Published OnlineFirst June 21, 2017; DOI: 10.1158/1055-9965.EPI-17-0211

Research Article

Cancer Epidemiology, Biomarkers & Prevention

Quantifying the Genetic Correlation between Multiple Cancer Types

Sara Lindström[1,2], Hilary Finucane[3,4], Brendan Bulik-Sullivan[5], Fredrick R. Schumacher[6,7], Christopher I. Amos[8], Rayjean J. Hung[9], Kristin Rand[10], Stephen B. Gruber[10], David Conti[10], Jennifer B. Permuth[11,12], Hui-Yi Lin[13], Ellen L. Goode[14], Thomas A. Sellers[11], Laufey T. Amundadottir[15], Rachael Stolzenberg-Solomon[15], Alison Klein[16,17], Gloria Petersen[14], Harvey Risch[18], Brian Wolpin[19,20], Li Hsu[2], Jeroen R. Huyghe[2], Jenny Chang-Claude[21,22], Andrew Chan[23], Sonja Berndt[15], Rosalind Eeles[24], Douglas Easton[25], Christopher A. Haiman[10], David J. Hunter[3,26], Benjamin Neale[5], Alkes L. Price[3], and Peter Kraft[3,27] on behalf of PanScan, GECCO and the GAME-ON Network: CORECT, DRIVE, ELLIPSE, FOCI, and TRICL-ILCCO

Abstract

Background: Many cancers share specific genetic risk factors, including both rare high-penetrance mutations and common SNPs identified through genome-wide association studies (GWAS). However, little is known about the overall shared heritability across cancers. Quantifying the extent to which two distinct cancers share genetic origin will give insights to shared biological mechanisms underlying cancer and inform design for future genetic association studies.

Methods: In this study, we estimated the pair-wise genetic correlation between six cancer types (breast, colorectal, lung, ovarian, pancreatic, and prostate) using cancer-specific GWAS summary statistics data based on 66,958 case and 70,665 control subjects of European ancestry. We also estimated genetic correlations between cancers and 14 noncancer diseases and traits.

Results: After adjusting for 15 pair-wise genetic correlation tests between cancers, we found significant ($P < 0.003$) genetic correlations between pancreatic and colorectal cancer (rg = 0.55, $P = 0.003$), lung and colorectal cancer (rg = 0.31, $P = 0.001$). We also found suggestive genetic correlations between lung and breast cancer (rg = 0.27, $P = 0.009$), and colorectal and breast cancer (rg = 0.22, $P = 0.01$). In contrast, we found no evidence that prostate cancer shared an appreciable proportion of heritability with other cancers. After adjusting for 84 tests studying genetic correlations between cancer types and other traits (Bonferroni-corrected P value: 0.0006), only the genetic correlation between lung cancer and smoking remained significant (rg = 0.41, $P = 1.03 \times 10^{-6}$). We also observed nominally significant genetic correlations between body mass index and all cancers except ovarian cancer.

Conclusions: Our results highlight novel genetic correlations and lend support to previous observational studies that have observed links between cancers and risk factors.

Impact: This study demonstrates modest genetic correlations between cancers; in particular, breast, colorectal, and lung cancer share some degree of genetic basis. *Cancer Epidemiol Biomarkers Prev; 26(9); 1427–35. ©2017 AACR.*

[1]Department of Epidemiology, University of Washington, Seattle, Washington. [2]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington. [3]Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. [4]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts. [5]The Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts. [6]Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio. [7]Seidman Cancer Center, University Hospitals, Cleveland, Ohio. [8]Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire. [9]Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario, Canada. [10]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California. [11]Department of Cancer Epidemiology, Moffitt Cancer Center and Research Institute, Tampa, Florida. [12]Department of Gastrointestinal Oncology, Moffitt Cancer Center and Research Institute, Tampa, Florida. [13]Department of Biostatistics and Bioinformatics, Moffitt Cancer Center, Tampa, Florida. [14]Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota. [15]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, U.S. Department of Health and Human Services, Bethesda, Maryland. [16]Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, Baltimore, Maryland. [17]Department of Pathology, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins School of Medicine, Baltimore, Maryland. [18]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut. [19]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. [20]Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. [21]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [22]University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [23]Division of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts. [24]Division of Genetics and Epidemiology, The Institute of Cancer Research, and Royal Marsden NHS Foundation Trust, London, United Kingdom. [25]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, United Kingdom. [26]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts. [27]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Sara Lindström, Department of Epidemiology, University of Washington, 1959 NE Pacific Street, Health Sciences Bldg, F-247B, Box 357236, Seattle, WA 98195. Phone: 206-221-3148; Fax: 206-543-8525; E-mail: saralind@uw.edu

doi: 10.1158/1055-9965.EPI-17-0211

Published OnlineFirst June 21, 2017; DOI: 10.1158/1055-9965.EPI-17-0211

Introduction

In the United States, cancer remains the second leading cause of death, with an estimated 1.69 million new cancer diagnoses and 600,000 cancer-related deaths in 2017 (1). Six cancer types—breast, colorectal, lung, ovarian, pancreatic, and prostate cancer—together constitute more than 50% of annual cancer diagnoses (1). The etiologies of these cancers are complex, and associated heritability estimates (2) from twin studies range between 0.15 (colon) and 0.57 (prostate), indicating genetic components. For pancreatic cancer, data have been too sparse to estimate heritability based on twin studies. However, Mucci and colleagues (2) did observe significant, albeit lower relative to other cancers, familial risks for pancreatic cancer. These results are supported by a meta-analysis based on more than 6,500 cases, where having a relative diagnosed with pancreatic cancer was associated with a 1.8-fold [95% confidence interval (CI) 1.5–2.1] risk increase of pancreatic cancer (3).

Various cancers share both environmental and genetic risk factors, including rare high-penetrant mutations in genes such as *BRCA2* which predisposes to breast, ovarian, lung, prostate, and pancreatic cancers (4). Genome-wide association studies (GWAS) have identified more 350 distinct (reported SNPs > 500-kb apart) genomic regions that are associated with cancer (5) of which several are shared between cancer types (6). For example, the 8q24 (7–9) and *TERT* (10–13) regions have been associated with multiple cancer types (pleiotropy), although specific alleles often differ. In contrast to pleiotropy, which does not take the direction of association into account, genetic correlation describes the genome-wide correlation in allele effects and thus, considers the allele-specific direction of association between two traits. Thus, the shared genetic etiology between two traits can be either due to a shared genetic variant (or variants) with nonequal non-zero effect sizes (pleiotropy) or via a correlation between effect sizes for causal variants (genetic correlation; ref. 14). Identifying genetic regions that are associated with multiple cancer types may be useful for determining mechanisms involved in global carcinogenesis. However, the benefit of simultaneously studying multiple cancer types relies on their genetic correlation.

