

Genome-Wide Association Study of Relative Telomere Length

Jennifer Prescott^{1,2}, Peter Kraft², Daniel I. Chasman³, Sharon A. Savage⁴, Lisa Mirabello⁴, Sonja I. Berndt⁴, Joel L. Weissfeld⁵, Jiali Han^{1,2,6}, Richard B. Hayes⁷, Stephen J. Chanock⁴, David J. Hunter^{1,2,8}, Immaculata De Vivo^{1,2*}

1 Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **2** Program in Molecular and Genetic Epidemiology, Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America, **3** Donald W. Reynolds Center for Cardiovascular Research, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **4** Division of Cancer Epidemiology and Genetics, Department of Health and Human Services, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, United States of America, **5** University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania, United States of America, **6** Department of Dermatology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **7** Division of Epidemiology, New York University Medical Center, New York, New York, United States of America, **8** Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, United States of America

Abstract

Telomere function is essential to maintaining the physical integrity of linear chromosomes and healthy human aging. The probability of forming proper telomere structures depends on the length of the telomeric DNA tract. We attempted to identify common genetic variants associated with log relative telomere length using genome-wide genotyping data on 3,554 individuals from the Nurses' Health Study and the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial that took part in the National Cancer Institute Cancer Genetic Markers of Susceptibility initiative for breast and prostate cancer. After genotyping 64 independent SNPs selected for replication in additional Nurses' Health Study and Women's Genome Health Study participants, we did not identify genome-wide significant loci; however, we replicated the inverse association of log relative telomere length with the minor allele variant [C] of rs16847897 at the *TERC* locus (per allele $\beta = -0.03$, $P = 0.003$) identified by a previous genome-wide association study. We did not find evidence for an association with variants at the *OBFC1* locus or other loci reported to be associated with telomere length. With this sample size we had >80% power to detect β estimates as small as ± 0.10 for SNPs with minor allele frequencies of ≥ 0.15 at genome-wide significance. However, power is greatly reduced for β estimates smaller than ± 0.10 , such as those for variants at the *TERC* locus. In general, common genetic variants associated with telomere length homeostasis have been difficult to detect. Potential biological and technical issues are discussed.

Citation: Prescott J, Kraft P, Chasman DI, Savage SA, Mirabello L, et al. (2011) Genome-Wide Association Study of Relative Telomere Length. PLoS ONE 6(5): e19635. doi:10.1371/journal.pone.0019635

Editor: Marie-Pierre Dubé, Université de Montreal, Canada

Received: November 30, 2010; **Accepted:** April 11, 2011; **Published:** May 10, 2011

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Funding: The NHS was supported by National Institutes of Health (NIH) grants CA49449, CA87969, CA065725, CA133914, and U01-CA98233. Dr. Prescott was supported by NIH training grant T32 CA 09001. The PLCO and work of Drs. Mirabello, Berndt, Hayes, Chanock, and Savage was supported by the Intramural Research Program of the National Institutes of Health and the National Cancer Institute. Primary financial support for the WGH genome-wide scans was provided by Amgen, with additional financial support provided by the National Heart, Lung, and Blood Institute (HL043851 and HL69757), National Cancer Institute (CA047988), the Donald W. Reynolds Foundation, and the Leducq Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: devivo@channing.harvard.edu

Introduction

Two major challenges are encountered in maintaining the physical integrity of linear chromosomes. First, cells must prevent the recognition of chromosome ends as double-stranded DNA breaks [1]. Second, telomeric ends are lost with each cell division as a result of the end replication problem [2]. On average, human telomeres lose 50 to 100 bp per mitotic division, which limits the replicative capacity of a cell [3]. To overcome these issues, eukaryotic cells evolved the telomere maintenance system, the almost universal mechanism used to protect chromosome ends [4].

Telomeres are complex dynamic nucleoprotein structures that consist of a long stretch of hexameric (TTAGGG)_n DNA repeats, a single-stranded G-rich 3' overhang, the telomerase enzyme complex, six telomere-specific proteins, and other telomere-

associated proteins including those of the DNA repair machinery [5]. In normal mitotic cells, telomeres switch between capped and uncapped states [6]. Telomere-associated proteins assist the single-stranded G-rich 3' overhang in invading the double-stranded telomeric tract to form a large terminal loop. This conformation contributes to the functional capped structure of the telomeric end, inhibiting the DNA damage response [7]. Telomerase, a highly conserved enzyme consisting of protein [telomerase reverse transcriptase (*TERT*), dyskerin (*DKC1*)] and RNA [telomerase RNA component (*TERC*)] components [8], appends chromosome ends with hexameric repeats to restore telomere length [9] and plays a role in stabilizing telomeres in the capped state [6].

The likelihood of a telomere existing in an uncapped vs. a capped state is dependent on telomere length. Longer telomeres are more likely to form the protective capped structure than

shorter telomeres [6]. In germ cells, strong telomerase expression maintains long telomere lengths [10]. Telomerase activity has been detected at lower levels in fibroblasts, peripheral blood leukocytes, skin, hair follicles, intestinal crypts, and endometrium. In these tissues, expression levels are not sufficient to prevent replication-associated telomere attrition [11–15]. As a result, telomere length declines with age [16–21]. As telomeres shorten and become dysfunctional, the uncapped chromosomal ends undergo nucleolytic decay, chromosomal end-to-end fusions, and atypical recombination triggering senescence or apoptosis [22].

Diseases characterized by telomere dysfunction highlight the importance of telomere maintenance for healthy human aging. Patients with dyskeratosis congenita, a rare inherited bone marrow failure and cancer predisposition syndrome, have very short germline telomeres (<1st percentile for age), and approximately one-half have a mutation in a known telomere biology gene [23]. Patients with the shortest telomeres exhibit the most severe phenotypes [24]. Telomere attrition, genomic instability, and premature senescence are also features of Hutchinson-Gilford progeria syndrome and Werner syndrome [25]. Shorter leukocyte telomere length has been implicated in a number of common aging-related diseases such as cancer [26–30] and cardiovascular disease [31–36] as well as increased mortality [37–41].

Despite the essential role telomeres play in the maintenance of normal cellular function, little is known about the common genetic determinants of telomere length. Heritability estimates from family and twin studies for telomere length range from 36% to 86% [17,40–45]. Telomere lengths in different tissues are significantly correlated, with far less variation between tissues from the same individual compared to variation between individuals of a particular tissue type [18,20,21,46–48]. Quantitative trait linkage analyses identified significant linkage to chromosomes 12p11.2 [43,49], and 14q23.2 [44], but these loci have not been replicated in independent studies [44,50–52]. Common genetic variants in known telomere maintenance genes do not exhibit strong influences on telomere length [45,53–55]. Genome-wide association studies (GWAS) have identified and replicated associations at the *TERC* locus, which account for no more than 1% of variation in telomere length [51,52]. Variants at the oligonucleotide/oligosaccharide-binding fold containing 1 (*OBFC1*) gene locus, which codes for a protein that interacts with telomere-specific proteins and is implicated in telomere length regulation [56], were also recently identified using a meta-analysis of GWAS cohorts [52]. To identify additional variants associated with telomere length, we conducted a GWAS using data from two large cohorts that took part in the National Cancer Institute Cancer Genetic Markers of Susceptibility (CGEMS) initiative for breast and prostate cancer.

