The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/97229

Please be advised that this information was generated on 2020-03-17 and may be subject to change.
Polymorphisms in Stromal Genes and Susceptibility to Serous Epithelial Ovarian Cancer: A Report from the Ovarian Cancer Association Consortium

Ernest K. Amank wah1, Qinggang Wang1, Joellen M. Schildkraut2, Ya-Yu Tsa i3, Susan J. Ramaus4, Brooke L. Fridley5, Jonathan Beesley6, Sharon E. Johnatty6, Penelope M. Webb6, Georgia Chenevix-Trench6, Australian Ovarian Cancer Study Group6,7, Laura C. Dale1, Diether Lambrechts8, Frederic Amant9, Evelyn Despierre9, Ignace Vergote9, Simon A. Gayther9, Aleksandra Gentry-Maharaj4, Usha Menon4, Jenny Chang-Claude10, Shan Wang-Gohrke11, Hoda Anton-Culver12, Argyrios Ziegas12, Thilo Dör k13, Matthias Dürst14, Natalia Antonen kova15, Natalia Bogdanova13,15, Robert Brown16, James M. Flanagan14, Stanley B. Kaye17, Department of Pathology, University of California, San Francisco, San Francisco, California, United States of America, 17Department of Pathology, University of Helsinki, Haartman Institute, Helsinki, Finland, 18Department of Obstetrics and Gynecology, University of Helsinki, Central Hospital, Helsinki, Finland, 19Department of Obstetrics and Gynecology, Helsinki University, Central Hospital, Helsinki, Finland, 20Department of Obstetrics and Gynecology, Helsinki University, Central Hospital, Helsinki, Finland, 21Department of Pathology, University of Melbourne, Parkville, Australia, 22Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, United Kingdom, 23Women's Cancer Research Institute at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States of America, 24Strangeways Research Laboratory, Cancer Research United Kingdom, Department of Oncology, University of Cambridge, Cambridge, United Kingdom, 25Danish Cancer Society, Copenhagen, Denmark, 26Gynecologic Clinic, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark, 27Danish Cancer Biobank, Copenhagen and Department of Pathology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark, 28Department of Pathology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark, 29Department of Epidemiology, Biostatistics, and Health Technology Assessment, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, 30Department of Obstetrics and Gynecology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, 31Cancer Control Research, British Columbia Cancer Agency, Vancouver, British Columbia, Canada, 32Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia, Canada, 33Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, British Columbia, Canada, 34Division of Epidemiology and Biostatistics, University of New Mexico, Albuquerque, New Mexico, United States of America, 35Department of Pathology and Laboratory Medicine, Mayo Clinic, Rochester, Minnesota, United States of America, 36Departments of Oncology and Medical Genetics, University of Calgary, Calgary, Alberta, Canada

1Department of Population Health Research, Alberta Health Services-Cancer Care, Calgary, Alberta, Canada, 2Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina, United States of America, 3Division of Cancer Prevention and Control, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, United States of America, 4Department of Gynaecological Oncology, UCL EGA Institute for Women's Health, University College London, London, United Kingdom, 5Department of Science Research, Mayo Clinic, Rochester, Minnesota, United States of America, 6The Queensland Institute of Medical Research, Post Office Royal Brisbane Hospital, University of Queensland, Queensland, Australia, 7Vassilis Research Center, VIB and KU Leuven, Leuven, Belgium, 9Department of Obstetrics and Gynecology, University Hospitals Leuven, Leuven, Belgium, 10Department of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany, 11Department of Obstetrics and Gynecology, University of Utrecht, Utrecht, The Netherlands, 12Department of Epidemiology, School of Medicine, University of California Irvine, Irvine, California, United States of America, 13Clinics of Obstetrics and Gynaecology, Hanover Medical School, Hanover, Germany, 14Department of Gynaecology, Jena University Hospital, Jena, Germany, 15Byelorussian Institute for Oncology and Medical Radiology Aleksandrov N.N., Minsk, Belarus, 16Epigenetics Unit, Department of Surgery and Cancer, Imperial College London, London, United Kingdom, 17Section of Medicine, Institute Cancer Research, Sutton, United Kingdom, 18The Beatson West of Scotland Cancer Centre, Glasgow University, Glasgow, United Kingdom, 19Department of Pathology, University of Helsinki, Haartman Institute, Helsinki, Finland, 20Department of Obstetrics and Gynecology, Helsinki University, Central Hospital, Helsinki, Finland, 21Department of Pathology, University of Melbourne, Parkville, Australia, 22Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, United Kingdom, 23Women's Cancer Research Institute at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States of America, 24Strangeways Research Laboratory, Cancer Research United Kingdom, Department of Oncology, University of Cambridge, Cambridge, United Kingdom, 25Danish Cancer Society, Copenhagen, Denmark, 26Gynecologic Clinic, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark, 27Danish Cancer Biobank, Copenhagen and Department of Pathology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark, 28Department of Pathology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark, 29Department of Epidemiology, Biostatistics, and Health Technology Assessment, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, 30Department of Obstetrics and Gynecology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, 31Cancer Control Research, British Columbia Cancer Agency, Vancouver, British Columbia, Canada, 32Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia, Canada, 33Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, British Columbia, Canada, 34Division of Epidemiology and Biostatistics, University of New Mexico, Albuquerque, New Mexico, United States of America, 35Department of Pathology and Laboratory Medicine, Mayo Clinic, Rochester, Minnesota, United States of America, 36Departments of Oncology and Medical Genetics, University of Calgary, Calgary, Alberta, Canada

