EBioMedicine 11 (2016) 296–301

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Racial and Socioeconomic Variation in Genetic Markers of Telomere Length: A Cross-Sectional Study of U.S. Older Adults

Rita Hamad [a,*], Shripad Tuljapurkar [b], David H. Rehkopf [a]

[a] Stanford University, Department of Medicine, 1070 Arastradero Road, Palo Alto, CA 94304, USA
[b] Stanford University, Biology Department, Herrin Labs 454, Stanford, CA 94305, USA

ARTICLE INFO

Article history:
Received 24 June 2016
Received in revised form 2 August 2016
Accepted 10 August 2016
Available online 11 August 2016

Keywords:
Telomere
Telomere length
Race
Socioeconomic position
Single nucleotide polymorphism
Polygenic risk score

ABSTRACT

Background: Shorter telomere length (TL) has been associated with stress and adverse socioeconomic conditions, yet U.S. blacks have longer TL than whites. The role of genetic versus environmental factors in explaining TL by race and socioeconomic position (SEP) remains unclear.
Methods: We used data from the U.S. Health and Retirement Study ($N = 11{,}934$) to test the hypothesis that there are differences in TL-associated SNPs by race and SEP. We constructed a TL polygenic risk score (PRS) and examined its association with race/ethnicity, educational attainment, assets, gender, and age.
Results: U.S. blacks were more likely to have a lower PRS for TL, as were older individuals and men. Racial differences in TL were statistically accounted for when controlling for population structure using genetic principal components. The GWAS-derived SNPs for TL, however, may not have consistent associations with TL across different racial/ethnic groups.
Conclusions: This study showed that associations of race/ethnicity with TL differed when accounting for population stratification. The role of race/ethnicity for TL remains uncertain, however, as the genetic determinants of TL may differ by race/ethnicity. Future GWAS samples should include racially diverse participants to allow for better characterization of the determinants of TL in human populations.

1. Introduction

Telomeres are DNA-protein structures that include a repeated nucleotide sequence at the ends of eukaryotic chromosomes, acting to prevent the degradation of functional DNA sequences during cellular replication (Olovnikov, 1973). Shortening of telomeres in human cells in vitro has been shown to lead to cellular dysfunction, senescence, and cell death (Allsopp and Harley, 1995; Blackburn, 2000). Telomere length (TL) has therefore been hypothesized to be a marker of human aging and chronic disease (Zhu et al., 2011a; Hamad et al., 2016). Observational studies in human populations have built on this molecular basis and have shown that chronic and acute stressors, including adverse socioeconomic conditions, are associated with shorter TL, suggesting that environmental factors may lead to variation in TL in later life (Chae et al., 2014; Epel et al., 2004; Needham et al., 2013). Thus, telomeres are thought to shorten with natural aging and with stress over the lifespan. It remains unclear whether these correlations are due to a causal relationship between stress and TL, or whether the association is confounded by other unobserved factors. Nevertheless, there is substantial evidence that TL acts as a marker of biological aging, even if it is not a causal relationship.

The observational literature contains many investigations of racial differences in human TL, with findings that seem to contradict the association between shorter TL and stress. The majority of studies have found that U.S. blacks have longer telomeres than whites, despite on average lower socioeconomic position (SEP) and presumably higher levels of stressors (Adler et al., 2013; Diaz et al., 2010; Hunt et al., 2008; Needham et al., 2013). Research has shown that this is likely the case beginning at birth: several studies have found longer telomeres among black newborns and adolescents, ostensibly before life stressors can accumulate (Rewak et al., 2014; Zhu et al., 2011b), although a single study found similar TL among black and white newborns (Okuda et al., 2002). A smaller number of studies suggest that blacks have shorter TL compared to whites (Diez roux et al., 2009; Geronimus et al., 2010), while others demonstrate more rapid shortening of telomeres among blacks over the life course (Hunt et al., 2008; Rewak et al., 2014).

While it is assumed that most demographic differences in TL are due to environmental factors, this prior evidence is difficult to reconcile with current knowledge of the determinants of TL. A growing body of research has begun to elucidate genetic markers – i.e., single nucleotide polymorphisms (SNPs) – associated with TL, many of which fall in genes known to act on telomere biology (Codd et al., 2013; Levy et al.,

* Corresponding author.
E-mail address: rhamad@stanford.edu (R. Hamad).

http://dx.doi.org/10.1016/j.ebiom.2016.08.015

2010). Yet the majority of these studies have been conducted among white or Asian populations, with few examining racial differences in the prevalence of these markers that might explain TL differences across the life course. One prior abstract examined racial differences in TL-associated genetic markers, finding that SNPs strongly associated with TL in whites were only weakly associated when replications were conducted in other racial groups (Kvale et al., 2012). Another included African American participants in a replication sample, but did not conduct comparative analyses as we do here (Levy et al., 2010). In fact, the lack of attention to cataloguing SNPs in populations of African ancestry has been identified as a major gap in the literature: only 4% of genome-wide association studies have been conducted among individuals of non-European descent (Bustamante et al., 2011), although prior research documents racial differences in a variety of genetic markers across the genome (Jorgenson et al., 2005).

