:375–404. [PubMed: 18039120]
5. Goldin LR, et al. Familial aggregation of Hodgkin lymphoma and related tumors. Cancer. 2004; 100:1902–8. [PubMed: 15112271]
6. Amiel, J. Study of the leukocyte phenotypes in Hodgkin's disease. In: Teraski, PI., editor. Histocompatibility testing. Munksgaard; Copenhagen: 1967. p. 79-81.
7. Klitz W, Aldrich CL, Fildes N, Horning SJ, Begovich AB. Localization of predisposition to Hodgkin disease in the HLA class II region. Am J Hum Genet. 1994; 54:497–505. [PubMed: 8116619]
8. Oza AM, et al. A clinical and epidemiological study of human leukocyte antigen-DPB alleles in Hodgkin's disease. Cancer Res. 1994; 54:5101–5. [PubMed: 7923125]
9. Hjalgrim H, 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. 107:6400–5. [PubMed: 20308568]
10. Niens M, et al. HLA-A*02 is associated with a reduced risk and HLA-A*01 with an increased risk of developing EBV+ Hodgkin lymphoma. Blood. 2007; 110:3310–5. [PubMed: 17630352]
11. Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. Am J Hum Genet. 1987; 40:1–14. [PubMed: 3468804]
12. Goldin LR, et al. A genome screen of families at high risk for Hodgkin lymphoma: evidence for a susceptibility gene on chromosome 4. J Med Genet. 2005; 42:595–601. [PubMed: 15994882]
13. Mack TM, et al. Concordance for Hodgkin's disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease. N Engl J Med. 1995; 332:413–8. [PubMed: 7824015]
14. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). Int J Epidemiol. 2006; 35:34–41. [PubMed: 16155052]

Europe PMC Funders Author Manuscripts

15. Clayton DG, et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat Genet. 2005; 37:1243–6. [PubMed: 16228001]
16. Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–9. [PubMed: 16862161]
17. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447:661–78. [PubMed: 17554300]
18. Cabannes E, Khan G, Aillet F, Jarrett RF, Hay RT. Mutations in the IkBa gene in Hodgkin's disease suggest a tumour suppressor role for IkappaBalpha. Oncogene. 1999; 18:3063–70. [PubMed: 10340377]
19. Emmerich F, et al. Overexpression of I kappa B alpha without inhibition of NF-kappaB activity and mutations in the I kappa B alpha gene in Reed-Sternberg cells. Blood. 1999; 94:3129–34. [PubMed: 10556199]
20. Schmitz R, et al. TNFAIP3 (A20) is a tumor suppressor gene in Hodgkin lymphoma and primary mediastinal B cell lymphoma. J Exp Med. 2009; 206:981–9. [PubMed: 19380639]
21. Barth TF, et al. Gains of 2p involving the REL locus correlate with nuclear c-Rel protein accumulation in neoplastic cells of classical Hodgkin lymphoma. Blood. 2003; 101:3681–6. [PubMed: 12511414]
22. Joos S, et al. Hodgkin's lymphoma cell lines are characterized by frequent aberrations on chromosomes 2p and 9p including REL and JAK2. Int J Cancer. 2003; 103:489–95. [PubMed: 12478664]
23. Martin-Subero JI, et al. Recurrent involvement of the REL and BCL11A loci in classical Hodgkin lymphoma. Blood. 2002; 99:1474–7. [PubMed: 11830502]
24. Villeneuve L, Rassart E, Jolicoeur P, Graham M, Adams JM. Proviral integration site Mis-1 in rat thymomas corresponds to the pvt-1 translocation breakpoint in murine plasmacytomas. Mol Cell Biol. 1986; 6:1834–7. [PubMed: 3785181]
25. Beck-Engeser GB, et al. Pvt1-encoded microRNAs in oncogenesis. Retrovirology. 2008; 5:4. [PubMed: 18194563]
26. Bakkus MH, Brakel-van Peer KM, Michiels JJ, van 't Veer MB, Benner R. Amplification of the c-myc and the pvt-like region in human multiple myeloma. Oncogene. 1990; 5:1359–64. [PubMed: 2216459]
27. Huppi K, Siwarski D, Skurla R, Klinman D, Mushinski JF. Pvt-1 transcripts are found in normal tissues and are altered by reciprocal(6;15) translocations in mouse plasmacytomas. Proc Natl Acad Sci U S A. 1990; 87:6964–8. [PubMed: 2402486]
28. Storlazzi CT, et al. Identification of a commonly amplified 4.3 Mb region with overexpression of C8FW, but not MYC in MYC-containing double minutes in myeloid malignancies. Hum Mol Genet. 2004; 13:1479–85. [PubMed: 15163636]
29. Crowther-Swanepoel D, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. Nat Genet. 42:132–6. [PubMed: 20062064]
30. Gudmundsson J, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet. 2007; 39:631–7. [PubMed: 17401366]
31. Easton DF, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007; 447:1087–93. [PubMed: 17529967]
32. Tomlinson I, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet. 2007; 39:984–8. [PubMed: 17618284]
33. Yeager M, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet. 2007; 39:645–9. [PubMed: 17401363]
34. Amundadottir LT, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet. 2006; 38:652–8. [PubMed: 16682969]
35. Kiemeney LA, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat Genet. 2008; 40:1307–12. [PubMed: 18794855]
36. Tuupanen S, et al. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat Genet. 2009; 41:885–90. [PubMed: 19561604]