With the introduction of GWAS, it is now possible to quantify the phenotypic variance explained by genotyped single nucleotide polymorphisms (SNP) in single-trait (h_g^2; refs. 15–17) and two-trait (r_g; ref. 18) settings by using variance component methods. Lu and colleagues (19) used this approach to estimate h_g^2 for 12 cancers and found significant nonzero genetic contribution for eight of them. However, their average sample size for each cancer was only 1,793 cases (range: 564–2,848) and 2,200 controls (range: 574–3,159), leading to imprecise heritability estimates. Sampson and colleagues (20) estimated the genetic correlation across 13 different cancers in an average sample size across cancers of 3,807 (range: 535–5,942) cases and 2,625 (range: 1,056–10,857) controls. Although no genetic correlation withstood multiple testing after adjusting for 91 tests, they observed the strongest genetic correlations (all $P < 0.01$) between kidney and testis cancer ($r_g = 0.73$, SE = 0.28), diffuse large B-cell lymphoma and osteosarcoma ($r_g = 0.53$, SE = 0.21), diffuse large B-cell lymphoma and chronic lymphocytic leukemia ($r_g = 0.51$, SE = 0.18), and bladder and lung ($r_g = 0.35$, SE = 0.14) cancer.

The sample sizes of these prior studies make interpretation of their findings difficult. Indeed, a drawback with the variance component approach is its requirement for individual-level data, which prohibits researchers from leveraging GWAS results based on meta-analyses which are often based on much larger sample sizes. The recently developed cross-trait linkage disequilibrium (LD) score regression approach overcomes this limitation by estimating the proportion of phenotypic variance explained by common SNPs (21) and the genetic correlation between two traits (22) using summary statistics only. Here, we set out to quantify pair-wise genetic correlation across breast, colorectal, lung, ovarian, pancreatic, and prostate cancer, capitalizing on summary statistics obtained from GWAS data in 66,958 case and 70,665 control subjects obtained from the GAME-ON, PanScan and GECCO consortia. In addition, we estimated the genetic correlation between each of these cancers and 14 noncancer traits which have all been suggested to be linked to cancer and for which we had access to GWAS summary statistics.

Materials and Methods

The GAME-ON network of consortia for post-GWA research, PanScan, and GECCO

We utilized three large-scale cancer genetic epidemiological consortia: GAME-ON, PanScan III, and GECCO. The Genetic Associations and Mechanisms in Oncology (GAME-ON) consortia is a network of cancer-specific post-GWAS initiatives. The five GAME-ON sites are breast (DRIVE), colorectal (CORECT), lung (TRICL-ILLCO), ovarian (FOCI), and prostate (ELLIPSE; ref. 23). One of the main goals with GAME-ON was to test hypotheses across the cancer types that might illuminate common mechanisms of susceptibility. PanScan is a part of The Pancreatic Cancer Cohort Consortium with the goal of conducting GWAS to identify susceptibility markers for pancreatic cancer. For this study, we utilized genome-wide summary statistics from PanScan III (24). The Genetic Epidemiology of Colorectal Cancer Consortium (GECCO) is a large collaborative consortium evaluating genetic and environmental risk factors for colorectal cancer (25). Details of GAME-ON, PanScan, GECCO and the participating studies are available at http://epi.grants.cancer.gov/gameon/, http://epi.grants.cancer.gov/PanScan/, and https://www.fredhutch.org/en/labs/phs/projects/cancer-prevention/projects/gecco.html. Sample sizes for each cancer is listed in Table 1. These studies have been described in detail previously (23). For each cancer type, genotyping was performed using Illumina or Affymetrix arrays of varying densities described elsewhere (23, 26, 27). For all studies except GECCO, imputation was performed using the 1,000 Genomes Project Phase 1 version 3 reference haplotypes, resulting in up to approximately 10 million SNPs available for the analysis for each cancer type. For GECCO, data was imputed using an in-house reference panel of 2,159 whole-genome sequenced European ancestry GECCO participants (28). As imputation quality scores were not readily accessible across all GWAS, we only included HapMap 3 SNPs as a proxy for well-imputed SNPs (29).

Noncancer traits

We also estimated the genetic correlations between the six cancer types and 14 noncancer traits (29) for which GWAS summary statistics were publicly available (Supplementary Table S1). The included traits were schizophrenia, bipolar disorder, coronary artery disease, type II diabetes, Crohn disease, ulcerative colitis, rheumatoid arthritis, ever/never smoked tobacco, height,

Published OnlineFirst June 21, 2017; DOI: 10.1158/1055-9965.EPI-17-0211

Table 1. Overview of cancers analyzed

Cancer type	Cases	Controls	h_g^2 (SE) – observed scale	h_g^2 (95% CI) – liability scale
Breast	15,748	18,084	0.12 (0.02)	0.14 (0.09–0.18)
Colorectal	15,716	18,154	0.13 (0.02)	0.11 (0.07–0.14)
Lung	12,160	16,838	0.14 (0.03)	0.13 (0.08–0.19)
Ovarian	4,369	9,123	0.10 (0.04)	0.07 (0.02–0.12)
Pancreatic	5,107	8,845	0.07 (0.04)	0.05 (0–0.10)
Prostate	14,160	12,724	0.25 (0.03)	0.27 (0.21–0.33)

NOTE: Number of subjects, observed heritability (standard error) and heritability (95% CI) on the liability scale explained by studied SNPs for each cancer. Cumulative risks used for calculating h_g^2 on the liability scale were obtained from SEER.

body mass index (BMI), fasting serum glucose, triglycerides, LDL cholesterol, and HDL cholesterol. The average sample size for the noncancer traits was 70,488 subjects.

Statistical analysis

Associations between SNPs and cancer risk were estimated by unconditional logistic regression adjusted for age, sex (when applicable), and top principal components (ranging from two to six across contributing GWAS) to adjust for potential population stratification. We used a newly developed method for estimating heritability due to common SNPs and genetic correlations that employs only summary statistics together with linkage disequilibrium (LD) information from a reference panel (21, 22). This method, known as LD score regression, relies on the fact that SNP-specific association statistics reflect the associations of all SNPs in LD with that SNP. Thus, for a polygenic trait, SNPs in high-LD regions will on average have higher χ^2 statistics than SNPs in low-LD regions and similarly, for two polygenic, genetically correlated traits with z-scores z_1 and z_2, the product $z_1 z_2$ will on average be higher for SNPs with high LD than SNPs with low LD. Formally, the relationship between the expected χ^2 statistic for SNP j and the LD score $l(j)$ for SNP j can be described by $E[\chi_j^2] \approx \frac{N_j h_g^2}{M} l_j + 1$, where N_j is the sample size, h_g^2 is the heritability due to included SNPs, M is the number of SNPs and $l(j) := \sum_k r^2(j,k)$ where $r^2(j,k)$ is the correlation between SNP j and all other SNPs i. As the observed SNP heritability estimates for binary traits are not directly comparable with more traditional estimates of heritability (e.g., from twin studies), we transformed the observed heritability to the liability scale which takes both ascertainment and disease prevalence into account as described previously (30). To obtain estimates of cancer-specific prevalence, we used SEER cumulative risks (31). We can extend the calculations of single-trait heritability to include two traits: $E[z_{1j} z_{2j}] = \frac{\sqrt{N_1 N_2} r_g}{M} l_j + \frac{N_s \rho}{N_1 N_2}$, where r_g is the genetic covariance, N_1 and N_2 are sample sizes for trait 1 and 2 respectively, N_s is the number of overlapping samples and ρ is the phenotypic correlation in the overlapping samples. We estimate r_g by the slope of the regression of $z_{1j} z_{2j}$ on the LD score and tested for its difference from 0, as described previously (22). In total, we conducted 15 pair-wise genetic correlation tests between cancers types at a significance threshold of $P = 0.05/15 = 0.003$. In our secondary analyses, estimating genetic correlations between cancer and noncancer traits, we conducted 84 tests and for these analyses, we considered P values less than $P = 0.05/84 = 0.0006$ significant. For all analysis, we only included SNPs from HapMap3.

