Variants near TERT and TERC influencing telomere length are associated with high-grade glioma risk

Kyle M Walsh1,2, Veryan Codd3,4, Ivan V Smirnov5, Terri Rice1, Paul A Decker6, Helen M Hansen1, Thomas Kollmeyer7, Matthew L Koef6, Annette M Molinaro8, Lucie S McCoy1, Paige M Bracci8, Belinda S Cabriga1, Melike Pekmezci9, Shichun Zheng1, Joseph L Wiemels1,10, Alexander R Pico1,11, Tarik Tihan9, Mitchell S Berger5, Susan M Chang5, Michael D Prados5, Daniel H Lachance12, Brian Patrick O’Neill12, Hugues Sicotte6, Jeanette E Eckel-Passow6, ENGAGE Consortium Telomere Group13, Pim van der Harst14,15, John K Wiencke1,10, Nilesh J Samani3,4, Robert B Jenkins7 & Margaret R Wrensch1,10

Glioma, the most common central nervous system cancer in adults, has poor prognosis. Here we identify a new SNP associated with glioma risk, rs1920116 (near TERC), that reached genome-wide significance in a meta-analysis of genome-wide association studies (GWAS) of high-grade glioma and replication data (1,644 cases and 7,736 controls). This region has previously been associated with mean leukocyte telomere length (LTL). We therefore examined the relationship between LTL and both this new risk locus and other previously established risk loci for glioma using data from a recent GWAS of LTL (n = 37,684 individuals). Alleles associated with glioma risk near TERC and TERT were strongly associated with longer LTL (P = 5.5 × 10−20 and 4.4 × 10−19, respectively). In contrast, risk-associated alleles near RTEL1 were inconsistently associated with LTL, suggesting the presence of distinct causal alleles. No other risk loci for glioma were associated with LTL. The identification of risk alleles for glioma near TERC and TERT that also associate with telomere length implicates telomerase in gliomagenesis.

Glioblastoma, the most common and aggressive form of glioma, has a median survival time of just 15 months. Both inherited and acquired genetic variation influence gliomagenesis. Previous glioma GWAS have identified seven risk loci, including two near genes involved in telomere dynamics (TERC and TERT). Findings for the other telomere-related glioma risk loci, TERT and RTEL1, in these data sets are highlighted in Table 1. Interestingly, the association peak for this newly discovered risk locus for glioma encompassed the TERC gene encoding the RNA component of telomerase (Fig. 1a).

To identify new risk loci for glioma, we imputed thirty 500-kb regions with at least one SNP having an association P value of <1.0 × 10−4 in our previous GWAS of high-grade glioma (Online Methods). The discovery phase meta-analysis combined new imputation data and existing genotyping array data from the University of California, San Francisco (UCSF) Adult Glioma Study (AGS)6.11 with new data from The Cancer Genome Atlas (TCGA)12 and the Wellcome Trust Case Control Consortium (WTCCC)13 (1,013 cases and 6,595 controls). The replication phase used criteria defined in the original GWAS for high-grade glioma, evaluating SNPs with an association P value of <1.0 × 10−4 in an additional 631 glioblastoma cases and 1,141 controls from the Mayo Clinic and UCSF. All individuals were of European ancestry. Study design and study populations are summarized in Supplementary Figure 1 and Supplementary Table 1.

In the discovery phase, only rs1920116 on chromosome 3q26.2 was associated with glioma risk at P < 1.0 × 10−6 (P = 8.7 × 10−7). The association at rs1920116 was also significant in the replication data set (P = 3.4 × 10−3). The combined P value for this SNP for all 1,644 cases and 7,736 controls in the discovery and replication data (P = 8.3 × 10−9) achieved genome-wide statistical significance (Table 1). Interestingly, the association peak for this newly discovered risk locus for glioma encompassed the TERC gene encoding the RNA component of telomerase (Fig. 1a).

Findings for the other telomere-related glioma risk loci, TERT and RTEL1, in these data sets are highlighted in Figure 1b,c and Table 1. Interestingly, alleles associated with glioma risk near TERC, TERT and RTEL1 had similar effect sizes and were the major allele

1Department of Neuroepidemiology, Department of Neurological Surgery, University of California, San Francisco, San Francisco, California, USA. 2Program in Cancer Genetics, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, California, USA. 3Department of Cardiovascular Sciences, University of Leicester, Leicester, UK. 4National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK. 5Department of Neurological Surgery, University of California, San Francisco, San Francisco, California, USA. 6Division of Biomedical Statistics and Informatics, Mayo Clinic College of Medicine, Rochester, Minnesota, USA. 7Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, Minnesota, USA. 8Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California, USA. 9Department of Pathology, University of California, San Francisco, San Francisco, California, USA. 10Institute for Human Genetics, University of California, San Francisco, San Francisco, California, USA. 11Department of Biomedical Informatics, Gladstone Institutes, San Francisco, California, USA. 12Department of Neurology, Mayo Clinic College of Medicine, Rochester, Minnesota, USA. 13A full list of members and affiliations appears in the Supplementary Note. 14Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. 15Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. Correspondence should be addressed to K.M.W. (kyle.walsh@ucsf.edu).

Received 4 November 2013; accepted 12 May 2014; published online 8 June 2014; doi:10.1038/ng.3004
in all three regions. SNPs in TERC, TERT and RTEL1 were independently associated with risk of high-grade glioma; modeling the effect of one SNP while controlling for the presence of the others did not result in attenuated associations. SNP-SNP interaction tests did not show any modification of effects (P_{min} = 0.11). Each additional risk-associated allele in TERC, TERT and RTEL1 (range of 0–6) contributed to glioma risk in a monotonic, increasing fashion (Supplementary Fig. 2). We previously observed that the effects of the TERT and RTEL1 SNPs on glioma risk increased with age^{14}. We observed modest evidence of this trend at rs1920116 near TERC (P = 0.098) (Supplementary Fig. 3).

