**LIN28B Polymorphisms Influence Susceptibility to Epithelial Ovarian Cancer**


**Abstract**

Defective microRNA (miRNA) biogenesis contributes to the development and progression of epithelial ovarian cancer (EOC). In this study, we examined the hypothesis that single nucleotide polymorphisms (SNP) in miRNA biogenesis genes may influence EOC risk. In an initial investigation, 318 SNPs in 18 genes were evaluated among 1,815 EOC cases and 1,900 controls, followed up by a replicative joint meta-analysis of data from an additional 2,172 cases and 3,052 controls. Of 23 SNPs from 9 genes associated with risk (empirical $P < 0.05$) in the initial investigation, the meta-analysis replicated 6 SNPs from the DROSHA, FAHR1, LIN28, and LIN28B genes, including rs12194974 (G>A), an SNP in a putative transcription factor binding site in the LIN28B promoter region (summary OR = 0.90, 95% CI: 0.82–0.98; $P = 0.015$) which has been recently implicated in age of menarche and other phenotypes. Consistent with reports that LIN28B overexpression in EOC contributes to tumorigenesis by repressing tumor suppressor let-7 expression, we provide data from luciferase reporter assays and quantitative RT-PCR to suggest that the inverse association among rs12194974 A allele carriers may be because of reduced LIN28B expression. Our findings suggest that variants in LIN28B and possibly other miRNA biogenesis genes may influence EOC susceptibility. Cancer Res; 71(11); 3896–903. ©2011 AACR

**Introduction**

Epithelial ovarian cancer (EOC) is an important cause of morbidity and mortality (1). Given that most cases present with advanced, incurable disease, high-risk women should be targeted for primary prevention and early detection. Identification of genetic susceptibility markers may help in this endeavor, and interindividual differences impacting posttranscriptional regulation is largely unexplored.

MicroRNAs (miRNA) are a class of evolutionarily conserved 19- to 25-nucleotide long nonprotein-encoding RNAs that function as posttranscriptional regulators by binding to the 3′-untranslated region (UTR) of target mRNA, causing translational inhibition and/or mRNA degradation (2, 3). Ovarian
tumors exhibit differential miRNA expression profiles compared with normal tissues (4), and defective miRNA biogenesis may influence miRNA expression, tumorigenesis, and clinical outcomes (2, 5). In EOC, downregulation of Dicer and DROSHA1 and upregulation of LIN28 and LIN28B have been associated with altered miRNA levels and reduced patient survival (6–9). Single nucleotide polymorphisms (SNP) in miRNA biogenesis genes may influence cancer risk by altering miRNA expression (10), and associations between such SNPs and cancer risk have been reported (11–15). While this article was under review, Liang and colleagues (15) published results on the association between EOC risk and 238 SNPs in 8 miRNA biogenesis genes (DROSHA, DCCR8, Dicer1, EIF2C1/AGO1, GEMIN4, RAN, RNASEN/DROSHA, and XPO5) and miRNA binding sites of 134 cancer-related genes in 339 EOC cases and 349 controls. Six SNPs from 4 miRNA biogenesis genes (DROSHA, DCCR8, GEMIN4, and XPO5) were found to be significantly associated with risk (P < 0.05; ref. 15).

In the current investigation, we used a 2-stage approach to identify alleles in miRNA biogenesis genes that may be associated with EOC risk. In stage 1, the Illumina Infinium 610K array was used to genotype 336 SNPs in 1,815 EOC cases (59% with serous carcinomas) and 1,900 controls of European ancestry. To enhance power to detect true susceptibility loci, we conducted a joint meta-analysis of data from all evaluated SNPs in an additional 2,172 invasive EOC cases (49% with serous carcinomas) and 3,052 controls of European ancestry from 8 studies that are part of the international Ovarian Cancer Association Consortium (OCAC).

Materials and Methods

Participants

Characteristics of participating studies are summarized in Supplementary Table S1. Stage 1 comprised 4 case–control studies (MAY, NCO, TBO, TOR). Cases had incident, pathologically confirmed primary invasive EOC. Controls had at least one ovary intact when interviewed, and were frequency-matched to cases on age group and race. To increase etiologic homogeneity, we excluded cases with nonepithelial or borderline tumors, nonwhites, known BRCA1 and BRCA2 mutation carriers, women of Ashkenazi Jewish background, and women with a prior history of ovarian, breast, endometrial, or early-onset colorectal cancer. The stage 2 replication included 8 additional studies from the OCAC whose eligibility criteria have been previously described (16–18). All studies had data on disease status, age at diagnosis/interview, self-reported ethnic group, and histologic subtype. The study protocol was approved by the institutional or ethics review board at each site, and all participants provided written informed consent.