Results

Cancer-specific GWAS summary statistics from data of 66,958 case and 70,665 control subjects across six cancer types were shared between the cancer sites. We first estimated the heritability due to common SNPs (h_g^2) on the observed and liability scale (Table 1) using cancer-specific cumulative risks (Supplementary Table S2) as reported in SEER (31). Heritability estimates on the liability scale ranged between 0.04 (pancreatic cancer) and 0.27 (prostate cancer). Using cross-trait LD score regression, we quantified the pairwise genetic correlations between breast, colorectal, lung, ovarian, pancreatic and prostate cancer. A summary of the results is displayed as a 6×6 matrix (Table 2) with the estimated pair-wise genetic correlations and associated SEs in the upper right corner of the matrix, and corresponding P values for each genetic correlation can be seen in the lower left corner of the matrix. We found significant positive genetic correlations between pancreatic and colorectal cancer ($r_g = 0.55$, $P = 0.003$), lung and colorectal cancer ($r_g = 0.31$, $P = 0.001$), as well as suggestive positive genetic correlations between lung and breast cancer ($r_g = 0.27$, $P = 0.009$), and colorectal and breast cancer ($r_g = 0.22$, $P = 0.01$).

We also estimated the genetic correlation between each of these six cancer types and 14 other diseases and traits implicated to be linked to cancer [schizophrenia, bipolar disorder, coronary artery disease, type II diabetes, Crohn disease, ulcerative colitis, rheumatoid arthritis, ever/never smoked tobacco, height, BMI, fasting serum glucose, triglycerides, LDL cholesterol and HDL cholesterol (Fig. 1; Table 3; Supplementary Tables S1 and S3)]. The strongest positive genetic correlation was, as expected, between lung cancer and smoking status (never/ever; $r_g = 0.41$,

Table 2. Genetic correlations between cancers

	Breast	Colorectal	Lung	Ovarian	Pancreatic	Prostate
Breast	1	**0.22 (0.091)**	**0.27 (0.11)**	0.26 (0.20)	0.17 (0.19)	0.06 (0.09)
Colorectal	0.014	1	**0.31 (0.097)**	−0.08 (0.13)	**0.55 (0.19)**	0.09 (0.07)
Lung	0.009	0.001	1	−0.17 (0.17)	0.32 (0.19)	0.095 (0.08)
Ovarian	0.18	0.57	0.32	1	−0.40 (0.29)	0.02 (0.14)
Pancreatic	0.37	0.003	0.08	0.17	1	−0.06 (0.14)
Prostate	0.52	0.2	0.25	0.89	0.68	1

NOTE: Genetic correlations with standard errors are in the upper right part of the table; corresponding P values are in the lower left part of the table. Nominally significant ($P < 0.05$) genetic correlations are highlighted in bold.

Published OnlineFirst June 21, 2017; DOI: 10.1158/1055-9965.EPI-17-0211

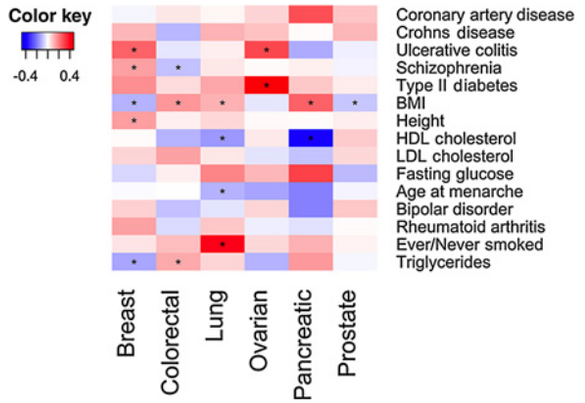

Figure 1.
Genetic correlations between cancers and noncancer traits. Nominally significant genetic correlations ($P < 0.05$) are highlighted with *.

$P = 1.03 \times 10^{-6}$), which was the only correlation to remain significant after adjusting for multiple testing. Other nominally significant genetic correlations with lung cancer were with BMI ($r_g = 0.12$, $P = 0.03$) and an inverse genetic correlation with HDL levels ($r_g = -0.15$, $P = 0.02$). For breast cancer, we observed positive genetic correlations with ulcerative colitis ($r_g = 0.24$, $P = 0.002$), schizophrenia ($r_g = 0.14$, $P = 0.004$), and height ($r_g = 0.14$, $P = 0.01$). In contrast, triglycerides ($r_g = -0.13$, $P = 0.03$) and BMI ($r_g = -0.11$, $P = 0.04$) both had inverse genetic correlations with breast cancer. For colorectal cancer, we observed positive genetic correlations with BMI ($r_g = 0.16$, $P = 6.2 \times 10^{-4}$) and triglycerides ($r_g = 0.13$, $P = 0.03$) and an inverse genetic correlation with schizophrenia ($r_g = -0.09$, $P = 0.05$). Ovarian cancer showed positive genetic correlations with type II diabetes ($r_g = 0.47$, $P = 0.01$) and ulcerative colitis ($r_g = 0.29$, $P = 0.03$). For pancreatic cancer, we observed a positive genetic correlation with BMI ($r_g = 0.24$, $P = 0.04$) and an inverse genetic correlation with HDL levels ($r_g = -0.41$, $P = 0.02$). Finally, prostate cancer showed an inverse correlation with BMI ($r_g = -0.08$, $P = 0.04$).

Discussion

It is well known that cancer tends to cluster in families, which has been attributed to shared environmental factors and genetics (2). Many lifestyle factors associated with cancer including smoking, obesity, and alcohol intake have been shown to influence risk of multiple cancer types, implying that different cancer types share an underlying biological mechanism. Similarly, genetic variation in genes such as *BRCA1/2* and *TERT* has been associated with risks of multiple different cancer types, providing empirical support that there are specific regions in the genome that harbor genetic variation influencing risk of multiple cancer sites. However, although many cancer types might share susceptibility loci (i.e., pleiotropy), their genetic correlation, which depends on specific risk alleles and their direction of associations, might not be strong. Here, we aimed to assess the latter among six cancers, breast, colorectal, lung, ovary, pancreatic and prostate, as well as between cancer types and seven additional disease and seven nondisease traits that have all been found to have heritable components.