To determine whether loci associated with glioma risk influenced telomere length in an independent data set, we examined SNP data from a recent GWAS of mean LTL conducted in 37,684 individuals of European descent\(^1\). LTL association data were available for lead SNPs associated with glioma risk near TERC, TERT, RTEL1, CCDC26, CDKN2B, PHLDB1 and EGFR\(^{3,4,6}\). The top alleles for glioma risk near TERC and TERT were strongly associated with longer LTL (P = 5.5 × 10^{-20} and 4.4 × 10^{-19}, respectively) (Table 2). In contrast, the risk allele for glioma at rs6010620 in RTEL1 was modestly associated with shorter LTL (P = 1.1 × 10^{-3}) (Table 2). Other loci known to be associated with glioma risk were not associated with LTL (Supplementary Table 2). The lack of association with LTL at these other loci supports the hypothesis that multiple mechanisms contribute to glioma risk, of which changes in telomere dynamics might be one\(^3\). The presence of genes causally implicated in carcinogenesis at risk loci for glioma not associated with LTL supports this hypothesis (for example, EGRF and CDKN2B).

Next, we comprehensively compared the direction and statistical significance of associations in the TERC, TERT and RTEL1 regions with high-grade glioma risk and mean LTL. For all SNPs, the allele associated with glioma risk was set as the reference allele. We then evaluated whether the allele for glioma risk was associated with longer or shorter LTL. Alleles in the TERC and TERT regions were consistently associated with both higher glioma risk and longer LTL (Fig. 2a,b).

![Figure 1](https://example.com/figure1.png)

**Figure 1** SNP association plots for high-grade glioma risk and mean LTL. (a–c) Association plots are shown for 3q26.2 (TERC) (a), 5p15.33 (TERT) (b) and 20q13.33 (RTEL1) (c). Associations with telomere length are from the telomere length genome-wide association meta-analysis (n = 37,684 individuals). Associations with glioma risk are from the genome-wide discovery meta-analysis, combining data from UCSF AGS, TCGA and WTCCC (n = 1,013 cases and 6,595 controls). Results are also shown for the combined glioma discovery and UCSC and Mayo Clinic replication analyses (n = 1,644 cases and 7,736 controls).

### Table 1 Associations of risk for high-grade glioma with SNPs near TERC, TERT and RTEL1 in the UCSF AGS, the Mayo Clinic glioma case-control study, TCGA and WTCCC cohorts

| SNP   | Chr. | Position | Risk allele | Gene   | Study          | Number of cases/controls | RAF in cases/controls | OR (95% CI) | P       |
|-------|------|----------|-------------|--------|----------------|--------------------------|-----------------------|-------------|---------|
| rs1920116 | 3    | 169,579,971 | G           | TERC   | Discovery GWAS | 1,013/5,959               | 0.77/0.72              | 1.31 (1.18–1.46) | 8.7 × 10^{-7} |
|        |      |           |             |        | Replication   | 631/1,141                | 0.76/0.71              | 1.27 (1.09–1.51) | 3.4 × 10^{-3} |
|        |      |           |             |        | Combined total | 1,644/7,736              | 0.77/0.72              | 1.30 (1.19–1.42) | 8.3 × 10^{-9} |
| rs2736100 | 5    | 1,286,516     | C           | TERT   | Discovery GWAS | 1,013/5,959               | 0.60/0.51              | 1.43 (1.30–1.57) | 2.9 × 10^{-13} |
|        |      |           |             |        | Replication   | 605/1,141                | 0.57/0.49              | 1.29 (1.12–1.49) | 5.8 × 10^{-4}  |
|        |      |           |             |        | Combined total | 1,618/7,736              | 0.59/0.51              | 1.39 (1.28–1.50) | 1.4 × 10^{-15} |
| rs6010620 | 20   | 62,309,839    | G           | RTEL1  | Discovery GWAS | 1,013/5,959               | 0.84/0.76              | 1.53 (1.36–1.72) | 7.0 × 10^{-13} |
|        |      |           |             |        | Replication   | 631/1,141                | 0.83/0.75              | 1.64 (1.37–1.97) | 1.4 × 10^{-7}  |
|        |      |           |             |        | Combined total | 1,644/7,736              | 0.84/0.76              | 1.56 (1.42–1.72) | 4.7 × 10^{-10} |

Chr., chromosome; RAF, risk allele frequency; CI, confidence interval.
glioma risk is associated with longer (blue) or shorter (orange) LTL. Similarly, of 67 SNPs at 3q26.2 associated with increased LTL, every 3q26.2 SNP that was associated with glioma at the genome-wide significance level (P < 5 × 10−5) was also associated with increased LTL (Table 2). The association peaks for glioma risk and LTL near RTEL1 showed substantially less overlap than those near TERC and TERT. Indeed, the RTEL1 SNPs most significantly associated with glioma risk were not in LD with the top SNPs associated with LTL, suggesting that different causal alleles influence the two phenotypes (Supplementary Fig. 4c). This finding is particularly intriguing because the RTEL1 protein interacts with proliferating cell nuclear antigen (PCNA), a processivity factor for DNA polymerase. The RTEL1-PCNA interaction is essential for replication fork stability and suppression of telomere fragility, but, in the absence of this interaction, RTEL1 can still disassemble telomere loops and inhibit telomere shortening. Thus, genetic variation that limits the RTEL1-PCNA interaction might influence glioma risk without directly influencing telomere length. The PCNA interaction motif (PIP box) is encoded by exon 34 of RTEL1, near the interaction site (Fig. 2c).

