Genetically predicted longer telomere length is associated with increased risk of B-cell lymphoma subtypes

Mitchell J. Machiela, National Cancer Institute
Qing Lan, National Cancer Institute
Susan L. Slager, Mayo Clinic
Roel C.H. Vermeulen, University of Utrecht
Lauren R. Teras, American Cancer Society,
Nicola J. Camp, University of Utah
James R. Cerhan, Mayo Clinic
John J. Spinelli, University of British Columbia
Sophia S. Wang, City of Hope Beckman Research Institute
Alexandra Nieters, University Medical Center Freiburg

Only first 10 authors above; see publication for full author list.

Journal Title: Human Molecular Genetics
Volume: Volume 25, Number 8
Publisher: Oxford University Press (OUP): Policy B - Oxford Open Option B | 2016-04-15, Pages 1663-1676
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1093/hmg/ddw027
Permanent URL: https://pid.emory.edu/ark:/25593/s1hhg

Final published version: http://dx.doi.org/10.1093/hmg/ddw027

Copyright information:
Published by Oxford University Press 2016.
Accessed September 16, 2017 10:44 AM EDT
Genetically predicted longer telomere length is associated with increased risk of B-cell lymphoma subtypes


†These authors contributed equally to this work.
‡These authors jointly directed this work.

Received: September 16, 2015. Revised: January 6, 2016. Accepted: January 27, 2016

Published by Oxford University Press 2016. This work is written by (a) US Government employee(s) and is in the public domain in the US.

1Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA, 2Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, 3Department of Health Sciences Research, Utrecht University, Utrecht, The Netherlands, 4Division of Medical Genetics and Epidemiology, University Medical Center Utrecht, Utrecht, The Netherlands, 5Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands, 6Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands, 7Division of Cancer Epidemiology and Genetics, National Cancer Institute, Gaithersburg, MD, USA, 8Division of Health Surveillance and Research, Statens Serum Institut, Copenhagen, Denmark, 9Department of Epidemiology, MD Anderson Cancer Center, Houston, TX, USA, 10Division of Biostatistics, School of Public Health and Imperial College London, London, UK.

‡Corresponding author.

‡‡Equal last authors.

‡‡‡Equal contributing authors.

‡‡‡‡Equal first authors.
Abstract

Evidence from a small number of studies suggests that longer telomere length measured in peripheral leukocytes is associated with an increased risk of non-Hodgkin lymphoma (NHL). However, these studies may be biased by reverse causation, confounded by unmeasured environmental exposures and might miss time points for which prospective telomere measurement would best reveal a relationship between telomere length and NHL risk. We performed an analysis of genetically inferred telomere length and NHL risk in a study of 10 102 NHL cases of the four most common B-cell histologic types and 9562 controls using a genetic risk score (GRS) comprising nine telomere length-associated single-nucleotide polymorphisms. This approach uses existing genotype data and estimates telomere length by weighing the number of telomere length-associated variant alleles an individual carries with the published change in kb of telomere length. The analysis of the telomere length GRS resulted in an association between longer telomere length and increased NHL risk [four B-cell histologic types combined; odds ratio (OR) = 1.49, 95% CI 1.22–1.82, P-value = 8.5 × 10^{-5}]. Subtype-specific analyses indicated that chronic lymphocytic leukemia or small lymphocytic lymphoma (CLL/SLL) was the principal NHL subtype contributing to this association (OR = 2.60, 95% CI 1.93–3.51, P-value = 4.0 × 10^{-5}). Significant interactions were observed across strata of sex for CLL/SLL and marginal zone...
lymphoma subtypes as well as age for the follicular lymphoma subtype. Our results indicate that a genetic background that favors longer telomere length may increase NHL risk, particularly risk of CLL/SLL, and are consistent with earlier studies relating longer telomere length with increased NHL risk.

### Introduction

Telomeres are repetitive AGGGTT nucleotide sequences that protect the ends of chromosomes from degradation and shorten during each round of cell division (1,2). Excessive telomere shortening may lead to cellular senescence, genetic instability and apoptosis (3). Excessively long telomere length and upregulated telomerase activity may result in immortalized cells with unlimited potential for growth and proliferation (4,5). The standard approach to measuring telomere length in large study populations is to measure circulating leukocyte telomere length by multiplex quantitative polymerase chain reactions (6). Family studies suggest that telomere length is highly heritable (7,8).

Recently, genome-wide association studies have identified nine common single-nucleotide polymorphisms (SNPs) that are robustly associated with circulating leukocyte telomere length (Table 1) (9–11). When extrapolating association data with qPCR-based telomere length measurements back to Southern blot data, SNPs tagging these loci could explain up to 731 bp of telomere length. This is equivalent to an approximately 20 year or greater difference in age-related telomere attrition (10).

While the total variance in telomere length explained by these variants is limited (~1%), recent studies suggest that genetic risk scores (GRSs) of these variants have utility as surrogate measures of peripheral leukocyte telomere length in deciphering associations with coronary artery disease (9), Alzheimer’s disease (12), melanoma (13) and lung cancer (14,15). A few studies have investigated the association of telomere length measured in peripheral circulating leukocytes with non-Hodgkin lymphoma (NHL) risk. A prospective study by Lan et al. (16) of a group of 107 incident male NHL cases and 107 matched controls found evidence that suggested longer relative telomere length, as measured by monochrome multiplex qPCR, might be associated with an increased risk of NHL and this association was similar across common NHL subtypes. A subsequent nested case–control study of 464 lymphoma cases and 464 matched controls from the European Prospective Investigation into Cancer (EPIC) cohort also found evidence that suggests longer telomere length is associated with an increased risk of B-cell lymphoma, particularly for diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL) (17). Further evidence from Epstein–Barr virus-infected B-lymphocytes suggests that these cells express a phenotype of progressively increasing telomere length accompanied by the accumulation of promyelocytic leukemia nuclear bodies (18). It is important to note that these studies investigated the relationship between telomere length and NHL risk, which is distinct from the studies that investigated the relationship between telomere length and NHL prognosis (19,20).