We used SEER estimates (31) to obtain cancer-specific cumulative risks, recognizing that not all study subjects came from U.S. populations. For comparison, we also calculated the heritability on the liability scale using cumulative risks obtained from Mucci and colleagues (2). No qualitative difference was observed for the liability estimates using the two different sources of cumulative risks. In concordance with previous studies, the cancer-specific heritability estimates observed here were lower than what has been observed in twin studies. This is not unexpected given that here, we are only estimating the additive genetic component based on common SNPs captured by GWAS, and thus, any contribution to the heritability based on factors such as gene–gene interactions, gene–environment interactions, structural variants or rare variants, will not be captured by our analysis. Among the studied cancers, we observed the largest heritability for prostate cancer ($h_g^2 = 0.27$) in agreement with previous twin studies (2). We also compared our cancer-specific heritability results to previous studies estimating heritability based on GWAS data. In general, our results were comparable with previous studies. Lu and colleagues (19) estimated h_g^2 for breast cancer to be 0.13 (95% CI: 0–0.56) compared with our estimate of 0.14 (95% CI:

Table 3. Nominally significant genetic correlations between cancers and noncancer traits

Cancer	Noncancer trait	r_g	SE	P	Epidemiologic observations	Reference
Breast cancer	Ulcerative colitis	0.243	0.08	0.002	+	55
Breast cancer	Schizophrenia	0.142	0.049	0.004	+	56
Breast cancer	Height	0.143	0.055	0.01	+	57
Breast cancer	Triglycerides	−0.13	0.061	0.033	−	58
Breast cancer	BMI	−0.112	0.053	0.035	− (premenopausal) + (postmenopausal)	35
Colorectal cancer	BMI	0.157	0.046	6.22E−04	+	35
Colorectal cancer	Triglycerides	0.126	0.058	0.029	+	59
Colorectal cancer	Schizophrenia	−0.091	0.046	0.048	No association	56
Lung cancer	Ever/never smoked	0.412	0.084	1.03E−06	+	60
Lung cancer	HDL	−0.151	0.063	0.017	−	59
Lung cancer	BMI	0.116	0.054	0.032	− (smokers) No association (nonsmokers)	35
Ovarian cancer	Type-2 diabetes	0.469	0.191	0.014	+	61
Ovarian cancer	Ulcerative colitis	0.291	0.137	0.034	No association	55
Pancreatic cancer	HDL	−0.405	0.167	0.015	Unknown	N/A
Pancreatic cancer	BMI	0.243	0.117	0.038	+	35
Prostate cancer	BMI	−0.083	0.04	0.039	−	35

NOTE: For each cancer-trait genetic correlation observed, current evidence from observational studies and corresponding reference are also listed. *P* values are not corrected for multiple testing.

Published OnlineFirst June 21, 2017; DOI: 10.1158/1055-9965.EPI-17-0211

0.09–0.18). For lung cancer, previous estimates have varied with Lu and colleagues (19) estimating h_g^2 to 0.10 (95% CI: 0–0.24) in European populations while Sampson and colleagues (20) estimated h_g^2 to be 0.21 (95% CI: 0.14–0.27), compared with our estimate of 0.13 (95% CI: 0.08–0.19). For ovarian cancer, we observed a small heritability ($h_g^2 = 0.07$, 95% CI: 0.02–0.12) compared to Lu and colleagues (ref. 19; $h_g^2 = 0.30$, 95% CI: 0.18–0.42). It is not clear why we observe this discrepancy in results. We also observed lower h_g^2 of pancreatic cancer (0.05, 95% CI: 0–0.10) than previously observed [0.18, 95% CI: 0.06–0.30 for Lu and colleagues (19) and 0.10, 95% CI: 0.04–0.16 for Sampson and colleagues (20)] but we note that the CIs are wide and overlap. For prostate cancer, we observed a heritability of 0.27 (95% CI: 0.21–0.37) compared with Lu (ref. 19; 0.81, 95% CI: 0.32–1), and Sampson (ref. 20; 0.29, 95% CI: 0.15–0.42). We note that heritability estimates reported here are all on the liability scale and were calculated using SEER rates for all three studies including ours.

We found that colorectal cancer showed significant genetic correlations with pancreatic and lung cancers, with the largest genetic correlation observed for the two gastrointestinal tract cancers: the r_g for colorectal and pancreatic cancer was 0.55 ($P = 0.003$). Amundadottir and colleagues studied cancer risk for first up to fifth degree relatives in an Icelandic population and observed an increased risk for pancreatic cancer among colon cancer patients (and vice versa) for first degree relatives but not beyond (32). Colorectal cancer patients have been observed to have a higher incidence of pancreatic cancer than the general population (33). Furthermore, Lynch syndrome, the most common hereditary colorectal cancer syndrome, has also been shown to increase risk for pancreatic cancer (34). Obesity is a well-established risk factor for both colorectal and pancreatic cancer (35, 36) and we observed nominally significant ($P < 0.05$) genetic correlations between BMI and both colorectal and pancreatic cancer.

Colorectal cancer also showed suggestive genetic correlation with breast cancer in agreement with the study from Amundadottir and colleagues (32). A recent study found that women diagnosed with breast cancer have a 1.59-fold (95% CI: 1.53–1.65) increased risk of developing colorectal cancer compared with the general population (37).

Breast and lung cancer showed a suggestive positive genetic correlation ($r_g = 0.27$, $P = 0.009$), supported by observational studies finding familial cosegregation of the two cancers (38, 39) as well as overlap in multiple susceptibility genes such as *BRCA2*, *CHEK2* (40) and *LSP1* (7). In contrast, Amundadottir and colleagues did not observe a significant cooccurrence among relatives (32). A recent cross-cancer GWAS based on the GAME-ON data (23) identified a pleiotropic locus at 1q22 that was associated with both breast and lung cancer.

We expect these results to generate testable hypotheses about mechanisms. There are data to support inflammation response, DNA repair and stress responses, to list just a few. For example, the genetic correlations between breast, colorectal and lung cancer might be driven in part by genetic variants in the inflammation pathway. Indeed, a recent analysis of genetic variation in the inflammation pathway from the GAME-ON consortium identified *SH2B3*, a key negative regulator of cytokine signaling to be associated with all three cancers (41). Furthermore, the authors found no evidence that genetic variation in inflammation-related genes was associated with prostate cancer risk. This observation corroborates our findings that prostate cancer does not share an appreciable genetic component with other cancers and supports inflammation to be a pathway in which genetic variation affects the risk of breast, colorectal, and lung cancer. Removing all SNPs with a χ^2 test statistic >25 in the individual cancer GWAS, did not change our results: colorectal–breast cancer ($r_g = 0.22$ for all SNPs and $r_g = 0.23$ excluding significant SNPs); colorectal–lung cancer ($r_g = 0.31$ for all SNPs and $r_g = 0.34$ excluding significant SNPs); colorectal–pancreatic cancer ($r_g = 0.55$ for all SNPs and $r_g = 0.58$ excluding significant SNPs); breast–lung cancer ($r_g = 0.27$ for all SNPs and $r_g = 0.33$ excluding significant SNPs). Thus, it is likely that the genetic correlations that we observe are due to yet unidentified SNPs, and future studies should focus on simultaneously study genetically correlated cancers with the goal of identifying SNPs that are associated with multiple cancer types.