37. Pomerantz MM, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet. 2009; 41:882–4. [PubMed: 19561607]
38. Faumont N, et al. c-Myc and Rel/NF-kappaB are the two master transcriptional systems activated in the latency III program of Epstein-Barr virus-immortalized B cells. J Virol. 2009; 83:5014–27. [PubMed: 19264782]
39. 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; 9:125–35. [PubMed: 19151747]
40. Li J, Qian CN, Zeng YX. Regulatory T cells and EBV associated malignancies. Int Immunopharmacol. 2009; 9:590–2. [PubMed: 19539572]
41. Atayar C, et al. Expression of the T-cell transcription factors, GATA-3 and T-bet, in the neoplastic cells of Hodgkin lymphomas. Am J Pathol. 2005; 166:127–34. [PubMed: 15632006]
42. La Rosa FA, Pierce JW, Sonenshein GE. Differential regulation of the c-myc oncogene promoter by the NF-kappa B rel family of transcription factors. Mol Cell Biol. 1994; 14:1039–44. [PubMed: 8289784]
43. Maurice D, Hooper J, Lang G, Weston K. c-Myb regulates lineage choice in developing thymocytes via its target gene Gata3. EMBO J. 2007; 26:3629–40. [PubMed: 17641686]
44. Papaemmanuil E, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet. 2009; 41:1006–10. [PubMed: 19684604]
45. Jundt F, et al. Loss of PU.1 expression is associated with defective immunoglobulin transcription in Hodgkin and Reed-Sternberg cells of classical Hodgkin disease. Blood. 2002; 99:3060–2. [PubMed: 11929801]
46. Hjalgrim H, Engels EA. Infectious aetiology of Hodgkin and non-Hodgkin lymphomas: a review of the epidemiological evidence. J Intern Med. 2008; 264:537–48. [PubMed: 19017178]
47. Stranger BE, et al. Genome-wide associations of gene expression variation in humans. PLoS Genet. 2005; 1:e78. [PubMed: 16362079]
48. Stranger BE, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science. 2007; 315:848–53. [PubMed: 17289997]
49. Alexander FE, 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; 107:298–302. [PubMed: 12949811]
50. Jarrett RF, et al. The Scotland and Newcastle epidemiological study of Hodgkin's disease: impact of histopathological review and EBV status on incidence estimates. J Clin Pathol. 2003; 56:811–6. [PubMed: 14600123]
51. 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; 18:130–6. [PubMed: 17099321]
52. Penegar S, et al. National study of colorectal cancer genetics. Br J Cancer. 2007; 97:1305–9. [PubMed: 17895893]
53. Hjalgrim H, et al. Infectious mononucleosis, childhood social environment, and risk of Hodgkin lymphoma. Cancer Res. 2007; 67:2382–8. [PubMed: 17332371]
54. Smedby KE, et al. Ultraviolet radiation exposure and risk of malignant lymphomas. J Natl Cancer Inst. 2005; 97:199–209. [PubMed: 15687363]
55. Sorensen KM, et al. Whole genome amplification on DNA from filter paper blood spot samples: an evaluation of selected systems. Genet Test. 2007; 11:65–71. [PubMed: 17394394]
56. van Leeuwen FE, et al. Roles of radiation dose, chemotherapy, and hormonal factors in breast cancer following Hodgkin's disease. J Natl Cancer Inst. 2003; 95:971–80. [PubMed: 12837833]
57. Broeks A, et al. Increased risk of breast cancer following irradiation for Hodgkin's disease is not a result of ATM germline mutations. Int J Radiat Biol. 2000; 76:693–8. [PubMed: 10866292]
58. De Bruin ML, et al. Breast cancer risk in female survivors of Hodgkin's lymphoma: lower risk after smaller radiation volumes. J Clin Oncol. 2009; 27:4239–46. [PubMed: 19667275]
59. Lake A, et al. Mutations of NFKBIA, encoding IkappaB alpha, are a recurrent finding in classical Hodgkin lymphoma but are not a unifying feature of non-EBV-associated cases. Int J Cancer. 2009; 125:1334–42. [PubMed: 19507254]

