Published in final edited form as:
Circ Cardiovasc Genet. 2014 June ; 7(3): 287–295. doi:10.1161/CIRCGENETICS.113.000412.

A Genome-Wide Association Study Identifies Variants in Casein Kinase II (*CSNK2A2*) to be Associated with Leukocyte Telomere Length in a Punjabi Sikh Diabetic Cohort

Richa Saxena, PhD[1,2], **Andrew Bjonnes, MS**[1,2], **Jennifer Prescott, PhD**[4,5], **Patrick Dib, MS**[3], **Praveen Natt, MS**[3], **Jacqueline Lane, PhD**[1,2], **Megan Lerner, HT**[6], **Jackie A. Cooper, MS**[7], **Yuanqing Ye, PhD**[8], **Ka Wah Li, MS**[7], **Cécilia G. Maubaret, PhD**[7], **Veryan Codd, PhD**[9], **Daniel Brackett, AS**[6], **Lisa Mirabello, PhD**[10], **Peter Kraft, PhD**[5,11], **Colin P. Dinney, MD**[12], **Donald Stowell, MD**[6], **Marvin Peyton, MD**[6], **Sarju Ralhan, MBBS, MS, MCh**[13], **Gurpreet S. Wander, MBBS, MD, DM**[13], **Narinder K. Mehra, PhD**[14], **Klelia D. Salpea, PhD**[7], **Jian Gu, PhD**[8], **Xifeng Wu, MD, PhD**[8], **Massimo Mangino, PhD**[9], **David J. Hunter, MBBS, ScD**[2,4,5], **Immaculata De Vivo, PhD**[4,5], **Steve E. Humphries, PhD**[7], **Nilesh J. Samani, MD**[9], **Tim D. Spector, MD**[9], **Sharon A. Savage, MD**[10], and **Dharambir K. Sanghera, PhD**[3]

[1]Center for Human Genetic Research, Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston

[2]Broad Institute of Massachusetts Institute of Technology & Harvard, Cambridge, MA

[3]Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK

[4]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital & Harvard Medical School, Boston

[5]Program in Molecular and Genetic Epidemiology, Department of Epidemiology, Harvard School of Public Health, Boston, MA

[6]Department of Surgery, University of Oklahoma Health Sciences Center, Oklahoma City, OK

[7]Cardiovascular Genetics, BHF Laboratories, Institute Cardiovascular Sciences, University College London, London, United Kingdom

[8]Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX

[9]National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, United Kingdom

Correspondence: Dharambir K. Sanghera, PhD, FAHA, Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences Center, 940 Stanton L. Young Blvd., Rm 317 BMSB, Oklahoma City, OK 73104, Tel: 405-271-8001 Ext. 43064, Fax: 405-271-6027, Dharambir-sanghera@ouhsc.edu.

URLs: IMPUTE - http://mathgen.stats.ox.ac.uk/impute/impute.html / METAL - http://www.sph.umich.edu/csg/abecasis/Metal / SNAP - http://www.broadinstitute.org/mpg/snap/ / LocusZoom - http://csg.sph.umich.edu/locuszoom/ / GenABEL - http://www.genabel.org/ / ProbABEL - http://www.genabel.org/packages/ProbABEL

Conflict of Interest Disclosures: None

[10]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD

[11]Department of Biostatistics, Harvard School of Public Health, Boston, MA

[12]Department of Urology, The University of Texas MD Anderson Cancer Center, Houston, TX

[13]Hero DMC Heart Institute, Ludhiana, Punjab

[14]All India Institute of Medical Sciences and Research, New Delhi, India

Abstract

Background—Telomere length is a heritable trait and short telomere length has been associated with multiple chronic diseases. We investigated the relationship of relative leukocyte telomere length (RTL) with cardiometabolic risk and performed the first GWAS and meta-analysis to identify variants influencing RTL in a population of Sikhs from South Asia.

Methods and Results—Our results revealed a significant independent association of shorter RTL with type 2 diabetes (T2D) and heart disease. Our discovery GWAS (n=1,616) was followed by Stage 1 replication of 25 top signals ($P<10^{-6}$) in an additional Sikhs (n=2,397). On combined discovery and Stage 1 meta-analysis (n= 4013), we identified a novel RTL locus at chromosome 16q21 represented by an intronic variant (rs74019828) in the *CSNK2A2* gene (β −0.38, $P=4.5\times10^{-8}$). We further tested 3 top variants by genotyping in UKCVD (Caucasians n=2,952) for Stage 2. Next we performed *in silico* replication of 139 top signals ($p<10^{-5}$) in UKTWIN, NHS, PLCO and MDACC (n=10,033) and joint meta-analysis (n=16,998). The observed signal in *CSNK2A2* was confined to South Asians and could not be replicated in Caucasians due to significant difference in allele frequencies ($P<0.001$). CSNK2A2 phosphorylates TRF1 and plays an important role for regulation of telomere length homoeostasis.

Conclusions—By identification of a novel signal in telomere pathway genes, our study provides new molecular insight into the underlying mechanism that may regulate telomere length and its association with human aging and cardiometabolic pathophysiology.

Keywords

telomere genetics; type 2 diabetes mellitus; Genome Wide Association Study; cardiovascular disease

Introduction

Leukocyte telomere length has been inversely associated with multiple diseases including osteoporosis, hypertension, myocardial infarction (MI), coronary heart disease (CHD), type 2 diabetes (T2D), and Alzheimer's disease[1–5]. Growing evidence suggests that telomere length plays a critical role in cellular aging[6, 7]. As loss of telomeric DNA is impacted by cell division and oxidative stress, and induces cellular senescence, telomere length has been postulated as a biomarker of senescence and human aging in many published studies[6, 8]. Inter-individual variation in telomere length at birth and subsequent years is attributed to by both genetic and environmental factors that start *in utero*[9]. Ethnic differences in the relative leukocyte telomere length (RTL) have been reported in multi-ethnic studies where

populations such as African Americans and Hispanics had significant age-associated differences in RTL than Caucasian participants[10]. Further, a cumulative impact of differential exposures to oxidative stress and other environmental stressors on telomere attrition in different ethnic groups have been shown to be predictors of cellular and biological aging that may impact race and ethnicity-related health outcomes[11–13].

Gene mapping and twin studies have confirmed the strong influence of genetic factors for controlling RTL with heritability estimates ranging from 0.36–0.84[14–16]. Genome-wide association studies (GWAS) and candidate gene studies have identified common genetic variation contributing to RTL in healthy and disease conditions[13, 17–20]. Most GWAS on RTL have been performed in populations of European ancestry. With exception of three small studies (comprised of 40 male and females from Chennai, India[12], 218 males from UK[21], and the largest with 238 patients who had undergone coronary artery bypass graft (CABG) and 238 controls from Mumbai, India), no previous study has comprehensively evaluated the association of RTL with cardiometabolic risk or the role of genetic factors on RTL in South Asians. The goals of this investigation were; 1) to test association of RTL with cardiometabolic traits in this diabetic sample with a high risk of CHD, 2) to confirm whether gene variants identified in earlier GWA studies replicate in a diabetic case-control cohort of Punjabi Sikhs, and 3) to identify new genomic regions associated with RTL by GWAS, replication, and meta-analysis studies in cohorts of South Asian and European ancestry.

Materials and Methods

Sample and characteristics

Our primary Punjabi Sikh discovery and replication study comprised 4,013 individuals including 1,616 in the GWAS (discovery) and 2,397 in the replication (Stage 1) (Supplementary Table 1, 2A and Figure 1). These subjects were part of the Asian Indian Diabetic Heart Study, also named the Sikh Diabetes Study (AIDHS/SDS) as described previously[22]. The AIDHS/SDS has unique characteristics that are ideal for genetic studies. Sikhs are *strictly* a non-smoking population and about 50% of participants are teetotalers and life-long vegetarians. All individuals for the GWAS discovery cohort were recruited from one geographical location. Diagnosis of T2D was confirmed by scrutinizing medical records for symptoms, use of medication, and measuring fasting glucose levels following the guidelines of the American Diabetes Association[23], as described previously[24]. Details of sample recruitment and clinical phenotypes are described previously (Saxena et al, 2013)[22]. The selection of controls was based on a fasting glucose of <100.8 mg/dL or a 2h glucose <141.0 mg/dL. All blood samples were obtained at the baseline visits.

CHD was considered if there was use of nitrate medication (nitroglycerine), electrocardiographic evidence of angina pain, coronary angiographic evidence of severe (greater than 50%) stenosis, or echocardiographic evidence of myocardial infarction. Diagnosis was based on date of CABG or angioplasty, and medication usage obtained from patient records as described previously[25]. In this study, approximately 11% of participants had CHD in discovery and 21% in replication cohorts. All participants signed a written informed consent for the investigations. The study was reviewed and approved by the

University of Oklahoma Health Sciences Center's Institutional Review Board, as well as the Human Subject Protection Committees at the participating institutes in India. Description of the datasets used in Stage 2 and Stage 3 replication are described in Supplementary Section. Clinical characteristics of the Stage 2 replication and Stage 3 and look-up cohorts are described in Supplementary Table 2B.

