International Journal of Epidemiology, 2021, 1325–1334
doi: 10.1093/ije/dyab042
Advance Access Publication Date: 22 March 2021
Original article

Genetics

Common genetic polymorphisms contribute to the association between chronic lymphocytic leukaemia and non-melanoma skin cancer

Caroline Besson [1†], Amy Moore,[2†] Wenting Wu,[3†] Claire M Vajdic,[4] Silvia de Sanjose,[5] Nicola J Camp,[6] Karin E Smedby,[7] Tait D Shanafelt,[8] Lindsay M Morton,[2] Jerry D Brewer,[9] Lydia Zablotska,[10] Eric A Engels,[2] James R Cerhan [11], Susan L Slager,[11] Jiali Han [12‡] and Sonja I Berndt[2‡]* on behalf of the InterLymph Consortium. Full authors list is given at the end of the manuscript

[1]Service d'hématologie et Oncologie, Centre Hospitalier de Versailles, Le Chesnay; Université Paris-Saclay, UVSQ, Inserm, Équipe "Exposome et Hérédité", CESP, 94805, Villejuif, France, [2]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA, [3]Department of Medical and Molecular Genetics, Indiana University School of Medicine, USA, [4]Centre for Big Data Research in Health, University of New South Wales, Sydney, New South Wales, Australia, [5]CIBER de Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain, [6]Department of Internal Medicine and Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT, USA, [7]Department of Medicine, Solna, Karolinska Institutet, Stockholm, Sweden, [8]Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA, [9]Department of Dermatology, Mayo Clinic, Rochester, MN, USA, [10]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA, [11]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA and [12]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

[†]These authors contributed equally to this work.

[‡]These authors jointly directed this work.

*Corresponding author. Division of Cancer Epidemiology & Genetics, Occupational and Environmental Epidemiology Branch, NCI Shady Grove, 9609 Medical Center Drive, MSC 9776, Bethesda, MD 20892, USA. E-mail: berndts@mail.nih.gov

Editorial decision 21 January 2021; Accepted 19 February 2021

Abstract

Background: Epidemiological studies have demonstrated a positive association between chronic lymphocytic leukaemia (CLL) and non-melanoma skin cancer (NMSC). We hypothesized that shared genetic risk factors between CLL and NMSC could contribute to the association observed between these diseases.

Methods: We examined the association between (i) established NMSC susceptibility loci and CLL risk in a meta-analysis including 3100 CLL cases and 7667 controls and (ii) established CLL loci and NMSC risk in a study of 4242 basal cell carcinoma (BCC) cases, 825 squamous cell carcinoma (SCC) cases and 12802 controls. Polygenic risk scores (PRS)

for CLL, BCC and SCC were constructed using established loci. Logistic regression was used to estimate odds ratios (ORs) and 95% confidence intervals (CIs).

Results: Higher CLL-PRS was associated with increased BCC risk ($OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.13$, 95% CI: 1.02–1.24, $P_{trend} = 0.009$), even after removing the shared 6p25.3 locus. No association was observed with BCC-PRS and CLL risk ($P_{trend} = 0.68$). These findings support a contributory role for CLL in BCC risk, but not for BCC in CLL risk. Increased CLL risk was observed with higher SCC-PRS ($OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.22$, 95% CI: 1.08–1.38, $P_{trend} = 1.36 \times 10^{-5}$), which was driven by shared genetic susceptibility at the 6p25.3 locus.

Conclusion: These findings highlight the role of pleiotropy regarding the pathogenesis of CLL and NMSC and shows that a single pleiotropic locus, 6p25.3, drives the observed association between genetic susceptibility to SCC and increased CLL risk. The study also provides evidence that genetic susceptibility for CLL increases BCC risk.

Key words: CLL, NMSC, polygenic risk score, pleiotropy

Key Messages

- We explored the hypothesis that shared genetic risk factors between chronic lymphocytic leukaemia (CLL) and non-melanoma skin cancer (NMSC) could contribute to the association observed between these diseases.
- Our findings provide evidence that genetic susceptibility for CLL increases basal cell carcinoma risk.
- The results also highlight the pleiotropy between CLL and NMSC, showing that a single locus, 6p25.3, drives the association between genetic susceptibility to squamous cell carcinoma and increased CLL risk.

Introduction

Genome-wide association studies (GWASs) have identified thousands of genetic variants associated with numerous traits and diseases.[1] These findings provide opportunities to investigate pleiotropy among multiple diseases. Seemingly disparate diseases can share common susceptibility single-nucleotide polymorphisms (SNPs) and may therefore share biological mechanisms or pathways[2]. The analysis of shared genetic associations can help explain previous observations between diseases and provide better insight into their aetiology.Epidemiological studies have demonstrated a consistent positive association between leukaemia, specifically chronic lymphocytic leukaemia (CLL), and non-melanoma skin cancer (NMSC).[3–10] In registry-based studies, estimates of the risk of developing CLL following NMSC range from 1.1 to 2.4.[4–6] Estimates of the risk of NMSC following CLL are generally larger and range between 2.4 and 8.6.[3–5] The underlying causes of the observed association between CLL and NMSC are not well understood. Several factors, including increased screening,[6] shared environmental or occupational risk factors, which may have immunotoxic effects,[11] genetic factors or common biological pathways have been hypothesized to play a role. Moreover, the increased risk of NMSC following CLL is thought to be due in part to immunosuppression induced by CLL per se or its treatments.[12] There are several types of NMSC with basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) constituting the majority (~99%) of NMSC. Previous GWAS of CLL, BCC and SCC have uncovered multiple genetic loci associated with susceptibility to these diseases.[13–29] Established susceptibility loci explain 17–25% of the familial risk for CLL,[13,14] 11% for BCC[21] and 6% for SCC.[20]

In the present study, we evaluated the hypothesis that shared genetic risk factors between CLL and NMSC, specifically BCC and SCC, may contribute to the observed association between these diseases. Because individual SNPs only explain a small portion of any shared risk, we constructed polygenic risk scores (PRS) to aggregate the genetic burden of disease risk into a single quantitative index for each disease and evaluated the genetic contribution of SCC- and BCC-associated SNPs on CLL risk and CLL-associated SNPs on SCC and BCC risk.

