addition of imputation of the 19 SNPs in the 6 regions under study here (28).

AUTHORS' CONTRIBUTIONS

Conceived and designed the experiments: Z.W., N.C., L.T.A. Performed the experiments: Z.W., B.Z., M.Z., H.P., J.J., C.C.C., J.N.S., J.W.H., A.H., L.B., A.I., C.H., L.T.A. Analyzed the data: Z.W., B.Z., M.Z., H.P., J.J., C.C.C., J.N.S., J.W.H., M.Y., N.C., L.T.A. Contributed reagents/materials/analysis tools: all authors. Wrote the paper: ZW and LTA. Contributed to the writing of the paper: all authors.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

The authors acknowledge the contribution of the staff of the Cancer Genomics Research Laboratory for their invaluable help throughout the project.

Conflict of Interest statement: None declared.

FUNDING

This work was supported by the Intramural Research Program and by contract number HHSN261200800001E of the US National Institutes of Health (NIH), National Cancer Institute. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services nor does mention of trade names, commercial products or organizations imply endorsement by the U.S. Government. Additional funding acknowledgements are listed in Supplementary Material. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

REFERENCES

1. Rothman, N., Garcia-Closas, M., Chatterjee, N., Malats, N., Wu, X., Figueroa, J.D., Real, F.X., Van Den Berg, D., Matullo, G., Baris, D. *et al.* (2010) A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. *Nat. Genet.*, **42**, 978–984.
2. Haiman, C.A., Chen, G.K., Vachon, C.M., Canzian, F., Dunning, A., Millikan, R.C., Wang, X., Ademuyiwa, F., Ahmed, S., Ambrosone, C.B. *et al.* (2011) A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. *Nat. Genet.*, **43**, 1210–1214.
3. Shete, S., Hosking, F.J., Robertson, L.B., Dobbins, S.E., Sanson, M., Malmer, B., Simon, M., Marie, Y., Boisselier, B., Delattre, J.Y. *et al.* (2009) Genome-wide association study identifies five susceptibility loci for glioma. *Nat. Genet.*, **41**, 899–904.
4. Wang, Y., Broderick, P., Webb, E., Wu, X., Vijayakrishnan, J., Matakidou, A., Qureshi, M., Dong, Q., Gu, X., Chen, W.V. *et al.* (2008) Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat. Genet.*, **40**, 1407–1409.
5. McKay, J.D., Hung, R.J., Gaborieau, V., Boffetta, P., Chabrier, A., Byrnes, G., Zaridze, D., Mukeria, A., Szeszenia-Dabrowska, N., Lissowska, J. *et al.* (2008) Lung cancer susceptibility locus at 5p15.33. *Nat. Genet.*, **40**, 1404–1406.
6. Broderick, P., Wang, Y., Vijayakrishnan, J., Matakidou, A., Spitz, M.R., Eisen, T., Amos, C.I. and Houlston, R.S. (2009) Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res.*, **69**, 6633–6641.
7. Landi, M.T., Chatterjee, N., Yu, K., Goldin, L.R., Goldstein, A.M., Rotunno, M., Mirabello, L., Jacobs, K., Wheeler, W., Yeager, M. *et al.* (2009) A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am. J. Hum. Genet.*, **85**, 679–691.
8. Beesley, J., Pickett, H.A., Johnatty, S.E., Dunning, A.M., Chen, X., Li, J., Michailidou, K., Lu, Y., Rider, D.N., Palmieri, R.T. *et al.* (2011) Functional polymorphisms in the TERT promoter are associated with risk of serous epithelial ovarian and breast cancers. *PLoS ONE*, **6**, e24987.
9. Rafnar, T., Sulem, P., Stacey, S.N., Geller, F., Gudmundsson, J., Sigurdsson, A., Jakobsdottir, M., Helgadottir, H., Thorlacius, S., Aben, K.K. *et al.* (2009) Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat. Genet.*, **41**, 221–227.
10. Stacey, S.N., Sulem, P., Masson, G., Gudjonsson, S.A., Thorleifsson, G., Jakobsdottir, M., Sigurdsson, A., Gudbjartsson, D.F., Sigurgeirsson, B., Benediktsdottir, K.R. *et al.* (2009) New common variants affecting susceptibility to basal cell carcinoma. *Nat. Genet.*, **41**, 909–914.
11. Yang, X., Yang, B., Li, B. and Liu, Y. (2013) Association between TERT-CLPTM1L rs401681[C] allele and NMSC cancer risk: a meta-analysis including 45,184 subjects. *Arch. Dermatol. Res.*, **1**, 49–52.
12. Petersen, G.M., Amundadottir, L., Fuchs, C.S., Kraft, P., Stolzenberg-Solomon, R.Z., Jacobs, K.B., Arslan, A.A., Bueno-de-Mesquita, H.B., Gallinger, S., Gross, M. *et al.* (2010) A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat. Genet.*, **42**, 224–228.
13. Kote-Jarai, Z., Olama, A.A., Giles, G.G., Severi, G., Schleutker, J., Weischer, M., Campa, D., Riboli, E., Key, T., Gronberg, H. *et al.* (2011) Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study. *Nat. Genet.*, **43**, 785–791.
14. Turnbull, C., Rapley, E.A., Seal, S., Pernet, D., Renwick, A., Hughes, D., Ricketts, M., Linger, R., Nsengimana, J., Deloukas, P. *et al.* (2010) Variants near DMRT1, TERT and ATF7IP are associated with testicular germ cell cancer. *Nat. Genet.*, **42**, 604–607.
15. Mocellin, S., Verdi, D., Pooley, K.A., Landi, M.T., Egan, K.M., Baird, D.M., Prescott, J., De Vivo, I. and Nitti, D. (2012) Telomerase reverse transcriptase locus polymorphisms and cancer risk: a field synopsis and meta-analysis. *J. Natl. Cancer Inst.*, **104**, 840–854.
16. Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L. and Shay, J.W. (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science*, **266**, 2011–2015.
17. Yamamoto, K., Okamoto, A., Isonishi, S., Ochiai, K. and Ohtake, Y. (2001) A novel gene, CRR9, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. *Biochem. Biophys. Res. Commun.*, **280**, 1148–1154.
18. James, M.A., Wen, W., Wang, Y., Byers, L.A., Heymach, J.V., Coombes, K.R., Girard, L., Minna, J. and You, M. (2012) Functional characterization of CLPTM1L as a lung cancer risk candidate gene in the 5p15.33 locus. *PLoS ONE*, **7**, e36116.
19. Ni, Z., Tao, K., Chen, G., Chen, Q., Tang, J., Luo, X., Yin, P. and Wang, X. (2012) CLPTM1L is overexpressed in lung cancer and associated with apoptosis. *PLoS ONE*, **7**, e52598.
20. Jia, J., Bosley, A.D., Thompson, A., Hoskins, J.W., Cheuk, A., Collins, I., Parikh, H., Xiao, Z., Ylaya, K., Dzyadyk, M. *et al.* (2014) CLPTM1L promotes growth and enhances aneuploidy in pancreatic cancer cells. *Cancer Res.*, **74**, 2785–2795.
21. James, M.A., Vikis, H.G., Tate, E., Rymaszewski, A.L. and You, M. (2014) CRR9/CLPTM1L Regulates Cell Survival Signaling and Is Required for Ras Transformation and Lung Tumorigenesis. *Cancer Res.*, **74**, 1116–1127.
22. Ballew, B.J. and Savage, S.A. (2013) Updates on the biology and management of dyskeratosis congenita and related telomere biology disorders. *Expert Rev. Hematol.*, **6**, 327–337.
23. Armanios, M. and Blackburn, E.H. (2012) The telomere syndromes. *Nat. Rev. Genet.*, **13**, 693–704.
24. Bhattacharjee, S., Rajaraman, P., Jacobs, K.B., Wheeler, W.A., Melin, B.S., Hartge, P., Yeager, M., Chung, C.C., Chanock, S.J. and Chatterjee, N. (2012) A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am. J. Hum. Genet.*, **90**, 821–835.

