nature genetics

Common variants near *TERC* are associated with mean telomere length

Veryan Codd[1,11], Massimo Mangino[2,11], Pim van der Harst[1,3,11], Peter S Braund[1], Michael Kaiser[1], Alan J Beveridge[1], Suzanne Rafelt[1], Jasbir Moore[1], Chris Nelson[1], Nicole Soranzo[2,4], Guangju Zhai[2], Ana M Valdes[2], Hannah Blackburn[4], Irene Mateo Leach[3], Rudolf A de Boer[3], Masayuki Kimura[5], Abraham Aviv[5], Wellcome Trust Case Control Consortium[10], Alison H Goodall[1], Willem Ouwehand[6], Dirk J van Veldhuisen[3], Wiek H van Gilst[3], Gerjan Navis[7], Paul R Burton[8], Martin D Tobin[8], Alistair S Hall[9], John R Thompson[8], Tim Spector[2,11] & Nilesh J Samani[1,11]

We conducted genome-wide association analyses of mean leukocyte telomere length in 2,917 individuals, with follow-up replication in 9,492 individuals. We identified an association with telomere length on 3q26 (rs12696304, combined $P = 3.72 \times 10^{-14}$) at a locus that includes *TERC*, which encodes the telomerase RNA component. Each copy of the minor allele of rs12696304 was associated with an ~75-base-pair reduction in mean telomere length, equivalent to ~3.6 years of age-related telomere-length attrition.

© 2010 Nature America, Inc. All rights reserved.

Telomeres are structures at the ends of eukaryotic chromosomes that are made up of a simple repetitive sequence (in humans, TTAGGG) and are involved in maintaining genomic stability and regulating cellular proliferation[1]. Telomere length is important in determining telomere function. In somatic cells, telomeres progressively shortens with each mitotic division because of the inability of DNA polymerase to fully replicate the 3′ end of the DNA strand. Cellular senescence and subsequent cell death often occur when the mean telomere length reaches a critical value[1]. Shorter mean leukocyte telomere length has been shown to be associated with risk of several age-related diseases and has been proposed as a marker of biological aging. Maintenance of telomere length is required in certain cell types (such as germ cells) and is effected by telomerase, a ribonucleoprotein consisting of a reverse transcriptase (TERT) and an RNA template (TERC) that facilitates addition of the telomere repeat sequence. Telomerase reactivation and telomere length maintenance also contribute to the pathogenesis of several cancers[1].

Telomere length has a strong genetic determination, with heritability estimates ranging from 44% to 80%[2,3]. Quantitative trait linkage (QTL) studies have mapped putative loci for telomere length to human chromosomes 3p26.1, 10q26.13, 12q12.22 and 14q23.2 (refs. 3–5). A recent genome-wide association study (GWAS) identified associations of two SNPs on chromosome 18q12.2 with telomere length, although the associations were not at a genome-wide significant level[6]. To identify additional variants that affect telomere length, we undertook genome-wide association analyses in two large European cohorts followed by replication of promising signals in three further European cohorts.

The discovery cohorts comprised 1,487 individuals with coronary artery disease from the British Heart Foundation Family Heart Study (BHF-FHS)[7] and 1,430 United Kingdom Blood Service donors (UKBS)[8] for whom there was genome-wide SNP genotype data available that was generated using the Affymetrix 500K array as part of the Wellcome Trust Case-Control Consortium (WTCCC) study[8]. Further details of the cohorts and genotyping procedures are given in the **Supplementary Methods** and **Supplementary Table 1**. Mean leukocyte telomere length was measured using a quantitative PCR–based technique[9] that expresses mean telomere length as a ratio (*T*/*S*) of telomere repeat length (*T*) to the copy number (*S*) of a single-copy gene, *36B4*, within each sample (**Supplementary Methods**). The *T*/*S* ratios were distributed normally (**Supplementary Fig. 1a**) and showed the expected age-related attrition in telomere length (**Supplementary Fig. 1b**) in both cohorts.