Anthropometric and metabolic measures

Body mass index (BMI) was calculated as (weight [kg]/height [meter]^2), and waist to hip ratio (WHR) was calculated as the ratio of abdomen or waist circumference to hip circumference Details of demographic, anthropometric and clinical traits are summarized in Supplementary Table 2A. Insulin was measured by radio-immuno assay (Diagnostic Products, Cypress, USA). Serum lipids (total cholesterol, low density lipoprotein cholesterol [LDL-C], high-density lipoprotein [HDL-C], very low-density lipoprotein cholesterol [VLDL-C], and triglycerides [TG]) were measured by using standard enzymatic methods (Roche, Basel, Switzerland), as described[26, 27]. C-peptide, leptin, amylin, and MCP-1 measures were simultaneously quantified using Millipore's Magnetic MILLIPLEX Human Metabolic panel (St. Charles, MO) and analyzed on a Bio-plex 200 multiplex system (Bio-Rad Hercules, CA), as described previously[28].

Punjabi Sikh Discovery GWAS

Study design for the RTL GWA study is shown in Figure 1. Clinical characteristics of GWAS (discovery) and Stage 1 (replication) in Punjabi Sikh cohorts are described in (Supplementary Table 2A). After quality control as described previously[22], 474,231 directly genotyped SNPs (MAF ≥5%) in 1,616 subjects (842 cases and 774 controls) from 1,850 total subjects were available for association testing. To increase genome coverage, genotypes were imputed for un-typed SNPs and in-dels using the 1kG multi-ethnic reference panel, yielding a total of 6,378,483 variants, and of these 5,904,251 with MAF≥ 5% were analyzed in the current investigation.

Measurement of relative telomere length (RTL)

Genomic DNA was extracted from blood buffy coats as described previously[29]. DNA was quantified and equilibrated using Quant-iTed with PicoGreen and using lambda DNA standard (Invitrogen, Eugene, OR, USA). Telomere length was measured on the Applied Biosystems 7900HT Genetic Analyzer using a modified version of Cawthon's quantitative PCR-based method[30]. Each individual DNA sample of 20 ng/ul was assayed in duplicate for measuring telomere length. Primer sequences used to amplify the single copy gene (36b4) and telomere repeats are listed in Supplementary Table 3. Each run also contained no template controls and 5–6 technical replicates previously determined to have short, medium, and long telomere. Samples from T2D and CHD patients and healthy controls were run blindly for RTL measurements both in discovery and Stage 1 replication. A standard curve was used with a range of concentrations (2-fold dilutions) and mixing multiple samples during initial optimization. Any sample that fell out of range was repeated and outliers (~2%) were discarded. RTL was calculated using T/S ratio of telomere repeats (T) to single copy gene (S). Assays were conducted blind to the disease status, age or gender of the individuals. The overall coefficient of variation (CV) for the telomere length variable in

AIDHS/SDS is 2.83% in telomere, 1.34% in single copy run and 7.80% in T/S ratio for the entire sample. The inter-plate mean CV was 4.84% in telomere, 3.26% in single copy and 7.41% in T/S ratio.

In most consortium cohorts (UKCVD, UKTWIN, MDACC, NHS and PLCO), the telomere length was measured similarly using quantitative PCR-based methods[19, 31–33]. Despite measured similarly, the mean (± SD) of RTL ranged from 1.19 (0.37) to 3.71 (0.69) and varied across studies (Supplementary Table 2B).

Genotyping and Statistical Analysis

Genome-wide genotyping and quality control: (Discovery)—Genome-wide genotyping was performed using the Human 660W Quad BeadChip panel (Illumina, Valencia, CA, USA) described in detail previously[22]. Briefly, samples with genotyping call rate <95%, cryptic relatedness, population outliers and extremes of heterozygosity (+/− 3 s.d.) were removed, and SNPs with genotyping call rate <95%, departures from Hardy-Weinberg equilibrium ($P<10^{-7}$) or minor allele frequency (MAF) <5% were excluded using the software PLINK[34] before association testing or imputation. As described previously in Saxena et al, 2013[22], the inbreeding coefficient and measures of autozygosity were determined using the PLINK. We identified runs of homozygosity using the metrics defined in Nalls et al.[35], evaluating 1 Mb autosomal regions with at least 50 adjacent SNPs, with a sliding window of 50 SNPs including no more than 2 SNPs with missing genotypes and 1 possible heterozygous genotype.

RTL Association to cardiometabolic traits—Associations between natural log transformed (ln) RTL and cardiometabolic traits including anthropometric traits (age, BMI, waist), Diabetes (fasting glucose, homeostasis model assessment for insulin resistance [HOMA-IR] and beta cell function [HOMA-B], C-peptide, leptin, and amylin etc), and cardiovascular traits (CHD, blood pressure and MCP-1) were assessed using linear and logistic regression in SPSS, adjusting for significant covariates, including age, gender and T2D status.

Sikh RTL genome-wide association analysis (Discovery)—Associations of SNPs with ln RTL were tested using linear regression and an additive genetic model. Age, gender, BMI, T2D status and five principal components to adjust for residual population stratification were included as covariates. As the existing HapMap2 or HapMap3 and 1kG data do not include Sikhs, the five principal components used for this correction were estimated using our Sikh population sample. After association analyses the genomic control inflation factor, [lambda] was 1.0, so no adjustments were made (Figure 2A).

In addition to analysis of directly genotyped SNPs, we performed imputation using the Impute 2 program[36–38]. Imputation was based on the entire multi-ethnic 1000 genomes reference panel of 28.3 M autosomal SNPs and short in-dels (release v2) in 1,092 individuals from Africa, Asia, Europe and the Americas[39]. Imputed SNPs were analyzed using linear regression for ln-RTL using SNPTEST[36, 38], adjusted for covariates age, gender, BMI, T2D and the five principal component (Supplementary Figure 1). Post-

imputation quality control included removal of SNPs with an information score of ≥0.5 and MAF <5%. The inflation factor, lambda, for imputed SNPs was 1.021 (Figure 2B)

Punjabi Sikh Discovery GWAS

Study design for the RTL GWA study is shown in Figure 1. Clinical characteristics of GWAS (discovery) and Stage 1 (replication) in Punjabi Sikh cohorts are described in (Supplementary Table 2A). Principal components analysis revealed little population structure (Supplementary Figure 1). After quality control as described previously[22], 474,231 directly genotyped SNPs (MAF ≥5%) in 1,616 subjects (842 cases and 774 controls) from 1,850 total subjects were available for association testing after removing samples showing cryptic relatedness through identity by descent sharing. Also, as reported previously in Saxena et al, 2013[22],average inbreeding coefficients (F=0.041 +/− 0.018) were comparable to other Indian populations but higher than European outbred populations. To increase genome coverage, genotypes were imputed for un-typed SNPs and in-dels using the 1kG multi-ethnic reference panel, yielding a total of 6,378,483 variants, and of these 5,904,251 with MAF≥ 5% were analyzed in the current investigation for association analysis. We performed a genome-wide association analysis for ln-RTL using multiple linear regression and adjusting for covariates age, gender, BMI, T2D and the five principal components of ancestry (Figures 3 and 4) (see Methods). No association signals exceeded genome-wide significance, but strong signals ($P<10^{-7}$) were seen at three loci: chr 16q21 (*CSNK2A2*), 5p13.2 (*C5ORF42*), and 5q21.3 (*FER*) (Figure 3, Supplementary Table 4). The association at the *CSNK2A2* locus remained strongly significant when performed separately in diabetic cases (rs74019828 [β±s.e.] −0.40± 0.13), *P*= 0.0019, and controls (rs74019828 [β±s.e.] −0.41± 0.11), *P*=0.00036.

Replication samples and characteristics—Recruitment and diagnostic details of T2D for Stage 1 (Sikh) replication sample are similar as described above for discovery cohort. Clinical and demographical details of the replication dataset used for the Stage 1 and Stage 2 and 3 replication cohorts are provided in the Supplementary Section and Supplementary Table 2B.

Replication through de novo genotyping (Stage 1)—Genotyping of 25 SNPs selected for validation in the Punjabi replication sample (n=2,397) was performed on the BioMark HD MX/HX (Fluidigm) using the Fluidigm 96.96 GT Dynamic Array chip and SNPtype assays (Fluidigm, San Francisco, CA, USA) at Rutgers's Core lab. Upon completion of PCR amplification, end-point data was collected using the Fluidigm BioMarkHD Genetic Analysis instrument. Individual genotype calls and data analysis was performed using the Fluidigm SNP Genotyping Analysis Software v 3.0.2.

Replication through de novo genotyping (Stage 2)—We further selected three SNPs from the two top independent loci identified by meta-analysis of the discovery and Stage 1 populations for replication in the UKCVD cohort (n=2,952). These three SNPs included rs7196068 and rs74019828 for *CSNK2A2* and rs78341307 for *FER* signals, and were genotyped using TaqMan assays designed by Applied Biosystems (Foster City, CA,

USA), and KASPar technology (http://www.kbioscience.co.uk/chemistry/) (KBiosciences, Herts, UK).

Statistical Analysis (Replication Studies)

In each replication sample, genetic association analysis for ln RTL was performed using a linear regression model, with SNPs coded in an additive genetic model and cohort-specific adjustment for covariates. In order to identify RTL association signals common to Punjabi and other ethnic populations, we analyzed the association of 139 top independent signals ($P<10^{-5}$) derived from the discovery cohort and Stage 1 meta-analysis using genotyping data available from four previously published GWA studies in leukocyte telomere length. These studies comprised a total of 10,033 individuals from the UKTWIN[17], Nurses Heart Study (NHS) and Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO)[19], MD Anderson Cancer Controls (MDACC) from lung, bladder, and kidney cancer study[32] as part of Stage 3 replication (Figure 1, Supplementary Table 1).

