White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/142007/

Version: Published Version

Article:
Jiang, X, Finucane, HK, Schumacher, FR et al. (2019) Shared heritability and functional enrichment across six solid cancers. Nature Communications, 10. 431. ISSN: 2041-1723

https://doi.org/10.1038/s41467-018-08054-4

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-018-08054-4 OPEN

Shared heritability and functional enrichment across six solid cancers

Xia Jiang et al.[#]

Quantifying the genetic correlation between cancers can provide important insights into the mechanisms driving cancer etiology. Using genome-wide association study summary statistics across six cancer types based on a total of 296,215 cases and 301,319 controls of European ancestry, here we estimate the pair-wise genetic correlations between breast, colorectal, head/neck, lung, ovary and prostate cancer, and between cancers and 38 other diseases. We observed statistically significant genetic correlations between lung and head/neck cancer ($r_g = 0.57$, $p = 4.6 \times 10^{-8}$), breast and ovarian cancer ($r_g = 0.24$, $p = 7 \times 10^{-5}$), breast and lung cancer ($r_g = 0.18$, $p = 1.5 \times 10^{-6}$) and breast and colorectal cancer ($r_g = 0.15$, $p = 1.1 \times 10^{-4}$). We also found that multiple cancers are genetically correlated with non-cancer traits including smoking, psychiatric diseases and metabolic characteristics. Functional enrichment analysis revealed a significant excess contribution of conserved and regulatory regions to cancer heritability. Our comprehensive analysis of cross-cancer heritability suggests that solid tumors arising across tissues share in part a common germline genetic basis.

Correspondence and requests for materials should be addressed to X.J. (email: xiajiang@hsph.harvard.edu) or to P.K. (email: pkraft@hsph.harvard.edu) or to S.Löm. (email: saralind@uw.edu). [#]A full list of authors and their affiliations appears at the end of the paper.

Inherited genetic variation plays an important role in cancer etiology. Large twin studies have demonstrated an excess familial risk for cancer sites including, but not limited to, breast, colorectal, head/neck, lung, ovary, and prostate with heritability estimates ranging between 9% (head/neck) to 57% (prostate)[1–3]. Data from nation-wide and multi-generation registries further show that elevated cancer risks go beyond nuclear families and isolated types, as family history of a specific cancer can increase risk for other cancers[4–6]. Additional evidence for a shared genetic component have been demonstrated by cross-cancer genome-wide association study (GWAS) meta-analyses, which set out to identify genetic variants associated with more than one cancer type. Fehringer et al. studied breast, colorectal, lung, ovarian, and prostate cancer, and identified a novel locus at 1q22 associated with both breast and lung cancer[7]. Kar et al. focused on three hormone-related cancers (breast, ovarian, and prostate), and identified seven novel susceptibility loci shared by at least two cancers[8].

Previous attempts to estimate the genetic correlation across cancers using GWAS data[9–12] have mostly relied on restricted maximum likelihood (REML) implemented in GCTA (genome-wide complex trait analysis)[13] and individual-level genotype data. However, these studies have had limited sample sizes, yielding inconclusive results. Sampson et al. quantified genetic correlations across 13 cancers in European ancestry populations and identified four cancer pairs with nominally significant genetic correlations (bladder–lung, testis–kidney, lymphoma–osteosarcoma, and lymphoma–leukemia)[9]. They did not observe any significant genetic correlations across common solid tumors including cancers of the breast, lung and prostate[9]. REML becomes computationally challenging for large sample sizes and is sensitive to technical artifacts. LD score regression (LDSC)[14,15] overcomes these issues by leveraging the relationship between association statistics and LD patterns across the genome. We recently used cross-trait LDSC to quantify genetic correlations across six cancers based on a subset of the data included here and found moderate correlations between colorectal and pancreatic cancer, as well as between lung and colorectal cancer[16]. However, the average sample size was only 11,210 cases and 13,961 controls per cancer, resulting in imprecise estimates with wide confidence intervals.

In addition to the development of novel analytical methods tailored to genomic data, several high-quality functional annotations have recently been released into the public domain through large-scale efforts. For example, the ENCODE consortium has built a comprehensive and informative parts list of functional elements in the human genome (http://www.nature.com/encode/#/threads), which allows for the analysis of components of SNP-heritability to unravel the functional architecture of complex traits.

Here, we use summary statistics from the largest-to-date European ancestry GWAS of breast, colorectal, head/neck, lung, ovary, and prostate cancer with an average sample size of 49,369 cases and 50,219 controls per cancer, to quantify genetic correlations between cancers and their subtypes. We also use GWAS summary statistics for 38 non-cancer traits (average N = 113,808 per trait), to quantify the genetic correlations between the six cancers and other diseases. Furthermore, we assessed the proportion of cancer heritability attributable to specific functional categories, with the goal of identifying functional elements that are enriched for SNP-heritability.

Our comprehensive analysis identifies statistically significant genetic correlations between lung and head/neck cancer, breast and ovarian cancer, breast and lung cancer, and breast and colorectal cancer. We also find multiple cancers to be genetically correlated with non-cancer traits including smoking, psychiatric diseases, and metabolic traits. Functional enrichment analysis reveals a significant contribution of conserved and regulatory regions to cancer heritability. Our results suggest that solid tumors arising across tissues share in part a common germline genetic basis.

Results

Heritability estimates across cancers. We first estimated cancer-specific heritability causally explained by common SNPs (h_g^2) using LDSC (note that this quantity is slightly different from the h_g^2 as defined in Yang et al.[17] which estimates the heritability due to genotyped and imputed SNPs) (see Methods). Estimates of h_g^2 on the liability scale ranged from 0.03 (ovarian) to 0.25 (prostate) (Supplementary Table 1). After removing genome-wide significant ($p < 5 \times 10^{-8}$) loci, defined as all SNPs within 500 kb of the most significant SNP in a given region (Supplementary Data 1), we observed an ~50% decrease in SNP-heritability for prostate and breast cancer, and ~20% decrease for lung, ovarian, and colorectal cancer, despite the fact that we were only excluding 1% (colorectal cancer) to 5% (breast cancer) of the genome. In contrast, the SNP-heritability for head/neck cancer was not affected by removing genome-wide significant loci (Fig. 1a). For most of the cancers, the GWAS significant loci for that particular cancer explained most of the heritability. For some cancers, however, significant GWAS loci of other cancers also explained a non-trivial part of its heritability. For example, the significant breast cancer GWAS loci explained 10%, 15%, and 22% heritability of colorectal, ovarian and prostate cancer, respectively; the significant colorectal cancer GWAS loci explained 11% heritability of prostate cancer; the significant lung cancer GWAS loci explained 10% heritability of head/neck cancer; and the significant prostate cancer GWAS loci explained 11 and 15% heritability of breast and ovarian cancer, respectively (Supplementary Table 2). Comparing the liability-scale SNP-heritability to corresponding estimates from twin studies suggests that common SNPs can almost entirely explain the classical heritability of head/neck cancer, whereas for other cancers, only 30–40% of heritability can be explained (Fig. 1b).

Genetic correlations between cancers. We then estimated the genetic correlation between cancers using cross-trait LDSC (see Methods). After adjusting for the number of tests ($p < 0.05/15 = 0.003$), we found multiple significant genetic correlations Fig. 1c and Supplementary Table 1), with the strongest result observed for lung and head/neck cancer ($r_g = 0.57$, se = 0.10). In addition, colorectal and lung cancer ($r_g = 0.28$, se = 0.06), breast and ovarian cancer ($r_g = 0.24$, se=0.06), breast and lung cancer ($r_g = 0.18$, se = 0.04), and breast and colorectal cancer ($r_g = 0.15$, se = 0.04) showed statistically significant genetic correlations. We also observed nominally significant genetic correlations ($p < 0.05$) between lung and ovarian cancer ($r_g = 0.16$, se = 0.08), prostate cancer and head/neck ($r_g = 0.15$, se = 0.08), colorectal ($r_g = 0.11$, se = 0.05), and breast cancer ($r_g = 0.07$, se = 0.03) (Fig. 1c). Some cancer pairs showed minimal correlations with estimates close to 0 (ovarian and prostate: $r_g = 0.02$, se = 0.07; lung and prostate: $r_g = -0.03$, se = 0.04; breast and head/neck: $r_g = 0.03$, se = 0.06). We further calculated the cross-cancer genetic correlation based on data after excluding the GWAS significant regions of each cancer. The estimates were mostly consistent with the results calculated based on all SNPs.

We conducted subtype-specific analysis for breast, lung, ovarian, and prostate cancer (Supplementary Table 1). Estrogen receptor positive (ER+) and negative (ER−) breast cancer showed a genetic correlation of 0.60 (se = 0.03), indicating that the genetic contributions to these two subtypes are in part distinct. The genetic correlation between the two common lung

Fig. 1 Estimates of SNP-heritability (h_g^2) and cross-cancer heritability (r_g) for the six cancer types. SNP-heritability and cross-cancer heritability are calculated based on HapMap3 SNPs using LD score regression (LDSC). **a** The solid bar represents overall SNP h_g^2 on the liability scale, calculated based on all HapMap3 SNPs. The dark green bar represents h_g^2 calculated based on non-significant SNPs—the remaining SNPs after excluding genome-wide significant hits ($p < 5 \times 10^{-8}$) ± 500 kb. The black bar with density texture indicates proportion of h_g^2 (as reflected by the percentages displayed on top of each bar) that could be explained by top hits ±500 kb surrounded areas. The orange error bars represent 95% confidence intervals. **b** The solid blue bar represents overall SNP h_g^2 in liability scale (no SNP exclusion), with black error bars indicating 95% confidence intervals. The red short lines correspond to classical estimates of h^2 measured in a twin study of Scandinavian countries (Mucci et al.[2]). **c** Genetic correlations between cancers. Estimates withstood Bonferroni corrections ($p < 0.05/15$) are marked with double asterisk (**), and nominal significant results ($p < 0.05$) are marked with single asterisk (*)

cancer subtypes adenocarcinoma and squamous cell carcinoma was similarly 0.58 (se = 0.10). Further, we observed a significantly larger genetic correlation of lung cancer with ER− ($r_g = 0.29$, se = 0.06) than with ER + breast cancer ($r_g = 0.13$, se = 0.04) ($p_{\text{difference}} = 0.002$). This also held true for lung squamous cell carcinoma, which showed statistically stronger genetic correlation with ER− ($r_g = 0.33$, se = 0.08) than with ER + breast cancer ($r_g = 0.11$, se = 0.05) ($p_{\text{difference}} = 0.0019$). We observed no other statistically significant differential genetic correlations across subtypes (all $p_{\text{difference}} > 0.1$).

We then estimated local genetic correlations between cancers using ρ-HESS, dividing the genome into 1703 regions (see Methods) (Fig. 2 and Supplementary Fig. 1). We found that although the genome-wide genetic correlation between breast and prostate cancer was modest ($r_g = 0.07$), chr10:123M (10q26.13, $p = 1.0 \times 10^{-7}$) and chr9:20–22 M (9p21, $p = 1.0 \times 10^{-6}$), two previously known pleiotropic regions[18], showed significant genetic correlations ($r_g = -0.00098$ and $r_g = 0.00046$). Similarly, although the genome-wide genetic correlation between lung and prostate cancer was negligible ($r_g = -0.03$), two previously identified pleiotropic regions (chr6:30–31 M or 6p21.33, $p = 5.7 \times 10^{-7}$ and chr20:62M or 20q13.33, $p = 2.8 \times 10^{-6}$) exhibited significant local genetic correlations ($r_g = -0.00060$ and $r_g = 0.00067$). Overall, local genetic correlation analysis reinforced shared effects for 44% (31/71) of previously reported pleiotropic cancer regions (Supplementary Data 2). It also identified novel pleiotropic signals. For example, the breast and prostate cancer pleiotropic region at 2q33.1 showed significant local genetic correlation between breast and ovarian cancer ($p = 2.3 \times 10^{-6}$). Additionally, 6p21.32, a region indicated for head/neck and prostate cancer, showed highly significant local genetic correlation for head/neck and lung cancer ($p = 8.6 \times 10^{-8}$).

Genetic correlations between cancer and other traits. Significant genetic correlations ($p < 0.05/228 = 0.0002$) between the six cancers and 38 non-cancer traits reflected several known associations (Fig. 3 and Supplementary Data 3). We observed a strong genetic correlation between smoking and lung cancer ($r_g = 0.56$, se = 0.06), and similarly for head/neck cancer ($r_g = 0.47$, se = 0.08), both cancers having smoking as its primary risk factor[19,20]. Educational attainment was negatively genetically correlated with colorectal ($r_g = -0.17$, se = 0.04), head/neck ($r_g = -0.42$, se = 0.07), and lung cancer ($r_g = -0.39$, se=0.04) (all $p < 5 \times 10^{-6}$). Body mass index (BMI) showed a positive genetic correlation with colorectal cancer ($r_g = 0.15$, se = 0.03) and also suggestive but weak negative correlations with prostate ($r_g = -0.07$, se = 0.03) and breast cancer ($r_g = -0.06$, se = 0.03). Lung cancer showed a negative genetic correlation with lung function ($r_g = -0.15$, se = 0.04) and age at natural menopause ($r_g = -0.25$, se = 0.05), and moderate positive genetic correlations with depressive symptoms ($r_g = 0.25$, se=0.06) and waist-to-hip ratio ($r_g = 0.16$, se = 0.04). Breast cancer showed a positive genetic correlation with schizophrenia ($r_g = 0.14$, se = 0.03).

We did not find evidence of genetic correlations between cancer and several previously suggested risk factors[21–23] including cardiovascular traits (coronary artery disease, hypertension, and blood pressure) or sleep characteristics (chronotype, duration, and insomnia). Further, we did not observe genetic correlations between cancer and circulating lipids (HDL, LDL, and triglycerides) or type 2 diabetes-related traits except a significant negative correlation between HDL and lung cancer ($r_g = -0.14$, se = 0.04). We observed no significant genetic correlation between breast cancer and age at menarche ($r_g = -0.03$, se = 0.03) or age at natural menopause ($r_g = -0.01$, se = 0.03). We also did not observe notable genetic correlations between cancer and autoimmune inflammatory diseases or height.

Fig. 2 Local genetic correlation between breast, lung and prostate cancer. The region-specific *p*-values for the local genetic covariance for breast and prostate cancer are shown in **a**, and for lung and prostate cancer in **b**. Each dot presents a specific genomic region. In the QQ plots, red color indicates significance after multiple corrections ($p < 0.05/1703$ regions compared), and blue color indicates nominal significance ($p < 0.05/15$ pairs of cancers compared). Manhattan-style plots showing the estimates of local genetic covariance for breast and prostate cancer (**c**), and for lung and prostate cancer (**d**). Although breast and prostate cancer only show modest genome-wide genetic correlation, two loci exhibit significant local genetic covariance. Similarly, albeit the negligible overall genetic correlation for lung and prostate cancer, three loci present significant local genetic covariance. In the Manhattan plots, red color indicates even number chromosomes and blue color indicates odd number chromosomes

Fig. 3 Cross-trait genetic correlation (r_g) analysis between cancers and non-cancer traits. The traits were divided into four categories: **a** Common phenotypes, **b** Metabolic or cardiovascular related traits, **c** Psychiatric traits, **d** Autoimmune inflammatory diseases. Pair-wise genetic correlations withstood Bonferroni corrections (228 tests) are marked with double asterisk (**), with estimates of correlation shown in the cells. Pair-wise genetic correlations with significance at $p < 0.01$ are marked with a single asterisk (*). The color of cells represents the magnitude of correlation

Subtype analysis revealed that smoking and educational attainment showed genetic correlations with all lung cancer subtypes (Supplementary Data 3). Educational attainment, forced vital capacity and depressive symptoms showed genetic correlations with ER− but not ER + breast cancer, whilst the observed genetic correlation between schizophrenia and breast cancer was limited to ER + disease, and the genetic correlation between depressive symptoms and lung cancer was observed only for lung squamous cell carcinoma.

We further assessed the support for mediated or pleiotropic causal models for non-cancer traits and cancer using the correlation between trait-specific effect sizes of genome-wide significant SNPs for pairs of phenotypes. We detected four putative directional genetic correlations (defined as $p < 0.05$ from a likelihood ratio (LR) comparing the best non-causal model to the best causal model) (Fig. 4), where SNPs associated with the non-cancer trait showed correlated effect estimates with cancer but the reverse was not true (circulating HDL concentrations and breast cancer, $LR_{non\text{-}causal\ vs.\ causal} = 0.04$, schizophrenia and breast cancer, $LR_{non\text{-}causal\ vs.\ causal} = 0.003$, age at natural menopause and breast cancer, $LR_{non\text{-}causal\ vs.\ causal} = 0.04$, and lupus and prostate cancer, $LR_{non\text{-}causal\ vs.\ causal} = 0.0006$).

Functional enrichment analysis of cancer heritability. Finally, we partitioned SNP-heritability of each cancer by using 24 genomic functional annotations (the baseline-LD model described in Gazal et al.[24]) and 220 cell-type-specific histone mark annotations (the cell-type-specific model described in Finucane et al.[14]). Meta-analysis across the six cancers revealed statistically significant enrichments for multiple functional categories. We observed the highest enrichment for conserved regions (Table 1, Supplementary Table 3) which overlapped with only 2.6% of SNPs but explained 25% of cancer SNP-heritability

Fig. 4 Putative directional relationships between cancers and traits. For each cancer–trait pair identified as candidates to be related in a causal manner, the plots show trait-specific effect sizes (beta coefficients) of the included genetic variants. Gray lines represent the relevant standard errors. **a** HDL and breast cancer. Trait-specific effect sizes for HDL and breast cancer are shown for SNPs associated with HDL levels (left) and breast cancer (right). **b** Schizophrenia and breast cancer. Trait-specific effect sizes for schizophrenia and breast cancer are shown for SNPs associated with schizophrenia (left) and breast cancer (right). **c** Age at natural menopause and breast cancer. Trait-specific effect sizes for age at natural menopause and breast cancer are shown for SNPs associated with age at natural menopause (left) and breast cancer (right). **d** Lupus and prostate cancer. Trait-specific effect sizes for lupus and prostate cancer are shown for SNPs associated with lupus (left) and prostate cancer (right)

(9.8-fold enrichment, $p = 2.3 \times 10^{-5}$). Transcription factor binding sites showed the second highest enrichment (4.0-fold, 13% of SNPs explaining 40% of SNP-heritability, $p = 1.4 \times 10^{-7}$). Further, super-enhancers (groups of putative enhancers in close genomic proximity with unusually high levels of mediator binding) showed a significant 2.6-fold enrichment ($p = 2.0 \times 10^{-20}$). Additional enhancers, including regular enhancers (3.2-fold), weak enhancers (3.1-fold) and FANTOM5 enhancers (3.1-fold), presented similar enrichments but were not statistically significant. In addition, multiple histone modifications of epigenetic markers H3K9ac, H3K4me3, and H3K27ac, were all significantly enriched for cancer heritability. Repressed regions exhibited depletion (0.34-fold, $p = 1.2 \times 10^{-6}$). Enrichment analysis of functional categories for each cancer and cancer subtype are shown in Fig. 5 and Supplementary Table 4.