Accumulating evidence indicates that longer telomere length may be associated with increased NHL risk although existing observational studies on telomere length can be confounded by insufficient adjustment for exposures that affect both NHL risk and telomere length; for example, Epstein–Barr infection affects telomere length (18) and NHL risk (21), but was not adjusted for in prospective studies on telomere length and NHL risk (16,17).

Existing studies on telomere length and NHL risk also might not capture the biological time points most relevant for determining NHL risk since single measurements of telomere length can be influenced by changes in psychological stress (22), exercise (23) and nutrition (24). In addition, the aforementioned studies (16,17) lacked the statistical power to consider potential subtype-specific associations with relative telomere length. Here, we used telomere length–associated genetic variants to investigate an association between genetically inferred telomere length and B-cell NHL in a large pooled study population that included four major histologic subtypes of B-cell NHL. Specifically, we used associations with individual telomere length–associated variants as well as an aggregate GRS to examine whether telomere length could modify the risk of four major types of B-cell NHL. Improved understanding of how telomere length is linked to NHL risk may provide a better insight into lymphomagenesis and could serve as a biomarker for future NHL risk assessment.

### Results

Our analysis consisted of a pooled sample of 10,102 NHL cases and 9562 controls from samples of European descent (Table 2). The NHL cases included 3104 chronic lymphocytic leukemia or small lymphocytic lymphoma (CLL/SLL), 3652 DLBCL, 2521 FL and 825 marginal zone lymphoma (MZL) cases. CLL and SLL are grouped together by REAL and WHO classifications since they are both different manifestations of the same neoplastic immunophenotype in which SLL is the solid phase and CLL is the circulating phase (25,26). There were approximately equal numbers of men and women for NHL cases overall; however, there were more men with CLL/SLL and fewer men with MZL than women. The overall age range of participants was 15–94 years, with an overall median age of 64 years.

An overall excess of SNP associations with P < 0.05 was observed for the four NHL types combined (exact binomial P-value = 0.008) and for CLL/SLL and FL subtypes (exact binomial P-value = 0.001 and 0.008, respectively; Table 3). Associations were found for NHL subtypes combined (rs10936599, rs9420907 and rs11125529), CLL/SLL (rs10936599, rs7236100, rs7675989 and rs9420907), DLBCL (rs30072234) and FL (rs10936599, rs9420907 and rs755017). No evidence of association was observed between any of the nine telomere length–associated variants and MZL risk. Associations generally suggested that longer telomere length increased risk for certain NHL subtypes, with the exception of the associations between NHL subtypes combined and rs11125529, and FL and rs755017 where associations decreased risk.

Aggregate tests that compared a null model having only sex and age with an expanded model that included all nine telomere length–associated variants were also fit for the combined NHL types and for each NHL subtype (Table 3). Statistically significant aggregate associations of the nine telomere length–associated variants were observed for the combined NHL subtypes and for CLL/SLL and FL subtypes (P-values = 5.0 × 10⁻⁷, 4.7 × 10⁻⁴ and 0.01, respectively). DLBCL and MZL subtypes were not significantly associated with the nine variants in aggregate. These tests of the association with NHL overall and CLL/SLL and FL subtypes suggest that, in aggregate, one or more of the nine telomere length–associated variants are associated with overall NHL, CLL/SLL and FL risk, but give no information about the individual SNPs driving the association, the direction of the association or the effect size.

Analyses investigating the associations between a telomere length–associated GRS and NHL risk were conducted in which a
higher GRS indicates longer circulating leukocyte telomere length and a lower GRS indicates shorter telomere length (Table 3). Adjusting for sex and age, positive associations of the GRS were observed (i.e. indicating an increased risk with longer telomere length) for the four NHL types combined and for CLL/SLL (P-values = 8.5 × 10⁻⁵ and 4.0 × 10⁻¹⁰, respectively). When modeled as a decile of telomere length-associated GRS, a per decile increase in GRS was significantly associated with an increased risk of the four NHL subtypes combined (per decile OR = 1.02, 95% CI 1.01–1.03, P-value = 0.0001) and with CLL/SLL (per decile OR = 1.05, 95% CI 1.03–1.06, P-value = 2.37 × 10⁻⁵), but not with the other NHL subtypes (Fig. 1). Comparing the highest with the lowest decile of GRS, the ORs for the combined NHL subtypes and for CLL/SLL were 1.25 (95% CI 1.10–1.42, P-value = 0.0006) and 1.60 (95% CI 1.32–1.93, P-value = 1.32 × 10⁻⁴), respectively. When excluding CLL/SLL subtypes and doing a combined DLBCL, FL and MZL analysis, the odds ratio (OR) for the telomere length-associated GRS was 1.16 (95% CI 0.93–1.45, P-value = 0.18).