60. Pettiti, D. Meta-analysis Decision Analysis and Cost-effectiveness Analysis. Oxford University Press; Oxford, New York: 1994.
61. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002; 21:1539–58. [PubMed: 12111919]
62. Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. A fine-scale map of recombination rates and hotspots across the human genome. Science. 2005; 310:321–4. [PubMed: 16224025]
63. Gabriel SB, et al. The structure of haplotype blocks in the human genome. Science. 2002; 296:2225–9. [PubMed: 12029063]
64. Matys V, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res. 2006; 34:D108–10. [PubMed: 16381825]
65. Ferretti V, et al. PReMod: a database of genome-wide mammalian cis-regulatory module predictions. Nucleic Acids Res. 2007; 35:D122–6. [PubMed: 17148480]
66. Hallikas O, et al. Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. Cell. 2006; 124:47–59. [PubMed: 16413481]
67. Cuzick J. A Wilcoxon-type test for trend. Stat Med. 1985; 4:87–90. [PubMed: 3992076]

Europe PMC Funders Author Manuscripts

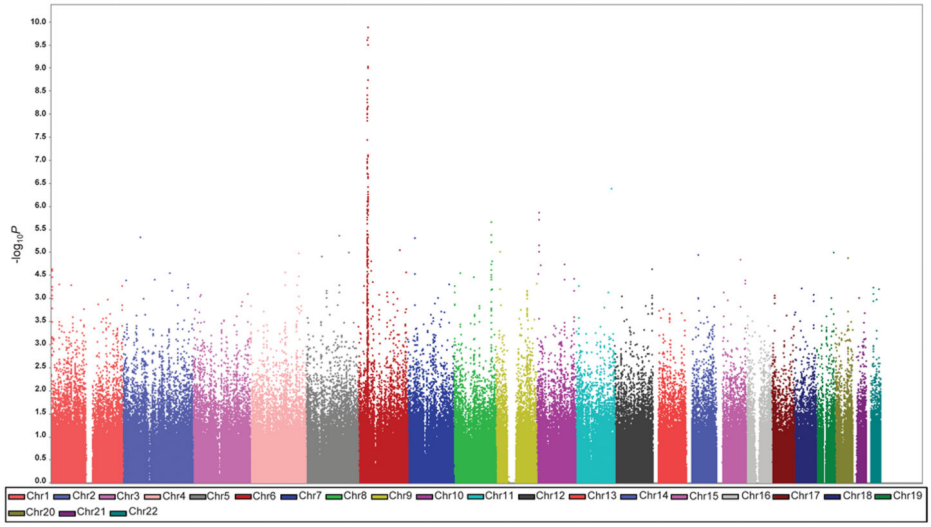
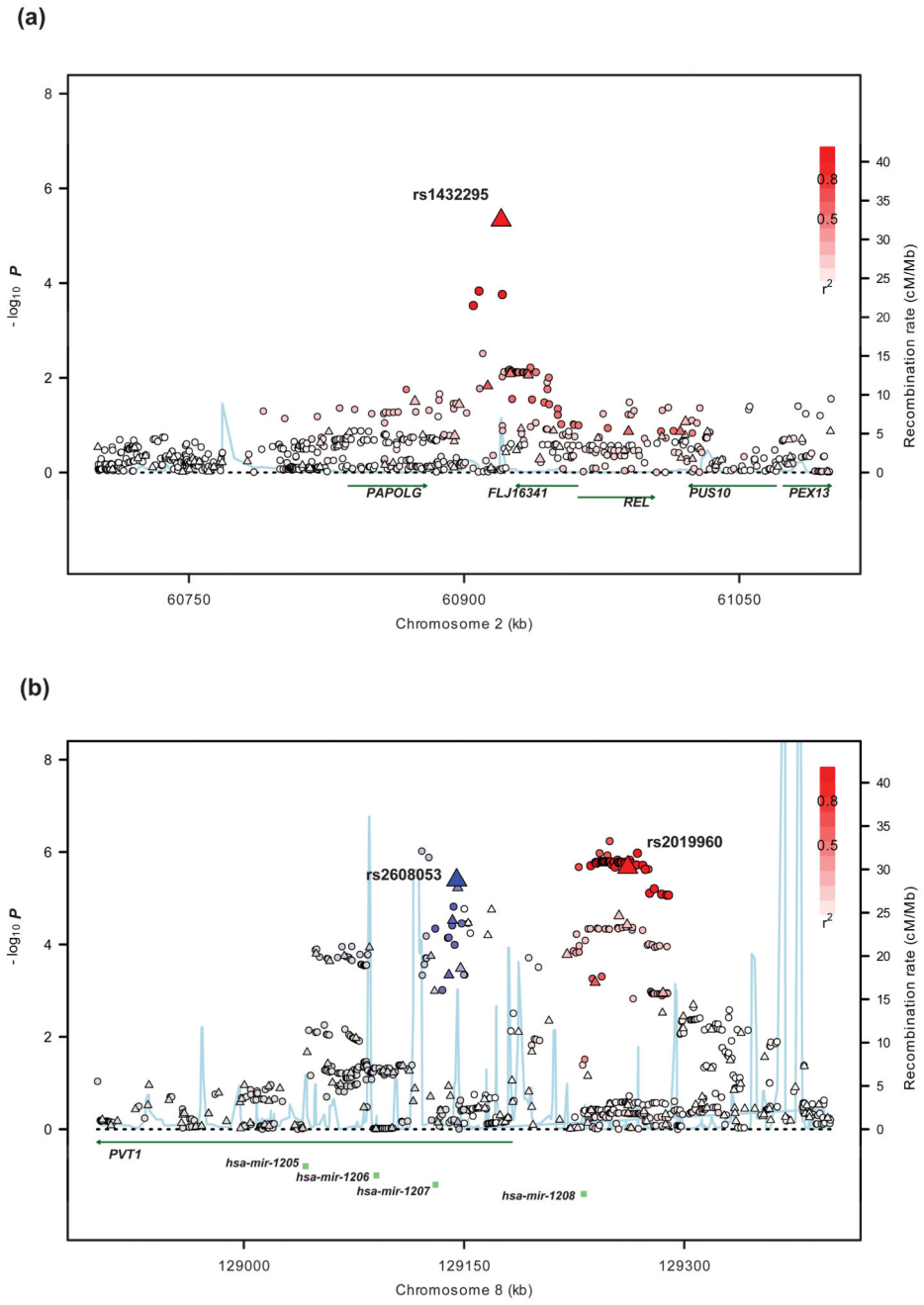