Methods

Study populations

We used data from previous genome-wide association studies to investigate shared genetic risk factors between NMSC and CLL.[14,20,21] The association between NMSC loci and CLL risk was evaluated using data from four CLL GWAS of European ancestry: National Cancer Institute NHL GWAS (NCI GWAS),[15] Utah Chronic Lymphocytic Leukemia GWAS (Utah), Genetic Epidemiology of CLL Consortium GWAS (GEC)[30] and Molecular Epidemiology of Non-Hodgkin Lymphoma GWAS (UCSF).[31] CLL cases for these studies were ascertained from clinics or hospitals, cancer registries or through self-report verified by medical and pathology reports. For the NCI GWAS, the International Lymphoma Epidemiology Consortium (InterLymph) Data Coordinating Center reviewed ICD codes and other available pathology/medical information and classified cases according to the hierarchical classification of the InterLymph Pathology Working Group based on the World Health Organization (WHO) classification (2008).[32] Individuals with ICD-O-2/3 morphology codes 9823 and 9670 were classified as CLL cases.

CLL cases and controls were genotyped using the Illumina OmniExpress or Omni2.5 (NCI GWAS), Affymetrix 6.0 (GEC), Illumina HumanHap 610 K (Utah) or Illumina HumanCNV370-Duo (UCSF). After quality-control metrics, including removal of samples with poor call rates, non-European ancestry, gender discordance, relatedness and abnormal heterozygosity, the genotype data from each GWAS were imputed separately using IMPUTE2 and the 1000 Genomes Project (March 2012 release) reference panel. Only SNPs with an info score >0.3 were considered for analysis. Across these four CLL GWAS, a total of 3100 CLL cases and 7667 controls were available for the analysis. This included 2179 cases and 6221 controls from the NCI GWAS, 387 cases and 294 controls from GEC, 213 cases and 747 controls from UCSF, and 321 cases and 405 controls from Utah (**Supplementary Table 1**, available as Supplementary data at *IJE* online). All studies obtained informed consent from their participants and approval from their respective institutional review boards.

The evaluation of the CLL loci on the risk of NMSC was conducted using GWAS data from the Nurses' Health Study (NHS) and Health Professionals Follow-up Study (HPFS).[20,21] Eligible cases were participants of European ancestry who self-reported BCC or SCC on one of the biennial questionnaires after enrolment in NHS or HPFS. For SCC, cases were limited to those with pathologically confirmed invasive SCC based on medical-record review. For BCC, no medical-record review was done; however, because the participants are health professionals, the validity of their self-reports is expected to be high.[33] For the BCC case–control study, controls were participants without reported BCC. For the SCC case–control study, controls were participants without reported SCC; persons without available information on SCC diagnosis were excluded. Only cases and controls with available genotype data from previous nested case–control GWAS were included in this study.[20,21] Data from these subjects with GWAS data were compiled into three analytic data sets based on their genotyping platform: Affymetrix (Affy 6.0), Illumina HumanHap (550 K, 610Q, 660) or Illumina Omni Express. Each data set underwent quality-control metrics and was imputed using the 1000 Genomes Project ALL Phase I Integrated Release Version 3 Haplotypes reference panel and Minimac (v.2012–08015). Only SNPs with a quality-control score $r^2 > 0.3$ were considered for analysis. There were 1197 BCC cases and 3706 controls genotyped on the Illumina Omni Express, 1268 BCC cases and 3685 controls with Illumina HumanHap data, and 1777 BCC cases and 5411 controls with Affymetrix data, resulting in a total of 4242 BCC cases and 12802 controls for analysis (**Supplementary Table 2**, available as Supplementary data at *IJE* online). For the SCC analysis, data were available for 238 SCC cases and 3164 controls genotyped on the Illumina Omni Express, 220 SCC cases and 2901 controls with Illumina HumanHap data, and 367 SCC cases and 5453 controls with Affymetrix data, resulting in a total of 825 SCC cases and 11518 controls (**Supplementary Table 2**, available as Supplementary data at *IJE* online). The study was approved by the institutional review boards at Brigham and Women's Hospital and the Harvard School of Public Health, and all participants provided informed consent.

Identification of susceptibility SNPs and measures of linkage disequilibrium (LD)

We downloaded the publicly available NHGRI-EBI Catalog of published GWAS in 2019[34] and identified the SNPs that were reported to be associated with CLL or NMSC at a genome-wide significance level ($P < 5 \times 10^{-8}$) in populations of European descent. We also searched PubMed for GWAS of CLL and NMSC. The search of NMSC-associated SNPs used the terms 'skin cancer', 'NMSC', 'basal cell carcinoma' or 'squamous cell carcinoma'. Studies of melanoma were excluded. SNPs were assumed to be independent for a particular outcome if they were >1 Mb away from each other or found to have a $r^2 < 0.05$ or shown to be independent in conditional

analyses. A total of 43 independent SNPs were identified for CLL, 35 for BCC and 14 for SCC (**Supplementary Tables 3 and 4**, available as Supplementary data at *IJE* online). One SCC-associated SNP (rs74899442) could not be found within the data sets for CLL and so a proxy (rs151267007, $D' = 1.0$, $r^2 = 0.43$) was utilized for analysis. Another SCC-associated SNP (rs192481803) was only adequately imputed in one of the four GWAS for CLL and so was not included in the PRS.