The Mendelian randomization analysis produced similar effect estimates as the GRS associations. In addition, the four NHL subtypes combined (likelihood-based method: OR = 1.50, 95% CI 1.22–1.85, P-value = 1.3 × 10⁻⁴) and CLL/SLL subtypes (likelihood-based method: OR = 2.68, 95% CI 1.96–3.67, P-value = 8.2 × 10⁻⁵) were significantly associated with increasing telomere length (Table 3 and Fig. 2). No significant Mendelian randomization effect estimates were observed for DLBCL, FL or MZL subtypes of NHL. Heterogeneity tests were conducted to assess if the telomere length-associated variant effects on telomere length were proportional to their effects on NHL risk. Significant evidence for heterogeneity was observed for the four NHL subtypes combined, CLL/SLL and FL subtypes (P-values = 1.9 × 10⁻⁴, 0.04 and 0.01, respectively; Table 3), with the rs755017 (RTTL1) and rs11125529 (ACY2P2) telomere length-associated variants displaying the largest departures from expectation.

In the ASSET analysis, we further explored which of the four examined subtypes of NHL are associated with telomere length-associated variants and identified CLL/SLL as the predominant B-cell NHL subtype included in the majority of the observed associations (Table 4). For example, CLL/SLL appeared in the optimal subset for six of the nine telomere length-associated SNPs in all four of the significant SNP associations. In addition, the strongest P-value association with the telomere length-associated GRS was observed for the CLL/SLL subtype.

To better characterize the association of telomere length with NHL risk, we investigated the magnitude of the association of the telomere length-associated GRS with NHL risk by sex and age groups. These associations were first investigated for all NHL subtypes combined. Men had a stronger OR for the telomere length-associated GRS (OR = 1.78, 95% CI 1.37–2.32, P-value = 2.0 × 10⁻⁵)
| SNP (ref/alternate) | All NHL subtypes | | | | | | | | | | CLL/SLL | OR | 95% CI | P-value | OR | 95% CI | P-value | DLBCL | OR | 95% CI | P-value | FL | OR | 95% CI | P-value | MZL | OR | 95% CI | P-value |
| rs10936599 (T/C)  | 1.10            | 1.05–1.15 | 5.67e–5* | 1.17 | 1.09–1.25 | 2.33e–5* | 1.04 | 0.98–1.11 | 0.2158 | 1.09 | 1.01–1.18 | 0.0250* | 1.06 | 0.94–1.20 | 0.3577 |
| rs2736100 (A/C)   | 1.03            | 0.99–1.07 | 0.1719 | 1.13 | 1.06–1.20 | 5.86e–5* | 1.01 | 0.95–1.06 | 0.8508 | 0.97 | 0.91–1.03 | 0.3081 | 1.03 | 0.92–1.14 | 0.6314 |
| rs7675998 (A/G)   | 1.04            | 0.99–1.10 | 0.0819 | 1.12 | 1.04–1.20 | 0.0023* | 1.04 | 0.98–1.12 | 0.2071 | 1.01 | 0.94–1.09 | 0.7602 | 1.00 | 0.88–1.13 | 0.9515 |
| rs9420907 (A/C)   | 1.12            | 1.05–1.18 | 0.0002* | 1.15 | 1.05–1.25 | 0.0014* | 1.02 | 0.94–1.11 | 0.6504 | 1.11 | 1.01–1.22 | 0.0243* | 1.11 | 0.96–1.28 | 0.1675 |
| rs8105767 (A/G)   | 0.99            | 0.95–1.04 | 0.7452 | 1.02 | 0.95–1.09 | 0.6254 | 1.00 | 0.94–1.06 | 0.8817 | 0.99 | 0.92–1.06 | 0.7947 | 0.97 | 0.86–1.09 | 0.6392 |
| rs755017 (A/G)    | 0.94            | 0.89–1.00 | 0.0561 | 0.97 | 0.88–1.06 | 0.5091 | 0.99 | 0.91–1.08 | 0.7729 | 0.85 | 0.77–0.94 | 0.0022* | 0.91 | 0.77–1.07 | 0.2553 |
| rs1112529 (C/A)   | 0.94            | 0.89–1.00 | 0.0352* | 0.96 | 0.88–1.05 | 0.3658 | 0.95 | 0.88–1.03 | 0.2514 | 0.97 | 0.88–1.06 | 0.4997 | 0.90 | 0.77–1.05 | 0.1803 |
| rs6772228 (A/T)   | 1.03            | 0.94–1.12 | 0.5516 | 1.05 | 0.92–1.19 | 0.5023 | 1.07 | 0.95–1.22 | 0.2729 | 1.02 | 0.88–1.18 | 0.8157 | 1.06 | 0.83–1.34 | 0.6381 |
| rs3027234 (T/C)   | 1.04            | 0.99–1.09 | 0.1076 | 1.05 | 0.97–1.12 | 0.2236 | 1.07 | 1.00–1.15 | 0.0413* | 1.01 | 0.93–1.09 | 0.8195 | 1.02 | 0.90–1.16 | 0.7738 |
| Aggregate test\(b) | –              | –          | 5.00e–7* | –     | –          | 4.71e–9* | –     | –          | 0.3327 | –     | –          | 0.0095* | –     | –          | 0.6644 |
| Genetic risk score\(b) | 1.49          | 1.22–1.82 | 8.54e–5* | 2.60 | 1.93–3.51 | 3.96e–10* | 1.28 | 0.97–1.70 | 0.0844 | 1.10 | 0.79–1.53 | 0.5640 | 1.16 | 0.69–1.95 | 0.5810 |
| MR (IVW)\(c)     | 1.49            | 1.22–1.81 | 8.10e–5* | 2.60 | 1.93–3.51 | 3.29e–10* | 1.28 | 0.97–1.70 | 0.0864 | 1.10 | 0.79–1.53 | 0.5659 | 1.17 | 0.69–1.98 | 0.5673 |
| MR (likelihood)\(d) | 1.50         | 1.22–1.85 | 1.27e–4* | 2.68 | 1.96–3.67 | 8.20e–10* | 1.29 | 0.97–1.71 | 0.0841 | 1.11 | 0.79–1.54 | 0.5572 | 1.17 | 0.69–1.99 | 0.5627 |
| Heterogeneity\(e) | –              | –          | 1.93e–4* | –     | –          | 0.0380* | –     | –          | 0.5180 | –     | –          | 0.0070* | –     | –          | 0.6170 |