Abstract

Alterations in stromal tissue components can inhibit or promote epithelial tumorigenesis. Decorin (DCN) and lumican (LUM) show reduced stromal expression in serous epithelial ovarian cancer (SEOC). We hypothesized that common variants in these genes associate with risk. Associations with SEOC among Caucasians were estimated with odds ratios (OR) among 397 cases and 920 controls in two U.S.-based studies (discovery set), 436 cases and 1,098 controls in Australia (replication set 1) and a consortium of 15 studies comprising 1,668 cases and 4,249 controls (replication set 2). The discovery set and replication set 1 (833 cases and 2,013 controls) showed statistically homogeneous (Pinteraction = 0.48) decreased risks of SEOC at four variants: DCN rs13188156, rs13112816 and rs516115, and LUM rs17018765 (OR = 0.6 to 0.9; Pinteraction = 0.001 to 0.03). Results from replication set 2 were statistically homogeneous (Pinteraction = 0.13) and associated with increased risks at DCN rs13188156 and LUM rs17018765: all ORs = 1.2; Pinteraction = 0.02. The ORs at the four variants were statistically heterogeneous across all 18 studies (Pinteraction = 0.03), which precluded combining. In post-hoc analyses, interactions were observed between each variant and recruitment period (Pinteraction = 0.003), age at diagnosis (Pinteraction = 0.04), and year of diagnosis (Pinteraction = 0.05) in the five studies with available information (1,044 cases, 2,469 controls). We conclude that variants in DCN and LUM are not directly associated with sEOC, and that confirmation of possible effect modification of the variants by non-genetic factors is required.
Introduction

Cancers at the ovary are the most lethal gynecologic cancer, with 21,630 new cases and 15,520 deaths in the U.S. in 2008 [1]. Most (>95%) ovarian cancers are epithelial in origin, affecting cells on the surface of the ovary [2], which are separated from the underlying ovarian stromal tissue by a basal lamina. The stroma is the supportive framework of biologic tissue consisting of an extracellular matrix (ECM) and soluble growth factors that mediate epithelial-stromal interactions and regulate intercellular communication [3]. Activation of oncogenes and inhibition of tumor suppressor genes in the epithelium were previously considered to be the only alterations required for the development of epithelial cancers [4]; however, alterations in stromal components that disrupt normal cell functions can lead to morphologic changes that manifest as tumors through perturbation of the epithelium [5]. For example, radiation-induced changes in the stromal microenvironment have been shown to contribute to neoplastic progression of initiated mammary epithelial cells in vivo [5], and may include processes that activate transforming growth factor-beta (TGF-β) and initiate ECM remodelling [6,7].

The ECM is composed of different proteins: decorin and lumican are members of the small leucine-rich proteoglycan family that bind to collagen in the stroma and are involved in matrix assembly and structure, and in the control of cell proliferation [8]. The expression of both decorin and lumican is altered in various cancers [9,10], including serous epithelial ovarian cancer [11–14]. Conceivably, factors that alter epithelial-stromal interactions or the cross-talk among growth factors like TGF-β may also influence expression and/or activity of decorin or lumican, or vice versa. Such factors may include inherited genetic susceptibility. This could be particularly germane to decorin, which binds to TGF-β and serves as a regulatory control for TGF-β release and activation [15].

In view of the important role of the stroma in epithelial cancers and the role of decorin and lumican in tumorigenesis, we tested the hypothesis that inherited variation in DCN and LUM may influence the risk of serous epithelial ovarian cancer in 18 independent study populations: a discovery set that included studies from Mayo Clinic (MAY) and the North Carolina Ovarian Cancer (NCO) study, replication set 1 from Australia (AUS), and replication set 2 comprised of 12 matched studies from the Ovarian Cancer Association Consortium (OCAC).

Results

The distributions of selected covariates between cases and controls in the discovery set and replication set 1 are listed in Table 1. Covariates were distributed similarly between the discovery set and replication set 1, including the proportion of serous carcinomas across tumor stage. The MAFs for the 10 tagSNPs in the discovery set ranged from 0.08 to 0.29 among controls and were similar in replication set 1 for those SNPs in common (Table S1).

