<table>
<thead>
<tr>
<th>Title</th>
<th>Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Kerin, Michael J.; Miller, Nicola</td>
</tr>
<tr>
<td>Publication Date</td>
<td>2013-04</td>
</tr>
<tr>
<td>Publisher</td>
<td>Nature Publishing Group (Macmillan)</td>
</tr>
<tr>
<td>Link to publisher's version</td>
<td><a href="http://dx.doi.org/10.1038/ng.2566">http://dx.doi.org/10.1038/ng.2566</a></td>
</tr>
<tr>
<td>Item record</td>
<td><a href="http://hdl.handle.net/10379/3883">http://hdl.handle.net/10379/3883</a></td>
</tr>
<tr>
<td>DOI</td>
<td><a href="http://dx.doi.org/10.1038/ng.2566">http://dx.doi.org/10.1038/ng.2566</a></td>
</tr>
</tbody>
</table>

Some rights reserved. For more information, please see the item record link above.
Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer

TERT-locus SNPs and leukocyte telomere measures are reportedly associated with risks of multiple cancers. Using the Illumina custom genotyping array iCOGs, we analyzed ~480 SNPs at the TERT locus in breast (n = 103,991), ovarian (n = 39,774) and BRCA1 mutation carrier (n = 11,705) cancer cases and controls. Leukocyte telomere measurements were also available for 53,724 participants. Most associations cluster into three independent peaks. The minor allele at the peak 1 SNP rs2736108 associates with longer telomeres (P = 5.8 × 10⁻⁷), lower risks for estrogen receptor (ER)-negative (P = 1.0 × 10⁻⁸) and BRCA1 mutation carrier (P = 1.1 × 10⁻⁵) breast cancers and altered promoter assay signal. The minor allele at the peak 2 SNP rs4705526 associates with longer telomeres (P = 2.3 × 10⁻¹⁴), higher risk of low-malignant-potential ovarian cancer (P = 1.3 × 10⁻¹⁵) and greater promoter activity. The minor alleles at the peak 3 SNPs rs10069690 and rs2242652 increase ER-negative (P = 1.2 × 10⁻¹²) and BRCA1 mutation carrier (P = 1.6 × 10⁻¹⁴) breast and invasive ovarian (P = 1.3 × 10⁻¹¹) cancer risks but not via altered telomere length. The cancer risk alleles of rs2242652 and rs10069690, respectively, increase silencing and generate a truncated TERT splice variant.

Chromosome ends are capped by telomeres, which protect them from inappropriate DNA repair and maintain genomic integrity. Telomeres consist of structural proteins combined with many hundreds of hexanucleotide DNA repeats, which are progressively shortened by normal cell division. Shortening restricts the proliferation of normal somatic cells but not cancer cells, which can maintain long telomeres, usually via telomerase and may divide indefinitely. The TERT gene at 5p15.33 (NCBI gene 7015) encodes the catalytic subunit of telomerase reverse transcriptase, a key component of telomerase. Germline mutations in TERT cause dyskeratosis congenita, a cancer susceptibility disorder characterized by exceedingly short telomeres. Although up to 80% of the variation of telomere length is estimated to be due to heritable factors, association studies of TERT SNPs and differences in leukocyte telomere length have so far been inconclusive. Furthermore, it is unclear whether telomere length, measured in leukocyte DNA, is predictive of cancer risk: retrospective studies report that cancer patients after diagnosis have shorter telomeres than unaffected controls, but prospective studies with DNA taken before diagnosis have been inconclusive. SNPs at 5p15.33 are reported to be associated with risks of several human cancers, including certain subtypes of both ovarian and breast cancers.

Resulting from a common interest, members of each of the constituent consortia in the Collaborative Oncological Gene-environment Study (COGS) nominated SNPs surrounding the TERT locus for inclusion on a genotyping array. Consequently, the iCOGS array design included a combination of individual TERT gene candidate SNPs, as well as a more comprehensive set to fine-scale map the entire locus, for shared use by all consortia. This study had three aims: to assess SNPs across the TERT locus for all detectable associations with mean telomere length and breast and ovarian cancer subtypes; to fine-scale map this locus to identify potentially causal variants for the observed associations; and to evaluate the functional effects of the strongest candidate causative variants.

RESULTS

One hundred and ten SNPs at the 5p15.33 locus (Build 37 positions 1,227,693–1,361,969) passed quality control tests in 103,991 breast cancer cases and controls from 52 Breast Cancer Association Consortium (BCAC) studies, of which 41 studies (89,050 individuals) were of European ancestry, 9 were of Asian ancestry (12,893 individuals) and 2 were of African-American ancestry (2,048 individuals). The same 110 SNPs passed quality control tests in 11,705 BRCA1 mutation carriers of European ancestry, recruited by 51 studies from the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA), and 108 SNPs passed quality control tests in 44,308 ovarian cancer cases and controls from 43 Ovarian Cancer Association Consortium (OCAC) studies. For OCAC, analysis was confined to the 39,774 participants of European ancestry, of whom 8,371 cases had invasive epithelial ovarian neoplasia and 986 had serous low-malignant-potential (LMP) neoplasia. For all study participants, genotype imputation, using the 110 genotyped SNPs together with the January 2012 release of the 1000 Genomes Project, was used to increase coverage to ~480 SNPs (imputation r² > 0.3; minor allele frequency (MAF) > 0.02) for each phenotype. Telomere length was initially measured in control subjects from two BCAC studies (Studies of Epidemiology And Risk factors in Cancer Heredity (SEARCH) and the Copenhagen City Heart Study (CCHS); combined n = 15,567) (Supplementary Note).
Manhattan plots are shown of the genotyped and well-imputed SNPs for the seven phenotypes analyzed, including mean telomere length (Fig. 1a), overall breast cancer risk (Fig. 1b), breast cancer in BRCA1 mutation carriers (Fig. 1c), ER-negative breast cancer (Fig. 1d), ER-positive breast cancer (Fig. 1e), serous LMP ovarian cancer (Fig. 1f) and serous invasive ovarian cancer (Fig. 1g). Conditional analyses within each of these phenotypes identified multiple independent SNP associations each for telomere length, overall breast cancer risk, ER-negative breast cancer and breast cancer in BRCA1 mutation carriers but only one peak each for ER-positive breast cancer, serous LMP ovarian cancer and invasive ovarian cancer (Table 1). Full results of all these SNP analyses are given in Supplementary Tables 1–3. All associations are consistent with a log-additive model.