In contrast, some alleles for glioma risk near RTEL1 were associated with longer LTL, whereas others were associated with shorter LTL (Fig. 2c). The associations with LTL near RTEL1 had substantially more modest P values than those observed near the telomerase components TERC and TERT.

The association peaks for glioma risk and LTL at 3q26.2, containing TERC and five additional genes, span a ~200-kb region with low recombination rates and strong linkage disequilibrium (LD) (Supplementary Fig. 4a). Although the association peak for LTL is located ~85 kb centromeric to the association peak for glioma risk, every 3q26.2 SNP that was associated with glioma at P < 0.01 (n = 54 SNPs) was also associated with increased LTL at P < 1.0 × 10−5. Similarly, of 67 SNPs at 3q26.2 associated with increased LTL at P < 1.0 × 10−10, 65 (97%) were associated with increased glioma risk at P < 0.05. These SNPs included the top association with LTL, rs10936599 (PGLMA = 7.6 × 10−3; odds ratio (OR) = 1.17). Although distinct causal alleles might underlie the association signals for LTL and glioma risk near TERC, such alleles likely reside on the same extended haplotype.

The OR estimates for risk of high-grade glioma appeared relatively uniform throughout the 3q26.2 region (Supplementary Fig. 5), similar to observations of the relationship of this region to colorectal cancer. Conditional analyses suggested that multiple 3q26.2 SNPs might influence glioma risk, as controlling for either the lead SNP for glioma (rs1920116) or the lead SNP for LTL (rs10936599) did not fully attenuate the association between glioma risk and nearby alleles (Supplementary Fig. 6). Differences in the statistical significance of the associations for LTL and glioma risk might also be related to differences in sample size and analytic technique (linear regression for LTL versus logistic regression for glioma) or to the tissue type under study (leukocytes for LTL versus astrocytes for glioma). Variants regulating TERC expression and possible downstream effects on telomere length might differ across tissues owing to differential transcription factor expression or other tissue-divergent regulatory mechanisms. A full analysis of the potential functional consequences of 3q26.2 SNPs is presented in Supplementary Table 3 and indicates that rs1920116 might have long-range effects on gene expression.

In the TERT region at 5p15.33, with high recombination rates and low LD (Supplementary Fig. 4b), rs2736100 was the most statistically significant association for both glioma risk and LTL. A previous study found multiple independent SNPs in TERT that influenced LTL and breast cancer risk. Our data suggest that, in the case of glioma, TERT variants are simultaneously associated with increased glioma risk and longer LTL.

The association peaks for glioma risk and LTL near RTEL1 showed substantially less overlap than those near TERC and TERT. Indeed, the RTEL1 SNPs most significantly associated with glioma risk were not in LD with the top SNPs associated with LTL, suggesting that different causal alleles influence the two phenotypes (Supplementary Fig. 4c). This finding is particularly intriguing because the RTEL1 protein interacts with proliferating cell nuclear antigen (PCNA), a processivity factor for DNA polymerase. The RTEL1-PCNA interaction is essential for replication fork stability and suppression of telomere fragility, but, in the absence of this interaction, RTEL1 can still disassemble telomere loops and inhibit telomere shortening. Thus, genetic variation that limits the RTEL1-PCNA interaction might influence glioma risk without directly influencing telomere length. The PCNA interaction motif (PIP box) is encoded by exon 34 of RTEL1, near the most significant association signals for glioma in our data set.

Because telomere maintenance is a universal requirement for oncogenic progression and telomere length shows substantial interindividual variability, telomere length is a promising epidemiological risk factor for cancer. Mean telomere length decreases with age (declining by 20–40 bp per year) and has been proposed as a biomarker of health. Although shorter telomere length is causally linked to increased risk of coronary artery disease, both longer

---

**Table 2** Association of SNPs near TERC, TERT and RTEL1 with mean LTL in the UK genome-wide association meta-analysis

| SNP      | Chr. | Position   | Allele⁴ | Gene | Number | AF  | β (s.e.)⁹ | P       |
|----------|------|------------|---------|------|--------|-----|----------|---------|
| rs1920116| 3    | 169,579,971| G       | TERC | 37,489 | 0.71| 0.0733 (0.0080) | 5.5 × 10⁻²⁰ |
| rs2736100| 5    | 1,286,516  | C       | TERT | 25,842 | 0.49| 0.0783 (0.0087) | 4.4 × 10⁻¹⁹ |
| rs6010620| 20   | 62,309,839 | G       | RTEL1| 37,607 | 0.77| −0.0278 (0.0085) | 1.1 × 10⁻³  |

⁴The allele listed is the risk allele for high-grade glioma. ⁹β estimates (standard error) are from regressing mean LTL on each additional copy of the glioma risk allele. Negative β values indicate that the glioma risk allele is associated with shorter LTL, whereas positive β values indicate that the glioma risk allele is associated with longer LTL.

---

© 2014 Nature America, Inc. All rights reserved.
and shorter telomere lengths have been associated with increased cancer risk.25–27

Although data connecting telomere length in astrocytes with LTL is lacking, telomere length in individuals is highly correlated across tissues.28,29 Recent research indicates that environmental insults alter telomere length and that the magnitude of these changes is linked to telomere length at baseline.30 This correlation suggests that stress-induced telomere extension occurs through the same pathway as canonical telomere extension. Therefore, heritable variation might influence telomere dynamics in both normal and premalignant cells.