To evaluate pairwise LD, we estimated both r^2 and |D'| based on data from the European populations present in the 1000 Genomes Project Phase 3 data using LDlink.[35] We estimated LD using these two metrics because, together, they provide a fuller picture of the LD structure. The metric r^2 is highly dependent on allele frequency and may give false negatives, especially with rare variants; although |D'| is robust to differences in allele frequencies, it may provide more false positives.

Statistical analysis

Individual SNP associations, including odds ratios (ORs) and 95% confidence intervals (95% CIs), were estimated for each disease separately by GWAS platform using logistic regression, adjusting for age, sex and principal components for ancestry. The PRS were constructed separately for each disease using the SNPs identified in the literature as being independently associated with CLL, BCC or SCC. The PRS for each cancer was calculated by multiplying the allele dosage for each SNP by the log OR of the SNP as reported in the literature and summing across all SNPs. In addition to the BCC- and SCC-specific PRS, an overall NMSC-PRS was created to capture susceptibility to NMSC more broadly and covered a total of 45 SNPs, including for loci common to both BCC and SCC, both index SNPs unless the correlation was modest/high ($r^2 > 0.25$). For each PRS, quartiles were constructed based on the distribution among the controls in each GWAS. P_{trend} was estimated based on the PRS as a continuous variable. GWAS-specific ORs for the PRS were estimated using logistic regression, adjusting for age, sex and principal components for ancestry. Combined ORs, *p*-values and 95% CIs were calculated by combining GWAS-specific ORs in a fixed-effects meta-analysis; no between-GWAS heterogeneity was observed for CLL ($P > 0.1$ for all). Adjustment for multiple testing was performed using the Benjamini–Hochberg false discovery rate (FDR) method.[36] Analyses for NMSC were performed using Plink. Analyses for CLL were conducted using R v3.1.2 and STATA.

Results

Shared chromosomal regions between CLL and NMSC

Of the 43 established genetic loci for CLL and 45 for NMSC, eight chromosomal regions contained at least one SNP associated with CLL and one SNP associated with NMSC within 1 Mb of each other: 1q42.13, 2q33.1, 3q28, 5p15.33, 6p21.32, 6p25.3, 9q21.3 and 19p13.3. The LD between the CLL and NMSC loci was weak for most of these regions (|D'| < 0.5, $r^2 < 0.15$), except for 3q28 (*LPP*) and 6p25.3 (*IRF4/EXOC2*). At the 6p25.3 locus, the SNP associated with SCC (rs12203592) and the SNP associated with CLL (rs872071) were in modest LD (|D'| = 0.74), although the correlation between them was quite low ($r^2 = 0.08$). The LD between the BCC SNP (rs12210050) at 6p25.3 and CLL SNP (rs872071) was weaker (|D'| = 0.36, $r^2 = 0.02$). At the 3q28 locus, the CLL SNP (rs9815073) was in strong LD with the BCC SNP (rs191177147, |D'| = 0.99, $r^2 = 0.69$) but not with the SCC SNP (rs6791479, |D'| = 0.07, $r^2 = 0.002$). Whereas the SCC risk allele, rs12203592-T, was positively correlated with the CLL risk allele at 6p25.3 (rs872071-G), the BCC risk allele at 3q28, rs191177147-T, was negatively correlated with risk allele for CLL, rs9815073-C.

CLL risk associated with individual NMSC loci and PRS

Of the 35 established susceptibility loci for BCC and 14 reported loci for SCC, eight index SNPs for BCC and one SNP for SCC were nominally associated with CLL risk (**Supplementary Table 3**, available as Supplementary data at *IJE* online). After adjustment for multiple testing, three BCC SNPs and one SCC SNP remained associated with CLL risk at a FDR < 5%: rs191177147 at 3q28 (*LPP*), rs12203592 and rs12210050 at 6p25.3 (*IRF4/EXOC2*), rs78378222 at 17p13.1 (*TP53*) (**Supplementary Figure 1**, available as Supplementary data at *IJE* online). Although chromosomes 3q28 and 6p25.3 are established CLL loci,[14,18] 17p13.1 has not been previously reported to be associated with CLL risk. For rs78378222 at 17p13.1, the BCC risk allele was associated with an increased risk of CLL (OR = 1.92, 95% CI: 1.45–2.55, $P = 6.30 \times 10^{-6}$).

No association was observed with the BCC-PRS ($OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.02$, 95% CI: 0.90–1.16), but a higher SCC-PRS was associated with an increased risk of CLL ($OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.22$, 95% CI: 1.08–1.38) (**Table 1** and **Supplementary Table 5**, available as Supplementary data at *IJE* online). To determine whether the association was driven by one or more chromosomal regions shared between SCC and CLL, we conducted

Table 1 Risk of CLL associated with polygenic risk scores of established NMSC loci[a]

	Quartile 1	Quartile 2		Quartile 3		Quartile 4		P_{trend}
	OR	OR[b] (95% CI)	*P*	OR[b] (95% CI)	*P*	OR[b] (95% CI)	*P*	
BCC-PRS	1.0 (ref.)	1.10 (0.98–1.25)	0.11	1.03 (0.91–1.17)	0.64	1.02 (0.90–1.16)	0.72	0.68
SCC-PRS	1.0 (ref.)	0.97 (0.85–1.09)	0.58	1.09 (0.96–1.24)	0.17	1.22 (1.08–1.38)	0.002	1.36×10^{-5}
NMSC-PRS	1.0 (ref.)	1.04 (0.92–1.18)	0.55	1.12 (0.99–1.26)	0.08	1.17 (1.03–1.33)	0.01	0.002

[a]Polygeneic risk scores (PRS) are based on previously published loci for basal cell carcinoma (BCC), squamous cell carcinoma (SCC) or any type of non-melanoma skin cancer (NMSC).