25. Genomes. (2010) A map of human genome variation from population-scale sequencing. *Nature*, **467**, 1061–1073.
26. Wang, Z., Jacobs, K.B., Yeager, M., Hutchinson, A., Sampson, J., Chatterjee, N., Albanes, D., Berndt, S.I., Chung, C.C., Diver, W.R. *et al.* (2012) Improved imputation of common and uncommon SNPs with a new reference set. *Nat. Genet.*, **44**, 6–7.
27. Rajaraman, P., Melin, B.S., Wang, Z., McKean-Cowdin, R., Michaud, D.S., Wang, S.S., Bondy, M., Houlston, R., Jenkins, R.B., Wrensch, M. *et al.* (2012) Genome-wide association study of glioma and meta-analysis. *Hum. Genet.*, **131**, 1877–1888.
28. Shi, J., Marconett, C.N., Duan, J., Hyland, P.L., Li, P., Wang, Z., Wheeler, W., Zhou, B., Campan, M., Lee, D.S. *et al.* (2014) Characterizing the genetic basis of methylome diversity in histologically normal human lung tissue. *Nat. Commun.*, **5**, 3365.
29. Cheung, A.L. and Deng, W. (2008) Telomere dysfunction, genome instability and cancer. *Front. Biosci.*, **13**, 2075–2090.
30. Cesare, A.J. and Reddel, R.R. (2010) Alternative lengthening of telomeres: models, mechanisms and implications. *Nat. Rev. Genet.*, **11**, 319–330.
31. Shay, J.W. and Bacchetti, S. (1997) A survey of telomerase activity in human cancer. *Eur. J. Cancer*, **33**, 787–791.
32. Hou, L., Zhang, X., Gawron, A.J. and Liu, J. (2012) Surrogate tissue telomere length and cancer risk: shorter or longer? *Cancer Lett.*, **319**, 130–135.
33. Pooley, K.A., Sandhu, M.S., Tyrer, J., Shah, M., Driver, K.E., Luben, R.N., Bingham, S.A., Ponder, B.A., Pharoah, P.D., Khaw, K.T. *et al.* (2010) Telomere length in prospective and retrospective cancer case-control studies. *Cancer Res.*, **70**, 3170–3176.
34. De Vivo, I., Prescott, J., Wong, J.Y., Kraft, P., Hankinson, S.E. and Hunter, D.J. (2009) A prospective study of relative telomere length and postmenopausal breast cancer risk. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 1152–1156.
35. Weischer, M., Nordestgaard, B.G., Cawthon, R.M., Freiberg, J.J., Tybjaerg-Hansen, A. and Bojesen, S.E. (2013) Short telomere length, cancer survival, and cancer risk in 47102 individuals. *J. Natl. Cancer Inst.*, **105**, 459–468.
36. Slagboom, P.E., Droog, S. and Boomsma, D.I. (1994) Genetic determination of telomere size in humans: a twin study of three age groups. *Am. J. Hum. Genet.*, **55**, 876–882.
37. Andrew, T., Aviv, A., Falchi, M., Surdulescu, G.L., Gardner, J.P., Lu, X., Kimura, M., Kato, B.S., Valdes, A.M. and Spector, T.D. (2006) Mapping genetic loci that determine leukocyte telomere length in a large sample of unselected female sibling pairs. *Am. J. Hum. Genet.*, **78**, 480–486.
38. Nan, H., Qureshi, A.A., Prescott, J., De Vivo, I. and Han, J. (2011) Genetic variants in telomere-maintaining genes and skin cancer risk. *Hum. Genet.*, **129**, 247–253.
39. Melin, B.S., Nordfjall, K., Andersson, U. and Roos, G. (2012) hTERT cancer risk genotypes are associated with telomere length. *Genet. Epidemiol.*, **36**, 368–372.
40. Hsu, C.P., Hsu, N.Y., Lee, L.W. and Ko, J.L. (2006) Ets2 binding site single nucleotide polymorphism at the hTERT gene promoter – effect on telomerase expression and telomere length maintenance in non-small cell lung cancer. *Eur. J. Cancer*, **42**, 1466–1474.
41. Horn, S., Figl, A., Rachakonda, P.S., Fischer, C., Sucker, A., Gast, A., Kadel, S., Moll, I., Nagore, E., Hemminki, K. *et al.* (2013) TERT promoter mutations in familial and sporadic melanoma. *Science*, **339**, 959–961.
42. Killela, P.J., Reitman, Z.J., Jiao, Y., Bettegowda, C., Agrawal, N., Diaz, L.A. Jr, Friedman, A.H., Friedman, H., Gallia, G.L., Giovanella, B.C. *et al.* (2013) 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.
43. Simon, M., Hosking, F.J., Marie, Y., Gousias, K., Boisselier, B., Carpentier, C., Schramm, J., Mokhtari, K., Hoang-Xuan, K., Idbaih, A. *et al.* (2010) Genetic risk profiles identify different molecular etiologies for glioma. *Clin. Cancer Res.*, **16**, 5252–5259.
44. Yang, P., Li, Y., Jiang, R., Cunningham, J.M., Zhang, F. and de Andrade, M. (2010) A rigorous and comprehensive validation: common genetic variations and lung cancer. *Cancer Epidemiol. Biomarkers Prev.*, **19**, 240–244.
45. Hu, Z., Wu, C., Shi, Y., Guo, H., Zhao, X., Yin, Z., Yang, L., Dai, J., Hu, L., Tan, W. *et al.* (2011) A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. *Nat. Genet.*, **43**, 792–796.
46. Miki, D., Kubo, M., Takahashi, A., Yoon, K.A., Kim, J., Lee, G.K., Zo, J.I., Lee, J.S., Hosono, N., Morizono, T. *et al.* (2010) Variation in TP63 is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations. *Nat. Genet.*, **42**, 893–896.
47. Pande, M., Spitz, M.R., Wu, X., Gorlov, I.P., Chen, W.V. and Amos, C.I. (2011) Novel genetic variants in the chromosome 5p15.33 region associate with lung cancer risk. *Carcinogenesis*, **32**, 1493–1499.
48. Johnatty, S.E., Beesley, J., Chen, X., Macgregor, S., Duffy, D.L., Spurdle, A.B., deFazio, A., Gava, N., Webb, P.M., Rossing, M.A. *et al.* (2010) Evaluation of candidate stromal epithelial cross-talk genes identifies association between risk of serous ovarian cancer and TERT, a cancer susceptibility 'hot-spot'. *PLoS Genet.*, **6**, e1001016.
49. Zhao, Y., Chen, G., Song, X., Chen, H., Mao, Y. and Lu, D. (2012) Fine-mapping of a region of chromosome 5p15.33 (TERT-CLPTM1L) suggests a novel locus in TERT and a CLPTM1L haplotype are associated with glioma susceptibility in a Chinese population. *Int. J. Cancer*, **131**, 1569–1576.
50. Shiraishi, K., Kunitoh, H., Daigo, Y., Takahashi, A., Goto, K., Sakamoto, H., Ohnami, S., Shimada, Y., Ashikawa, K., Saito, A. *et al.* (2012) A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population. *Nat. Genet.*, **44**, 900–903.
51. Palmer, J.R., Ruiz-Narvaez, E.A., Rotimi, C.N., Cupples, L.A., Cozier, Y.C., Adams-Campbell, L.L. and Rosenberg, L. (2012) Genetic susceptibility loci for subtypes of breast cancer in an African American population. *Cancer Epidemiol. Biomarkers Prev.*, **22**, 127–34.
52. Barrett, J.H., Iles, M.M., Harland, M., Taylor, J.C., Aitken, J.F., Andresen, P.A., Akslen, L.A., Armstrong, B.K., Avril, M.F., Azizi, E. *et al.* (2011) Genome-wide association study identifies three new melanoma susceptibility loci. *Nat. Genet.*, **43**, 1108–1113.
53. Bojesen, S.E., Pooley, K.A., Johnatty, S.E., Beesley, J., Michailidou, K., Tyrer, J.P., Edwards, S.L., Pickett, H.A., Shen, H.C., Smart, C.E. *et al.* (2013) Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat. Genet.*, **45**, 371–384.
54. Kote-Jarai, Z., Saunders, E.J., Leongamornlert, D.A., Tymrakiewicz, M., Dadaev, T., Jugurnauth-Little, S., Ross-Adams, H., Al Olama, A.A., Benlloch, S., Halim, S. *et al.* (2013) Fine-mapping identifies multiple prostate cancer risk loci at 5p15, one of which associates with TERT expression. *Hum. Mol. Genet.*, **22**, 2520–2528.
55. Grundberg, E., Meduri, E., Sandling, J.K., Hedman, A.K., Keildson, S., Buil, A., Busche, S., Yuan, W., Nisbet, J., Sekowska, M. *et al.* (2013) Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *Am. J. Hum. Genet.*, **93**, 876–890.
56. Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M. *et al.* (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.*, **41**, 178–186.
57. Jjingo, D., Conley, A.B., Yi, S.V., Lunyak, V.V. and Jordan, I.K. (2012) On the presence and role of human gene-body DNA methylation. *Oncotarget*, **3**, 462–474.
58. Rideout, W.M. III, Coetzee, G.A., Olumi, A.F. and Jones, P.A. (1990) 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science*, **249**, 1288–1290.
59. Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y. *et al.* (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, **466**, 253–257.
60. Shen, H. and Laird, P.W. (2013) Interplay between the cancer genome and epigenome. *Cell*, **153**, 38–55.
61. Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M. and Smoller, J.W. (2013) Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.*, **14**, 483–495.
62. Figueroa JD, Y.Y., Siddiq, A., Garcia-Closas, M., Chatterjee, N. *et al.* (2014) Genome-wide association study identifies multiple loci associated with bladder cancer risk. *Hum. Mol. Genet.*, **23**, 1387–1398.
63. Hunter, D.J., Kraft, P., Jacobs, K.B., Cox, D.G., Yeager, M., Hankinson, S.E., Wacholder, S., Wang, Z., Welch, R., Hutchinson, A. *et al.* (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat. Genet.*, **39**, 870–874.
64. Siddiq, A., Couch, F.J., Chen, G.K., Lindstrom, S., Eccles, D., Millikan, R.C., Michailidou, K., Stram, D.O., Beckmann, L., Rhie, S.K. *et al.* (2012) A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum. Mol. Genet.*, **21**, 5373–5384.