Another SNP near TERC has been shown to significantly associate with colorectal cancer risk (lead SNP rs10936599).31 The allele for colorectal cancer risk was also associated with increased glioma risk (OR = 1.17; P = 7.6 × 10⁻³) and longer telomeres (P = 2.5 × 10⁻³¹) in our data. Idiopathic pulmonary fibrosis risk has been shown to significantly associate with SNPs near TERC (rs6793295) and TERT (rs2736100).32 For both SNPs, the allele associated with risk of pulmonary fibrosis was also associated with significantly shorter telomere length (P = 1.1 × 10⁻²⁸ and 4.4 × 10⁻¹⁹, respectively) and with decreased glioma risk (P = 2.1 × 10⁻³ and 2.9 × 10⁻¹¹) in our data. These observations suggest that both longer and shorter telomere lengths might be pathogenic, depending on the disease under consideration.

In summary, we identify a new susceptibility locus for high-grade glioma near TERC and demonstrate that the alleles for glioma risk near TERC and TERT might also influence telomere length. Top SNPs associated with glioma risk near TERC, TERT and RETEA all have risk allele frequencies of greater than 50% among our control sample. Considering that increased telomere length protects against cardiovascular disease33 but may concurrently increase risk for some cancers, these alleles might be influenced simultaneously by positive and negative selective pressures. This balancing selection underscores the complexity of telomere-based pathways of disease susceptibility.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

Work at UCSF was supported by the US National Institutes of Health (grants R25CA112355, R01CA52689, P50CA097257, R01CA126831 and R01CA139020), as well as by the National Brain Tumor Foundation, the UCSF Lewis Chair in Brain Tumor Research, the UCSF Robert Magnin Newman Chair in Neuro-Oncology and donations from the families and friends of John Berardi, Helen Glaser, Elvera Olsen, Raymond E. Cooper and William Martinussen. Work at the Mayo Clinic was supported by the US National Institutes of Health (grants P50CA108961 and P30CA15083), the National Institute of Neurological Disorders and Stroke (grant R1C1NS068222Z), the Bernie and Edith Waterman Foundation, and the Ting Tsung and Wei Fong Chao Family Foundation. Work at the University of Leicester was undertaken under the European Union Framework Programme 7 ENGAGE Project (HEALTH-F4-2007-201413). V.C. and N.J.S. are supported by the British Heart Foundation.

This project was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, US National Institutes of Health, through UCSF Clinical and Translational Science Institute grant UL1RR024131. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the US National Institutes of Health.

The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; by the National Cancer Institute’s Surveillance, Epidemiology and End Results Program under contract HHSN261201000040C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California and contract HHSN261201000034C awarded to the Public Health Institute; and by the Centers for Disease Control and Prevention National Program of Cancer Registries, under agreement USDP003862-01 awarded to the California Department of Public Health. The ideas and opinions expressed herein are those of the authors, and endorsement by the State of California Department of Public Health, the National Cancer Institute and the Centers for Disease Control and Prevention or their contractors and subcontractors is not intended nor should be inferred.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

