HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2015 August 05.

Published in final edited form as:

Nat Genet. 2009 February ; 41(2): 221–227. doi:10.1038/ng.296.

Sequence variants at the *TERT- CLPTM1L* locus associate with many cancer types

A full list of authors and affiliations appears at the end of the article.

Abstract

The common sequence variants that have recently been associated with cancer risk are particular to a single, or at most two, cancer types. Following up on our genome-wide scan of basal cell carcinoma[1], we identified rs401681(C) on chromosome 5p15.33 satisfying our threshold for genome-wide significance (OR=1.25, $P=3.7\times10^{-12}$). We tested rs401681 for association with sixteen additional cancer types in over 30,000 cancer cases and 45,000 controls and found association with lung cancer (OR=1.15, $P=7.2\times10^{-8}$) and urinary bladder, prostate and cervix cancer (ORs 1.07–1.31, all $P<4\times10^{-4}$). However, rs401681(C) appears to confer protection against cutaneous melanoma (OR=0.88, $P=8.0\times10^{-4}$). Interestingly, most of these cancer types have a strong environmental component to their risk. Investigation of the region led us to rs2736098(A), that showed stronger association with some cancer types. However, neither variant could fully account for the association of the other. Rs2736098 corresponds to A305A in the telomerase reverse transcriptase (TERT) protein while rs401681 is in an intron of the *CLPTM1L* gene.

Cancer is caused by a complex interplay between genetic and environmental factors. Highly penetrant mutations explain only a small fraction of cancer cases and the majority of genetic cancer risk is thought to be due to the contribution of many common sequence variants of low penetrance. Recently, genome-wide association (GWA) studies have yielded common sequence variants that associate with cancer risk of the prostate, breast, colon and rectum, lung, urinary bladder and skin[1–14]. Notably, in most cases the variants reported appear to be specific to the particular cancer type under study. This tissue specificity holds true even in the region on chromosome 8q24 where several independent variants have been found that associate with risk of cancer of the prostate, breast and bladder[2, 4, 6, 14–16]. Only one of the

#To whom correspondence should be addressed; kstefans@decode.is (KS) or thorunn.rafnar@decode.is (TR).
*These authors contributed equally to this work

Supplementary information is linked to the online version of the paper.

Author contributions

The study was designed and results were interpreted by TR, PS, SNS, FG, JG, JRG, DFG, AK, ST, UT, and KS. Statistical analysis was carried out by PS, FG, DFG, MF, GT and AK. Patient ascertainment, recruitment, biological material collection and collection of clinical and lifestyle information was organized and carried out by TT, SGS, RR, BS, KT, JHO, SJ, HH, TG, HI, EJ, TJ, GVE, RBB, KRB, BAA, HS, KO, AS, JH, VH, EN, S. Polidoro, S. Porru, RB-E, RK, KH, PR, KK, EG, GS, DTB, AEK, MC, EK, MPZ, PV, PdV, GM, AF, DI, MJV, RA, BS, PJ, JB, SN, AT, DK, AL, FdV, FB, WJC, JAS, HFMH, HJS, RAT, EO, OH, KKHA, JIM, LAK. Principal collaborators for the non-Icelandic populations were LAK (The Netherlands), JIM (Zaragoza, Spain), AEK (UK), GM and PV (Torino), SP (Brescia), MPZ and FB (Belgium), RK (Eastern Europe), GS (bladder cancer, Sweden), JH (melanoma, Sweden), EN (Valencia, Spain), WJC (Chicago, USA). Genotyping and laboratory experiments were carried out by AS, TB, MJ, HH, HB, MA, KTK and SM. Bioinformatic analysis was carried out by PS, TR, AS, TT, GM, TB and GT. Authors TR, PS, DFG, AK, UT and KS drafted the manuscript. All authors contributed to the final version of the paper.

The authors at deCODE genetics declare competing financial interests.

prostate cancer variants has been shown also to associate with risk of another cancer; colorectal cancer[7].

We previously performed a GWA study of basal cell carcinoma (BCC) of the skin where we found two signals that reach genome-wide significance in a combined analysis of Icelandic and European sample sets[1]. Here, we followed up the initial GWA scan, increasing the effective sample size by genotyping additional Icelandic BCC cases (total of 1025 cases) using the Illumina HumanCNV370-duo chip. Furthermore, we used a method where known genotypes of relatives are used to provide information on BCC cases not genotyped (*in silico* genotyping) to add genotypes that are equivalent to an additional 480 BCC cases[17]. Analysis of this larger dataset confirmed the two previously reported loci as containing the strongest signals (Supplementary Figure 1). The third strongest signal was on chromosome 5p15.33 in an area of high linkage disequilibrium (LD) and was represented by two correlated SNPs, rs401681 and rs31489 ($D'=1$ and $r^2=0.87$ in data from HapMap CEU and the Icelandic controls) (Figure 1). The allele C of rs401681 had an odds ratio (OR) of 1.25 (95% CI 1.15–1.36; $P=2.3\times10^{-7}$) and allele C of rs31489 had an OR of 1.25 (95% CI 1.15–1.36; $P=1.9\times10^{-7}$) (Supplementary Table 1). We selected rs401681 for follow up genotyping in an additional 744 BCC cases from Iceland, including some that were utilized in the initial analysis through *in silico* genotyping, and 525 BCC cases and 515 controls from Eastern Europe (for a summary of cases and controls used in this paper, see Table 1). In the combined analysis of the Icelandic and Eastern European samples, the association of rs401681(C) with BCC reached genome-wide significance (OR=1.25; $P=3.7\times10^{-12}$) (Table 2). We did not observe heterogeneity of the ORs between the Icelandic and East European groups ($P=0.35$). Furthermore, based on results from both groups, the association between rs401681(C) and BCC did not show significant deviation from the multiplicative risk model ($P>0.05$).