65. De Vivo, I., Prescott, J., Setiawan, V.W., Olson, S.H. and Wentzensen, N., Australian National Endometrial Cancer Study Group, Attia, J., Black, A., Brinton, L., Chen, C. *et al.* (2014) Genome-wide association study of endometrial cancer in E2C2. *Hum. Genet. 133*, 211–24.
66. Abnet, C.C., Freedman, N.D., Hu, N., Wang, Z., Yu, K., Shu, X.O., Yuan, J.M., Zheng, W., Dawsey, S.M., Dong, L.M. *et al.* (2010) A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nat. Genet.*, **42**, 764–767.
67. Hinch, A.G., Tandon, A., Patterson, N., Song, Y., Rohland, N., Palmer, C.D., Chen, G.K., Wang, K., Buxbaum, S.G., Akylbekova, E.L. *et al.* (2011) The landscape of recombination in African Americans. *Nature*, **476**, 170–175.
68. Hsiung, C.A., Lan, Q., Hong, Y.C., Chen, C.J., Hosgood, H.D., Chang, I.S., Chatterjee, N., Brennan, P., Wu, C., Zheng, W. *et al.* (2010) The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. *PLoS Genet.*, **6**.
69. Lan, Q., Hsiung, C.A., Matsuo, K., Hong, Y.C., Seow, A., Wang, Z., Hosgood, H.D. 3rd, Chen, K., Wang, J.C., Chatterjee, N. *et al.* (2012) Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. *Nat. Genet.*, **44**, 1330–1335.
70. Savage, S.A., Mirabello, L., Wang, Z., Gastier-Foster, J.M., Gorlick, R., Khanna, C., Flanagan, A.M., Tirabosco, R., Andrulis, I.L., Wunder, J.S. *et al.* (2013) Genome-wide association study identifies two susceptibility loci for osteosarcoma. *Nat. Genet.*, **45**, 799–803.
71. Bolton, K.L., Tyrer, J., Song, H., Ramus, S.J., Notaridou, M., Jones, C., Sher, T., Gentry-Maharaj, A., Wozniak, E., Tsai, Y.Y. *et al.* (2010) Common variants at 19p13 are associated with susceptibility to ovarian cancer. *Nat. Genet.*, **42**, 880–884.
72. Amundadottir, L., Kraft, P., Stolzenberg-Solomon, R.Z., Fuchs, C.S., Petersen, G.M., Arslan, A.A., Bueno-de-Mesquita, H.B., Gross, M., Helzlsouer, K., Jacobs, E.J. *et al.* (2009) Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat. Genet.*, **41**, 986–990.
73. Wu, C., Miao, X., Huang, L., Che, X., Jiang, G., Yu, D., Yang, X., Cao, G., Hu, Z., Zhou, Y. *et al.* (2012) Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. *Nat. Genet.*, **44**, 62–66.
74. Thomas, G., Jacobs, K.B., Yeager, M., Kraft, P., Wacholder, S., Orr, N., Yu, K., Chatterjee, N., Welch, R., Hutchinson, A. *et al.* (2008) Multiple loci identified in a genome-wide association study of prostate cancer. *Nat. Genet.*, **40**, 310–315.
75. Schumacher, F.R., Berndt, S.I., Siddiq, A., Jacobs, K.B., Wang, Z., Lindstrom, S., Stevens, V.L., Chen, C., Mondul, A.M., Travis, R.C. *et al.* (2011) Genome-wide association study identifies new prostate cancer susceptibility loci. *Hum. Mol. Genet.*, **20**, 3867–3875.
76. Purdue, M.P., Johansson, M., Zelenika, D., Toro, J.R., Scelo, G., Moore, L.E., Prokhortchouk, E., Wu, X., Kiemeney, L.A., Gaborieau, V. *et al.* (2011) Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. *Nat. Genet.*, **43**, 60–65.
77. Schumacher, F.R., Wang, Z., Skotheim, R.I., Koster, R., Chung, C.C., Hildebrandt, M.A., Kratz, C.P., Bakken, A.C., Timothy Bishop, D., Cook, M.B. *et al.* (2013) Testicular germ cell tumor susceptibility associated with the UCK2 locus on chromosome 1q23. *Hum. Mol. Genet.*, **22**, 48–53.
78. Marchini, J. and Howie, B. (2010) Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.*, **11**, 499–511.
79. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.*, **39**, 906–913.
80. Fearnhead, P. (2006) SequenceLDhot: detecting recombination hotspots. *Bioinformatics*, **22**, 3061–3066.
81. Fearnhead, P. and Donnelly, P. (2002) Approximate likelihood methods for estimating local recombination rates. *J. R. Stat. Soc.*, **64**, 657–680.
82. Li, N. and Stephens, M. (2003) Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, **165**, 2213–2233.
83. Crawford, D.C., Bhangale, T., Li, N., Hellenthal, G., Rieder, M.J., Nickerson, D.A. and Stephens, M. (2004) Evidence for substantial fine-scale variation in recombination rates across the human genome. *Nat. Genet.*, **36**, 700–706.
84. Browning, B.L. and Browning, S.R. (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.*, **84**, 210–223.
85. Ward, L.D. and Kellis, M. (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.*, **40**, D930–934.
86. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S. *et al.* (2012) Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.*, **22**, 1790–1797.
87. Grundberg, E., Small, K.S., Hedman, A.K., Nica, A.C., Buil, A., Keildson, S., Bell, J.T., Yang, T.P., Meduri, E., Barrett, A. *et al.* (2012) Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat. Genet.*, **44**, 1084–1089.
88. Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P. and Brunak, S. (1996) Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. *Nucleic Acids Res.*, **24**, 3439–3452.
89. Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B*, **57**, 289–300.