Bold values designate statistically significant associations with a P-value < 0.05.

\(\text{ref.},\) short allele; alternate, long allele.

\(a\)Aggregate test is a log likelihood ratio test comparing a model having all telomere length-associated SNPs and covariates with a null model with only sex and age.

\(b\)Genetic risk score ORs refer to a 1-kb increase in telomere length.

\(c\)Inverse-variance weighted Mendelian randomization estimate for a 1-kb increase in telomere length.

\(d\)Pooled estimate for the likelihood-based Mendelian randomization method for a 1-kb increase in telomere length.

\(e\)Test for significant heterogeneity across the nine SNP instruments used in the Mendelian randomization analysis.

\(*\)Statistical significance at the P < 0.05 level.
than women did (OR = 1.17, 95% CI 0.86–1.59, P-value = 0.33); however, the interaction between telomere length-associated GRS and sex was marginally non-significant (P-value = 0.064; Fig. 3A). This sex difference in the GRS-NHL association was more pronounced for the CLL/SLL subtype, for which men had an OR of 3.47 (95% CI 2.36–5.09, P-value = 2.25 × 10^{-10}), whereas women had an OR of 1.70 (95% CI 1.05–2.76, P-value = 0.03), with a statistically significant interaction (P-value = 0.02). An interaction between the telomere length-associated GRS and sex was also detected for the MZL subtype (P-value = 0.004), for which men had a large telomere length-associated GRS effect (OR = 2.76, 95% CI 1.24–6.15, P-value = 0.01), whereas for women the effect was significantly lower (OR = 0.59, 95% CI 0.29–1.29, P-value = 0.14). Additionally, age-related effect modification was suggested across quartiles of age in which older age groups had a higher OR for the telomere length-associated GRS association with a risk of combined NHL subtypes (interaction P-value = 0.006; Fig. 3B). In the youngest age quartile (individuals aged 15–55), the OR for telomere length-associated GRS was 1.09 (95% CI 0.73–1.62, P-value = 0.69), whereas in the oldest age quartile (individuals aged 72–94) the OR was 2.23 (95% CI 1.47–3.83, P-value = 0.0002). This interaction of telomere length-associated GRS and age was also statistically significant for the FL subtype (interaction P-value = 0.007).

Discussion

Our study of genetically inferred telomere length provides a proxy measure of circulating leukocyte telomere length that suggests a positive association between telomere length and risk of four major types of B-cell NHL. While evidence suggests that this association may be present in multiple NHL subtypes, our associations with telomere length were most consistent and strongest for CLL/SLL. Subset analyses by sex and age indicate that telomere length may be particularly important for risk of some NHL subtypes for men and for older individuals although further replication is needed.

Recent genetic studies investigating the relationship between telomere length and risk of cancer have suggested that longer telomere length is associated with increased cancer risk. A prior study by our group and another independent group found evidence linking longer genetically inferred telomere length with increased risk of lung cancer (14,15) and another group has linked longer genetically predicted telomere length with increased risk of melanoma (13). We now provide evidence that longer genetically predicted telomere length may also be associated with increased NHL risk, particularly for CLL/SLL. Excessively long telomere length and upregulated telomerase activity may be important for cancer risk since longer telomeres result in greater replicative potential (4,5). Cells with a genetic ability to maintain long telomeres may result in phenotypes with increased potential for growth and proliferation which, when not adequately regulated by cellular growth mechanisms, could lead to greater carcinogenic potential.

Studies of measured telomere length are at risk for reverse causation bias by early undetected disease and can be confounded by shared environmental exposures if the mechanism
of the exposure disease relationship is not mediated through alterations in telomere length. To our knowledge, our study is the first to investigate the relationship between telomere length and NHL risk using genetic proxies of telomere length. By using a combination of nine telomere length-associated variants to infer telomere length, our study suggests a genetic background that favors longer telomere length to be associated with increased NHL risk. Our genetic proxy of telomere length estimates exposure to telomere length over an individual’s lifetime and may have advantages over traditional measures of telomere length that generally focus on one sample collection time point. Our study has the added advantage of using variants associated with leukocyte telomere length, the same progenitor cell type for NHL. Also, by using a genetic proxy for telomere length, it may be possible to isolate the effect of the genetically determined variance in TL from the effects of the variance in TL due to non-genetic factors, such as aging, oxidative damage and other relevant processes caused by environmental exposures, or unknown genetic factors. Such studies could be instrumental for identifying environmental or genetic risk factors that when appropriately targeted by a focused intervention could reduce NHL risk by its associated effect on telomere length.