Rs401681 resides in an LD block that contains the *CLPTM1L* (cisplatin resistance related protein CRR9p) gene and the 5′ end of the *TERT* (human telomerase reverse transcriptase) gene. CLPTM1L is a predicted transmembrane protein that is expressed in a range of normal and malignant tissues including skin, lung, breast, ovary and cervix. Expression of CLPTM1L has been shown to sensitize ovarian cancer cells to cisplatin-induced apoptosis[18]. The *TERT* gene encodes the catalytic subunit of the telomerase ribonucleoprotein complex (telomerase). The major role of telomerase is to catalyse the *de novo* addition of telomeric repeat sequences onto chromosome ends and thereby counterbalance telomere-dependent replicative aging[19]. Several studies have reported an association between short telomeres and increased risk of cancer at several sites, including BCC, cancers of the lung, head and neck, bladder, kidney, oesophagus and breast, as well as lymphoma[20–24]. Furthermore, this region is frequently amplified in many types of cancer such as lung and cervical cancer[25, 26].

Given the relevance of this genomic region to cancer biology, we assessed the association of rs401681(C) with 16 additional cancer types in individuals of European ancestry (Table 1). Altogether, we directly genotyped rs401681 on approximately 20,500 cases and 36,000 controls. Using *in silico* genotyping, information corresponding to genotypes of 4,265 Icelandic cancer cases could also be added (Table 1 and Supplementary Table 2). The results from the directly- and *in silico* genotyped samples were combined with summary level data

from publicly available GWA datasets, specifically the dataset on colorectal cancer from the U.K. Institute of Cancer Research (ICR)[27], the dataset on lung cancer from the International Agency for Research on Cancer (IARC)[11] and the datasets on prostate and breast cancer from the Cancer Genetics Markers of Susceptibility (CGEMS) study group[4]. In total, we assessed the association of rs401681(C) to 17 individual cancer sites using more than 33,800 cancer cases and 45,800 controls.

Of the 16 cancer sites tested in addition to BCC, 5 sites showed nominally significant association ($P<0.05$) with rs401681(C) with the same direction of the effect (Table 2 and Supplementary Table 3). Of those, 4 cancer sites showed significant association after accounting for the 16 additional cancer sites tested ($P < 0.05/16 = 0.003$) (Table 2). These cancer sites are lung, urinary bladder, prostate and cervix of the uterus. The strongest association of rs401681(C), following BCC, was with lung cancer, with an OR of 1.15 ($P=7.2\times10^{-8}$) in the combined analysis of samples from Iceland, the Netherlands and Spain, in addition to the lung cancer dataset from the IARC. We observed an association between rs401681(C) and bladder cancer with a combined OR of 1.12 ($P=5.7\times10^{-5}$) for the 9 European case control groups tested. For prostate cancer, we were able to analyze data from 5 groups (over 9,000 cases) and demonstrated a significant effect that is consistent among the groups tested with combined OR=1.07 ($P=3.6\times10^{-4}$). For cervical cancer, where we only had samples from Iceland, we detected a significant association to rs401681(C) (OR=1.31, $P=2.6\times10^{-4}$). No signs of heterogeneity or deviation from the multiplicative model were observed.

We did not detect an association between rs401681(C) and breast cancer (OR=0.98; P=0.340) even though a large sample set was used (3,645 cases and 30,030 controls) (Supplementary Table 3). Endometrial cancer showed a trend that did not reach significance after adjusting for the number of tests (OR=1.21, $P=5.5\times10^{-3}$). Interestingly, we observed a significant association between rs401681(C) and protection against cutaneous melanoma (OR=0.88, $P=8.0\times10^{-4}$) in a sample set consisting of 2,443 melanoma cases and 30,839 controls from Iceland, Sweden and Spain. We note that a recently published study of telomere length in individuals with skin cancers showed that while short telomeres are associated with increased risk of BCC, long telomeres are associated with increased risk of melanoma[24]. The rs401681(C) variant was also marginally associated with protection against colorectal cancer (OR=0.95, $P=8.4\times10^{-3}$) although this was not significant after taking into account the number of cancer sites tested. We observed no association with cancers of the kidney, stomach, thyroid, ovary, pancreas, lymphoma, multiple myeloma or SCC of the skin. However, the moderate sizes of the sample sets tested do not allow us to draw definitive conclusion about the lack of association and further assessment will be needed in larger sample sets.

To explore the potential contributions from other variants in the region, we indirectly tested SNPs that are in the HapMap CEU database, but not on the HumanHap300 or HumanCNV370-duo chips (Supplementary Methods). Using the chip-genotyped Icelandic samples, rs2736100 and rs4975616 were used to tag rs2736098 and perform an indirect test of its association with BCC, giving a P of 3.9×10^{-8}. Rs2736098 is a synonymous coding SNP (Ala305Ala) in the second exon of the *TERT* gene (Figure 1). We proceeded to directly

genotype rs2736098 in all available cases of the 5 cancer types showing association with rs401681, in a subset of all the other cancer cases (total of 14,389 cancer cases), in all available non-Icelandic controls and in a subset of the Icelandic controls (total of 9,703 controls). With the exception of cervical cancer, rs2736098(A) was significantly associated with each of the other four cancers that associated significantly with rs401681(C) (Table 3). Allele A of rs2736098 is positively correlated with rs401681(C) and the D′ between rs2736098 and rs401681 is very high (0.94), but the value of r^2 (0.39) while substantial, is more moderate. For three of the cancers that associated significantly with rs2736098(A), lung, bladder and prostate, the estimated OR for rs2736098(A) was higher than that for rs401681(C). In the case of BCC, the OR for rs2736098(A) was, however, lower than that for rs401681(C) mainly because the latter did not show association with the disease in the Eastern European samples (heterogeneity P=0.0035).

We examined the joint effect of rs401681(C) and rs2736098(A), for each of the 5 cancers, using only samples typed for both SNPs (Table 4). After adjusting for rs2736098(A), the association of rs401681(C) remained significant in all except prostate cancer. After adjusting for rs401681(C), rs2736098(A) remained significant for 3 cancers, lung, bladder and prostate. Overall, these results indicate that neither rs401681(C) nor rs2736098(A) can, by themselves, fully account for the association observed between sequence variants in this region and the 5 cancer types. This suggests that a unique variant capturing the effect of both rs401681(C) and rs2736098(A) remains to be discovered or, alternatively, that the region contains more than one variant that predisposes to cancers at the same or different sites, analogous to the region on 8q24 where independent variants have been found that associate with different cancer types. We analyzed the association between the 27 SNPs depicted in Figure 1 and the 17 cancer types studied using the Icelandic sample sets and found that 15 sites showed an association with one or more of these SNPs at the P<0.05 level (Supplementary Table 4).