Results

We used a two-stage GWAS to identify common genetic variants associated with log relative telomere length. Age, smoking, and mean log relative telomere length values of each population are shown in **Table 1**. The GWAS discovery set included genotypes for 519,076 SNPs on a total of 3554 individuals from the Nurses' Health Study (NHS) and the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. After adjusting for the top principal components of genetic variation from each cohort, the distribution of the observed P values does not suggest any inflation in Type 1 error rates due to population stratification or other sources of bias (**Figure 1**).

None of the P values observed in the first stage of our GWAS reached the genome-wide significance level of $P < 5.0 \times 10^{-8}$ [57].

The SNP most significantly associated with log relative telomere length in this discovery set was rs1978423, located upstream of the pyroglutamyl-peptidase I (*PGPEP1*) and LSM4 homolog, U6 small nuclear RNA associated (*LSM4*) genes on chromosome 19 ($P = 3.8 \times 10^{-7}$). We selected the 64 independent SNPs most significantly associated with log relative telomere length to genotype in additional samples in the NHS (N = 1345) and the Women's Genome Health Study (WGHS) (N = 1115) replication populations. Estimates for all SNPs exhibited moderate to high between-study heterogeneity ($I^2 \geq 0.50$; **Table S1**). rs1978423 was not associated with log relative telomere length in either replication population ($P \geq 0.12$). Two SNPs, rs4073054 at the alipoprotein A-II (*APOA2*)/nuclear receptor subfamily 1, group I, member 3 (*NRII3*) locus on chromosome 1 and rs975504 near dehydrogenase/reductase (SDR family) member 12 (*DHRS12*) and a predicted non-coding RNA on chromosome 13, displayed estimates similar to the associations from the initial GWAS population with P values <0.05 in the NHS replication population, but not within the WGHS (**Table S1**; joint P values of GWAS and replication populations are 1.7×10^{-4} and 9.5×10^{-4} , respectively). rs7210405, ~24 kb downstream of the cytohesin 1 (*CYTH1*) gene on chromosome 17, displayed the smallest P value (0.01) within the WGHS population with a similar estimate. However, this SNP was not associated with log relative telomere length in the NHS replication population (joint $P = 1.4 \times 10^{-4}$).

We examined SNPs at the *TERC* and *OBFC1* loci, which were recently identified by published GWAS of telomere length [51,52]. Within our discovery GWAS and WGHS replication populations, we examined associations with genotyped and imputed SNPs (rs12696304, rs10936599, rs3772190, rs16847897) at the *TERC* locus reported as significantly associated with telomere length in prior GWAS. In our study populations, we found evidence for associations between log relative telomere length and SNPs at this locus, with the most significant association observed for the minor allele [C] of rs16847897 (per allele $\beta = -0.03$, joint $P = 0.003$). Meta-analysis of our results for rs16847897 compared with study-specific estimates from published GWAS generated a $P = 1.6 \times 10^{-13}$ and displayed virtually no between-study heterogeneity (**Table 2**). Whereas Levy et al. [52] reported genome-wide significant associations with SNPs in the *OBFC1* gene, we did not find evidence for associations between genotyped or imputed SNPs at the *OBFC1* locus and log relative telomere length in our study (joint $P \geq 0.12$). Meta-analysis of our results for SNPs at the *OBFC1* locus with those reported by Levy et al. demonstrated high between-study variability (**Table 2**). As we previously reported, we did not find associations between log relative telomere length and common genetic variants at the *TERT-CLPTM1L*, *BICD1*, and *DDX11* loci [55].

Discussion

Despite the high heritability of telomere length [17,40–45], common variants associated with large effects have remained elusive. We were unable to identify novel variants associated with log relative telomere length with a high degree of confidence among a total of 6,014 participants. The two most promising SNPs identified by our study were rs4073054 (NHS GWAS, $\beta = -0.04$, $P = 1.6 \times 10^{-4}$; PLCO GWAS, $\beta = -0.04$, $P = 0.041$; NHS replication, $\beta = -0.03$, $P = 0.01$) and rs7210405 (NHS GWAS, $\beta = 0.03$, $P = 0.006$; PLCO GWAS, $\beta = 0.05$, $P = 0.005$; WGHS replication, $\beta = 0.03$, $P = 0.01$), as 3 of our 4 study populations displayed evidence of an association with log relative telomere length. We had >80% power to detect β estimates as small as

Table 1. Characteristics of GWAS and replication populations.

	NHS GWAS	PLCO GWAS	NHS Replication	WHS Replication
No. subjects	1944	1610	1345	1115
Mean age (range)	58.5 (43–70)	63.9 (55–74)	58.4 (43–69)	56.0 (45–82)
Smoking status (%)				
Never	46.2	39.9	46.2	49.7
Former	40.5	50.1	41.9	36.7
Current	13.3	9.9	12.0	13.6
Pack-years (%)*				
1–9	32.3	21.1	33.8	43.8
10–29	34.9	36.5	34.9	39.8
30+	32.7	42.4	31.2	16.4
Mean log(RTL) (SD)				
Cases	2.78 (0.35)	2.59 (0.55)	2.31 (0.35)	2.32 (0.26)
Controls	2.79 (0.35)	2.58 (0.58)	2.30 (0.34)	2.31 (0.25)

*among participants who have ever smoked.

doi:10.1371/journal.pone.0019635.t001

± 0.10 for SNPs with minor allele frequencies (MAF) of ≥ 0.15 in our discovery GWAS population ($N = 3554$). However, published associations with relative telomere length suggest effect estimates may be quite a bit smaller than ± 0.10 . Our study was underpowered to detect genome-wide significant associations between log relative telomere length and SNPs of much weaker effect. For example, adequate power to detect genome-wide significant associations between SNPs at loci such as *TERC* that have a MAF ~ 0.25 , $\beta \sim 0.03$, and log relative telomere length as measured in our study would require a sample size of $\sim 25,000$ unrelated individuals. Even to reach the suggestive threshold of $\alpha = 1 \times 10^{-5}$, to detect an association with such a SNP would require a sample size of 18,000 individuals, roughly 5 x greater than the size of our initial GWAS population. Discovery of additional loci with small effect sizes will be feasible once meta-analyses of existing and future GWAS of telomere length are conducted.

To date, only the *TERC* locus has been identified and replicated in multiple independent populations for its association with telomere length [51,52]. Although SNPs within the *TERC* locus did not reach genome-wide significance in the discovery GWAS population ($N = 2917$) of Codd et al. [51], rs16847897 fell marginally outside of their pragmatic significance threshold ($P < 1 \times 10^{-5}$) for replication. Due to the biological significance of the locus, the authors examined and confirmed associations between rs16847897 and a second SNP at the locus, rs12696304, with telomere length in 3 replication populations. Levy et al. found additional evidence for associations between SNPs at the *TERC* locus and telomere length [52]. We found evidence for associations with SNPs at the *TERC* locus, observing the strongest association between the minor allele [C] of rs16847897 and log relative telomere length (joint $P = 0.003$) with an effect size and direction consistent with that of Codd et al. Similar to Codd et al., we did not find associations between log relative telomere length and SNPs at loci previously reported to be associated with telomere length [43,49,50,58], including the *OBFC1* locus recently identified by a similarly sized GWAS of telomere length [52].