UICC global cancer control

IJC
International Journal of Cancer

Genetic variants associated with longer telomere length are associated with increased lung cancer risk among never-smoking women in Asia: a report from the female lung cancer consortium in Asia

Mitchell J. Machiela[1], Chao Agnes Hsiung[2], Xiao-Ou Shu[3,4], Wei Jie Seow[1], Zhaoming Wang[5], Keitaro Matsuo[6], Yun-Chul Hong[7], Adeline Seow[8], Chen Wu[9,10], H. Dean Hosgood III[11], Kexin Chen[12], Jiu-Cun Wang[13,14], Wanqing Wen[3,4], Richard Cawthon[15], Nilanjan Chatterjee[1], Wei Hu[1], Neil E. Caporaso[1], Jae Yong Park[16], Chien-Jen Chen[17], Yeul Hong Kim[18], Young Tae Kim[19], Maria Teresa Landi[1], Hongbing Shen[20,21], Charles Lawrence[22], Laurie Burdett[5], Meredith Yeager[5], I-Shou Chang[23], Tetsuya Mitsudomi[24], Hee Nam Kim[25], Gee-Chen Chang[26,27], Bryan A. Bassig[1,28], Margaret Tucker[1], Fusheng Wei[29], Zhihua Yin[30], She-Juan An[31], Biyun Qian[32], Victor Ho Fun Lee[33], Daru Lu[13], Jianjun Liu[34,35], Hyo-Sung Jeon[36], Chin-Fu Hsiao[2,37], Jae Sook Sung[38], Jin Hee Kim[39], Yu-Tang Gao[40], Ying-Huang Tsai[41], Yoo Jin Jung[19], Huan Guo[42], Zhibin Hu[20,21], Amy Hutchinson[5], Wen-Chang Wang[2], Robert J. Klein[43], Charles C. Chung[1], In-Jae Oh[44,45], Kuan-Yu Chen[46], Sonja I. Berndt[1], Wei Wu[30], Jiang Chang[9], Xu-Chao Zhang[31], Ming-Shyan Huang[47], Hong Zheng[12], Junwen Wang[48,49], Xueying Zhao[13], Yuqing Li[50], Jin Eun Choi[51], Wu-Chou Su[52], Kyong Hwa Park[18], Sook Whan Sung[53],

Key words: association study, genetics, lung cancer, telomere length, genetic risk score

Abbreviations: DNA: deoxyribonucleic acid; GRS: genetic risk score; GWAS: genome-wide association study; SKAT: snp-set kernel association test; SNP: single nucleotide polymorphism; T/S: telomere/single copy gene; WBC: white blood cell

Conflicts of interest: Richard Cawthon holds a patent for the polymerase chain reaction method of measuring telomere length that is used in this study, and licensed that method for commercial use. However, throughout this study his laboratory has been blind as to the age, genotype, and outcomes of the subjects in the study.

M.J.M., C.A.H., X.-O.S., W.J.S., Z.W., K.M., Y.-C.H., A.S., C.W., H.D.H., K.C. and J.-C.W. contributed equally to this work.

T.W., M.P.W., Y-L.W., P-C.Y., B.Z., M-H.S., J.F.F., W.Z., D.L., S.J.C., N.R., Q.L. jointly directed this work.

Grant sponsor: Ministry of Health); **Grant number:** 201002007; **Grant sponsor:** Ministry of Science and Technology; **Grant number:** 2011BAI09B00; **Grant sponsor:** National S&T Major Special Project; **Grant number:** 2011ZX09102-010-01; **Grant sponsor:** China National High-Tech Research and Development Program; **Grant numbers:** 2012AA02A517, 2012AA02A518; **Grant sponsor:** National Science Foundation of China; **Grant number:** 30890034; **Grant sponsor:** National Basic Research Program; **Grant number:** 2012CB944600; **Grant sponsor:** Scientific and Technological Support Plans from Jiangsu Province; **Grant number:** BE2010715; **Grant sponsor:** Foundation of Guangdong Science and Technology Department); **Grant numbers:** 2006B60101010, 2007A032000002, 2011A030400010; **Grant sponsor:** Guangzhou Science and Information Technology Bureau; **Grant number:** 2011Y2-00014; **Grant sponsor:** Chinese Lung Cancer Research Foundation, National Natural Science Foundation; **Grant number:** 81101549; **Grant sponsor:** Natural Science Foundation of Guangdong Province; **Grant number:** S2011010000792; **Grant sponsor:** National Research Program on Genomic Medicine in Taiwan); **Grant number:** DOH98-TD-G-111-015; **Grant sponsor:** National Research Program for Biopharmaceuticals in Taiwan; **Grant numbers:** DOH 100-TD-PB-111-TM013, MOST 103-2325-B-400-011; **Grant sponsor:** National Science Council, Taiwan; **Grant number:** NSC 100-2319-B-400-001; **Grant sponsor:** National Medical Research Council Singapore; **Grant number:** NMRC/0897/2004, NMRC/1075/2006; **Grant sponsor:** Agency for Science, Technology and Research; **Grant sponsor:** General Research Fund of Research Grant Council, Hong Kong); **Grant numbers:** 781511M, 17121414M; **Grant sponsor:** National Science Foundation, China; **Grant number:** 91229105; **Grant sponsor:** Scientific Research on Priority Areas and on Innovative Area from the Ministry of Education, Science, Sports, Culture and Technology of Japan; **Grant number:** NCI R01-CA121210; **Grant sponsor:** China National High-Tech Research and Development Program; **Grant number:** 2009AA022705; **Grant sponsor:** National Key Basic Research Program; **Grant number:** 2011CB503805; **Grant sponsor:** National Research Foundation of Korea; **Grant number:** NRF-2014R1A2A2A05003665; **Grant sponsor:** Ministry of Health &Welfare, Republic of Korea; **Grant numbers:** 0720550-2, A010250; **Grant sponsor:** National Key Basic Research and Development Program); **Grant number:** 2011CB503800; **Grant sponsor:** National Nature Science Foundation of China); **Grant number:** 81102194; **Grant sponsor:** Liaoning Provincial Department of Education; **Grant number:** LS2010168; **Grant sponsor:** China Medical Board; **Grant number:** 00726; **Grant sponsor:** National Institutes of Health); **Grant number:** R37 CA70867; **Grant sponsor:** National Cancer Institute Intramural Research Program; **Grant number:** N02 CP1101066

DOI: 10.1002/ijc.29393

History: Received 30 July 2014; Accepted 22 Oct 2014; Online 16 Dec 2014

Correspondence to: Mitchell J. Machiela, Advanced Technology Center, Rm 225, 8717 Grovemont Circle, Bethesda, MD 20892, USA, Tel.: 301-435-6391, Fax: 301-402-3134, E-mail: mitchell.machiela@nih.gov