We assessed known missense variants in *CLPTM1L* and *TERT*, the potentially functional *TERT* promoter variant (−1327T/C), as well as the variable number of tandem repeats (VNTRs) in introns 2, 6 and 12 of *TERT* and found that none of these variants associated significantly with any of the 5 cancer sites or could account for the association observed with rs401681 or rs2736098 (Supplementary Note and Supplementary Tables 5 and 6). Furthermore, no association was observed between rs401681(C) or rs2736098(A) and the RNA expression of *TERT* or *CLPTM1L* in whole blood (N=991) or adipose tissue (N=662) (Supplementary Note and Supplementary Table 7).

We postulated that the cancer-associated sequence variants in the *TERT* gene might be associated with shorter telomeres. In order to test this hypothesis, we examined the association between rs401681 and rs2736098 and telomere length in DNA from whole blood, using a quantitative PCR assay. To limit variability, we took into account several factors that have been reported to affect telomere length, including age, gender and smoking status[28, 29] and selected from our database 276 females born between 1925 and 1935 who reported to have never smoked and who had not been diagnosed with cancer. To maximize the contrast, only women homozygous for allele C or allele T at rs401681 were included in the test. In these subjects, rs401681(C) and rs2736098(A) were associated with shorter

telomeres with nominal significance (P=0.017 and 0.027, respectively) (Supplementary Figure 2, Supplementary Table 8). However, when we tested telomere length in a group of 260 younger women (selected by the same criteria regarding smoking and cancer, but born between 1940 and 1950), there was no association between telomere length and the risk alleles. Indeed, the effect estimates, while insignificant (P=0.08 and 0.28 for rs401681 and rs2736098, respectively) were in the opposite direction (Supplementary Figure 2, Supplementary Table 8). These results suggest that the variants may lead to an increase in the gradual shortening of telomeres over time, the effect only becoming apparent after a certain age. Further testing of these hypotheses is warranted.

We assessed the association of rs401681(C)) and rs2736098(A) with the major histological types of lung cancer (Supplementary Table 9). For all histological types except carcinoids, the frequency of the risk variants was higher than in controls with the highest frequencies found in squamous cell carcinomas. Among the cervix cancer cases, squamous cell carcinomas showed an insignificant trend towards stronger association with the risk variants than adenocarcinomas. In prostate cancer, the variants were not associated with age at diagnosis or disease aggressiveness, as defined by Gleason score ≥7 and/or stage T3 or higher and/or node positive and/or metastatic disease. We found no association between rs401681(C) and either smoking status or nicotine addiction or any of several pigment phenotypes assessed (hair and eye colour, freckling or Fitzpatrick skin-type score) (Supplementary Note). Furthermore, neither rs401681(C) nor rs2736098(A) were more strongly associated with bladder cancer risk in smokers (current or former, n=4,346) than in non-smokers (n=556). Finally, rs401681(C) did not associate with longevity (P=0.50) (Supplementary Note).

It is of interest that 4 of the 5 cancers associated with the risk variants are cancer types that have strong environmental contribution to risk, i.e. smoking and occupational exposures for lung and bladder cancer, UV irradiation for BCC and infection with human papillomavirus for cervical cancer. The majority of cancers in these organs arise in the epithelial layer that is in closest contact with the environment. Although no strong environmental risk factors are currently known for prostate cancer, several external factors such as diet, physical activity and inflammation may have an effect on disease risk. Although telomere length is partly inherited[30] various environmental factors such as smoking and radiation also affect telomere length[28].

In conclusion, we have discovered sequence variants in the region of the *TERT* and *CLPTM1L* genes that associate with risk of many types of cancer. The biology of the *TERT* gene makes it a compelling candidate for a gene that predisposes to many cancers. Further investigations of the potential effects of genetic variants at the 5p15.33 locus on the functions and expression of *TERT* and *CLPTM1L* are clearly warranted.

Methods

A fully referenced copy of the Methods can be found in the Supplementary Methods.

Genotyping

Detailed information on all case control sample sets are found in the Supplementary Methods. Whole-genome association studies have been performed on the following cancers in the Icelandic population; prostate cancer, breast cancer, lung cancer, BCC, melanoma, urinary bladder cancer and colorectal cancer[1–3, 9, 14]. All cases and controls were assayed using genotyping systems and specialized software from Illumina (Human Hap300 and HumanCNV370-duo Bead Arrays, Illumina). Furthermore, all Dutch bladder cancer cases and controls have been genotyped with the HumanCNV370-duo Bead Arrays [14]. These chips provide about 75% genomic coverage in the Utah CEPH (CEU) HapMap samples for common SNPs at r2>0.8. SNP data were discarded if the minor allele frequency in the combined case and control was <0.001 or had less than 95% yield or showed a very significant distortion from Hardy-Weinberg equilibrium in the controls ($P<1\times10^{-10}$). Any chips with a call rate below 98% of the SNPs were excluded from the genome-wide association analysis.

All single SNP genotyping was carried out applying the Centaurus (Nanogen) platform. The quality of each Centaurus SNP assay was evaluated by genotyping each assay in the CEU HapMap samples and comparing the results with the HapMap publicly released data. Assays with >1.5% mismatch rate were not used and a linkage disequilibrium (LD) test was used for markers known to be in LD. Approximately 10% of the Icelandic case samples that were genotyped on the Illumina platform were also genotyped using the Centaurus assays and the observed mismatch rate was lower than 0.5%. All genotyping was carried out at deCODE Genetics.

Assessment of telomere length

We selected whole blood as the tissue for analyzing telomere length for its accessibility but studies have shown that the length of telomeres is very similar within different tissues of the same individual but vary significantly between individuals. Telomeres were measured utilizing a quantitative Taqman® PCR assay. RNAseP endogenous control assay (Cat.no. 4316844) (Applied Biosystems Inc.) was used to correct for DNA input. This quantitative PCR method has been shown to give consistent results as Southern blot and FISH based telomere measurments. All reactions were run on ABI7900TH real time PCR system (Applied Biosystems Inc.). All assays were done in duplicate and repeated in an independent experiment. Primers and probes used are listed in Supplementary Table 10. The use of RNAseP is a standard procedure in gene dosage measurements with real time qRT-PCR. The main limitation of the method is that it measures relative telomere length rather than actual telomere length. However, in our study the relative telomere length is sufficient for determining if there is a difference in telomere length between individuals depending on their genotype.