Figure 1. Genome-wide association results from the initial GWAS

Shown are the genome-wide *P*-values obtained using the Cochran-Armitage trend test from 504,374 autosomal SNPs in 589 HL cases and 5,199 controls. *P*-values ($-\log_{10}P$, y axis) are plotted against their respective chromosomal positions (x axis). Each chromosome is depicted in a different color. The points with $P<10^{-10}$ were truncated; the smallest *P* value is 8.12×10^{-21}.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts



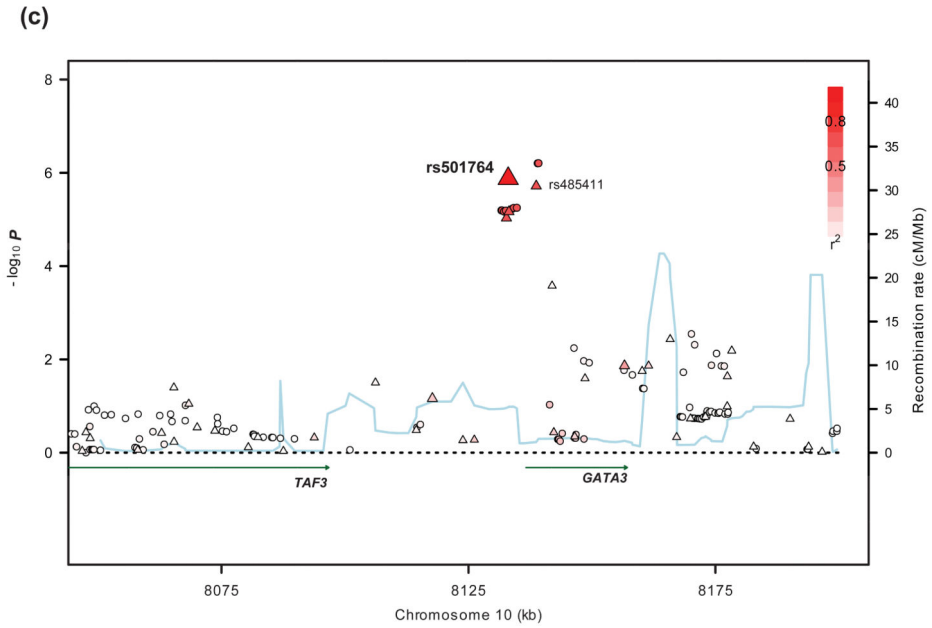


Figure 2. Regional plots of association results and recombination rates for 2p16.1, 8q24.21, and 10p14 susceptibility loci