[b]Odds ratios (ORs) and 95% confidence intervals (95% CIs) for chronic lymphocytic leukemia (CLL) are based on fixed-effect meta-analysis ($N = 3100$ cases, 7667 controls) and are adjusted for age, sex and principal components.

sensitivity analyses. After removal of the SNPs at 6p25.3, no association was observed with CLL risk for the SCC-PRS ($OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.01$, 95% CI: 0.89–1.14). Similar results were observed after removal of all SNPs located within 1 Mb of established CLL loci ($OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 0.97$, 95% CI: 0.86–1.10). We explored the combined contribution of SCC and BCC by creating a NMSC-PRS using loci for both BCC and SCC. An increased risk of CLL was observed with higher NMSC-PRS ($OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.17$, 95% CI: 1.03–1.33).

NMSC risk associated with individual CLL loci and PRS

Of the 43 established susceptibility loci for CLL, two SNPs (4q24 and 18q21.3) were nominally associated with risk of both BCC and SCC and five SNPs (1q42.13, 5p15.33, 6p21.32, 6p25.3 and 11q23.2) with either BCC or SCC risk (**Supplementary Table 4**, available as Supplementary data at *IJE* online), but none had a FDR < 5%. Higher CLL-PRS was associated with increased risk of BCC (**Table 2**, $OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.13$, 95% CI: 1.02–1.24, $P = 0.02$). Removal of 6p25.3 (*IRF4/EXOC2*) SNP or the 3q28 (*LPP*) SNP from the CLL-PRS did substantially not modify these results ($OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.11$, 95% CI: 1.00–1.22; and $OR_{4th\text{-}quartile\text{-}vs\text{-}1st\text{-}quartile} = 1.14$, 95% CI: 1.03–1.26, respectively). Although the ORs were elevated, no trend was observed between increased CLL-PRS and SCC risk ($P_{trend} = 0.45$).

Discussion

We investigated the extent to which genetic factors may contribute to the observed association between CLL and NMSC. Using data from several large GWAS, we identified multiple shared loci and constructed PRS of established loci for the diseases. We found that higher CLL-PRS was associated with an increased risk of BCC providing evidence that increased genetic susceptibility to CLL contributes to higher BCC risk. We also discovered that higher SCC-PRS was associated with increased CLL risk; interestingly, a single pleiotropic locus, the 6p25.3 locus, appeared to be the main determinant of the association.

Epidemiological studies have shown a consistent association between NMSC and CLL.[3–6] Although the underlying aetiology for this association is largely unknown, several hypotheses have been put forward. One early hypothesis was that the association was due to increased exposure to ultraviolet radiation among individuals developing skin cancer, but large population-based studies have largely rebuked the hypothesis that there is an association between sun exposure and CLL risk.[37,38] A second hypothesis is that the increased risk of NMSC after CLL is due to immune suppression induced by CLL or its

Table 2 Risk of NMSC associated with polygenic risk score of established CLL loci[a]

	Quartile 1	Quartile 2		Quartile 3		Quartile 4		P_{trend}
	OR	OR[b] (95% CI)	*P*	OR[b] (95% CI)	*P*	OR[b] (95% CI)	*P*	
Basal cell carcinoma								
CLL-PRS	1.0 (ref.)	1.03 (0.93–1.13)	0.62	1.09 (0.98–1.20)	0.10	1.13 (1.02–1.24)	0.02	0.009
Squamous cell carcinoma								
CLL-PRS	1.0 (ref.)	1.38 (1.06–1.81)	0.02	1.06 (0.80–1.40)	0.70	1.23 (0.94–1.62)	0.14	0.45

[a]Polygeneic risk scores (PRS) are based on established loci for chronic lymphocytic leukaemia (CLL).

[b]Odds ratios (ORs) and 95% confidence intervals (95% CIs) for basal cell carcinoma ($N = 4242$ cases, 12 802 controls) and squamous cell carcinoma ($N = 449$ cases, 11 518 controls) are adjusted for age, sex and principal components.

NMSC, non-melanoma skin cancer.

treatment. Immune suppression is a known risk factor for NMSC[12] and CLL can lead to reduced cancer recognition and antitumor immune activity.[39] Evidence suggests that the incidence of secondary cancers is similar in treated and untreated CLL patients,[40] but some chemotherapy agents, such as fludarabine, may also lead to immunosuppression. Immunosuppression is an attractive hypothesis for explaining the higher risk of NMSC after CLL; however, it may not explain the increased risk of CLL observed following a diagnosis of NMSC. Other hypotheses include increased detection of these cancers due to reinforced medical surveillance after the diagnosis of a first cancer.

We evaluated the hypothesis of shared genetic susceptibility between BCC, SCC and CLL. Although we observed several BCC loci to be individually associated with CLL risk, the BCC-PRS was not associated CLL risk. Of the 35 BCC loci evaluated, the allele that increases BCC risk was only positively associated with CLL risk (e.g. OR ≥ 1) for 17 SNPs. For the remainder, the BCC risk allele was associated with a reduced risk of CLL (OR < 1). Even among the three SNPs with a FDR $< 5\%$, the BCC risk allele was only positively associated with CLL risk for two SNPs [rs12210050 (*EXOC2*), rs78378222 (*TP53*)]; the third SNP [rs191177147 (*LPP*)] was negatively associated with CLL risk. Thus, when the SNPs were combined together in a PRS, there was no association with CLL risk, suggesting that genetic susceptibility to BCC, as a whole, does not increase CLL risk, even though individual variants may be associated with risk.