AUTHOR CONTRIBUTIONS

K.M.W., M.R.W. and J.K.W. led the study at UCSE, R.B.I. led the study at the Mayo Clinic, and N.I.S. led the study at the University of Leicester. K.M.W., V.C., R.B.I., M.R.W. and T.R. contributed to manuscript preparation. Study coordination was the responsibility of T.K. at the Mayo Clinic and T.R. and L.S.M. at UCSF. K.M.W. and V.C. coredicted and conducted biostatistics and bioinformatics analyses with additional support from F.A.D., J.E.E.-P., M.L.K., A.M.M., P.M.R., T.R., H.S., A.R.P., I.V.S., Pvd.H. and the ENGAGE Consortium. Laboratory work was performed by T.K. under the direction of R.B.I. at the Mayo Clinic and by H.M.H., S.Z. and B.S.C. under the direction of J.K.W. and J.L.W. at UCSF. Pathology support was provided by T.T. Subject enrollment or clinical record review was performed or facilitated by M.D.P., S.M.C., M.S.B., B.P.O., D.H.L. and Pvd.H.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Codd, V. et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat. Genet. 45, 422–427 (2013).
2. Stupp, R. et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone in survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol. 10, 459–466 (2009).
3. Sanson, M. et al. Chromosome 7p11.2 (EGFR) variation influences glioma risk. Hum. Mol. Genet. 20, 2897–2904 (2011).
4. Shehe, S. et al. Genome-wide association study identifies five susceptibility loci for glioma. Nat. Genet. 41, 899–904 (2009).
5. Stacey, S.N. et al. A germline variant in the TPS3 polyadenylation signal confers glioblastoma susceptibility. Nat. Genet. 43, 1098–1103 (2011).
6. Wrensch, M. et al. Variants in the CDK4N28 and RETEL1 regions are associated with high-grade glioma susceptibility. Nat. Genet. 41, 905–908 (2009).
7. Jenkins, R.B. et al. A low-frequency variant at 8q24.21 is strongly associated with risk of oligodendrogliomas and astrocytomas with IDH1 or IDH2 mutation. Nat. Genet. 44, 1122–1125 (2012).
8. Killela, P.J. et al. TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. Proc. Natl. Acad. Sci. USA 110, 6021–6026 (2013).
9. Horn, S. et al. TERT promoter mutations in familial and sporadic melanoma. Science 339, 957–959 (2013).
10. Huang, F.W. et al. Highly recurrent TERT promoter mutations in human melanoma. Science 339, 957–959 (2013).
11. Walsh, K.M. et al. Analysis of 60 reported glioma risk SNPs replicates published GWAS findings but fails to replicate associations from published candidate-gene studies. Genet. Epidemiol. 37, 222–228 (2013).
12. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455, 1061–1068 (2008).
13. Welcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661–678 (2007).
14. Walsh, K.M. et al. Genetic variants in telomerase-related genes are associated with an older age at diagnosis in glioma patients: evidence for distinct pathways of glioma progression. Neuro-oncology 15, 1041–1047 (2013).
15. Mein, B.S., Nordfjall, K., Anderson, U. & Roos, G. hTERT cancer risk genotypes are associated with telomere length. Genet. Epidemiol. 36, 368–372 (2012).
16. Jones, A.M. et al. TERC polymorphisms are associated both with susceptibility to colorectal cancer and with longer telomeres. Gut 61, 248–254 (2012).
17. Bojesen, S.E. et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat. Genet. 45, 371–384 (2013).
18. Vannier, J.B. et al. RTEL1 is a replisome-associated helicase that promotes telomere and genome-wide replication. Science 342, 239–242 (2013).
19. Chang, S., Khoo, C.M., Naylor, M.L., Maser, R.S. & DePinho, R.A. Telomere-based crisis: functional differences between telomerase activation and ALT in tumor progression. Genes Dev. 17, 88–100 (2003).
20. Vasa-Nicotera, M. et al. Mapping of a major locus that determines telomere length in humans. Am. J. Hum. Genet. 76, 147–151 (2005).
21. Fitzpatrick, A.L. et al. Leukocyte telomere length and cardiovascular disease in the Cardiovascular Health Study. Am. J. Epidemiol. 165, 14–21 (2007).
22. Codd, V. et al. Common variants near TERC are associated with mean telomere length. Nat. Genet. 42, 197–199 (2010).
23. Denham, J. et al. Longer leukocyte telomeres are associated with ultra-endurance exercise independent of cardiovascular risk factors. PLoS ONE 8, e69377 (2013).
24. Brouilette, S.W. 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 369, 107–114 (2007).
25. Ma, H. et al. Shortened telomere length is associated with increased risk of cancer: a meta-analysis. PLoS ONE 6, e20466 (2011).
26. Wentzensen, I.M., Mirabello, L., Pfeiffer, R.M. & Savage, S.A. The association of telomere length and cancer: a meta-analysis. Cancer Epidemiol. Biomarkers Prev. 20, 1238–1250 (2011).
27. Hou, L., Zhang, X., Gawron, A.J. & Liu, J. Surrogate tissue telomere length and cancer risk: shorter or longer? Cancer Lett. 319, 130–135 (2012).
28. Wilson, W.R. et al. Blood leucocyte telomere DNA content predicts vascular telomere DNA content in humans with and without vascular disease. Eur. Heart J. 29, 2689–2694 (2008).
29. Okuda, K. et al. Telomere length in the newborn. Pediatr. Res. 52, 377–381 (2002).
30. Romaro, G.H. et al. Environmental stresses disrupt telomere length homeostasis. PLoS Genet. 9, e1003721 (2013).
31. Houlston, R.S. et al. Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. Nat. Genet. 42, 973–977 (2010).
32. Fingerlin, T.E. et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. Nat. Genet. 45, 613–620 (2013).
33. Samani, N.J. & van der Harst, P. Biological ageing and cardiovascular disease. Heart 94, 537–539 (2008).
ONLINE METHODS

Ethics statement. Glioma studies were approved by the University of California, San Francisco Committee on Human Research and the Mayo Clinic Office for Human Research Protection. Informed consent was obtained from all study participants. The genome-wide meta-analysis of mean LTL obtained approval from local ethics committees as previously outlined1.

Study design. The overall study design is summarized in Supplementary Figure 1. The study had four stages. Stage 1 and stage 2 were part of a combined discovery phase, while stage 3 involved replication and stage 4 examined the function of top associated loci in an independent data set with LTL measurements.

Stage 1 was a GWAS of high-grade glioma cases (n = 690) and controls (n = 3,992), the results of which were first published in a 2009 study8. This GWAS analyzed UCSF AGS cases (n = 620) and controls (n = 602), TCGA cases genotyped on Illumina arrays (n = 70) and Illumina iControls (n = 3,390). Results for this pooled analysis were used to generate a list of regions that had not undergone attempted replication in the 2009 publication (association \( P > 1 \times 10^{-6} \)) but that showed suggestive evidence of association (\( P < 1 \times 10^{-6} \)). This analysis yielded 30 independent loci containing at least 1 SNP with an association \( P \) value between 1.0 \( \times 10^{-4} \) and 1.0 \( \times 10^{-6} \) for additional study. Imputation to 1000 Genomes Project data was performed for 500-kb regions centered on the most statistically significant SNP for each of the 30 regions (Supplementary Table 4). EIGENSTRAT-adjusted case-control associations were calculated for genotyped and imputed SNPs.

Stage 2 analyzed Affymetrix 6.0 SNP array data from new TCGA glioblastoma cases (n = 323) and WTCCC controls (n = 2,603). Imputation to 1000 Genomes Project data was performed for the thirty 500-kb regions identified in stage 1, and case-control comparisons were made for genotyped and imputed SNPs. EIGENSTRAT-adjusted case-control association results from stage 1 and stage 2 were combined using meta-analysis to identify SNPs meeting the original a priori threshold for attempted replication (\( P < 1 \times 10^{-6} \)).