Associations with telomere length
SNPs in two distinct regions (hereafter denoted peaks 1 and 2) were strongly associated with telomere length (Fig. 1a, Tables 1 and 2 and Supplementary Fig. 1a). Imputed SNP rs7705526 (peak 2, position 1,285,974, TERT intron 2) had the largest effect, with a change in relative telomere length of 1.026-fold per allele (95% confidence interval (CI) = 1.019–1.033; \(P = 4.0 \times 10^{-4}\)) evidence for heterogeneity among odds ratios (ORs) or hazard ratios (HRs) between studies for any of the top SNPs was observed (Supplementary Fig. 2). The strongest association with overall breast cancer risk in BCAC was with peak 1 SNP rs2242652 (Fig. 1a, Tables 1 and 2 and Supplementary Fig. 1b). There was also good evidence for an association with SNPs in peak 2 and weaker evidence that an additional SNP, outside the three main association peaks, was independently associated with breast cancer risk (Table 1 and Supplementary Table 1). The most strongly associated SNPs in BRCA1 mutation carriers were located in introns 2–4 (hereafter denoted peak 3), including rs10069690 (Fig. 1c, Tables 1 and 2 and Supplementary Fig. 2c) and rs2242652 (correlation with rs10069690, \(r^2 = 0.70\)). The latter SNP also showed the strongest association with breast cancer risk (Fig. 1d, Tables 1 and 2 and Supplementary Fig. 1d) but showed little evidence of association with ER-positive breast cancer (Table 2). Stepwise regression analysis in CIMBA studies indicated two independent associations with breast cancer risk in BRCA1 mutation carriers (conditional \(P = 5 \times 10^{-5}\) for rs2736108 in peak 1 and \(P = 4.8 \times 10^{-13}\) for rs10069690 in peak 3).
A very similar pattern was observed for ER-negative breast cancer in BCAC (conditional \( P = 6 \times 10^{-6} \) for rs3215401 in peak 1 and \( P = 4.3 \times 10^{-9} \) for rs2242652 in peak 3; Table 1).

The most strongly associated SNP with ER-positive breast cancer was rs2736107 in peak 1 (Fig. 1e, Tables 1 and 2 and Supplementary Fig. 3c). Weak associations between the key SNPs and risk for BRCA2 mutation carriers were also observed, but the sample size was too small to draw definitive conclusions (data not shown).

### Associations with ovarian cancer risk

The strongest association observed for risk of LMP ovarian cancer was with peak 2 SNP rs7705526, and this was the only SNP retained in the stepwise regression analysis (Fig. 1f, Tables 1 and 2 and Supplementary Fig. 1f). The strongest observed association for serous invasive ovarian cancer was with peak 3 SNP rs10069690 (Fig. 1g, Tables 1 and 2 and Supplementary Fig. 1g). No other independent association was observed for serous invasive ovarian cancer (Table 1). We also analyzed SNP associations with endometrioid, mucinous, clear-cell invasive and mucinous LMP ovarian cancers but found no associations at \( P < 1 \times 10^{-4} \) (Supplementary Table 4). We attempted analysis of invasive serous ovarian cancer stratified by grade, but, again, statistical power was low (Supplementary Fig. 3).

### Three main peaks of association within the TERT locus

The above results indicate that the majority of observed associations with all seven tested phenotypes fall into association peaks 1–3.

Correlated SNPs in the TERT promoter (peak 1) were associated with telomere length, ER-positive breast cancer, ER-negative breast cancer and breast cancer in BRCA1 mutation carriers. SNPs in peak 2, spanning TERT introns 2–4, were independently associated with telomere length, overall breast cancer risk and serous LMP ovarian cancer. SNPs in peak 3, also spanning TERT introns 2–4, showed strong associations with ER-negative breast cancer, breast cancer risk for BRCA1 mutation carriers and serous invasive ovarian cancer but not with telomere length (Tables 1 and 2). Although peaks 2 and 3 overlap physically, they define distinct sets of SNPs that are only partially correlated (for example, correlation between rs10069690 and rs7705526 was weak, \( r^2 = 0.33 \); Fig. 2). Some SNP-phenotype associations in peak 2 were clearly weaker than those in peak 3 (for example, with ER-negative breast cancer) and became nonsignificant after adjustment for SNP rs2242652 in peak 3. Conversely, the associations with telomere length and serous LMP ovarian cancer were stronger for SNPs in peak 2, indicating that the associations in peaks 2 and 3 are not being driven by the same causal variants.

The strongest candidates for causation within each peak were identified by computing likelihood ratios; the SNPs listed in Tables 1 and 2 are those that cannot be excluded from being causal candidates at a likelihood ratio of >1:100 fold compared to the top hit in the peak. The statistical power to exclude SNPs differed between phenotypes: in peak 1, all but seven SNPs could be excluded from being causal for relative telomere length, breast cancer risk in BRCA1 mutation carriers and ER-negative breast cancer risk, but an additional SNP could be excluded for

---

**Table 1** Independently associated SNPs for each phenotype

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr. position</th>
<th>TERT peak</th>
<th>Source</th>
<th>Effect (95% CI)</th>
<th>( P_{\text{stepwise}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2736108</td>
<td>1,297,488</td>
<td>1</td>
<td>Genotyped</td>
<td>1.010 (1.004–1.016)</td>
<td>0.0004</td>
</tr>
<tr>
<td>rs7705526</td>
<td>1,285,974</td>
<td>2</td>
<td>Imputed</td>
<td>1.019 (1.014–1.025)</td>
<td>2.47 \times 10^{-11}</td>
</tr>
<tr>
<td>rs3215401</td>
<td>1,296,255</td>
<td>1</td>
<td>Imputed</td>
<td>0.94 (0.91–0.96)</td>
<td>9.91 \times 10^{-10}</td>
</tr>
<tr>
<td>rs7734992</td>
<td>1,280,128</td>
<td>2</td>
<td>Imputed</td>
<td>1.06 (1.04–1.08)</td>
<td>1.73 \times 10^{-7}</td>
</tr>
<tr>
<td>rs56963355</td>
<td>1,251,503</td>
<td>None</td>
<td>Imputed</td>
<td>0.90 (0.84–0.95)</td>
<td>1.95 \times 10^{-5}</td>
</tr>
</tbody>
</table>

Independent associated SNPs are shown for each phenotype, including overall breast cancer and ER subgroups in European individuals in BCAC and invasive and LMP subgroups in OCAC following forward conditional stepwise logistic regression analysis, relative change in telomere length in SEARCH and CCHS combined data following forward stepwise linear regression analysis and breast cancer in BRCA1 mutation carriers in CIMBA following forward stepwise Cox regression. These analyses were performed on all SNPs with MAF > 0.02 and association \( P < 1 \times 10^{-4} \) in the single-SNP analyses.
Table 2 Association between TERT SNPs and the seven studied phenotypes