If yeast genetic mapping studies are any indication [59,60], hundreds of genes regulate human telomere length homeostasis. Given that known telomere maintenance genes are highly

conserved [61] and mutations have deleterious consequences [23], most functional common variants ($> 5\%$) will likely have small effects. Additionally, recent studies demonstrate that while telomere length of offspring is significantly correlated with maternal [62,63] and paternal [41,64,65] telomere lengths, the paternal contribution appears to be much stronger [41,64,65] and positive correlations have been observed between paternal age at birth and offspring telomere length [41,66–68]. An influence of imprinted genes in regulating telomere length could account for these observations. By not controlling for parental origin of alleles at imprinted loci, SNP effects, which are likely to be small to begin with, would be underestimated and the power to detect associations reduced [69]. For many GWAS populations, including ours, determining the parental origin of alleles is not feasible.

Evidence suggests cis-acting factors in subtelomeric regions regulate telomere length. While chromosome-specific telomeres tend to be similar between different individuals [48,70–73], allele specific lengths are inherited [74] and may differ by more than 6 kb at the same telomere [75]. A substantial amount of genetic diversity is provided by subtelomeric regions, which are riddled with repetitive DNA, segmental duplications, and large copy number polymorphisms including variable copy numbers of functional genes. Knowledge of subterminal sequences is still incomplete and the complexity within these regions has complicated sequencing efforts [76] resulting in low coverage of such regions on commercially available genome-wide genotyping arrays [77].

Technical challenges further complicate the search for genetic variants of telomere length. While the quantitative PCR-based assay is the most economical, high-throughput method for telomere length measurements in large epidemiologic studies [78], values are relative representations of the average telomere length. The telomere restriction fragment (TRF) assay provides absolute values for the average, shortest, and longest telomere lengths. However, TRF measurements include variable amounts of the subtelomeric region, are time and labor intensive, and require much more genomic DNA than the PCR-based assay [79]. The single telomere length analysis (STELA) assay is currently the most sensitive assay,

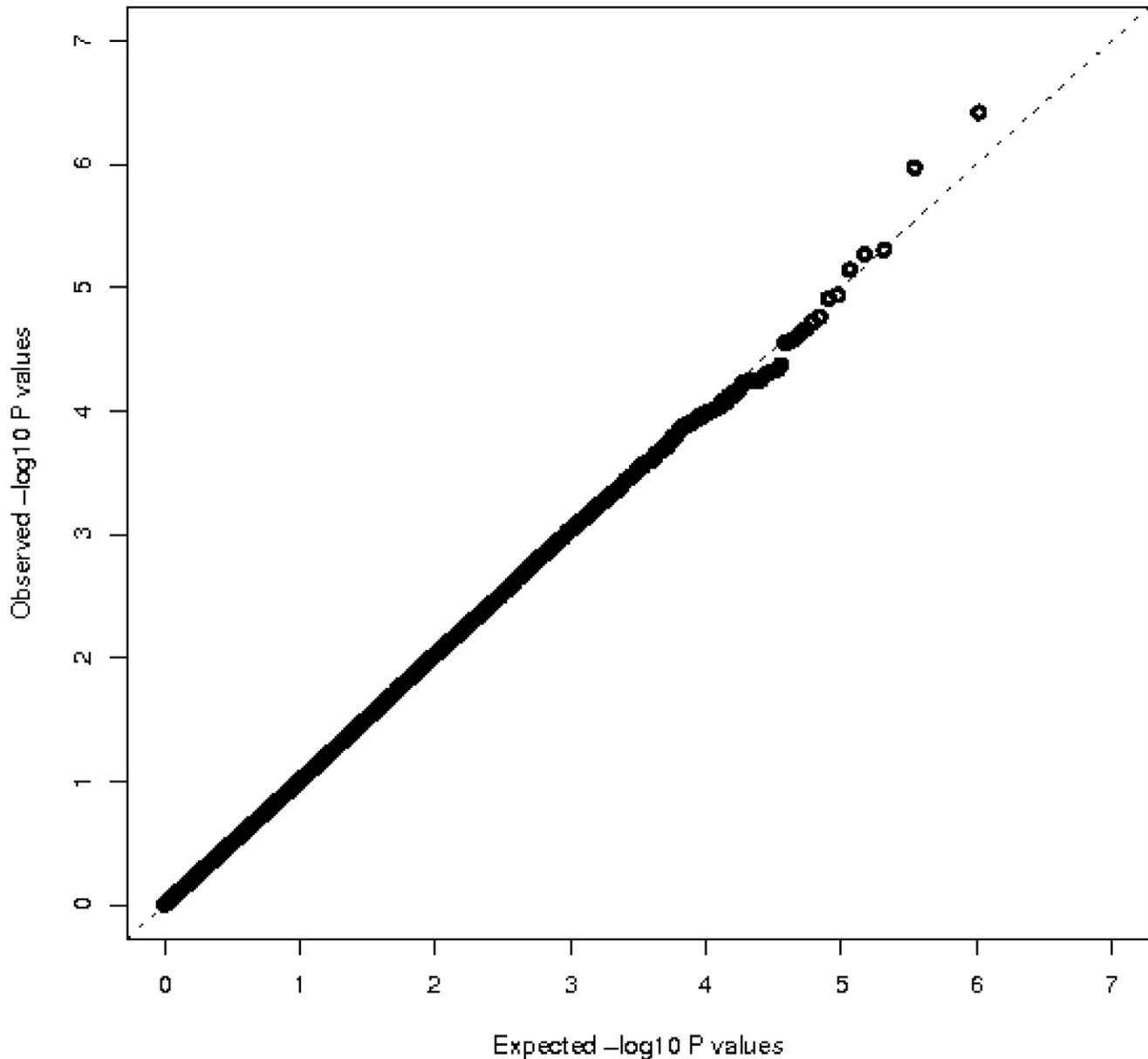


Figure 1. Log Quantile-Quantile (Q-Q) P value plot. The observed $-\log_{10}$ P values (Y-axis) of 519,076 SNPs from the pooled NHS and PLCO GWAS data set adjusted for the principal components of genetic variation plotted against the expected $-\log_{10}$ quantile (X-axis). The dashed line represents imputed P values.

doi:10.1371/journal.pone.0019635.g001

developed with the potential to provide allele-specific telomere lengths [75]. To date, telomere and allele-specific STELA assays have only been developed for a very small fraction of chromosome ends due to incomplete knowledge of subtelomeric regions [73].

The inability to capture the complexity of telomere length regulation with a single measurement is a major obstacle to GWAS efforts. Even the well established decline in telomere length with increasing age is not a simple linear relationship. Relatively fast rates of telomere attrition are observed during childhood and adolescence. Attrition rates slow creating a plateau from ~ 30 to 50 years of age followed by another decline in telomere length from about age 50 to 80 [19,64,80,81]. It is unknown whether temporally regulated genes and/or environmental exposures are responsible for the change in attrition rates.

The age range of individuals in our GWAS population (**Table 1**) and those of previously published studies [50–52] span these different age periods potentially diluting telomere length associations with variants that influence temporally regulated genes.

Oxidative stress is thought to accelerate telomere attrition as a result of damage to telomeric DNA, which is less efficiently repaired [82]. Systemic exposures that contribute to oxidative stress such as smoking [62,83,84] and obesity [54,84–87] have been associated with shorter telomeres. Whereas, healthy lifestyle choices promote a more stable telomere length [81,86,88], possibly by increasing telomerase activity in cells [89]. Evidence suggests environmental factors may take on a more prominent role as determinants of telomere length with increasing age [65,90], but the relationship with these

Table 2. Relative telomere length associations with SNPs at loci identified by published GWAS.