Meta-analysis—A fixed-effect, inverse variance meta-analysis (as implemented in METAL)[40] was the primary approach used to combine the results for individual studies. A random-effects approach in METASOFT[41] and subset-based approach for heterogeneous traits in ASSET[42] were also performed to allow for heterogeneity between populations under study and between measurements of telomere length.

Results

Association of RTL with cardiometabolic traits in Sikhs

Our results revealed a significant independent association of shorter RTL with T2D and CHD. T2D patients had shorter RTL (1.57±0.26) compared with controls (3.11± 0.19; $P= 4.2 \times 10^{-14}$), and CHD had shorter RTL (1.83±0.16) compared with non-CHD subjects (2.12 ±0.34), ($P= 2.2\times 10^{-3}$). Shorter RTL was also associated with elevated systolic blood pressure (β=−1.03; $P=2\times 10^{-3}$), diastolic blood pressure (β=−0.68; $P=2\times 10^{-3}$), arterial pressure (β=−0.83; $P=5.8\times 10^{-2}$), and MCP-1 (β=−0.05; $P=7\times 10^{-7}$) in analysis adjusted for age, gender, BMI and T2D (Supplementary Table 5). Interestingly, the results were similar when restricted to the control group for all these phenotypes (data not shown). In addition, mean RTL showed a gradual decline from healthy subjects (with no disease) to individuals with T2D and CHD showing respective mean RTL of 2.67± 0.16 in healthy, 2.08±0.14 in CHD, 1.83±0.34 in T2D, and 0.77±0.14 in T2D+CHD individuals showing highly significant difference between healthy subjects versus T2D and CHD patients ($P=3\times 10^{-15}$) (Figure 5). Mean RTL levels were also significantly lower ($P=0.014$) in Sikh males (1.82±0.11) compared with females (2.29±0.15), irrespective of the disease status, consistent with observations in South Asians and other ethnic groups[21, 31, 43, 44] (Supplementary Figure 2).

We performed a genome-wide association analysis for ln-RTL using multiple linear regression and adjusting for covariates age, gender, BMI, T2D and five principal components of ancestry (Figures 3 and 4) (see Methods). No association signals exceeded genome-wide significance, but strong signals ($P<10^{-7}$) were seen at three loci: chr 16q21

(*CSNK2A2*), 5p13.2 (*C5ORF42*), and 5q21.3 (*FER*) (Figure 3, Supplementary Table 4). The association at the *CSNK2A2* locus remained strongly significant when performed separately in diabetic cases (rs74019828 [β±s.e.] −0.40± 0.13), *P*=0.0019, and controls (rs74019828 [β ±s.e.] −0.41± 0.11), *P*=0.00036. A consistent allelic effect was observed for 3 of 6 previously replicated RTL association signals from GWAS[20] (Supplementary Table 6); the results for the most significant SNP for the seventh signal (*RTEL1*) did not pass quality control based on low frequency in the Sikh population. Notably, at the *TERC* locus, we observed moderately significant association, but the allelic effect of the previously identified index SNP (rs10936599) was in the opposite direction (Supplementary Table 6), perhaps due to population differences in linkage disequilibrium (LD) between the marker and causal SNP at this locus. A combined genotype risk score of previously associated variants from the 6 known RTL loci from Codd et al, 2013[20] trended towards association with RTL (β ±s.e.)(−0.80 ±0.45, *P*=0.075). The differences in the LD patterns between the Sikh population and European populations (in which previous index SNPs were identified), suggests the possibility that independent signals at these loci may exist in our population. For instance, using other variants from the same previously associated loci, in 5 out of 9 regions, we observed association effects in Sikhs with *P* values ranging from $10^{-.03}$ to $10^{-.07}$. The most significant association was seen in a variant representing *C5orf42* (~35 Mb) from *TERT* (β=−0.33; $P=2.1\times10^{-7}$) and a variant at *NAF1* (different from the reported[33]) in Sikhs (β=0.26, 8.9×10^{-5}) (Supplementary Figure 3).

Two-stage replication and meta-analysis

We undertook a two-stage replication including T2D case-control samples of Punjabi Sikh ancestry (Stage 1) and genotyping or *in silico* replication in five studies of European ancestry (Stages 2 and 3; Figure 1). The analyses were adjusted for age, gender, BMI, and T2D.

In Stage 1 replication, top SNPs representing 25 putatively novel signals with $P<10^{-4}$ from the discovery GWAS were direct genotyped and analyzed for association with RTL in 2,397 additional Punjabi Sikhs comprising 1,108 T2D cases and 1,289 controls (Supplementary Table 1). The analyses were adjusted for age, gender, BMI, and disease. In discovery and Stage 1 meta-analysis (n=4,013) we identified a novel signal at *CSNK2A2* (16q21) (located in intron 4) and *represented* by rs74019828 to be associated with shorter RTL (β=−0.38 ± SE 0.06; $P=4.5\times10^{-8}$) (Table 1, Figures 4 and 6). Five additional independent signals showed suggestive association ($P<10^{-6}$ to $<10^{-7}$): chromosome 5p13.2 *C5orf42* (rs2098713), 5q21.3 *FER* (rs78869517), and 5q35.2 (an uncharacterized gene) *LOC101928726* (rs244731), 4q34.2 *SPATA4* (rs10004325), 12p11.23 *PTF1BP1* (rs4409879), and 1p31.2 an unknown gene (rs9988609) (Table 1, Supplementary Table 7). We also performed sensitivity analysis by removing T2D and BMI covariates from the model. Our results looked very similar to the previous findings after excluding T2D from model and including only age and gender as covariates (Supplementary Table 8A). In the BMI stratified analyses including age, gender and T2D as covariates, the association signals for RTL for FER at chromosome 5 and chromosome 8 were significantly improved at BMI <25, however, the strongest p value for the SNP association remained consistent at *CSNK2A2* region (Supplementary Table 8A–C, 9A–C).

Next, we analyzed our data using alternative models (dominant, recessive, and co-dominant) for the top variant in *CSNK2A2* and other two variants in chromosome 5 near *TERT* and *FER* regions These analyses did not improve the earlier outcome using additive model (Supplementary Table 10).

For Stage 2 replication, we directly genotyped 3 lead SNPs representing 2 top novel independent (r^2<0.25) association signals *(CSNK2A2* and *FER)* in a UK CVD cohort available with RTL measures on 2,952 MI patients and healthy controls. Combined AIDHS/SDS and UKCVD meta-analysis revealed rs74019828 *(CSNK2A2)* and rs112020835 *(FER)* to be significantly associated with RTL (Figure 6, Supplementary Table 11). However, even though allelic affects were in the same direction, the association of the top variant (rs7401928) did not achieve the GWAS significance threshold (P=3.2×10^{-4}), partly because of lower frequency in UKCVD (Table 1).

In Stage 3 lookup in RTL GWAS of European ancestry, none of our top independent signals could be confirmed. In joint multi-ethnic meta-analysis on individuals from all studies, our top signal in the *CSNK2A2* (rs1393203 used as proxy for rs74019828) (r2 CEU=1.00), did not reach GWAS significance, although there was a moderately significant trend in the same direction (β −0.03 ± SE 0.01; P=2.1×10^{-3}; (Table 1, Supplementary Table 12, and Figure 6) along with significant heterogeneity (I^2=73.41%; P=2×10^{-4}). Using a random effects model, nominal significance was retained (β −0.04 ± SE 0.02); P=0.069). Meta-analysis allowing for heterogeneity of phenotypic measurements (Bhattacharjee et al[42]) revealed a negative z-score comprising our discovery GWAS, Sikh Stage 1 and UKCVD Stage 2 replication studies (β −0.05; P=3.5×10^{-6}) and a positive z-score subset comprising the Stage 3 replication studies (β +0.04; P=1). The gene-based analysis of each GWAS replication dataset could not confirm association of any SNP within 20Kb+/− of the *CSNK2A2* locus with log RTL.

Association of *CSNK2A2* and previously identified variants with cardiometabolic traits in the Sikh population: We first examined the relationship between top SNP in the *CSNK2A2* (rs74019828 or proxy rs1393203) with cardiovascular and metabolic traits. We did not observe any association of these variants with T2D, T2D age of onset or cardiovascular traits (Supplementary Table 15). We tested for association of the lead *CSNK2A2* SNP rs74019828 or proxy rs1393203 and other previously established loci with cardiometabolic traits (LDL-cholesterol, HDL-cholesterol, triglycerides, CHD, systolic and diastolic blood pressure, mean arterial pressure, pulse pressure and T2D) in Sikhs and found marginal association of the telomere shortening allele at rs10936599 *(TERC)* with lower total cholesterol (P=0.0152) and triglycerides (P=0.0188), rs2736100 *(TERT)* with lower blood pressure measures (P=0.0027) and rs9420907 *(OBCF1)* with increased triglycerides (P=0.0225) (Supplementary Table 13). A combined genotype risk score of 6 established shorter telomere length alleles from GWAS[20] showed a trend towards association with increased total cholesterol levels in the entire sample and with increased pulse pressure in non-diabetic controls (Supplementary Table 13).