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr. 5 position</th>
<th>Major/minor allele</th>
<th>Telomere length change</th>
<th>Overall breast cancer risk</th>
<th>Risk of breast cancer in BRCA2 mutation carriers</th>
<th>Risk of breast cancer in ER-negative breast cancer carriers</th>
<th>Risk of breast cancer in ER-positive breast cancer carriers</th>
<th>Risk of ovarian cancer carriers</th>
<th>Risk of ovarian cancer carriers</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2736107</td>
<td>1,297,854</td>
<td>C/T</td>
<td>8.33 × 10^{-7}</td>
<td>0.94 (0.91–0.96)</td>
<td>0.75  (0.65–0.85)</td>
<td>0.88 (0.83–0.93)</td>
<td>1.41 × 10^{-8} (0.92–0.97)</td>
<td>0.95 (0.91–0.98)</td>
<td>0.98 (0.91–1.02)</td>
</tr>
<tr>
<td>rs2736108a</td>
<td>1,297,488</td>
<td>C/T</td>
<td>5.81 × 10^{-7}</td>
<td>0.94 (0.92–0.95)</td>
<td>0.63  (0.54–0.73)</td>
<td>0.89 (0.84–0.94)</td>
<td>1.05 × 10^{-3} (0.93–0.89)</td>
<td>0.96 (0.89–0.98)</td>
<td>0.98 (0.89–1.02)</td>
</tr>
<tr>
<td>rs7255896</td>
<td>1,297,081</td>
<td>CCAC</td>
<td>7.13 × 10^{-4}</td>
<td>0.94 (0.92–0.96)</td>
<td>1.20  (0.84–1.93)</td>
<td>0.88 (0.84–0.93)</td>
<td>1.04 × 10^{-5} (0.89–0.93)</td>
<td>0.96 (0.91–0.98)</td>
<td>0.98 (0.91–1.02)</td>
</tr>
<tr>
<td>rs1297077</td>
<td>1,297,077</td>
<td>ACCA</td>
<td>7.13 × 10^{-4}</td>
<td>0.94 (0.92–0.96)</td>
<td>1.31  (0.84–1.94)</td>
<td>0.88 (0.84–0.94)</td>
<td>1.05 × 10^{-6} (0.89–0.93)</td>
<td>0.96 (0.91–0.98)</td>
<td>0.98 (0.91–1.02)</td>
</tr>
<tr>
<td>rs3215401</td>
<td>1,296,255</td>
<td>AAG</td>
<td>1.63 × 10^{-6}</td>
<td>0.90 (0.91–0.96)</td>
<td>1.05  (0.84–0.94)</td>
<td>0.88 (0.84–0.92)</td>
<td>4.68 × 10^{-10} (0.91–0.93)</td>
<td>0.95 (0.91–0.98)</td>
<td>0.98 (0.91–1.01)</td>
</tr>
<tr>
<td>rs2853669</td>
<td>1,294,349</td>
<td>AGA</td>
<td>2.95 × 10^{-6}</td>
<td>0.91 (0.92–0.96)</td>
<td>1.00  (0.84–0.94)</td>
<td>0.88 (0.84–0.92)</td>
<td>4.01 × 10^{-9} (0.91–0.93)</td>
<td>0.95 (0.91–0.98)</td>
<td>0.98 (0.91–1.01)</td>
</tr>
<tr>
<td>rs2730989</td>
<td>1,294,086</td>
<td>C/T</td>
<td>6.65 × 10^{-10}</td>
<td>0.91 (0.91–0.96)</td>
<td>1.17  (0.84–0.94)</td>
<td>0.89 (0.84–0.93)</td>
<td>4.22 × 10^{-5} (0.92–0.97)</td>
<td>0.95 (0.91–0.98)</td>
<td>0.97 (0.91–1.01)</td>
</tr>
</tbody>
</table>

Peak 1: promoter

rs2736107 1,297,854 C/T 0.28 1.015 8.33 × 10^{-7} 0.93 7.95 × 10^{-10} 0.90 0.00013 0.88 1.41 × 10^{-8} 0.95 3.32 × 10^{-6} 1.01 0.909 0.98 0.301
rs2736108a 1,297,488 C/T 0.29 1.017 5.81 × 10^{-7} 0.94 6.73 × 10^{-9} 0.89 1.05 × 10^{-5} 0.89 1.01 × 10^{-8} 0.96 0.00051 1.00 0.996 0.98 0.238
rs7255896 1,297,081 CCAC 0.27 1.016 7.13 × 10^{-4} 0.94 1.20 × 10^{-8} 0.88 1.04 × 10^{-5} 0.88 1.21 × 10^{-8} 0.96 0.00083 1.02 0.762 0.98 0.324
rs1297077 1,297,077 ACCA 0.27 1.016 7.13 × 10^{-4} 0.94 1.91 × 10^{-8} 0.89 1.90 × 10^{-6} 0.88 1.37 × 10^{-8} 0.96 0.00051 1.02 0.769 0.98 0.351
rs3215401 1,296,255 AAG 0.30 1.016 1.63 × 10^{-6} 0.94 4.68 × 10^{-10} 0.90 3.05 × 10^{-5} 0.88 1.19 × 10^{-9} 0.95 0.0012 1.01 0.908 0.98 0.188
rs2853669 1,294,349 AGA 0.30 1.016 2.95 × 10^{-6} 0.94 4.01 × 10^{-9} 0.90 4.26 × 10^{-5} 0.88 2.60 × 10^{-9} 0.96 0.00034 1.01 0.893 0.97 0.204
rs2730989 1,294,086 C/T 0.27 1.017 1.35 × 10^{-7} 0.93 6.65 × 10^{-10} 0.89 4.22 × 10^{-5} 0.88 2.41 × 10^{-8} 0.95 4.58 × 10^{-5} 1.00 0.935 0.97 0.152

Peak 2: introns 2-4

Single-SNP estimates for the most significant SNPs per peak per phenotype in the TERT fine-mapping region of chromosome 5 from positions 1,227,014 to 1,361,964 (Genome Build 37). The major and minor alleles are from the genotyping assay and are not necessarily from the coding strand. Change in telomere length per copy of an increase/decrease (for BRCA) or increase/decrease (for breast cancer risk) is given for each SNPs tested. Independent peaks 1–3 were localized using forward conditional analyses (Table 1). The most significant hit in each region for each phenotype is shown in bold for each peak.