Locus	SNP	WT/VT	CGEMS GWAS*			WGHS replication†			Meta-analysis‡				
			MAF	β (s.e.)	P value	MAF	β (s.e.)	P value	β (s.e.)	P value	Q	I ²	P _{heterogeneity}
TERC	rs12696304	C/G	0.27	-0.001 (0.01)	9.1E-01	0.26	-0.03 (0.01)	3.4E-02	-0.03 (0.004)	1.5E-14	15.35	0.54	3.2E-02
TERC	rs10936599	C/T	0.25	-0.02 (0.01)	1.9E-01	0.24	-0.03 (0.01)	3.2E-02	-0.03 (0.008)	2.8E-04	5.39	0.63	6.7E-02
TERC	rs3772190	G/A	0.25	-0.02 (0.01)	1.9E-01	0.24	-0.03 (0.01)	3.8E-02	-0.03 (0.008)	3.3E-04	5.40	0.63	6.7E-02
TERC	rs16847897	G/C	0.28	-0.03 (0.01)	2.1E-03	0.28	-0.01 (0.01)	2.7E-01	-0.03 (0.004)	1.6E-13	5.83	0.00	4.4E-01
OBFC1	rs11191839	G/A	0.35	-0.001 (0.01)	9.2E-01	0.36	-0.006 (0.01)	6.0E-01	-0.003 (0.008)	0.66	0.10	0.00	7.5E-01
OBFC1	rs2902638	T/C	0.25	0.002 (0.01)	8.6E-01	0.25	-0.03 (0.01)	2.6E-02	-0.01 (0.009)	0.16	3.03	0.67	8.2E-02
OBFC1	rs10748858	T/G	0.41	-0.004 (0.01)	7.1E-01	0.41	0.006 (0.01)	6.1E-01	0.0005 (0.008)	0.95	0.40	0.00	5.3E-01
OBFC1	rs2067832	T/C	0.50	-0.002 (0.01)	8.4E-01	N/A							
OBFC1	rs9325507	T/C	0.50	-0.002 (0.01)	8.7E-01	0.50	0.006 (0.01)	6.1E-01	0.02 (0.007)	1.1E-02	19.45	0.74	1.6E-03
OBFC1	rs1980653	G/A	0.50	-0.003 (0.01)	8.1E-01	0.50	0.006 (0.01)	6.1E-01	0.008 (0.007)	2.3E-01	25.17	0.80	1.0E-04
OBFC1	rs4918069	T/G	0.27	0.004 (0.01)	7.4E-01	0.29	-0.03 (0.01)	4.3E-02	-0.01 (0.008)	0.26	2.92	0.66	8.8E-02
OBFC1	rs2984132	T/C	0.40	-0.002 (0.01)	8.5E-01	0.40	-0.009 (0.01)	4.2E-01	-0.005 (0.008)	0.49	0.21	0.00	6.5E-01
OBFC1	rs2487999	C/T	0.10	0.002 (0.02)	9.2E-01	0.10	0.04 (0.02)	3.4E-02	0.04 (0.01)	1.1E-04	19.94	0.75	1.3E-03
OBFC1	rs3752949	A/G	0.23	-0.0006 (0.01)	9.6E-01	0.22	0.03 (0.01)	4.4E-02	0.01 (0.009)	0.18	2.24	0.55	1.3E-01
OBFC1	rs11191865	G/A	0.50	0.001 (0.01)	9.1E-01	0.49	-0.006 (0.01)	6.1E-01	-0.002 (0.008)	0.80	0.21	0.00	6.5E-01
OBFC1	rs1265164	G/A	0.14	-0.008 (0.02)	6.0E-01	0.07	0.02 (0.02)	4.3E-01	-0.0003 (0.01)	0.98	0.88	0.00	3.5E-01
OBFC1	rs9419958	C/T	0.14	-0.006 (0.02)	7.1E-01	0.14	0.03 (0.02)	1.0E-01	0.03 (0.01)	5.3E-04	23.72	0.79	2.0E-04
OBFC1	rs9420907	A/C	0.14	-0.004 (0.02)	7.8E-01	0.14	0.03 (0.02)	1.0E-01	0.03 (0.01)	4.3E-04	23.27	0.79	3.0E-04
OBFC1	rs4387287	C/A	0.15	-0.005 (0.02)	7.3E-01	0.14	0.03 (0.02)	6.3E-02	0.04 (0.01)	1.1E-05	26.55	0.77	2.0E-04
OBFC1	rs17813699	G/A	0.10	-0.005 (0.02)	7.5E-01	0.12	-0.02 (0.02)	4.5E-01	-0.01 (0.01)	0.48	0.18	0.00	6.7E-01
OBFC1	rs1342214	C/T	0.40	0.003 (0.01)	7.9E-01	0.41	-0.007 (0.01)	5.2E-01	-0.002 (0.008)	0.79	0.42	0.00	5.2E-01

*β estimates and P values derived from linear regression adjusted for age at blood collection, gender, an age-gender interaction term, smoking status, pack-year categories of smoking, disease status, and principal components of genetic variation.
 †β estimates and P values derived from linear regression adjusted for age at blood collection, smoking status, pack-year categories of smoking, disease status, and principal components of genetic variation; N/A = SNP could not be imputed.
 ‡Combined effect sizes and P values are calculated using a fixed-effect meta-analysis. Estimates in bold are based upon meta-analysis of the current study populations along with those published by Codd et al. and Levy et al. Study-specific values from the published studies were used where available.
 doi:10.1371/journal.pone.0019635.t002

exposures are not yet clearly defined as many studies fail to find significant correlations between telomere length and smoking [33,40,44,54,91], obesity [92,93], and/or physical activity [92,94]. Since pack-years of smoking were significantly correlated with relative telomere length among men of the PLCO GWAS population, we adjusted for smoking characteristics in all of our regression models. Although relative telomere length in our prospectively collected blood samples was not associated with breast [93] or prostate cancer [88], we adjusted

for case status to control for any potential effects of pre-clinical disease on telomere length.

In summary, improved sequence maps and technical capabilities are necessary to increase success in identifying and validating common genetic variants associated with telomere length homeostasis. Efforts will likely require meta-analyses of existing and future telomere length GWAS to increase power to detect common variants of small effects while stratifying by age groups defined by attrition rate and controlling for environmental factors.

Materials and Methods

Ethics Statement

The NHS and WGHS study protocols were approved by the Committee on Use of Human Subjects of the Brigham and Women's Hospital, Boston, MA. Institutional review boards at the U.S. National Cancer Institute and the 10 screening centers approved the PLCO protocol.

NHS breast cancer GWAS population

The NHS is a prospective cohort study of 121,700 female registered nurses in 11 states in the United States who were 30–55 years of age at enrollment. In 1976 and biennially thereafter, detailed information from participants was collected by self-administered questionnaires. Participants in this study were selected for a nested case-control study of telomere length and postmenopausal breast cancer risk from the subcohort of 32,826 women who donated a blood sample in 1989–90 [93]. Eligible cases consisted of postmenopausal women of European ancestry with pathologically confirmed incident invasive breast cancer diagnosed anytime after blood collection up to June 1, 2004 with no prior diagnosis of cancer. Controls were randomly selected postmenopausal women free of cancer and matched to cases according to age, blood collection, and ethnicity. Completion of the questionnaire and submission of the blood sample was considered to imply informed consent. No significant difference was observed for log relative telomere length between 1,122 cases and 1,147 controls and no significant association with postmenopausal breast cancer risk [93]; therefore, we included telomere length data from both breast cancer cases and controls for the current study.