Discussion

In this study, we report an independent association of shorter RTL in the Punjabi Sikh population with T2D, CHD, and other cardiometabolic traits including inverse association with blood pressure, MCP-1 and positive association with HOMA-B. The Sikh males had significantly shorter mean RTL compared to Sikh females, P=0.014. These findings are in agreement with several previous studies from multiple ethnic populations reporting presence of shorter RTL associated with age-related cardiometabolic diseases including T2D, insulin resistance, MI, and CHD [31,43, 44], containing three previous studies on South Asians [12,21, 45]. We also report discovery of a novel signal represented by an intronic variant in *CSNK2A2* gene associated with shorter RTL (P=4.52×10^{-8}) in Punjabi Sikhs (Supplementary Figure 4). The observed association was confined to South Asian Sikhs and was not replicated in GWAS of European ancestry, and therefore, could be population-specific. The significance of association of the key variant (rs74019828) remained unchanged after controlling for MCP-1, blood pressure, pulse pressure, CHD and T2D in Sikhs, implies independent effect of *CSNK2A2* genetic variants on RTL.

Interestingly, the LD patterns in the region (~1.2Mb) surrounding *CSNK2A2* varied between South Asian Sikhs and HapMap founder populations including Caucasians (CEU), Gujarati Indians (GIH), East Asians (JPT), Yorubans (YRI) (Supplementary Figure 5). The frequency of the susceptibility allele ('A') of our key variant (rs74019828) at the *CSNK2A2* locus was 0.17 in Punjabi Sikhs and 0.20 in GIH, 0.09 in JPT, 0.03–0.07 in CEU, and 0.05 in YRI populations. The difference in allele frequencies between Sikhs and Europeans could have contributed to non-replication of this association. Also, the association of the established *TERC* variant (rs10936599) is in opposite direction in Sikhs from the Europeans (see Supplementary Table 6), suggests a further evidence that there may be population-specific *causal* variant not in LD with these SNPs from these genes. A similar population-specific association was observed in our T2D GWAS in which a novel signal for T2D susceptibility represented by a directly genotyped SNP in the *SGCG* gene (rs9552911, P=1.82×10^{-8}) found in the Sikhs was monomorphic in subjects of European ancestry (Saxena et al)[22]. The Sikh sample comprising discovery and replication datasets (n=4,013) has over 92% power to detect the *CSNK2A2* SNP association to RTL with genome-wide significance (Supplementary Table 14). Furthermore, since the discovery effect size is influenced by the winner's curse, the actual effect size may be smaller, requiring even larger sample sizes for European datasets to observe significant replication.

CSNK2A2 encodes an enzyme, casein kinase II subunit alpha, that phosphorylates a large number of substrates and regulates numerous cellular processes, such as cell cycle progression, apoptosis and transcription[46, 47]. It is affiliated with the members of the shelterin complex involved in chromosome-end protection, telomere regulation, and maintenance[48] Interestingly, the telomeric repeat binding factor 1 (TRF1) serves as a substrate for CSNK2A2, which phosphorylates and initiates its binding to telomeres[48, 49] (Figure 7). Partial knock-down of CSNK2A2 with small interfering RNA resulted in removal of TRF1 from telomeres and degradation of TRF1[49]. CSNK2A2 also influences Wnt signaling via beta-catenin phosphorylation and the PI3-K signaling pathway via the

phosphorylation of Akt. CSNK2A2 also interacts with multiple genes and miRNAs in pathway controlling telomere length and CHD[50].

There are only a few studies connecting CSNK2A2 to telomere length. Our results substantiate the need for a deeper examination and characterization of genetic variation in *CSNK2A2* in conjunction with environmental influences for affecting cardiometabolic risk. As none of the variants within the *CSNK2A2* locus revealed any independent association with CHD in Sikhs, (Supplementary Table 15), it is also possible that the CSNK2A2 does not have any direct role in T2D/CAD pathophysiology.

Limitations of our study include multiple sources of inter-study variability. First, variability across studies due to well-known limitations of telomere length measurement techniques, could have contributed to non-replication in Caucasians; the mean RTL varied widely across studies (Supplementary Table 2B). Second, the presence of population heterogeneity and variation in observed telomere length has resulted in significant heterogeneity in global meta-analysis. Notably, most replication cohorts had 9–20% smokers whereas our Sikh population is a strictly non-smoking population, and gene-environment interactions with smoking may have obscured true association[11, 51, 52] signals (Supplementary Table 2B). Third, a significant proportion of gender bias existed in most Caucasian replication cohorts including 91% females in UKTWIN, 100% females in NHS, 100% males in PLCO and UKCVD, and 72% males in MDACC compared to AIDHS/SDS (55% males) which could have partly contributed in non-replication as shorter RTL correlates remarkably with male gender in Sikhs as well as in previous studies[53] (see Supplementary Table 2B). Finally, performing analysis from findings derived from T2D case control study with cohorts including prostate cancer survivors and post-menopausal women etc. could have resulted in replication bias in addition to other factors. For example, shorter telomeres in cancer cases due to disease-related secondary effects may reduce power to detect a genetic effect. Differences in LD between marker and causal SNPs in Sikhs and non-Sikh replication cohorts has certainly contributed to non-replication in the European sample and to the limited association of 6 European SNPs in Sikh populations. However, despite these limitations, our original association results were replicated in an independent validation cohort of Punjabi Sikhs

Conclusions

We report association of shorter RTL with T2D and cardiometabolic risk in Punjabi Sikhs from South Asia. Our GWAS and meta-analysis identified a new signal within the *CSNK2A2* gene associated with RTL in South Asian Sikhs. CSNK2A2 phosphorylates TRF1 which initiates and regulates its binding to telomere. *CSNK2A2* also interacts with multiple genes and miRNAs in pathways controlling telomere length and cardiovascular disease. Thus far, no other GWAS has been conducted for telomere length and association of genome-wide variants with T2D and cardiovascular traits in conjunction with shorter RTL in South Asian populations. Therefore, future confirmation of our findings in other South Asian populations will be necessary to evaluate this population-specific association signal. Also, targeted resequencing *of CSNK2A2* in Sikhs and other multi-ethnic populations and functional studies may provide clinically important insights into the causal relationship

of possible rare variants in *CSNK2A2* with telomere attrition and cardiometabolic risk in diabetes and other age-dependent chronic disorders.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Authors thank all the participants of AIDHS/SDS and are grateful for their contribution in this study. We also thank Ms. Anuradha Subramanian for her technical and administrative help in the manuscript preparation. Study-specific acknowledgments are provided in the Supplementary Online Material.

Funding Sources: This work was partly supported by NIH grants -R01DK082766 funded by the National Institute of Diabetes and Digestive and Kidney Diseases and NOT-HG-11-009 funded by National Genome Research Institute, and VPR Bridge grant from University of Oklahoma Health Sciences Center, Oklahoma City, USA. The genotyping and analysis for UKCVD was supported by the British Heart Foundation (PG08/008). Study-specific funding sources for the replication studies are provided in the Supplementary Online Material.

References

1. Samani NJ, Boultby R, Butler R, Thompson JR, Goodall AH. Telomere shortening in atherosclerosis. Lancet. 2001; 358:472–473. [PubMed: 11513915]
2. Valdes AM, Richards JB, Gardner JP, Swaminathan R, Kimura M, Xiaobin L, et al. Telomere length in leukocytes correlates with bone mineral density and is shorter in women with osteoporosis. Osteoporos Tntl. 2007; 18:1203–1210.
3. Thomas P, NJ OC, Fenech M. Telomere length in white blood cells, buccal cells and brain tissue and its variation with ageing and alzheimer's disease. Mech Ageing Dev. 2008; 129:183–190. [PubMed: 18242664]
4. Tentolouris N, Nzietchueng R, Cattan V, Poitevin G, Lacolley P, Papazafiropoulou A, et al. White blood cells telomere length is shorter in males with type 2 diabetes and microalbuminuria. Diab care. 2007; 30:2909–2915.
5. Brouilette SW, Moore JS, McMahon AD, Thompson JR, Ford I, Shepherd J, et al. Telomere length, risk of coronary heart disease, and statin treatment in the west of scotland primary prevention study: A nested case-control study. Lancet. 2007; 369:107–114. [PubMed: 17223473]
6. Sanders JL, Newman AB. Telomere length in epidemiology: A biomarker of aging, age-related disease, both, or neither? Epidemiol Rev. 2013
7. Mather KA, Jorm AF, Parslow RA, Christensen H. Is telomere length a biomarker of aging? A review. The J Gerontol. Series A, Biological sciences and medical sciences. 2011; 66:202–213.
8. Vaziri H, Schachter F, Uchida I, Wei L, Zhu X, Effros R, et al. Loss of telomeric DNA during aging of normal and trisomy 21 human lymphocytes. Am J Hum Genet. 1993; 52:661–667. [PubMed: 8460632]
9. Okuda K, Bardeguez A, Gardner JP, Rodriguez P, Ganesh V, Kimura M, et al. Telomere length in the newborn. Pediatric research. 2002; 52:377–381. [PubMed: 12193671]
10. Carroll JE, Diez-Roux AV, Adler NE, Seeman TE. Socioeconomic factors and leukocyte telomere length in a multi-ethnic sample: Findings from the multi-ethnic study of atherosclerosis (mesa). Brain Behav Immun. 2013; 28:108–114. [PubMed: 23142704]
11. Diez Roux AV, Ranjit N, Jenny NS, Shea S, Cushman M, Fitzpatrick A, et al. Race/ethnicity and telomere length in the multi-ethnic study of atherosclerosis. Aging cell. 2009; 8:251–257. [PubMed: 19302371]
12. Adaikalakoteswari A, Balasubramanyam M, Mohan V. Telomere shortening occurs in asian indian type 2 diabetic patients. Diabetic medicine : J. Brit Diabet Assoc. 2005; 22:1151–1156.