*Genotyped SNPs. All three peaks contain at least one directly genotyped variant. †Results for the most significant SNPs in each block in the forward conditional analyses (described further in Table 1).
ER-positive breast cancer risk (Table 2). In peak 2, the greatest power was for the telomere length phenotype, where all but three SNPs could be excluded, whereas five or six remained for cancer risk. For peak 3, three putative causal SNPs remained for ER-negative breast cancer risk, two for serous invasive ovarian cancer risk and just one for breast cancer risk in BRCA1 mutation carriers. Results in each peak are compatible with a single causative variant being responsible for the multiple phenotype associations (notably, in peak 3, SNPs rs2242652 and rs10069690 were equally compatible with being the single causal variant). However, the possibilities of different causal variants being responsible for different phenotypes or of the associations being due to haplotype effects cannot be ruled out.

**Asian and African-American studies**

We tested all SNPs (n = 341) with MAF > 0.02 and imputation r² > 0.3 for association with breast cancer in the 9 BCAC Asian studies (comprising 6,269 cases and 6,624 controls) for association, but none reached formal levels of significance. Furthermore, none of the top SNPs in individuals of European ancestry showed more than borderline levels of significance in Asians (Supplementary Table 5). Peak 3 SNP rs10069690 was directly genotyped in 2 BCAC African-American studies (1,116 cases and 932 controls), as well as in the above-mentioned Asian studies, and had estimated effects on ER-negative breast cancer similar to those in European populations (per-allele OR = 1.19, 95% CI = 1.06–1.31, P = 0.009 in African-Americans and OR = 1.09, 95% CI = 1.00–1.19, P = 0.07 in Asian women). Within OCAC, there were too few women of Asian and African ancestry to draw meaningful conclusions (Supplementary Table 6).

**Chromatin analysis**

Analysis of Encyclopedia of DNA Elements (ENCODE) data showed no evidence of regulatory elements or open chromatin coinciding with any risk-associated SNPs in normal breast epithelial cells or the other represented tissues (Supplementary Fig. 4). Data for ovarian tissues are not included in ENCODE. We therefore performed site-specific formaldehyde-assisted isolation of regulatory elements (FAIRE) in ovarian cancer precursor tissues to identify regulatory elements in a 1 Mb region centered on peak 3. In fallopian tube secretory and ovarian surface epithelial cells, we detected FAIRE peaks coinciding with the CLPTM1L promoter but not the TERT promoter (Supplementary Fig. 4). In silico analyses additionally indicated that TERT introns 4 and 5 (within and beyond peak 3) contained regions showing regulatory potential and vertebrate conservation. We performed site-specific FAIRE analyses on a ~1 kb region centered on the peak 3 SNP rs10069690 in normal tissue samples from breast reduction mammoplasty (n = 4), ovarian cancer precursor tissues (n = 4) and ovarian cancer cell lines (n = 4). Breast cells from each woman were sorted into four enriched fractions on the basis of differential expression of cell surface markers (myoepithelial/stem, luminal progenitor, mature luminal and stromal cells), and assays were performed on each fraction (Fig. 3). Chromatin was in a closed configuration in all ovarian, breast luminal progenitor and mature luminal cell fractions. However, in two of four stromal cell fractions, we detected ~600 bp of open chromatin of varying amplitude, covering the position of SNP rs10069690 but not of rs2242652, and in three of four myoepithelial/stem cell fractions, we detected ~800 bp of open chromatin, covering the positions of both SNPs rs10069690 and rs2242652.

**Luciferase reporter assays**

The regulatory capabilities of the DNA in each of the three peaks and the effects of most of the strongest candidate causative variants in each one were examined in luciferase reporter assays, using a construct containing 3,915 bp of the TERT promoter sequence. The effects of peak 1 TERT promoter variants were examined via five haplotype constructs differing at rs2736107, rs2736108 and rs2736109 (ref. 25) (Fig. 4a): one with all three major alleles (wild-type TERT), another with all three minor alleles and three each with a single minor allele of the SNPs. Relative promoter activity was determined in ER-positive (MCF7) and ER-negative (MDA-MB-468) breast cancer cell lines and in an ovarian cancer cell line (A2780). The construct containing all three minor alleles consistently generated the lowest luciferase signals, close to baseline. To determine whether the risk-associated variants in peaks 2 and 3 fell within putative cis-acting regulatory elements (PREs), we cloned ~3 kb of sequence surrounding each SNP. Constructs of PRE-A (peak 2) had no significant effect on the activity of either the wild-type (TERTwt) promoter or the promoter with three minor alleles (TERT) (Fig. 4a). However, inclusion of the minor
alleles of rs7705526 resulted in ~30% higher TERT promoter activity in all three cell lines, suggesting that it can act as a transcriptional enhancer. Higher promoter activity was also observed with this construct in A2780 ovarian cancer cells but not in the two breast cancer cell lines. Constructs of PRE-B (peak 3) consistently acted as strong transcriptional silencers, leading to 40–50% lower activity, specifically in constructs containing the wild-type TERT promoter. Notably, inclusion of the minor allele of rs2242652 in PRE-B constructs decreased relative wild-type TERT promoter activity by a further ~20% compared to the silencer containing the major allele, but the minor allele of the highly correlated SNP rs10069690 did not generate this effect (Fig. 4a).

**Alternative splicing of TERT**

Several alternatively spliced variants of TERT have been found to affect telomerase activity.\(^{44,45}\) To determine the role of PRE-B (peak 3) SNPs in TERT alternative splicing, we inserted intron 4 sequence into a full-length TERT cDNA mini-gene construct and confirmed accurate splicing. Cancer risk–associated alleles for rs10069690 and rs2242652 were generated individually and in combination within the mini-gene. RT-PCR, using primers spanning intron 4, showed that all SNP permutations in all cell lines produced comparable levels of both wild-type transcript and an INS1 alternatively spliced variant, which is predicted to generate a severely truncated protein product, which is likely to affect telomerase activity (Supplementary Fig. 5b).

**Gene expression and methylation analyses in ovarian tissue**

We used The Cancer Genome Atlas (TCGA)\(^{46}\) data to examine gene expression of the 11 protein-coding genes and 1 microRNA (MIR4457) located within 1 Mb of peak 3 SNP rs10069690. Most genes showed higher expression in ovarian tumors compared with normal tissues (Supplementary Fig. 4 and Supplementary Table 7). We observed no association between rs10069690 and the expression levels of any of the genes in any of the cells tested (Supplementary Fig. 5 and Supplementary Tables 7 and 8). There was some evidence of association between rs10069690 and tumor methylation with probes cg23827991 (TERT CpG island, \(P = 1.3 \times 10^{-6}\)) and cg06550200 (CLPTM1L, \(P = 6.9 \times 10^{-4}\)) out of the 935 probes tested. Both regions showed lower methylation with the minor, cancer risk–associated allele (Supplementary Table 9), but this did not correlate with changes in expression.