We did not observe evidence that either ovarian or prostate cancer shared an appreciable amount of heritability with other cancers, although our sample size for ovarian cancer was relatively small, leading to wide CIs. We had higher statistical power to detect genetic correlations involving prostate cancer, but no estimate was >0.1, suggesting that prostate cancer has a genetic contribution that is distinct from that of other cancer types studied here. Witte and Hoffmann used polygenic risk scores to investigate a potential shared heritability between breast and prostate cancer, and although they observed a potential common polygenic model between nonaggressive prostate cancer and breast cancer, they observed no evidence of a common model between overall prostate cancer and breast cancer, consistent with our results here (42). It is important to note that our analysis does not capture rare higher penetrance mutations, and thus, the observed increased risk of multiple cancers among relatives to prostate cancer cases (32) is likely to at least in part be attributable to rare variants such as *BRCA2*.

In addition, we examined the genetic correlation between these cancer types and 14 noncancer diseases and traits. The only genetic correlation between cancer and noncancer traits that withstood correction for multiple testing was smoking status and lung cancer. The strongest lung cancer susceptibility locus is in the 15q25 region which contains genes encoding the nicotinic acetylcholine receptor subunits *CHRNA5*, *CHRNA3*, and *CHRNB4* and has also been associated with smoking behavior with associations in the same direction (43–45). We found that BMI showed nominally significant genetic correlations with all cancers except ovarian cancer. While BMI showed positive genetic correlation with colorectal, lung, and pancreatic cancer, it showed negative genetic correlations with breast and prostate cancer. These results mirror recent Mendelian randomization (MR) studies of BMI and colorectal (46), breast and lung cancer (47, 48), providing further evidence that obesity is involved in cancer development. An MR study of prostate cancer found a nonsignificant lower risk associated with a BMI genetic score (OR = 0.98; 95% CI: 0.96–1.00; $P = 0.07$; ref. 49). Although the positive genetic correlation between BMI and lung cancer seems to contradict results from observational studies, the observational association between BMI and lung cancer might be due to residual confounding by smoking (35). BMI and smoking behavior have been shown to share a genetic basis ($r_g = 0.20$, $P = 8.3 \times 10^{-7}$; ref. 22) and further, cell-type enrichment heritability analysis have shown that both smoking behavior and BMI are enriched for central nervous system–related cell types. Therefore, it is

Published OnlineFirst June 21, 2017; DOI: 10.1158/1055-9965.EPI-17-0211

possible that smoking and BMI to some extent affect lung cancer risk through the same biological mechanism.

We observed very few genetic correlations with prostate cancer compared with the other cancers of comparable sample size and no genetic correlation with noncancer traits were >0.1. We also note the lack of genetic correlation between prostate cancer and type II diabetes ($r_g = 0.03$, 95% CI −0.10–0.15, $P = 0.67$). The epidemiologic inverse association between prostate cancer and type II diabetes is well-documented (50–53) and a previous study showed that 10 of 36 type II diabetes SNPs were associated (2 with increased risk and 8 with decreased risk) with advanced prostate cancer (54). The different directions of significant prostate cancer associations across type II diabetes SNPs are consistent with the lack of genetic correlation (which is sensitive to direction of effects) observed in this study.

Sampson and colleagues previously used individual-level GWAS data to estimate genetic correlations between 13 cancers in 49,492 cancer cases and 34,131 controls, including estrogen receptor negative (ER−) breast cancer, lung cancer, pancreatic, and prostate cancer (20). Although, they did not observe the statistically significant genetic correlations between the cancers studied here, their SEs were in general large, making it difficult to compare the results. We note that the ER^- breast cancer and pancreatic datasets they used are a subset of the data analyzed here. While we had access to GWAS summary statistics based on cancer subtypes including ER^- breast cancer, squamous cell lung cancer, lung adenocarcinoma, serous, clear cell, and endometrioid ovarian cancer and aggressive prostate cancer, sample sizes for these subsets were too small for meaningful analysis. On the basis of our experience, LD score regression requires at least 10,000 cases for adequately stable estimates at these heritabilities. We note a few limitations with only having access to summary statistics data compared to individual-level data. Most importantly, the SEs associated with the estimated genetic correlations based on LD score regression are larger than what is obtained by similar methods using individual-level data. In addition, we are not able to conduct any subtype analysis on the original traits that might be of interest, for example, it might have been of interest to study the genetic correlation between BMI and breast cancer stratified by menopausal status. LD score regression leverages summary statistics rather than individual-level data and thereby overcome many of the issues associated with relying on individual-level data. Moreover, appropriate quality control steps were conducted as part of the cancer-specific GWAS meta-analysis. Furthermore, we limited our analysis to HapMap 3 SNPs to ensure well-imputed data.

In summary, our results indicate that some cancers show modest genetic correlations; in particular, breast, colorectal and lung cancer share some degree of genetic basis. In contrast, prostate cancer appears to have a unique genetic architecture that is not shared with breast, lung, ovarian, and pancreatic cancer. Furthermore, a number of cancer types show genetic correlations with obesity, highlighting the involvement of adiposity-related processes in cancer. As GWAS sample sizes continue to increase and GWAS summary statistics from other traits become available, we will be able to additionally characterize the shared heritability between cancer types including histologic subtypes as well as with noncancer traits.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Disclaimer

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 CORECT, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government or CORECT.

Authors' Contributions

Conception and design: S. Lindström, F.R. Schumacher, C.I. Amos, E.L. Goode, D.J. Hunter, P. Kraft

Development of methodology: H. Finucane, B. Bulik-Sullivan, B. Neale, A. Price, P. Kraft

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): F.R. Schumacher, C.I. Amos, R.J. Huang, S.B. Gruber, E.L. Goode, T.A. Sellers, L.T. Amundadottir, R.Z. Stolzenberg-Solomon, A.P. Klein, G.M. Petersen, H.A. Risch, B. Wolpin, J.R. Huyghe, J.C. Claude, A. Chan, S.I. Berndt, R. Eeles, D.F. Easton, C.A. Haiman, D.J. Hunter, P. Kraft

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S. Lindström, H. Finucane, B. Bulik-Sullivan, F.R. Schumacher, R.J. Huang, S.B. Gruber, D.V. Conti, L.T. Amundadottir, H.A. Risch, L. Hsu, B. Neale, P. Kraft