Overall, cell-type-specific analysis of histone marks identified significant enrichments specific to individual cancers (Supplementary Fig. 2). For breast cancer, 3 out of 8 statistically significant tissues were adipose nuclei (H3K4me1, H3K9ac) and breast myoepithelial (H3K4me1) cells. For colorectal cancer, 15 out of the 18 statistically significant enrichments were observed in either colon or rectal tissues (colon/rectal mucosa, duodenum mucosa, small/large intestine, and colon smooth muscle). We observed no significant enrichments for head/neck, lung, and ovarian cancer, but we noted that for both lung (9 out of 10) and ovarian cancer (6 out of 10), the most enriched cell types were immune cells; while in head/neck cancer, 6 out of 10 most highly enriched cell types belonged to CNS (Supplementary Fig. 3, Supplementary Data 4). Cell-type-specific analysis for cancer subtypes are shown in Supplementary Data 5. Comparing

Table 1 Significant enrichment estimates of genomic functional categories, meta-analyzed across six cancer sites

Category	Enrichment (95% CI)	*P*-value
Conserved region	9.78 (5.72–13.84)	2.28×10^{-5}
TFBS	4.04 (2.91–5.17)	1.43×10^{-7}
H3K9ac	3.41 (2.14–4.69)	2.04×10^{-4}
H3K4me3	3.23 (2.47–4.00)	8.91×10^{-9}
Super Enhancer	2.56 (2.23–2.89)	1.99×10^{-20}
H3K27ac (PGC)	2.36 (1.91–2.80)	2.12×10^{-9}
H3K27ac (Hnisz)	1.90 (1.65–2.15)	1.86×10^{-12}
H3K4me1	1.84 (1.56–2.12)	2.57×10^{-9}
Repressed region	0.34 (0.07–0.61)	1.15×10^{-6}

The meta-analysis was performed based on the enrichment estimates and standard errors calculated using LD score regression in each individual cancer type. *P*-values were significant after Bonferroni correction ($P < 0.05/24$)
TFBS transcription factor binding sites

cell-type-specific enrichment for cancers to the additional 38 non-cancer traits revealed notably differential clustering patterns (Supplementary Fig. 4). Breast, colorectal, and prostate cancer showed enrichment mostly for adipose and epithelial tissues, in contrast to autoimmune diseases (enriched for immune/hematopoietic cells) or psychiatric disorders (enriched for brain tissues).

Discussion

We performed a comprehensive analysis quantifying the heritability and genetic correlation of six cancers, leveraging summary statistics from the largest cancer GWAS conducted to date. Our study demonstrates shared genetic components across multiple cancer types. These results contrast with a prior study conducted by Sampson et al. which reported an overall negligible genetic correlation among common solid tumors[9]. Our results are, however, in line with a recent study,[16] which analyzed a subset of the data included here, and identified a significant genetic correlation between lung and colorectal cancer.

Our data support, and for the first time quantify, the strong genetic correlation ($r_g = 0.57$) between lung and head/neck cancer, two cancers linked to tobacco use[20,25]. We also for the first time observed a significant genetic correlation between breast and ovarian cancer ($r_g = 0.24$), two cancers that are known to share rare genetic factors including *BRCA1/2* mutations, and environmental exposures associated with endogenous and exogenous hormone exposures[26]. Prostate cancer is also considered as hormone-dependent and associated with *BRCA1/2* mutations, but interestingly, we only observed a nominally significant and modest ($r_g = 0.07$) genetic correlation between breast and prostate cancer, while ovarian and prostate cancer showed no genetic correlation ($r_g = 0.02$, se $= 0.07$).

Our large sample sizes allowed us to conduct well-powered analyses for cancer subtypes. While head/neck cancer showed negligible genetic correlation with overall ($r_g = 0.03$, se $= 0.06$) and ER + breast cancer ($r_g = -0.02$, se $= 0.07$), it showed a stronger genetic correlation with ER− breast cancer ($r_g = 0.21$, se $= 0.09$). Similarly, lung cancer showed a statistically more pronounced genetic correlation with ER− ($r_g = 0.29$, se $= 0.06$) than ER + breast cancer ($r_g = 0.13$, se $= 0.04$). A recent pooled analysis of smoking and breast cancer risk demonstrated a smoking-related increased risk for ER + but not for ER− breast cancer[27], and thus it is unlikely that the stronger genetic correlation between ER− subtype and lung and head/neck cancer is due to smoking behavior. Perhaps surprisingly, despite literature suggesting substantial similarities between ER− breast cancer and serous ovarian cancer in particular[28], we did not observe statistically significant different genetic correlations between ER− or ER + breast cancer and serous ovarian cancer ($r_g = 0.17$, se $= 0.08$ *vs.* $r_g = 0.11$, se $= 0.06$). This suggests that rare high penetrance variants may play a more important role in driving the similarities behind ER− breast cancer and serous ovarian cancer than common genetic variation.

Heritability analysis confirms that common cancers have a polygenic component that involves a large number of variants. Although susceptibility variants identified at genome-wide significance explain an appreciable fraction of the heritability for some cancers, we estimate that the majority of the polygenic effect is attributable to other, yet undiscovered variants, presumably with effects that are too weak to have been identified with current sample sizes. We found the genetic component that could be attributed to genome-wide significant loci varied greatly from ~0% for head/neck cancer to ~50% for breast and prostate cancer. These results reflect in part the strong correlation between number of GWAS-identified loci and sample size, as we had more than twice as many breast and prostate cancer samples compared to the other cancers. One corollary is that larger GWAS are likely to identify new susceptibility loci that could help our understanding of disease development, improve prediction power of genetic risk scores and hence contribute to screening and personalized risk prediction[29].

Among the genetic correlations between cancer and non-cancer traits, we observed positive correlations for psychiatric disorders (depressive symptoms, schizophrenia) with lung and breast cancer, where findings from epidemiological studies have been suggestive but inconclusive. It has been proposed that the linkage between psychiatric traits and cancers are more likely to be mediated through cancer-associated risk phenotypes such as smoking, excessive alcohol consumption in depressed populations[30], and reduced fertility patterns (e.g., nulliparous) in psychiatric populations[31]. Detailed analyses considering confounding traits like reproductive history and smoking are needed to make inference about the mechanisms involved. GWAS have identified pleiotropic regions influencing both lung cancer and nicotine dependence, such as 15q25.1[32,33]. In line with those results, we identified a strong genetic correlation between smoking and both lung ($r_g = 0.56$) and head/neck cancer ($r_g = 0.47$). It remains unclear whether this genetic correlation is completely explained by the direct influence of smoking or if the shared genetic component affects the traits through separate pathways. Interestingly, a genetic correlation ($r_g = 0.35$, se $= 0.14$) between lung and bladder cancer, another smoking-associated cancer, has been identified previously[9]. Due to the small numbers of GWAS-identified smoking-associated SNPs, we were unable to assess a directional correlation between smoking and cancer, but we expect such analyses to become feasible as additional smoking-related SNPs are identified. We found modest positive, yet significant genetic correlations between adiposity-related measures (as reflected by waist-to-hip ratio, circulating HDL levels and BMI) and both colorectal and lung cancer, but negative genetic correlations between BMI and prostate and breast cancer, consistent with previous reported findings[34] and reinforce the complex dynamics between obesity and cancer where multiple factors including age, smoking, endogenous hormones and reproductive status play a role.

We did not observe genetic correlations between breast cancer and age at menarche or age at natural menopause. These null observations were largely driven by ER + breast cancer (ER +: $r_g = 0.006$, se $= 0.03$ vs. ER−: $r_g = -0.09$, se $= 0.04$ for age at menarche. ER +: $r_g = 0.0005$, se $= 0.04$ vs. ER−: $r_g = -0.10$, se $= 0.05$ for age at natural menopause), and were unexpected given that both factors play pivotal roles in breast cancer etiology[35] and previous Mendelian randomization (MR) analyses have identified a link[36,37]. An important difference between genetic

Fig. 5 Enrichment *p*-values of 24 non-cell-type-specific functional categories over six cancer types. The *x*-axis represents each of the 24 functional categories, *y*-axis represents log-transformed *p*-values of enrichment. Annotations with statistical significance after Bonferroni corrections ($p < 0.05/24$) were plotted in orange, otherwise blue. The horizontal gray dash line indicates *p*-threshold of 0.05; horizontal red dash line indicates *p*-threshold of 0.05/24. From top to bottom are six panels representing six cancers: breast cancer, colorectal cancer, head/neck cancer, lung cancer, ovarian cancer, and prostate cancer. TSS transcription start site, UTR untranslated region, TFBS transcription factor binding sites, DHS DNase I hypersensitive sites, DGF digital genomic foot printing, CTCF CCCTC-binding factor

correlation and MR analyses is that the latter only considers genome-wide significant SNPs while the former incorporates the entire genome. It is possible that a relatively small overlap in strongly associated SNPs can result in significant MR results despite low evidence of an overall genetic correlation. Indeed, the directional genetic correlations we observed for age at natural menopause, schizophrenia, and HDL with breast cancer, and for lupus with prostate cancer, highlight again that although an overall genetic correlation may be negligible, there can still be genetic links between traits. It is important to note that we cannot rule out unmeasured confounding, including the possibility that these genetic variants affect an intermediate phenotype that is pleiotropic for both target traits. Given the observational nature of our data, these putative causal directions should be interpreted with caution.

Pan-cancer tumor-based studies have demonstrated that different cancers are sometimes driven by similar somatic functional events such as specific copy number abnormalities and mutations[38,39]. Our enrichment results of germline genetic across functional annotation data shed new light on the biological mechanisms leading to cancer development. The more pronounced enrichment identified for conserved regions compared with coding regions provides evidence for the biological importance of the former, which has been shown to be true for multiple traits[14,40]. Even though the biochemical function of many conserved regions remains uncharacterized, transcribed ultra-conserved regions have been found to be frequently located at fragile sites. Compared to normal cells, cancer cells have a unique spectrum of transcribed ultra-conservative regions, suggesting that variation in expression of these regions are involved in the malignant process[41,42]. These results bridge the link between germline and somatic genetics in cancer development, which was also observed in a recent breast cancer GWAS that has demonstrated a strong overlap between target genes for GWAS hits and somatic driver genes in breast tumors[43]. We also found a four-fold enrichment for transcription factor binding sites and a three-fold enrichment for super-enhancers, consistent with prior observations that breast cancer GWAS loci fall in enhancer regions involved in distal regulation of target genes[43]. Cell-type-specific analysis of histone marks demonstrated the importance of tissue specificity, primarily for colorectal and breast cancer. Further, our results suggest that immune cells are important for ovarian and lung cancer whilst CNS is important to head/neck cancer. Unfortunately, we did not have data on prostate-specific tissues, but we note that tissue-specific enrichment of prostate cancer heritability for epigenetic markers has been observed previously[10]. We note that generation of rich functional annotation is ongoing and we expect to include additional tissue-specific functional elements in our future work.

Our study has several strengths. We were able to robustly quantify pair-wise genetic correlations between multiple cancers using the largest available cancer GWAS, comprising almost 600,000 samples across six major cancers and subtypes. We were also able to systematically assess the genetic correlations between cancer and 38 non-cancer traits. Notwithstanding the large sample sizes, several limitations need to be acknowledged. We did not have the sample sizes required to assess relevant cancer subgroups including oropharyngeal cancer, clear cell, mucinous and endometrioid ovarian cancer, or lung cancer among never smokers (each with ~2000 cases). In addition, we did not have access to GWAS summary statistics for pre- vs. post-menopausal breast cancer. We were not able to consider all cancer risk factors when selecting non-cancer traits, since some of the well-established risk factors such as infection were either not available, showed no evidence of heritability or were not based on adequate sample sizes for robust analyses. SNP-heritability varies with minor allele frequency, linkage disequilibrium, and genotype certainty; we note that approaches to estimate heritability leveraging GWAS data are constantly evolving. We also note that estimate variability needs to be taken into account when comparing the SNP-heritability with the classical twin-heritability, in particular for cancers with small sample sizes such as head/neck cancer (SNP-heritability varied between 5–14% and twin-heritability varied between 0–60%, although both point estimates were 9%). Further, our data were based on GWAS meta-analysis from multiple individual GWAS across European ancestry populations from Europe, Australia and the US. Intra-European ancestry differences are likely to be a source of bias. However, since we limited our analysis to SNPs with MAF > 1% and HapMap3 SNPs (which have proven to be well imputed across European ancestry populations), we believe that any population structure across cancers will have minimal effect on our results. Finally, as more non-European and multi-ethnic GWAS data become available, it is important to examine trans-ethnic genetic correlation in cancer.

In conclusion, results from our comprehensive analysis of heritability and genetic correlations across six cancer types indicate that solid tumors arising from different tissues share common germline genetic influences. Our results also demonstrate evidence for common genetic risk sharing between cancers and smoking, psychiatric, and metabolic traits. In addition, functional components of the genome, particularly conserved and regulatory regions, are significant contributors to cancer heritability across multiple cancer types. Our results provide a basis and direction for future cross-cancer studies aiming to further explore the biological mechanisms underlying cancer development.

Methods

Studies and quality control. We used summary statistics from six cancer GWASs based on a total of 597,534 participants of European ancestry. Cancer-specific sample sizes were: breast cancer: 122,977 cases/105,974 controls; colorectal cancer: 36,948/30,864; head/neck cancer (oral and oropharyngeal cancers): 5452/5984; lung cancer: 29,266/56,450; ovarian cancer: 22,406/40,941; prostate cancer: 79,166/61,106. These data were generated through the joint efforts of multiple consortia. Details on study characteristics and subjects contributed to each cancer-specific GWAS summary dataset have been described elsewhere[43–49]. SNPs were imputed to the 1000 Genomes Project reference panel (1KGP) using a standardized protocol for all cancer types[18]. We included autosomal SNPs with a minor allele frequency (MAF) larger than 1% and present in HapMap3 (N_{SNPs} = ~1 million) because those SNPs are usually well imputed in most studies (note that excluding sex chromosomes could reduce the overall heritability estimates). A brief overview of the quality control in each cancer dataset are presented in Supplementary Table 5. For some of the cancers, we further obtained summary statistics data on subtypes (ER + and ER− breast cancer; lung adenocarcinoma, and squamous cell carcinoma; serous invasive ovarian cancer and advanced stage prostate cancer, defined as metastatic disease or Gleason score ≥ 8 or PSA > 100 or prostate cancer death). Sample sizes and more details shown in Supplementary Table 1.

We additionally assembled European ancestry GWAS summary statistics from 38 traits, which spanned a wide range of phenotypes including anthropometric (e.g., height and body mass index (BMI)), psychiatric disorder (e.g., depressive symptoms and schizophrenia), and autoimmune disease (e.g., rheumatoid arthritis and celiac disease) (Supplementary Table 6). We calculated trait-specific SNP-heritability and restricted our analysis to traits with a heritable component (Supplementary Table 7)[14]. We removed the major histocompatibility complex (MHC) region from all analysis because of its unusual LD and genetic architecture.

Estimation of SNP-heritability and genetic correlation. We estimated the SNP-heritability due to genotyped and imputed SNPs (h_g^2, the proportion of phenotypic variance causally explained by common SNPs) of each cancer using LDSC[15]. Briefly, this method is based on the relationship between LD score and χ^2-statistics:

$$E\left[\chi_j^2\right] \approx \frac{N_j h_g^2}{M} l_j + 1 \quad (1)$$

where $E\left[\chi_j^2\right]$ denotes the expected χ^2-statistics for the association between the outcome and SNP j, N_j is the study sample size available for SNP j, M is the total numbers of variants and l_j denotes the LD score of SNP j defined as $l_j = \sum_k r^2(j,k)$ (k denotes other variants within the LD region). Note that the quantity estimated by LDSC is the causal heritability of common SNPs, which is different from the

SNP-heritability as defined in Yang et al.[17]. To estimate h_g^2 attributable to undiscovered loci, we identified SNPs that were associated with a given cancer at genome-wide significance ($p < 5 \times 10^{-8}$) and removed all SNPs within (+/− 500,000 base-pairs of those loci prior to calculation (number of regions (+/− 500 kb) for each cancer that reach the 5×10^{-8} threshold and measures of effect size are shown in Supplementary Data 1). We also converted the SNP-heritability from observed scale to liability scale by incorporating sample prevalence (P) and population prevalence (F) of each cancer:

$$h^2_{\text{liability}} = h^2_{\text{observed}} \frac{F(1-F)}{\phi(\Phi^{-1}(F))^2} \frac{F(1-F)}{P(1-P)} \quad (2)$$

We subsequently calculated the genome-wide genetic correlations (r_g) between different cancers, and between cancers and non-cancer traits, using an algorithm[14]:

$$E\left[\beta_j \gamma_j\right] = \frac{\sqrt{N_1 N_2} r_g}{M} l_j + \frac{N_s r}{\sqrt{N_1 N_2}} \quad (3)$$

where β_j and γ_j are the effect sizes of SNP j on traits 1 and 2, r_g is the genetic covariance, M is number of SNPs, N_1 and N_2 are the sample sizes for trait 1 and 2, N_s is the number of overlapping samples, r is the phenotypic correlation in overlapping samples and l_j is the LD score defined as above. For genetic correlation between 6 cancers, the significance level is $0.05/15 = 0.003$; for genetic correlation between 6 cancers and 38 traits, the significance level is $0.05/(6 \times 38) = 0.0002$.

Overall genetic correlations as estimated by LDSC are based on aggregated information across all variants in the genome. It is possible that even though two traits show negligible overall genetic correlation, there are specific regions in the genome that contribute to both traits. We therefore examined local genetic correlations between cancer pairs using ρ-HESS[50], an algorithm which partitions the whole genome into 1703 regions based on LD-pattern of European populations and quantifies correlation between pairs of traits due to genetic variation restricted to these genomic regions. Local genetic correlation was considered statistically significant if $p < 0.05/1{,}703 = 2.9 \times 10^{-5}$. In particular, we assessed the local genetic correlations for previously reported pleiotropic regions[18,51] known to harbor SNPs affecting multiple cancers.

Directional genetic correlation analysis. In addition to the genetic correlation analysis, which reflects overall genetic overlaps, we also attempted to identify directions of potential genetic correlations using a subset of SNPs as proposed by Pickrell et al.[52]. The method adopts the following assumption: if a trait X influences trait Y, then SNPs influencing X should also influence Y, and the SNP-specific effect sizes for the two traits should be correlated. Further, since Y does not influence X, but could be influenced by mechanisms independent of X, genetic variants that influence Y do not necessarily influence X. Based on this assumption, the method proposes two causal models and two non-causal models; and calculates the relative likelihood ratio (LR) of the best non-causal model compared to the best causal model. We determined significant SNPs for each given cancer or trait in two independent ways, (1) LD pruned SNPs: we selected genome-wide significant ($p < 5 \times 10^{-8}$) SNPs and pruned on LD-pattern in the European populations in Phase1 of 1KGP; (2) posterior probability of association (PPA) SNPs: we used a method implemented in fgwas[53], which splits the genome into independent blocks based on LD patterns in 1KGP and estimates the prior probability that any block contains an association. The model outputs posterior probability that the region contains a variant that influences the trait. We selected the lead SNP from each of the regions with a PPA of at least 0.9. We scanned through all pairs of cancers and traits to identify directional correlations. Only pairs of traits with evidence of directional correlations (LR comparing the best non-causal model over the best causal model < 0.05) and without evidence of heteroscedasticity (pleiotropic effects)[54] were reported as relatively more likely to exhibit mediated causation.

Functional partitioning of SNP-heritability. To assess the importance of specific functional annotations in SNP-heritability across cancers, we partitioned the cancer-specific heritability using stratified-LDSC[14]. This method partitions SNPs into functional categories and calculates category-specific enrichments based on the assumption that a category of SNPs is enriched for heritability if SNPs with high LD to that category have higher χ^2 statistics than SNPs with low LD to that category. The analysis was performed using two models[14,24].