Stage 3 attempted replication of the single SNP from stage 2 that had an association \( P \) value less than 1.0 \( \times 10^{-6} \) (rs1920116 near \( TERC, P = 8.7 \times 10^{-7} \)). Targeted genotyping of this SNP using either TaqMan or Sequenom assays was performed in an additional 328 Mayo Clinic glioblastoma cases, 303 UCSF glioblastoma cases, 766 Mayo Clinic controls and 375 UCSC controls. SNP data for these samples were also available for previously established risk loci for high-grade glioma in \( TERT \) and \( RET11 \) that emerged from stage 1 and were first published in the 2009 study9.

Stage 4 combined data from a GWAS meta-analysis of mean LTL (n = 37,684 individuals), published in a 2013 study4, with our high-grade glioma association data to investigate the direction and statistical significance of associations for glioma risk alleles and mean LTL.

Study population. Characteristics of cases and controls from UCSF, the Mayo Clinic, Illumina, TCGA and WTCCC are summarized in Supplemental Table 1. All individuals were of European ancestry and were older than 18 years of age. UCSF cases and controls were taken from the San Francisco Bay Area AGS. A total of 923 high-grade glioma cases and 977 controls from the UCSF AGS were used in this study (stages 1 and 3). A total of 328 glioblastoma cases and 766 controls from the Mayo Clinic were used in this study (stage 3). A total of 37,684 individuals of European descent, aged >18 years, were included in the genome-wide analyses of mean LTL (stage 4).

Subject recruitment. All UCSF AGS cases were adults (>18 years of age) with newly diagnosed, histologically confirmed glioma (International Classification of Disease for Oncology, morphology codes 9380–9481). Population-based cases from six San Francisco Bay Area counties were ascertained using the early case ascertainment system of the Cancer Prevention Institute of California from May 1997 to August 1999 (series 2), from November 2001 to September 2005 (series 3) and from September 2006 to September 2009 (series 4). Clinic-based cases diagnosed between 2002 and 2006 (series 3), between 2006 and 2010 (series 4) and between 2009 and 2012 (series 5) of the same histology were recruited from the UCSF Neuro- oncology Clinic, regardless of place of residence. From 1991 to 2010, population-based controls from the same residential area as the population-based cases were identified using random-digit dialing and were frequency matched to population-based cases on the basis of age, sex and ancestry. After 2010, all controls were selected from the UCSF Laboratory Medicine at Parnassus clinic. These controls were matched to clinic-based glioma cases on the basis of age, sex and ancestry. Tumor specimens and pathology reports were reviewed by UCSF neuropathologists. Consent forms provided blood, buccal and/or saliva specimens and information during an in-person or telephone interview.

The Mayo Clinic cases included individuals with newly diagnosed glioblastoma identified between 2005 and 2012. Cases were identified within 24 h of diagnosis, except for those who had their initial diagnosis elsewhere, and cases were then verified at the Mayo Clinic. Pathologic diagnosis was confirmed by review of the primary surgical material for all cases by two Mayo Clinic neuropathologists. The control group consisted of consented individuals who had a general medical exam at the Mayo Clinic. Matching variables were sex, date of birth (within 2.5 years), self-identified ancestry (Hispanic white, non-Hispanic white, American Indian, African American, Asian, Pacific Islander or other) and residence. Geographic region of residence was matched in three zones on the basis of distance to the Mayo Clinic Rochester: Olmsted County; the rest of Minnesota, Wisconsin, Iowa, North Dakota and South Dakota; and the rest of the United States and Canada. Individuals under 18 years of age and those with a history of a brain tumor were excluded.

A total of 37,684 individuals from 15 cohorts were used in the GWAS meta-analysis of mean LTL. Samples were collected and analyzed by the European Network for Genetic and Genomic Epidemiology (ENGAGE) Consortium Telomere Group and minor collaborators. These institutions are distinct from those collecting the glioma case-control samples. Details on the ENGAGE Consortium and subject recruitment procedures can be found in Supplementary Table 1 of Codd et al.1.

Measurement of telomere length. Telomere measurements were performed on blood-derived DNA. Assays were conducted at one of four centralized laboratories to ensure data harmonization. Mean LTL was measured using a quantitative PCR–based technique in all samples34,35. This method expresses telomere length as a ratio (T/S) of telomere repeat length (T) to copy number for a single-copy gene (S) in each sample. To standardize across plates, either a calibrator sample or a standard curve was used for quantification. LTL measurements were made in five separate laboratories. Details on the methods used are provided elsewhere1. The majority of the samples (67%) were run in a single laboratory, with mean inter-run coefficients of variation for LTL measurements in individual cohorts ranging between 2.7 and 3.9%. Ranges in T/S ratios were found to vary for cohorts measured in different laboratories, largely owing to differences in the calibrator or standard DNA used. Standardized LTL in each cohort was generated using a Z-transformation approach.

Sample preparation and genotyping. Stage 1. For the UCSF AGS discovery samples used in stage 1 of the study, DNA was isolated from whole blood using the Gentra PureGene DNA isolation kit (Qiagen) and quantified using PicoGreen reagent (Invitrogen). Genotyping was conducted by deCODE genetics. Samples were randomized before plating on specimen plates provided by deCODE genetics. The genotyping assay panel was used by Illumina HumanCNV370-Duo BeadChip. In addition to randomizing samples and the quality control measures provided by deCODE genetics, we included two duplicate samples per plate and one CEPH24 trio (parents and child) per plate. A total of 51 duplicate samples were plated and showed average concordance of >99%. DNA was re-extracted for any sample with a call rate of <98% and genotyped again. Samples with call rates of <98% were excluded from analysis (n = 1), as were samples with mismatched reported and genotyped sex (n = 1). Although all subjects were of self-reported European ancestry, one subject was removed because of not clustering with subjects of European ancestry in EIGENSTRAT analyses. The assay panel contained a total of 353,202 SNP probes, of which 326,506 biallelic SNPs had call rates of >98% and Hardy–Weinberg equilibrium \( P \) values of >0.0001 (among controls) and were included in analyses.