Writing, review, and/or revision of the manuscript: S. Lindström, F.R. Schumacher, C.I. Amos, R.J. Huang, S.B. Gruber, J.B. Permuth, H.-Y. Lin, E.L. Goode, T.A. Sellers, L.T. Amundadottir, R.Z. Stolzenberg-Solomon, A.P. Klein, H.A. Risch, B. Wolpin, L. Hsu, J.R. Huyghe, J.C. Claude, A. Chan, S.I. Berndt, R. Eeles, D.F. Easton, D.J. Hunter, B. Neale, P. Kraft

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C.I. Amos, K.A. Rand, S.B. Gruber, R.Z. Stolzenberg-Solomon, H.A. Risch, R. Eeles, C.A. Haiman, D.J. Hunter, P. Kraft

Study supervision: R.Z. Stolzenberg-Solomon, J.R. Huyghe, J.C. Claude, A. Chan, R. Eeles, D.F. Easton, D.J. Hunter, P. Kraft

Acknowledgments

Members of CORECT

CORECT acknowledges the following investigators: Kendra Blalock, Peter T. Campbell, Graham Casey, David V. Conti, Christopher K. Edlund, Jane Figueiredo, W. James Gauderman, Jian Gong, Roger C. Green, Stephen B. Gruber, John F. Harju, Tabitha A. Harrison, Eric J. Jacobs, Mark A. Jenkins, Shuo Jiao, Li Li, Yi Lin, Frank J. Manion, Victor Moreno, Bhramar Mukherjee, Ulrike Peters, Leon Raskin, Fredrick R. Schumacher, Daniela Seminara, Gianluca Severi, Stephanie L. Stenzel, and Duncan C. Thomas.

Members of DRIVE

DRIVE acknowledges the following GWASs and investigators that shared genome-wide summary data as part of the breast-cancer GWAS meta-analysis: the Australian Breast Cancer Family Study (ABCFS; John L. Hopper, Melissa C. Southey, Enes Makalic, Daniel F. Schmidt), the British Breast Cancer Study (BBCS; Olivia Fletcher, Julian Peto, Lorna Gibson, Isabel dos Santos Silva), the Breast and Prostate Cancer Cohort Consortium (BPC3) (David J. Hunter, Sara Lindström, Peter Kraft), the Breast Cancer Family Registries (BCFR; Habib Ahsan, Alice Whittemore), the Dutch Familial Bilateral Breast Cancer Study (DFBBCS; Quinten Waisfisz, Hanne Meijers-Heijboer, Muriel Adank, Rob B. van der Luijt, Andre G. Uitterlinden, Albert Hofman), German Consortium for Hereditary Breast and Ovarian Cancer (GC-HBOC; Alfons Meindl, Rita K. Schmutzler, Bertram Müller-Myhsok, Peter Lichtner), the Helsinki Breast Cancer Study (HEBCS; Heli Nevanlinna, Taru A. Muranen, Kristiina Aittomäki, Carl Blomqvist), the Mammary Carcinoma Risk Factor Investigation (MARIE; Jenny Chang-Claude, Rebecca Hein, Norbert Dahmen, Lars Beckman), SardiNIA (Laura Crisponi), the Singapore and Sweden Breast Cancer Study (SASBAC; Per Hall, Kamila Czene, Astrid Irwanto, Jianjun Liu), and the UK2 (Douglas F. Easton, Clare Turnbull, Nazneen Rahman).

Members of ELLIPSE

ELLIPSE acknowledges the following GWASs and investigators that shared genome-wide summary data as part of the prostate cancer GWAS meta-analysis: CRUK (Rosalind Eeles, Douglas F. Easton, Zsofia Kote-Jarai, Kenneth Muir, Graham Giles, Gianluca Severi, David Neal, Jenny L. Donovan, Freddie C. Hamdy), CAPS1 and CAPS2 (Fredrik Wiklund, Henrik Gronberg), BPC3-MEC (Christopher Haiman, Fred Schumacher), BPC3-EPIC (Ruth Travis, Elio Riboli), BPC3-Harvard (Peter Kraft, David Hunter), BPC3-ACS (Susan Gapstur), PEGASUS (Sonja Berndt, Stephen Chanock).

Members of TRICL

TRICL acknowledges the following investigators: Younghun Han, Li Su, Yongyue Wei, Rayjean J. Hung, Yonathan Brhane, John McLaughlin, Paul Brennan, James D. McKay, Heike Bickeböller, Albert Rosenberger, Richard S. Houlston, Neil Caporaso, Maria Teresa Landi, Joachim Heinrich, Angela Risch, Xifeng Wu, Yuanqing Ye, David C. Christiani, Christopher I. Amos.

PanScan: The authors wish to thank all participants and participating institutions, investigators, and staff that helped made this project possible. The authors acknowledge the contribution of the staff of the Cancer Genomics Research Laboratory (CGR) at the National Cancer Institute, National Institutes of Health, Bethesda, MD, for their help throughout the project.

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, Utz Benscheid, Muhabbet Celik and Ursula Eilber for excellent technical assistance.

GECCO: The authors would like to thank all those at the GECCO Coordinating Center for helping bring together the data and people that made this project possible. The authors acknowledge Dave Duggan and team members at TGEN (Translational Genomics Research Institute), the Broad Institute, and the Génome Québec Innovation Center for genotyping DNA samples of cases and controls, and for scientific input for GECCO.

HPFS, NHS, and PHS: We would like to 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: 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.

PLCO: The authors 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 and staff, SAIC-Frederick. Most importantly, we acknowledge the study participants for their contributions to 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: The authors would like to thank the study participants and staff of the Hormones and Colon Cancer study.

WHI: The authors 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 http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf.

Grant Support

This work was supported by the Genetic Associations and Mechanisms in Oncology Network, GAME-ON (http://epi.grants.cancer.gov/gameon/), which includes the following consortia: CORECT (grant number U19 CA148107); DRIVE (U19 CA148065); ELLIPSE (grant number U19 CA148537); FOCI (U19 CA148112); TRICL (U19 CA148127). This work was also supported by NIH grants CA194393, CA173785 and CA189532. R. Eeles is supported by NIHR funding to the Biomedical Research Centre at The Institute of Cancer Research and Royal Marsden NHS Foundation Trust. HKF is supported by the Fannie and John Hertz Foundation.

PanScan is supported by the Intramural Research Program of the NIH, National Cancer Institute.

GECCO was supported by National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services (U01 CA137088; U01 CA164930). Genotyping/Sequencing services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the NIH to The Johns Hopkins University, contract number HHSN268201200008I.

ASTERISK: A Hospital Clinical Research Program (PHRC-BRD09/C) from the University Hospital Center of Nantes (CHU de Nantes) 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: NIH (R01 CA60987).

CCFR: This work was supported by grant UM1 CA167551 from the National Cancer Institute and through cooperative agreements with the following CCFR centers: Ontario Familial Colorectal Cancer Registry (U01/U24 CA074783) and Seattle Colorectal Cancer Family Registry (U01/U24 CA074794).