**DISCUSSION**

Our comprehensive examination of the TERT locus has answered some long-standing questions and raised several new ones. We have identified two independent regions associated with telomere length in leukocyte DNA; these provide definitive evidence for genetic control of telomere length by common TERT variants. For rs2736108, the most significant SNP in promoter peak 1, the minor allele is associated with a 1.7% increase in telomere length. This is equal to a telomere length change of ~60 bp, which, because telomere length decreases by approximately 19 bp per year,\(^{46}\) is equivalent in magnitude to an age difference of 3.1 years. We estimate that rs2736108 explains 0.08% of the variance in telomere length in men and 0.06% of the variance in women. SNPs in peak 2 have a stronger effect on telomere length, with each additional A (minor) allele of rs7705526 associated with a 2.6% increase. This is equal to a ~90 bp change in telomere length and, correspondingly, to 4.7 years of age. We estimate that rs7705526 explains 0.31% of the variance in telomere length in men and 0.16% of the variance in women. The only other reported associations with telomere length reaching genome-wide significance involve TERC-locus SNP rs1269304 (ref. 51) and OBF1-locus SNP rs4387287 (ref. 52), which have similar effects on telomere length (75 bp and 115 bp per allele, respectively).

Our only findings consistent with the hypothesis that shorter telomeres predispose to increased cancer risk\(^{23}\) (equivalent to longer telomeres being protective) are those from the peak 1 SNPs. However, a regulatory element construct containing the longer telomere–associated alleles of three highly correlated SNPs, rs2736108, rs2736107 and rs2736109 (reconstructing a haplotype with 25% frequency in Europeans\(^{35}\)), virtually abolished promoter activity in a reporter assay. This finding leaves a seemingly paradoxical association.
between lower enhancer activity and greater telomere length (Fig. 4). Control of telomerase activity is currently poorly understood, and this finding clearly merits further investigation.

SNPs within peak 3 (TERT introns 2–4) show strong associations with hormone-related cancers: peak 3 SNP rs10069690 is associated with risk of ER-negative breast cancer (MCF7) and ovarian cancer (A2780) cell lines. Error bars represent standard error from at least three independent experiments. (a) Luciferase reporter assays after transient transfection of cells with pGL2-Control (SV40 promoter and enhancer), pGL2-Basic (no promoter or enhancer) or the pGL2-Basic (3.9 kb of TERT promoter), TERT promoter with the minor (T) allele of rs2736107, rs2736108 or rs2736109 or TERT promoter with the T allele at all sites. The results of comparisons with wild-type TERT performed using one-way ANOVA with post-hoc Dunnett’s tests are shown (**P < 0.001, *P < 0.005). (b) PRE-A or PRE-B was cloned downstream of either the TERTwt or TERTth promoter-driven reporters with and without the minor alleles of SNPs rs10069690, rs2242652 and rs7705526, respectively. The results of comparisons with wild-type TERTwt or TERTth performed using one-way ANOVA with post-hoc Dunnett’s tests are shown (***P < 0.0001, **P < 0.001). WT, wild type.

Figure 4 TERT promoter and PRE activity. Luciferase reporter assays following transient transfection of ER-negative breast cancer (MCF7), ER-positive breast cancer (MCF7) and ovarian cancer (A2780) cell lines. Error bars represent standard error from at least three independent experiments. (a) Luciferase reporter assays after transient transfection of cells with pGL2-Control (SV40 promoter and enhancer), pGL2-Basic (no promoter or enhancer) or the TERT promoter vectors with TERTwt (3.9 kb of TERT promoter), TERT promoter with the minor (T) allele of rs2736107, rs2736108 or rs2736109 or TERT promoter with the T allele at all sites. The results of comparisons with wild-type TERT performed using one-way ANOVA with post-hoc Dunnett’s tests are shown (***P < 0.001, **P < 0.005). (b) PRE-A or PRE-B was cloned downstream of either the TERTwt or TERTth promoter-driven reporters with and without the minor alleles of SNPs rs10069690, rs2242652 and rs7705526, respectively. The results of comparisons with wild-type TERTwt or TERTth performed using one-way ANOVA with post-hoc Dunnett’s tests are shown (***P < 0.0001, **P < 0.001). WT, wild type.

SNPs within peak 3 (TERT introns 2–4) show strong associations with hormone-related cancers: peak 3 SNP rs10069690 is associated with risk of ER-negative breast cancer and breast cancer in BRCA1 mutation carriers, consistent with the observation that the majority of breast cancers arising in BRCA1 mutation carriers are ER negative. This variant has been reported to be associated with prostate risk of melanoma. Although SNPs in peaks 2 and 3 overlap on a physical map, the SNPs most strongly associated with cancer risk or telomere length were not highly correlated with each other (ρ2 between rs10069690 and rs7705526 = 0.33; Fig. 2b). This observation suggests that either the associations observed with multiple cancers and SNPs in peak 3 are mediated via a mechanism distinct from control of telomere length or that telomere length in breast, prostate and ovarian cells is under the control of a different set of SNPs from those controlling telomere length in leukocytes. Luciferase reporter assays show that peak 3 contains a silencer of the TERT promoter and that the minor allele of peak 3 SNP rs2242652 further reduces expression. Consistent with this finding, Kote-Jarai et al.54 report that the minor allele of this SNP is associated with reduced TERT expression in benign prostate tissue. However, we were unable to identify comparable associations in ovarian or breast tumor tissue, possibly because TERT expression is severely dysregulated in most tumors. Taken together, our luciferase assays indicate that either reduced signal from regulatory elements in peaks 1 and 3 or increased signal from peak 2 increases risk of specific cancer types.
codon (Supplementary Fig. 6). We do not yet know whether this isoform affects canonical telomerase activity or how it changes activity. We further identified new open chromatin signatures overlapping rs10069690 in breast stromal and myoepithelial/stem cell fractions but not in progenitor or differentiated luminal epithelial cell fractions. Senescent stromal cells can stimulate malignant transformation of epithelial cells in *in vitro* and *in vivo* models35,56, and the mechanisms mediated by these SNPs merit investigation in future studies.

The SNPs originally reported to be associated with risk of lung (rs402710)57 and breast (rs3816659)58 cancers (Supplementary Table 10) were not associated with any cancer in this study. Moreover, SNP rs2736100 in peak 2 has been reported to be associated with glioma and lung and testicular cancers27,28,31,57,59–62, whereas nearby SNP rs2853677 was reported to be associated with glioma in the Han Chinese population63. Despite their physical proximity, these SNPs are not highly correlated with rs7705526 ($r^2 = 0.52$ and 0.18, respectively), nor do they show independent associations with telomere length after adjustment for rs7705526. Thus, variants underlying susceptibility to different cancer types are different from the set of variants in the *TERT* locus mediating changes in telomere length.