We did, however, find that elevated genetic susceptibility to SCC was associated with an increased risk of CLL. This association was driven by the shared 6p25.3 (*IRF4/EXOC2*) risk locus and no association was observed after its removal from the PRS. In addition to NMSC and CLL, previous studies have reported SNPs at *IRF4/EXOC2* to be associated with skin and hair colour.[41,42] Consistently with our findings, a previous study based on a systematic search for similarity between traits and diseases through a SNP association database reported a connection between pigmentary characteristics and both CLL and NMSC genetic networks and highlighted the *IRF4* gene as the common pathway.[2] The members of the interferon regulatory factor (IRF) family are transcriptional regulators with multiple biologic functions. *IRF4* is broadly expressed in lymphocytes and skin cells. It is a key factor in the regulation of the differentiation and activation of lymphocytes, and IRF–/–Vh11 mice have been shown to develop CLL.[43] *IRF4* polymorphisms may play a role in NMSC by curtailing the host immune response against atypical keratinocytes in the skin.[28] The SCC risk allele rs12203592-T has been shown to reduce *IRF4* transcription in melanocytes[44] and the CLL risk allele rs872071-G has been reported to be associated with reduced *IRF4* mRNA expression in EBV-transformed lymphocytes.[18] Although both risk alleles reduce expression, evidence suggests that *IRF4* expression is regulated by different sets of regulatory elements in lymphocytes and melanocytes.[44] Thus, the shared genetic risk observed at the 6p25.3 locus may be due to LD between the two variants as opposed to a single functional variant.

In contrast to the findings for SCC and CLL risk, the positive association between CLL-PRS and BCC risk was maintained after removal of the shared 6p25.3 locus. Although we did not observe an association between the CLL-PRS and SCC risk, the OR estimate was >1 and the lack of statistically significant association may have been due to a lack of power with the smaller sample size. These results provide evidence that CLL confers an increased risk of NMSC. In population-based studies, an increased incidence of both BCC and SCC has been observed after CLL diagnosis.[5] More broadly, CLL has been associated with an increased risk of bladder, breast, kidney, lung and prostate cancers, and a 1.2- to 2-fold increased risk of second cancers in general.[3,45–48] Although some of the excess risk could be due to increased cancer screening and surveillance, the association between CLL and multiple cancers suggests that shared genetic factors may play a role. Variants at many of the known CLL loci, such as *CDKN2B*, which encodes for the multiple tumour suppressor 2, and *TERT*, the telomerase reverse transcriptase, are also known to be associated with multiple cancers.[49,50] Although CLL and NMSC share some genetic loci, the fact that we observed a positive association with CLL-PRS and BCC risk even after removal of known loci suggests that genetic susceptibility to CLL contributes to increased risk of BCC. The proportion of BCC risk that can be attributed to genetic susceptibility to CLL is likely small, but it would be worthwhile to examine it in more depth as better genetic prediction models are developed in the future.

Our study had several strengths and weaknesses. Our sample size for SCC was relatively small and so our power to detect association with SCC was limited; however, we had larger sample sizes for CLL and BCC. The validity of using PRS to account for the cumulative effect of genetic exposure has been discussed previously[51] and PRS have been used successfully to investigate the genetic contributions of different traits and diseases, such as height and colorectal cancer.[52] The PRS approach has the advantage that genotypes are randomly distributed at birth and therefore unlikely to be confounded by lifestyle, environmental or CLL treatment; however, it is possible that individuals modified some aspects of their behaviour, such as unprotected sun exposure, in response to their skin colour and genetic determinants of pigmentation. Finally, the PRS

used only explain a fraction of the genetic risk of NMSC and CLL, and do not account for the possible impact of rare variants on the risk of these diseases. Despite these limitations, PRS can still provide insight into observed associations between complex diseases and suggest underlying biological mechanisms.[53]

In conclusion, we found evidence that genetic susceptibility to SCC was associated with CLL risk and that the association was driven by a single pleiotropic locus, 6p25.3. However, we found no association between BCC susceptibility and CLL risk. In contrast, we observed that genetic susceptibility to CLL is associated with an increased risk of BCC, providing evidence that CLL increases subsequent NMSC risk and that risk is due in part to underlying genetic factors. As our knowledge of the genetic architecture of CLL grows, we will be able to gain further insight into the underlying biological pathways that contribute to subsequent NMSC risk.

Full author list with affiliations

Caroline Besson,[1]*† Amy Moore,[2]† Wenting Wu,[3]† Claire M Vajdic,[4] Silvia de Sanjose,[5] Nicola J Camp,[6] Karin E Smedby,[7] Tait D Shanafelt,[8] Lindsay M Morton,[2] Jerry D Brewer,[9] Lydia Zablotska,[10] Shengchao A Li,[11] Charles C Chung,[11] Lauren R Teras,[12] Geffen Kleinstern,[13,53] Alain Monnereau,[14] Eleanor Kane,[15] Yolanda Benavente,[5] Mark P Purdue,[2] Brenda M Birmann,[16] Brian K Link,[17] Roel CH Vermeulen,[18] John J Spinelli,[19] Demetrius Albanes,[2] Alan A Arslan,[20] Lucia Miligi,[21] Thierry J Molina,[22] Christine F Skibola,[23] Yawei Zhang,[24] Wendy Cozen,[25] Anthony Staines,[26] Neil E Caporaso,[2] Graham G Giles,[27] Melissa C Southey,[28] Roger L Milne,[27] Lesley F Tinker,[29] Richard K Severson,[30] Mads Melbye,[31] Hans-Olov Adami,[32] Bengt Glimelius,[33] Paige M Bracci,[10] Lucia Conde,[34] Martha Glenn,[6] Karen Curtin,[6] Qing Lan,[2] Tongzhang Zheng,[35] Stephanie Weinstein,[2] Angela R Brooks-Wilson,[36] W Ryan Diver,[12] Jacqueline Clavel,[14] Paolo Vineis,[37] Elisabete Weiderpass,[38] Nikolaus Becker,[39] Paolo Boffetta,[40] Paul Brennan,[41] Lenka Foretova,[42] Marc Maynadie,[43] J Brice Weinberg,[44] Sonia Sanna,[45] Angela Gambelunghe,[46] Rebecca D Jackson,[47] Henrik Hjalgrim,[48] Kari E North,[49] James McKay,[41] Kenneth Offit,[50] Joseph Vijai,[50] Alexandra Nieters,[51] Sophia Wang,[52] Eric A Engels,[2] Stephen J Chanock[2] Nathaniel Rothman[2], James R Cerhan,[13] Susan L Slager[13], Jiali Han[16‡] and Sonja I Berndt[2‡]