The relationship between genetically predicted telomere length and NHL reported in our study for NHL overall and CLL/SLL subtypes is consistent with prior evidence from studies using telomere length measured in white blood cell DNA. For NHL overall, we detected quartile ORs (95% CI) of 1.0, 1.07 (0.98–1.16), 1.07 (0.99–1.16) and 1.17 (1.08–1.26). A prior study of measured telomere length

Figure 2. Plots of the effect of each variant on telomere length and NHL risk overall and by the subtype. The X axis (G–X association) plots the previously published linear regression β estimates for a 1 kb change in telomere length for each telomere length-associated variant (Table 1). The Y axis (G–Y association) plots the β estimate from the logistic regression model for the association of each variant with NHL risk overall and by the subtype (Table 3). Error bars around each β estimate indicate the uncertainty of effect estimates. A best fit regression line and 95% confidence interval are plotted for NHL overall and each subtype using Mendelian randomization likelihood-based estimates. P-values are from the Mendelian randomization likelihood-based method.
and NHL risk by Lan et al. (16) (107 NHL cases and 107 matched controls) found quartile ORs (95% CI) of 1.0, 1.1 (0.4–2.7), 1.8 (0.7–4.9) and 3.6 (1.4–8.9). Likewise, a study of measured telomere length and NHL risk in the EPIC cohort (464 NHL cases and 464 matched controls) found quartile ORs (95% CI) of 1.00, 1.66 (0.99–2.78), 1.80 (1.05–3.11) and 3.20 (1.71–5.98) (17). While our quartile estimates for the relationship between telomere length and NHL were of lesser magnitude, as one might expect given that a relatively small amount of variance in measured telomere length explained by the GRS, the trend was consistent with that seen in studies of measured telomere length and the 95% confidence intervals generally overlapped. Differences in sample size, sample collection time points, measurement techniques and covariates included in the model may account for the differences observed between the studies that directly measured telomere length and our approach that used genetically predicted telomere length.

Our investigation is one of the first to establish a connection between telomere length-associated genetic variants and risk of NHL. A recent CLL GWAS detected evidence for a signal at the TERC locus (rs10936599) (27). The SNP associated with CLL at this locus is the same telomere length-associated SNP reported by Codd et al. (9) (rs10936599) and was included in our telomere length-associated GRS. Our study provides replication of this association between rs10936599 and CLL/SLL risk (P-value = 2.3 × 10⁻⁵). Additionally, the TERT locus (rs10069690) has also been associated with CLL/SLL risk (27,28). Our investigation also observed an association at the TERT locus for CLL/SLL (rs2736100, P-value = 5.9 × 10⁻⁵), but the two TERT variants are weakly correlated and may tag different TERT signals (D‘ = 0.85, R² = 0.28). When performing a follow-up analysis that removes the TERC and TERT SNPs from the GRS and separately adjusts for them as covariates in the analysis, the seven SNP telomere length-associated GRS remains statistically significant (OR = 1.93, 95% CI 1.29–2.89, P-value = 0.0014), suggesting that an additional novel signal associated with CLL risk remains in the telomere length-associated GRS. To our knowledge, no prior associations with telomere length-associated variants have been reported for DLBCL, FL and MZL subtypes of NHL. Additional variation in telomere length important for NHL risk may be

### Table 4. ASSET meta-analysis of all possible NHL subtype subsets

<table>
<thead>
<tr>
<th>SNP (ref./alternate)</th>
<th>OR</th>
<th>95% CI</th>
<th>P-value</th>
<th>NHL subsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10936599 (T/C)</td>
<td>1.13</td>
<td>1.06–1.20</td>
<td>4.33 × 10⁻⁵</td>
<td>CLL/SLL, FL, MZL</td>
</tr>
<tr>
<td>rs2736100 (A/C)</td>
<td>1.12</td>
<td>1.05–1.19</td>
<td>9.76 × 10⁻⁴</td>
<td>CLL/SLL</td>
</tr>
<tr>
<td>rs7675998 (A/G)</td>
<td>1.11</td>
<td>1.01–1.22</td>
<td>0.04</td>
<td>CLL/SLL</td>
</tr>
<tr>
<td>rs9420907 (A/C)</td>
<td>1.18</td>
<td>1.08–1.28</td>
<td>1.57 × 10⁻⁴</td>
<td>CLL/SLL, MZL</td>
</tr>
<tr>
<td>rs8105767 (A/G)</td>
<td>0.96</td>
<td>0.00–Inf.</td>
<td>1.00</td>
<td>MZL</td>
</tr>
<tr>
<td>rs755017 (A/G)</td>
<td>0.89</td>
<td>0.79–1.01</td>
<td>0.06</td>
<td>FL, MZL</td>
</tr>
<tr>
<td>rs11125529 (C/A)</td>
<td>0.93</td>
<td>0.85–1.02</td>
<td>0.11</td>
<td>CLL/SLL, DLBCL, MZL</td>
</tr>
<tr>
<td>rs6772228 (A/T)</td>
<td>1.05</td>
<td>0.00–Inf.</td>
<td>1.00</td>
<td>DLBCL, MZL</td>
</tr>
<tr>
<td>rs3027234 (T/C)</td>
<td>1.05</td>
<td>0.96–1.16</td>
<td>0.30</td>
<td>CLL/SLL, DLBCL</td>
</tr>
<tr>
<td>Genetic risk score</td>
<td>2.47</td>
<td>1.82–3.35</td>
<td>7.19 × 10⁻⁹</td>
<td>CLL/SLL</td>
</tr>
</tbody>
</table>