PLCO prostate cancer GWAS population

The PLCO Cancer Screening Trial is an ongoing randomized trial with 154,942 persons aged 55 to 74 enrolled between September 1993 and July 2001 from 10 screening centers nationwide [95]. Detailed questionnaire data was collected from all subjects at baseline. Participants provided blood and tissue samples for etiologic studies of cancer [96] and all participants provided written informed consent. Participants in this study were male subjects selected for a nested case-control study of telomere length, prostate cancer risk, and life-style variables [88]. Eligible cases and controls consisted of non-Hispanic white men aged 55 to 74 who had been screened for prostate cancer (PSA test) prior to October 1, 2003, completed a baseline questionnaire of cancer risk factors, provided a blood sample 1 month to 3 years prior to prostate cancer diagnosis for cases, and did not have a personal history of cancer prior to study entry. All cases had pathologically confirmed incident aggressive prostate cancer and a Gleason score of ≥ 7 . No significant difference was observed for relative telomere length between 616 cases and 1,061 matched controls and no significant association with aggressive prostate cancer risk [88]; therefore, we included telomere length data from both prostate cancer cases and controls for the current study.

Replication populations

NHS participants in the replication phase of this study were selected for a nested case-control study of telomere length and skin cancer risk from the blood subcohort [97]. Eligible cases were women of European ancestry with skin cancer diagnosed anytime after blood collection up to June 1, 2000 with no prior diagnosis of skin cancer. A common control series was randomly selected from participants who gave a blood sample and were free of diagnosed skin cancer up to and including the questionnaire cycle in which

the case was diagnosed. Controls were matched to cases by age and ethnicity. The nested case-control study consisted of 218 melanoma cases, 285 cases with squamous cell carcinoma, 300 cases with basal cell carcinoma, and 870 matched controls. The study protocol was approved by the Committee on Use of Human Subjects of the Brigham and Women's Hospital, Boston, MA.

The Women's Genome Health Study (WGHS) is a prospective cohort of female North American health care professionals representing participants in the Women's Health Study (WHS) who provided a blood sample at baseline and consent for blood-based analyses [98]. The WHS was a 2×2 trial beginning in 1992–1994 of vitamin E and low dose aspirin in prevention of cancer and cardiovascular disease with about 10 years of follow-up. Follow-up continues in observational mode. Participants in the WHS were 45 or older at enrollment and free of cardiovascular disease, cancer or other major chronic illness and were followed prospectively for the influence of random allocation of vitamin E on cancer [99–101]. Additional information related to health and lifestyle were collected by questionnaire throughout the WHS trial and continuing observational follow-up. Participants in the current analysis are individuals selected for a breast cancer case-control study nested within the WGHS. Eligible cases consisted of women of European ancestry diagnosed with pathologically confirmed incident invasive breast cancer until March 7, 2000. Controls were randomly selected participants who had given a blood sample, and were free of diagnosed cancer. Controls were matched to cases according to age, menopausal status, postmenopausal hormone use at time of blood draw, and race/ethnicity. Written informed consent was obtained from all women before their entry into the trial.

Genotyping and quality control

As part of the CGEMS initiative, genotyping of the NHS samples was conducted for the first stage of a three-stage GWAS of breast cancer susceptibility using the Illumina HumanHap550 Infinium assay (Illumina, San Diego, CA), which contains SNPs derived from the HapMap phase I and II data [102]. Genotyping of PLCO samples for the prostate cancer susceptibility project occurred in two parts [103]. Phase 1A used Illumina's Sentrix HumanHap300 assay and Phase 1B used the Sentrix HumanHap240 assay. For each cohort, samples with call rates $< 90\%$ and single nucleotide polymorphism (SNP) assays with call rates under 90% were removed. Polymorphisms with a minor allele frequency of $< 1\%$ were removed.

Genotyping of WGHS samples was performed using the Illumina HumanHap300 Duo “+” chips or the combination of the HumanHap300 Duo and iSelect chips. In either case, the custom SNP content was the same; these custom SNPs were chosen without regard to minor allele frequency (MAF) to saturate candidate genes for cardiovascular disease as well as to increase coverage of SNPs with known or suspected biological function, e.g. disease association, non-synonymous changes, substitutions at splice sites, etc. For quality control, all samples were required to have successful genotyping using Illumina's BeadStudio v. 3.3 software for at least 98% of the SNPs. Self-reported European ancestry was verified on the basis of multidimensional scaling analysis of identity by state using 1443 ancestry informative markers in PLINK v. 1.06. The final data set retained SNPs with MAF $> 1\%$, successful genotyping in 90% of the subjects, and deviations from Hardy-Weinberg equilibrium not exceeding $P = 10^{-6}$ in significance. Among the final 23,294 individuals of verified European ancestry, genotypes for a total of 2,608,509 SNPs were imputed from the experimental genotypes and LD relationships implicit in the HapMap r. 22 CEU samples.

Genotyping of SNPs selected for replication in the NHS nested skin cancer case-control data set was performed at the Dana Farber/Harvard Cancer Center High-Throughput Genotyping Core. Whole genome amplified DNA was genotyped using the Taqman® OpenArray® Real-Time qPCR system (Applied Biosystems Inc, Foster City, CA). Of the 64 SNPs selected for replication, one failed to genotype. We removed 1 SNP with a call rate <90%. All SNPs were tested for deviation from Hardy-Weinberg equilibrium within the entire skin cancer case-control data set. We observed significant deviations for 3 SNPs at the Bonferroni-adjusted $P < 0.0008$, which were excluded from analyses. We included 5% blinded quality control samples to validate genotyping procedures; concordance for blinded samples was >99%.

Relative telomere length measurement

Genomic DNA was extracted from peripheral blood leukocytes using the QIAmp (Qiagen, Chatsworth, CA) 96-spin blood protocol. DNA was quantified using either the Molecular Devices 96-well spectrophotometer (NHS, WGHS) or the Nanodrop SD-1000 spectrophotometer (PLCO), and subsequently dried down and resuspended to ensure accurate and uniform DNA concentrations. Relative telomere length was measured using a previously described modified version [27] of the quantitative PCR-based telomere assay [78]. Briefly, 5 ng of genomic DNA was dried down in a 384-well plate and resuspended in 10 μ L of either the telomere (T) or 36B4 (S; single copy gene) PCR reaction mixture. The Telomere reaction mixture consists of 1x QuantiTect® SYBR® Green PCR Master Mix (Qiagen), 2 mM of DTT, 270 nM Tel-1b primer, and 900 nM Tel-2b primer. The Telomere thermal cycling profile proceeds as follows: 95°C for 10 minutes then 30 cycles consisting of 95°C for 15 seconds and 54°C for 2 minutes. The 36B4 reaction mixture consists of 1x QuantiTect® SYBR® Green PCR Master Mix, 300 nM 36B4u primer, and 500 nM 36B4d primer. The 36B4 thermal cycling profile proceeds as follows: 95°C for 10 minutes then 30 cycles consisting of 95°C for 15 seconds and 58°C for 1 minute and 10 seconds. The threshold cycle (Ct) value for each reaction represents the number of PCR cycles required to detect a signal over background fluorescence and is inversely proportional to the amount of starting DNA. Assuming 100% PCR efficiency, the amount of PCR product exactly doubles with each cycle. Triplicate reactions of each assay were performed on each sample and the average of the 3 measurements was used for analyses. Relative telomere length is calculated as the exponentiated ratio of the telomere repeat copy number to single-gene copy number ($2^{-T/S}$) and represents the average telomeric DNA signal per genome of an individual. Coefficients of variation (CV) for the telomere and single-gene assay ranged from 0.66% to 3.02% and 0.56% to 2.07%, respectively.