In the discovery set, decreased risks were associated with serous epithelial ovarian cancer under both co-dominant and ordinal models at DCN rs3138165, DCN rs13312816, DCN rs516115 and LUM rs17018765 (all four SNPs: $P_{	ext{trend}} = 0.009$) (Table 2). Associations at all SNPs interrogated in the discovery set are in Table S2. No statistically significant associations were found in haplotype analyses (Table S3).

In replication set 1, decreased risk associations were found under both co-dominant and ordinal models at DCN rs3138165 ($P_{	ext{trend}} = 0.009$), DCN rs13312816 ($P_{	ext{trend}} = 0.01$) and LUM rs17018765 ($P_{	ext{trend}} = 0.008$) but not at DCN rs516115 ($P_{	ext{trend}} = 0.20$) or DCN rs741212 ($P_{	ext{trend}} = 0.61$) (Table 2). Imputed genotypes tended to assume similar risk associations as typed SNPs, likely from high LD among these variants (Figure S1). The squared correlation between imputed and true genotypes was 0.73 for DCN rs3138165, 0.73 for DCN rs13312816, and 0.68 for LUM rs17018765.

The discovery set and replication set 1 were combined in the absence of OR heterogeneity ($P_{	ext{heterogeneity}} = 0.48$) to increase statistical power. The decreased risk associations remained evident at DCN rs3138165 (OR = 0.7; $P_{	ext{trend}} = 0.002$), DCN rs13312816 (OR = 0.7; $P_{	ext{trend}} = 0.002$), DCN rs516115 (OR = 0.9; $P_{	ext{trend}} = 0.03$) and LUM rs17018765 (OR = 0.6; $P_{	ext{trend}} = 0.001$) under both co-dominant and ordinal models (Table 2).

Associations at these four SNPs were tested further in the OCAC replication set 2, and DCN rs3138165 and rs13312816, and LUM rs17018765 showed statistically significant increased risks (Figure 1, Table S4). Within the OCAC replication set 2, the ORs were statistically homogeneous ($P_{	ext{heterogeneity}} = 0.13$), but not when combined with the discovery set and replication set 1 ($P_{	ext{heterogeneity}} = 0.001$ to 0.03). Heterogeneous ORs were not due to errors in allele coding.

Information on age at diagnosis (or age at interview for controls) and on years of study recruitment was available for all studies and was used to test for SNP interactions in post-hoc analyses in an effort to explain the OR heterogeneity across studies (Table 3). For example, the interaction between DCN rs3138165 and age group was suggestive ($P_{	ext{interaction with age}} = 0.04$) and per-minor allele associations were highest among women age <40 years (OR = 2.1; $P = 0.01$; 104 cases) and lowest among women age ≥70 years.
particularly true for definition of an important change-in-estimate effect [16]. This is interaction term were 19% for interaction term. Likewise, changes in the coefficient for the 0.003). However, there was a 20% change in the coefficient for the (interaction for continuous years=0.07 and Pinteraction for binary variable=0.002 to 0.01). Because of the recruitment before 2000, and 0.9 (P = 0.07; 1,494 cases) after 2000 1.3 (P = 0.01; 1,007 cases) for studies with a median year of recruitment before 2000 (OR = 1.3; P = 0.03; 612 cases), and after 2000 (OR = 0.8; P = 0.24; 510 cases). Results were similar for the three other SNPs (Pinteraction = 0.31), parity (Pinteraction = 0.16), BMI (Pinteraction = 0.12), menopausal status (Pinteraction = 0.41), age at menarche (Pinteraction = 0.62) and family history (Pinteraction = 0.24). Per-minor allele associations by diagnosis year mimicked those of recruitment period (diagnoses prior to year 2000: OR = 1.5; P = 0.14; 148 cases and diagnoses after year 2000: OR = 0.8; P = 0.02; 896 cases). Although the interactions were not statistically significant, per-minor allele associations at DCN rs3138165 were associated with lower risk among women who ever took OC hormones (OR = 0.7; P = 0.03; 604 cases), among women with ≥3 full-term births (OR = 0.7; P = 0.01; 573 cases), and among lean women (OR = 0.6; P = 0.03; 272 cases). Associations were similar for the three other SNPs (data not shown). Diagnosis year (1993–1999 vs 2000–2006) was associated with OC use (P < 0.0001), parity (P < 0.0001), BMI (P < 0.0001), menopausal status (P < 0.0001), age at menarche (P < 0.0001) and family history (P < 0.0001), but not with age group (P = 0.32) or DCN rs3138165 (P = 0.54), in unadjusted models. In a model fitting all covariates including DCN rs3138165 and site, only site (P < 0.0001) and age at menarche (P = 0.02) were significantly associated with diagnosis year (data not shown).