Yuh-Min Chen[54,55], Li Liu[56], Chang Hyun Kang[19], Lingmin Hu[20,21], Chung-Hsing Chen[23], William Pao[57], Young-Chul Kim[44,45], Tsung-Ying Yang[27], Jun Xu[58], Peng Guan[30], Wen Tan[9], Jian Su[31], Chih-Liang Wang[59], Haixin Li[12], Alan Dart Loon Sihoe[60], Zhenhong Zhao[13], Ying Chen[8], Yi Young Choi[51], Jen-Yu Hung[47], Jun Suk Kim[61], Ho-Il Yoon[62], Qiuyin Cai[3,4], Chien-Chung Lin[52], In Kyu Park[19], Ping Xu[63], Jing Dong[20,21], Christopher Kim[1], Qincheng He[30], Reury-Perng Perng[54], Takashi Kohno[64], Sun-Seog Kweon[65,66], Chih-Yi Chen[67], Roel C.H Vermeulen[68], Junjie Wu[13], Wei-Yen Lim[8], Kun-Chieh Chen[27], Wong-Ho Chow[1], Bu-Tian Ji[1], John K. C. Chan[69], Minjie Chu[20,21], Yao-Jen Li[17], Jun Yokota[64,70], Jihua Li[71], Hongyan Chen[13], Yong-Bing Xiang[72], Chong-Jen Yu[46], Hideo Kunitoh[73], Guoping Wu[29], Li Jin[13], Yen-Li Lo[2], Kouya Shiraishi[64], Ying-Hsiang Chen[2], Hsien-Chih Lin[2], Tangchun Wu[42], Maria Pik Wong[74], Yi-Long Wu[31], Pan-Chyr Yang[75], Baosen Zhou[30], Min-Ho Shin[66], Joseph F. Fraumeni Jr.[1], Wei Zheng[3,4], Dongxin Lin[9], Stephen J. Chanock[1], Nathaniel Rothman[1] and Qing Lan[1]

[1] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD
[2] Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan
[3] Department of Medicine, Vanderbilt University Medical Center, Nashville, TN
[4] Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN
[5] Cancer Genomics Research Laboratory, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD
[6] Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan
[7] Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea
[8] Saw Swee Hock School of Public Health, National University of Singapore, Republic of Singapore
[9] Department of Etiology and Carcinogenesis, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[10] State Key Laboratory of Molecular Oncology, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[11] Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY
[12] Department of Epidemiology and Biostatistics, Key Laboratory of Cancer Prevention and Therapy, National Clinical Research Center of Cancer, Medical University Cancer Institute and Hospital, Tianjin, China
[13] Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, China
[14] State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai, China
[15] Department of Human Genetics, University of Utah, Salt Lake City, UT
[16] Lung Cancer Center, Kyungpook National University Medical Center, Daegu, Republic of Korea
[17] Genomic Research Center, Taipei, Taiwan
[18] Department of Internal Medicine, Division of Oncology/Hematology, College of Medicine, Korea University Anam Hospital, Seoul, Republic of Korea
[19] Department of Thoracic and Cardiovascular Surgery, Cancer Research Institute, Seoul National University College of Medicine, Seoul, Republic of Korea
[20] Jiangsu Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, China
[21] Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, Cancer Center, School of Public Health, Nanjing Medical University, Nanjing, China
[22] Westat, Rockville, MD
[23] National Institute of Cancer Research, National Health Research Institutes, Zhunan, Taiwan
[24] Division of Thoracic Surgery, Kinki University School of Medicine, Sayama, Japan
[25] Center for Creative Biomedical Scientists, Chonnam National University, Gwangju, Republic of Korea
[26] Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan
[27] Department of Internal Medicine, Division of Chest Medicine, Taichung Veterans General Hospital, Taichung, Taiwan
[28] Division of Environmental Health Sciences, Yale School of Public Health, New Haven, CT
[29] China National Environmental Monitoring Center, Beijing, China
[30] Department of Epidemiology, School of Public Health, China Medical University, Shenyang, China
[31] Guangdong Lung Cancer Institute, Guangdong Provincial Key Laboratory of Translational Medicine in Lung Cancer, Guangdong General Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China
[32] Department of Epidemiology and Biostatistics, School of Public Health, Shanghai Jiao Tong University, Shanghai, China
[33] Department of Clinical Oncology, Li Ka Shing Faculty of Medicine, the University of Hong Kong, HongKong, China
[34] Department of Human Genetics, Genome Institute of Singapore, Republic of Singapore
[35] School of Life Sciences, Anhui Medical University, Hefei, China
[36] Molecular Diagnostics and Imaging Center, Kyungpook National University, Daegu, Republic of Korea
[37] Taiwan Lung Cancer Tissue/Specimen Information Resource Center, National Health Research Institutes, Zhunan, Taiwan
[38] Cancer Research Institute, Korea University, Seoul, Republic of Korea
[39] Department of Environmental Health, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea
[40] Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China
[41] Department of Respiratory Thearpy, Chang Gung Memorial Hospital, Chiayi, Taiwan
[42] Institute of Occupational Medicine and Ministry of Education Key Lab for Environment and Health, School of Public Health, Huazhong University of Science and Technology, Wuhan, China
[43] Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, Icahn Institute for Genomics and Multiscale Biology, New York, NY
[44] Lung and Esophageal Cancer Clinic, Chonnam National University Hwasun Hospital, Hwasun-Eup, Republic of Korea

[45] Department of Internal Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea
[46] Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan
[47] Department of Internal Medicine, Kaohsiung Medical University Hospital, School of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan
[48] Centre for Genomic Sciences, Li Ka Shing Faculty of Medicine, the University of Hong Kong, Hong Kong, China
[49] Department of Biochemistry, Li Ka Shing Faculty of Medicine, the University of Hong Kong, Hong Kong, China
[50] Cancer Prevention Institute of California, Fremont, CA
[51] Cancer Research Center, Kyungpook National University Medical Center, Daegu, Republic of Korea
[52] Department of Internal Medicine, Cancer Center, National Cheng Kung University Hospital, National College of Medicine, Cheng Kung University, Tainan, Taiwan
[53] Department of Thoracic and Cardiovascular Surgery, Seoul St Mary's Hospital, Seoul, Republic of Korea
[54] Department of Chest Medicine, Taipei Veterans General Hospital, Taipei, Taiwan
[55] Taipei Cancer Center, Taipei Medical University, Taipei, Taiwan
[56] Cancer Center, Union Hospital, Huazhong University of Science and Technology, Wuhan, China
[57] Division of Hematology and Oncology, Vanderbilt University Medical Center, Nashville, TN
[58] School of Public Health, the University of Hong Kong, Hong Kong
[59] Department of Pulmonary and Critical Care, Chang Gung Memorial Hospital, Taoyuan, Taiwan
[60] Department of Surgery, Division of Cardiothoracic Surgery, Queen Mary Hospital, Hong Kong, China
[61] Department of Internal Medicine, Division of Medical Oncology, College of Medicine, Korea University Guro Hospital, Seoul, Republic of Korea
[62] Department of Internal Medicine, Seoul National University Bundang Hospital, , Seongnam, Republic of Korea
[63] Department of Oncology, Wuhan Iron and Steel Corporation Staff Worker Hospital, Wuhan, China
[64] Division of Genome Biology, National Cancer Center Research Institute, Tokyo, Japan
[65] Jeonnam Regional Cancer Center, Chonnam National University Hwasun Hospital, Hwasun-Eup, Republic of Korea
[66] Department of Preventive Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea
[67] Institute of Medicine, Division of Thoracic Surgery, Department of Surgery, Chung Shan Medical University, Chung Shan Medical University Hospital, Taichung, Taiwan
[68] Division of Environmental Epidemiology, Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands
[69] Department of Pathology, Queen Elizabeth Hospital, Hong Kong, China
[70] Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Barcelona, Spain
[71] Qujing Center for Diseases Control and Prevention, Sanjiangdadao, Qujing, China
[72] Department of Epidemiology, Shanghai Cancer Institute, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China
[73] Department of Medical Oncology, Japanese Red Cross Medical Center, Tokyo, Japan
[74] Department of Pathology, Li Ka Shing Faculty of Medicine, the University of Hong Kong, Hong Kong, China
[75] Department of Internal Medicine, National Taiwan University College of Medicine, Taipei, Taiwan