References

1. d'Adda di Fagnana F, Teo SH, Jackson SP (2004) Functional links between telomeres and proteins of the DNA-damage response. *Genes Dev* 18: 1781–1799.
2. Olovnikov AM (1973) A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J Theor Biol* 41: 181–190.
3. Allsopp RC, Vaziri H, Patterson C, Goldstein S, Younglai EV, et al. (1992) Telomere length predicts replicative capacity of human fibroblasts. *Proc Natl Acad Sci U S A* 89: 10114–10118.
4. de Lange T (2004) T-loops and the origin of telomeres. *Nat Rev Mol Cell Biol* 5: 323–329.
5. de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* 19: 2100–2110.
6. Blackburn EH (2000) Telomere states and cell fates. *Nature* 408: 53–56.
7. Greider CW (1999) Telomeres do D-loop-T-loop. *Cell* 97: 419–422.

Statistical analyses

Continuous relative telomere length was natural logarithm transformed to satisfy the assumption of normality. We used multivariable linear regression to analyze the additive effect of each SNP (0, 1, or 2 copies of minor allele) on log relative telomere length using the pooled NHS and PLCO GWAS data set. As log relative telomere length is inversely associated with age in all populations within this study [88,93,97] and significantly inversely associated with smoking in the PLCO population, which included a higher proportion of heavier smokers (**Table 1**) [88], the regression model included age as a continuous variable, smoking status, and pack-year categories of smoking (never, <10, 10 to <30, 30+) in addition to disease status. Since most published studies find shorter telomere lengths among adult men compared to women [42,43,62,65,92,104,105] as well as potentially faster rates of telomere attrition among men [66,67,92], we included gender and an age-gender interaction term in the initial GWAS linear regression analyses. To control for potential confounding by population stratification in the NHS GWAS, PLCO GWAS, and WGHS replication data sets, we additionally adjusted for the top principal components of genetic variation chosen for each study after excluding any admixed individuals clearly not of European descent. Principal components of genetic variation were calculated with EIGENSTRAT software [106] as described in Hunter et al, 2007 [102]. All P values are two-sided. Statistical analyses were performed with SAS (version 9.1; SAS Institute, Cary, NC) and PLINK [107]. Power calculations were performed using QUANTO [108].

Supporting Information

Table S1 Relative telomere length association results from the GWAS and replication populations. (XLSX)

Acknowledgments

We are grateful to the NHS, PLCO, and WGHS participants for their valued participation. We thank H. Ranu and P. Soule for technical assistance and C. Chen for programming support. We thank the laboratory of the Core Genotyping Facility at the N.C.I. for DNA standardization. We thank J. Miletich and A. Parker as well as the technical staff at Amgen for their collaboration and scientific support in performing the genotyping for the WGHS.

Author Contributions

Conceived and designed the experiments: DJH IDV. Performed the experiments: JP DIC SJC DJH. Analyzed the data: JP PK DIC IDV. Contributed reagents/materials/analysis tools: PK DIC SAS LM SIB JLW JH RBH SJC DJH IDV. Wrote the paper: JP PK DIC SAS LM SIB JLW DJH IDV.