Regression Analysis of Telomere Length Data

A total of 528 females were analyzed in two batches, each batch consisting of 3 DNA plates, batch 1 included 268 women with a mean age at blood sampling of 72.8 (SD 5.0) years, batch 2 included 260 women with a mean age at blood sampling of 57.8 (SD 4.6) years. The relationship between the SNPs showing association and telomerase length was analyzed by

multiple regression. The logarithm of the ratio between telomerase and RNAseP was taken as dependent variable, and the covariates age at blood sampling and DNA plates were included in the models. SNPs showing association were analyzed using multiple linear regression. The experiments were carried out at two different points in time and were analyzed separately.

hTERT minisatellite genotyping

Genotyping of the five variable number tandem repeat (VNTRs) polymorphisms in the intronic regions of the hTERT gene was performed by size fractionation of PCR products by gel electrophoresis. Two polymerase chain reaction systems were used, Recombinant Taq (Fermentas Inc.) and Extensor-High Fidelity Master Mix (ABgene Inc). PCR conditions were according to the manufacturer instructions. DNA input material was 30 ng. Primers used are listed in Supplementary Table 10.

Analysis of RNA expression

Samples of RNA from human adipose and whole blood were hybridized to Agilent Technologies Human 25k microarrays. Expression changes between two samples were quantified as the mean logarithm (log10) expression ratio (MLR) compared to a reference pool RNA sample. The array probes for *CLPTM1L* and *TERT* genes were in the 3′ untranslated regions of the genes.

Association analysis

A likelihood procedure implemented in the NEMO software was used for the association analyses. An attempt was made to genotype all individuals and all SNPs reported, and for each of the SNPs, the yield was higher than 95% in every study group. We tested the association of an allele to cancer using a standard likelihood ratio statistic that, if the subjects were unrelated, would have asymptotically a χ^2 distribution with one degree of freedom under the null hypothesis. Allelic frequencies rather than carrier frequencies are presented for the markers in the main text. Allele-specific ORs and associated P values were calculated assuming a multiplicative model for the two chromosomes of an individual. Results from multiple case-control groups were combined using a Mantel-Haenszel model in which the groups were allowed to have different population frequencies for alleles, haplotypes and genotypes but were assumed to have common relative risks. All P values are reported as two-sided.

For the analysis of the Icelandic samples, the same set of cancer free controls used in the BCC discovery analysis was used for all other cancer types, introducing a potential bias. However, due to the lack of association with common cancers like breast and colorectal cancer and also because of the modest effect sizes for the cancers associating with rs401681(C), the frequency of the variant is not substantially different in the Icelandic cancer free controls (0.545) compared to the whole group of Icelanders (N=36,139) genotyped with the BeadChips (0.547) which includes all cancer cases. Therefore, the potential bias introduced into the estimation of the association of the sixteen cancers with rs401681(C) is small. Furthermore, this effect is confined to the Icelandic part of our study.

Test of un-genotyped Hapmap markers

To test for SNP that are in the CEU section of the Hapmap database, but that are absent on the Illumina chip, we use a method based on haplotypes of two markers on the chip. We computed associations with a linear combination of the different haplotypes chosen to act as surrogates to HapMap markers in the regions. In the 5p13.33 region displayed in Figure 1 (corresponding to a 200 kb interval), we tested with this method 95 markers in addition to the ones on the chip. These calculations were based on 1,025 BCC cases and 28,890 controls genotyped on chip. Of those markers, rs2736098 had the most significantly association with BCC.

Genomic control and inflation factors

To adjust for possible population stratification and the relatedness amongst individuals, we divided the χ^2 statistics from the initial scan of basal cell carcinoma in Iceland, using the method of genomic control, i.e. the 304 thousand test statistics were divided by their means, which was 1.22. In the cases where the method of genomic control is not directly applicable (i.e. if the genome wide association results are not available for the same groups), we used the genealogy to estimate the inflation factor. Since some of the Icelandic patients and controls are related to each other, both within and between groups, the χ^2 statistics have a mean >1. We estimated the inflation factor by simulating genotypes through the Icelandic genealogy, as described previously, and corrected the χ^2 statistics for Icelandic OR's accordingly. The estimated inflation factor for different analyses is presented in Supplementary table 11.

In silico genotyping of un-genotyped individuals

We extended the classical SNP case-control association study design by including un-genotyped cases with genotyped relatives, using information from genotyped individuals in the Icelandic population and the genealogy of all Icelanders. For a detailed description of the calculations of the probability of genotypes see Supplementary Methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Thorunn Rafnar[1,*,#], Patrick Sulem[1,*], Simon N. Stacey[1], Frank Geller[1], Julius Gudmundsson[1], Asgeir Sigurdsson[1], Margret Jakobsdottir[1], Hafdis Helgadottir[1], Steinunn Thorlacius[1], Katja K.H. Aben[2,3], Thorarinn Blöndal[1], Thorgeir E. Thorgeirsson[1], Gudmar Thorleifsson[1], Kristleifur Kristjansson[1], Kristin Thorisdottir[4], Rafn Ragnarsson[5], Bardur Sigurgeirsson[4], Halla Skuladottir[6], Tomas Gudbjartsson[7,43], Helgi J. Isaksson[8], Gudmundur V. Einarsson[9], Kristrun R. Benediktsdottir[8,43], Bjarni A. Agnarsson[8,43], Karl Olafsson[10], Anna Salvarsdottir[10], Hjordis Bjarnason[1], Margret Asgeirsdottir[1], Kari T. Kristinsson[1], Sigurborg Matthiasdottir[1], Steinunn G. Sveinsdottir[11], Silvia Polidoro[12,13], Veronica Höiom[14], Rafael Botella-Estrada[15], Kari Hemminki[16], Peter Rudnai[17], D. Timothy Bishop[18],