Recent evidence from several relatively small nested case-control studies in prospective cohorts shows an association between longer telomere length measured phenotypically in peripheral white blood cell (WBC) DNA and increased lung cancer risk. We sought to further explore this relationship by examining a panel of seven telomere-length associated genetic variants in a large study of 5,457 never-smoking female Asian lung cancer cases and 4,493 never-smoking female Asian controls using data from a previously reported genome-wide association study. Using a group of 1,536 individuals with phenotypically measured telomere length in WBCs in the prospective Shanghai Women's Health study, we demonstrated the utility of a genetic risk score (GRS) of seven telomere-length associated variants to predict telomere length in an Asian population. We then found that GRSs used as instrumental variables to predict longer telomere length were associated with increased lung cancer risk (OR = 1.51 (95% CI = 1.34–1.69) for upper *vs.* lower quartile of the weighted GRS, p value = 4.54×10^{-14}) even after removing rs2736100 (p value = 4.81×10^{-3}), a SNP in the *TERT* locus robustly associated with lung cancer risk in prior association studies. Stratified analyses suggested the effect of the telomere-associated GRS is strongest among younger individuals. We found no difference in GRS effect between adenocarcinoma and squamous cell subtypes. Our results indicate that a genetic background that favors longer telomere length may increase lung cancer risk, which is consistent with earlier prospective studies relating longer telomere length with increased lung cancer risk.

Telomeres are specialized chromatin structures that shorten during each round of cellular division in mammalian cells. Prolonged erosion of telomere length can lead to genetic instability, cellular senescence and apoptosis.[1] Earlier studies, mainly retrospective, on peripheral white blood cells (WBCs) have suggested increased cancer risk associated with shorter telomere length.[2–6] These studies may suffer from disease bias in which telomere shortening was a consequence of tumor growth and progression rather than a risk factor for tumorigenesis. Recent, primarily prospective studies indicate that, contrary to expectation, longer telomere length may be associated with cancer risk,[7–14] particularly for lung cancer.[15–18]

What's new?

The possibility for a relationship between telomere length and cancer is intriguing, but many questions remain, including whether short or long telomeres are involved. Here, a genetic risk score derived from seven telomere-length associated genetic variants revealed a positive association between telomere length and lung cancer risk in Asian women who never smoked. The genetic proxy was unaffected by reverse-causation bias or environmental exposures. The differences in telomere length captured by the variants could aid in the identification of biological mechanisms that underlie the association between longer telomere length and increased lung cancer risk.

Telomere length has historically been measured in peripheral WBC by multiplex quantitative polymerase chain reaction.[19] A recent genome-wide association study (GWAS) on telomere length has identified 7 loci robustly associated with WBC telomere length.[20] Although genetic variants at these loci explain a small proportion of the total biological variation in telomere length, the age-related shortening per variant risk allele was equivalent to 1.9–3.9 years of attrition in telomere/single copy gene (*T/S*) ratio, equating to ~57–117 base pairs in telomere length per risk allele. Furthermore, the authors demonstrated the utility of genetic risk scores (GRS) of these variants to replicate a well-established association between shorter mean peripheral WBC telomere length and coronary artery disease. This suggests that by using telomere-length associated GRS as an instrument to approximate telomere shortening or lengthening, causal relationships with telomere length can be investigated in etiologically complex diseases that include environmental risk factors associated with both disease risk and telomere length.

We herein report an investigation of the seven identified telomere-length associated variants in a sample of lung cancer cases and controls from a population of never-smoking Asian females. Our investigation uses data generated as part of a previously reported genome-wide association study (GWAS) conducted by the Female Lung Cancer Consortium in Asia.[21] Our objectives are to (*i*) validate the utility of these seven telomere-length associated variants discovered in a primarily European population to predict measured telomere length in an Asian population; (*ii*) characterize overall and individual associations of telomere-length associated variants with lung cancer risk; (*iii*) investigate the ability of GRSs of these variants to predict lung cancer risk; and (*iv*) describe the direction of the associations observed between telomere-length associated variants and lung cancer risk.

Material and Methods

Study subjects were from a published GWAS investigating lung cancer susceptibility risk in female Asian non-smokers drawn from 14 studies from mainland China, South Korea, Japan, Singapore, Taiwan, and Hong Kong.[21] Cases had histologically confirmed lung cancer. Each study was approved by the Institutional Review Board of the investigator's institution, and all participants provided written informed consent.

Genotyping was performed in the Cancer Genomics Research Laboratory of the National Cancer Institute's Division of Cancer Epidemiology and Genetics (Gaithersburg, MD); Gene-Square Biotech, Inc. (Beijing, China); GeneTech Biotech Co. (Taiwan); deCODE Genetics (Iceland); Memorial Sloan-Kettering Cancer Center (New York, NY); and Genome Institute of Singapore (Singapore). Genotyping was carried out on commercially available Illumina Infinium BeadArray human assays (Illumina 370k, Illumina 610Q, and Illumina 660W SNP microarrays) following standard procedures. The methods and quality control metrics applied to genotyping with SNP microarrays have been previously published.[21] Briefly, samples were excluded with low completion rates, extreme heterozygosity values, gender discordance, low Asian ancestry (<86%), and first degree relatives were removed. After quality control filtering, a total of 5,510 cases and 4,544 controls had genetic data available for analysis.

To address potential population substructure, principal components were calculated using the GLU struct.pca module (http://code.google.com/p/glu-genetics/) using 33,165 SNPs with low pairwise correlation ($R^2 < 0.01$).

Genotype imputation was performed to ensure complete data existed for all seven telomere-length associated variants. The IMPUTE2 program (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) was used with the March 2012 release of the 1,000 Genomes Project data[22] and the DCEG Imputation Reference Set[23] as merged references for imputation. The DCEG reference set serves as a supplement to the 1,000 Genomes reference and includes 2.8 million autosomal polymorphic SNPs for 1,249 individuals, of which 162 individuals are of Asian ancestry. Because the genotyping data was on NCBI Build 36, all genotyped variant coordinates were converted to NCBI Build 37 using UCSC's liftOver utility (http://hgdownload.cse.ucsc.edu/downloads.html) before performing genotype imputation. Recommended IMPUTE2 default settings were used and all imputed SNPs (rs7675998, rs8105767, rs755017, rs11125529) achieved INFO scores >0.99. There was no evidence for significant departures from Hardy–Weinberg proportions (p value > 0.05).