The Colon CFR GWAS was supported by funding from the National Cancer Institute, NIH (U01 CA122839 and R01 CA143237 to Graham Casey). 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 Colon Cancer Family Registry (CCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the CCFR.

CORECT: National Cancer Institute, NIH under RFA # CA-09-002 (U19 CA148107).

CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study-II (CPS-II) cohort. This study was conducted with Institutional Review Board approval.

DACHS: German Research Council (Deutsche Forschungsgemeinschaft, BR 1704/6-1, BR 1704/6-3, BR 1704/6-4 and CH 117/1-1), and the German Federal Ministry of Education and Research (01KH0404 and 01ER0814).

DALS: NIH (R01 CA48998 to M.L. Slattery).

HPFS is supported by the NIH (P01 CA055075, UM1 CA167552, R01 CA137178, R01 CA151993, R35 CA197735, K07 CA190673, and P50 CA127003), NHS and NHS II by the NIH (R01 CA137178, P01 CA087969, UM1 CA186107, UM1 CA176726, R01 CA67262, R01 CA151993, R01 CA49449, R35 CA197735, K07 CA190673, and P50 CA127003), and PHS by the NIH (R01 CA042182).

MEC: NIH (R37 CA54281, P01 CA033619, and R01 CA63464).

OFCCR: NIH, through funding allocated to the Ontario Registry for Studies of Familial Colorectal Cancer (U01 CA074783); see previous CCFR section. Additional funding toward genetic analyses of OFCCR includes the Ontario Research Fund, the Canadian Institutes of Health Research, and 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. In addition, a subset of control samples were genotyped as part of the Cancer Genetic Markers of Susceptibility (CGEMS) Prostate Cancer GWAS (Yeager, M and colleagues Genome-wide association study of prostate cancer identifies a second risk locus at 8q24 [Nat Genet 2007 May;39(5):645–9], CGEMS pancreatic cancer scan (PanScan; Amundadottir and colleagues, Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer; [Nat Genet. 2009 Sep;41(9):986–90], and Petersen and colleagues [A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet. 2010 Mar;42(3):224–8], and the Lung Cancer and Smoking study [Landi and colleagues, A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma; Am J Hum Genet 2009 Nov;85 (5):679–91]. 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 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, and the Johns Hopkins University Center for Inherited Disease Research conducted genotyping.

PMH: NIH (R01 CA076366 to P.A. Newcomb).

VITAL: NIH (K05 CA154337).

WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, NIH, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 8, 2017; revised May 3, 2017; accepted June 6, 2017; published OnlineFirst June 21, 2017.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. CA Cancer J Clin 2017;67:7–30.
2. Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial risk and heritability of cancer among twins in nordic countries. JAMA 2016;315:68–76.
3. Permuth-Wey J, Egan KM. Family history is a significant risk factor for pancreatic cancer: results from a systematic review and meta-analysis. Fam Cancer 2009;8:109–17.
4. Maxwell KN, Domchek SM. Cancer treatment according to BRCA1 and BRCA2 mutations. Nat Rev Clin Oncol 2012;9:520–8.
5. The NHGRI-EBI catalog of published genome-wide association studies. Available from: https://www.ebi.ac.uk/gwas/.
6. Sakoda LC, Jorgenson E, Witte JS. Turning of COGS moves forward findings for hormonally mediated cancers. Nat Genet 2013;45: 345–8.
7. Park SL, Fesinmeyer MD, Timofeeva M, Caberto CP, Kocarnik JM, Han Y, et al. Pleiotropic associations of risk variants identified for other cancers with lung cancer risk: the PAGE and TRICL consortia. J Natl Cancer Inst 2014;106:dju061.
8. Witte JS. Multiple prostate cancer risk variants on 8q24. Nat Genet 2007;39:579–80.
9. Wokolorczyk D, Gliniewicz B, Sikorski A, Zlowocka E, Masojc B, Debniak T, et al. A range of cancers is associated with the rs6983267 marker on chromosome 8. Cancer Res 2008;68:9982–6.
10. Bojesen SE, Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet 2013; 45:371–84.
11. McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet 2008; 40:1404–6.
12. Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, Sigurdsson A, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nat Genet 2009;41:221–7.
13. Yin J, Li Y, Yin M, Sun J, Liu L, Qin Q, et al. TERT-CLPTM1L polymorphism rs401681 contributes to cancers risk: evidence from a meta-analysis based on 29 publications. PLoS One 2012;7:e50650.
14. Pasaniuc B, Price AL. Dissecting the genetics of complex traits using summary association statistics. Nat Rev Genet 2017;18:117–27.
15. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. Am J Hum Genet 2011;88:294–305.
16. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 2010;42:565–9.
17. Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, et al. Genome partitioning of genetic variation for complex traits using common SNPs. Nat Genet 2011;43:519–25.
18. Cross-Disorder Group of the Psychiatric Genomics Consortium, Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet 2013;45:984–94.
19. Lu Y, Ek WE, Whiteman D, Vaughan TL, Spurdle AB, Easton DF, et al. Most common `sporadic' cancers have a significant germline genetic component. Hum Mol Genet 2014;23:6112–8.
20. Sampson JN, Wheeler WA, Yeager M, Panagiotou O, Wang Z, Berndt SI, et al. Analysis of heritability and shared heritability based on genome-wide association studies for thirteen cancer types. J Natl Cancer Inst 2015;107: djv279.
21. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015;47:291–5.
22. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015;47:1236–41.
23. Fehringer G, Kraft P, Pharoah PD, Eeles RA, Chatterjee N, Schumacher FR, et al. Cross-cancer genome-wide analysis of lung, ovary, breast, prostate and colorectal cancer reveals novel pleiotropic associations. Cancer Res 2016;76:5103–14.
24. Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. Nat Genet 2014;46:994–1000.
25. Peters U, Jiao S, Schumacher FR, Hutter CM, Aragaki AK, Baron JA, et al. Identification of genetic susceptibility loci for colorectal tumors in a genome-wide meta-analysis. Gastroenterology 2013;144:799–807.
26. Kaminski BM ACI, DeRycke E, Gillanders EM, Gruber SB, Henderson BE, Hunter DJ, et al. Genetic Associations and Mechanisms in Oncology (GAME-ON): A network approach to post-GWAS research. Cancer Epidemiol Biomarkers Prev 2013;21:78.
27. Zhang C, Doherty JA, Burgess S, Hung RJ, Lindström S, Kraft P, et al. Genetic determinants of telomere length and risk of common cancers: a Mendelian randomization study. Hum Mol Genet 2015;24:5356–66.
28. Huyghe J, Chen S, Kang HM, Harrison TA, Berndt SI, Bézieau S, Brenner H, et al. Large scale whole genome sequencing with imputation into GWAS improves our understanding of the genetic architecture of colorectal cancer [abstract]. In: Proceedings of the 107th Annual Meeting of the American Association for Cancer Research; 2016 Apr 16–20; New Orleans, LA. Philadelphia (PA): AACR; 2016. Abstract nr 5230.
29. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet 2015;47:1228–35.
30. Zaitlen N, Kraft P. Heritability in the genome-wide association era. Hum Genet 2012;131:1655–64.
31. Lifetime Risk (Percent) of being diagnosed with Cancer by Site and Race/Ethnicity Both Sexes, 18 SEER Areas, 2012–2014; 2017. Available from: https://seer.cancer.gov/csr/1975_2014/results_merged/topic_lifetime_risk.pdf.
32. Amundadottir LT, Thorvaldsson S, Gudbjartsson DF, Sulem P, Kristjansson K, Arnason S, et al. Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. PLoS Med 2004;1:e65.
33. Rahimi E, Batra S, Thosani N, Singh H, Guha S. Increased incidence of second primary pancreatic cancer in patients with prior colorectal cancer: a population-based US study. Dig Dis Sci 2016;61:1652–60.
34. Underhill ML, Germansky KA, Yurgelun MB. Advances in hereditary colorectal and pancreatic cancers. Clin Ther 2016;38:1600–21.
35. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. Lancet 2014;384:755–65.
36. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet 2008;371:569–78.
37. Lu Y, Segelman J, Nordgren A, Lindstrom L, Frisell J, Martling A. Increased risk of colorectal cancer in patients diagnosed with breast cancer in women. Cancer Epidemiol 2016;41:57–62.
38. Dite GS, Whittemore AS, Knight JA, John EM, Milne RL, Andrulis IL, et al. Increased cancer risks for relatives of very early-onset breast cancer cases with and without BRCA1 and BRCA2 mutations. Br J Cancer 2010;103: 1103–8.
39. Gorlova OY, Weng SF, Zhang Y, Amos CI, Spitz MR. Aggregation of cancer among relatives of never-smoking lung cancer patients. Int J Cancer 2007;121:111–8.
40. Wang Y, McKay JD, Rafnar T, Wang Z, Timofeeva MN, Broderick P, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat Genet 2014;46:736–41.
41. Hung RJ, Ulrich CM, Goode EL, Brhane Y, Muir K, Chan AT, et al. Cross cancer genomic investigation of inflammation pathway for five common cancers: lung, ovary, prostate, breast, and colorectal cancer. J Natl Cancer Inst 2015; 107(11). pii: djv246.
42. Witte JS, Hoffmann TJ. Polygenic modeling of genome-wide association studies: an application to prostate and breast cancer. OMICS 2011;15: 393–8.
43. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet 2008;40:616–22.
44. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 2008;452:633–7.
45. Saccone NL, Culverhouse RC, Schwantes-An TH, Cannon DS, Chen X, Cichon S, et al. Multiple independent loci at chromosome 15q25.1 affect