13. Salpea KD, Talmud PJ, Cooper JA, Maubaret CG, Stephens JW, Abelak K, et al. Association of telomere length with type 2 diabetes, oxidative stress and ucp2 gene variation. Atherosclerosis. 2010; 209:42–50. [PubMed: 19889414]
14. Jeanclos E, Schork NJ, Kyvik KO, Kimura M, Skurnick JH, Aviv A. Telomere length inversely correlates with pulse pressure and is highly familial. Hypertension. 2000; 36:195–200. [PubMed: 10948077]
15. Andrew T, Aviv A, Falchi M, Surdulescu GL, Gardner JP, Lu X, et al. Mapping genetic loci that determine leukocyte telomere length in a large sample of unselected female sibling pairs. Am J Hum Genet. 2006; 78:480–486. [PubMed: 16400618]
16. Aviv A. Genetics of leukocyte telomere length and its role in atherosclerosis. Mutat Res. 2012; 730:68–74. [PubMed: 21600224]
17. Mangino M, Richards JB, Soranzo N, Zhai G, Aviv A, Valdes AM, et al. A genome-wide association study identifies a novel locus on chromosome 18q12.2 influencing white cell telomere length. J Med Genet. 2009; 46:451–454. [PubMed: 19359265]
18. Levy D, Neuhausen SL, Hunt SC, Kimura M, Hwang SJ, Chen W, et al. Genome-wide association identifies obfc1 as a locus involved in human leukocyte telomere biology. Proc Natl Acad Sc (USA). 2010; 107:9293–9298. [PubMed: 20421499]
19. Prescott J, Kraft P, Chasman DI, Savage SA, Mirabello L, Berndt SI, et al. Genome-wide association study of relative telomere length. PloS one. 2011; 6:e19635. [PubMed: 21573004]
20. Codd V, Mangino M, van der Harst P, Braund PS, Kaiser M, Beveridge AJ, et al. Common variants near terc are associated with mean telomere length. Nat Genet. 2010; 42:197–199. [PubMed: 20139977]
21. Harte AL, da Silva NF, Miller MA, Cappuccio FP, Kelly A, O'Hare JP, et al. Telomere length attrition, a marker of biological senescence, is inversely correlated with triglycerides and cholesterol in south asian males with type 2 diabetes mellitus. Exprt Diabet Res. 2012; 2012:895185.
22. Saxena R, Saleheen D, Been LF, Garavito ML, Braun T, Bjonnes A, et al. Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in sikhs of punjabi origin from india. Diabetes. 2013; 62:1746–1755. [PubMed: 23300278]
23. Diagnosis and classification of diabetes mellitus. Diab care. 2004; 27(Suppl 1):S5–S10.
24. Sanghera DK, Ortega L, Han S, Singh J, Ralhan SK, Wander GS, et al. Impact of nine common type 2 diabetes risk polymorphisms in asian indian sikhs: Pparg2 (pro12ala), igf2bp2, tcf7l2 and fto variants confer a significant risk. BMC Med Genet. 2008; 9:59. [PubMed: 18598350]
25. Schierer A, Been LF, Ralhan S, Wander GS, Aston CE, Sanghera DK. Genetic variation in cholesterol ester transfer protein, serum cetp activity, and coronary artery disease risk in asian indian diabetic cohort. Pharmacogenet and Genom. 22:95–104.
26. Sanghera DK, Been LF, Ralhan S, Wander GS, Mehra NK, Singh JR, et al. Genome-wide linkage scan to identify loci associated with type 2 diabetes and blood lipid phenotypes in the sikh diabetes study. PLoS One. 6:e21188. [PubMed: 21698157]
27. Sanghera DK, Been L, Ortega L, Wander GS, Mehra NK, Aston CE, et al. Testing the association of novel meta-analysis-derived diabetes risk genes with type ii diabetes and related metabolic traits in asian indian sikhs. J Hum Genet. 2009; 54:162–168. [PubMed: 19247373]
28. Braun TRBL, Blackett PR, Sanghera DK. Vitamin d deficiency and cardio-metabolic risk in a north indian community with highly prevalent type 2 diabetes. J Diabetes Metab. 2012; 3:213.
29. Sanghera DK, Nath SK, Ortega L, Gambarelli M, Kim-Howard X, Singh JR, et al. Tcf7l2 polymorphisms are associated with type 2 diabetes in khatri sikhs from north india: Genetic variation affects lipid levels. Ann Hum Genet. 2008; 72:499–509. [PubMed: 18397358]
30. Cawthon RM. Telomere measurement by quantitative pcr. Nucleic Ac. Res. 2002; 30:e47.
31. Salpea KD, Nicaud V, Tiret L, Talmud PJ, Humphries SE. The association of telomere length with paternal history of premature myocardial infarction in the european atherosclerosis research study ii. J Mol Med (Berl). 2008; 86:815–824. [PubMed: 18414821]
32. Gu J, Chen M, Shete S, Amos CI, Kamat A, Ye Y, et al. A genome-wide association study identifies a locus on chromosome 14q21 as a predictor of leukocyte telomere length and as a

marker of susceptibility for bladder cancer. Cancer Prev Res (Phila). 2011; 4:514–521. [PubMed: 21460395]

33. Codd V, Nelson CP, Albrecht E, Mangino M, Deelen J, Buxton JL, et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat Genet. 2013; 45:422–427. 427e421–427e422. [PubMed: 23535734]
34. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. Plink: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]
35. Nalls MA, Simon-Sanchez J, Gibbs JR, Paisan-Ruiz C, Bras JT, Tanaka T, et al. Measures of autozygosity in decline: Globalization, urbanization, and its implications for medical genetics. PLoS Genet. 2009; 5:e1000415. [PubMed: 19282984]
36. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007; 39:906–913. [PubMed: 17572673]
37. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. Genes, Genom, Genet. 2011; 1:457–470.
38. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev. Genet. 11:499–511. [PubMed: 20517342]
39. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]
40. Willer CJ, Li Y, Abecasis GR. Metal: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 26:2190–2191. [PubMed: 20616382]
41. Han B, Eskin E. Interpreting meta-analyses of genome-wide association studies. PLoS Genet. 2012; 8:e1002555. [PubMed: 22396665]
42. Bhattacharjee S, Rajaraman P, Jacobs KB, Wheeler WA, Melin BS, Hartge P, et al. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. Am J Hum Genet. 2012; 90:821–835. [PubMed: 22560090]
43. Wang YY, Chen AF, Wang HZ, Xie LY, Sui KX, Zhang QY. Association of shorter mean telomere length with large artery stiffness in patients with coronary heart disease. Aging Male. 2011; 14:27–32. [PubMed: 21067315]
44. Zee RY, Castonguay AJ, Barton NS, Germer S, Martin M. Mean leukocyte telomere length shortening and type 2 diabetes mellitus: A case-control study. Translational research :J Lab Clin Med. 2010; 155:166–169.
45. Mukherjee M, Brouilette S, Stevens S, Shetty KR, Samani NJ. Association of shorter telomeres with coronary artery disease in indian subjects. Heart. 2009; 95:669–673. [PubMed: 19028761]
46. Ruzzene M, Penzo D, Pinna LA. Protein kinase ck2 inhibitor 4,5,6,7-tetrabromobenzotriazole (tbb) induces apoptosis and caspase-dependent degradation of haematopoietic lineage cell-specific protein 1 (hs1) in jurkat cells. Biochem J. 2002; 364:41–47. [PubMed: 11988074]
47. Litchfield DW. Protein kinase ck2: Structure, regulation and role in cellular decisions of life and death. Biochem J. 2003; 369:1–15. [PubMed: 12396231]
48. Oganesian L, Karlseder J. Telomeric armor: The layers of end protection. J Cell Sci. 2009; 122:4013–4025. [PubMed: 19910493]
49. Kim MK, Kang MR, Nam HW, Bae YS, Kim YS, Chung IK. Regulation of telomeric repeat binding factor 1 binding to telomeres by casein kinase 2-mediated phosphorylation. J Biol Chem. 2008; 283:14144–14152. [PubMed: 18347021]
50. Gao Y, Wang HY. Casein kinase 2 is activated and essential for wnt/beta-catenin signaling. J Biol Chem. 2006; 281:18394–18400. [PubMed: 16672224]
51. Sanghera DK, Bhatti JS, Bhatti GK, Ralhan SK, Wander GS, Singh JR, et al. The khatri sikh diabetes study (sds): Study design, methodology, sample collection, and initial results. Hum Biol. 2006; 78:43–63. [PubMed: 16900881]
52. Yang Z, Ye J, Li C, Zhou D, Shen Q, Wu J, et al. Drug addiction is associated with leukocyte telomere length. Sci Rep. 2013; 3:1542. [PubMed: 23528991]

53. Drury SS, Theall K, Gleason MM, Smyke AT, De Vivo I, Wong JY, et al. Telomere length and early severe social deprivation: Linking early adversity and cellular aging. Mol Psychol. 2012; 17:719–727.