Bold values designate statistically significant associations with a P-value < 0.05

**Figure 3.** Telomere length GRS associations by strata of sex and age quartile. P-values are for interactions between telomere length-associated GRS and sex or age group.
attributable to other genetic variants that remain to be discovered. An interesting observation from our study is that not all published telomere length-associated variants have associations with NHL subtypes that are proportional to their associations with telomere length. This is particularly true for the rs755017 (RTEL1) and rs11125529 (ACYP2) variants. Such variants may impact telomere length through mechanisms that are not important for risk of the four common NHL subtypes we studied and as such the Mendelian randomization method to estimate the effect of overall telomere length on NHL risk that uses all nine telomere length-associated variants may not be the optimal approach to accurately estimate the magnitude of the true underlying causal effect of telomere length on NHL risk. Future studies are needed to further refine effect estimates for the rs755017 (RTEL1) and rs11125529 (ACYP2) variants and leukocyte telomere length as well as explore the heterogeneity we observed in the association of the nine telomere length-associated variants and risk of the four common NHL subtypes. In addition to telomere length, other aspects of telomeres, such as maintenance of genome stability, chromosomal repair or distinct biological processes tagged by the nine telomere length-associated variants, may also be important for NHL risk.

NHL comprises several subtypes that may have different associations with telomere length. A clear association was observed with longer telomere length and increased CLL/SLL risk in our analysis. DLBCL and FL results were less clear. For DLBCL, a single SNP association was detected for one telomere length-associated variant, and a marginally statistically significant association was observed for the GRS association. The FL subtype showed evidence for three SNP associations and had a significant aggregate association test, but the GRS association was null. No evidence of single SNP associations was observed for telomere length-associated variants with MZL risk, as perhaps we were underpowered to detect associations for this subtype. Due to sample size limitations, we were also not able to examine potential associations of telomere length-associated polymorphisms with other less common B-cell or T-cell NHL subtypes.

Associations were found to differ by sex and age. If replicated, these results may be useful for screening populations where telomere length is more strongly associated with NHL risk. The association of the telomere length-associated GRS with risk of NHL was particularly strong for men, especially for the CLI/SLL and MZL subtypes. Additionally, results from individuals aged 72 years and over suggest that older individuals have elevated estimates for their GRS associations. The biological rationale for these observations is not well understood. Future studies of telomere length and NHL risk that sample men and women across a wide range of ages are needed to confirm these findings. The abundance of single SNP associations with NHL risk, the dose–response relationship by the decile of GRS and the agreement in directionality of the GRS association with prior evidence from prospective studies suggest that variation in telomere length tagged by the nine telomere length-associated variants is important for four of the most common B-cell lymphoma subtypes. Future functional studies investigating the biological mechanisms in telomere length captured by these genetic variants and their haplotypes will be instrumental in better understanding telomere biology. In addition, an improved understanding of the key molecular pathways responsible for telomere length may also be instrumental in identifying important preventive strategies for NHL as well.

**Materials and Methods**

Participants and data for this study originate from four previously published subtype-specific NHL genome-wide association studies (GWASs) which included CLI/SLL (28), DLBCL (29), FL (30) and MZL (31). Each NHL subtype-specific GWAS was a collection of cases of European descent from 22 studies of NHL, including 9 prospective cohort studies, 8 population-based case–control studies and 5 hospital or clinic-based case–control studies (Supplementary Material, Table S1). All studies obtained informed consent from the participants and were approved by the respective Institutional Review Boards.

In total, 3104 cases of CLI/SLL, 3906 cases of DLBCL, 2731 cases of FL and 825 cases of MZL were extracted from stage 1 of the respective GWAS analyses. NHL diagnoses were verified by medical and pathology reports to meet InterLymph and World Health Organization (WHO) criteria (32,33). A set of 9562 cancer-free controls was also extracted from stage 1 of the NHL GWASs. Controls included cancer-free representatives from each participating NHL study population as well as additional cancer-free individuals from the α-Tocopherol, β-Carotene Lung Cancer Prevention Study (ATBC), the American Cancer Society Cancer Prevention Study II Cohort (CPS-II) and the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO). Further details on study makeup and participants are described in the original GWAS publications (28–31).

Briefly, DNA was isolated from participants and genotyped on commercially available Illumina SNP genotyping microarrays at participating study centers. Standard quality control and filtering was applied to called genotypes to ensure well-performing samples and high-quality genotypes were reported. Genotype imputation was performed with IMPUTE2 (34) using a hybrid of the 1000 Genomes Project version 2 (February 2012 release) (35) and the Division of Cancer Epidemiology and Genetics (DCEG) European reference panels (36).

Genotypes were extracted for the nine previously identified common SNPs associated with circulating leukocyte telomere length (rs10936599, rs2736100, rs7675998, rs9420907, rs8105767, rs755017, rs11125529, rs6772228 and rs3027234). Depending on the genotyping platform of the contributing study different combinations of telomere length-associated SNPs needed to be imputed for each study. All imputed SNPs had IMPUTE2 info scores >0.75 (Supplementary Material, Table S2), indicating that the imputation for these variants had a high degree of accuracy. There was no evidence of significant departures from Hardy–Weinberg proportions for control participants (P-values >0.01).