[1]Service d'hématologie et oncologie, Centre Hospitalier de Versailles, Le Chesnay, Université Paris-Saclay, UVSQ, Inserm, Équipe "Exposome et Hérédité", CESP, 94805, Villejuif, France, [2]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA, [3]Department of Medical and Molecular Genetics, Indiana University School of Medicine, USA, [4]Centre for Big Data Research in Health, University of New South Wales, Sydney, New South Wales, Australia, [5]CIBER de Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain, [6]Department of Internal Medicine and Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT, USA, [7]Department of Medicine, Solna, Karolinska Institutet, Stockholm, Sweden, [8]Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA, [9]Department of Dermatology, Mayo Clinic, Rochester, MN, USA, [10]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA, [11]Cancer Genomics Research Laboratory, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Gaithersburg, MD, USA, [12]Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA, [13]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, [14]Epidemiology of childhood and adolescent cancers Group, Inserm, Center of Research in Epidemiology and Statistics Sorbonne Paris Cité (CRESS), Paris, France, [15]Department of Health Sciences, University of York, York, United Kingdom, [16]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA, [17]Department of Internal Medicine, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA, [18]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands, [19]Cancer Control Research, BC Cancer Agency, Vancouver, British Columbia, Canada, [20]Department of Obstetrics and Gynecology, New York University School of Medicine, New York, NY, USA, [21]Environmental and Occupational Epidemiology Branch, Cancer Risk Factors and Lifestyle Epidemiology Unit, Institute for Cancer Research Prevention and Clinical Network-ISPRO, Florence, Italy, [22]Department of Pathology, AP-HP, Necker Enfants malades, Université Paris Descartes, EA 7324, Sorbonne Paris Cité, France, [23]Department of Hematology and Medical Oncology, Emory University School of Medicine, Atlanta, GA, USA, [24]Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT, USA, [25]Department of Preventive Medicine, USC Keck School of Medicine, University of Southern California, Los Angeles, CA, USA, [26]School of Nursing and Human Sciences, Dublin City University, Dublin, Ireland, [27]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia, [28]Epidemiology Laboratory, Department of Pathology, University of Melbourne, Melbourne, Victoria, Australia, [29]Division of Public Health Sciences, Fred Hutchinson Cancer Research

Center, Seattle, WA, USA, [30]Department of Family Medicine and Public Health Sciences, Wayne State University, Detroit, MI, USA, [31]Department of Epidemiology Research, Division of Health Surveillance and Research, Statens Serum Institut, Copenhagen, Denmark, [32]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, [33]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden, [34]Bill Lyons Informatics Centre, UCL Cancer Institute, University College London, London, United Kingdom, [35]Department of Epidemiology, Brown University, Providence, RI, USA, [36]Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada, [37]MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, United Kingdom, [38]International Agency for Research on Cancer, World Health Organization, [39]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Baden-Württemberg, Germany, [40]The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA, [41]International Agency for Research on Cancer (IARC), Lyon, France, [41]Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic, [43]INSERM U1231, Registre des Hémopathies Malignes de Côte d'Or, University of Burgundy and Dijon University Hospital, Dijon, France, [44]Department of Medicine, Duke University and VA Medical Centers, Durham, NC, USA, [45]Department of Biomedical Science, University of Cagliari, Monserrato, Cagliari, Italy, [46]Department of Medicine, University of Perugia, Perugia, Italy, [47]Division of Endocrinology, Diabetes and Metabolism, The Ohio State University, Columbus, OH, USA, [48]Department of Haematology, National University Hospital Rigshospitalet, Copenhagen, Denmark, [49]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, [50]Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA, [51]Institute for Immunodeficiency, Center for Chronic Immunodeficiency, University Medical Center Freiburg, Freiburg, Baden-Württemberg, Germany, [52]City of Hope Beckman Research Institute, Division of Health Analytics, Department of Computational and Quantitative Medicine, City of Hope Comprehensive Cancer Center Duarte, CA, USA and [53]School of Public Health, University of Haifa, Haifa, Israel

Supplementary data

Supplementary data are available at *IJE* online.

Funding

This work was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health. Caroline Besson thanks the Monahan Foundation for providing her a research travel grant. A complete list of study-specific acknowledgements can be found in the Supplementary Material.