We analyzed the association of *T*/*S* ratio, adjusted for age and gender, with genotype individually in the BHF-FHS (**Supplementary Table 2a**) and UKBS (**Supplementary Table 2b**) cohorts and also in a combined analysis of the two cohorts. The quantile-quantile plots for each cohort are shown in **Supplementary Figure 1c** and the power to detect associations in **Supplementary Figure 1d**. The genomic inflation control factors for the BHF-FHS and UKBS analyses were 1.02 and 0.99, respectively. We screened the results from the combined analysis for SNPs that showed concordance in their results from the individual analyses (that is, were at least nominally significant ($P < 0.05$) in both cohorts, with β coefficients having the same direction in each). These criteria identified 180 SNPs at $P < 1 \times 10^{-3}$ and 24 SNPs at $P < 1 \times 10^{-4}$. SNPs achieving a combined $P < 5 \times 10^{-5}$ are shown in **Supplementary Table 2c**. Notably, neither the previously reported SNPs on chromosome 18q12.2 (ref. 6) nor SNPs at any of the loci identified by QTL analysis[3–5] showed a consistent association with telomere length in our cohorts (**Supplementary Table 3a**). We set a pragmatic significance threshold of $P < 1 \times 10^{-5}$ in the combined analysis to determine which SNPs to take forward for replication. This identified two SNPs: rs610160 (chr. 11q22) and rs10511887 (chr. 9p21.1) (**Table 1**). However, one SNP, rs16847897 (chr. 3q26), that fell marginally outside this threshold (combined $P = 1.03 \times 10^{-5}$) was the lead SNP from a locus that includes

[1]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK. [2]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [3]Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [4]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. [5]The Center of Human Development and Aging, New Jersey Medical School, University of Medicine and Dentistry of New Jersey, Newark, New Jersey, USA. [6]Department of Hematology, University of Cambridge, Cambridge, UK. [7]Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [8]Departments of Health Sciences and Genetics, University of Leicester, Leicester, UK. [9]Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK. [10]A full list of members is provided is provided in the **Supplementary Note**. [11]These authors contributed equally to this work. Correspondence should be addressed to N.J.S. (njs@le.ac.uk).

Received 25 August 2009; accepted 12 January 2010; published online 7 February 2010; corrected after print 16 April 2010 (details online); doi:10.1038/ng.532

© 2010 Nature America, Inc. All rights reserved.

Table 1 Telomere length association results from the GWAS studies and in replication studies for the 3q26.2, 9p21.1 and 11q22.3 loci

			BHF-FHS GWAS		UKBS GWAS		BHF-FHS+UKBS		GRAPHIC		TwinsUK			PREVEND		
Chr.	SNP	Minor allele	MAF	β (s.e.m.)	*P* value	MAF	β (s.e.m.)	*P* value	MAF	β (s.e.m.)	*P* value	MAF	β (s.e.m.)	*P* value	MAF	β (s.e.m.)	*P* value
3	rs12696304	G	0.26	−0.039 (0.009)	9.73×10^{-5}	0.26	−0.022 (0.013)	8.34×10^{-2}	0.26	−0.030 (0.008)	9.33×10^{-5}	0.26	−0.028 (0.009)	1.39×10^{-3}	0.25	−0.075 (0.019)	8.13×10^{-5}	0.30	−0.030 (0.008)	3.21×10^{-4}
3	rs16847897	C	0.26	−0.033 (0.009)	1.25×10^{-4}	0.27	−0.035 (0.013)	6.64×10^{-3}	0.26	−0.034 (0.008)	1.03×10^{-5}	0.26	−0.029 (0.009)	9.34×10^{-4}	0.26	−0.064 (0.019)	6.44×10^{-4}	0.30	−0.024 (0.009)	6.40×10^{-3}
9	rs10511887	G	0.24	−0.037 (0.011)	7.41×10^{-4}	0.24	−0.052 (0.017)	2.10×10^{-3}	0.24	−0.044 (0.010)	8.86×10^{-6}	0.25	−0.001 (0.009)	0.91		—			—	
11	rs610160	C	0.13	0.030 (0.011)	8.89×10^{-3}	0.12	0.064 (0.017)	2.03×10^{-4}	0.13	0.046 (0.010)	7.05×10^{-6}	0.11	−0.007 (0.013)	0.57		—			—	