We downloaded HumanHap550 platform genotyping data from blood specimen DNA for 72 glioblastoma cases from TCGA for use in stage 1. Although all 72 subjects were identified as white, our analyses showed that 1 had non-European ancestry and another appeared to duplicate an AGS case, leaving 70 TCGA cases in stage 1 analyses.
We assembled an independent control genotype data set of 3,390 non-redundant European-ancestry controls from the Illumina iControlDB. The subjects are anonymous, with information available only on age, sex and ancestry. We checked for evidence of non-European ancestry and sample duplication or related subjects (identity by state (IBS) > 1.6) among AGS samples, TCGA cases and Illumina iControls by performing multidimensional scaling (MDS) analysis on 20 bootstrap samples of 1,000 random autosomal biallelic SNPs. After filtering on the basis of these quality assessment measures, we obtained a total of 3,390 European-ancestry Illumina iControls from 3 different panels with up to 306,154 autosomal SNPs overlapping the HumanHap370duo panel used for the AGS subjects. The iControls samples were genotyped on Illumina HumanHap300 (n = 319 subjects), HumanHap550v1 (n = 1,519) and HumanHap550v3 (n = 1,552) arrays.

Stage 2. Genome-wide SNP data for 323 TCGA glioblastoma cases, not included in stage 1 analyses, were downloaded from TCGA for use in stage 2 analyses.

Stage 3. DNA for cases and controls was isolated from whole-blood or saliva samples. UCSF samples were whole-genome amplified as previously described. UCSF and Mayo Clinic samples were genotyped using a TaqMan assay from Applied Biosystems (C_26010927_10). Cases and controls were randomized on 96-well plates containing HapMap trios and five duplicate samples per plate. All trio genotypes displayed mendelian consistency, and duplicates showed genotype concordance. Genotypes for a small subset of UCSF samples were generated on a custom Sequenom panel (n = 25 cases). All cluster plots were visually inspected. To exclude poorly genotyped SNPs, any SNP with a Hardy-Weinberg equilibrium P value of <0.001 in controls, stratified by site, was removed from further analyses. The TERT and RETE1 genotypes for samples in stage 3 were generated on an Illumina GoldenGate custom genotyping array, as previously described. No SNP failed Hardy-Weinberg equilibrium checks in stage 3.

Stage 4. All discovery cohorts for the genome-wide meta-analysis of LTL had genotype information generated on a standard genotyping platform from Illumina or Affymetrix and included imputed genotypes based on HapMap 2 CEU (Utah residents of Northern and Western European ancestry) reference data. Detailed information about individual genotyping platforms is provided elsewhere. All study-specific files underwent extensive quality control procedures before meta-analysis. All files were checked for completeness. Allele frequencies were checked for compliance with HapMap. In addition to undergoing study-specific quality control filters, SNP results from each study were included in the meta-analysis only if the imputation quality score was >0.5 and if the minor allele frequency was >1%. Only SNPs for which data were available in >50% of the total sample across all studies were analyzed, resulting in a total of 2,362,330 SNPs in the meta-analysis.

Statistical analyses. Stage 1. Genome-wide SNP data for high-grade glioma cases and controls were analyzed using logistic regression, adjusted for the first two ancestry-informative principal components, using R and EIGENSTRAT. Single-SNP association results were computed assuming an allelic additive model for zero, one or two copies of the minor allele (equivalent to a Cochran-Armitage test for trend).

Imputation was performed for 500-kb regions centered on 30 top SNPs using IMPUTE2 v2.1.2.2 software and its standard Markov chain Monte-Carlo algorithm and default settings for targeted imputation. All 1000 Genomes Project Phase I interim release haplotypes were provided as the imputation reference panel. SNPs with imputation quality (info) scores less than 0.80 or posterior probabilities less than 0.90 were excluded to remove poorly imputed SNPs. Any SNP with a minor allele frequency of <1% in cases was excluded from association tests. Association statistics for imputed and directly genotyped SNPs were calculated using logistic regression in SNPTEST2v2, applying an allelic additive model. A missing-data likelihood score test was applied to the imputed variants to produce standard errors that accounted for the additional uncertainty inherent in the analysis of imputed genotypes. The effect of individual SNPs on glioma risk was calculated while adjusting for the first two principal components from EIGENSTRAT.

Stage 2. As in stage 1, thirty 500-kb regions were imputed using IMPUTE2 v2.1.2.2 software and the 1000 Genomes Project Phase I reference panel. SNPs with imputation quality (info) scores less than 0.80 or posterior probabilities less than 0.90 were excluded to remove poorly imputed SNPs. Association statistics for imputed and directly genotyped SNPs were calculated using logistic regression in SNPTEST2v2, applying an allelic additive model. Association statistics were adjusted for the first two principal components, generated by EIGENSTRAT. SNP association statistics from stage 1 and stage 2 analyses were combined using the program META. As a measure of between-study heterogeneity, I² was calculated. For SNPs with I² ≤ 40%, fixed-effects models were applied; random-effects models were applied for SNPs with I² > 40%.