Marcello Campagna[19], Eliane Kellen[20,21], Maurice P. Zeegers[22,23], Petra de Verdier[24], Ana Ferrer[25], Dolores Isla[25], Maria Jesus Vidal[25], Raquel Andres[25], Berta Saez[26], Pablo Juberias[27], Javier Banzo[28], Sebastian Navarrete[29], Alejandro Tres[25,30], Donghui Kan[31], Annika Lindblom[24], Eugene Gurzau[32], Kvetoslava Koppova[33], Femmie de Vegt[3], Jack A. Schalken[34], Henricus F.M. van der Heijden[35], Hans J. Smit[36], René A. Termeer[37], Egbert Oosterwijk[34], Onno van Hooij[34], Eduardo Nagore[15], Stefano Porru[19], Gunnar Steineck[14,38], Johan Hansson[14], Frank Buntinx[20,39], William J. Catalona[31], Giuseppe Matullo[12,13], Paolo Vineis[12,40], Anne E. Kiltie[41], José I. Mayordomo[25,30], Rajiv Kumar[16], Lambertus A. Kiemeney[2,3,34], Michael L. Frigge[1], Thorvaldur Jonsson[7,43], Hafsteinn Saemundsson[10], Rosa B. Barkardottir[8], Eirikur Jonsson[9], Steinn Jonsson[42,43], Jon H. Olafsson[4,43], Jeffrey R. Gulcher[1], Gisli Masson[1], Daniel F. Gudbjartsson[1], Augustine Kong[1], Unnur Thorsteinsdottir[1,43], and Kari Stefansson[1,43,#]

Affiliations

[1]deCODE Genetics, Sturlugata 8, 101 Reykjavik, Iceland [2]Comprehensive Cancer Center East, 6501 BG Nijmegen, The Netherlands [3]Dept. of Epidemiology, Biostatistics and HTA, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands [4]Dept. of Dermatology, University of Iceland, 101 Reykjavik, Iceland [5]Dept. of Plastic Surgery, Landspitali-University Hospital, 101 Reykjavik, Iceland [6]Dept. of Medical Oncology, Landspitali-University Hospital, 101 Reykjavik, Iceland [7]Dept. of Surgery, Landspitali-University Hospital, 101 Reykjavik, Iceland [8]Dept. of Pathology, Landspitali-University Hospital, 101 Reykjavik, Iceland [9]Dept. of Urology, Landspitali-University Hospital, 101 Reykjavik, Iceland [10]Dept. of Obstetrics and Gynecology, Landspitali-University Hospital, 101 Reykjavik, Iceland [11]Clinical Research Centre, Krokhals 5D, 110 Reykjavik, Iceland [12]ISI Foundation (Institute for Scientific Interchange), 10133 Torino, Italy [13]Dept. of Genetics, Biology and Biochemistry, University of Torino, 10126 Torino, Italy [14]Dept. of Oncology-Pathology, Cancer Centre Karolinska, Karolinska Institutet, Karolinska University Hospital Solna Stockholm, S-171 76, Sweden [15]Dept. of Dermatology, Instituto Valenciano de Oncologia, 46009, Valencia, Spain [16]Division of Molecular Genetic Epidemiology, German Cancer Research Centre, Heidelberg, D-69120, Germany [17]National Institute of Environmental Health, H-1450 Budapest, Hungary [18]Section of Epidemiology & Biostatistics, Leeds Institute of Molecular Medicine, St. James's University Hospital, LS9 7TF Leeds, UK [19]Dept. of Experimental and Applied Medicine - Section of Occupational Medicine and Industrial Hygiene, University of Brescia, 25125 Brescia, Italy [20]Dept. of General Practice, Catholic University of Leuven, 3000 Leuven, Belgium [21]Leuven University Centre for Cancer Prevention (LUCK), 3000 Leuven, Belgium [22]Unit of Genetic Epidemiology, Dept. of Public Health and Epidemiology. University of Birmingham, B15 2TT Birmingham, UK [23]Dept. of Complex Genetics, Cluster of Genetics and Cell Biology, Nutrition and Toxicology Research Institute, Maastricht University, 6200 MD Maastricht, The Netherlands [24]Dept. of Molecular Medicine and Surgery, Karolinska Institutet, S171 76 Stockholm, Sweden [25]Division of Medical Oncology, University Hospital, 50009 Zaragoza, Spain [26]Health Science Institute. 50009 Zaragoza, Spain [27]Division of

Dermatology. University Hospital. 50009 Zaragoza, Spain [28]Division of Nuclear Medicine. University Hospital. 50009 Zaragoza, Spain [29]Division of Radiation Oncology. University Hospital. 50009 Zaragoza, Spain [30]Health Science Institute, Nanotechnology Institute of Aragon, 50009, Zaragoza, Spain [31]Dept. of Urology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA [32]Environmental Health Center, 400166 Cluj Napoca, Romania [33]State Health Institute, SK-975 56 Banska Bystrica, Slovakia [34]Dept. of Urology, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500 HB, Nijmegen, the Netherlands [35]Dept. of Pulmonary Diseases, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500 HB, Nijmegen, the Netherlands [36]Dept. of Pulmonary Medicine, Rijnstate Hospital, Arnhem, the Netherlands [37]Dept. of Pulmonary Diseases, Canisius-Wilhelmina Hospital, Weg door Jonkerbos 100, 6532 SZ, Nijmegen, The Netherlands [38]Dept. of Oncology, Sahlgrenska University Hospital, S-413 45 Goteborg, Sweden [39]Dept. of General Practice, Maastricht University, 6200 MD Maastricht, The Netherlands [40]Dept. of Epidemiology and Public Health, Imperial College, W2 1PG London, UK [41]Section of Experimental Oncology, Leeds Institute of Molecular Medicine, St. James's University Hospital, LS9 7TF Leeds, UK [42]Dept. of Medicine, Landspitali-University Hospital, 101 Reykjavik, Iceland [43]Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland

Acknowledgments

We thank the individuals that participated in the study and whose contribution made this work possible. We also thank the nurses at deCODE's participant recruitment center and the personnel at deCODE's core facilities. We acknowledge the Icelandic Cancer Registry for assistance in the ascertainment of the Icelandic cancer patients. The project was funded in part by the European Commission (POLYGENE: LSHC-CT-2005-018827 and GENADDICT: LSHM-CT-2004-005166), the National Institute of Health (R01-DA017932) and a research investment grant of the Radboud University Nijmegen Medical Centre. The Leeds Bladder Cancer Study was funded by Cancer Research UK and Yorkshire Cancer Research. Torino Bladder Cancer Case Control Study was supported by a grant to ECNIS (Environmental Cancer Risk, Nutrition and Individual Susceptibility), a network of excellence operating within the European Union 6th Framework Program, Priority 5: "Food Quality and Safety" (Contract No 513943); by a grant of the compagnia di San Paolo, the Italian Association for Cancer Research, Italy and the Piedmont Region Progetti di Ricerca Sanitaria Finalizzata. JH is supported by the Swedish Cancer Society, The Radiumhemmet Research Funds, The Swedish Research Council and the Karolinska Institutet Research Funds.