A group of subjects included in previous nested case-control studies of various cancers in the prospective Shanghai Women's Health Study ($N = 1,536$) had both genotyping data and experimentally measured peripheral WBC telomere length that we used to validate the telomere-length associated

Table 1. Age distribution, by study, of lung cancer cases and controls among never-smoking females in Asia

		Asian lung cancer study			Age group distribution				
Study group[1]	Region	N	Cases	Controls	≤39	40–49	50–59	60–69	70+
CAMSCH	Mainland China	744	541	203	8%	21%	33%	28%	10%
CNULCS	South Korea	1,036	585	451	4%	9%	30%	35%	22%
FLCS	Mainland China	522	196	326	2%	7%	39%	38%	14%
GDS	Mainland China	648	533	115	14%	19%	27%	25%	13%
GELAC	Taiwan	2,043	1,065	978	3%	13%	31%	33%	21%
GEL-S	Singapore	483	187	296	2%	9%	26%	27%	36%
HKS	Hong Kong	342	132	210	6%	13%	23%	29%	29%
JLCS	Japan	886	405	481	5%	10%	28%	51%	7%
SKLCS	South Korea	557	334	223	3%	11%	23%	41%	22%
SLCS	Mainland China	951	556	395	10%	18%	27%	34%	11%
SWHS	Mainland China	161	88	73	0%	17%	35%	47%	1%
TLCS	Mainland China	438	299	139	4%	18%	41%	27%	9%
WLCS	Mainland China	61	30	31	3%	30%	38%	20%	10%
YLCS	Mainland China	1,078	506	572	9%	24%	33%	23%	10%
		9,950	5,457	4,493	6%	15%	30%	33%	16%

[1]CAMSCH: Chinese Academy of Medical Sciences Cancer Hospital Study; CNULCS: Chonnam National University Lung Cancer Study; FLCS: Fudan Lung Cancer Study; GDS: Guangdong Study; GELAC: Genetic Epidemiological Study of Lung Adenocarcinoma (in Taiwan); GEL-S: Genes and Environment in Lung Cancer, Singapore study; HKS: Hong Kong Study; JLCS: Japanese Female Lung Cancer Collaborative Study; SKLCS: South Korea Multi-Center Lung Cancer Study (includes Seoul National University Study; Korean University Medical Center Study; and Kyungpook National University Hospital Study); SLCS: Shenyang Lung Cancer Study; SWHS: Shanghai Women's Health Study; TLCS: Tianjin Lung Cancer Study; WLCS: Wuhan Lung Cancer Study; YLCS: Yunnan Lung Cancer Study.

variants in an Asian population. Multiplex quantitative polymerase chain reactions were used to quantify telomere length. *T/S* values were extracted for the analysis and log transformed to improve normality.

All plotting and statistical analyses were performed on a 64-bit Windows build of R version 3.0.1 "Good Sport."[24] Only subjects with complete genotyping, histology and covariate information were included in the analysis (5,457 cases and 4,493 controls). Models investigating lung cancer risk were adjusted for study indicator variable, 10-year age group indicator variables (<40, 40–49, 50–59, 60–69 and 70+), and significant principal components (EV1, EV2 and EV4), unless otherwise noted. Likelihood-ratio and SNP-set kernel association test (SKAT) linear kernel tests[25,26] were used to assess statistical significance of aggregations of telomere-length associated variants on lung cancer risk by comparing null models to fitted models containing combinations of the 7 telomere-length associated variants. The SKAT linear kernel test aggregates a set of SNP score test statistics and efficiently computes an overall p value.[26]

Both unweighted and weighted genetic risk scores (GRS) were calculated for telomere-length associated variants. To calculate GRS for the ith subject from the seven telomere-length associated variants the following formula was used:

$$\mathrm{GRS}_i = \sum_{j=1}^{7} w_j x_{ij} \quad (1)$$

Here x_{ij} is the number of risk alleles for the jth SNP in the ith subject ($x_{ij} = 0$, 1 or 2) and w_j is the weight or coefficient for the jth SNP. Unweighted genetic risk scores simply counted the number of alleles associated with longer telomere length an individual carried across all seven telomere-length associated variants, thus giving an equal weight to all risk alleles ($w_j = 1$). Weighted genetic risk scores were calculated likewise, with the addition of assigning previously published telomere-length associated beta estimates[20] as w_j for each telomere-length associated SNP allele count. Weighting normally results in more specificity of the GRS by assigning more weight to variants with stronger effects.

Results

Our dataset consisted of a sample of 5,457 lung cancer cases and 4,493 controls from a population of never-smoking Asian females (Table 1). The participants were drawn from 14 contributing studies with collection areas in mainland China, South Korea, Japan, Singapore, Taiwan, and Hong Kong. Age, a major factor associated with telomere attrition, was available in 10-year age-groups for all participants. Most participants were between 50 and 70 years of age (63%) with 6% of subjects younger than 40 years of age.

Measured and imputed genotypes were available for the seven telomere-length associated variants (Table 2). Alleles associated with longer telomere length were denoted the risk allele and risk allele frequencies from our dataset were compared to those previously reported by Codd *et al.*[20] Risk allele frequency differences between our Asian lung cancer study and the Codd *et al.* study of a population of primarily

Table 2. Relationship between genetic risk alleles, previously reported measured telomere length in peripheral white blood cells from a large study in a European population, and risk of lung cancer among never-smoking females in Asia

					Genetic variants and telomere length in a European population (Codd *et al.*)[1]				Genetic variants and lung cancer association in a female non-smoking Asian population				
					RAF	Telomere length association			RAF	Lung cancer association[3]			
SNP	Chr	Position (GRCh37/hg19)	Nearby gene	Alleles Risk[2]	Nonrisk	Controls	β	95% CI	*p* value	Controls	Cases	OR	95% CI	*p* value
rs10936599	3	169,492,101	*TERC*	C	T	0.748	0.097	(0.081–0.113)	2.54×10^{-31}	0.431	0.451	1.08	(1.01–1.14)	0.014
rs2736100	5	1,286,516	*TERT*	C	A	0.486	0.078	(0.060–0.096)	4.38×10^{-19}	0.397	0.480	1.38	(1.30–1.47)	1.09×10^{-26}
rs7675998	4	164,007,820	*NAF1*	G	A	0.783	0.074	(0.056–0.092)	4.35×10^{-16}	0.826	0.833	1.05	(0.97–1.14)	0.202
rs9420907	10	105,676,465	*OBFC1*	C	A	0.135	0.069	(0.049–0.089)	6.90×10^{-11}	0.009	0.009	0.92	(0.67–1.26)	0.603
rs8105767	19	22,215,441	*ZNF208*	G	A	0.291	0.048	(0.032–0.064)	1.11×10^{-9}	0.318	0.326	1.04	(0.98–1.11)	0.191
rs755017	20	62,421,622	*RTEL1*	G	A	0.131	0.062	(0.040–0.084)	6.71×10^{-9}	0.447	0.443	0.98	(0.93–1.04)	0.536
rs11125529	2	54,475,866	*ACYP2*	A	C	0.142	0.056	(0.036–0.076)	4.48×10^{-8}	0.166	0.178	1.07	(1.00–1.16)	0.066

[1]Data extracted from Table 1 of the Codd *et al.*[20]
[2]Risk allele is the allele associated with longer telomere length.
[3]Association adjusted for age group, contributing study, and significant eigenvectors.
Abbreviations: Chr: chromosome; RAF: risk allele frequency.

European descent likely reflect differences in ancestral allele frequencies.