1. A full baseline-LD model including 24 publicly available annotations that are not specific to any cell type. When performing this model, we adjusted for MAF via MAF-stratified quantile-normalized LD score, and other LD-related annotations such as predicted allele age and recombination rate, as implemented by Gazal et al.[24]. Briefly, the 24 annotations included coding, 3′UTR and 5′UTR, promoter and intronic regions, obtained from UCSC Genome Browser and post-processed by Gusev et al.[55]; the histone marks mono-methylation (H3K4me1) and tri-methylation of histone H3 at lysine 4 (H3K4me3), acetylation of histone H3 at lysine 9 (H3K9ac) processed by Trynka et al.[56–58] and two versions of acetylation of histone H3 at lysine 27 (H3K27ac, one version processed by Hnisz et al.[59], another used by the Psychiatric Genomics Consortium (PGC)[60]); open chromatin, as reflected by DNase I hypersensitivity sites (DHSs and fetal DHSs)[55], obtained as a combination of ENCODE and Roadmap Epigenomics data, processed by Trynka et al.[58]; combined chromHMM and Segway predictions obtained from Hoffman et al.[61], which make use of many annotations to produce a single partition of the genome into seven underlying chromatin states (The CCCTC-binding factor (CTCF), promoter-flanking, transcribed, transcription start site (TSS), strong enhancer, weak enhancer categories, and the repressed category); regions that are conserved in mammals, obtained from Lindblad-Toh et al.[40] and post-processed by Ward and Kellis[62]; super-enhancers, which are large clusters of highly active enhancers, obtained from Hnisz et al.[59]; FANTOM5 enhancers with balanced bi-directional capped transcripts identified using cap analysis of gene expression in the FANTOM5 panel of samples, obtained from Andersson et al.[63]; digital genomic footprint (DGF) and transcription factor binding site (TFBS) annotations obtained from ENCODE and post-processed by Gusev et al.[55]
2. In addition to the baseline-LD model, we also performed analyses using 220 cell-type-specific annotations for the four histone marks H3K4me1, H3K4me3, H3K9ac, and H3K27ac. Each cell-type-specific annotation corresponds to a histone mark in a single cell type (for example, H3K27ac in CD19 immune cells), and there were 220 such annotations in total. We further divided these 220 cell-type-specific annotations into 10 groups (adrenal and pancreas, central nervous system (CNS), cardiovascular, connective and bone, gastrointestinal, immune and hematopoietic, kidney, liver, skeletal muscle, and other) by taking a union of the cell-type-specific annotations within each group (for example, SNPs with any of the four histone modifications in any hematopoietic and immune cells were considered as one big category). When generating the cell-type-specific models, we added annotations individually to the baseline model, creating 220 separate models.

We performed a random-effects meta-analysis of the proportion of heritability over six cancers for each functional category. We set significance thresholds for individual annotations at $p < 0.05/24$ for baseline model and at $p < 0.05/220$ for cell-type-specific annotation.

Data availability

The datasets generated during and/or analyzed during the current study are available from the authors on request. Breast cancer: summary results for all variants are available at http://bcac.ccge.medschl.cam.ac.uk/. Requests for further data should be made through the Data Access Coordination Committee (http://bcac.ccge.medschl.cam.ac.uk/). Ovarian cancer: summary results are available from the Ovarian Cancer Association Consortium (OCAC) (http://ocac.ccge.medschl.cam.ac.uk/). Requests for further data can be made to the Data Access Coordination Committee (http://cimba.ccge.medschl.cam.ac.uk/). Prostate cancer: summary results are publicly available at the PRACTICAL website (http://practical.icr.ac.uk/blog/). Lung cancer: genotype data for lung cancer are available at the database of Genotypes and Phenotypes (dbGaP) under accession phs001273.v1.p1. Readers interested in obtaining a copy of the original data can do so by completing the proposal request form at http://oncoarray.dartmouth.edu/. Head/neck cancer: genotype data for the oral and pharyngeal OncoArray study have been deposited at the database of Genotypes and Phenotypes (dbGaP) under accession phs001202.v1.p1. Colorectal cancer: genotype data have been deposited at the database of Genotypes and Phenotypes (dbGaP) under accession number phs001415.v1.p1 and phs001078.v1.p1.

Received: 30 July 2018 Accepted: 10 December 2018
Published online: 25 January 2019

References

1. Lichtenstein, P. et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **343**, 78–85 (2000).
2. Mucci, L. A. et al. Familial risk and heritability of cancer among twins in nordic countries. *JAMA* **315**, 68 (2016).
3. Polderman, T. J. C. et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
4. Amundadottir, L. T. et al. Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. *PLoS Med.* **1**, e65 (2004).
5. Yu, H., Frank, C., Sundquist, J., Hemminki, A. & Hemminki, K. Common cancers share familial susceptibility: implications for cancer genetics and counselling. *J. Med. Genet.* **54**, 248–253 (2017).
6. Frank, C., Sundquist, J., Yu, H., Hemminki, A. & Hemminki, K. Concordant and discordant familial cancer: familial risks, proportions and population impact. *Int. J. Cancer* **140**, 1510–1516 (2017).

7. Fehringer, G. et al. Cross-cancer genome-wide analysis of lung, ovary, breast, prostate, and colorectal cancer reveals novel pleiotropic associations. *Cancer Res.* **76**, 5103–5114 (2016).
8. Kar, S. P. et al. Genome-wide meta-analyses of breast, ovarian, and prostate cancer association studies identify multiple new susceptibility loci shared by at least two cancer types. *Cancer Discov.* **6**, 1052–1067 (2016).
9. Sampson, J. N. et al. Analysis of heritability and shared heritability based on genome-wide association studies for thirteen cancer types. *J. Natl. Cancer Inst.* **107**, djv279 (2015).
10. Gusev, A. et al. Atlas of prostate cancer heritability in European and African-American men pinpoints tissue-specific regulation. *Nat. Commun.* **7**, 10979 (2016).
11. Jiao, S. et al. Estimating the heritability of colorectal cancer. *Hum. Mol. Genet.* **23**, 3898–3905 (2014).
12. Lu, Y. et al. Most common 'sporadic' cancers have a significant germline genetic component. *Hum. Mol. Genet.* **23**, 6112–6118 (2014).
13. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
14. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
15. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
16. Lindström, S. et al. Quantifying the genetic correlation between multiple cancer types. *Cancer Epidemiol. Biomark. Prev.* **26**, 1427–1435 (2017).
17. Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
18. Amos, C. I. et al. The OncoArray Consortium: a network for understanding the genetic architecture of common cancers. *Cancer Epidemiol. Biomark. Prev.* **26**, 126–135 (2017).
19. SMOKING and health. Joint report of the Study Group on Smoking and Health. *Science* **125**, 1129–1133 (1957).
20. Shaw, R. & Beasley, N. Aetiology and risk factors for head and neck cancer: United Kingdom National Multidisciplinary Guidelines. *J. Laryngol. Otol.* **130**, S9–S12 (2016).
21. Koene, R. J., Prizment, A. E., Blaes, A. & Konety, S. H. Shared risk factors in cardiovascular disease and cancer. *Circulation* **133**, 1104–1114 (2016).
22. Thompson, C. L. et al. Short duration of sleep increases risk of colorectal adenoma. *Cancer* **117**, 841–847 (2011).
23. Sigurdardottir, L. G. et al. Sleep disruption among older men and risk of prostate cancer. *Cancer Epidemiol. Prev. Biomark.* **22**, 872–879 (2013).
24. Gazal, S. et al. Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat. Genet.* **49**, 1421–1427 (2017).
25. Field, R. W. & Withers, B. L. Occupational and environmental causes of lung cancer. *Clin. Chest Med.* **33**, 681–703 (2012).
26. Hulka, B. S. Epidemiologic analysis of breast and gynecologic cancers. *Prog. Clin. Biol. Res.* **396**, 17–29 (1997).
27. Gaudet, M. M. et al. Pooled analysis of active cigarette smoking and invasive breast cancer risk in 14 cohort studies. *Int. J. Epidemiol.* **46**, 881–893 (2017).
28. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
29. Maas, P. et al. Breast cancer risk from modifiable and nonmodifiable risk factors among white women in the United States. *JAMA Oncol* **2**, 1295–1302 (2016).
30. Nakaya, N. et al. Personality traits and cancer risk and survival based on finnish and swedish registry data. *Am. J. Epidemiol.* **172**, 377–385 (2010).
31. Oksbjerg Dalton, S., Munk Laursen, T., Mellemkjaer, L., Johansen, C. & Mortensen, P. B. Schizophrenia and the risk for breast cancer. *Schizophr. Res.* **62**, 89–92 (2003).
32. Hung, R. J. et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* **452**, 633–637 (2008).
33. Amos, C. I. et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat. Genet.* **40**, 616–622 (2008).
34. Gao, C. et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int. J. Epidemiol.* **45**, 896–908 (2016).
35. Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol.* **13**, 1141–1151 (2012).
36. Day, F. R. et al. Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nat. Genet.* **47**, 1294–1303 (2015).
37. Day, F. R. et al. Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat. Genet.* **49**, 834–841 (2017).
38. Zack, T. I. et al. Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **45**, 1134–1140 (2013).
39. Ciriello, G. et al. Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* **45**, 1127–1133 (2013).
40. Lindblad-Toh, K. et al. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476–482 (2011).
41. Calin, G. A. et al. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell* **12**, 215–229 (2007).
42. Peng, J. C., Shen, J. & Ran, Z. H. Transcribed ultraconserved region in human cancers. *RNA Biol.* **10**, 1771–1777 (2013).
43. Michailidou, K. et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* **551**, 92–94 (2017).
44. Milne, R. L. et al. Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nat. Genet.* **49**, 1767–1778 (2017).
45. McKay, J. D. et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat. Genet.* **49**, 1126–1132 (2017).
46. Schmit, S. L. et al. Novel Common Genetic Susceptibility Loci for Colorectal Cancer. *J. Natl. Cancer Inst.* **111**, djy099 (2019).
47. Phelan, C. M. et al. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat. Genet.* **49**, 680–691 (2017).
48. Lesseur, C. et al. Genome-wide association analyses identify new susceptibility loci for oral cavity and pharyngeal cancer. *Nat. Genet.* **48**, 1544–1550 (2016).
49. Schumacher, F. R. et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat. Genet.* **50**, 928–936 (2018).
50. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local genetic correlation gives insights into the shared genetic architecture of complex traits. *Am. J. Hum. Genet.* **101**, 737–751 (2017).
51. Sakoda, L. C., Jorgenson, E. & Witte, J. S. Turning of COGS moves forward findings for hormonally mediated cancers. *Nat. Genet.* **45**, 345–348 (2013).
52. Pickrell, J. K. et al. Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48**, 709–717 (2016).
53. Pickrell, J. K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am. J. Hum. Genet.* **94**, 559–573 (2014).
54. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
55. Gusev, A. et al. Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *Am. J. Hum. Genet.* **95**, 535–552 (2014).
56. Roadmap Epigenomics Consortium. et al. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
57. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
58. Trynka, G. et al. Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat. Genet.* **45**, 124–130 (2013).
59. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
60. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
61. Hoffman, M. M. et al. Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res.* **41**, 827–841 (2013).
62. Ward, L. D. & Kellis, M. Evidence of abundant purifying selection in humans for recently acquired regulatory functions. *Science* **337**, 1675–1678 (2012).
63. Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).

Acknowledgements

The authors in this manuscript were working on behalf of BCAC, CCFR, CIMBA, CORECT, GECCO, OCAC, PRACTICAL, CRUK, BPC3, CAPS, PEGASUS, TRICL-ILCCO, ABCTB, APCB, BCFR, CONSIT TEAM, EMBRACE, GC-HBOC, GEMO, HEBON, kConFab/AOCS Mod SQuaD, and SWE-BRCA. The breast cancer genome-wide association analyses: BCAC is funded by Cancer Research UK [C1287/A16563, C1287/A10118], the European Union's Horizon 2020 Research and Innovation Programme (grant numbers 634935 and 633784 for BRIDGES and B-CAST, respectively), and by the European Community's Seventh Framework Programme under grant agreement number 223175 (grant number HEALTH-F2-2009-223175) (COGS). The EU Horizon 2020 Research and Innovation Programme funding source had no role in study design, data collection, data analysis, data interpretation, or writing of the report. Genotyping of the OncoArray was funded by the NIH Grant U19 CA148065, and Cancer UK Grant C1287/A16563 and the PERSPECTIVE project supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research (grant GPH-129344) and, the Ministère de l'Économie, Science et Innovation du Québec through Genome Québec and the PSR-SIIRI-701 grant, and the Quebec Breast Cancer Foundation. Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework

Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978), and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065, and 1U19 CA148112—the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, and Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund. The DRIVE Consortium was funded by U19 CA148065. The Australian Breast Cancer Family Study (ABCFS) was supported by grant UM1 CA164920 from the National Cancer Institute (USA). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia), and the Victorian Breast Cancer Research Consortium. J.L.H. is a National Health and Medical Research Council (NHMRC) Senior Principal Research Fellow. M.C.S. is a NHMRC Senior Research Fellow. The ABCS study was supported by the Dutch Cancer Society [grants NKI 2007-3839; 2009 4363]. The Australian Breast Cancer Tissue Bank (ABCTB) is generously supported by the National Health and Medical Research Council of Australia, The Cancer Institute NSW and the National Breast Cancer Foundation. The ACP study is funded by the Breast Cancer Research Trust, UK. The AHS study is supported by the intramural research program of the National Institutes of Health, the National Cancer Institute (grant number Z01-CP010119), and the National Institute of Environmental Health Sciences (grant number Z01-ES049030). The work of the BBCC was partly funded by ELAN-Fond of the University Hospital of Erlangen. The BBCS is funded by Cancer Research UK and Breast Cancer Now and acknowledges NHS funding to the NIHR Biomedical Research Centre, and the National Cancer Research Network (NCRN). The BCEES was funded by the National Health and Medical Research Council, Australia and the Cancer Council Western Australia and acknowledges funding from the National Breast Cancer Foundation (JS). For the BCFR-NY, BCFR-PA, and BCFR-UT this work was supported by grant UM1 CA164920 from the National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the BCFR. For BIGGS, ES is supported by NIHR Comprehensive Biomedical Research Centre, Guy's & St. Thomas' NHS Foundation Trust in partnership with King's College London, United Kingdom. IT is supported by the Oxford Biomedical Research Centre. BOCS is supported by funds from Cancer Research UK (C8620/A8372/A15106) and the Institute of Cancer Research (UK). BOCS acknowledges NHS funding to the Royal Marsden/Institute of Cancer Research NIHR Specialist Cancer Biomedical Research Centre. The BREast Oncology GAlician Network (BREOGAN) is funded by Acción Estratégica de Salud del Instituto de Salud Carlos III FIS PI12/02125/Cofinanciado FEDER; Acción Estratégica de Salud del Instituto de Salud Carlos III FIS Intrasalud (PI13/01136); Programa Grupos Emergentes, Cancer Genetics Unit, Instituto de Investigacion Biomedica Galicia Sur. Xerencia de Xestion Integrada de Vigo-SERGAS, Instituto de Salud Carlos III, Spain; Grant 10CSA012E, Consellería de Industria Programa Sectorial de Investigación Aplicada, PEME I + D e I + D Suma del Plan Gallego de Investigación, Desarrollo e Innovación Tecnológica de la Consellería de Industria de la Xunta de Galicia, Spain; Grant EC11-192. Fomento de la Investigación Clínica Independiente, Ministerio de Sanidad, Servicios Sociales e Igualdad, Spain; and Grant FEDER-Innterconecta. Ministerio de Economia y Competitividad, Xunta de Galicia, Spain. The BSUCH study was supported by the Dietmar-Hopp Foundation, the Helmholtz Society and the German Cancer Research Center (DKFZ). The CAMA study was funded by Consejo Nacional de Ciencia y Tecnología (CONACyT) (SALUD-2002-C01-7462). Sample collection and processing was funded in part by grants from the National Cancer Institute (NCI R01CA120120 and K24CA169004). CBCS is funded by the Canadian Cancer Society (grant # 313404) and the Canadian Institutes of Health Research. CCGP is supported by funding from the University of Crete. The CECILE study was supported by Fondation de France, Institut National du Cancer (INCa), Ligue Nationale contre le Cancer, Agence Nationale de Sécurité Sanitaire, de l'Alimentation, de l'Environnement et du Travail (ANSES), Agence Nationale de la Recherche (ANR). The CGPS was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council, and Herlev and Gentofte Hospital. The CNIO-BCS was supported by the Instituto de Salud Carlos III, the Red Temática de Investigación Cooperativa en Cáncer and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923 and PI12/00070). COLBCCC is supported by the German Cancer Research Center (DKFZ), Heidelberg, Germany. D.T. was in part supported by a postdoctoral fellowship from the Alexander von Humboldt Foundation. The American Cancer Society funds the creation, maintenance, and updating of the CPS-II cohort. The CTS was initially supported by the California Breast Cancer Act of 1993 and the California Breast Cancer Research Fund (contract 97-10500) and is currently funded through the National Institutes of Health (R01 CA77398, UM1 CA164917, and U01 CA199277). Collection of cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885. H.A.C eceives support from the Lon V Smith Foundation (LVS39420). The University of Westminster curates the DietCompLyf database funded by Against Breast Cancer Registered Charity No. 1121258 and the NCRN. The coordination of EPIC is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by: Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF) (Germany); the Hellenic Health Foundation, the Stavros Niarchos Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); Health Research Fund (FIS), PI13/00061 to Granada, PI13/01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020) (Spain); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom). The ESTHER study was supported by a grant from the Baden Württemberg Ministry of Science, Research and Arts. Additional cases were recruited in the context of the VERDI study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe). FHRISK is funded from NIHR grant PGfAR 0707-10031. The GC-HBOC (German Consortium of Hereditary Breast and Ovarian Cancer) is supported by the German Cancer Aid (grant no 110837, coordinator: Rita K. Schmutzler, Cologne). This work was also funded by the European Regional Development Fund and Free State of Saxony, Germany (LIFE - Leipzig Research Centre for Civilization Diseases, project numbers 713-241202, 713-241202, 14505/2470, and 14575/2470). The GENICA was funded by the Federal Ministry of Education and Research (BMBF) Germany grants 01KW9975/5, 01KW9976/8, 01KW9977/0, and 01KW0114, the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany. The GEPARSIXTO study was conducted by the German Breast Group GmbH. The GESBC was supported by the Deutsche Krebshilfe e. V. [70492] and the German Cancer Research Center (DKFZ). GLACIER was supported by Breast Cancer Now, CRUK and Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London. The HABCS study was supported by the Claudia von Schilling Foundation for Breast Cancer Research, by the Lower Saxonian Cancer Society, and by the Rudolf-Bartling Foundation. The HEBCS was financially supported by the Helsinki University Central Hospital Research Fund, Academy of Finland (266528), the Finnish Cancer Society, and the Sigrid Juselius Foundation. The HERPACC was supported by MEXT Kakenhi (No. 170150181 and 26253041) from the Ministry of Education, Science, Sports, Culture and Technology of Japan, by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from Ministry Health, Labour and Welfare of Japan, by Health and Labour Sciences Research Grants for Research on Applying Health Technology from Ministry Health, Labour and Welfare of Japan, by National Cancer Center Research and Development Fund, and "Practical Research for Innovative Cancer Control (15ck0106177h0001)" from Japan Agency for Medical Research and development, AMED, and Cancer Bio Bank Aichi. The HMBCS was supported by a grant from the Friends of Hannover Medical School and by the Rudolf Bartling Foundation. The HUBCS was supported by a grant from the German Federal Ministry of Research and Education (RUS08/017), and by the Russian Foundation for Basic Research and the Federal Agency for Scientific Organizations for support the Bioresource collections and RFBR grants 14-04-97088, 17-29-06014, and 17-44-020498. ICICLE was supported by Breast Cancer Now, CRUK, and Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London. Financial support for KARBAC was provided through the regional agreement on medical training and clinical research (A.L. F.) between Stockholm County Council and Karolinska Institutet, the Swedish Cancer Society, The Gustav V Jubilee foundation and Bert von Kantzows foundation. The KARMA study was supported by Märit and Hans Rausings Initiative Against Breast Cancer. The KBCP was financially supported by the special Government Funding (E.V. O.) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, and by the strategic funding of the University of Eastern Finland. kConFab is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. Financial support for the AOCS was provided by the United States Army Medical Research and Materiel Command [DAMD17-01-1-0729], Cancer Council Victoria, Queensland Cancer Fund, Cancer Council New South Wales, Cancer Council South Australia, The Cancer Foundation of Western Australia, Cancer Council Tasmania and the National Health and Medical Research Council of Australia (NHMRC; 400413, 400281, 199600). G.C.-T. and P.W. are supported by the NHMRC. RB was a Cancer Institute NSW Clinical Research Fellow. The KOHBRA study was partially supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), and the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (HI16C1127; 1020350; 1420190). LAABC is supported by grants (1RB-0287, 3PB-0102, 5PB-0018, 10PB-0098) from the California Breast Cancer Research Program. Incident breast cancer cases were collected by the USC Cancer Surveillance Program (CSP) which is supported under subcontract by the California Department of Health. The