References

1. Stacey SN, et al. Common variants on 1p36 and 1q42 are associated with cutaneous basal cell carcinoma but not with melanoma or pigmentation traits. Nat Genet. 2008; 40:1313–8. [PubMed: 18849993]
2. Gudmundsson J, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet. 2007; 39:631–7. [PubMed: 17401366]
3. Stacey SN, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet. 2007; 39:865–9. [PubMed: 17529974]
4. Yeager M, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet. 2007; 39:645–9. [PubMed: 17401363]
5. Haiman CA, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet. 2007; 39:638–44. [PubMed: 17401364]
6. Easton DF, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007; 447:1087–93. [PubMed: 17529967]

7. Tomlinson I, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet. 2007; 39:984–8. [PubMed: 17618284]
8. Gudbjartsson DF, et al. ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. Nat Genet. 2008; 40:886–91. [PubMed: 18488027]
9. Thorgeirsson TE, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature. 2008; 452:638–42. [PubMed: 18385739]
10. Eeles RA, et al. Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet. 2008; 40:316–21. [PubMed: 18264097]
11. Hung RJ, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature. 2008; 452:633–7. [PubMed: 18385738]
12. Amos CI, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet. 2008; 40:616–22. [PubMed: 18385676]
13. Thomas G, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet. 2008; 40:310–5. [PubMed: 18264096]
14. Kiemeney LA, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat Genet. 2008; 40:1307–12. [PubMed: 18794855]
15. Amundadottir LT, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet. 2006; 38:652–8. [PubMed: 16682969]
16. Haiman CA, et al. A common genetic risk factor for colorectal and prostate cancer. Nat Genet. 2007; 39:954–6. [PubMed: 17618282]
17. Gudbjartsson DF, et al. Many sequence variants affecting diversity of adult human height. Nat Genet. 2008; 40:609–15. [PubMed: 18391951]
18. Yamamoto K, Okamoto A, Isonishi S, Ochiai K, Ohtake Y. A novel gene, CRR9, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. Biochem Biophys Res Commun. 2001; 280:1148–54. [PubMed: 11162647]
19. Blackburn EH. Switching and signaling at the telomere. Cell. 2001; 106:661–73. [PubMed: 11572773]
20. Wu X, et al. Telomere dysfunction: a potential cancer predisposition factor. J Natl Cancer Inst. 2003; 95:1211–8. [PubMed: 12928346]
21. Shen J, et al. Short telomere length and breast cancer risk: a study in sister sets. Cancer Res. 2007; 67:5538–44. [PubMed: 17545637]
22. Widmann TA, Herrmann M, Taha N, Konig J, Pfreundschuh M. Short telomeres in aggressive non-Hodgkin's lymphoma as a risk factor in lymphomagenesis. Exp Hematol. 2007; 35:939–46. [PubMed: 17533048]
23. Risques RA, et al. Leukocyte telomere length predicts cancer risk in Barrett's esophagus. Cancer Epidemiol Biomarkers Prev. 2007; 16:2649–55. [PubMed: 18086770]
24. Han J, et al. A Prospective Study of Telomere Length and the Risk of Skin Cancer. J Invest Dermatol. 2008
25. Zhang A, et al. Genetic alterations in cervical carcinomas: frequent low-level amplifications of oncogenes are associated with human papillomavirus infection. Int J Cancer. 2002; 101:427–33. [PubMed: 12216070]
26. Kang JU, Koo SH, Kwon KC, Park JW, Kim JM. Gain at chromosomal region 5p15.33, containing TERT, is the most frequent genetic event in early stages of non-small cell lung cancer. Cancer Genet Cytogenet. 2008; 182:1–11. [PubMed: 18328944]
27. Tomlinson IP, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. Nat Genet. 2008; 40:623–30. [PubMed: 18372905]
28. Valdes AM, et al. Obesity, cigarette smoking, and telomere length in women. Lancet. 2005; 366:662–4. [PubMed: 16112303]
29. Frenck RW Jr, Blackburn EH, Shannon KM. The rate of telomere sequence loss in human leukocytes varies with age. Proc Natl Acad Sci U S A. 1998; 95:5607–10. [PubMed: 9576930]
30. Slagboom PE, Droog S, Boomsma DI. Genetic determination of telomere size in humans: a twin study of three age groups. Am J Hum Genet. 1994; 55:876–82. [PubMed: 7977349]