CLL GWAS data used in this manuscript have been deposited in the dbGaP repository with accession code phs000801.v2.p1. All other relevant data generated for this study are available upon request from the authors. All studies were approved by the respective institutional review boards as listed. These are ATBC: (NCI Special Studies Institutional Review Board); BCCA: UBC BC Cancer Agency Research Ethics Board; CPS-II: American Cancer Society; ELCCS: Northern and Yorkshire Research Ethics Committee; ENGELA: IRB00003888—Comite d' Evaluation Ethique de l'Inserm IRB # 1; EPIC: Imperial College London; EpiLymph: International Agency for Research on Cancer, HPFS: Harvard School of Public Health (HSPH) Institutional Review Board; Iowa-Mayo SPORE: University of Iowa Institutional Review Board; Italian GxE: Comitato Etico Azienda Ospedaliero Universitaria di Cagliari; Mayo Clinic Case–Control: Mayo Clinic Institutional Review Board; Mayo GEC: Mayo Clinic Institutional Review Board, #07–005788-03; MCCS: Cancer Council Victoria's Human Research Ethics Committee; MD Anderson: University of Texas MD Anderson Cancer Center Institutional Review Board; MSKCC: Memorial Sloan-Kettering Cancer Center Institutional Review Board; NCI-SEER (NCI Special Studies Institutional Review Board); NHS: Partners Human Research Committee, Brigham and Women's Hospital, NSW: NSW Cancer Council Ethics Committee; NYU-WHS: New York University School of Medicine Institutional Review Board; PLCO: (NCI Special Studies Institutional Review Board); SCALE: Scientific Ethics Committee for the Capital Region of Denmark; SCALE: Regional Ethical Review Board in Stockholm (Section 4) IRB#5; UCSF2: University of California San Francisco Committee on Human Research; UTAH: University of Utah; WHI: Fred Hutchinson Cancer Research Center; Yale: Human Investigation Committee, Yale University School of Medicine. Informed consent was obtained from all participants.

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

Conflict of interest

None declared.

References

1. Visscher PM, Wray NR, Zhang Q *et al.* 10 years of GWAS discovery: biology, function, and translation. *Am J Hum Genet* 2017;**101**:5–22.
2. Li L, Ruau DJ, Patel CJ *et al.* Disease risk factors identified through shared genetic architecture and electronic medical records. *Sci Transl Med* 2014;**6**:234ra57.

3. Mellemgaard A, Geisler CH, Storm HH. Risk of kidney cancer and other second solid malignancies in patients with chronic lymphocytic leukemia. *Eur J Haematol* 2009;**53**:218–22.
4. Adami J, Frisch M, Yuen J, Glimelius B, Melbye M. Evidence of an association between non-Hodgkin's lymphoma and skin cancer. *BMJ* 1995;**310**:1491–95.
5. Levi F, Randimbison L, Te VC, La VC. Non-Hodgkin's lymphomas, chronic lymphocytic leukaemias and skin cancers. *Br J Cancer* 1996;**74**:1847–50.
6. van den Broek EC, Liu L, Posthuma EF, Janssen-Heijnen ML, Coebergh JW, Soerjomataram I. Increased risk of chronic lymphocytic leukaemia among cancer survivors in the Netherlands: increased detection, causal factors or both? *Ann Hematol* 2014; **93**:157–62.
7. Maitra SK, Gallo H, Rowland-Payne C, Robinson D, Moller H. Second primary cancers in patients with squamous cell carcinoma of the skin. *Br J Cancer* 2005;**92**:570–71.
8. Teppo L, Pukkala E, Saxen E. Multiple cancer-an epidemiologic exercise in Finland. *J Natl Cancer Inst* 1985;**75**:207–17.
9. Robsahm TE, Karagas MR, Rees JR, Syse A. New malignancies after squamous cell carcinoma and melanomas: a population-based study from Norway. *BMC Cancer* 2014;**14**:210.
10. Dong C, Hemminki K. Second primary neoplasms among 53 159 haematolymphoproliferative malignancy patients in Sweden, 1958-1996: a search for common mechanisms. *Br J Cancer* 2001;**85**:997–1005.
11. Veraldi A, Costantini AS, Bolejack V *et al.* Immunotoxic effects of chemicals: a matrix for occupational and environmental epidemiological studies. *Am J Ind Med* 2006;**49**:1046–55.
12. Collins L, Quinn A, Stasko T. Skin cancer and immunosuppression. *Dermatol Clin* 2019;**37**:83–94.
13. Law PJ, Berndt SI, Speedy HE, C *et al.* Genome-wide association analysis implicates dysregulation of immunity genes in chronic lymphocytic leukaemia. *Nat Commun* 2017;**8**:14175.
14. Berndt SI, Camp NJ, Skibola CF *et al.* Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia. *Nat Commun* 2016;**7**:10933.
15. Berndt SI, Skibola CF, Joseph V *et al.* Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nat Genet* 2013;**45**:868–76.
16. Speedy HE, Di Bernardo MC, Sava GP, D *et al.* A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. *Nat Genet* 2014;**46**:56–60.
17. Sava GP, Speedy HE, Di Bernardo MC *et al.* Common variation at 12q24.13 (OAS3) influences chronic lymphocytic leukemia risk. *Leukemia* 2015;**29**:748–51.
18. Di Bernardo MC, Crowther-Swanepoel D, Broderick P *et al.* A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nat Genet* 2008;**40**: 1204–10.
19. Crowther-Swanepoel D, Broderick P, Di Bernardo MC *et al.* Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. *Nat Genet* 2010;**42**: 132–36.
20. Chahal HS, Lin Y, Ransohoff KJ *et al.* Genome-wide association study identifies novel susceptibility loci for cutaneous squamous cell carcinoma. *Nat Commun* 2016;**7**:12048.
21. Chahal HS, Wu W, Ransohoff KJ *et al.* Genome-wide association study identifies 14 novel risk alleles associated with basal cell carcinoma. *Nat Commun* 2016;**7**:12510.
22. Siiskonen SJ, Zhang M, Li WQ, L *et al.* A genome-wide association study of cutaneous squamous cell carcinoma among European descendants. *Cancer Epidemiol Biomarkers Prev* 2016;**25**:714–20.
23. Asgari MM, Wang W, Ioannidis NM *et al.* Identification of susceptibility loci for cutaneous squamous cell carcinoma. *J Invest Dermatol* 2016;**136**:930–37.
24. Rafnar T, Sulem P, Stacey SN *et al.* Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat Genet* 2009;**41**:221–27.
25. Stacey SN, Gudbjartsson DF, Sulem P *et al.* Common variants on 1p36 and 1q42 are associated with cutaneous basal cell carcinoma but not with melanoma or pigmentation traits. *Nat Genet* 2008;**40**:1313–18.
26. Stacey SN, Sulem P, Gudbjartsson DF *et al.* Germline sequence variants in TGM3 and RGS22 confer risk of basal cell carcinoma. *Hum Mol Genet* 2014;**23**:3045–53.
27. Stacey SN, Sulem P, Masson G *et al.* New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009; **41**:909–14.
28. Han J, Qureshi AA, Nan H *et al.* A germline variant in the interferon regulatory factor 4 gene as a novel skin cancer risk locus. *Cancer Res* 2011;**71**:1533–39.
29. Nan H, Xu M, Kraft P *et al.* Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet* 2011;**20**:3718–24.
30. Slager SL, Rabe KG, Achenbach SJ *et al.* Genome-wide association study identifies a novel susceptibility locus at 6p21.3 among familial CLL. *Blood* 2011;**117**:1911–16.
31. Conde L, Halperin E, Akers NK *et al.* Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat Genet* 2010;**42**:661–64.
32. Turner JJ, Morton LM, Linet MS *et al.* InterLymph hierarchical classification of lymphoid neoplasms for epidemiologic research based on the WHO classification (2008): update and future directions. *Blood* 2010;**116**:e90–98.
33. Hunter DJ, Colditz GA, Stampfer MJ, Rosner B, Willett WC, Speizer FE. Risk factors for basal cell carcinoma in a prospective cohort of women. *Ann Epidemiol* 1990;**1**:13–23.
34. Buniello A, MacArthur JAL, Cerezo M *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019;**47**:D1005–12.
35. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 2015;**31**:3555–57.
36. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Ser B* 1995;**57**:289–300.
37. Adami J, Gridley G, Nyrén O *et al.* Sunlight and non-Hodgkin's lymphoma: a population-based cohort study in Sweden. *Int J Cancer* 1999;**80**:641–45.