CSP is also part of the National Cancer Institute's Division of Cancer Prevention and Control Surveillance, Epidemiology, and End Results Program, under contract number N01CN25403. L.M.B.C. is supported by the 'Stichting tegen Kanker'. D.L. is supported by the FWO. The MABCS study is funded by the Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov" and supported by the German Academic Exchange Program, DAAD. The MARIE study was supported by the Deutsche Krebshilfe e.V. [70-2892-BR I, 106332, 108253, 108419, 110826, 110828], the Hamburg Cancer Society, the German Cancer Research Center (DKFZ) and the Federal Ministry of Education and Research (BMBF) Germany [01KH0402]. MBCSG is supported by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects "5 × 1000"). The MCBCS was supported by the NIH grants CA192393, CA116167, CA176785 an NIH Specialized Program of Research Excellence (SPORE) in Breast Cancer [CA116201], and the Breast Cancer Research Foundation and a generous gift from the David F. and Margaret T. Grohne Family Foundation. MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057 and 396414, and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database. The MEC was support by NIH grants CA63464, CA54281, CA098758, CA132839, and CA164973. The MISS study is supported by funding from ERC-2011-294576 Advanced grant, Swedish Cancer Society, Swedish Research Council, Local hospital funds, Berta Kamprad Foundation, Gunnar Nilsson. The MMHS study was supported by NIH grants CA97396, CA128931, CA116201, CA140286, and CA177150. MSKCC is supported by grants from the Breast Cancer Research Foundation and Robert and Kate Niehaus Clinical Cancer Genetics Initiative. The work of MTLGEBCS was supported by the Quebec Breast Cancer Foundation, the Canadian Institutes of Health Research for the "CIHR Team in Familial Risks of Breast Cancer" program – grant # CRN-87521 and the Ministry of Economic Development, Innovation and Export Trade – grant # PSR-SIIRI-701. MYBRCA is funded by research grants from the Malaysian Ministry of Higher Education (UM.C/HlR/MOHE/06) and Cancer Research Malaysia. MYMAMMO is supported by research grants from Yayasan Sime Darby LPGA Tournament and Malaysian Ministry of Higher Education (RP046B-15HTM). The NBCS has been supported by the Research Council of Norway grant 193387/V50 (to A.-L. Børresen-Dale and V.N. Kristensen) and grant 193387/H10 (to A.-L. Børresen-Dale and V.N. Kristensen), South Eastern Norway Health Authority (grant 39346 to A.-L. Børresen-Dale and 27208 to V.N. Kristensen) and the Norwegian Cancer Society (to A.-L. Børresen-Dale and 419616 - 71248 - PR-2006-0282 to V.N. Kristensen). It has received funding from the K.G. Jebsen Centre for Breast Cancer Research (2012-2015). The NBHS was supported by NIH grant R01CA100374. Biological sample preparation was conducted the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The Northern California Breast Cancer Family Registry (NC-BCFR) and Ontario Familial Breast Cancer Registry (OFBCR) were supported by grant UM1 CA164920 from the National Cancer Institute (USA). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR. The Carolina Breast Cancer Study was funded by Komen Foundation, the National Cancer Institute (P50 CA058223, U54 CA156733, and U01 CA179715), and the North Carolina University Cancer Research Fund. The NGOBCS was supported by Grants-in-Aid for the Third Term Comprehensive Ten-Year Strategy for Cancer Control from the Ministry of Health, Labor and Welfare of Japan, and for Scientific Research on Priority Areas, 17015049 and for Scientific Research on Innovative Areas, 221S0001, from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. The NHS was supported by NIH grants P01 CA87969, UM1 CA186107, and U19 CA148065. The NHS2 was supported by NIH grants UM1 CA176726 and U19 CA148065. The OBCS was supported by research grants from the Finnish Cancer Foundation, the Academy of Finland (grant number 250083, 122715 and Center of Excellence grant number 251314), the Finnish Cancer Foundation, the Sigrid Juselius Foundation, the University of Oulu, the University of Oulu Support Foundation, and the special Governmental EVO funds for Oulu University Hospital-based research activities. The ORIGO study was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16). The PBCS was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. Genotyping for PLCO was supported by the Intramural Research Program of the National Institutes of Health, NCI, Division of Cancer Epidemiology and Genetics. The PLCO is supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics and supported by contracts from the Division of Cancer Prevention, National Cancer Institute, National Institutes of Health. The POSH study is funded by Cancer Research UK (grants C1275/A11699, C1275/C22524, C1275/A19187, C1275/A15956, and Breast Cancer Campaign 2010PR62, 2013PR044. PROCAS is funded from NIHR grant PGfAR 0707-10031. The RBCS was funded by the Dutch Cancer Society (DDHK 2004-3124, DDHK 2009-4318). The SASBAC study was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the US National Institute of Health (NIH) and the Susan G. Komen Breast Cancer Foundation. The SBCGS was supported primarily by NIH grants R01CA64277, R01CA148667, UMCA182910, and R37CA70867. Biological sample preparation was conducted the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The scientific development and funding of this project were, in part, supported by the Genetic Associations and Mechanisms in Oncology (GAME-ON) Network U19 CA148065. The SBCS was supported by Sheffield Experimental Cancer Medicine Centre and Breast Cancer Now Tissue Bank. The SCCS is supported by a grant from the National Institutes of Health (R01 CA092447). Data on SCCS cancer cases used in this publication were provided by the Alabama Statewide Cancer Registry; Kentucky Cancer Registry, Lexington, KY; Tennessee Department of Health, Office of Cancer Surveillance; Florida Cancer Data System; North Carolina Central Cancer Registry, North Carolina Division of Public Health; Georgia Comprehensive Cancer Registry; Louisiana Tumor Registry; Mississippi Cancer Registry; South Carolina Central Cancer Registry; Virginia Department of Health, Virginia Cancer Registry; Arkansas Department of Health, Cancer Registry, 4815 W. Markham, Little Rock, AR 72205. The Arkansas Central Cancer Registry is fully funded by a grant from National Program of Cancer Registries, Centers for Disease Control and Prevention (CDC). Data on SCCS cancer cases from Mississippi were collected by the Mississippi Cancer Registry which participates in the National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention (CDC). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the CDC or the Mississippi Cancer Registry. SEARCH is funded by Cancer Research UK [C490/A10124, C490/A16561] and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. The University of Cambridge has received salary support for PDPP from the NHS in the East of England through the Clinical Academic Reserve. SEBCS was supported by the BRL (Basic Research Laboratory) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2012-0000347). SGBCC is funded by the NUS start-up Grant, National University Cancer Institute Singapore (NCIS) Centre Grant and the NMRC Clinician Scientist Award. Additional controls were recruited by the Singapore Consortium of Cohort Studies-Multi-ethnic cohort (SCCS-MEC), which was funded by the Biomedical Research Council, grant number: 05/1/21/19/425. The Sister Study (SISTER) is supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01-ES044005 and Z01-ES049033). The Two Sister Study (2SISTER) was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01-ES044005 and Z01-ES102245), and, also by a grant from Susan G. Komen for the Cure, grant FAS0703856. SKKDKFZS is supported by the DKFZ. The SMC is funded by the Swedish Cancer Foundation. The SZBCS was supported by Grant PBZ_KBN_122/P05/2004. The TBCS was funded by The National Cancer Institute, Thailand. The TNBCC was supported by a Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), a grant from the Breast Cancer Research Foundation, a generous gift from the David F. and Margaret T. Grohne Family Foundation. The TWBCS is supported by the Taiwan Biobank project of the Institute of Biomedical Sciences, Academia Sinica, Taiwan. The UCIBCS component of this research was supported by the NIH [CA58860, CA92044] and the Lon V Smith Foundation [LVS39420]. The UKBGS is funded by Breast Cancer Now and the Institute of Cancer Research (ICR), London. ICR acknowledges NHS funding to the NIHR Biomedical Research Centre. The UKOPS study was funded by The Eve Appeal (The Oak Foundation) and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. The US3SS study was supported by Massachusetts (K.M.E., R01CA47305), Wisconsin (P.A.N., R01 CA47147) and New Hampshire (L.T.-E., R01CA69664) centers, and Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The USRT Study was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The WAABCS study was supported by grants from the National Cancer Institute of the National Institutes of Health (R01 CA89085 and P50 CA125183 and the D43 TW009112 grant), Susan G. Komen (SAC110026), the Dr. Ralph and Marian Falk Medical Research Trust, and the Avon Foundation for Women. The WHI program is funded by the National Heart, Lung, and Blood Institute, the US National Institutes of Health and the US Department of Health and Human Services (HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C). This work was also funded by NCI U19 CA148065-01. D.G.E. is supported by the all Manchester NIHR Biomedical research center Manchester (IS-BRC-1215-20007). HUNBOCS, Hungarian Breast and Ovarian Cancer Study was supported by Hungarian Research Grant KTIA-OTKA CK-80745, NKFI_OTKA K-112228. C.I. received support from the Nontherapeutic Subject Registry Shared Resource at Georgetown University (NIH/NCI P30-CA-51008) and the Jess and Mildred Fisher Center for Hereditary Cancer and Clinical Genomics Research. K.M. is supported by CRUK C18281/A19169. City of Hope Clinical Cancer Community Research Network and the Hereditary Cancer Research Registry, supported in part by Award Number RC4CA153828 (PI: J Weitzel) from the National Cancer Institute and the office of the Directory, National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The colorectal cancer genome-wide association analyses: Colorectal Transdisciplinary Study (CORECT): The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the CORECT Consortium, nor does mention of trade names, commercial products or organizations imply endorsement by the US Government or the CORECT Consortium. We are incredibly grateful for the contributions of Dr. Brian Henderson and Dr. Roger Green over the course of this study and acknowledge them in memoriam. We are also grateful for support from Daniel and Maryann Fong. ColoCare: we thank the many investigators and staff who made this

research possible in ColoCare Seattle and ColoCare Heidelberg. ColoCare was initiated and developed at the Fred Hutchinson Cancer Research Center by Drs. Ulrich and Grady. CCFR: the Colon CFR graciously thanks the generous contributions of their study participants, dedication of study staff, and financial support from the U.S. National Cancer Institute, without which this important registry would not exist. Galeon: GALEON wishes to thank the Department of Surgery of University Hospital of Santiago (CHUS), Sara Miranda Ponte, Carmen M Redondo, and the staff of the Department of Pathology and Biobank of CHUS, Instituto de Investigación Sanitaria de Santiago (IDIS), Instituto de Investigación Sanitaria Galicia Sur (IISGS), SERGAS, Vigo, Spain, and Programa Grupos Emergentes, Cancer Genetics Unit, CHUVI Vigo Hospital, Instituto de Salud Carlos III, Spain. MCCS: this study was made possible by the contribution of many people, including the original investigators and the diligent team who recruited participants and continue to work on follow-up. We would also like to express our gratitude to the many thousands of Melbourne residents who took part in the study and provided blood samples. SEARCH: We acknowledge the contributions of Mitul Shah, Val Rhenius, Sue Irvine, Craig Luccarini, Patricia Harrington, Don Conroy, Rebecca Mayes, and Caroline Baynes. The Swedish low-risk colorectal cancer study: we thank Berith Wejderot and the Swedish low-risk colorectal cancer study group. Genetics & Epidemiology of Colorectal Cancer Consortium (GECCO): we thank all those at the GECCO Coordinating Center for helping bring together the data and people that made this project possible. ASTERISK: we are very grateful to Dr. Bruno Buecher without whom this project would not have existed. We also thank all those who agreed to participate in this study, including the patients and the healthy control persons, as well as all the physicians, technicians and students. DACHS: we thank all participants and cooperating clinicians, and Ute Handte-Daub, Renate Hettler-Jensen, Utz Benscheid, Muhabbet Celik, and Ursula Eilber for excellent technical assistance. HPFS, NHS and PHS: we acknowledge Patrice Soule and Hardeep Ranu of the Dana-Farber Harvard Cancer Center High-Throughput Polymorphism Core who assisted in the genotyping for NHS, HPFS, and PHS under the supervision of Dr. Immaculata Devivo and Dr. David Hunter, Qin (Carolyn) Guo, and Lixue Zhu who assisted in programming for NHS and HPFS and Haiyan Zhang who assisted in programming for the PHS. We thank the participants and staff of the Nurses' Health Study and the Health Professionals Follow-Up Study, for their valuable contributions as well as the following state cancer registries for their help: A.L., A.Z., A.R., C.A., C.O., C.T., D.E., F.L., G.A., I.D., I.L., I.N., I.A., K.Y., L.A., M.E., M.D., M.A., M.I., N.E., N.H., N.J., N.Y., N.C., N.D., O.H., O.K., O.R., P.A., R.I., S.C., T.N., T.X., V.A., W.A., W.Y. In addition, this study was approved by the Connecticut Department of Public Health (DPH) Human Investigations Committee. Certain data used in this publication were obtained from the DPH. We assume full responsibility for analyses and interpretation of these data. PLCO: we thank Drs. Christine Berg and Philip Prorok, Division of Cancer Prevention, National Cancer Institute, the Screening Center investigators and staff or the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial, Mr. Tom Riley and staff, Information Management Services Inc., Ms. Barbara O'Brien and staff, Westat Inc. and Drs. Bill Kopp, Wen Shao and staff, SAIC-Frederick. Most importantly, we acknowledge the study participants for their contributions for making this study possible. The statements contained herein are solely those of the authors and do not represent or imply concurrence or endorsement by NCI. PMH: we thank the study participants and staff of the Hormones and Colon Cancer study. WHI: we thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at https://cleo.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short20List.pdf. CORECT: The CORECT Study was supported by the National Cancer Institute, National Institutes of Health (NCI/NIH), U.S. Department of Health and Human Services (grant numbers U19 CA148107, R01 CA81488, P30 CA014089, R01 CA197350; P01 CA196569; and R01 CA201407) and National Institutes of Environmental Health Sciences, National Institutes of Health (grant number T32 ES013678). The ATBC Study was supported by the US Public Health Service contracts (N01-CN-45165, N01-RC-45035, N01-RC-37004, and HHSN261201000006C) from the National Cancer Institute. The Cancer Prevention Study-II Nutrition Cohort is funded by the American Cancer Society. ColoCare: This work was supported by the National Institutes of Health (grant numbers R01 CA189184, U01 CA206110, 2P30CA015704-40 (Gilliland)), the Matthias Lackas-Foundation, the German Consortium for Translational Cancer Research, and the EU TRANSCAN initiative. Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO): funding for GECCO was provided by the National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services (grant numbers U01 CA137088, R01 CA059045, and U01 CA164930). This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA015704. The Colon Cancer Family Registry (CFR) Illumina GWAS was supported by funding from the National Cancer Institute, National Institutes of Health (grant numbers U01 CA122839, R01 CA143247). The Colon CFR/CORECT Affymetrix Axiom GWAS and OncoArray GWAS were supported by funding from National Cancer Institute, National Institutes of Health (grant number U19 CA148107 to S.G.). The Colon CFR participant recruitment and collection of data and biospecimens used in this study were supported by the National Cancer Institute, National Institutes of Health (grant number UM1 CA167551) and through cooperative agreements with the following Colon CFR centers: Australasian Colorectal Cancer Family Registry (NCI/NIH grant numbers U01 CA074778 and U01/U24 CA097735), USC Consortium Colorectal Cancer Family Registry (NCI/NIH grant numbers U01/U24 CA074799), Mayo Clinic Cooperative Family Registry for Colon Cancer Studies (NCI/NIH grant number U01/U24 CA074800), Ontario Familial Colorectal Cancer Registry (NCI/NIH grant number U01/U24 CA074783), Seattle Colorectal Cancer Family Registry (NCI/NIH grant number U01/U24 CA074794), and University of Hawaii Colorectal Cancer Family Registry (NCI/NIH grant number U01/U24 CA074806), Additional support for case ascertainment was provided from the Surveillance, Epidemiology and End Results (SEER) Program of the National Cancer Institute to Fred Hutchinson Cancer Research Center (Control Nos. N01-CN-67009 and N01-PC-35142, and Contract No. HHSN2612013000121), the Hawai'i Department of Health (Control Nos. N01-PC-67001 and N01-PC-35137, and Contract No. HHSN26120100037C, and the California Department of Public Health (contracts HHSN261201000035C awarded to the University of Southern California, and the following state cancer registries: A.Z., C.O., M.N., N.C., N.H., and by the Victoria Cancer Registry and Ontario Cancer Registry. ESTHER/VERDI was supported by grants from the Baden–Württemberg Ministry of Science, Research and Arts and the German Cancer Aid. MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. GALEON: FIS Intrasalud (PI13/01136). The MCCS was further supported by Australian NHMRC grants 509348, 209057, 251553, and 504711 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database. MSKCC: the work at Sloan Kettering in New York was supported by the Robert and Kate Niehaus Center for Inherited Cancer Genomics and the Romeo Milio Foundation. Moffitt: This work was supported by funding from the National Institutes of Health (grant numbers R01 CA189184, P30 CA076292), Florida Department of Health Bankhead-Coley Grant 09BN-13, and the University of South Florida Oehler Foundation. Moffitt contributions were supported in part by the Total Cancer Care Initiative, Collaborative Data Services Core, and Tissue Core at the H. Lee Moffitt Cancer Center & Research Institute, a National Cancer Institute-designated Comprehensive Cancer Center (grant number P30 CA076292). SEARCH: Cancer Research UK (C490/A16561). The Spanish study was supported by Instituto de Salud Carlos III, co-funded by FEDER funds –a way to build Europe– (grants PI14-613 and PI09-1286), Catalan Government DURSI (grant 2014SGR647), and Junta de Castilla y León (grant LE22A10-2). The Swedish Low-risk Colorectal Cancer Study: the study was supported by grants from the Swedish research council; K2015-55 × -22674-01-4, K2008-55 × -20157-03-3, K2006-72 × -20157-01-2 and the Stockholm County Council (ALF project). CIDR genotyping for the Oncoarray was conducted under contract 268201200008I (to K.D.), through grant 101HG007491-01 (to C.I.A.). The Norris Cotton Cancer Center - P30CA023108, The Quantitative Biology Research Institute - P20GM103534, and the Coordinating Center for Screen Detected Lesions - U01CA196386 also supported efforts of C.I.A. This work was also supported by the National Cancer Institute (grant numbers U01 CA1817700, R01 CA144040). ASTERISK: a Hospital Clinical Research Program (PHRC) and supported by the Regional Council of Pays de la Loire, the Groupement des Entreprises Françaises dans la Lutte contre le Cancer (GEFLUC), the Association Anne de Bretagne Génétique and the Ligue Régionale Contre le Cancer (LRCC). COLO2&3: National Institutes of Health (grant number R01 CA060987). DACHS: This work was supported by the German Research Council (BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, CH 117/1-1, HO 5117/2-1, HE 5998/2-1, KL 2354/3-1, RO 2270/8-1, and BR 1704/17-1), the Interdisciplinary Research Program of the National Center for Tumor Diseases (NCT), Germany, and the German Federal Ministry of Education and Research (01KH0404, 01ER0814, 01ER0815, 01ER1505A, and 01ER1505B). DALS: National Institutes of Health (grant number R01 CA048998 to M.L. S). HPFS is supported by National Institutes of Health (grant numbers P01 CA055075, UM1 CA167552, R01 137178, and P50 CA127003), NHS by the National Institutes of Health (grant numbers UM1 CA186107, R01 CA137178, P01 CA087969, and P50 CA127003), NHSII by the National Institutes of Health (grant numbers R01 050385CA and UM1 CA176726), and PHS by the National Institutes of Health (grant number R01 CA042182). MEC: National Institutes of Health (grant numbers R37 CA054281, P01 CA033619, and R01 CA063464). OFCCR: National Institutes of Health, through funding allocated to the Ontario Registry for Studies of Familial Colorectal Cancer (grant number U01 CA074783); see Colon CFR section above. As subset of ARCTIC, OFCCR is supported by a GL2 grant from the Ontario Research Fund, the Canadian Institutes of Health Research, and the Cancer Risk Evaluation (CaRE) Program grant from the Canadian Cancer Society Research Institute. T.J.H. and B.W.Z. are recipients of Senior Investigator Awards from the Ontario Institute for Cancer Research, through generous support from the Ontario Ministry of Research and Innovation. PLCO: Intramural Research Program of the Division of Cancer Epidemiology and Genetics and supported by contracts from the Division of Cancer Prevention, National Cancer Institute, NIH, DHHS. Additionally, a subset of control samples was genotyped as part of the Cancer Genetic Markers of Susceptibility (CGEMS) Prostate Cancer GWAS, Colon CGEMS pancreatic cancer scan (PanScan), and the Lung Cancer and Smoking study. The prostate and PanScan study datasets were accessed with appropriate approval through the dbGaP online resource (http://cgems.cancer.gov/data/) accession numbers phs000207.v1.p1 and phs000206.v3.p2, respectively, and the lung datasets were accessed from the dbGaP website (http://www.ncbi.nlm.nih.gov/gap) through accession number phs000093.v2.p2. Funding for the Lung Cancer and Smoking study was provided by National Institutes of Health (NIH), Genes, Environment and Health Initiative (GEI) Z01 CP 010200, NIH U01 HG004446, and NIH GEI U01 HG 004438. For the lung study, the GENEVA Coordinating Center provided assistance with genotype cleaning and general study coordination, 23 and the Johns Hopkins University Center for Inherited Disease Research conducted genotyping. PMH: National Institutes of Health (grant number R01 CA076366). VITAL: National Institutes of Health (grant number K05-CA154337). WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201600018C,

HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN26 8201600004C. The head and neck cancer genome-wide association analyses: The study was supported by NIH/NCI: P50 CA097190, and P30 CA047904, Canadian Cancer Society Research Institute (no. 020214) and Cancer Care Ontario Research Chair to R.H. The Princess Margaret Hospital Head and Neck Cancer Translational Research Program is funded by the Wharton family, Joe's Team, Gordon Tozer, Bruce Galloway and the Elia family. Geoffrey Liu was supported by the Posluns Family Fund and the Lusi Wong Family Fund at the Princess Margaret Foundation, and the Alan B. Brown Chair in Molecular Genomics. This publication presents data from Head and Neck 5000 (H&N5000). H&N5000 was a component of independent research funded by the UK National Institute for Health Research (NIHR) under its Programme Grants for Applied Research scheme (RP-PG-0707-10034). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. Human papillomavirus (HPV) in H&N5000 serology was supported by a Cancer Research UK Programme Grant, the Integrative Cancer Epidemiology Programme (grant number: C18281/A19169). National Cancer Institute (R01-CA90731); National Institute of Environmental Health Sciences (P30ES10126). The authors thank all the members of the GENCAPO team/The Head and Neck Genome Project (GENCAPO) was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Grant numbers 04/12054-9 and 10/51168-0). CPS-II recruitment and maintenance is supported with intramural research funding from the American Cancer Society. Genotyping performed at the Center for Inherited Disease Research (CIDR) was funded through the U.S. National Institute of Dental and Craniofacial Research (NIDCR) grant 1 × 01HG007780-0. The University of Pittsburgh head and neck cancer case-control study is supported by National Institutes of Health grants P50 CA097190 and P30 CA047904. The Carolina Head and Neck Cancer Study (CHANCE) was supported by the National Cancer Institute (R01-CA90731). The Head and Neck Genome Project (GENCAPO) was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Grant numbers 04/12054-9 and 10/51168-0). The authors thank all the members of the GENCAPO team. The HN5000 study was funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research scheme (RP-PG-0707-10034), the views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The Toronto study was funded by the Canadian Cancer Society Research Institute (020214) and the National Cancer Institute (U19-CA148127) and the Cancer Care Ontario Research Chair. The alcohol-related cancers and genetic susceptibility study in Europe (ARCAGE) was funded by the European Commission's 5th Framework Program (QLK1-2001-00182), the Italian Association for Cancer Research, Compagnia di San Paolo/FIRMS, Region Piemonte, and Padova University (CPDA057222). The Rome Study was supported by the Associazione Italiana per la Ricerca sul Cancro (AIRC) IG 2011 10491 and IG2013 14220 to S.B., and Fondazione Veronesi to S.B. The IARC Latin American study was funded by the European Commission INCO-DC programme (IC18-CT97-0222), with additional funding from Fondo para la Investigacion Cientifica y Tecnologica (Argentina) and the Fundação de Amparo à Pesquisa do Estado de São Paulo (01/01768-2). We thank Leticia Fernandez, Instituto Nacional de Oncologia y Radiobiologia, La Habana, Cuba and Sergio and Rosalina Koifman, for their efforts with the IARC Latin America study São Paulo center. The IARC Central Europe study was supported by European Commission's INCO-COPERNICUS Program (IC15- CT98-0332), NIH/National Cancer Institute grant CA92039, and the World Cancer Research Foundation grant WCRF 99A28. The IARC Oral Cancer Multicenter study was funded by grant S06 96 202489 05F02 from Europe against Cancer; grants FIS 97/0024, FIS 97/0662, and BAE 01/5013 from Fondo de Investigaciones Sanitarias, Spain; the UICC Yamagiwa-Yoshida Memorial International Cancer Study; the National Cancer Institute of Canada; Associazione Italiana per la Ricerca sul Cancro; and the Pan-American Health Organization. Coordination of the EPIC study is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The lung cancer genome-wide association analyses: Transdisciplinary Research for Cancer in Lung (TRICL) of the International Lung Cancer Consortium (ILCCO) was supported by (U19-CA148127, CA148127S1, U19CA203654, and Cancer Prevention Research Institute of Texas RR170048). The ILCCO data harmonization is supported by Cancer Care Ontario Research Chair of Population Studies to R. H. and Lunenfeld-Tanenbaum Research Institute, Sinai Health System. The TRICL-ILCCO OncoArray was supported by in-kind genotyping by the Centre for Inherited Disease Research (26820120008i-0-26800068-1). The CAPUA study was supported by FIS-FEDER/Spain grant numbers FIS-01/310, FIS-PI03-0365, and FIS-07-BI060604, FICYT/Asturias grant numbers FICYT PB02-67 and FICYT IB09-133, and the University Institute of Oncology (IUOPA), of the University of Oviedo and the Ciber de Epidemiologia y Salud Pública. CIBERESP, SPAIN. The work performed in the CARET study was supported by the National Institute of Health/National Cancer Institute: UM1 CA167462 (PI: Goodman), National Institute of Health UO1-CA6367307 (PIs Omen, Goodman); National Institute of Health R01 CA111703 (PI Chen), National Institute of Health 5R01 CA151989-01A1(PI Doherty). The Liverpool Lung project is supported by the Roy Castle Lung Cancer Foundation. The Harvard Lung Cancer Study was supported by the NIH (National Cancer Institute) grants CA092824, CA090578, CA074386. The Multi-ethnic Cohort Study was partially supported by NIH Grants CA164973, CA033619, CA63464, and CA148127. The work performed in MSH-PMH study was supported by The Canadian Cancer Society Research Institute (020214), Ontario Institute of Cancer and Cancer Care Ontario Chair Award to R.J.H. and G.L. and the Alan Brown Chair and Lusi Wong Programs at the Princess Margaret Hospital Foundation. NJLCS was funded by the State Key Program of National Natural Science of China (81230067), the National Key Basic Research Program Grant (2011CB503805), the Major Program of the National Natural Science Foundation of China (81390543). The Norway study was supported by Norwegian Cancer Society, Norwegian Research Council. The Shanghai Cohort Study (SCS) was supported by National Institutes of Health R01 CA144034 (PI: Yuan) and UM1 CA182876 (PI: Yuan). The Singapore Chinese Health Study (SCHS) was supported by National Institutes of Health R01 CA144034 (PI: Yuan) and UM1 CA182876 (PI: Yuan). The work in TLC study has been supported in part the James & Esther King Biomedical Research Program (09KN-15), National Institutes of Health Specialized Programs of Research Excellence (SPORE) Grant (P50 CA119997), and by a Cancer Center Support Grant (CCSG) at the H. Lee Moffitt Cancer Center and Research Institute, an NCI designated Comprehensive Cancer Center (grant number P30-CA76292). The Vanderbilt Lung Cancer Study—BioVU dataset used for the analyses described was obtained from Vanderbilt University Medical Center's BioVU, which is supported by institutional funding, the 1S10RR025141-01 instrumentation award, and by the Vanderbilt CTSA grant UL1TR000445 from NCATS/NIH. Dr. Aldrich was supported by NIH/National Cancer Institute K07CA172294 (PI: Aldrich) and Dr. Bush was supported by NHGRI/NIH U01HG004798 (PI: Crawford). The Copenhagen General Population Study (CGPS) was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council and Herlev Hospital. The NELCS study: Grant Number P20RR018787 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). The Kentucky Lung Cancer Research Initiative was supported by the Department of Defense [Congressionally Directed Medical Research Program, U.S. Army Medical Research and Materiel Command Program] under award number: 10153006 (W81XWH-11-1-0781). Views and opinions of, and endorsements by the author(s) do not reflect those of the US Army or the Department of Defense. This research was also supported by unrestricted infrastructure funds from the UK Center for Clinical and Translational Science, NIH grant UL1TR000117 and Markey Cancer Center NCI Cancer Center Support Grant (P30 CA177558) Shared Resource Facilities: Cancer Research Informatics, Biospecimen and Tissue Procurement, and Biostatistics and Bioinformatics. The M.D. Anderson Cancer Center study was supported in part by grants from the NIH (P50 CA070907, R01 CA176568) (to X.W.), Cancer Prevention & Research Institute of Texas (RP130502) (to X. W.), and The University of Texas MD Anderson Cancer Center institutional support for the Center for Translational and Public Health Genomics. The deCODE study of smoking and nicotine dependence was funded in part by a grant from NIDA (R01- DA017932). The study in Lodz center was partially funded by Nofer Institute of Occupational Medicine, under task NIOM 10.13: Predictors of mortality from non-small cell lung cancer—field study. Genetic sharing analysis was funded by NIH grant CA194393. The research undertaken by M.D.T., L.V.W., and M.S.A. was partly funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. M.D.T. holds a Medical Research Council Senior Clinical Fellowship (G0902313). The work to assemble the FTND GWAS meta-analysis was supported by the National Institutes of Health (NIH), National Institute on Drug Abuse (NIDA) grant number R01 DA035825 (Principal Investigator [PI]: DBH). The study populations included COGEND (dbGaP phs000092.v1.p1 and phs000404.v1.p1), COPDGene (dbGaP phs000179.v3.p2), deCODE Genetics, EAGLE (dbGaP phs000093.vs.p2), and SAGE. dbGaP phs000092.v1.p1). See Hancock et al. Transl Psychiatry 2015 (PMCID: PMC4930126) for the full listing of funding sources and other acknowledgments. The Resource for the Study of Lung Cancer Epidemiology in North Trent (ReSoLuCENT)study was funded by the Sheffield Hospitals Charity, Sheffield Experimental Cancer Medicine Centre and Weston Park Hospital Cancer Charity. The ovarian cancer genome-wide association analysis: The Ovarian Cancer Association Consortium (OCAC) is supported by a grant from the Ovarian Cancer Research Fund thanks to donations by the family and friends of Kathryn Sladek Smith (PPD/RPCI.07). The scientific development and funding for this project were in part supported by the US National Cancer Institute GAME-ON Post-GWAS Initiative (U19-CA148112). This study made use of data generated by the Wellcome Trust Case Control consortium that was funded by the Wellcome Trust under award 076113. The results published here are in part based upon data generated by The Cancer Genome Atlas Pilot Project established by the National Cancer Institute and National Human Genome Research Institute (dbGap accession number phs000178.v8.p7). The OCAC OncoArray genotyping project was funded through grants from the U.S. National Institutes of Health (CA1X01HG007491-01 (C.I.A.), U19-CA148112 (T.A.S.), R01-CA149429 (C.M.P.), and R01-CA058598 (M.T.G.); Canadian Institutes of Health Research (MOP-86727 (L.E.K.) and the Ovarian Cancer Research Fund (A.B.). The COGS project was funded through a European Commission's Seventh Framework Programme grant (agreement number 223175 - HEALTH-F2-2009-223175) and through a grant from the U.S. National Institutes of Health (R01-CA122443 (E.L.G)). Funding for individual studies: AAS: National Institutes of Health (RO1-CA142081); AOV: The Canadian Institutes for Health Research (MOP-86727); AUS: The Australian Ovarian Cancer Study Group was supported by the U.S. Army Medical Research and Materiel Command (DAMD17-01-1-0729), National Health & Medical Research Council of Australia (199600, 400413 and 400281), Cancer Councils of New South Wales, Victoria, Queensland, South Australia and Tasmania and Cancer Foundation of Western Australia (Multi-State Applications 191, 211, and 182). The Australian Ovarian Cancer Study gratefully acknowledges additional support from Ovarian Cancer Australia and the Peter MacCallum Foundation; BAV: ELAN Funds of the University of Erlangen-Nuremberg; BEL: National Kankerplan; BGS: Breast Cancer Now, Institute of Cancer Research; BVU: Vanderbilt CTSA grant from the National Institutes of Health (NIH)/National Center for Advancing Translational Sciences

(NCATS) (ULTR000445); CAM: National Institutes of Health Research Cambridge Biomedical Research Centre and Cancer Research UK Cambridge Cancer Centre; CHA: Innovative Research Team in University (PCSIRT) in China (IRT1076); CNI: Instituto de Salud Carlos III (PI12/01319); Ministerio de Economía y Competitividad (SAF2012); COE: Department of Defense (W81XWH-11-2-0131); CON: National Institutes of Health (R01-CA063678, R01-CA074850; and R01-CA080742); DKE: Ovarian Cancer Research Fund; DOV: National Institutes of Health R01-CA112523 and R01-CA87538; EMC: Dutch Cancer Society (EMC 2014-6699); EPC: The coordination of EPIC is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF) (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); ERC-2009-AdG 232997 and Nordforsk, Nordic Centre of Excellence programme on Food, Nutrition and Health (Norway); Health Research Fund (FIS), PI13/00061 to Granada, PI13/01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020) (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom); GER: German Federal Ministry of Education and Research, Programme of Clinical Biomedical Research (01 GB 9401) and the German Cancer Research Center (DKFZ); GRC: This research has been co-financed by the European Union (European Social Fund—ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF)—Research Funding Program of the General Secretariat for Research & Technology: SYN11_10_19 NBCA. Investing in knowledge society through the European Social Fund; GRR: Roswell Park Cancer Institute Alliance Foundation, P30 CA016056; HAW: U.S. National Institutes of Health (R01-CA58598, N01-CN-55424, and N01-PC-67001); HJO: Intramural funding; Rudolf-Bartling Foundation; HMO: Intramural funding; Rudolf-Bartling Foundation; HOC: Helsinki University Research Fund; HOP: Department of Defense (DAMD17-02-1-0669) and NCI (K07-CA080668, R01-CA95023, P50-CA159981 MO1-RR000056 R01-CA126841); HUO: Intramural funding; Rudolf-Bartling Foundation; JGO: JSPS KAKENHI grant; JPN: Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labour and Welfare; KRA: This study (Ko-EVE) was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), and the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (HI16C1127; 0920010); LAX: American Cancer Society Early Detection Professorship (SIOP-06-258-01-COUN) and the National Center for Advancing Translational Sciences (NCATS), Grant UL1TR000124; LUN: ERC-2011-AdG 294576-risk factors cancer, Swedish Cancer Society, Swedish Research Council, Beta Kamprad Foundation; MAC: National Institutes of Health (R01-CA122443, P30-CA15083, P50-CA136393); Mayo Foundation; Minnesota Ovarian Cancer Alliance; Fred C. and Katherine B. Andersen Foundation; Fraternal Order of Eagles; MAL: Funding for this study was provided by research grant R01- CA61107 from the National Cancer Institute, Bethesda, MD, research grant 94 222 52 from the Danish Cancer Society, Copenhagen, Denmark; and the Mermaid I project; MAS: Malaysian Ministry of Higher Education (UM.C/HlR/MOHE/06) and Cancer Research Initiatives Foundation; MAY: National Institutes of Health (R01-CA122443, P30-CA15083, and P50-CA136393); Mayo Foundation; Minnesota Ovarian Cancer Alliance; Fred C. and Katherine B. Andersen Foundation; MCC: Cancer Council Victoria, National Health and Medical Research Council of Australia (NHMRC) grants number 209057, 251533, 396414, and 504715; MDA: DOD Ovarian Cancer Research Program (W81XWH-07-0449); MEC: NIH (CA54281, CA164973, CA63464); MOF: Moffitt Cancer Center, Merck Pharmaceuticals, the state of Florida, Hillsborough County, and the city of Tampa; NCO: National Institutes of Health (R01-CA76016) and the Department of Defense (DAMD17-02-1-0666); NEC: National Institutes of Health R01-CA54419 and P50-CA105009 and Department of Defense W81XWH-10-1-02802; NHS: UM1 CA186107, P01 CA87969, R01 CA49449, R01-CA67262, UM1 CA176726; NJO: National Cancer Institute (NIH-K07 CA095666, R01-CA83918, NIH-K22-CA138563, and P30-CA072720) and the Cancer Institute of New Jersey; If Sara Olson and/or Irene Orlow is a co-author, please add NCI CCSG award (P30-CA008748) to the funding sources; NOR: Helse Vest, The Norwegian Cancer Society, The Research Council of Norway; NTH: Radboud University Medical Centre; OPL: National Health and Medical Research Council (NHMRC) of Australia (APP1025142) and Brisbane Women's Club; ORE: OHSU Foundation; OVA: This work was supported by Canadian Institutes of Health Research grant (MOP-86727) and by NIH/NCI 1 R01CA160669-01A1; PLC: Intramural Research Program of the National Cancer Institute; POC: Pomeranian Medical University; POL: Intramural Research Program of the National Cancer Institute; PVD: Canadian Cancer Society and Cancer Research Society GRePEC Program; RBH: National Health and Medical Research Council of Australia; RMH: Cancer Research UK, Royal Marsden Hospital; RPC: National Institute of Health (P50-CA159981, R01-CA126841); SEA: Cancer Research UK (C490/A10119 C490/A10124); UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge; SIS: NIH, National Institute of Environmental Health Sciences, Z01-ES044005 and Z01-ES049033; SMC: The bbSwedish Research Council-SIMPLER infrastructure; the Swedish Cancer Foundation; **SON:** National Health Research and Development Program, Health Canada, grant 6613-1415-53; SRO: Cancer Research UK (C536/A13086, C536/A6689) and Imperial Experimental Cancer Research Centre (C1312/A15589); STA: NIH grants U01 CA71966 and U01 CA69417; SWE: Swedish Cancer foundation, WeCanCureCancer and VårKampMotCancer foundation; SWH: NIH (NCI) grant R37-CA070867; TBO: National Institutes of Health (R01-CA106414-A2), American Cancer Society (CRTG-00-196-01-CCE), Department of Defense (DAMD17-98-1-8659), Celma Mastery Ovarian Cancer Foundation; TOR: NIH grants R01-CA063678 and R01 CA063682; UCI: NIH R01-CA058860 and the Lon V Smith Foundation grant LVS39420; UHN: Princess Margaret Cancer Centre Foundation-Bridge for the Cure; UKO: The UKOPS study was funded by The Eve Appeal (The Oak Foundation) and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre; **UKR:** Cancer Research UK (C490/A6187), UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge; USC: P01CA17054, P30CA14089, R01CA61132, N01PC67010, R03CA113148, R03CA115195, N01CN025403, and California Cancer Research Program (00-01389V-20170, 2II0200); VAN: BC Cancer Foundation, VGH & UBC Hospital Foundation; VTL: NIH K05-CA154337; WMH: National Health and Medical Research Council of Australia, Enabling Grants ID 310670 & ID 628903. Cancer Institute NSW Grants 12/RIG/1-17 & 15/RIG/1-16; WOC: National Science Centren (N N301 5645 40). The Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland. The University of Cambridge has received salary support for PDPP from the NHS in the East of England through the Clinical Academia Reserve. The prostate cancer genome-wide association analyses: we pay tribute to Brian Henderson, who was a driving force behind the OncoArray project, for his vision and leadership, and who sadly passed away before seeing its fruition. We also thank the individuals who participated in these studies enabling this work. The ELLIPSE/PRACTICAL (http//:practical.icr.ac.uk) prostate cancer consortium and his collaborating partners were supported by multiple funding mechanisms enabling this current work. ELLIPSE/PRACTICAL Genotyping of the OncoArray was funded by the US National Institutes of Health (NIH) (U19 CA148537 for ELucidating Loci Involved in Prostate Cancer SuscEptibility (ELLIPSE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN268201200008I). Additional analytical support was provided by NIH NCI U01 CA188392 (F.R.S.). Funding for the iCOGS infrastructure came from the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, and C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065, and 1U19 CA148112; the GAME-ON initiative), the Department of Defense (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund. This work was supported by the Canadian Institutes of Health Research, European Commission's Seventh Framework Programme grant agreement n° 223175 (HEALTH-F2-2009-223175), Cancer Research UK Grants C5047/A7357, C1287/A10118, C1287/A16563, C5047/A3354, C5047/A10692, C16913/A6135, C5047/A21332 and The National Institute of Health (NIH) Cancer Post-Cancer GWAS initiative grant: No. 1 U19 CA148537-01 (the GAME-ON initiative). We also thank the following for funding support: The Institute of Cancer Research and The Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK (now Prostate Action), The Orchid Cancer Appeal, The National Cancer Research Network UK, and The National Cancer Research Institute (NCRI) UK. We are grateful for support of NIHR funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. The Prostate Cancer Program of Cancer Council Victoria also acknowledge grant support from The National Health and Medical Research Council, Australia (126402, 209057, 251533, 396414, 450104, 504700, 504702, 504715, 623204, 940394, and 614296), VicHealth, Cancer Council Victoria, The Prostate Cancer Foundation of Australia, The Whitten Foundation, PricewaterhouseCoopers, and Tattersall's. E.A.O., D.M.K., and E.M.K. acknowledge the Intramural Program of the National Human Genome Research Institute for their support. The BPC3 was supported by the U.S. National Institutes of Health, National Cancer Institute (cooperative agreements U01-CA98233 to D.J.H., U01-CA98710 to S.M.G., U01-CA98216 to E.R., and U01-CA98758 to B.E.H., and Intramural Research Program of NIH/National Cancer Institute, Division of Cancer Epidemiology and Genetics). CAPS GWAS study was supported by the Swedish Cancer Foundation (grant no 09-0677, 11-484, 12-823), the Cancer Risk Prediction Center (CRisP; www.crispcenter.org), a Linneus Centre (Contract ID 70867902) financed by the Swedish Research Council, Swedish Research Council (grant no K2010-70 × -20430-04-3, 2014-2269). The Hannover Prostate Cancer Study was supported by the Lower Saxonian Cancer Society. PEGASUS was supported by the Intramural Research Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health. RAPPER was supported by the NIHR Manchester Biomedical Research Center, Cancer Research UK (C147/A25254, C1094/A18504) and the EU's 7th Framework Programme Grant/Agreement no 60186. Overall: this research has been conducted using the UK Biobank Resource (application number 16549). NHS is supported by UM1 CA186107 (NHS cohort infrastructure grant), P01 CA87969, and R01 CA49449. NHSII is supported by UM1 CA176726 (NHSII cohort infrastructure grant),