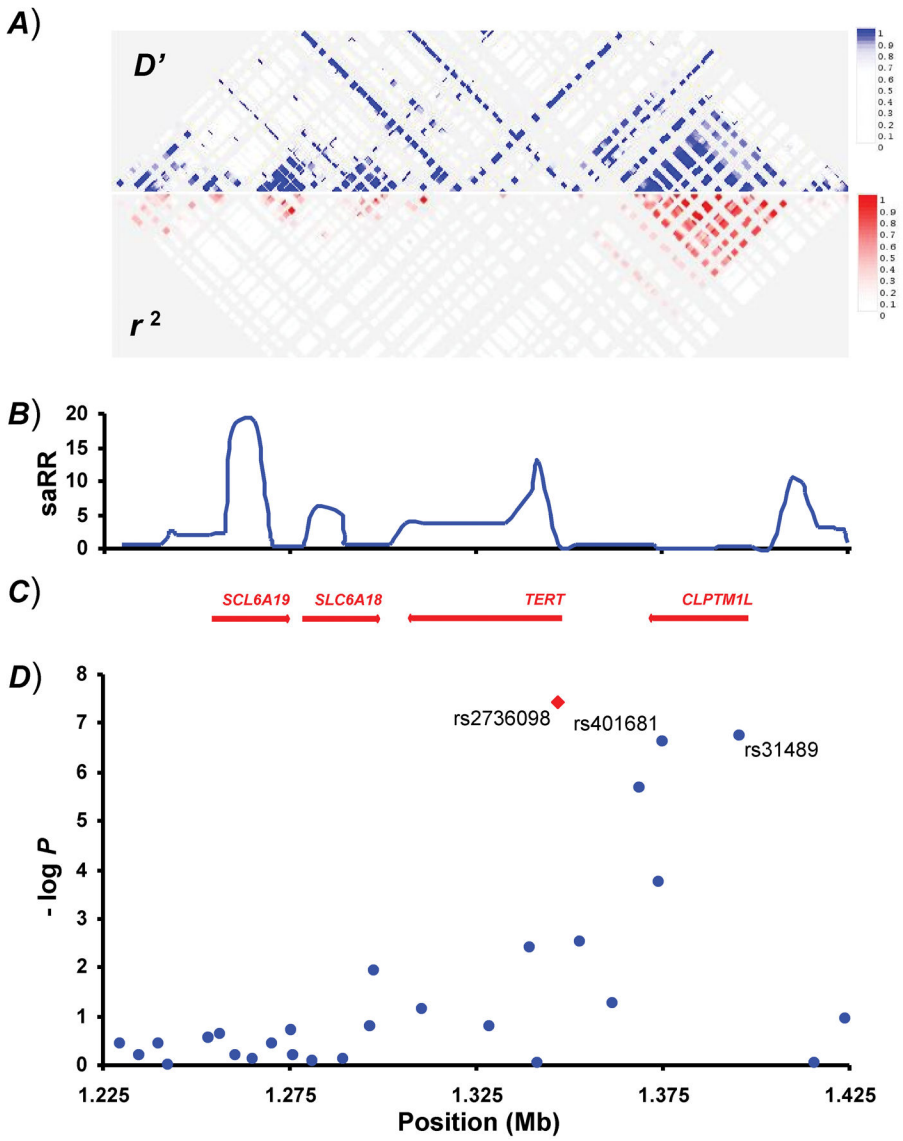

Figure 1.

A schematic view of the association results and LD-structure in a region on chromosome 5p15.33. A) The pair-wise correlation structure in a 200 kb interval (1.225 – 1.425 Mb, NCBI B36) on chromosome 5. The upper plot shows pair-wise D′ for 100 common SNPs (with MAF > 5%) from the HapMap (v22) CEU dataset. The lower plot shows the corresponding r^2 values. B) Estimated recombination rates (saRR) in cM/Mb from the HapMap Phase II data. C) Location of known genes in the region. D) Schematic view of the association with BCC in the Icelandic sample set consisting of cases genotyped by chip or *in silico* (blue dots). Red triangle shows the location of rs2736098 and corresponding significance of association to BCC, testing for the HapMap CEU markers absent on the chip for individuals directly genotyped on chip.

Table 1

Cancer cases and controls used in the study

Cancer site	N sample sets	N cases		N controls
		directly genotyped	genotyped *in silico**	
Basal cell carcinoma (skin)	2	2,294	271	29,405
Lung	4	3,613	652	34,666
Bladder	9	3,945	202	34,988
Prostate	4	8,951	522	37,901
Cervix	1	276	93	28,890
Breast	2	3,089	556	30,030
Colon and rectum	2	1,966	529	29,817
Melanoma	3	2,381	62	30,839
Endometrium	1	387	83	28,890
Kidney	2	784	203	30,722
Lymphoma	1	178	70	28,890
Multiple Myeloma	1	64	62	28,890
Ovary	1	363	134	28,890
Pancreas	1	75	226	28,890
Squamous cell carcinoma (skin)	1	547	ND	28,890
Stomach	1	277	485	28,890
Thyroid	1	413	115	28,890
Total		**29,603**	**4265**	**45,846**

* Effective sample size, see methods

ND= Not done

Table 2

Association of rs401681 (C) on 5p15.33 to basal cell carcinoma and cancers of the lung, bladder, prostate and cervix

Study population	Number		Frequency		OR	95% CI	*P* value
	Cases	Controls	Cases	Controls			
Basal cell carcinoma							
Iceland all[a]	2,040	28,890	0.604	0.545	1.27	1.19–1.36	9.5×10^{-12}
Iceland genotyped[b]	1,769	28,890	0.602	0.545	1.26	1.17–1.35	4.3×10^{-10}
Eastern Europe	525	515	0.616	0.575	1.16	0.97–1.39	0.098
All combined[c]	**2,565**	**29,405**	**0.610**	**0.560**	**1.25**	**1.18–1.34**	$\mathbf{3.7\times10^{-12}}$
Lung cancer							
Iceland all[a]	1,449	28,890	0.575	0.545	1.13	1.04–1.23	3.6×10^{-3}
Iceland genotyped[b]	797	28,890	0.584	0.545	1.18	1.06–1.32	2.8×10^{-3}
The Netherlands	529	1,832	0.610	0.570	1.18	1.02–1.35	0.021
Spain	367	1,427	0.582	0.538	1.19	1.01–1.41	0.034
IARC	1,920	2,517	0.617	0.586	1.16	1.06–1.27	8×10−4
All combined[c]	**4,265**	**34,666**	**0.596**	**0.560**	**1.15**	**1.10–1.22**	$\mathbf{7.2\times10^{-8}}$
Bladder cancer							
Iceland all[a]	780	28,890	0.583	0.545	1.16	1.05–1.29	4.5×10^{-3}
Iceland genotyped[b]	578	28,890	0.583	0.545	1.17	1.03–1.32	0.012
The Netherlands	1,277	1,832	0.584	0.570	1.06	0.96–1.17	0.27
UK	707	506	0.564	0.514	1.23	1.04–1.44	0.014
Italy-Torino	329	379	0.550	0.545	1.02	0.84–1.24	0.84
Italy-Brescia	122	156	0.574	0.564	1.04	0.74–1.46	0.82
Belgium	199	378	0.603	0.554	1.22	0.95–1.56	0.11
Eastern Europe	214	515	0.619	0.575	1.20	0.96–1.51	0.12
Sweden	346	905	0.545	0.521	1.10	0.92–1.31	0.30
Spain	173	1,427	0.546	0.538	1.03	0.83–1.29	0.78
All combined[c]	**4,147**	**34,988**	**0.578**	**0.535**	**1.12**	**1.06–1.18**	$\mathbf{5.7\times10^{-5}}$
Prostate cancer							
Iceland all[a]	2,276	28,890	0.569	0.545	1.10	1.03–1.17	3.75×10^{-3}