**Discussion**

We could not confirm the associations of SNPs in two stromal genes, DCN and LUM, with the risk of serous epithelial ovarian cancer, the most common histological type of ovarian cancer, using a multi-stage replication approach within the OCAC. We found decreased risks at four SNPs in the discovery set and replication set 1, and increased risks in a larger sample of cases in the OCAC replication set 2. The heterogeneity in associations...
across studies was statistically significant and was explained, in part, by the period of case recruitment, with the four SNPs imparting up to a 30% increased risk for diagnoses before the year 2000 and up to a 20% decreased risk after the year 2000. Weaker interactions were seen with age at diagnosis and with diagnosis year in post-hoc analyses. Non-statistically significant modifying effects of OC use, parity and BMI were also observed.

Age-adjusted incidence rates of epithelial ovarian cancer have been trending lower in most of North America and Europe since the 1990s [17,18], and since our gene pool is not changing over such a short period, we speculated that our results reflect changes in the environment. As expected, there was no statistically significant association of diagnosis year with DCN rs3138165, although there were significant associations with several of the covariates and one of these, age at menarche, remained statistically significant with diagnosis year in the multivariable-adjusted model. Our findings may suggest that temporal changes in risk factors are modifiers of inherited susceptibility in DCN and LUM. However, we cannot exclude the role of unmeasured factors that are related to diagnosis year or to study site, or that our findings are due to chance. Several of the 15 studies in replication set 2 are new to OCAC, and epidemiological variables for subjects have not yet been submitted to the central Data Coordinating Center. We were, therefore, under-powered to evaluate gene-environment interactions and can only speculate that age, increasing obesity [19], changing trends in OC hormone preparation and use [20,21], or increasing age at pregnancy/decreasing parity [22] may be obvious candidates for future testing of temporal changes that may modify risk associations of these SNPs. Each of these hormonally-related factors is associated with ovarian cancer [23–26].

### Table 2. Odds ratios (OR) and 95% confidence intervals (CI)\(^a\) between variants in DCN and LUM genes and serous epithelial ovarian cancer risk.

<table>
<thead>
<tr>
<th>Gene/SNP rsID</th>
<th>Discovery Set</th>
<th>Replication Set 1</th>
<th>Combined Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAY-NCO</td>
<td>AUS</td>
<td>MAY-NCO and AUS</td>
</tr>
<tr>
<td></td>
<td>Case/control</td>
<td>OR (95%CI)</td>
<td>P</td>
</tr>
<tr>
<td>DCN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs3138165(^b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>353/778</td>
<td>1.0 (Ref)</td>
<td>395/939</td>
</tr>
<tr>
<td>GA</td>
<td>43/138</td>
<td>0.7 (0.5–1.0)</td>
<td>40/152</td>
</tr>
<tr>
<td>AA</td>
<td>1/4</td>
<td>0.7 (0.1–6.2)</td>
<td>1/7</td>
</tr>
<tr>
<td>Per allele</td>
<td>0.7 (0.5–1.0)</td>
<td>0.06(^c)</td>
<td>0.6 (0.4–0.9)</td>
</tr>
<tr>
<td>rs13312816(^c)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>353/778</td>
<td>1.0 (Ref)</td>
<td>389/928</td>
</tr>
<tr>
<td>AT</td>
<td>43/138</td>
<td>0.7 (0.5–1.0)</td>
<td>46/163</td>
</tr>
<tr>
<td>TT</td>
<td>1/4</td>
<td>0.7 (0.1–6.2)</td>
<td>1/7</td>
</tr>
<tr>
<td>Per allele</td>
<td>0.7 (0.5–1.0)</td>
<td>0.06(^c)</td>
<td>0.7 (0.5–0.9)</td>
</tr>
<tr>
<td>rs516115</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>224/462</td>
<td>1.0 (Ref)</td>
<td>229/543</td>
</tr>
<tr>
<td>AG</td>
<td>148/390</td>
<td>0.8 (0.6–1.0)</td>
<td>164/442</td>
</tr>
<tr>
<td>GG</td>
<td>23/67</td>
<td>0.7 (0.4–1.2)</td>
<td>37/109</td>
</tr>
<tr>
<td>Per allele</td>
<td>0.7 (0.7–1.0)</td>
<td>0.06(^c)</td>
<td>0.9 (0.8–1.1)</td>
</tr>
<tr>
<td>rs741212</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>313/699</td>
<td>1.0 (Ref)</td>
<td>325/830</td>
</tr>
<tr>
<td>AG</td>
<td>81/210</td>
<td>0.9 (0.6–1.2)</td>
<td>103/247</td>
</tr>
<tr>
<td>GG</td>
<td>3/11</td>
<td>0.7 (0.2–2.4)</td>
<td>8/19</td>
</tr>
<tr>
<td>Per allele</td>
<td>0.9 (0.7–1.1)</td>
<td>0.26(^c)</td>
<td>1.1 (0.8–1.3)</td>
</tr>
<tr>
<td>LUM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17018765(^a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>354/776</td>
<td>1.0 (Ref)</td>
<td>395/939</td>
</tr>
<tr>
<td>GA</td>
<td>42/141</td>
<td>0.7 (0.5–1.0)</td>
<td>40/152</td>
</tr>
<tr>
<td>GG</td>
<td>1/3</td>
<td>0.9 (0.1–8.6)</td>
<td>1/7</td>
</tr>
<tr>
<td>Per allele</td>
<td>0.7 (0.5–1.0)</td>
<td>0.06(^c)</td>
<td>0.6 (0.4–0.9)</td>
</tr>
</tbody>
</table>