and R01-CA67262. A.L.K. is supported by R01 MH107649. We would like to thank the participants and staff of the NHS and NHSII for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data.

Author Contribution

All authors reviewed and commented on the manuscript, as well as approved the submission. Writing group: X.J., H.K.F., F.R.S., S.L.S., J.P.T., Y. Han., K. Michailidou, C.L., K.B.K., J.D., D.V.C., G. Casey, M.M.G., J. Huyghe, D. Thomas, R.J. Hung, B.D., J.M., U.P., L.H., M. Garcia-Closas, R.A.E., G. Chenevix-Trench, P.J.B., C.A.H., J. Schleutker, D.F.E., S.B.G., P.D.P., A.L.P., B.P., C.I.A., P.K., S. Lindström. Interpret results: A.A., I.L.A., A.C.A., N.N.A., S.A., B.K.A., R.B.B., J. Batra, A.B., S.I.B., S.A.B., C.B., S.E.B., M.K.B., J. Benitez, R.B., G. Cadoni, T.C., P.T.C., G. Cancel-Tassin, D.C., A.T.C., J. Chang-Claude, D.C.C., J.M. Collee, F.J.C., A.C., J.M. Cunningham, C. Chen, M.B.D., P.D., O.D., J.L.D., T.D., E.D., C.K.E., D.V.E., D.G.E., P.A.F., R.T.F., F.F., S.F., E.F., J. Garber, S.A.G., G.G.G., D.E.G., M.T.G., G.G., K.G., P.G., U.H., J. Huyghe, F.H., R. Herrero, P. Hall, M.H., D.G.H., G.I., E.N.I., C.I., P.J., M.A. Jakubowska, A. Joshi, L.E.K., T.K., E.K., A.S.K., L.A.K., J. Kim, M.K., V.N.K., P.L., N.D.L., F. Loupakis, H.L., G. Liu, D. Lambrechts, D.A.L., C.I.L., A.L., N.M.L., G. Leslie, J. Lester, L. Maehle, C.M., L.L.M., S.M., L. McGuffog, A. Mannermaa, P.M., F.M., M.M., V.M., K.B.M., L. Mucci, K. Muir, F.C.N., B.G.N., R.L.N., K.O., O.I.O., H.O., A.O., H.P., J.Y.P., M.T.P., T.P., C.M.P., A.I.P., D. Plaseska-Karanfilska, M.P., K. Stefansson, S.J.R., L.R., H.S.R., M.J. Riggan, M.A.R., K.D.R., E.S., E.J.S., M.B.S., M.K.S., V.W.S., E.M.S., M.L.S., K.D.S., M.C.S., A.B.S., V.L.S., S.S., J. Stone, K. Sundfeldt, A.T., J.A.T., M.R.T., M.B.T., K.L.T., L.B., A.E.T., P.A.T., R.C.T., N.T., C.V., A.V., Q.W., S.W., J.N.W., E.W., A.S.W., F.W., R.W., X.W., D.Y., W.Z., A.Z., J.M.L. Oversee consortium dataset: F.R.S., M.K.B., J.M.L., R. Saxena, R.J. Hung, U.P., R.A.E., G. Chenevix-Trench, D.F.E., S.B.G., C.I.A., P.K. Develop and review the analysis plan and statistical analysis: X.J., H.K.F., A.L.P., B.P., C.I.A., P.K., S. Lindström. Design and manage individual study: D.A., M.C.A., I.L.A., H.A.-C., N.N.A., K.J.A., E.V.B., D.D.B., J. Brenton, M.W.B., S. Benlloch, H. Bickeboller, S. Boccia, N.V.B., H. Brauch, H. Brenner, J. Brunet, M.B., H. Brunnstrom, D.R.B., B.B., M.A.C., I.C., L.C., N.C., A.L.C., J. Clements, S.J.C., C. Cybulski, K.B.C., F.C., J. Chang-Claude, M.C.C., K.C., A.d., M. Gago-Dominguez, J.L.D., A.M.D., M.D., D.M.E., C.E., R.L.F., T.L., J.C.F., O.F., S.J.G., P.A.G., J.A.G., G.G.G., A.K.G., M.S.G., E.L.G., M.T.G., K.G., M.H.G., H.G., J. Gronwald, N.H., P. Hillemanns, F.C.H., R.J. Hamilton, J. Hampe, A.H., E.H., Y. Hong, J.L.H., R. Houlston, P.J.H., D.J.H., E.N.I., S.A.I., A.J., P.J., Mattias. Johansson, Mikael. Johansson, E.M.J., R.K., B.Y.K., K.K., S.K.K., J.A.K., Z.K., S.K., J. Kupryjanczyk, M.L., S. Lam, D. Lessel, M.T.L., E.L., J. Lubinski, L.L., F. Lejbkowicz, J. Lubinski, A. Meindl, T.M., P.M., A. Miller, R.L.M., R.J.M., M.M., A.M.M., K.L.N., D.E.N., A.R.N., S.L.N., H.N., P.A.N., L.F.N., L.N., K.O., E.O., A.A.A.O., O.I.O., A.F.O., H.P., J.Y.P., N.P., K.L.P., W.H.P., C.M.P., D. Prokofieva, P.R., G.R., E.J.v.R., H.A.R., A.R., M.J. Roobol, B.R., K.D.R., D.P.S., J. Simard, V.W.S., H.S., E.M.S., W.S., C.F.S., J.L.S., V.L.S., R. Sutphen, A.J.S., E.H.T., C.M.T., M.D.T., S.N.T., M. Thomassen, M. Tischkowitz, D. Torres, S.S.T., C.M.U., N.U., E.V.N., M.E.A., P.M.W., C.R.W., S.W., M.C.W., C.W., H.W., F.W., A.W., P.W., M.W., A.H.W., X.W., S.Z., K.K.Z., R. Saxena.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-08054-4.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019

Xia Jiang[1,2], Hilary K. Finucane[3,4], Fredrick R. Schumacher[5,6], Stephanie L. Schmit[7,8], Jonathan P. Tyrer[9], Younghun Han[10], Kyriaki Michailidou[11,12], Corina Lesseur[13,14], Karoline B. Kuchenbaecker[15,16], Joe Dennis[11], David V. Conti[17], Graham Casey[18,19], Mia M. Gaudet[20], Jeroen R. Huyghe[21], Demetrius Albanes[22], Melinda C. Aldrich[23], Angeline S. Andrew[24], Irene L. Andrulis[25,26], Hoda Anton-Culver[27], Antonis C. Antoniou[11], Natalia N. Antonenkova[28], Susanne M. Arnold[29], Kristan J. Aronson[30], Banu K. Arun[31], Elisa V. Bandera[32], Rosa B. Barkardottir[33,34], Daniel R. Barnes[11], Jyotsna Batra[35,36], Matthias W. Beckmann[37], Javier Benitez[38,39], Sara Benlloch[11,40], Andrew Berchuck[41], Sonja I. Berndt[22], Heike Bickeböller[42], Stephanie A. Bien[21,43], Carl Blomqvist[44,45], Stefania Boccia[46,47], Natalia V. Bogdanova[28,48,49], Stig E. Bojesen[50,51,52], Manjeet K. Bolla[11], Hiltrud Brauch[53,54,55], Hermann Brenner[55,56,57], James D. Brenton[58], Mark N. Brook[40], Joan Brunet[59], Hans Brunnström[60,61], Daniel D. Buchanan[62,63,64], Barbara Burwinkel[65,66], Ralf Butzow[67], Gabriella Cadoni[46,47], Trinidad Caldés[68], Maria A. Caligo[69], Ian Campbell[70,71], Peter T. Campbell[20], Géraldine Cancel-Tassin[72,73], Lisa Cannon-Albright[74,75], Daniele Campa[76,77], Neil Caporaso[22], André L. Carvalho[78,79], Andrew T. Chan[80,81], Jenny Chang-Claude[76,82], Stephen J. Chanock[22], Chu Chen[83], David C. Christiani[3], Kathleen B.M. Claes[84], Frank Claessens[85], Judith Clements[35,36], J. Margriet Collée[86], Marcia Cruz Correa[87], Fergus J. Couch[88], Angela Cox[89], Julie M. Cunningham[88], Cezary Cybulski[90], Kamila Czene[91], Mary B. Daly[92], Anna deFazio[93,94], Peter Devilee[95,96], Orland Diez[97], Manuela Gago-Dominguez[98,99], Jenny L. Donovan[100], Thilo Dörk[49], Eric J. Duell[101], Alison M. Dunning[9], Miriam Dwek[102], Diana M. Eccles[103], Christopher K. Edlund[104],

Digna R Velez Edwards[105], Carolina Ellberg[106], D. Gareth Evans[107], Peter A. Fasching[37,108], Robert L. Ferris[109], Triantafillos Liloglou[110], Jane C. Figueiredo[111,112], Olivia Fletcher[113], Renée T. Fortner[76], Florentia Fostira[114], Silvia Franceschi[115], Eitan Friedman[116,117], Steven J. Gallinger[118,119,120], Patricia A. Ganz[121], Judy Garber[122], José A. García-Sáenz[68], Simon A. Gayther[123,124,125], Graham G. Giles[126,127,128], Andrew K. Godwin[129], Mark S. Goldberg[130,131], David E. Goldgar[132], Ellen L. Goode[133], Marc T. Goodman[134,135], Gary Goodman[136], Kjell Grankvist[137], Mark H. Greene[138], Henrik Gronberg[91], Jacek Gronwald[90], Pascal Guénel[139], Niclas Håkansson[140], Per Hall[91,141], Ute Hamann[142], Freddie C. Hamdy[143], Robert J. Hamilton[144], Jochen Hampe[145], Aage Haugen[146], Florian Heitz[147,148], Rolando Herrero[149], Peter Hillemanns[49], Michael Hoffmeister[56], Estrid Høgdall[150,151], Yun-Chul Hong[152], John L. Hopper[127], Richard Houlston[153], Peter J. Hulick[154,155], David J. Hunter[1], David G. Huntsman[156,157,158], Gregory Idos[17], Evgeny N. Imyanitov[159], Sue Ann Ingles[17], Claudine Isaacs[160], Anna Jakubowska[90,161], Paul James[71,162], Mark A. Jenkins[62,127], Mattias Johansson[14], Mikael Johansson[163], Esther M. John[164], Amit D. Joshi[3,165], Radka Kaneva[166], Beth Y. Karlan[167], Linda E. Kelemen[168], Tabea Kühl[169], Kay-Tee Khaw[170], Elza Khusnutdinova[171,172], Adam S. Kibel[173], Lambertus A. Kiemeney[174], Jeri Kim[175], Susanne K. Kjaer[150,176], Julia A. Knight[177,178], Manolis Kogevinas[39,179,180,181], Zsofia Kote-Jarai[40], Stella Koutros[182], Vessela N. Kristensen[183,184,185], Jolanta Kupryjanczyk[186], Martin Lacko[187], Stephan Lam[188], Diether Lambrechts[189,190], Maria Teresa Landi[191], Philip Lazarus[192], Nhu D. Le[193], Eunjung Lee[123], Flavio Lejbkowicz[194], Heinz-Josef Lenz[104], Goska Leslie[11], Davor Lessel[195], Jenny Lester[167], Douglas A. Levine[196,197], Li Li[198,199], Christopher I. Li[200], Annika Lindblom[201], Noralane M. Lindor[202], Geoffrey Liu[203], Fotios Loupakis[204], Jan Lubiński[90], Lovise Maehle[205], Christiane Maier[206], Arto Mannermaa[207,208,209], Loic Le Marchand[210], Sara Margolin[211], Taymaa May[212], Lesley McGuffog[11], Alfons Meindl[213], Pooja Middha[76,214], Austin Miller[215], Roger L. Milne[126,127], Robert J. MacInnis[126,127], Francesmary Modugno[216,217], Marco Montagna[218], Victor Moreno[219], Kirsten B. Moysich[220], Lorelei Mucci[3], Kenneth Muir[221,222], Anna Marie Mulligan[223,224], Katherine L. Nathanson[225], David E. Neal[58,143,226], Andrew R. Ness[227], Susan L. Neuhausen[228], Heli Nevanlinna[229], Polly A. Newcomb[21,43], Lisa F. Newcomb[21,230], Finn Cilius Nielsen[231], Liene Nikitina-Zake[232], Børge G. Nordestgaard[50,51,52], Robert L. Nussbaum[233], Kenneth Offit[234,235], Edith Olah[236], Ali Amin Al Olama[11,237], Olufunmilayo I. Olopade[238], Andrew F. Olshan[239,240], Håkan Olsson[106], Ana Osorio[38,39], Hardev Pandha[241], Jong Y. Park[242], Nora Pashayan[243,244], Michael T. Parsons[245], Tanja Pejovic[246,247], Kathryn L. Penney[81], Wilbert H M. Peters[248], Catherine M. Phelan[242], Amanda I. Phipps[21,249], Dijana Plaseska-Karanfilska[250], Miranda Pring[251], Darya Prokofyeva[171], Paolo Radice[252], Kari Stefansson[253], Susan J. Ramus[254,255], Leon Raskin[256], Gad Rennert[257], Hedy S. Rennert[257], Elizabeth J. van Rensburg[258], Marjorie J. Riggan[41], Harvey A. Risch[259], Angela Risch[260,261,262], Monique J. Roobol[263], Barry S. Rosenstein[264,265], Mary Anne Rossing[83,266], Kim De Ruyck[267], Emmanouil Saloustros[268], Dale P. Sandler[269], Elinor J. Sawyer[270], Matthew B. Schabath[242], Johanna Schleutker[271,272,273], Marjanka K. Schmidt[274,275], V. Wendy Setiawan[276], Hongbing Shen[277], Erin M. Siegel[7], Weiva Sieh[278], Christian F. Singer[279], Martha L. Slattery[280], Karina Dalsgaard Sorensen[281,282], Melissa C. Southey[283,284], Amanda B. Spurdle[245], Janet L. Stanford[21,249], Victoria L. Stevens[20], Sebastian Stintzing[285], Jennifer Stone[127,286], Karin Sundfeldt[287], Rebecca Sutphen[288], Anthony J. Swerdlow[40,289], Eloiza H. Tajara[290,291], Catherine M. Tangen[292], Adonina Tardon[293], Jack A. Taylor[269,294], M. Dawn Teare[295], Manuel R. Teixeira[296,297], Mary Beth Terry[298], Kathryn L. Terry[299,300], Stephen N. Thibodeau[88], Mads Thomassen[301], Line Bjørge[302,303], Marc Tischkowitz[304,305], Amanda E. Toland[306], Diana Torres[142,307], Paul A. Townsend[308], Ruth C. Travis[309], Nadine Tung[310], Shelley S. Tworoger[3,242], Cornelia M. Ulrich[21,311], Nawaid Usmani[312,313], Celine M. Vachon[133],

Els Van Nieuwenhuysen[314], Ana Vega[39,315], Miguel Elías Aguado-Barrera[315], Qin Wang[11], Penelope M. Webb[316], Clarice R. Weinberg[317], Stephanie Weinstein[22], Mark C. Weissler[318], Jeffrey N. Weitzel[319], Catharine M.L. West[320], Emily White[321,322], Alice S. Whittemore[323,324], H-Erich Wichmann[325,326,327], Fredrik Wiklund[91], Robert Winqvist[328,329], Alicja Wolk[140,330], Penella Woll[331], Michael Woods[332], Anna H. Wu[123], Xifeng Wu[333], Drakoulis Yannoukakos[114], Wei Zheng[256], Shanbeh Zienolddiny[146], Argyrios Ziogas[27], Kristin K. Zorn[334], Jacqueline M. Lane[4,335], Richa Saxena[4,335], Duncan Thomas[123], Rayjean J. Hung[177,178], Brenda Diergaarde[336,337], James McKay[338], Ulrike Peters[249], Li Hsu[21], Montserrat García-Closas[22], Rosalind A. Eeles[40,339], Georgia Chenevix-Trench[245], Paul J. Brennan[14], Christopher A. Haiman[17], Jacques Simard[340], Douglas F. Easton[9,11], Stephen B. Gruber[123], Paul D.P. Pharoah[9,11], Alkes L. Price[1,3,4], Bogdan Pasaniuc[341], Christopher I. Amos[342], Peter Kraft[1,3] & Sara Lindström[21,249]