The association findings with mean leukocyte telomere length of the four SNPs identified in the combined analyses of the British Heart Foundation Family Heart study (BHF-FHS) and the United Kingdom Blood Service (UKBS) donors genome-wide association studies (GWAS) and taken through to replication in the GRAPHIC study and subsequently for rs12696304 and rs16847897 into the TwinsUK and PREVEND studies. The minor allele frequency (MAF), beta coefficient (*β*), standard error (s.e.m.) and *P* value for association for each cohort individually and for the combined analysis of the BHF-FHS and UKBS subjects are shown. SNP associations were analyzed using an additive model. The *β* coefficients equate to the mean change in telomere length for the particular measure of telomere length (T/S ratio in BHF-FHS, UKBS, GRAPHIC and PREVEND and Southern blot determined telomere length in TwinsUK) per copy of the minor allele for the specific SNP. A negative *β* indicates a shorter mean telomere in those carrying the minor allele.

TERC. Another SNP in the locus, rs12696304, with a relatively weak linkage disequilibrium (LD) ($r^2 = 0.49$) to rs16847897 also showed moderate association (combined $P = 9.33 \times 10^{-5}$). Owing to the presence of a candidate gene at this locus, these two SNPs were also taken forward for replication analyses (**Table 1**). None of the SNPs taken forward showed any association with coronary artery disease.

Details of the replication cohorts are given in the **Supplementary Methods** and **Supplementary Table 1**. Initial replication was performed in 2,020 subjects of the GRAPHIC study[10]. Leukocyte mean telomere length was determined in these subjects using the same PCR technique used in the discovery cohorts. Neither rs610160 nor rs10511887 showed any evidence of replication. However, both SNPs on chromosome 3q26 showed significant associations in the same direction as in the discovery cohorts (**Table 1**). Next, we analyzed data on leukocyte mean telomere length and genotypes for rs16847897 and rs12696304 obtained in 3,256 subjects from the TwinsUK study[11]. Mean telomere length in this cohort was measured by DNA blotting (**Supplementary Methods**). Despite this difference in measurement method, there was still a significant and similar association of both SNPs with telomere length (**Table 1** and **Supplementary Table 3b**). Finally, we examined the association of the 3q26 SNPs with telomere length in 4,216 individuals from the PREVEND study[12]. We measured telomere length in these subjects using a variation of the PCR-based method used in the discovery and GRAPHIC cohorts (**Supplementary Methods**). Both rs16847897 and rs12696304 again showed significant associations with telomere length (**Table 1**). To assess the combined evidence for the association of the 3q26 locus with telomere length, we used Fisher's method (**Supplementary Methods**). We also *z*-transformed the individual telomere length measurements in each study to obtain comparable results (**Supplementary Table 4a**) and performed a meta-analysis using METAL (**Supplementary Methods** and URLs). This gave $P = 2.79 \times 10^{-12}$ for rs16847897 and $P = 3.72 \times 10^{-14}$ for rs12696304.

The association signal across the 3q26 locus from the combined BHF-FHS and UKBS analysis (**Fig. 1**) indicates that rs12696304 and rs16847897 are located toward opposite ends of an ~87-kb region

Figure 1 A map of the chromosome 3q26 region showing genetic association with mean leukocyte telomere length. Association of individual SNPs from the combined BHF-FHS and UKBS analysis is plotted as $-\log_{10} P$ against chromosomal base pair position. Results of both genotyped and imputed SNPs are provided. rs12696304 is shown in blue and the LD relationship of the other markers (including rs16847897) to rs12696304 is indicated by color: red, $r^2 > 0.8$; orange, $r^2 > 0.5$; yellow, $r^2 > 0.2$; white, $r^2 < 0.2$. The location of known genes in the region *ARPM1* (actin-related protein M1), *MYNN* (myoneurin) and three members of the *LRRC* (leucine-rich repeat-containing) superfamily (*LRRC34*, *LRRC31* and *LRRIQ4*) are shown, as well as the location of recombination hot spots (blue peaks). * indicates the position of the *TERC* coding sequence (451 bp).

showing significant association with telomere length. Notably, despite the relatively weak LD between rs12696304 and rs16847897, conditional analyses showed that their associations are not independent. Inclusion of rs16847897 as a covariate in the association analysis of rs12696304 negated the latter's association with telomere length ($P = 0.40$) and vice versa ($P = 0.10$). Haplotype analysis of SNPs at the locus also suggested the presence of only a single signal (**Supplementary Methods** and **Supplementary Table 4b**).