One limitation of this study is the incomplete representation of all SNPs at 5p15.33 on the iCOGS chip, which was designed in March 2010 using SNPs catalogued in HapMap 3 together with those from the pilot study of the 1000 Genomes Project53. To help fill known gaps on the iCOGS chip, additional SNPs were genotyped from the October 2010 1000 Genomes Project data release, and imputation was based on the most recent January 2012 release. However, several gaps remain across the *TERT* locus, and the existence of these gaps, coupled with the low linkage disequilibrium across the region (Fig. 2), raises the possibility that there could be more independent associations that we have not yet detected. Furthermore, the incomplete SNP catalog at the time of study design means that we cannot assume with certainty that the true causal variants, directly responsible for the observed association peaks, were captured in our analysis. It is also possible that additional rare variants not specifically investigated in this study could have functional effects within this locus. Further resequencing of this region is needed to uncover the full spectrum of variation and phenotype associations. Another limitation is that telomere length was measured in DNA from leukocytes rather than from breast or ovarian tissue. Whereas we obtained suitable blood DNA for measurements in >53,000 subjects (a necessary sample size for adequate statistical power), obtaining comparable qualities and quantities of DNA from normal breast or ovary cells would be almost impossible. Telomere lengths measured in different tissues within one individual have been shown to be highly correlated64–66, meaning that leukocyte telomere lengths are likely to be good surrogates for the corresponding lengths in other tissues. Furthermore, one of our aims was to investigate whether the previously reported associations between mean telomere length and cancer risk might be mediated by *TERT* variants, and such studies have been based on telomere length measured in blood cell DNA. Another limitation was that we were unable to stratify OCAC ovarian cancer cases by *BRCA1* and *BRCA2* mutation status because this information was not available; nor was there sufficient power to evaluate ovarian cancer risk in mutation carriers in CIMBA.

Our findings provide evidence relevant to the hypothesis that shorter telomeres increase cancer risks: associations in the *TERT* promoter (peak 1) fit this hypothesis best, whereas those in peaks 2 and 3 (*TERT* introns 2–4) and other reported 5p15.33 SNP cancer associations (Supplementary Table 10) do not. Thus, it would seem that the majority of cancer associations within the *TERT* locus are mediated via alternative mechanisms involving the *TERT* gene. The protein product of *TERT* has functions beyond the telomerase-mediated extension of telomeres67. These non-canonical functions of *TERT* strongly resemble those mediated by MYC and WNT78, which are upstream regulators of proliferation, differentiation and migration. TERT also modulates WNT/β-catenin signaling69, and ectopic TERT expression induces increased cell division and decreased apoptosis in primary mammary cells, independent of telomere elongation70.

In conclusion, this study provides definitive evidence for genetic control of telomere length by common genetic variants in the *TERT* locus. Additionally, we report multiple, independent *TERT* SNP associations with breast cancer risk, confirming previously reported associations and identifying new associations in both the general population and in *BRCA1* mutation carriers. We also provide, for the first time to our knowledge, highly significant evidence for the association of distinct *TERT* SNPs with serous LMP and invasive ovarian cancer risks. Our results show that the relationships between *TERT* genotype, telomere length and cancer risk are complex and that the *TERT* locus may influence cancer risk through multiple mechanisms.


**METHODS**

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

**ACKNOWLEDGMENTS**

We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out. COGS is funded through a grant from the European Commission’s Seventh Framework Programme (agreement 223175–HEALTH-F2-2009-223175). BCAC is funded by Cancer Research UK (C1287/A10118 and C1287/A12014). BCAC meetings have been funded by the European Union Cooperation in Science and Technology (COST) programme (BM0606). Telomere length measurement and analysis were funded by Cancer Research UK project grant C1287/A9540 and Chief Physician Johan Boserup and Lise Boserup’s Fund. CIMBA data management and analysis were supported by Cancer Research UK grants C12292/A11174 and C1287/A10118. OCAC is supported by a grant from the Ovarian Cancer Research Fund thanks to the family and friends of Kathryn Sladek Smith (PPID/RPC1.07). Genotyping of the iCOGS array was funded by the European Union (HEALTH-F2-2009-223175), Cancer Research UK (C1287/A10710), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer program (J.S. and D.E.) and the Ministry of Economic Development, Innovation and Export Trade of Quebec (grant PSR-SIIRI-701; J.S., D.E. and F. Hall). Scientific development and funding of the OCAC portion of this project were supported by Genetic Associations and Mechanisms in Oncology (GAME-ON; U19-CA148112). CIMBA genotyping was supported by US National Institutes of Health (NIH) grant CA128979, a National Cancer Institute (NCI) Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), a US Department of Defense Ovarian Cancer Idea award (WB1XW10-1-0341) and grants from the Breast Cancer Research Foundation and the Komen Foundation for the Cure. This study made use of data generated by The Wellcome Trust Case Control Consortium (funding was provided by Wellcome Trust award 076113) and the TCGA Pilot Project established by NCI and the National Human Genome Research Institute.

**AUTHOR CONTRIBUTIONS**


**ARTICLES**

© 2013 Nature America, Inc. All rights reserved.