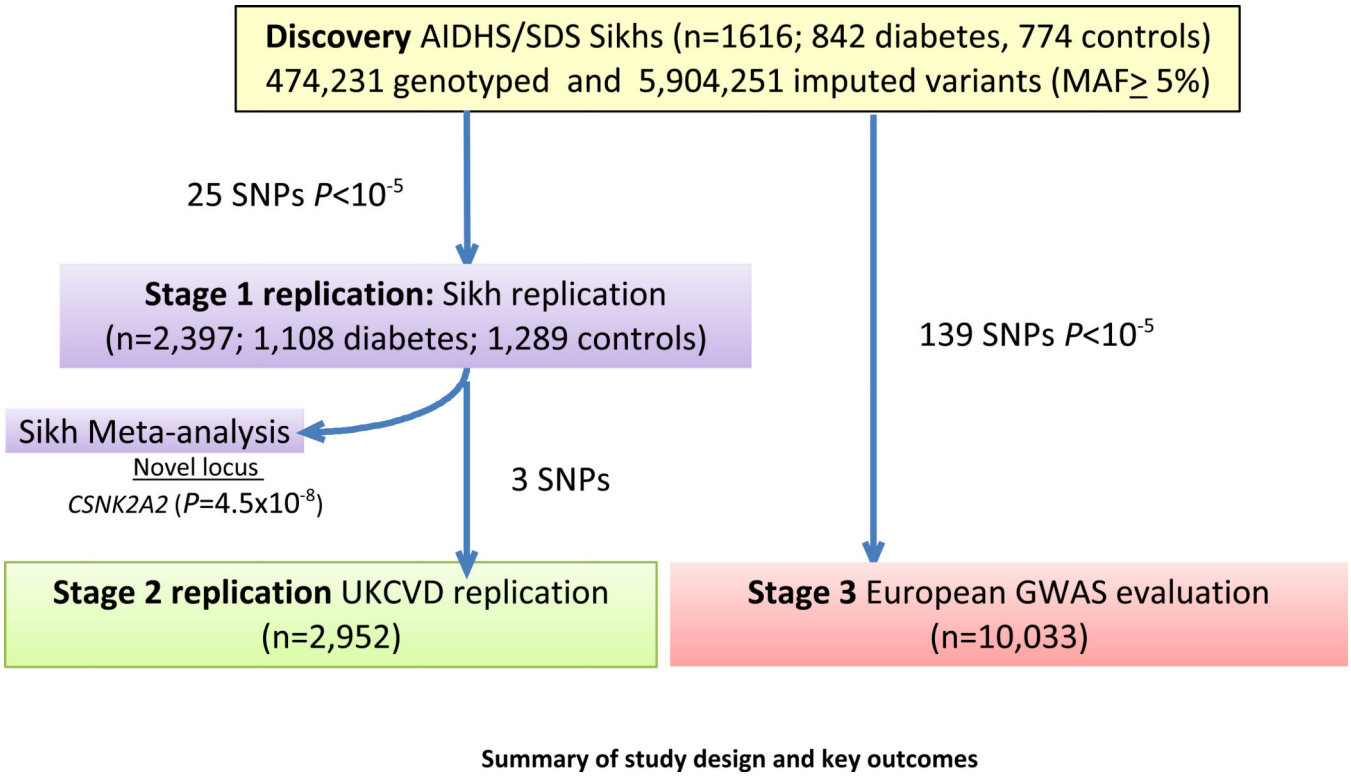

Figure 1.
Summary of study design and key outcomes

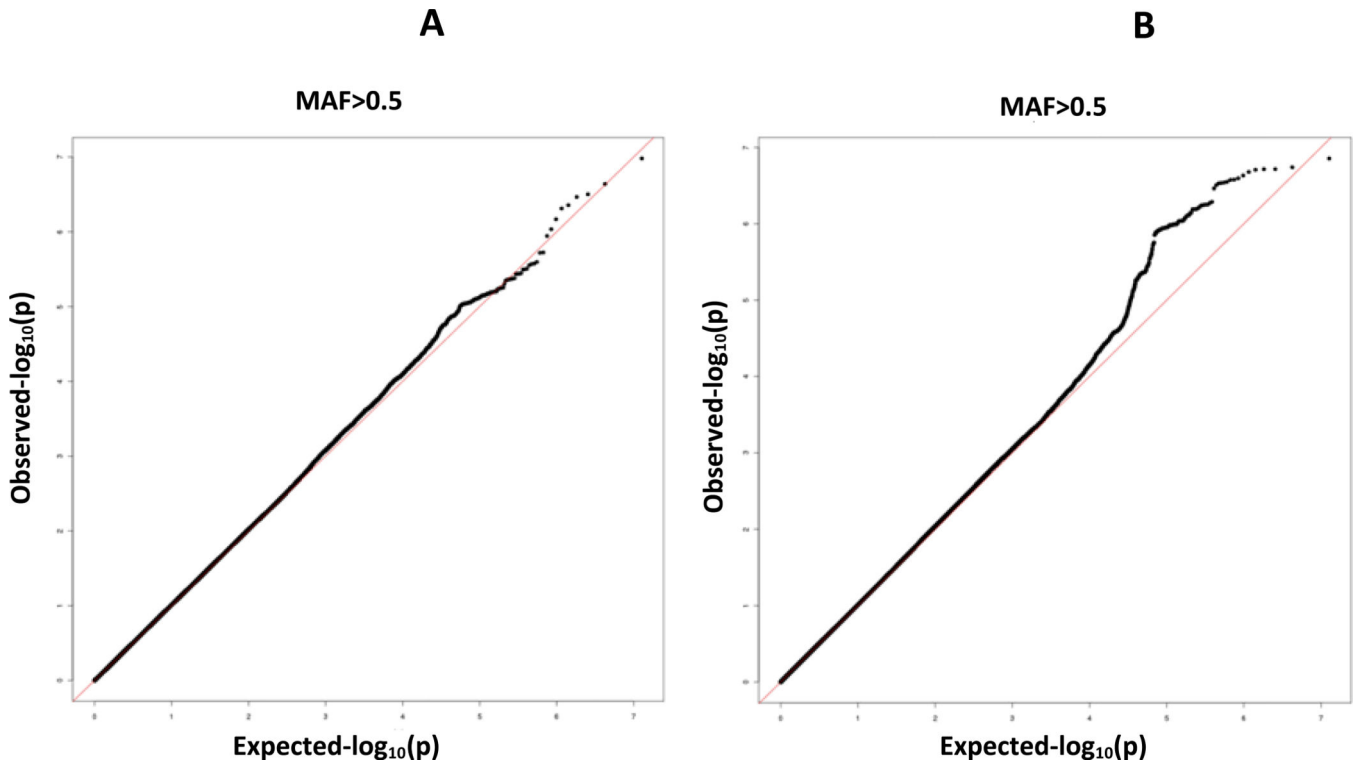

Figure 2.

(A) QQ plot of RTL GWAS of the Sikh discovery cohort after quality control of directly genotyped (474,231), and **(B)** imputed variants 5,904,251 (MAF≥5%) from the 1kG reference panel of 1092 world-wide subjects

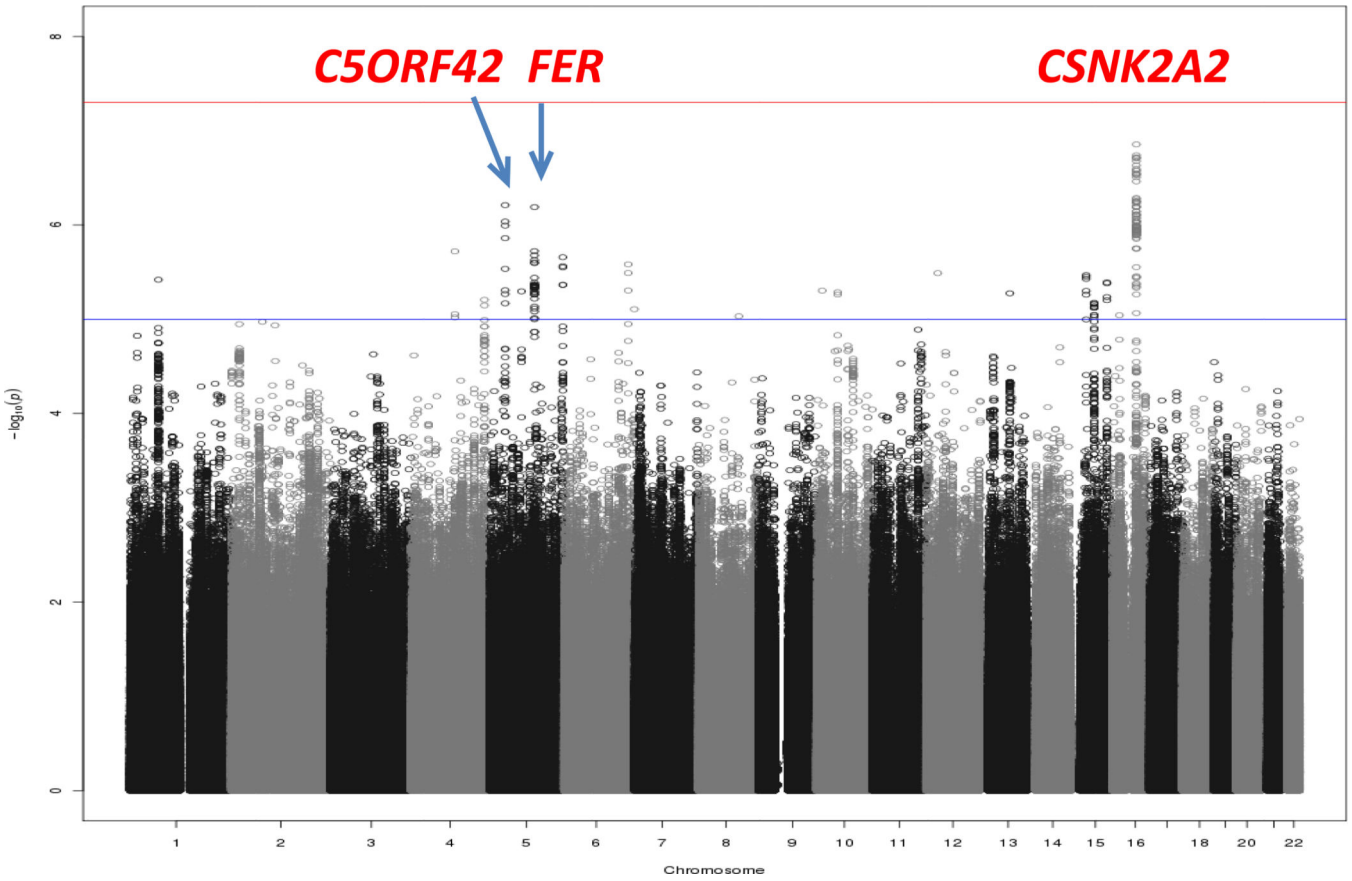

Figure 3.