38. Kricker A, Armstrong BK, Hughes AM, for the InterLymph Consortium *et al.* Personal sun exposure and risk of non Hodgkin lymphoma: a pooled analysis from the InterLymph Consortium. *Int J Cancer* 2008;**122**:144–54.
39. Cantwell M, Hua T, Pappas J, Kipps TJ. Acquired CD40-ligand deficiency in chronic lymphocytic leukemia. *Nat Med* 1997;**3**:984–89.
40. Falchi L, Vitale C, Keating MJ *et al.* Incidence and prognostic impact of other cancers in a population of long-term survivors of chronic lymphocytic leukemia. *Ann Oncol* 2016;**27**:1100–06.
41. Han J, Kraft P, Nan H *et al.* A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008;**4**:e1000074.
42. Gathany AH, Hartge P, Davis S *et al.* Relationship between interferon regulatory factor 4 genetic polymorphisms, measures of sun sensitivity and risk for non-Hodgkin lymphoma. *Cancer Causes Control* 2009;**20**:1291–302.
43. Shukla V, Ma S, Hardy RR, Joshi SS, Lu R. A role for IRF4 in the development of CLL. *Blood* 2013;**122**:2848–55.
44. Praetorius C, Grill C, Stacey SN *et al.* A polymorphism in IRF4 affects human pigmentation through a tyrosinase-dependent MITF/TFAP2A pathway. *Cell* 2013;**155**:1022–33.
45. Travis LB, Curtis RE, Hankey BF, Fraumeni JF Jr., Second cancers in patients with chronic lymphocytic leukemia. *J Natl Cancer Inst* 1992;**84**:1422–27.
46. Tsimberidou AM, Wen S, McLaughlin P *et al.* Other malignancies in chronic lymphocytic leukemia/small lymphocytic lymphoma. *J Clin Oncol* 2009;**27**:904–10.
47. Royle JA, Baade PD, Joske D, Girschik J, Fritschi L. Second cancer incidence and cancer mortality among chronic lymphocytic leukaemia patients: a population-based study. *Br J Cancer* 2011; **105**:1076–81.
48. Morton LM, Curtis RE, Linet MS *et al.* Second malignancy risks after non-Hodgkin's lymphoma and chronic lymphocytic leukemia: differences by lymphoma subtype. *J Clin Oncol* 2010;**28**:4935–44.
49. Fletcher O, Houlston RS. Architecture of inherited susceptibility to common cancer. *Nat Rev Cancer* 2010;**10**:353–61.
50. Wang Z, Zhu B, Zhang M, P *et al.* Imputation and subset-based association analysis across different cancer types identifies multiple independent risk loci in the TERT-CLPTM1L region on chromosome 5p15.33. *Hum Mol Genet* 2014;**23**:6616–33.
51. F D. Power and predictive accuracy of polygenic risk scores. *PLoS Genet* 2013;**9**:e1003348.
52. Thrift AP, Gong J, Peters U C, GIANT Consortium *et al.* Mendelian randomization study of height and risk of colorectal cancer. *Int J Epidemiol* 2015;**44**:662–72.
53. Pingault J-B, O'Reilly PF, Schoeler T, Ploubidis GB, Rijsdijk F, Dudbridge F. Using genetic data to strengthen causal inference in observational research. *Nat Rev Genet* 2018;**19**:566–80.