The variance in telomere length explained by the 3q26 locus ranged from 0.32% in PREVEND to 1.0% in BHF-FHS for rs12696304. The TwinsUK study provided a direct estimate in base pairs (bp) of the effect of the locus on telomere length. Each minor allele of rs12696304 was associated with a mean leukocyte telomere length that was ~75 bp shorter in this cohort (**Table 1**). To put this in context, in cross-sectional analyses, telomere length decreased by ~21 bp per year in the TwinsUK cohort; therefore, each minor allele of rs12696304 was associated with a shorter mean telomere length equivalent to ~3.6 years of average age-related telomere attrition. Similar analyses of the combined BHF-FHS and UKBS data and the PREVEND study gave effect sizes per minor allele of rs12696304 equaling 4.0 years and 6.6 years of age-related attrition, respectively.

rs12696304 lies 1.5 kb downstream of *TERC* (**Fig. 1**). To investigate the possibility that the genotyped variants are markers for causal variants within *TERC* that affect its function, we sequenced the coding region of *TERC* (spanning 451 bp) and approximately 1 kb upstream and 1.5 kb downstream of it in individuals from the GRAPHIC study who were homozygous either for both minor alleles ($n = 16$) or for both major alleles ($n = 16$) of rs12696304 and rs16847897 (**Supplementary Methods**). No variants were identified within the *TERC* coding sequence. However, this does not exclude the possibility that the association with telomere length is mediated by an effect on *TERC* expression. Alternately, the effect on telomere length could occur through one of the other genes in the 3q26 locus. These include *ARPM1* (encoding actin-related protein M1), *MYNN* (encoding myoneurin) and three members of the *LRCC* (leucine-rich repeat–containing) superfamily (*LRRC34, LRRC31* and *LRRIQ4*). *ARPM1* is a nuclear protein thought to have a role in organization of the sperm-specific nucleus in mouse[13]. Although very little is directly known about *LRRC34, LRRC31* and *LRRIQ4*, the LRRC superfamily consists of members with diverse roles, including DNA repair, cell-cycle regulation, apoptosis and chromosomal stability[14,15].

We report here a locus on 3q26 that affects telomere length in humans. Given the importance of telomeres in nuclear and cellular function and the central role of telomere length in determining telomere function, our findings could have broad relevance for both normal and pathological age-associated processes.

URLs. METAL, http://www.sph.umich.edu/csg/abecasis/metal/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

The BHF-FHS Study and the GRAPHIC Study were supported by the BHF. The WTCCC and recruitment of the UKBS donors were funded by the Wellcome Trust. The TwinsUK study was funded by the Wellcome Trust and Biotechnology and Biological Sciences Research Council project grant G20234. The PREVEND study is supported by the Dutch Kidney Foundation (grant E033), The Netherlands Heart Foundation (grant 2006B140, 2006T003) and the EU project grant Genomic Strategies for Treatment and Prevention of Cardiovascular Death in Uraemia and End-stage Renal Disease (GENECURE) (FP-6 LSHM CT 2006 037697). We thank J. Marchini for generating the imputed genotypes for the WTCCC cohorts and J. Gardner for generating the telomere length data in the TwinsUK study. V.C., A.H.G. and N.J.S. are supported by the BHF and the Leicester National Institute of Health Research (NIHR) Biomedical Research Unit in Cardiovascular Disease. M.M. and T.D.S. are supported by NIHR Biomedical Research Centre award to Guy's & St Thomas' National Health Service Foundation Trust in partnership with King's College London. P.vd.H. is supported by the Netherlands Organisation for Scientific Research (NWO) VENI grant 916.76.170 and the Interuniversity Cardiology Institute of the Netherlands (ICIN). Collaboration between the WTCCC Study telomere GWAS and the TwinsUK study was established under the European Framework 7 (FP7/2007-2013) European Network for Genetic and Genomic Epidemiology (ENGAGE) project (HEALTH-F4-2007- 201413).