\(^a\)Adjusted for region of residence (Minnesota, Iowa, Wisconsin, Illinois, North Dakota, South Dakota and North Carolina) for MAY and NCO studies; and adjusted for site (MAY, NCO and AUS) for the combined analysis. Sample is 1,317 Caucasian subjects in the discovery set and 1,534 Caucasian subjects in replication set 1.

\(^b\)Imputed in replication set 1 (AUS).

\(^c\)Imputed in discovery set (MAY and NCO).

\(^d\)P-value for two degrees of freedom test.

\(^e\)P-value for the ordinal model.

\(^f\)doi:10.1371/journal.pone.0019642.t002
primates receiving progestin alone had higher frequencies of apoptotic ovarian epithelial cells compared to control animals or those receiving estrogen alone [27]. Furthermore, the protective effect of parity may be from exposure to pregnancy hormones such as progesterone that have been speculated to clear the ovarian epithelium of precancerous cells [28]. In the macaque study, the increase in apoptotic cells correlated highly with a shift in expression from TGF-β1 isoform to TGF-β2 and -β3 isoforms in the ovarian surface epithelium [29]. TGF-β isoforms appear to be regulated by a variety of steroid hormones in a tissue-specific manner (reviewed in [29]). In contrast, ovarian carcinomas are frequently resistant to TGF-β-mediated growth inhibition [30,31] and express higher levels of TGF-β1 and TGF-β3, and lower levels of TGF-β2, than normal human ovarian specimens [30], the significance of which is unclear.

Decorin and lumican have multiple biological roles including control of cell proliferation [32]. Interestingly, decorin belongs to the family of secretory glycoproteins known as latent TGF-β-binding proteins (LTBPs) that sequesters the pro-hormone or latent form of TGF-β and prevents it from interacting with its signaling receptors, TβR1 and TβRII [15,33]. LTBPs may facilitate the secretion, storage or activation of latent TGF-βs and serve as a reservoir for concentrated delivery of TGF-βs to receptors [15]. TGF-βs [33] and decorin [34] have been implicated as potent tumor suppressors; however, the diverse array of cellular processes regulated by TGF-βs seems to depend on the microenvironment: for example, promoting apoptosis and inhibiting epithelial growth in normal cells and promoting proliferation and angiogenesis in various cancer models [15,35]. The link between progesterone, TGF-βs and decorin is particularly intriguing within the context of our SNP-environment associations. However, this investigation was not designed to examine SNP-environment interactions and the power to detect significant effect modification with the available sample size was low.

We also compared our results to unpublished findings from two recent genome wide association (GWA) studies of ovarian cancer, but there was no clear support for associations at the SNPs. For example, the four SNPs were not associated statistically with serous epithelial ovarian cancer (ORs = 0.93–0.96; P = 0.28–0.76).

Figure 1. Forest plots for DCN and LUM SNPs and serous epithelial ovarian cancer. Associations represent ORs (95% CI) for the individual study (squares) and study-adjusted pooled (diamonds) estimates. Models are ordinal genetic risk model. HAN-HJO and HAN-HMO were combined for presentation.

doi:10.1371/journal.pone.0019642.g001
The strengths of this study include the multi-stage replication strategy, representing 2,501 total serous epithelial ovarian cancer cases. To reduce the impact of population stratification, our analyses were restricted to known or presumed Caucasians. Although one study (SRO) consisted of mostly Caucasians, our results were unchanged when this study was excluded in sensitivity analyses. The characteristics of the samples from the discovery set and replication set 1 were similarly distributed, as was the period of recruitment, thus reducing the impact of effect modification on the SNP-disease associations in these three studies. By restricting the samples to serous epithelial ovarian cancers, we reduced etiologic heterogeneity that may exist among different histological types of ovarian cancer [38]. Finally, we used statistical techniques to impute untyped SNPs as an efficient approach to include these SNPs in a combined analysis of samples from the discovery set and replication set 1.

The major limitation of this investigation is the absence of epidemiological information for most of the OCAC studies included in this report. Thus, our findings in the post-hoc analyses, while intriguing, require a tempered interpretation. Although MAFs of SNPs were generally similar across OCAC analyses, while intriguing, require a tempered interpretation. The discrepancies in results likely reflect similar challenges in interpretation as the OCAC results and underscore the importance of understanding the distribution of individual-level environmental exposures in genetic studies [37].