14. Kyo S, Takakura M, Kohama T, Inoue M (1997) Telomerase activity in human endometrium. *Cancer Res* 57: 610–614.
15. Yasumoto S, Kunimura C, Kikuchi K, Tahara H, Ohji H, et al. (1996) Telomerase activity in normal human epithelial cells. *Oncogene* 13: 433–439.
16. Hastie ND, Dempster M, Dunlop MG, Thompson AM, Green DK, et al. (1990) Telomere reduction in human colorectal carcinoma and with ageing. *Nature* 346: 866–868.
17. Slagboom PE, Droog S, Boomsma DI (1994) Genetic determination of telomere size in humans: a twin study of three age groups. *Am J Hum Genet* 55: 876–882.
18. Butler MG, Tilburt J, DeVries A, Muralidhar B, Aue G, et al. (1998) Comparison of chromosome telomere integrity in multiple tissues from subjects at different ages. *Cancer Genet Cytogenet* 105: 138–144.
19. Frencq RW, Jr., Blackburn EH, Shannon KM (1998) The rate of telomere sequence loss in human leukocytes varies with age. *Proc Natl Acad Sci U S A* 95: 5607–5610.
20. Friedrich U, Griese E, Schwab M, Fritz P, Thon K, et al. (2000) Telomere length in different tissues of elderly patients. *Mech Ageing Dev* 119: 89–99.
21. Takubo K, Izumiya-Shimomura N, Honma N, Sawabe M, Arai T, et al. (2002) Telomere lengths are characteristic in each human individual. *Exp Gerontol* 37: 523–531.
22. Stewart SA, Weinberg RA (2006) Telomeres: cancer to human aging. *Annu Rev Cell Dev Biol* 22: 531–557.
23. Savage SA, Alter BP (2009) Dyskeratosis congenita. *Hematol Oncol Clin North Am* 23: 215–231.
24. Vulliamy TJ, Marrone A, Knight SW, Walne A, Mason PJ, et al. (2006) Mutations in dyskeratosis congenita: their impact on telomere length and the diversity of clinical presentation. *Blood* 107: 2680–2685.
25. Burtner CR, Kennedy BK (2010) Progeria syndromes and ageing: what is the connection? *Nat Rev Mol Cell Biol* 11: 567–578.
26. Wu X, Amos CI, Zhu Y, Zhao H, Grossman BH, et al. (2003) Telomere dysfunction: a potential cancer predisposition factor. *J Natl Cancer Inst* 95: 1211–1218.
27. McGrath M, Wong JY, Michaud D, Hunter DJ, De Vivo I (2007) Telomere length, cigarette smoking, and bladder cancer risk in men and women. *Cancer Epidemiol Biomarkers Prev* 16: 815–819.
28. Hosgood HD, 3rd, Cawthon R, He X, Chanock S, Lan Q (2009) Genetic variation in telomere maintenance genes, telomere length, and lung cancer susceptibility. *Lung Cancer* 66: 157–161.
29. Jang JS, Choi YY, Lee WK, Choi JE, Cha SI, et al. (2008) Telomere length and the risk of lung cancer. *Cancer Sci* 99: 1385–1389.
30. Mirabello L, Garcia-Closas M, Cawthon R, Lissowska J, Brinton LA, et al. (2010) Leukocyte telomere length in a population-based case-control study of ovarian cancer: a pilot study. *Cancer Causes Control* 21: 77–82.
31. Samani NJ, Boulby R, Butler R, Thompson JR, Goodall AH (2001) Telomere shortening in atherosclerosis. *Lancet* 358: 472–473.
32. Brouillette S, Singh RK, Thompson JR, Goodall AH, Samani NJ (2003) White cell telomere length and risk of premature myocardial infarction. *Arterioscler Thromb Vasc Biol* 23: 842–846.
33. Fitzpatrick AL, Kronmal RA, Gardner JP, Psaty BM, Jenny NS, et al. (2007) Leukocyte telomere length and cardiovascular disease in the cardiovascular health study. *Am J Epidemiol* 165: 14–21.
34. Brouillette SW, Moore JS, McMahon AD, Thompson JR, Ford I, et al. (2007) Telomere length, risk of coronary heart disease, and statin treatment in the West of Scotland Primary Prevention Study: a nested case-control study. *Lancet* 369: 107–114.
35. Zee RY, Michaud SE, Germer S, Ridker PM (2009) Association of shorter mean telomere length with risk of incident myocardial infarction: a prospective, nested case-control approach. *Clin Chim Acta* 403: 139–141.
36. Willeit P, Willeit J, Brandstätter A, Ehrlénbach S, Mayr A, et al. (2010) Cellular aging reflected by leukocyte telomere length predicts advanced atherosclerosis and cardiovascular disease risk. *Arterioscler Thromb Vasc Biol* 30: 1649–1656.
37. Cawthon RM, Smith KR, O'Brien E, Sivatchenko A, Kerber RA (2003) Association between telomere length in blood and mortality in people aged 60 years or older. *Lancet* 361: 393–395.
38. Kimura M, Hjelmborg JV, Gardner JP, Bathum L, Brimacombe M, et al. (2008) Telomere length and mortality: a study of leukocytes in elderly Danish twins. *Am J Epidemiol* 167: 799–806.
39. Epel ES, Merkin SS, Cawthon R, Blackburn EH, Adler NE, et al. (2009) The rate of leukocyte telomere shortening predicts mortality from cardiovascular disease in elderly men. *Ageing (Albany NY)* 1: 81–88.
40. Bakaysa SL, Mucci LA, Slagboom PE, Boomsma DI, McClearn GE, et al. (2007) Telomere length predicts survival independent of genetic influences. *Ageing Cell* 6: 769–774.
41. Njajou OT, Cawthon RM, Damcott CM, Wu SH, Ott S, et al. (2007) Telomere length is paternally inherited and is associated with parental lifespan. *Proc Natl Acad Sci U S A* 104: 12135–12139.
42. Jeanclous E, Schork NJ, Kyvik KO, Kimura M, Skurnick JH, et al. (2000) Telomere length inversely correlates with pulse pressure and is highly familial. *Hypertension* 36: 195–200.
43. Vasa-Nicotera M, Brouillette S, Mangino M, Thompson JR, Braund P, et al. (2005) Mapping of a major locus that determines telomere length in humans. *Am J Hum Genet* 76: 147–151.
44. Andrew T, Aviv A, Falchi M, Surdulescu GL, Gardner JP, et al. (2006) Mapping genetic loci that determine leukocyte telomere length in a large sample of unselected female sibling pairs. *Am J Hum Genet* 78: 480–486.
45. Atzmon G, Cho M, Cawthon RM, Budagov T, Katz M, et al. (2010) Evolution in health and medicine Sackler colloquium: Genetic variation in human telomerase is associated with telomere length in Ashkenazi centenarians. *Proc Natl Acad Sci U S A* 107 Suppl 1: 1710–1717.
46. Okuda K, Bardeguet A, Gardner JP, Rodriguez P, Ganesh V, et al. (2002) Telomere length in the newborn. *Pediatr Res* 52: 377–381.
47. Youngren K, Jeanclous E, Aviv H, Kimura M, Stock J, et al. (1998) Synchrony in telomere length of the human fetus. *Hum Genet* 102: 640–643.
48. Martens UM, Zijlmans JM, Poon SS, Dragowska W, Yui J, et al. (1998) Short telomeres on human chromosome 17p. *Nat Genet* 18: 76–80.
49. Mangino M, Brouillette S, Braund P, Tirmizi N, Vasa-Nicotera M, et al. (2008) A regulatory SNP of the BICD1 gene contributes to telomere length variation in humans. *Hum Mol Genet* 17: 2518–2523.
50. Mangino M, Richards JB, Soranzo N, Zhai G, Aviv A, et al. (2009) A genome-wide association study identifies a novel locus on chromosome 18q12.2 influencing white cell telomere length. *J Med Genet* 46: 451–454.
51. Codd V, Mangino M, van der Harst P, Braund PS, Kaiser M, et al. (2010) Common variants near TERC are associated with mean telomere length. *Nat Genet* 42: 197–199.
52. Levy D, Neuhausen SL, Hunt SC, Kimura M, Hwang SJ, et al. (2010) Genome-wide association identifies OBF1 as a locus involved in human leukocyte telomere biology. *Proc Natl Acad Sci U S A* 107: 9293–9298.
53. Varadi V, Brendle A, Brandt A, Johansson R, Enquist K, et al. (2009) Polymorphisms in telomere-associated genes, breast cancer susceptibility and prognosis. *Eur J Cancer* 45: 3008–3016.
54. Prescott J, McGrath M, Lee IM, Buring JE, De Vivo I (2010) Telomere length and genetic analyses in population-based studies of endometrial cancer risk. *Cancer*.
55. Mirabello L, Yu K, Kraft P, De Vivo I, Hunter DJ, et al. (2010) The association of telomere length and genetic variation in telomere biology genes. *Hum Mutat*.
56. Wan M, Qin J, Songyang Z, Liu D (2009) OB fold-containing protein 1 (OBF1), a human homolog of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation. *J Biol Chem* 284: 26725–26731.
57. Pe'er I, Yelensky R, Altshuler D, Daly MJ (2008) Estimation of the multiple testing burden for genome-wide association studies of nearly all common variants. *Genet Epidemiol* 32: 381–385.
58. Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, et al. (2009) Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat Genet* 41: 221–227.
59. Askree SH, Yehuda T, Smolnikov S, Gurevich R, Hawk J, et al. (2004) A genome-wide screen for *Saccharomyces cerevisiae* deletion mutants that affect telomere length. *Proc Natl Acad Sci U S A* 101: 8658–8663.
60. Gatbonton T, Imbesi M, Nelson M, Akey JM, Ruderfer DM, et al. (2006) Telomere length as a quantitative trait: genome-wide survey and genetic mapping of telomere length-control genes in yeast. *PLoS Genet* 2: e35.
61. Savage SA, Stewart BJ, Eckert A, Kiley M, Liao JS, et al. (2005) Genetic variation, nucleotide diversity, and linkage disequilibrium in seven telomere stability genes suggest that these genes may be under constraint. *Hum Mutat* 26: 343–350.
62. Nawrot TS, Staessen JA, Gardner JP, Aviv A (2004) Telomere length and possible link to X chromosome. *Lancet* 363: 507–510.
63. Akkad A, Hastings R, Konje JC, Bell SC, Thurston H, et al. (2006) Telomere length in small-for-gestational-age babies. *BJOG* 113: 318–323.
64. Nordfjall K, Larefalk A, Lindgren P, Holmberg D, Roos G (2005) Telomere length and heredity: Indications of paternal inheritance. *Proc Natl Acad Sci U S A* 102: 16374–16378.
65. Nordfjall K, Svenson U, Norrback KF, Adolfsson R, Roos G (2010) Large-scale parent-child comparison confirms a strong paternal influence on telomere length. *Eur J Hum Genet* 18: 385–389.
66. Unryn BM, Cook LS, Riabowol KT (2005) Paternal age is positively linked to telomere length of children. *Ageing Cell* 4: 97–101.
67. De Meyer T, Rietzschel ER, De Buyere ML, De Bacquer D, Van Criekinge W, et al. (2007) Paternal age at birth is an important determinant of offspring telomere length. *Hum Mol Genet* 16: 3097–3102.
68. Kimura M, Cherkas LF, Kato BS, Demissie S, Hjelmborg JB, et al. (2008) Offspring's leukocyte telomere length, paternal age, and telomere elongation in sperm. *PLoS Genet* 4: e37.
69. Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, et al. (2009) Parental origin of sequence variants associated with complex diseases. *Nature* 462: 868–874.
70. Londono-Vallejo JA, Der Sarkissian H, Cazes L, Thomas G (2001) Differences in telomere length between homologous chromosomes in humans. *Nucleic Acids Res* 29: 3164–3171.
71. Graakjaer J, Bischoff C, Korsholm L, Holstebro S, Vach W, et al. (2003) The pattern of chromosome-specific variations in telomere length in humans is determined by inherited, telomere-near factors and is maintained throughout life. *Mech Ageing Dev* 124: 629–640.
72. Goldman F, Bouarich R, Kulkarni S, Freeman S, Du HY, et al. (2005) The effect of TERC haploinsufficiency on the inheritance of telomere length. *Proc Natl Acad Sci U S A* 102: 17119–17124.