Manhattan plot of primary RTL GWAS analysis of the Sikh discovery cohort using directly genotyped and imputed SNPs on X axis and −log10 p-value on Y axis.

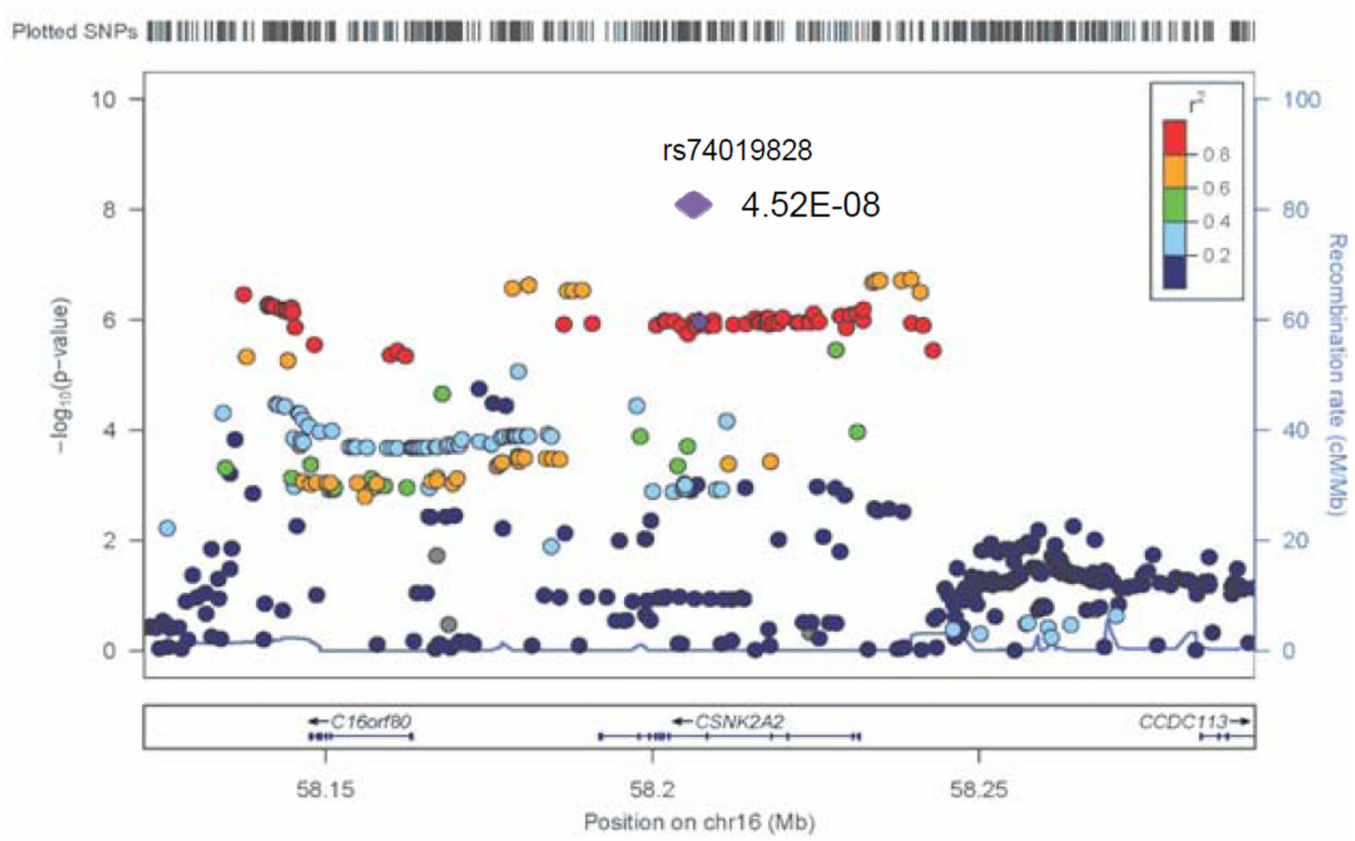

Figure 4.

Regional association plot for a novel GWAS locus associated with RTL represented by rs74019828 in the *CSNK2A2* gene in discovery GWAS and in combined analysis of GWAS and replication studies (meta-analysis result shown as purple diamond).

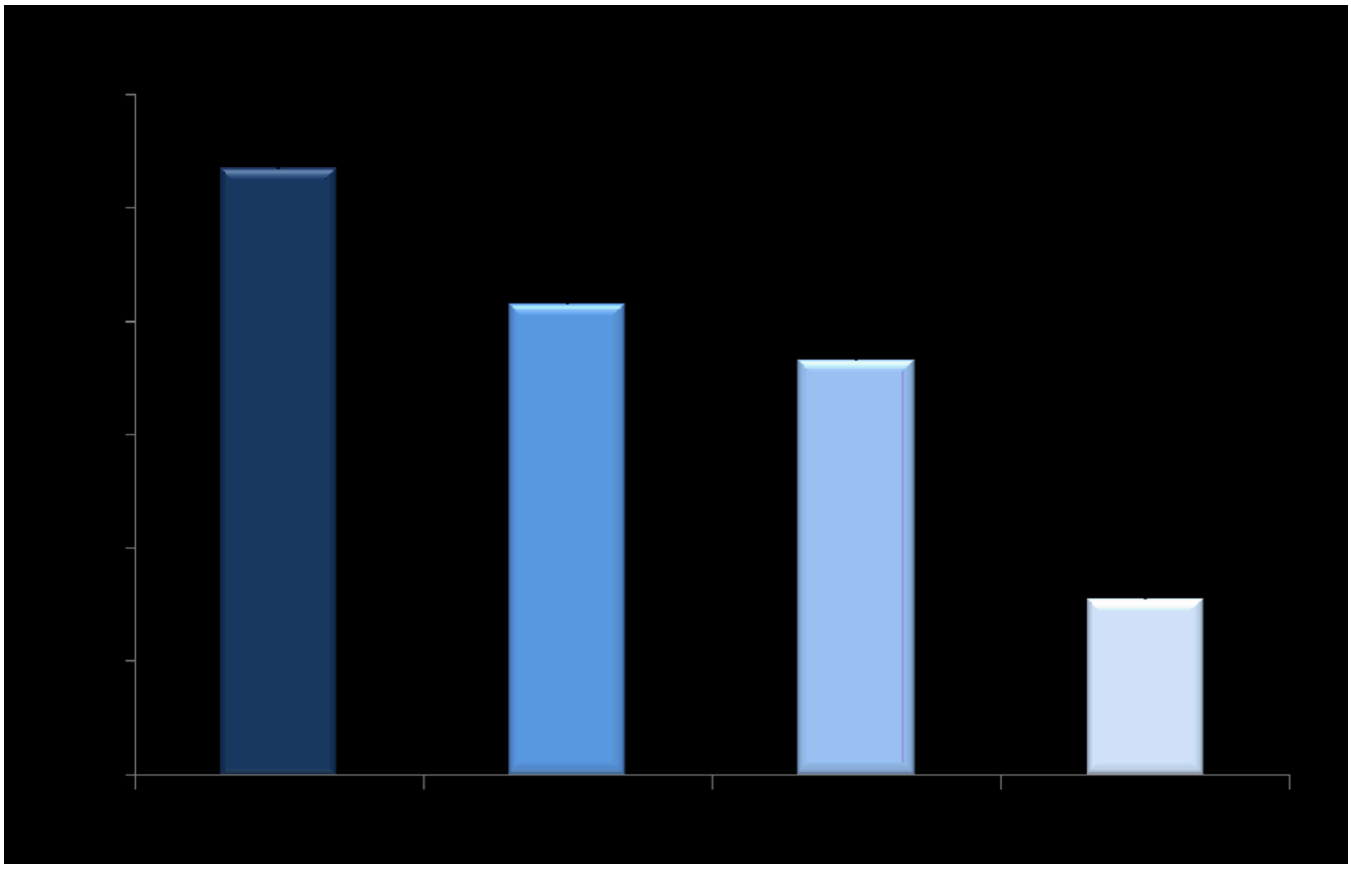

Figure 5.
Distribution of relative telomere length (RTL) in healthy controls, patients with coronary heart disease (CHD) and type 2 diabetes (T2D).