In summary, our multi-stage replication investigation suggests that SNPs in DCN and LUM are not associated with serous epithelial ovarian cancer. Verification of possible effect modification by age and other unconfirmed temporal effects is underway in an OCAC investigation of 10,000 cases and 10,000 controls.

Materials and Methods

Ethics statement

Participants in all the studies provided written informed consent and each site’s institutional review board approved the study.
protocol (Text S1), including Ethics Committees of the Queensland Institute of Medical Research and Peter MacCallum Cancer Centre (AUS); the local Ethical Committee (Commissie Medische Ethiek UZ Leuven, Belgium) (BEL); and Ethics Committee of the University of Heidelberg (GER).

Discovery set

The discovery set consisted of a combination of two individual studies of epithelial ovarian cancer from MAY and NCO in the United States. Details of the study protocols have been published previously [39]. Briefly, participants included Caucasians and African-Americans enrolled between June 1999 and March 2006 (see Table S5 for detailed study descriptions). In both studies, cases were newly diagnosed, histologically-confirmed, either borderline or invasive, and enrolled within one year of diagnosis. Controls had at least one intact ovary, no history of ovarian cancer and were frequency matched to cases on age (5-yr age categories), race and state of residence.

Replication set 1

The Australian Ovarian Cancer Study recruited cases diagnosed between January 2002 and June 2006 from surgical treatment centers and cancer registries throughout Australia [40]; recruitment through the New South Wales and Victorian Cancer Registries was conducted under the Australian Cancer Study [41] (together, they form the AUS study). Controls were population-based and were randomly selected from the Australian electoral roll and frequency matched to cases on age and state of residence (Table S5).

Ovarian Cancer Association Consortium (OCAC) replication set 2

Fifteen studies from Belgium (BEL), Canada (OVA), Denmark (MAL, PVD), Finland (HOC), Germany (GER, HAN-HJO, HAN-HMO), Netherlands (NTH), the United Kingdom (SEA, SOC, SRO, UKO) and the United States (LAX, UCI), comprising 4,536 primary epithelial ovarian cancer cases and 4,622 controls for whom genotype data were available, were used as a second replication set (Table S5). Most were case-control studies, although some of these studies (PVD, SOC, SRO and LAX) consisted of cases-only and were matched by region within country to unique controls from other OCAC studies. Thus, SOC cases were matched to UKO controls, SRO cases to UKO and SEA controls, LAX cases to UCI controls and PVD cases to MAL controls, resulting in 12 matched studies for analysis.

Information on established risk factors (reproductive history, family history of cancer, medical history, and lifestyle habits) including diagnosis year was collected in the discovery set and replication set 1 and was available for two replication set 2 studies (GER, UCI).

SNP selection, genotyping and quality control (QC)

Discovery set.

Tag single nucleotide polymorphisms (SNPs) were chosen from unrelated Caucasian samples within HapMap Consortium’s release 22 [42] as previously described [43], and also for their predicted likelihood of successful genotyping using the Illumina Golden Gate Assay™. We identified six tagSNPs from among 22 DCN SNPs and seven tagSNPs from among 16 LUM SNPs with minor allele frequency (MAF) $\geq 0.05$ and pairwise linkage disequilibrium (LD) of $r^2 \geq 0.8$. One tagSNP in each gene was predicted to assay poorly (design score = 0), was a singleton in its bin and could not be replaced. This left five tagSNPs in DCN (rs3138165, rs516115, rs10492230, rs741212, and rs3138268) and six tagSNPs in LUM (rs17714469, rs1920790, rs2268578, rs10859110, rs10745553, and rs17018765), including five putative functional SNPs, for genotyping. The SNPs were located within, and 5 kb upstream and downstream of, each gene region. The two genes comprise a contiguous segment on chromosome 12 of approximately 80 kb.

The 11 tagSNPs were genotyped as part of a larger investigation of 1,152 SNPs in a variety of pathways using the Illumina Golden Gate Assay™ and Illumina BeadStudio software [44]. Genotyping was attempted on 897 DNA samples from MAY and 1,279 samples from NCO (total = 2,176 including 129 duplicate samples) and 65 laboratory controls. Case status and duplicate samples were blinded to laboratory personnel who performed the genotyping. Of these samples, we excluded 44 with call rates <90% and Illumina QC (GenCall) scores <0.25, and 22 ineligible or mislabelled samples, resulting in 1,981 unique samples that were successfully genotyped. The sample call rate was 99.74% and...
the concordance for duplicate samples was 99.99%. DCN rs3138268, a nonsynonymous SNP, was monomorphic and was excluded from further analyses. The remaining 10 tagSNPs were genotyped successfully.