Stage 3. Any SNP with a P value in the stage 2 meta-analysis of 30 regions surpassing our a priori significance threshold of 1 × 10⁻⁸ was selected for replication. For the single SNP that met this threshold, replication was attempted in 328 glioblastoma cases and 766 controls from the Mayo Clinic and 303 cases and 375 controls from UCSE. UCSE cases included in stage 3 were not included in the original stage 1 GWAS owing to date of recruitment or insufficient DNA quantity. Case-control associations were evaluated using logistic regression, assuming an allelic additive model. Associations were calculated separately for each site (Mayo Clinic and UCSE) and then combined using fixed-effects meta-analysis with META software. A summary meta-analysis of all high-grade glioma cases and controls was performed by combining the results from stage 1, stage 2 and the two replication series from stage 3 using fixed-effects meta-analysis with META software. SNPs showing association with risk of high-grade glioma at P < 5.0 × 10⁻⁸, corresponding to a Bonferroni correction for 1 million independent tests, were considered to have achieved genome-wide statistical significance by current convention.

Stage 4. In the GWAS of LTL, SNP array data from each of 15 cohorts underwent genome-wide SNP imputation using HapMap 2 CEU as the reference, as previously described. Within each cohort, SNP associations with LTL were analyzed by linear regression assuming additive effects, with adjustment for age, sex and study-specific covariates where appropriate (for example, smoking). This approach ensured that all regression analyses were performed uniformly but permitted studies to adjust for additional, relevant site-specific covariates (for example, ancestry-related principal components). Standard errors for each study were corrected using study-specific λ estimates for the genomimic control procedure. Meta-analysis of all associations from individual studies was conducted using inverse-variance weighting in STATA. As a measure of between-study heterogeneity, I² was calculated. For SNPs with I² ≤ 40%, fixed-effects models were applied; random-effects models were applied for SNPs with I² > 40%. The overall λ inflation factor for the meta-analysis results was 1.007, and results were further corrected for this factor using the genomic control procedure.

The glioma and LTL SNP data sets were harmonized by mapping all positions to Build37/hg19 and setting the allele that increased glioma risk as the reference allele in association tests for both traits. As a result, all OR estimates for association between SNPs and high-grade glioma were >1.0. Similarly, associations between SNPs and LTL were calculated with the risk allele for glioma set as the reference. A positive β value from linear regression of mean LTL indicated that the risk allele for glioma was associated with decreased telomere length, whereas a negative β value from linear regression of mean LTL indicated that the risk allele for glioma was associated with increased telomere length. Loci for glioma risk showing association with LTL significant association.

Testing for the independence of glioma risk loci and for the presence of interaction. We tested for evidence of gene-gene interactions for the top SNPs.
in the telomere-associated genes TERC, TERT and RTEL1. Associations of rs1920116, rs2736100 and rs6010620 with risk for high-grade glioma were modeled jointly to assess whether the observed associations were independent or possibly synergistic in nature. First, all three SNPs were included in a single logistic regression model to determine whether controlling for glioma risk variants in related genes attenuated any of the SNP associations. Inclusion of an allelic additive interaction term in the model (rs1920116-rs2736100, rs1920116-rs6010620 or rs2736100-rs6010620) was performed to test for the presence of any significant modification of effect ($P < 0.05$). Combining the risk allele dosage into a single ordinal variable representing the total number of risk-conferring SNPs in a telomere-associated gene (range of 0–6) was carried out to determine whether the alleles contributed to glioma risk in a monotonic, increasing fashion (Supplementary Fig. 2). To test whether age modified the effect of these SNPs on glioma risk, an allelic additive interaction term was included in the model (for example, rs1920116-age). Additionally, OR estimates for glioma were calculated in case-control analyses stratified into 10-year intervals of age, as previously described (Supplementary Fig. 3).

Genomic annotation of SNPs. Fifty-four SNPs at 3q26.2 that had association $P$ values of <0.01 in the glioma discovery data set were annotated for their potential to regulate nearby genes. We used HaploReg and RegulomeDB to examine whether any of these SNPs were located in putative promoters, enhancers, DNase I hypersensitivity sites or transcription factor binding sites. Sequence conservation was assessed using Genomic Evolutionary Rate Profiling (GERP). Associations between SNP genotype and mRNA expression levels were assessed with Spearman’s $\rho$ and analyzed using the Stranger et al. data set and tools available in Genevar (3.3.0). Analyses were restricted to HapMap CEPH lymphoblastoid samples, and a $P$-value threshold of 0.05 was used.

34. Cawthon, R.M. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Res. 37, e21 (2009).
35. Cawthon, R.M. Telomere measurement by quantitative PCR. Nucleic Acids Res. 30, e47 (2002).
36. Hansen, H.M., Wiemels, J.L., Wrensch, M. & Wiencke, J.K. DNA quantification of whole genome amplified samples for genotyping on a multiplexed bead array platform. Cancer Epidemiol. Biomarkers Prev. 16, 1686–1690 (2007).
37. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 5, e1000529 (2009).
38. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 467, 1061–1073 (2010).
39. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. Nat. Rev. Genet. 11, 499–511 (2010).
40. Price, A.L. et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38, 904–909 (2006).
41. Liu, J.Z. et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. Nat. Genet. 42, 436–440 (2010).
42. Higgins, J.P., Thompson, S.G., Deeks, J.J. & Altman, D.G. Measuring inconsistency in meta-analyses. Br. Med. J. 327, 557–560 (2003).
43. Pe’er, I., Yelensky, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet. Epidemiol. 32, 381–385 (2008).
44. Ward, L.D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 40, D930–D934 (2012).
45. Boyle, A.P. et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 22, 1790–1797 (2012).
46. Cooper, G.M. et al. Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 15, 901–913 (2005).
47. Stranger, B.E. et al. Patterns of cis-regulatory variation in diverse human populations. PLoS Genet. 8, e1002639 (2012).
48. Yang, T.P. et al. Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. Bioinformatics 26, 2474–2476 (2010).