GRSs were calculated for telomere length-associated variants. To calculate a GRS for the ith individual from the nine telomere length-associated variants we used the following formula:

\[
GRS_i = \sum_{j=1}^{9} w_j x_{ij}
\]

where \( x_{ij} \) is the number of risk alleles for the jth SNP of the ith subject (\( x_{ij} = 0, 1 \) or 2) and \( w_j \) is the weight or coefficient for the jth SNP. For the weighted coefficients, \( w_j \) of each telomere length-associated allele, we used the previously published telomere length-associated \( \beta \)-estimates scaled to \( \text{kb} \) of telomere length per length allele (Table 1) (9–11). Weighting typically results in greater specificity of the GRS by assigning more weight to variants with stronger effects.

Only participants with complete genotyping, histology (for cases) and covariate (age and sex) information were included in the analysis (10 102 NHL cases and 9562 controls). Logistic regression models calculating ORs and 95% confidence intervals (95% CI) to investigate combined and subtype-specific NHL risk were adjusted for sex and a continuous variable for age, unless...
otherwise noted. Likelihood ratio tests were used to assess statistical significance of aggregations of the nine telomere length-associated variants on NHL risk by comparing null models with fitted models of the telomere length-associated variants.

In addition to the GRS approach, we estimated the effect of telomere length on risk of the four subtypes of NHL using two different Mendelian randomization methods that use summary association for each SNP: an inverse-variance weighting method and a likelihood-based method. The methods use average summary association estimates to estimate causal effects and are described in greater detail by Burgess et al. (37). Both the inverse-variance and likelihood-based methods give similar estimates and precision to the least squares method for individual-level data, but have the advantage of using effect estimates from other studies (37). We accessed the online web tool by Burgess et al. at https://sb452.shinyapps.io/summarized/ on 16 December 2015 to calculate both Mendelian randomization estimates. Several conditions are necessary for these Mendelian randomization effect estimates to have a causal interpretation: (1) the telomere length-associated variants need to be associated with telomere length in leukocytes, (2) the telomere length-associated variants are not associated with other factors that are associated with both telomere length and NHL risk and (3) the telomere length-associated SNPs only affect NHL risk through telomere length. While these conditions are not readily testable, tests of heterogeneity can be conducted to assess if a telomere length-associated variant’s effect on telomere length is proportional to its effect on NHL risk.

An additional analysis using ASSET (38) meta-analysis software further investigates which of the four subtypes of NHL were associated with each telomere length-associated SNP as well as the overall GRS. ASSET finds the optimal subset of NHL subtypes that are associated with a SNP or GRS by performing meta-analyses that span all possible combinations of NHL subtypes and efficiently adjusts for multiple comparisons. All ASSET analyses were adjusted for age and sex.

All plotting and statistical analyses were performed on a 64-bit Windows build of R version 3.0.1 ‘Good Sport’ (39). Exact binomial tests were carried out using the binom.test function in R. All statistical tests were two-sided with \( P < 0.05 \) considered statistically significant.

**Supplementary Material**

Supplementary Material is available at HMG online.

**Acknowledgements**

Support for individual studies:

ATBC: This research was supported in part by the Intramural Research Program of the NIH and the National Cancer Institute. Additionally, this research was supported by US Public Health Service contracts N01-CN-45165, N01-RC-45035, N01-RC-37004 and HHSN262201000006C from the National Cancer Institute, Department of Health and Human Services.

BC: Canadian Institutes for Health Research (CIHR); Canadian Cancer Society and Michael Smith Foundation for Health Research.

CPS-II: The Cancer Prevention Study II (CPS-II) Nutrition Cohort is supported by the American Cancer Society. Genotyping for all CPS-II samples were supported by the Intramural Research Program of the National Institutes of Health, NCI, Division of Cancer Epidemiology and Genetics. The authors would also like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, and Cancer Registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program.

ELCS: Leukaemia & Lymphoma Research.

ENGELA: Association pour la Recherche contre le Cancer (ARC), Institut National du Cancer (INCa), Fondation de France, Fondation contre la Leucémie, Agence nationale de sécurité sanitaire de l’alimentation, de l’environnement et du travail (ANSES).

EPIC: Coordinated Action (contract #006438, SP23-CT-2005-006438); HuGeP (Human Genetics Foundation), Torino, Italy; Cancer Research UK.

Epilymph: European Commission (grant references QLK4-CT-2000-00422 and FOOD-CT-2006-023103); the Spanish Ministry of Health (grant references CIBERESP, PI11/01810, PI14/01219, RCESP C03/09, RTICESP C03/10 and RTIC RD06/0020/0095), the Marató de TV3 Foundation (grant reference 051210), the Agència de Gestió d’Ajustes Universitaris de Recerca – Generalitat de Catalunya (grant reference 2014SGR756) who had no role in the data collection, analysis or interpretation of the results; the NIH (contract NO1-CO-12400); the Compagnia di San Paolo—Programma Oncologia; the Federal Office for Radiation Protection grants StSch4261 and StSch4420; the José Carreras Leukemia Foundation grant DJCLS-K12/23, the German Federal Ministry for Education and Research (BMBF-01-EO-1303); the Health Research Board, Ireland Cancer Research Ireland; Czech Republic supported by MFF CZ – DRO (MMCI, 00298050) and RECAMO, CZ.1.05/2.1.00/ 03.0101; Fondation de France and Association de Recherche Contre le Cancer.