**Replication set 1.** Three tagSNPs in DCN (rs13312816, rs516115, and rs741212) were genotyped as part of a larger assay of 1,536 SNPs in AUS (LUM SNPs were not genotyped). Genotyping was attempted on 1,674 samples using the Illumina GoldenGate™ assay and Illumina BeadStudio software. One non-template control and two DNA samples per 96-well plate were blindly duplicated (n = 18). Samples with call rates <95% and SNPs with call rates <98% were excluded. SNPs with GenTrain scores (a metric of genotype clustering) <0.5 were manually checked and adjusted according to Illumina guidelines. Greater than 97% of SNPs passed this initial QC and >84% of all SNPs passed all QC criteria, resulting in 550 cases and 1,101 controls (93% Caucasians) with genotype data on 1,292 SNPs, including the three tagSNPs in DCN included in this analysis.

**Imputation.** SNPs genotyped in the discovery set were not necessarily the same SNPs genotyped in replication set 1 and vice versa (e.g., only DCN rs516115 and DCN rs741212 were genotyped in both datasets). Genotypes at SNPs showing significant associations with ovarian cancer in either dataset were imputed so that datasets could be combined for analysis. Thus, we imputed DCN rs13312816 in the discovery set and DCN rs138165 and LUM rs17018765 in replication set 1 using the MACH software. Briefly, genotype data from the discovery set and replication set 1 were combined with the phase II HapMap data for Caucasian samples and the unobserved genotypes were then inferred probabilistically using a hidden Markov model.

**OCAC replication set 2.** We genotyped four SNPs (DCN SNPs rs3138165, rs13312816 and rs516115 and LUM rs17018765) showing significant associations in the discovery set and replication set 1 in the 12 matched OCAC studies using the Fluidigm EP1 system (Fluidigm, San Francisco, CA) at a central laboratory. Genotyping was performed on 96.96 dynamic arrays in a run of 96 SNPs using inventoried and Custom Assay-by-Design TaqMan probes (Applied Biosystems, Foster City, CA). Genotyping used 10 ng DNA following the manufacturer’s conditions using the pre-amplification protocol. Analysis was performed using Genotyping SNP Analysis software. Samples with call rates <80% were excluded immediately. The following criteria were used as measures of acceptable genotyping for each SNP and each matched study set: i) ≥2% sample duplicates included, ii) concordance for duplicate samples ≥96% and overall concordance for duplicate samples across all SNPs ≥90%, iii) pass rate per plate of ≥90%, iv) <25% overall failed plates, v) overall SNP call rate by study ≥95%, and vi) a difference in call rate between cases and controls of <5%. Studies failing one of these criteria were excluded for that particular SNP, resulting in 3,886 unique samples (4,419 cases and 4,467 controls) that were successfully genotyped. Excellent concordance (100%) in genotype calls was found between study samples and those of 95 HapMap genotyped DNAs (Coriell, Camden, NJ, USA). For all studies, genotyping quality was further assessed using tests for Hardy-Weinberg equilibrium (HWE). SNPs with significant deviations from HWE in Caucasian controls (P < 0.001) were assessed and excluded if the clustering was suboptimal. SNPs with HWE P < 0.001 were excluded from analysis.

**Statistical Analysis.** We restricted analyses to subjects who were self-reported or presumed Caucasian and cases with invasive epithelial ovarian cancer of serous histology, resulting in a final sample size of 1,317 participants in the discovery set (397 cases and 920 controls), 1,534...
participants in replication set 1 (436 cases and 1,098 controls) and 5,917 participants in OCAC replication set 2 (1,660 cases and 4,249 controls). Genotypes were used to estimate allele frequencies and pair-wise LD between SNPs was estimated with r^2 values using Haplovew view version 4.1 [46]. We estimated odds ratios (OR) and 95% confidence intervals (CI) at each SNP using unconditional logistic regression under both co-dominant and ordinal genetic models. In the discovery set only, we also estimated haplotype frequencies for each gene and tested the global statistical significance (P<0.05) for haplotype association [47]. Individual haplotype associations evaluated the risk of serous epithelial ovarian cancer compared to all other haplotypes combined.

Prior to combining data, statistical tests of heterogeneity in the ORs between studies were evaluated. Where heterogeneity existed, statistical significance of interaction in post-hoc analyses was assessed with the Wald test in models that included a product term for the ordinal coding of genotype and categories of age or period of recruitment (before the year 2000 or after the year 2000 based on the median year of the recruitment duration for each study, Figure 3) while adjusting for study. Among the five studies (AUS, GER, MAY, NCO and UCI) with detailed information on covariates, we also examined SNP interactions with diagnosis year, oral contraceptive (OC) use, parity, body mass index (BMI), menopausal status, age at menarche and family history of breast or ovarian cancer in first degree relatives. Missing observations were represented as a separate category within each variable. Associations representing the ordinal genetic model at each SNP were then stratified by the covariate. All models were adjusted for region of residence (discovery set only) or study. Additional adjustment for age categories did not alter associations, so models are presented without age.