AUTHOR CONTRIBUTIONS

N.J.S. and T.S. conceived the study. V.C., M.M. and P.v.d.H. designed the laboratory work and conducted the analyses. V.C., P.S.B., M. Kaiser, J.M., I.M.L. and R.A.d.B. undertook the laboratory work. A.J.B. provided bioinformatics support and S.R., C.N. and N.S. undertook statistical support. A.S.H. and N.J.S. recruited and provided samples and data from the BHF Family Heart Study; W.T.C.C.C. and W.O. from the UKBS samples; A.H.G., P.R.B., M.D.T. and N.J.S. from the GRAPHIC study; G.Z., A.M.V., H.B., M. Kimura, A.A. and T.S. from the TwinsUK study; and D.J.v.V., W.H.v.G. and G.N. from the PREVEND Study. J.R.T. oversaw the statistical analysis. The paper was written by V.C., M.M. and N.J.S. All authors contributed to the final version of the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Blackburn, E.H., Greider, C.W. & Szostak, J.W. *Nat. Med.* **12**, 1133–1138 (2006).
2. Njajou, O.T. *Proc. Natl. Acad. Sci. USA* **104**, 12135–12139 (2007).
3. Vasa-Nicotera, M. *et al. Am. J. Hum. Genet.* **76**, 147–151 (2005).
4. Andrew, T. *et al. Am. J. Hum. Genet.* **78**, 480–486 (2006).
5. Mangino, M. *et al. Hum. Mol. Genet.* **17**, 2518–2523 (2008).
6. Mangino, M. *et al. J. Med. Genet.* **46**, 451–454 (2009).
7. Samani, N.J. *et al. Am. J. Hum. Genet.* **77**, 1011–1020 (2005).
8. Wellcome Trust Case Control Consortium. *Nature* **447**, 661–678 (2007).
9. Cawthon, R.M. *Nucleic Acids Res.* **30**, e47 (2002).
10. Tobin, M.D. *et al. Hypertension* **51**, 1658–1664 (2008).
11. Spector, T.D. & Williams, F.M. *Twin Res. Hum. Genet.* **9**, 899–906 (2006).
12. Hillege, H.L. *et al. Circulation* **106**, 1777–1782 (2002).
13. Hara, Y. *et al. FEBS Lett.* **582**, 2998–3004 (2008).
14. Kobe, B. & Kajava, A.V. *Curr. Opin. Struct. Biol.* **11**, 725–732 (2001).
15. Rabenau, K.E. *et al. Oncogene* **23**, 5056–5067 (2004).

© 2010 Nature America, Inc. All rights reserved.

Corrigendum: Common variants near *TERC* are associated with mean telomere length

Veryan Codd, Massimo Mangino, Pim van der Harst, Peter S Braund, Michael Kaiser, Alan J Beveridge, Suzanne Rafelt, Jasbir Moore, Chris Nelson, Nicole Soranzo, Guangju Zhai, Ana M Valdes, Hannah Blackburn, Irene Mateo Leach, Rudolf A de Boer, Wellcome Trust Case Control Consortium, Alison H Goodall, Willem Ouwehand, Dirk J van Veldhuisen, Wiek H van Gilst, Gerjan Navis, Paul R Burton, Martin D Tobin, Alistair S Hall, John R Thompson, Tim Spector & Nilesh J Samani
***Nat. Genet.* 42, 197–199 (2010); published online 7 February 2010; corrected after print 16 April 2010.**

Dr. Abraham Aviv and Dr. Masayuki Kimura were inadvertently omitted from the author list and the author contribution statement in the version of the manuscript initially published on February 7, 2010. This has been corrected in the online html and PDF.

© 2010 Nature America, Inc. All rights reserved.