GEC/Mayo GWAS: National Institutes of Health (CA118444, CA148690 and CA92153). Intramural Research Program of the NIH, National Cancer Institute. Veterans Affairs Research Service. Data collection for Duke University was supported by a Leukemia & Lymphoma Society Career Development Award, the Bernstein Family Fund for Leukemia and Lymphoma Research, and the National Institutes of Health (K08CA134919), National Center for Advancing Translational Science (UL1 TR000135).

HPFS: The HPFS was supported in part by the National Institutes of Health grants CA167552, CA149445, CA098122, CA098566 and K07 CA115687 (B.M.B.). We thank the participants and staff of the Health Professionals Follow-up Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA and WY. The authors assume full responsibility for analyses and interpretation of these data.

Iowa-Mayo SPORCE: NCI Specialized Programs of Research Excellence (SPORE) in Human Cancer (P50 CA97274); National Cancer Institute (P30 CA086862 and P30 CA086862) and Henry J. Predolin Foundation.

Italian GxE: Italian Association for Cancer Research (AIRC, Investigator Grant 11855) (PC); Fondazione Banco di Sardegna 2010–2012, and Regione Autonoma della Sardegna (LR7 CRP-59812/2012) (MGE).

Mayo Clinic Case-Control: National Institutes of Health (R01 CA92153) and National Cancer Institute (P30 CA015083).

MCCS: The Melbourne Collaborative Cohort Study recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database.
MD Anderson: Institutional support to the Center for Translational and Public Health Genomics.
MSKCC: Geoffrey Beebe Cancer Research Grant, Lymphoma Foundation (LF5541); Barbara K. Lipman Lymphoma Research Fund (74419); Robert and Kate Niehaus Clinical Cancer Genetics Research Initiative (57470); U01 HG007033; ENCODE; U01 HG007033.
NCI-SEER: Intramural Research Program of the National Cancer Institute, National Institutes of Health and Public Health Service (N01-PC-65064, N01-PC-67008, N01-PC-67009, N01-PC-67010 and N02-PC-71105).
NHS: The NHS was supported in part by National Institutes of Health grants CA186107, CA87969, CA49449, CA149445, CA098122, CA098566 and K07 CA115887. We thank the participants and staff of the Nurses’ Health Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA and WY. The authors assume full responsibility for analyses and interpretation of these data.
NSW: NSW was supported by grants from the Australian National Health and Medical Research Council (ID990920), the Cancer Council NSW and the University of Sydney Faculty of Medicine.
NYU-WHS: National Cancer Institute (R01 CA098661 and P30 CA160887); National Institute of Environmental Health Sciences (ES000269).
PLCO: This research was supported by the Intramural Research Program of the National Cancer Institute and by contracts from the Division of Cancer Prevention, National Cancer Institute, NIH, DHHS.
SCALE: Swedish Cancer Society (2009/659), Stockholm County Council (20110209) and the Strategic Research Program in Epidemiology at Karolinska Institute. Swedish Cancer Society grant (02 6661). National Institutes of Health (5R01 CA69669-02); Plan Denmark.
UCSF2: The UCSF studies were supported by the NCI, National Institutes of Health, CA104628 and CA154643. The collection of cancer incidence data used in this study was supported by the California Department of Health Services as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute’s Surveillance, Epidemiology, and End Results Program under contract HHSN26120100014C awarded to the Cancer Prevention Institute of California, contract HHSN262120100035C awarded to the University of Southern California, and contract HHSN26121000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention’s National Program of Cancer Registries, under agreement #1U58 DP000807-01 awarded to the Public Health Institute. The ideas and opinions expressed here are those of the authors, and endorsement by the State of California, the California Department of Health Services, the National Cancer Institute or the Centers for Disease Control and Prevention or their contractors and subcontractors is not intended nor should be inferred.
UTAH: National Institutes of Health CA134674. Partial support for data collection at the Utah site was made possible by the Utah Population Database (UPDB) and the Utah Cancer Registry (UCR). Partial support for all datasets within the UPDB is provided by the Huntsman Cancer Institute (HCI) and the HCI Cancer Center Support grant P30 CA42014. The UCR is supported in part by NIH contract HHSN261201000026C from the National Cancer Institute SEER Program with additional support from the Utah State Department of Health and the University of Utah. Partial support for data collection in Sheffield, UK, was made possible by funds from Yorkshire Cancer Research. We thank the NCRI Haematology Clinical Studies Group, colleagues in the North Trent Cancer Network the North Trent Haematology-oncology Database.
WHI: WHI investigators are: Program Office—(National Heart, Lung, and Blood Institute, Bethesda, Maryland) Jacques Rossouw, Shari Ludlam, Dale Burwen, Joan McGowan, Leslie Ford and Nancy Geller; Clinical Coordinating Center—(Fred Hutchinson Cancer Research Center, Seattle, WA) Garnet Anderson, Ross Prentice, Andrea LaCroix and Charles Kooperberg; Investigators and Academic Centers—(Brigham and Women’s Hospital, Harvard Medical School, Boston, MA) JoAnn E. Manson; (MedStar Health Research Institute/Harvard University, Washington, DC) Barbara V. Howard; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Arizona, Tucson/Phoenix, AZ) Cynthia A. Thomson; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Iowa, Iowa City/Davenport, IA) Robert Wallace; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker; Women’s Health Initiative Memory Study—(Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker. The WHI Program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C and HHSN268201100004C. YALE: National Cancer Institute (CA62006) and National Cancer Institute (CA169523).

Conflict of Interest statement. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute of Health. The authors report no conflicts of interest.

References