1Copenhagen General Population Study, Herlev Hospital, Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark. 2Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark. 3Centre for Cancer Genetic Epidemiology, Department of Primary Health and Primary Care, University of Cambridge, Cambridge, UK. 4Department of Genetics, Queensland Institute of Medical Research, Brisbane, Queensland, Australia. 5Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK. 6School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Queensland, Australia. 7Cancer Research Unit, Children’s Medical Research Institute, Westmead, New South Wales, Australia. 8Sydney Medical School, University of Sydney, Sydney, New South Wales, Australia. 9Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA. 10University of Queensland, UQ Centre for Clinical Research (UQCCR) Royal Brisbane and Women’s Hospital, Herston, Queensland, Australia. 11Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Rockville, Maryland, USA. 12Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, USA. 13McGill University and Génome Québec Innovation Centre, Montréal, Quebec, Canada. 14A list of members is provided in the Supplementary Note. 15Department of Obstetrics and Gynaecology, Division of Gynecologic Oncology, University of Washington, Seattle, Washington, USA. 16Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. 17Section of Biostatistics and Epidemiology, The Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire, USA. 18Department of Health Science Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota, USA. 19Kansai IDeA Network of Biomedical Research Excellence Bioinformatics Core, The University of Kansas Cancer Center, Kansas City, Kansas, USA. 20Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. 21Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. 22Department of Health Science Research, Division of Epidemiology, Mayo Clinic, Rochester, Minnesota, USA. 23Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK. 24Department of Medical Oncology, Mayo Clinic, Rochester, Minnesota, USA. 25Division of Molecular Pathology, Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands. 26Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands. 27Department of Obstetrics and Gynecology, Mayo Clinic, Rochester, Minnesota, USA. 28Division of Obstetrics and Gynecology, Helsinki University Central Hospital, University of Helsinki, Helsinki, Finland. 29Division of Hematology and Oncology, Department of Medicine, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, California, USA. 30Laboratory for Translational Genetics, Department of Medicine, University of Leuven, Leuven, Belgium. 31Salignas Research Center (VRC), VIB, Leuven, Belgium. 32Institut National de la Santé et de la Recherche Médicale (INSERM) U1018, CESP (Center for Research in Epidemiology and Population Health), Environmental Epidemiology of Cancer, Villejuif, France. 33University Paris–Sud, Unité Mixte de Recherche Scientifique (UMRS) 1018, Villejuif, France. 34Genetic and Molecular Epidemiology Group, Human Cancer Genetics Program, CNIO, Madrid, Spain. 35Department of Laboratory Medicine and Pathology, Division of Experimental Pathology, Mayo Clinic, Rochester, Minnesota, USA. 36Cancer Research UK/Yorkshire Cancer Research Sheffield Cancer Research Centre, Department of Oncology, University of Sheffield, Sheffield, UK. 37Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg, Germany. 38Molecular Epidemiology Group, DKFZ, Heidelberg, Germany. 39National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany. 40Primärmazinmedizinische Versorgung (PMV) Research Group at the Department of Child and Adolescent Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany. 41Division of Cancer Studies, National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre, Guy’s & St. Thomas’ National Health Service (NHS) Foundation Trust in partnership with King’s College London, London, UK. 42Department of Obstetrics and Gynecology, University of Ulm, Ulm, Germany. 43Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. 44Ontario Cancer Genetics Network, Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada. 45Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, New York, USA. 46Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Victoria, Australia. 47Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. 48Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Victoria, Australia. 49Division of Clinical Epidemiology and Aging Research, DKFZ, Heidelberg, Germany. 50Cancer Genomics Laboratory, Centre Hospitalier Universitaire de Québec and Laval University, Quebec City, Quebec, Canada. 51Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii, USA. 52Department of Gynecology and Obstetrics, University Breast Center for Francia Erlangen University Hospital, Erlangen, Germany. 53Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Istituto Nazionale Tumori (INT), Milan, Italy. 54Division of Genetics and Epidemiology, The Institute of Cancer Research, Sutton, UK. 55Division of Breast Cancer Research, The Institute of Cancer Research, Sutton, UK. 56Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. 57Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany. 58Faculty of Medicine, University of Tübingen, Tübingen, Germany. 59Breakthrough Breast Cancer Research Centre, Division of Breast Cancer Research, The Institute of Cancer Research, London, UK. 60Clinics of Obstetrics and Gynaecology, Hannover Medical School, Hannover, Germany. 61Laboratory of Cancer Genetics and Tumor Biology, Department of Clinical Genetics, University of Oulu, Oulu University Hospital, Oulu, Finland. 62Biocenter Oulu, University of Oulu, Oulu, Finland. 63Department of Gynecology, Jena University Hospital, Jena, Germany. 64Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands. 65Department of Pathology, Leiden University Medical
Online methods

SNP selection and genotyping. Most SNPs were genotyped on the iCOGS custom array[36,37,71]. SNPs at 5p15.33 (Build 36 positions 1,280,000–1,415,000; Build 37 positions 1,227,693–1,361,969) were selected on the basis of published cancer associations, from the March 2010 release of the 1000 Genomes Project[35]. These included all known SNPs with MAF > 0.02 in Europeans and $r^2 > 0.1$ with the then-known cancer-associated SNPs (rs4020710 (ref. 57) and/or rs3816659 (ref. 58)), plus a tagging set for all known SNPs in the linkage disequilibrium blocks encompassing the genes in the region (SLC6A18, TERT and CLPTM1L). An additional 30 SNPs in TERT were selected through a telomere length gene approach. In total, 134 SNPs were selected, 121 of which were successfully manufactured; 110 of those passed quality control[56] in BCAC and CIMBA, and 108 passed quality control in OCAC (Supplementary Tables 1–3). After genotyping, these SNPs were complemented with 22 SNPs, selected from the October 2010 release of the 1000 Genomes Project to improve coverage. These were genotyped in two BCAC studies, SEARCH[72] and CCHS[73,74], using a Fluidigm array according to the manufacturer’s instructions. To improve SNP density further, comprehensive genotype data for the locus were imputed for all subjects on the basis of the January 2012 1000 Genomes Project release. The genotype imputation process is described in refs. 36–38. All participants provided written informed consent. Ethical approval for each study/consortium is described in detail in refs. 36–38.

Samples and quality control. Study characteristics, iCOGS methodology and quality control for cancer risk analyses are detailed elsewhere[36–38]. We measured telomere length in 6,766 control samples from the SEARCH study; 1,569 of these were accrued by SEARCH itself[56], 793 were collected as part of the Sisters in Breast Screening (SIBS) study[15], and 4,404 were cancer-free participants in the European Prospective Investigation into Cancer (EPIC)-Norfolk study[19]. We also measured telomere length in 8,841 participants in CCHS[73,74] and in 38,145 participants in the Copenhagen General Population Study (CGPS)[75,76]. Genotype clusters were visually inspected for the most strongly associated SNPs (Supplementary Fig. 2). For all studies, ancestry was assigned using HapMap (release 22) genotype data for European, African and Asian populations as reference (for BCAC and CIMBA, using multidimensional scaling; for OCAC, using LAMP[77]). All CIMBA analyses were restricted to individuals of European ancestry. For BCAC, separate estimates for individuals of east Asian and African-American ancestry were also derived. For OCAC, limited analyses of non-European ancestry groups were also performed. A subset of BCAC and OCAC cases and controls was used in previous breast and ovarian cancer association studies of individual SNPs[78]. However, the associations with the key SNPs (rs10069690, rs2376108 and rs7705526) remained significant after excluding this subset of cases and controls from analysis, demonstrating similar ORs.

Telomere length measurement. Telomere length was measured in SEARCH using a modified version of the protocol described elsewhere[19,79]. Twelve percent of samples were run in duplicate. Failed PCR reactions were not repeated. Telomere length was measured in CCHS and CGPS with a modified version of the protocol described elsewhere[60,80]. Each individual was measured in quadruplicate. After exclusion of outliers, average cycle threshold (Ct) values of the remaining samples were calculated. Failed measurements were repeated up to twice. For meta-analysis, telomere length measurements from SEARCH were converted to the same scale as that used for the CCHS and CGPS measurements on the basis of parameters from the linear regression between corresponding 5-percentile groups (including the 5th, 10th, 15th, 20th, 25th, 30th, 35th, 40th, 45th, 50th, 55th, 60th, 65th, 70th, 75th, 80th, 85th, 90th, 95th, 97th and 98th percentiles) in each 10-year age group of women from CCHS and SEARCH (Supplementary Fig. 7). This measure of telomere length was used for all the analyses and then converted into fold change (RTL) to aid interpretation (Supplementary Fig. 7).