This study is one of the first assessments of differences in the demographic distribution of TL-related SNPs identified in genome-wide association studies (GWAS), examining whether genetic makeup explains the variation in TL among various subgroups. We link genetic and survey data from the Health and Retirement Study, a nationally representative diverse U.S. sample, to test the hypothesis that there are differences in the prevalence of TL-associated SNPs in different racial and SEP subgroups.

2. Methods

2.1. Data Set

We used data from the U.S. Health and Retirement Study (HRS), a longitudinal panel study that has collected data biennially since 1992 among a representative sample of over 26,000 men and women over 50 years of age, with an over-sampling of older individuals. The survey also included data on respondents' spouses, which includes individuals under 50 years of age. Details on the HRS, including survey design, have been previously described (Juster and Suzman, 1995). We restricted our analyses to individuals for whom we have data on genotype or TL ($N = 11{,}934$).

2.2. Measures

The primary outcome variable for this study was a polygenic risk score (PRS) composed of seven SNPs that were previously associated with TL in a genome-wide meta-analysis (Codd et al., 2013). Genetic data collected from respondents during the 2006 and 2008 study waves were used to construct this score ($N = 11{,}143$). Subjects provided DNA samples using a mouthwash technique, and genotyping was conducted by the NIH Center for Inherited Disease Research using the Illumina Human Omni-2.5 Quad beadchip, which includes roughly 2.5 million SNPs. Further information on quality control procedures is available from HRS. Of the almost 36,986 individuals interviewed by HRS since 1992, 8151 died before genetic testing became available. These individuals were more likely to be male, white, and less educated. Meanwhile, 17,692 of those who survived did not provide genetic samples because testing was not offered to them or because they refused. These individuals were more likely to be male, non-white, and less educated.

We constructed the PRS for each individual by forming the weighted sum of alleles for these seven SNPs. Weights were assigned using log-odds ratios for each SNP as reported by Codd et al. (Supplemental Table 1), according to the following formula:

$$PRS_i = \sum_1^7 \beta_k allele_{ki}$$

Here, β_k is the log-odds ratio for each of the seven SNPs k, and $allele_k$ is the number of alleles present of SNP k for a given individual i. We standardized the PRS to have a mean of zero and a standard deviation of one so that associations are more easily interpretable and comparable to other studies; the score itself is only slightly left-skewed. A higher value means that an individual is genetically predisposed to shorter telomeres. Prior research has suggested that analyses using a PRS may be more successful in predicting disease risk than use of individual genetic markers (Dudbridge, 2013). Of note, the meta-analysis by Codd et al. included only individuals of European descent.

The secondary outcome variable was mean TL, which was obtained in 2008 from HRS participants who consented to provide a salivary sample ($N = 5808$). Samples were analyzed by Telome Health using a standard quantitative polymerase chain reaction (PCR) assay. TL was measured in standard fashion, using the telomere-to-single copy gene (T/S) ratio. This ratio was determined by comparing the telomere sequence copy number (T) with a single-copy gene copy number (S). The equation for conversion of the T/S ratio to TL varies by lab, and for this study was: base pairs $= (T/S) * 2400$.

Covariates included race, age at genotype testing, gender, educational attainment, and total assets. Racial categories were self-reported and included non-Hispanic white (reference group), black, Hispanic, and other. Educational attainment was constructed as a categorical variable with four levels: less than high school education (reference group), high school or GED completed, some college, and college completed. Total assets were highly right-skewed with some negative values, and therefore a Z-score was created by standardizing with a mean of zero and standard deviation of one.

To account for population stratification, we constructed four principal components to represent genetic structure within the sample (Price et al., 2006). We included these as covariates in some models to control for biased estimates that could result from differences in genetic structure between populations being compared.

For each of the variables used, <3% of values were missing, so imputation was not conducted.

2.3. Data Analysis

We first examined the racial and socioeconomic predictors of the PRS using linear regressions. Each of these predictors was included in bivariate analyses, and then a full multivariable model was constructed using all covariates. In secondary analyses, we also included the four genetic principal components.

Next, we conducted similar bivariate and multivariable linear regressions to examine racial and socioeconomic correlates of TL in the smaller sample of participants who had provided samples. A secondary analysis additionally controlled for genetic principal components. In all models, robust standard errors were clustered at the household level.

We also conducted several supplemental analyses. For each SNP, we graphically examined whether there were differences in the number of risk alleles by race; our study was not powered to examine these individual differences statistically. Next, we examined the association between self-reported race and principal components for genetic structure. Finally, we assessed the association between individual TL-associated SNPs and TL by race.

2.4. Ethics Approval

Ethics approval for this study was provided by the Stanford University Institutional Review Board (protocol # 25818). Approval for the HRS was provided by the University of Michigan Health Services Institutional Review Board.

3. Results

3.1. Sample Characteristics

The sample consisted of HRS participants who had provided genetic or telomere data ($N = 11{,}934$). Participants were diverse with respect to demographic and socioeconomic characteristics (Table 1).