73. Britt-Compton B, Rowson J, Locke M, Mackenzie I, Kipling D, et al. (2006) Structural stability and chromosome-specific telomere length is governed by cis-acting determinants in humans. *Hum Mol Genet* 15: 725–733.
74. Graakjaer J, Der-Sarkissian H, Schmitz A, Bayer J, Thomas G, et al. (2006) Allele-specific relative telomere lengths are inherited. *Hum Genet* 119: 344–350.
75. Baird DM, Rowson J, Wynford-Thomas D, Kipling D (2003) Extensive allelic variation and ultrashort telomeres in senescent human cells. *Nat Genet* 33: 203–207.
76. Mefford HC, Trask BJ (2002) The complex structure and dynamic evolution of human subtelomeres. *Nat Rev Genet* 3: 91–102.
77. Li M, Li C, Guan W (2008) Evaluation of coverage variation of SNP chips for genome-wide association studies. *Eur J Hum Genet* 16: 635–643.
78. Cawthon RM (2002) Telomere measurement by quantitative PCR. *Nucleic Acids Res* 30: e47.
79. Lin KW, Yan J (2005) The telomere length dynamic and methods of its assessment. *J Cell Mol Med* 9: 977–989.
80. Iwama H, Ohyashiki K, Ohyashiki JH, Hayashi S, Yahata N, et al. (1998) Telomeric length and telomerase activity vary with age in peripheral blood cells obtained from normal individuals. *Hum Genet* 102: 397–402.
81. Nettleton JA, Diez-Roux A, Jenny NS, Fitzpatrick AL, Jacobs DR, Jr. (2008) Dietary patterns, food groups, and telomere length in the Multi-Ethnic Study of Atherosclerosis (MESA). *Am J Clin Nutr* 88: 1405–1412.
82. von Zglinicki T (2002) Oxidative stress shortens telomeres. *Trends Biochem Sci* 27: 339–344.
83. Garcia CK, Wright WE, Shay JW (2007) Human diseases of telomerase dysfunction: insights into tissue aging. *Nucleic Acids Res* 35: 7406–7416.
84. Valdes AM, Andrew T, Gardner JP, Kimura M, Oelsner E, et al. (2005) Obesity, cigarette smoking, and telomere length in women. *Lancet* 366: 662–664.
85. Nordfjall K, Eliasson M, Stegmayr B, Melander O, Nilsson P, et al. (2008) Telomere length is associated with obesity parameters but with a gender difference. *Obesity (Silver Spring)* 16: 2682–2689.
86. Cherkas LF, Hunkin JL, Kato BS, Richards JB, Gardner JP, et al. (2008) The association between physical activity in leisure time and leukocyte telomere length. *Arch Intern Med* 168: 154–158.
87. Kim S, Parks CG, DeRoo LA, Chen H, Taylor JA, et al. (2009) Obesity and weight gain in adulthood and telomere length. *Cancer Epidemiol Biomarkers Prev* 18: 816–820.
88. Mirabello L, Huang WY, Wong JY, Chatterjee N, Reding D, et al. (2009) The association between leukocyte telomere length and cigarette smoking, dietary and physical variables, and risk of prostate cancer. *Aging Cell* 8: 405–413.
89. Ornish D, Lin J, Daubenmier J, Weidner G, Epel E, et al. (2008) Increased telomerase activity and comprehensive lifestyle changes: a pilot study. *Lancet Oncol* 9: 1048–1057.
90. Huda N, Tanaka H, Herbert BS, Reed T, Gilley D (2007) Shared environmental factors associated with telomere length maintenance in elderly male twins. *Aging Cell* 6: 709–713.
91. Collerton J, Martin-Ruiz C, Kenny A, Barras K, von Zglinicki T, et al. (2007) Telomere length is associated with left ventricular function in the oldest old: the Newcastle 85+ study. *Eur Heart J* 28: 172–176.
92. Bekaert S, De Meyer T, Rietzschel ER, De Buyzere ML, De Bacquer D, et al. (2007) Telomere length and cardiovascular risk factors in a middle-aged population free of overt cardiovascular disease. *Aging Cell* 6: 639–647.
93. De Vivo I, Prescott J, Wong JY, Kraft P, Hankinson SE, et al. (2009) A prospective study of relative telomere length and postmenopausal breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 18: 1152–1156.
94. Cassidy A, De Vivo I, Liu Y, Han J, Prescott J, et al. (2010) Associations between diet, lifestyle factors, and telomere length in women. *Am J Clin Nutr* 91: 1273–1280.
95. Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, et al. (2000) Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 21: 273S–309S.
96. Gohagan JK, Prorok PC, Hayes RB, Kramer BS (2000) The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. *Control Clin Trials* 21: 251S–272S.
97. Han J, Qureshi AA, Prescott J, Guo Q, Ye L, et al. (2009) A prospective study of telomere length and the risk of skin cancer. *J Invest Dermatol* 129: 415–421.
98. Ridker PM, Chasman DI, Zee RY, Parker A, Rose L, et al. (2008) Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin Chem* 54: 249–255.
99. Cook NR, Lee IM, Gaziano JM, Gordon D, Ridker PM, et al. (2005) Low-dose aspirin in the primary prevention of cancer: the Women's Health Study: a randomized controlled trial. *JAMA* 294: 47–55.
100. Lee IM, Cook NR, Gaziano JM, Gordon D, Ridker PM, et al. (2005) Vitamin E in the primary prevention of cardiovascular disease and cancer: the Women's Health Study: a randomized controlled trial. *JAMA* 294: 56–65.
101. Ridker PM, Cook NR, Lee IM, Gordon D, Gaziano JM, et al. (2005) A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women. *N Engl J Med* 352: 1293–1304.
102. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, et al. (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 39: 870–874.
103. Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, et al. (2007) Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 39: 645–649.
104. Nordfjall K, Svenson U, Norrback KF, Adolffsson R, Lenner P, et al. (2009) The individual blood cell telomere attrition rate is telomere length dependent. *PLoS Genet* 5: e1000375.
105. Willeit P, Willeit J, Mayr A, Weger S, Oberhollenzer F, et al. (2010) Telomere length and risk of incident cancer and cancer mortality. *JAMA* 304: 69–75.
106. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909.
107. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
108. Gauderman W, Morrison J (2006) QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. <http://hydra.usc.edu/gxe>.