Published OnlineFirst June 21, 2017; DOI: 10.1158/1055-9965.EPI-17-0211

smoking quantity: a meta-analysis and comparison with lung cancer and COPD. PLoS Genet 2010; 6(8): e1001053.
46. Thrift AP, Gong J, Peters U, Chang-Claude J, Rudolph A, Slattery ML, et al. Mendelian randomization study of body mass index and colorectal cancer risk. Cancer Epidemiol Biomarkers Prev 2015;24:1024–31.
47. Carreras-Torres R, Haycock PC, Relton CL, Martin RM, Smith GD, Kraft P, et al. The causal relevance of body mass index in different histological types of lung cancer: a Mendelian randomization study. Sci Rep 2016;6:31121.
48. Gao C, Patel CJ, Michailidou K, Peters U, Gong J, Schildkraut J, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. Int J Epidemiol 2016;45:896–908.
49. Davies NM, Gaunt TR, Lewis SJ, Holly J, Donovan JL, Hamdy FC, et al. The effects of height and BMI on prostate cancer incidence and mortality: a Mendelian randomization study in 20,848 cases and 20,214 controls from the PRACTICAL consortium. Cancer Causes Control 2015;26:1603–16.
50. Fall K, Garmo H, Gudbjornsdottir S, Stattin P, Zethelius B. Diabetes mellitus and prostate cancer risk; a nationwide case-control study within PCBaSe Sweden. Cancer Epidemiol Biomarkers Prev 2013;22:1102–9.
51. Gong Z, Neuhouser ML, Goodman PJ, Albanes D, Chi C, Hsing AW, et al. Obesity, diabetes, and risk of prostate cancer: results from the prostate cancer prevention trial. Cancer Epidemiol Biomarkers Prev 2006;15:1977–83.
52. Tsilidis KK, Allen NE, Appleby PN, Rohrmann S, Nothlings U, Arriola L, et al. Diabetes mellitus and risk of prostate cancer in the European Prospective Investigation into Cancer and Nutrition. Int J Cancer 2015;136:372–81.
53. Waters KM, Henderson BE, Stram DO, Wan P, Kolonel LN, Haiman CA. Association of diabetes with prostate cancer risk in the multiethnic cohort. Am J Epidemiol 2009;169:937–45.
54. Machiela MJ, Lindström S, Allen NE, Haiman CA, Albanes D, Barricarte A, et al. Association of type 2 diabetes susceptibility variants with advanced prostate cancer risk in the Breast and Prostate Cancer Cohort Consortium. Am J Epidemiol 2012;176:1121–9.
55. Hemminki K, Li X, Sundquist J, Sundquist K. Cancer risks in ulcerative colitis patients. Int J Cancer 2008;123:1417–21.
56. Catala-Lopez F, Suarez-Pinilla M, Suarez-Pinilla P, Valderas JM, Gomez-Beneyto M, Martinez S, et al. Inverse and direct cancer comorbidity in people with central nervous system disorders: a meta-analysis of cancer incidence in 577,013 participants of 50 observational studies. Psychother Psychosom 2014;83:89–105.
57. Zhang B, Shu XO, Delahanty RJ, Zeng C, Michailidou K, Bolla MK, et al. Height and breast cancer risk: evidence from prospective studies and mendelian randomization. J Natl Cancer Inst 2015;107(11). pii: djv219.
58. Ni H, Liu H, Gao R. Serum lipids and breast cancer risk: a meta-analysis of prospective cohort studies. PLoS One 2015;10:e0142669.
59. Chandler PD, Song Y, Lin J, Zhang S, Sesso HD, Mora S, et al. Lipid biomarkers and long-term risk of cancer in the Women's Health Study. Am J Clin Nutr 2016;103:1397–407.
60. Doll R, Hill AB. Lung cancer and other causes of death in relation to smoking; a second report on the mortality of British doctors. Br Med J 1956;2:1071–81.
61. Lee JY, Jeon I, Kim JW, Song YS, Yoon JM, Park SM. Diabetes mellitus and ovarian cancer risk: a systematic review and meta-analysis of observational studies. Int J Gynecol Cancer 2013;23:402–12.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Quantifying the Genetic Correlation between Multiple Cancer Types

Sara Lindström, Hilary Finucane, Brendan Bulik-Sullivan, et al.

Cancer Epidemiol Biomarkers Prev 2017;26:1427-1435. Published OnlineFirst June 21, 2017.