[1]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA 02115, USA. [2]Unit of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Nobels vagen 13, 17177 Stockholm, Sweden. [3]Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA 02115, USA. [4]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, 75 Ames St, Cambridge, MA 02142, USA. [5]Department of Population and Quantitative Health Sciences, Case Western Reserve University, 10900 Eucid Avenue, Cleveland, OH 44106, USA. [6]Seidman Cancer Center, University Hospitals, Cleveland, OH 44106, USA. [7]Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Dr. MRC-CANCONT, Tampa, FL 33612, USA. [8]Department of Gastrointestinal Oncology, H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Dr. MRC-CANCONT, Tampa, FL 33612, USA. [9]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, 2 Worts' Causeway, Cambridge CB1 8RN, UK. [10]Department of Biomedical Data Science, The Geisel School of Medicine at Dartmouth, 1 Medical Center Drive, Lebanon, NH 03756, USA. [11]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, 2 Worts' Causeway, Cambridge CB1 8RN, UK. [12]Department of Electron Microscopy/Molecular Pathology, The Cyprus Institute of Neurology and Genetics, 1683 Nicosia, Cyprus. [13]Genetic Epidemiology Group, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69008 Lyon, France. [14]Section of Genetics, International Agency for Research on Cancer, 150 cours Albert Thomas, 69008 Lyon, France. [15]Division of Psychiatry, University College London, Maple House, 149 Tottenham Court Road, London W1T 7NF, UK. [16]UCL Genetics Institute, University College London, Gower Street, London WC1E 6BT, UK. [17]Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, CA 48109, USA. [18]Public Health Sciences, University of Virginia, P.O. Box 800717 Charlottesville, VI 22908, USA. [19]Center for Public Health Genomics, University of Virginia, P.O. Box 800717 Charlottesville, VI 22908, USA. [20]Epidemiology Research Program, American Cancer Society, 250 Williams Street NW, Atlanta, GA 30303, USA. [21]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., Seattle, WA 98109-1024, USA. [22]Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Dr, Rockville, MD 20850, USA. [23]Department of Thoracic Surgery, Division of Epidemiology, Vanderbilt University Medical Center, 609 Oxford House, Nashville, TN 37232, USA. [24]Department of Neurology, Dartmouth-Hitchcock Medical Center, 7927 Rubin Building, Room 860, One Medical Center Drive, Lebanon, NH 3756, USA. [25]Fred ALitwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, 600 University Avenue, Toronto, ON M5G1X5, Canada. [26]Department of Molecular Genetics, University of Toronto, 1 King's College Circle, Toronto, ON M5S1A8, Canada. [27]Department of Epidemiology, Genetic Epidemiology Research Institute, University of California Irvine, 224 Irvine Hall, Irvine, CA 92617, USA. [28]NNAlexandrov Research Institute of Oncology and Medical Radiology, Settlement of Lesnoy-2, 223040 Minsk, Belarus. [29]Markey Cancer Center, University of Kentucky, 800 Rose Street, cc445, Lexington, KY 40508, USA. [30]Department of Public Health Sciences, and Cancer Research Institute, Queen's University, 10 Stuart Street, Kingston, ON K7L 3N6, Canada. [31]Department of Breast Medical Oncology, University of Texas MD Anderson Cancer Center, 1155 Pressler St, Houston, TX 77030, USA. [32]Cancer Prevention and Control Program, Rutgers Cancer Institute of New Jersey, 195 Little Albany Street, Room 5568, New Brunswick, NJ 08903, USA. [33]Department of Pathology, Landspitali University Hospital, Hringbraut, Reykjavik 101, Iceland. [34]BMC (Biomedical Centre), Faculty of Medicine, University of Iceland, Vatnsmyrarvegi 16, Reykjavik 101, Iceland. [35]Australian Prostate Cancer Research Centre-Qld, Translational Research Institute, 37 Kent St, Woolloongabba, QLD 4102, Australia. [36]Institute of Health and Biomedical Innovation and School of Biomedical Science, Queensland University of Technology, 60 Musk Ave, Kelvin Grove, QLD 4059, Australia. [37]Department of Gynecology and Obstetrics, Comprehensive Cancer Center Erlangen Nuremberg, University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg, Universitaetsstrasse 21-23, 91054 Erlangen, Germany. [38]Human Cancer Genetics Programme, Spanish National Cancer Research Centre (CNIO), Calle de Melchor Fernández Almagro, 3, 28029 Madrid, Spain. [39]Biomedical Network on Rare Diseases (CIBERER), AvMonforte de Lemos, 3-5Pabellón 11Planta 0, 28029 Madrid, Spain. [40]Division of Genetics and Epidemiology, The Institute of Cancer Research, 15 Cotswold Road, London SM2 5NG, UK. [41]Department of Obstetrics and Gynecology, Duke University Medical Center, 25171 Morris Bldg, Durham, NC 27710, USA. [42]Department of Genetic Epidemiology, University Medical Center Goettingen, Humboldtallee 32, 37073 Goettingen, Germany. [43]School of Public Health, University of Washington, 1959 NE Pacific Street, Health Science Buidling, F-350, Seattle, WA 98195, USA. [44]Department of Oncology, Helsinki University Hospital, University of Helsinki, Haartmaninkatu 4, 00290 Helsinki, Finland. [45]Department of Oncology, Örebro University Hospital, 70185 Örebro, Sweden. [46]Fondazione Policlinico Universitario AGemelli IRCCS, 00168 Roma, Italy. [47]Università Cattolica del Sacro Cuore, 00168 Roma, Italy. [48]Department of Radiation Oncology, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany. [49]Gynaecology Research Unit, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany. [50]Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark. [51]Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark. [52]Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark. [53]DrMargarete Fischer-Bosch-Institute of Clinical Pharmacology, Auerbachstr112, 70376 Stuttgart, Germany. [54]University of Tübingen, Geschwister-Scholl-Platz, 72074 Tübingen, Germany. [55]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. [56]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280,

69120 Heidelberg, Germany. [57]Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. [58]Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Robinson Way, CB2 0RE Cambridge, UK. [59]Genetic Counseling Unit, Hereditary Cancer Program, IDIBGI (Institut d'Investigació Biomèdica de Girona), Catalan Institute of Oncology, CIBERONC, AvFrança s/n, 17007 Girona, Spain. [60]Clinical Sciences, Lund University, Box 117, 221 00 Lund, Sweden. [61]Department of Genetics and Pathology, Division of Laboratory Medicine, 221 85 Lund, Sweden. [62]University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, VIC 3010, Australia. [63]Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Parkville, VIC 3010, Australia. [64]Genomic Medicine and Family Cancer Clinic, Royal Melbourne Hospital, Parkville, VIC 3010, Australia. [65]Department of Obstetrics and Gynecology, University of Heidelberg, Im Neuenheimer Feld 440, 69120 Heidelberg, Germany. [66]Molecular Epidemiology Group, C080, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. [67]Department of Pathology, University of Helsinki and Helsinki University Hospital, Biomedicum Helsinki 4th floor, Haartmaninkatu 8, 00029 Helsinki, Finland. [68]Medical Oncology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria San Carlos (IdISSC), Centro Investigación Biomédica en Red de Cáncer (CIBERONC), Calle del Prof Martín Lagos, 28040 Madrid, Spain. [69]Section of Genetic Oncology, Department of Laboratory Medicine, University and University Hospital of Pisa, via Roma 67, 56126 Pisa, Italy. [70]Peter MacCallum Cancer Center, 305 Grattan Street, Melbourne, VIC 3000, Australia. [71]Sir Peter MacCallum Department of Oncology, The University of Melbourne, 305 Grattan Street, Melbourne, VIC 3000, Australia. [72]Sorbonne Université, GRC N°5 ONCOTYPE-URO, Tenon Hospital, 75020 Paris, France. [73]CeRePP, Tenon Hospital, 75020 Paris, France. [74]Division of Genetic Epidemiology, Department of Medicine, University of Utah School of Medicine, Salt Lake City, UT 84112, USA. [75]George EWahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT 84112, USA. [76]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. [77]Department of Biology, University of Pisa, 56126 Pisa, Italy. [78]Molecular Oncology Research Center, Barretos Cancer Hospital, Rua Antenor Duarte Villela, 1331, Barretos, SP 784-400, Brazil. [79]Head and Neck Surgery Department, Barretos Cancer Hospital, Pio XII, 1331, Antenor Duarte Villela St, Barretos, SP 14784-400, Brazil. [80]Division of Gastroenterology, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA. [81]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 181 Longwood Avenue, Boston, MA 02115, USA. [82]Cancer Epidemiology Group, University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany. [83]Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle, WA 98109, USA. [84]Centre for Medical Genetics, Ghent University, De Pintelaan 185, 9000 Gent, Belgium. [85]Molecular Endocrinology Laboratory, Department of Cellular and Molecular Medicine, KU Leuven, Leuven 3000, Belgium. [86]Department of Clinical Genetics, Erasmus University Medical Center, Wytemaweg 80, 3015 Rotterdam, CN, The Netherlands. [87]University of Puerto Rico Medical Sciences Campus and Comprehensive Cancer Center, San Juan, PR 00936, USA. [88]Department of Laboratory Medicine and Pathology, Mayo Clinic, 200 First StSW, Rochester, MN 55905, USA. [89]Sheffield Institute for Nucleic Acids (SInFoNiA), Department of Oncology and Metabolism, University of Sheffield, Western Bank, Sheffield S10 2TN, UK. [90]International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, ulUnii Lubelskiej 1, 71-252 Szczecin, Poland. [91]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Karolinska Univ Hospital, 171 76 Stockholm, Sweden. [92]Department of Clinical Genetics, Fox Chase Cancer Center, 333 Cottman Ave, Philadelphia, PA 19111, USA. [93]Centre for Cancer Research, The Westmead Institute for Medical Research, The University of Sydney, 176 Hawkesbury Rd, Sydney, NSW 2145, Australia. [94]Department of Gynaecological Oncology, Westmead Hospital, Hawkesbury Rd & Darcy Rd, Sydney, NSW 2145, Australia. [95]Department of Pathology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands. [96]Department of Human Genetics, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands. [97]Oncogenetics Group, Clinical and Molecular Genetics Area, Vall d'Hebron Institute of Oncology (VHIO), University Hospital, Vall d'Hebron, Passeig de la Vall d'Hebron 119-129, 08035 Barcelona, Spain. [98]Genomic Medicine Group, Galician Foundation of Genomic Medicine, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Complejo Hospitalario Universitario de Santiago, SERGAS, Travesía da Choupana S/N, 15706 Santiago de Compostela, Spain. [99]Moores Cancer Center, University of California, San Diego, 3855 Health Sciences Drive, La Jolla, CA 92037, USA. [100]School of Social and Community Medicine, University of Bristol, Bristol BS8 1TH, UK. [101]Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology (ICO-IDIBELL), AvGran Via 199-203, L'Hospitalet de Llobregat, 08908 Barcelona, Spain. [102]Department of Biomedical Sciences, Faculty of Science and Technology, University of Westminster, 309 Regent Street, London W1B 2HW, UK. [103]Cancer Sciences Academic Unit, Faculty of Medicine, University of Southampton, Tremona Road, Southampton SO16 6YD, UK. [104]Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA. [105]Vanderbilt Epidemiology Center, Vanderbilt Genetics Institute, Department of Obstetrics and Gynecology, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 600, Nashville, TN 37203, USA. [106]Department of Cancer Epidemiology, Clinical Sciences, Lund University, Barngatan 4, Skånes universitetssjukhus, 222 42 Lund, Sweden. [107]Manchester Centre for Genomic Medicine, Division of Evolution and Genomic Sciences, University of Manchester, St Mary's Hospital, Central Manchester University Hospitals NHS Foundation Trust, Oxford Road, Manchester M13 9WL, UK. [108]David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, 10833 Le Conte Ave, Los Angeles, CA 90095, USA. [109]Department of Otolaryngology, UPMC Hillman Cancer Center, Cancer Pavilion, University of Pittsburgh, Suite 500, 5150 Centre Avenue, Pittsburgh, PA 15232, USA. [110]Molecular and Clinical Cancer Medicine, Roy Castle Lung Cancer Research Programme, The University of Liverpool Institute of Translational Medicine, The Wiliam Duncan Building, 6 West Derby Street, Liverpool L69 3BX, UK. [111]Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, CA 90048, USA. [112]Keck School of Medicine, University of Southern California, 1450 Biggy Street, Los Angeles, CA 90033, USA. [113]The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, 123 Old Brompton Road, London SW7 3RP, UK. [114]Molecular Diagnostics Laboratory, INRASTES, National Centre for Scientific Research 'Demokritos', Neapoleos 10, AgParaskevi, Athens 15310, Greece. [115]Section of Infections, International Agency for Research on Cancer, 150 cours Albert Thomas, 69008 Lyon, France. [116]The Susanne Levy Gertner Oncogenetics Unit, Chaim Sheba Medical Center, Emek HaEla St 1, 52621 Ramat Gan, Israel. [117]Sackler Faculty of Medicine, Tel Aviv University, Haim Levanon 30, 69978 Ramat Aviv, Israel. [118]Department of Surgery, Mount Sinai Hospital, 600 University Avenue, Toronto, ON M5G 1X5, Canada. [119]Samuel Lunenfeld Research Institute, 600 University Avenue, Toronto, ON M5G 1X5, Canada. [120]University Health Network Toronto General Hospital, 200 Elizabeth St, Toronto, ON M5G 2C4, Canada. [121]Schools of Medicine and Public Health, Division of Cancer Prevention & Control Research, Jonsson Comprehensive Cancer Centre, UCLA, 650 Charles Young Drive South, Los Angeles, CA 90095-6900, USA. [122]Cancer Risk and Prevention Clinic, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215, USA. [123]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, 1975 Zonal Ave, Los Angeles, CA 90033, USA. [124]Center for Cancer Prevention and Translational Genomics, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Spielberg Building, 8725 Alden Dr, Los Angeles, CA 90048, USA. [125]Department of Biomedical Sciences, Cedars-Sinai Medical Center, Spielberg Building, 8725 Alden Dr, Los Angeles, CA 90048, USA. [126]Cancer Epidemiology & Intelligence Division, Cancer Council Victoria, 615 St Kilda Road, Melbourne, VIC 3004, Australia. [127]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Level 1, 723 Swanston Street, Melbourne, VIC 3010, Australia. [128]Department

of Epidemiology and Preventive Medicine, Monash University, Melbourne, VIC, Australia. [129]Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, 3901 Rainbow Blvd, Kansas City, KS 66160, USA. [130]Department of Medicine, McGill University, 1001 Decarie Boulevard, Montréal, QC H4A3J1, Canada. [131]Division of Clinical Epidemiology, Royal Victoria Hospital, McGill University, 1001 Decarie Boulevard, Montréal, QC H4A3J1, Canada. [132]Department of Dermatology, Huntsman Cancer Institute, University of Utah School of Medicine, 2000 Circle of Hope, Salt Lake City, UT 84112, USA. [133]Department of Health Sciences Research, Mayo Clinic, 200 First StSW, Rochester, MN 55905, USA. [134]Cancer Prevention and Control, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, 8700 Beverly Blvd, Room 1S37, Los Angeles, CA 90048, USA. [135]Community and Population Health Research Institute, Department of Biomedical Sciences, Cedars-Sinai Medical Center, 8700 Beverly Blvd, Room 1S37, Los Angeles, CA 90048, USA. [136]Public Health Sciences Division, Swedish Cancer Institute, 1221 Madison StSte 300, Seattle, WA 98109, USA. [137]Unit of Clinical Chemistry, Department of Medical Biosciences, Umeå University, By 6M van 2, Sjukhusomradet, Umea universitet, 901 85 Umea, Sweden. [138]Clinical Genetics Branch, National Cancer Institute, DCEG, 9609 Medical Center Dr, Bethesda, MD 20850-9772, USA. [139]Cancer & Environment Group, Center for Research in Epidemiology and Population Health (CESP), INSERM, University Paris-Sud, University Paris-Saclay, 94805 Villejuif, France. [140]Department of Environmental Medicine, Division of Nutritional Epidemiology, Karolinska Institutet, Nobels väg 13, SE-171 77, SE-171 Stockholm, Sweden. [141]Department of Oncology, Södersjukhuset, Sjukhusbacken 10, 118 83 Stockholm, Sweden. [142]Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany. [143]Nuffield Department of Surgical Sciences, Faculty of Medical Science, John Radcliffe Hospital, University of Oxford, Oxford OX1 2JD, UK. [144]Department of Surgical Oncology, Princess Margaret Cancer Centre, 610 University Avenue, Toronto, Ontario M5G2M9, Canada. [145]Department of Internal Medicine 1, University Hospital Dresden, Technische Universität Dresden (TU Dresden), 01307 Dresden, Germany. [146]National Institute of Occupational Health (STAMI), Gydas vei 8, 0033 Oslo, Norway. [147]Department of Gynecology and Gynecologic Oncology, DrHorst Schmidt Kliniken Wiesbaden, Ludwig-Erhard-Straße 100, 65199 Wiesbaden, Germany. [148]Department of Gynecology and Gynecologic Oncology, Kliniken Essen-Mitte/ EvangHuyssens-Stiftung/ Knappschaft GmbH, Henricistrasse 92, 45136 Essen, Germany. [149]Early Detection and Prevention Section, International Agency for Research on Cancer, 150 cours Albert Thomas, 69008 Lyon, France. [150]Department of Virus, Lifestyle and Genes, Danish Cancer Society Research Center, Strandboulevarden 49, DK-2100 Copenhagen, Denmark. [151]Molecular Unit, Department of Pathology, Herlev Hospital, University of Copenhagen, Herlev Ringvej 75, DK-2730 Herlev, Denmark. [152]Preventive Medicine, Seoul National University College of Medicine, 1 Gwanak-ro, Gwanak-gu, Seoul 151 742, Korea. [153]German Research Center for Environmental Health, Institute for Cancer Research, Ingolstadter Landstr1, London SM2 5NG, UK. [154]Center for Medical Genetics, NorthShore University HealthSystem, 1000 Central St, Evanston, IL 60201, USA. [155]The University of Chicago Pritzker School of Medicine, 924 E 57th St, Chicago, IL 60637, USA. [156]British Columbia's Ovarian Cancer Research (OVCARE) Program, Vancouver General Hospital, BC Cancer Agency and University of British Columbia, #3427-600 West 10th Avenue, Vancouver, BC V5Z 4E6, Canada. [157]Department of Molecular Oncology, BC Cancer Agency Research Centre, #3427-600 West 10th Avenue, Vancouver, BC V5Z 4E6, Canada. [158]Department of Pathology and Laboratory Medicine, University of British Columbia, #3427-600 West 10th Avenue, Vancouver, BC V5Z 4E6, Canada. [159]NNPetrov Institute of Oncology, Leningradskaya ul, 68, StPetersburg, Russia 197758. [160]Lombardi Comprehensive Cancer Center, Georgetown University, 3800 Reservoir Road, Washington, DC 20007, USA. [161]Independent Laboratory of Molecular Biology and Genetic Diagnostics, Pomeranian Medical University, Rybacka 1, 70-204 Szczecin, Poland. [162]Parkville Familial Cancer Centre, Peter MacCallum Cancer Center, 305 Grattan Street, Melbourne, VIC 3000, Australia. [163]Department of Radiation Sciences, Umeå University, By 6M van 2, Sjukhusomradet, Umea universitet, 901 85 Umea, Sweden. [164]Department of Medicine, Division of Oncology and Stanford Cancer Institute, Stanford University School of Medicine, 780 Welch Rd, Stanford, CA 94304, USA. [165]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, 02114 Boston, MA, USA. [166]Molecular Medicine Center, Department of Medical Chemistry and Biochemistry, Medical Faculty, Medical University of Sofia, Sofia 1504, Bulgaria. [167]Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, CA 90048, USA. [168]Hollings Cancer Center and Department of Public Health Sciences, Medical University of South Carolina, 68 President Street Bioengineering Building, MSC955, Charleston, SC 29425, USA. [169]Cancer Epidemiology, University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany. [170]Clinical Gerontology, Department of Public Health and Primary Care, University of Cambridge, 2 Worts' Causeway, Cambridge CB1 8RN, UK. [171]Department of Genetics and Fundamental Medicine, Bashkir State University, ulZaki Validi 32, Ufa, Russia 450076. [172]Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, 71 prosp Oktyabrya, Ufa, Russia 450054. [173]Division of Urologic Surgery, Brigham and Womens Hospital, Boston, Massachusettes 02115, USA. [174]Radboud Institute for Health Sciences, Radboud University Medical Center, Geert Grooteplein 21, 6525 EZ Nijmegen, The Netherlands. [175]Department of Genitourinary Medical Oncology, University of Texas MD Anderson Cancer Center, 1155 Pressler St, Houston, TX 77030, USA. [176]Department of Gynaecology, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, DK-2100 Copenhagen, Denmark. [177]Prosserman Centre for Population Health Research, Lunenfeld-Tanenbaum Research Institute, Sinai Health System, 60 Murray Street, Toronto, Ontario M5T 3L9, Canada. [178]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, ON M5T3M7, Canada. [179]Centre for Research in Environmental Epidemiology (CREAL), ISGlobal, 08036 Barcelona, Spain. [180]IMIM (Hospital del Mar Research Institute), Barcelona 08003, Spain. [181]Universitat Pompeu Fabra (UPF), Barcelona 08002, Spain. [182]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Department of Health and Human Services, National Institutes of Health, 9609 Medical Center Dr, Bethesda, MD 20892, USA. [183]Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Ullernchausseen 70, 0379 Oslo, Norway. [184]Faculty of Medicine, Institute of Clinical Medicine, University of Oslo, Kirkeveien 166, 0450 Oslo, Norway. [185]Department of Clinical Molecular Biology, Oslo University Hospital, University of Oslo, Kirkeveien 166, 0450 Oslo, Norway. [186]Department of Pathology and Laboratory Diagnostics, the Maria Sklodowska-Curie Institute - Oncology Center, Roentgena 5, 02-781 Warsaw, Poland. [187]Head and Neck Surgery, Department of Otorhinolaryngology, Maastricht University Medical Center, PDebyelaan 25, POBox 5800, 6202 AZ Maastricht, The Netherlands. [188]Department of Integrative Oncology, British Columbia Cancer Agency, Room 10-111 675 West 10th Avenue, Vancouver, BC V5Z1L3, Canada. [189]VIB Center for Cancer Biology, VIB, Herestraat 49, 3001 Leuven, Belgium. [190]Laboratory for Translational Genetics, Department of Human Genetics, University of Leuven, Oude Markt 13, 3000 Leuven, Belgium. [191]Integrative Tumor Epidemiology Branch, DCEG, National Cancer Institute, 9609 Medical Center Drive, Room SG/7E106, Rockville, MD 20850, USA. [192]College of Pharmacy, Washington State University, PBS 431 PO Box 1495, Spokane, WA 99210-1495, USA. [193]Cancer Control Research, BC Cancer Agency, 675 West 10th Avenue, Vancouver, BC V5Z 1L3, Canada. [194]Clalit Health Services, Clalit National Israeli Cancer Control Center, Carmel Medical Center, 2 Horev Street, 3436212 Haifa, Israel. [195]Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany. [196]Gynecology Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. [197]Gynecologic Oncology, Laura and Isaac Pearlmutter Cancer Center, NYU Langone Medical Center, 240 East 38th Street 19th Floor, New York, NY 10016, USA. [198]Department of Family Medicine and Community Health, Mary Ann Swetland Center for Environmental Health, Case Western Reserve University, Cleveland, OH 44106, USA. [199]Servicio Galego de Saude (SERGAS), Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), 15706 Santiago De Compostela, Spain. [200]Translational Research Program, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA. [201]Department of