(a-c) Association results of both genotyped (triangles) and imputed (circles) SNPs in the GWAS samples and recombination rates within the three loci: (a) 2p16.1; (b) 8q24.21; (c) 10p14. For each plot, $-\log_{10}P$ values (y-axis) of the SNPs are shown according to their chromosomal positions (x-axis). The top genotyped SNP in the combined analysis is labeled by rs ID. The color intensity of each symbol reflects the extent of LD with the top genotyped SNP – red/blue (r^2>0.8) through to white (r^2<0.2). Genetic recombination rates (cM/Mb), estimated using HapMap CEU samples, are shown with a light blue line. Physical positions are based on build 36 (NCBI) of the human genome. Also shown are the relative positions of genes and transcripts mapping to each region of association. Genes and miRNAs have been redrawn to show the relative positions; therefore, maps are not to physical scale.

Europe PMC Funders Author Manuscripts

Table 1

Summary results for six SNPs associated with classical Hodgkin's lymphoma risk

Chr	SNP	Location (bps)[a]	Gene[b]	Risk allele	RAF[c] control	GWAS		Replication studies		Combined			
						OR[d] (95% CI)	*P* *	OR (95% CI)	*P*	OR (95% CI)	*P* †	P_{het}	I^2
2p16.1	rs1432295	60,920,170	*REL*	G	0.40	1.33 (1.18-1.51)	4.69×10^{-6}	1.17 (1.08-1.27)	1.91×10^{-4}	1.22 (1.14-1.30)	1.91×10^{-8}	0.25	26%
6p21.32	rs6903608	32,536,263	*HLA-DRA*	G	0.27	1.81 (1.60-2.05)	8.12×10^{-21}	1.65 (1.52-1.80)	4.95×10^{-32}	1.70 (1.58-1.82)	2.84×10^{-50}	0.12	46%
8q24.21	rs2608053	129,145,014	*PVT1*	G	0.52	1.33 (1.18-1.50)	4.06×10^{-6}	1.15 (1.06-1.24)	8.38×10^{-4}	1.20 (1.12-1.28)	1.16×10^{-7}	0.10	48%
8q24.21	rs2019960	129,261,453	*PVT1*	G	0.23	1.38 (1.21-1.58)	2.14×10^{-6}	1.31 (1.19-1.44)	8.92×10^{-9}	1.33 (1.23-1.44)	1.26×10^{-13}	0.89	0%
10p14	rs501764	8,133,040	*GATA3*	C	0.19	1.42 (1.23-1.63)	1.33×10^{-6}	1.18 (1.07-1.30)	1.28×10^{-3}	1.25 (1.15-1.36)	7.05×10^{-8}	0.09	51%
10p14	rs485411	8,133,191	*GATA3*	A	0.25	1.35 (1.18-1.54)	6.83×10^{-6}	1.17 (1.07-1.28)	8.68×10^{-4}	1.22 (1.13-1.32)	1.29×10^{-7}	0.25	26%

Detailed data including genotype counts are shown in Supplementary Table 3.

[a] Chromosome location based on NCBI Human Genome Build 36 coordinates.

[b] Putative candidate genes mapping within 50 kb of respective SNPs.

[c] Risk allele frequency.

[d] Odds ratio with 95% Confidence Interval.

* EIGENSTRAT-adjusted *P*-values: rs1432295, $P = 8.87 \times 10^{-6}$; rs6903608, $P = 2.93 \times 10^{-17}$; rs2608053, $P = 4.20 \times 10^{-6}$; rs2019960, $P = 9.14 \times 10^{-7}$; rs501764, $P = 1.67 \times 10^{-6}$; rs485411, $P = 1.25 \times 10^{-5}$.

† Combined P-values using adjusted data rs1432295, $P = 5.02 \times 10^{-9}$; rs6903608, $P = 1.86 \times 10^{-46}$; rs2608053, $P = 1.84 \times 10^{-8}$; rs2019960, $P = 4.01 \times 10^{-14}$; rs501764, $P = 1.80 \times 10^{-8}$; rs485411, $P = 4.51 \times 10^{-8}$.