Study population	Number		Frequency		OR	95% CI	*P* value
	Cases	Controls	Cases	Controls			
Iceland genotyped[b]	1,754	28,890	0.564	0.545	1.08	1.00–1.16	0.042
The Netherlands	994	1,832	0.576	0.570	1.02	0.92–1.14	0.67
Chicago, US	635	693	0.581	0.568	1.06	0.90–1.23	0.49
Spain	459	1,427	0.559	0.538	1.09	0.94–1.26	0.27
CGEMS	5,109	5,059	0.558	0.543	1.06	1.00–1.11	0.036
All combined[c]	**9,473**	**37,901**	**0.569**	**0.553**	**1.07**	**1.03–1.11**	**3.6×10^{-4}**
Cervical cancer							
Iceland all[a]	369	28,890	0.611	0.545	1.31	1.13–1.51	2.6×10^{-4}
Iceland genotyped[b]	276	28,890	0.611	0.545	1.31	1.03–1.32	1.9×10^{-3}

Shown are the numbers of cases and controls (N), allelic frequencies, the allelic ORs with P values based on the multiplicative model.

[a] Results obtained from combining data from individuals genotyped directly or *in silico*. See Supplementary Material.

[b] Results from directly genotyped individuals only

[c] For the combined study populations, the reported control frequency was the average, unweighted control frequency of the individual populations, while the OR and the P value were estimated using the Mantel-Haenszel model.

Table 3

Association of rs2736098 (A) on 5p15.33 to basal cell carcinoma, cancers of the lung, bladder, prostate and cervix

Study population	Number		Frequency		OR	95% CI	*P* value
	Cases	Controls	Cases	Controls			
Basal cell carcinoma							
Iceland	1,600	3,667	0.327	0.272	1.30	1.18–1.43	1.4×10^{-7}
Eastern Europe	496	491	0.249	0.264	0.93	0.77–1.12	0.45
All combined[a]	**2,096**	**4,158**	**0.288**	**0.268**	**1.21**	**1.11–1.32**	$\mathbf{1.3 \times 10^{-5}}$
Lung cancer							
Iceland	687	3,667	0.306	0.272	1.18	1.03–1.35	0.014
The Netherlands	525	1,740	0.326	0.286	1.21	1.04–1.41	0.014
Spain	365	1,384	0.271	0.229	1.25	1.04–1.51	0.019
All combined[a]	**1,577**	**6,791**	**0.301**	**0.262**	**1.20**	1.10–1.31	$\mathbf{3.2 \times 10^{-5}}$
Bladder cancer							
Iceland	460	3,667	0.283	0.272	1.05	0.90–1.22	0.53
The Netherlands	1,212	1,740	0.308	0.286	1.11	0.90–1.24	0.066
UK	677	486	0.313	0.270	1.24	1.03–1.49	0.023
Italy-Torino	322	375	0.278	0.249	1.16	0.91–1.48	0.23
Italy-Brescia	99	132	0.272	0.227	1.28	0.83–1.97	0.26
Belgium	188	365	0.293	0.271	1.11	0.84–1.46	0.46
Eastern Europe	206	491	0.323	0.264	1.33	1.03–1.71	0.026
Sweden	332	436	0.294	0.227	1.41	1.12–1.77	0.0031
Spain	173	1,384	0.249	0.229	1.11	0.86–1.43	0.42
All combined[a]	**3,669**	**9,076**	**0.290**	**0.255**	**1.16**	**1.08–1.23**	$\mathbf{1.3 \times 10^{-4}}$
Prostate cancer							
Iceland	1,640	3,667	0.290	0.272	1.09	0.99–1.20	0.076
The Netherlands	983	1,740	0.319	0.286	1.17	1.04–1.32	0.0096
Chicago, US	627	679	0.305	0.268	1.20	1.01–1.43	0.039
Spain	449	1,384	0.252	0.229	1.13	0.95–1.34	0.17
All combined[a]	**3,699**	**7,470**	**0.291**	**0.264**	**1.13**	**1.06–1.21**	$\mathbf{1.3 \times 10^{-4}}$
Cervical cancer							

Study population	Number		Frequency		OR	95% CI	*P* value
	Cases	Controls	Cases	Controls			
Iceland	249	3,667	0.295	0.272	1.12	0.91–1.37	0.28

Shown are the corresponding numbers of cases and controls (N), allelic frequencies of variants, the allelic odds-ratio (OR) with P values based on the multiplicative model.

[a]For the combined study populations, the reported control frequency was the average, unweighted control frequency of the individual populations, while the OR and the P value were estimated using the Mantel-Haenszel model.

Table 4

Joint analysis of rs401681(C) and rs2736098(A) of BCC and cancers of the lung, bladder, prostate and cervix

Cancer type	# populations	rs401681(C) adjusted for rs2736098(A)			rs2736098(A) adjusted for rs401681(C)		
		OR	95% CI	*P* value	OR	95% CI	*P* value
Basal cell carcinoma	2	1.20	1.10–1.31	7.8×10^{-5}	1.09	0.99 –1.21	0.091
Lung cancer	3	1.11	1.01–1.21	0.024	1.14	1.03–1.25	0.010
Bladder cancer	9	1.07	1.00–1.16	0.036	1.12	1.04–1.20	0.0034
prostate cancer	4	1.01	0.95–1.08	0.68	1.13	1.05–1.21	0.0015
Cervical cancer	1	1.27	1.03–1.55	0.022	0.97	0.77–1.22	0.80