Molecular Medicine and Surgery, Karolinska Institutet, Karolinska Univ Hospital, 171 76 Stockholm, Sweden. [202]Health Sciences Research, Mayo Clinic Arizona, 13400 EShea Blvd, Scottsdale, AZ 85259, USA. [203]Epidemiology Division, Princess Margaret Cancer Centre, 610 University Avenue, Toronto, ON M5G2M9, Canada. [204]Unit of Oncology 1, Department of Clinical and Experimental Oncology, Istituto Oncologico Veneto IRCCS, 35122 Padua, Italy. [205]Department of Medical Genetics, Oslo University Hospital, Kirkeveien 166, 0450 Oslo, Norway. [206]Institute of Human Genetics, University Hospital Ulm, Prittwitzstrasse 43, 89075 Ulm, Germany. [207]Translational Cancer Research Area, University of Eastern Finland, Yliopistonranta 1, 70210 Kuopio, Finland. [208]Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, KuopioYliopistonranta 1, 70210, Finland. [209]Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Puijonlaaksontie 2, 70210 Kuopio, Finland. [210]Epidemiology Program, University of Hawaii Cancer Center, 701 Ilalo St, Honolulu, HI 96813, USA. [211]Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet, Stockholm 17177, Sweden. [212]Division of Gynecologic Oncology, University Health Network, Princess Margaret Hospital, 610 University Avenue, OPG Wing, 6-811, Toronto, ON M5G 2M9, Canada. [213]Division of Gynaecology and Obstetrics, , Technische Universität München, Arcisstraße 21, 80333 Munich, Germany. [214]Faculty of Medicine, University of Heidelberg, In Neuenheimer Feld 672, 69120 Heidelberg, Germany. [215]NRG Oncology, Statistics and Data Management Center, Roswell Park Cancer Institute, Elm & Carlton Streets, Buffalo, NY 14263, USA. [216]Womens Cancer Research Center, Magee-Womens Research Institute and Hillman Cancer Center, Pittsburgh, PA 15213, USA. [217]Division of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine, 300 Halket Street, Pittsburgh, PA 15213, USA. [218]Immunology and Molecular Oncology Unit, Veneto Institute of Oncology IOV - IRCCS, Via Gattamelata 64, Padua 35128, Italy. [219]Catalan Institute of Oncology, Bellvitge Biomedical Research Institute (IDIBELL), Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP) and University of Barcelona, Barcelona 08908, Spain. [220]Division of Cancer Prevention and Control, Roswell Park Cancer Institute, Elm & Carlton Streets, Buffalo, NY 14263, USA. [221]Division of Population Health, Health Services Research and Primary Care, University of Manchester, Oxford Road, Manchester M13 9PL, UK. [222]Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK. [223]Department of Laboratory Medicine and Pathobiology, University of Toronto, 1 King's College Circle, Toronto, ON M5S1A8, Canada. [224]Laboratory Medicine Program, University Health Network, 200 Elizabeth Street, Toronto, ON M5G2C4, Canada. [225]Department of Medicine, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, 3400 Civic Center Boulevard, Philadelphia, PA 19104, USA. [226]Department of Oncology, Addenbrooke's Hospital, University of Cambridge, Cambridge CB1 8RN, UK. [227]NIHR Bristol Biomedical Research Centre Nutrition Theme, University of Bristol, Upper Maudlin Street, Bristol BS2 8AE, UK. [228]Department of Population Sciences, Beckman Research Institute of City of Hope, 1500 E Duarte, Duarte, CA 91010, USA. [229]Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, Haartmaninkatu 8, 00290 Helsinki, Finland. [230]Department of Urology, University of Washington, Seattle, Washington 98195, USA. [231]Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, DK-2100 Copenhagen, Denmark. [232]Latvian Biomedical Research and Study Centre, Ratsupites str 1, Riga LV-1067, Latvia. [233]Cancer Genetics and Prevention Program, University of California San Francisco, 1600 Divisadero St, San Francisco, CA 94143-1714, USA. [234]Clinical Genetics Research Lab, Department of Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. [235]Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. [236]Department of Molecular Genetics, National Institute of Oncology, Ráth György u7-9, 1122 Budapest, Hungary. [237]Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0QQ, UK. [238]Center for Clinical Cancer Genetics, The University of Chicago, 5841S Maryland Ave, Chicago, IL 60637, USA. [239]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, 135 Dauer Dr, Chapel Hill, NC 27599-7435, USA. [240]UNC Lineberger Comprehensive Cancer Center, 450 West Dr, Chapell Hill, NC 27599, USA. [241]The University of Surrey, Guildford, Surrey GU2 7XH, UK. [242]Department of Cancer Epidemiology, HLee Moffitt Cancer Center and Research Institute, 12902 Magnolia Drive, Tampa, FL 33612, USA. [243]Department of Applied Health Research, University College London, 1-19 Torrington Place, London WC1E 6BT, UK. [244]Centre for Cancer Genetic Epidemiology, Department of Oncology, Strangeways Laboratory, University of Cambridge, Cambridge CB1 8RN, UK. [245]Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, 300 Herston Road, Brisbane, QLD 4006, Australia. [246]Department of Obstetrics and Gynecology, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, L-466, Portland, OR 97239, USA. [247]Knight Cancer Institute, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, L-466, Portland, OR 97239, USA. [248]Department of Gastroenterology, Radboud University Nijmegen Medical Center, Geert Grooteplein Zuid 10, Internal BOBox 433, 6525 GA Nijmegen, The Netherlands. [249]Department of Epidemiology, University of Washington School of Public Health, 1959 NE Pacific St, Seattle, WA 98195, USA. [250]Research Centre for Genetic Engineering and Biotechnology 'Georgi DEfremov', Macedonian Academy of Sciences and Arts, Boulevard Krste Petkov Misirkov, 1000 Skopje, Republic of Macedonia. [251]Bristol Dental School, University of Bristol, Lower Maudlin Street, Bristol BS1 2LY, UK. [252]Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Research, Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale dei Tumori (INT), Via Giacomo Venezian 1, 20133 Milan, Italy. [253]Decode genetics, Sturlugata 8, IS-101 ReykjavikReykjavikIceland, Iceland. [254]School of Women's and Children's Health, Faculty of Medicine, University of NSW Sydney, 18 High St, Sydney, NSW 2052, Australia. [255]The Kinghorn Cancer Centre, Garvan Institute of Medical Research, 384 Victoria Street, Sydney, NSW 2010, Australia. [256]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, 1161 21st Ave S # D3300, Nashville, TN 37232, USA. [257]Clalit National Cancer Control Center, Carmel Medical Center and Technion Faculty of Medicine, 7 Michal Street, 34362 Haifa, Israel. [258]Department of Genetics, University of Pretoria, Private Bag X323, Arcadia 0007, South Africa. [259]Department of Chronic Disease Epidemiology, Yale School of Public Health, 60 College St, New Haven, CT 06510, USA. [260]Cancer Center Cluster Salzburg at PLUS, Department of Molecular Biology, University of Salzburg, Billrothstr11, 5020 Salzburg, Austria. [261]Division of Epigenomics and Cancer Risk Factors, DKFZ – German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. [262]Member of the German Center for Lung Research (DZL), Translational Lung Research Center Heidelberg (TLRC-H), 69120 Heidelberg, Germany. [263]Department of Urology, Erasmus University Medical Center, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands. [264]Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA. [265]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, USA. [266]Department of Epidemiology, University of Washington, M4 C308, 1100 Fairview Ave N, Seattle, WA 98109, USA. [267]Faculty of Medicine and Health Sciences, Basic Medical Sciences, Ghent University, De Pintelaan 185, 9000 Gent, Belgium. [268]Hereditary Cancer Clinic, University Hospital of Heraklion, Voutes, 711 10 Heraklion, Greece. [269]Epidemiology Branch, National Institute of Environmental Health Sciences, NIH, 111TWAlexander Drive, Research Triangle Park, NC 27709, USA. [270]Research Oncology, Guy's Hospital, King's College London, Guy's Hospital Great Maze Pond, London SE1 9RT, UK. [271]Institute of Biomedicine, University of Turku, 20014 Turku, Finland. [272]Division of Laboratory, Department of Medical Genetics, Turku University Hospital, 20014 Turku, Finland. [273]Prostate Cancer Research Center, Faculty of Medicine and Life Sciences and BioMediTech Institute, University of Tampere, 33014 Tampere, Finland. [274]Division of Molecular Pathology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. [275]Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. [276]Department of Preventive Medicine, Keck School of Medicine, University of Southern

California, 1450 Biggy Street, Los Angeles, CA 90033, USA. [277]Department of Epidemiology and Biostatistics, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, 101 Longmian Ave, Jiangning District, 211166 Nanjing, People's Republic of China. [278]Department of Genetics and Genomic Sciences, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, 2nd floor, New York, NY 10029, USA. [279]Dept of OB/GYN and Comprehensive Cancer Center, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria. [280]Department of Internal Medicine, University of Utah Health Sciences Center, 295 Chipeta Way, Salt Lake City, UT 84132, USA. [281]Department of Molecular Medicine, Aarhus University Hospital, DK-8200 Aarhus, Denmark. [282]Department of Clinical Medicine, Aarhus University, DK-8200 Aarhus, Denmark. [283]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, 246 Clayton Road, Clayton, VIC 3168, Australia. [284]Department of Clinical Pathology, The University of Melbourne, Cnr Grattan Street and Royal Parade, Melbourne, VIC 3010, Australia. [285]Department of Medicine III, University Hospital, LMU Munich, Marchioninistr15, 81377 Munich, Germany. [286]The Curtin UWA Centre for Genetic Origins of Health and Disease, Curtin University and University of Western Australia, 35 Stirling Hwy, Perth, WA 6000, Australia. [287]Department of Obstetrics and Gynecology, Sahlgrenska Cancer Center, Inst Clinical Scienses, University of Gothenburg, Blå stråket 6, 41345 Gothenburg, Sweden. [288]Epidemiology Center, College of Medicine, University of South Florida, 3650 Spectrum Blvd, Suite 100, Tampa, FL 33612, USA. [289]Division of Breast Cancer Research, The Institute of Cancer Research, London SW7 3RP, UK. [290]Department of Molecular Biology, School of Medicine of São José do Rio Preto, Av Brig Faria Lima 5416 Vila São Pedro, São José do Rio Preto, SP 15090-000, Brazil. [291]Department of Genetics and Evolutive Biology, Institute of Biosciences, University of São Paulo, Rua do Matão, 321, São Paulo, SP 05508-090, Brazil. [292]SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA. [293]Faculty of Medicine, University of Oviedo and CIBERESP, Campus del Cristo s/n, 33006 Oviedo, Spain. [294]Epigenetic and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, NIH, 111TWAlexander Drive, Research Triangle Park, NC 27709, USA. [295]Medical Statistics Group, School of Health and Related Research (ScHARR), University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK. [296]Department of Genetics, Portuguese Oncology Institute, Rua DrAntónio Bernardino de Almeida 62, 4220-072 Porto, Portugal. [297]Biomedical Sciences Institute (ICBAS), University of Porto, RJorge de Viterbo Ferreira 228, 4050-013 Porto, Portugal. [298]Department of Epidemiology, Mailman School of Public Health, Columbia University, 722 West 168th Street, New York, NY 10032, USA. [299]Obstetrics and Gynecology Epidemiology Center, Brigham and Women's Hospital, 221 Longwood Avenue RFB 368, Boston, MA 02115, USA. [300]Harvard THChan School of Public Health, 221 Longwood Avenue RFB 368, Boston, MA 02115, USA. [301]Department of Clinical Genetics, Odense University Hospital, Sonder Boulevard 29, 5000 Odence C, Denmark. [302]Department of Gynecology and Obstetrics, Haukeland University Hospital, 5021 Bergen, Norway. [303]Centre for Cancer Biomarkers CCBIO, Department of Clinical Science, University of Bergen, 5021 Bergen, Norway. [304]Program in Cancer Genetics, Departments of Human Genetics and Oncology, McGill University, 1001 Decarie Boulevard, Montréal, QC H4A3J1, Canada. [305]Department of Medical Genetics, Cambridge University, Hills Road, Cambridge CB2 0QQ, UK. [306]Department of Cancer Biology and Genetics, The Ohio State University, 460W12th Avenue, Columbus, OH 43210, USA. [307]Institute of Human Genetics, Pontificia Universidad Javeriana, Carrera 7 No40-90, Bogota, Colombia. [308]Division of Cancer Sciences, Manchester Cancer Research Centre, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, NIHR Manchester Biomedical Research Centre, Health Innovation Manchester, University of Manchester, Manchester M20 4GJ, UK. [309]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK. [310]Department of Medical Oncology, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215, USA. [311]Huntsman Cancer Institute and Department of Population Health Sciences, University of Utah, 2000 Circle of Hope, Rm 4125, Salt Lake City, UT 84112, USA. [312]Department of Oncology, Cross Cancer Institute, University of Alberta, 116 St & 85 Ave, Edmonton AB T6G 2R3, Canada. [313]Division of Radiation Oncology, Cross Cancer Institute, University of Alberta, 116 St & 85 Ave, Edmonton AB T6G 2R3, Canada. [314]Division of Gynecologic Oncology, Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospitals Leuven, Herestraat 49, 3000 Leuven, Belgium. [315]Fundación Pública Galega Medicina Xenómica & Instituto de Investigación Sanitaria de Santiago de Compostela, calle Choupana s/n, 15706 Santiago De Compostela, Spain. [316]Population Health Department, QIMR Berghofer Medical Research Institute, 300 Herston Road, Brisbane, QLD 4006, Australia. [317]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, NIH, 111TWAlexander Drive, Research Triangle Park, NC 27709, USA. [318]Department of Otolaryngology/Head and Neck Surgery, University of North Carolina at Chapel Hill, Chapel Hill 27514 NC, USA. [319]City of Hope Clinical Cancer Genomics Community Research Network, 1500 East Duarte Road, Duarte, CA 91010, USA. [320]Division of Cancer Sciences, University of Manchester, Manchester Cancer Research Centre, Manchester Academic Health Science Centre,, The Christie Hospital NHS Foundation Trust, Manchester M13 9PL, UK. [321]Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle, WA 98109, USA. [322]Department of Epidemiology, University of Washington, 1100 Fairview Ave N, Seattle, WA 98109, USA. [323]Department of Health Research and Policy - Epidemiology, Stanford University School of Medicine, 259 Campus Drive, Stanford, CA 94305, USA. [324]Department of Biomedical Data Science, Stanford University School of Medicine, 259 Campus Drive, Stanford, CA 94305, USA. [325]Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig Maximilians University, Neuherberg D-85764, Munich 803539 Bavaria, Germany. [326]Helmholtz Zentrum Munchen, German Research Center for Environmental Health (GmbH), Institute of Epidemiology, Ingolstadter Landstr1, 85764 Neuherberg, Germany. [327]Institute of Medical Statistics and Epidemiology, Technical University Munich, Munich 80333, Germany. [328]Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit, Biocenter Oulu, University of Oulu, Aapistie 5A, 90220 Oulu, Finland. [329]Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre Oulu, Aapistie 5A, 90220 Oulu, Finland. [330]Department of Surgical Sciences, Uppsala University, 751 85 Uppsala, Sweden. [331]Academic Unit of Clinical Oncology, University of Sheffield, Weston Park Hospital, Whitham Road, Sheffield S10 2SJ, UK. [332]Discipline of Genetics, Memorial University of Newfoundland, StJohn's, NL A1C 5S7, Canada. [333]Department of Epidemiology, Division of Cancer Prevention and Population Science, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA. [334]Magee-Womens Hospital, University of Pittsburgh School of Medicine, 300 Halket St, Pittsburgh, PA 15213, USA. [335]Center for Genomic Medicine and Department of Anasthesia, Massachusetts General Hospital, Boston, MA 02114, USA. [336]Human Genetics, Graduate School of Public Health, University of Pittsburgh, UPMC Cancer Pavilion, Suite 4C, Office # 467, 5150 Centre Avenue, Pittsburgh, PA 15232, USA. [337]UPMC Hillman Cancer Center, Pittsburgh 15232 PA, USA. [338]Genetic Cancer Susceptibility Group, International Agency for Research on Cancer, 150 cours Albert Thomas, 69008 Lyon, France. [339]Oncogenetics Team, The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, Downs Road, Sutton SM2 5NG, UK. [340]Genomics Center, Centre Hospitalier Universitaire de Québec - Université Laval Research Center, 2705 Laurier Boulevard, Québec City, QC G1V4G2, Canada. [341]UCLA Path and Lab Med, University of California, 10833 Le Conte Ave, Los Angeles, CA 190095, USA. [342]Department of Medicine, Epidemiology Section, Institute for Clinical and Translational Research, Baylor Medical College, One Baylor Plaza, MS: BCM451, Suite 100D, Houston, TX 77030-3411, USA