All statistical tests were two-sided with an alpha level<0.05 considered statistically significant, and were implemented with SAS (SAS Institute, NC).

Supporting Information

Figure S1 Linkage disequilibrium blocks for tagSNPs in DCN and LUM. Analysis is based on total number of controls from the discovery set and replication set 1. The two genes comprise a contiguous segment on chromosome 12 of approximately 80 kb. Numbers in the squares on the LD block indicate the correlation (r^2) between SNPs. * indicates SNPs that were significantly associated with serous ovarian cancer. (TIF)

Figure S2 Sensitivity analysis for DCN rs13312816 and serous epithelial ovarian cancer stratified by case recruitment period. Forest plots represent associations represent ORs (95% CI) for individual study (squares) and study-adjusted pooled (diamonds) estimates. Models are ordinal genetic risk models. HAN-HJO and HAN-HMO were combined for presentation. P_{het} refers to P value for heterogeneity in odds ratios among studies. (TIF)

Figure S3 Sensitivity analysis for DCN rs516115 and serous epithelial ovarian cancer stratified by case recruitment period. Forest plots represent associations represent ORs (95% CI) for individual study (squares) and study-adjusted pooled (diamonds) estimates. Models are ordinal genetic risk models. HAN-HJO and HAN-HMO were combined for presentation. P_{het} refers to P value for heterogeneity in odds ratios among studies. (TIF)

Figure S4 Sensitivity analysis for LUM rs17018765 and serous epithelial ovarian cancer stratified by case recruitment period. Forest plots represent associations represent ORs (95% CI) for individual study (squares) and study-adjusted pooled (diamonds) estimates. Models are ordinal genetic risk models. HAN-HJO and HAN-HMO were combined for presentation. P_{het} refers to P value for heterogeneity in odds ratios among studies. (TIF)

Table S1 SNP and location, HWE test P-value and MAF for variants in DCN and LUM in 920 controls in the discovery set and 1,098 controls in replication set 1. (DOC)

Table S2 Odds ratios (OR) and 95% confidence intervals (CI) for the association between genetic polymorphisms in DCN and LUM and serous epithelial ovarian cancer risk among 1,317 Caucasian subjects in the discovery set. (DOC)

Table S3 Haplotype analysis of decorin and lumican genes and invasive serous epithelial ovarian cancer risk among 1,317 Caucasian subjects in the discovery set. (DOC)

Table S4 Genotype counts, MAF and HWE statistics and associations between genotypes and risk of ovarian carcinoma for DCN and LUM SNPs among Caucasian subjects in OCAC replication set 2. (DOC)

Table S5 Overview of 18 OCAC studies with serous epithelial ovarian cancer cases and controls. (DOC)

Text S1 Ethics statement. (DOC)

Acknowledgments

We thank Amber Burt for assistance with OCAC data management and Dr Harvey Rich for insightful comments on the hormonal etiology of ovarian cancer. The AOCS Management Group (D Bowtell, G Chenexv-Trench, A deFazio, D Gertig, A Green, and P Webb) gratefully acknowledges the contribution of all the clinical and scientific collaborators (see http://www.aocsstudy.org/). The AOCS and ACS Management Group (A Green, P Parsons, N Hayward, P Webb, and D Whiteman) thank all of the project staff, collaborating institutions and study participants. The HAN-HJO study gratefully acknowledges the contribution of Drs Frauke Kramer and Wen Zheng to the recruitment of patients at Hannover Medical School. The HAN-HMO study gratefully acknowledges the help of Lena Gacucova in sample preparation. The OVA study gratefully acknowledges Barbara Jamieson, Donna Kan and Rozmim Janoo-Gilani for their assistance with data management and sample preparation. The SEA investigators thank the SEARCH team and the Eastern Cancer Registration and Information Centre for patient recruitment. SRO gratefully acknowledges support from Scottish Gynaecological Clinical Trials Group investigators. The UKO investigators thank Ian Jacobs, Eva Wozniak, Andy Ryan, Jeremy Ford, Nayala Balogun and Martin Widschwendter for their contribution to the study, and Clarissa Gandia, Eva Wozniak and Chris Jones who preformed the replication set 2 genotyping and bioinformatics.

Author Contributions

Conceived and designed the experiments: TAS ELG JMS AB PMW GC-T. Performed the experiments: JMC JB SEJ SJR LCD RAV ESI. Collected the samples: DL FA ED IV AG-M UM JC-C SW-G. Analyzed the data: QW Y-YT BLF SJR LCD RAV ESI LEK. Wrote the manuscript: SAG PDPP LEK. Conceived and designed the experiments: TAS ELG JMS AB PMW GC-T. Performed the experiments: JMC JB SEJ SJR. Analyzed the data: QW Y-YT BLF SJR LCD RAV ESI LEK. Wrote the manuscript: SAG PDPP LEK. Collected the samples: DL FA ED IV AG-M UM JC-C SW-G. Acknowledgments...
References