To ensure the telomere-length associated variants, discovered in a population of primarily European ancestry, were a valid surrogate for telomere length in our Asian population, we carried out an analysis on a set of 1,536 Asian females with both measured telomere length and genotype data from the prospective Shanghai Women's Health Study. When testing for an association of each of the seven telomere length associated variants with measured telomere length, only the *TERT* variant (rs2736100) had a significant association with measured telomere length (p value = 0.03); however, our sample size was substantially smaller than the Codd *et al.* analysis ($N = 48{,}423$), and although insignificant, six of the seven variants had beta estimates in the correct direction. A weighted GRS with all seven telomere-length associated variants was calculated and the association with telomere length was also investigated. In the overall sample, the telomere-length associated GRS was significantly associated with measured telomere length (p value = 0.001, Fig. 1*a*), the estimated effect was in the positive direction (beta = 0.15), and explained the same percent of total telomere length variance as in Codd *et al.* ($R^2 = 0.01$).[20] For the cancer cases in this sample, the mean time between blood sample collection and cancer diagnosis was 5.34 years with 75% of cases having blood collected >3 years prior to cancer diagnosis. When restricting the analysis to controls ($N = 533$), the association remained significant (p value = 0.04) with similar effect size and variance explained (Fig. 1*b*). Together, this provides evidence the weighted GRS of telomere-length associated variants has utility in predicting measured telomere length in Asian populations.

Overall association tests were conducted to investigate if, in aggregate, all seven telomere-length associated variants were associated with lung cancer risk. A likelihood ratio test comparing a null model adjusting for 10-year age group, contributing study, and significant principal components to the same model plus all seven telomere-length associated variants indicated that in aggregate the telomere-length associated variants were significantly associated with lung cancer risk (p value = 9.64×10^{-25}). Furthermore, a linear SKAT found a highly significant association between the seven telomere-length associated variants and lung cancer (p value = 3.19×10^{-27}).

Each telomere-length associated variant from Codd *et al.*[20] was tested for an individual association with lung cancer risk. All seven telomere-length associated variants were included in the same logistic regression model and covariates were included to adjust for 10-year age-group, contributing study, and significant principal components. Two of the seven telomere-length associated variants (rs2736100 and rs10936599) exhibited association p values <0.05, significantly <0.4 variants expected by chance (p value = 0.04) (Table 2). The rs2736100 variant, located in the first intron of the *TERT* gene, has previously been associated by GWAS

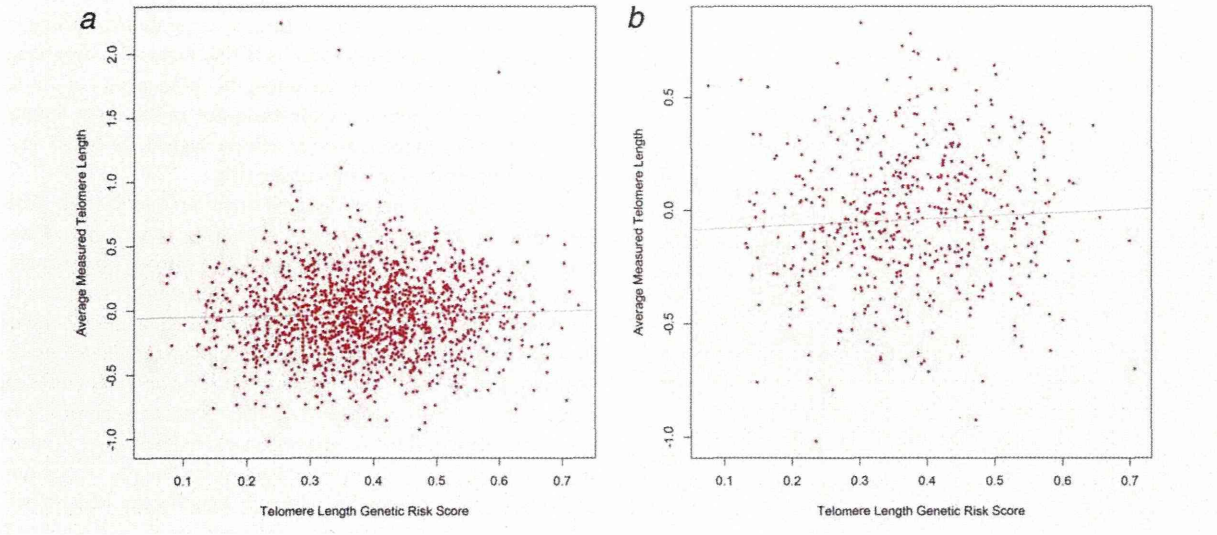

Figure 1. Relation of telomere-length associated variants with measured telomere length in peripheral white blood cell DNA from 1,536 women included in previous nested case-control studies of various cancers in the Shanghai Women's Health Study. A best-fit line (solid gray line) is drawn for the relationship of measured log-transformed telomere length with telomere-length associated weighted genetic risk score for (*a*) cancer cases and controls ($R^2 = 0.01$, p value = 0.001) and (*b*) controls ($N = 533$) only ($R^2 = 0.01$, p value = 0.04). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

with lung cancer risk.[27] Interestingly, five of the seven telomere-length associated variants show effects in the same direction for both the Codd *et al.* telomere-length association[20] and lung cancer association suggesting enrichment for variants that are associated with both longer telomere length and increased lung cancer risk (Table 2).

Both unweighted and weighted GRSs were calculated as measures of predicted telomere length for each study participant and association with lung cancer risk was tested by logistic regression models that adjusted for 10-year age group, contributing study, and significant principal components. The unweighted telomere-length associated GRS was significantly associated with lung cancer risk (p value = 1.90×10^{-12}), indicating scores associated with longer telomere length were also associated with increased lung cancer risk. The odds ratio comparing individuals in the upper quartile of GRS to those in the lower quartile of GRS was 1.47 (95% CI = 1.31–1.65). The beta weighted telomere-length associated GRS demonstrated greater specificity for the lung cancer association with greater evidence for association between longer telomere length and lung cancer risk (p value = 4.54×10^{-14}). A higher odds ratio of 1.51 (95% CI = 1.34–1.69) was observed for individuals in the upper quartile of the weighted GRS compared to those in the lower quartile. The association of the weighted GRS across contributing study was homogeneous (homogeneity p value = 0.34) and produced an overall meta-analysis odds ratio of 1.51 (95% CI = 1.34–1.71, p value = 1.53×10^{-11}) comparing individuals in the upper quartile of weighted GRS to those in the lower quartile of GRS (Fig. 2). When investigating deciles of the weighted GRS, the effect of weighted GRS on lung cancer risk appeared to be

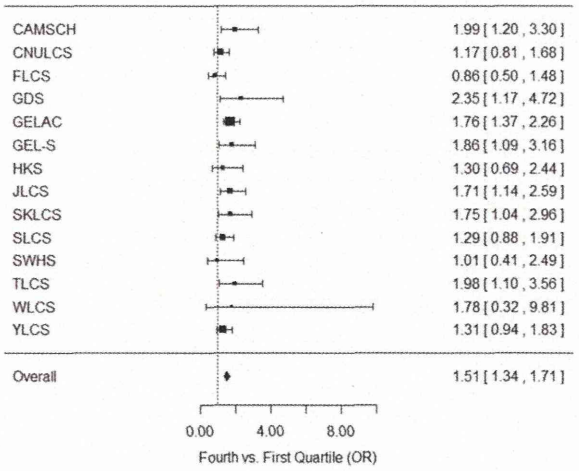

Figure 2. Adjusted odds ratios for risk of lung cancer among never-smoking females in Asia comparing upper quartile to lower quartile of weighted telomere length genetic risk scores, by study. Lung cancer risk was positively associated with increasing weighted GRS (p value = 1.53×10^{-11}) with no significant evidence for heterogeneity of effect (p value = 0.34).

monotonic with no threshold indicating a substantial change in risk (Fig. 3). Furthermore, to assess if rs2736100 was the only SNP accounting for the association between the weighted GRS and lung cancer risk, the weighted GRS was recomputed with the exclusion of rs2736100, and rs2736100 used as a separate covariate in the regression model. The weighted GRS minus rs2736100 remained significantly associated with increased lung cancer risk, although the p value was greatly attenuated (p value = 4.81×10^{-3}).