Statistical analyses. SNP associations with telomere length were evaluated using linear regression to model the fold change in telomere length per minor allele, adjusted for age, 384-well plate, sex, 7 principal components and study. Each SNP was coded as the number of minor alleles (0, 1 or 2 for genotyped SNPs and the inferred genotype for imputed SNPs). The test of association was based on the 1-degree-of-freedom trend test statistic. We also performed separate analyses (SEARCH, CCHS females, CCHS males, CGPS females and CGPS males) and combined the parameter estimates in a fixed-effect meta-analysis in STATA (StataCorp). Associations with breast and ovarian cancer risks in BCAC and OCAC were evaluated by comparing genotype frequencies in cases and controls using unconditional logistic regression. Analyses were adjusted for study and by seven principal components in BCAC[56] and five principal components in OCAC[77]. Nine OCAC studies with case-only genotype data were paired with case-control studies from similar geographic regions, resulting in 34 analysis study strata. The principal analysis fitted each SNP as an allelic dose and tested for association using a 1-degree-of-freedom trend test, but genotype-specific risks were also obtained. Associations between genotypes and breast cancer risk in CIMBA studies (BRCA1 mutation carriers) were evaluated using a 1-degree-of-freedom per-allele trend score test, based on modeling the retrospective likelihood of the observed genotypes conditional on breast cancer phenotypes[81]. To allow for non-independence among related individuals, an adjusted version of the score test was used in which the variance of the score was derived, taking into account the correlation between the genotypes by estimating the kinship coefficient for each pair of individuals using the available genotype data[82]. Per-allele HR estimates were obtained by maximizing the retrospective likelihood. All analyses were stratified by country of residence. US and Canadian strata were further stratified on the basis of reported Ashkenazi Jewish ancestry.

Conditional analyses were performed to identify SNPs independently associated with each phenotype. To identify the most parsimonious model, all SNPs with marginal P value < 0.001 were included in forward selection regression analyses with a threshold for inclusion of P < 1 × 10−4 and with terms for age (for telomere length only), principal components and study. Similarly, forward selection Cox regression analysis was performed for BRCA1 mutation carriers, stratified by country of residence, using the same P-value thresholds. This approach provides valid association tests, although the estimates can be biased[83,84]. Parameter estimates for the most parsimonious model were obtained using the retrospective likelihood approach.

FACS. Normal breast tissue was donated by women undergoing reduction mammoplasty surgery. These individuals provided written consent, and all work was performed with full local institutional human ethics approval. Tissue was dissociated as described previously[85]. Cells were prepared for flow cytometry as described previously[42] by staining with a cocktail of Lin+ markers (CD31−, CD45− and CD235a−) and Sytox Blue. Cells were then processed by a BD FACS Aria II Cell Sorter, and live cells negative for immunostaining of Lin+ markers were sorted into four subpopulations on the basis of their EpCAM-FITC and CD49F−PE-Cys5 fluorescence.

FAIRE analysis. Cell pellets derived from FACS fractionation of breast tissue samples were cross-linked in 1% formaldehyde and lysed in 200 μl of Tris-buffered 1% SDS lysis buffer containing protease inhibitors. Lysates were sonicated using a QSONICA Model Q125 Ultra Sonic Processor to shear chromatin to fragments of 200 bp to 1 kb in length. Insoluble cell material was removed through centrifugation, and supernatants were equally divided into 100-μl input and FAIRE samples. Input samples were incubated overnight at 65 °C to reverse cross-linking. All samples were purified through two rounds of phenol-chloroform extraction, and DNA was recovered through ethanol precipitation and resuspended in water for use as PCR template. Sequences for PCR primers are listed in Supplementary Table 11.

Plasmid construction and luciferase assays. TERT promoter variants were introduced into pGL3−TERT−3915 (ref. 43) by site-directed mutagenesis (Agilent Technologies). TERT PRE−A (hg19: chr 5: 1,284,900–1,287,087) and PRE−B (chr 5: 1,279,401–1,282,763) were PCR amplified using KAPAHiFi DNA polymerase (Geneworks) and cloned into pGL3−TERT−3915 or the vector encoding the minor alleles of rs2736107, rs2736108 and rs2736109. Individual SNPs were incorporated using overlap extension PCR. Sequences for PCR primers are listed in Supplementary Table 11. Cells were transfected with equimolar amounts of luciferase reporter plasmids and 50 ng of pRLTK using siPORT NeoFX Transfection Agent (Ambion), according to the manufacturer’s instructions, and harvested after 48 h. Luminescence activity was measured with a Wallac Victor3 1420 multilabel counter, and data from three replicates per construct were analyzed by one-way ANOVA with post-hoc Dunnett’s tests.
Mini-gene construction and quantitative RT-PCR analysis. TERT intron 4 was synthesized by GenScript and subcloned into pIRES-TERT<sup>4</sup>. The minor alleles at rs10069690 and rs2242652 were introduced by site-directed mutagenesis (Agilent Technologies). The resultant plasmids, designated pIRES-TERTint4-WT (wild-type intron 4), pIRES-TERTint4-rs10069690, pIRES-TERTint4-rs2242652 and pIRES-TERTint4-DM (minor alleles at both sites), were transfected into cells using siPORT NeoFX Transfection Agent, and cells were harvested after 24 h. Total RNA was extracted using the RNeasy Mini kit (Qiagen) and digested with DNase I (Invitrogen). cDNA was synthesized from 1 µg of RNA by random priming using SuperScript III reverse transcriptase (Invitrogen). Samples were screened for the presence of TERT splice variants by RT-PCR. Sequences for PCR primers are listed in Supplementary Table 11.

Molecular correlations at the 5p15.33 locus. For each gene within 1 Mb of the TERT locus, we performed the following assays: (i) gene expression analysis in ovarian cancer cell lines (n = 50) compared to ovarian surface epithelial and fallopian tube secretory cell lines (n = 73) and tissues from high-grade serous ovarian cancers; (ii) methylation analysis in high-grade serous ovarian cancers compared to normal tissues and methylation quantitative trait locus (mQTL) analysis to evaluate genotype–gene expression associations in normal high-grade serous ovarian cancer precursor tissues. We also evaluated these genes in silico in the somatic data from TCGA. We profiled the spectrum of noncoding regulatory elements in ovarian surface epithelial and fallopian tube secretory cell lines using a combination of FAIRE sequencing (FAIRE-seq<sup>49</sup>) and RNA sequencing (RNA-seq).