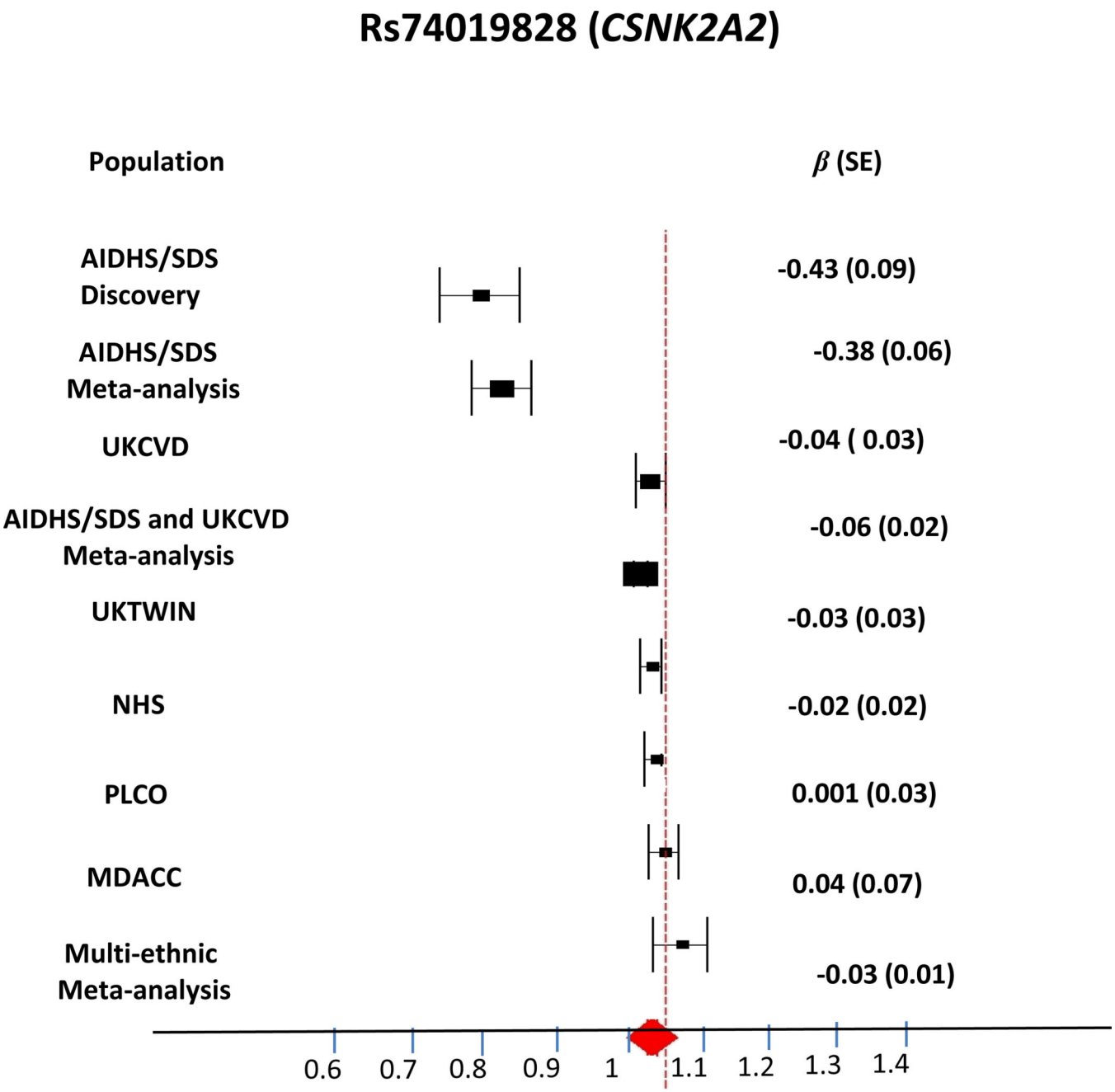

Figure 6.

Forest plot showing the association of lead SNP in the *CSNK2A2* (rs74019828) gene with RTL. The meta-analysis of Sikhs, Europeans and Multiethnic studies are shown. Meta-analysis in Sikhs shows a significant association of rs74019828 with RTL (β 0.38±0.06, $P = 4.52 \times 10^{-08}$).

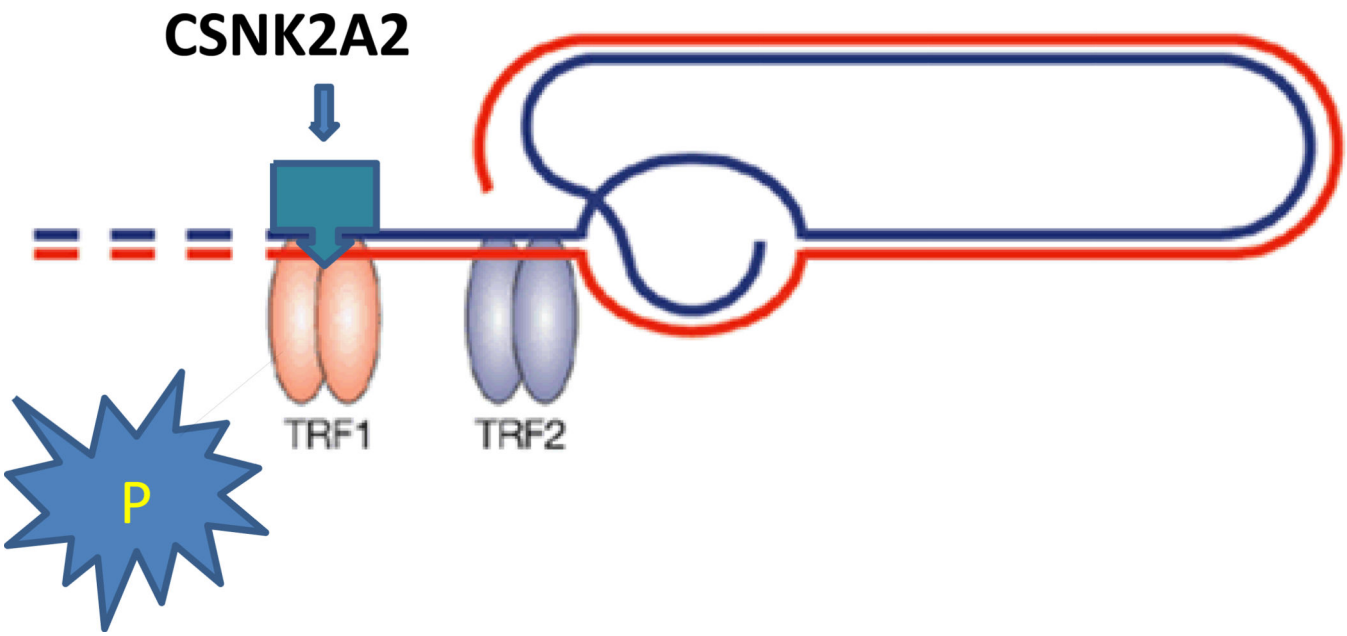

Figure 7.
Overview of the telomere assembly showing binding of CSNK2A2 with Telomere Repeat Binding Factor 1 (TERF1) as substrate.

Table 1

GWAS, replication, and meta-analysis results of RTL loci identified in Punjabi Sikhs

SNP	Chromosome Position	Nearest gene	Effect/ Other Allele	EAF SIKH CEU	Punjabi Sikh Discovery GWAS β,SE *P* value	Punjabi Sikh Replication (Stage 1) β,SE *P* value	Sikh Meta-Analysis β,SE *P* value	Sikh-UK CVDS Meta-Analysis β,SE *P* value	Multi-Ethnic Meta-Analysis β,SE *P* value
rs74019828	16 58209274	***CSNK2A2***	A/G	0.16 0.03	−0.43,0.09 1.02×10^{-06}	−0.27, 0.13 2.62×10^{-02}	−0.38,0.06 **4.52×10^{-08}**	−0.06, 0.02 3.20×10^{-04}	−0.03, 0.01 2.10×10^{-03}
rs2098713	5 37144574	***C5orf42***	T/C	0.47 0.35	−0.33,0.07 6.15×10^{-07}	−0.15,0.09 1.26×10^{-01}	−0.25,0.05 **3.35×10^{-06}**	N/A	−0.01,0.01 4.93×10^{-01}
rs10004325	4 177106289	***SPATA4***	G/C	0.84 0.75	−0.42,0.09 6.21×10^{-06}	−0.04,0.13 7.27×10^{-01}	−0.31,0.08 **1.95×10^{-05}**	N/A	−0.02,0.01 6.00×10^{-02}
rs78869517	5 108050074	***FER***	C/G	0.92 0.94	−0.59,0.13 2.10×10^{-06}	0.09,0.17 5.89×10^{-01}	−0.41,0.09 **2.78×10^{-05}**	0.02,0.02 7.80×10^{-02}	N/A
rs4409879	12 27739993	***PTFIBP1***	A/G	0.17 0.32	−0.45,0.09 3.24×10^{-06}	−0.08,0.13 5.31×10^{-01}	−0.31,0.07 **3.36×10^{-05}**	N/A	−0.01,0.001 1.37×10^{-01}
rs244731	5 176539679	***unknown***	A/G	0.23 0.28	0.38,0.08 2.20×10^{-06}	0.07,0.11 4.88×10^{-01}	0.25,0.06 **4.82×10^{-05}**	N/A	N/A
rs9988609	1 69315228	***unknown***	A/G	0.10 0.11	−0.47,0.11 2.67×10^{-05}	−0.22,0.16 1.72×10^{-01}	−0.36,0.08 **5.78×10^{-05}**	N/A	N/A

All P values are two sided. CEU- Euro-Caucasians, EAF-Effect Allele Frequency, SE-Standard Error

*data from rs1393203 is a proxy for rs74019828