Genotyping

Genomic DNA was extracted from blood, processed, and stored by using standard procedures. All stage 1 samples were genotyped with the Illumina Infinium 610K Array at the Mayo Clinic Genotyping Shared Resource Facility (Rochester, MN) by laboratory personnel blinded to case–control status. Each 96-well plate contained 375 ng DNA of random mixtures of case and control samples, 2 blind duplicates, and 2 replicates of a CEPH family trio (mother, father, and child) from the Coriell Institute. A quality assurance (QA) panel of 96 SNPs was run on the Illumina Bead Express platform to test sample performance and ensure concordance of replicate samples. Illumina Genome Studio software was used to perform automated clustering and calling for more than 550,000 beadtypes. SNPs were excluded from further analysis if (a) the call rate was <95%, (b) they were monomorphic on manual clustering, or (c) there were unresolved replicate errors. Among 81 pairs of replicate samples, the concordance rate was 99.93%. The overall genotype call rate was 99.7%.

Of 4,150 eligible unique subjects that we attempted to genotype in stage 1, we applied the following exclusions: generation of genotypes at less than 95% of loci (n = 394), failure on the QA panel (n = 15), ambiguous gender (n = 7), unresolved identical genotypes (n = 8), self-reported non-European ancestry (n = 2), or less than 80% European ancestry (n = 9) using STRUCTURE (19). This resulted in a final sample size of 3,715 subjects (1,815 cases and 1,900 controls). As shown in Supplementary Table S1, the second stage replication included the New England Case–Control Study (NEC), the Polish Ovarian Cancer Study (POL), 4 U.K.-based collections of EOC cases, and 2 collections of controls from U.K. genome-wide association studies of other phenotypes. Genotyping for NEC was done at the National Institute of Aging (Bethesda, Maryland) with the Illumina 317K and 370K Arrays, and genotyping for POL was done at the Core Genotyping Facility, SAIC-Frederick, Inc. (NCI-Frederick, Maryland) with the Illumina Human 660w-Quad Array. As described previously (16), genotyping for the U.K. cases was conducted using the Illumina Infinium 610K array at the Illumina Corporation. U.K. control data come from the Wellcome Trust Case–Control Consortium 1958 Birth Cohort (20) and a national colorectal control study (21) that used the Illumina 550K array for genotyping. Genotyping quality control procedures for these studies have been described in detail (16).

Identification of candidate miRNA biogenesis genes and SNPs

Genes involved in miRNA biogenesis were identified through a comprehensive PubMed search. A total of 19 genes were selected for evaluation in this investigation: AGO1, AGO2, DCP1, DCCR8, Dicer, DROSHA, FMR1, GEMIN3, GEMIN4, GW182, HIWI, LIN28, LIN28B, PACT, RAN, SMAD4, TRBP, TRIM32, and XPO5. A total of 357 SNPs in all but one gene of interest (TRBP) are included on the 610K array and of these, 336 (94.1%) had call rates more than 95%. Of the 336 SNPs, 318 had minor allele frequency (MAF) of at least 1% and were included in statistical analyses.

Statistical analyses

Descriptive statistics were generated by using frequencies and percents for categorical variables, and mean and SD for continuous variables. Distributions of covariates among cases and controls were compared within each study site by using χ² tests and t tests for categorical and continuous variables.
respectively. Genotype frequencies among controls were tested for Hardy–Weinberg equilibrium (HWE) by using $\chi^2$ goodness-of-fit tests. Statistical analysis was carried out by using SAS Version 9.1 (SAS Inc.) and PLINK software (22).

To examine associations between each SNP and EOC risk, unconditional logistic regression was used to estimate OR and 95% CI. In stage 1, log-additive, dominant, and recessive genetic models were fit for each SNP; the major allele was considered to be the reference allele during modeling. All models were adjusted for age, study site, and a quantitative variable for the first principal component representing European ancestry. Wald $\chi^2$ tests were used to obtain $P$ values for dominant and recessive SNP effects, and the Cochran–Armitage trend test was used to estimate $P$ for trend for log-additive effects. For each SNP, the minimum $P$ value over genetic models was used to represent the best-fitting model. $P < 0.05$ was considered the threshold of significance (as opposed to $P \leq 10^{-7}$ required for genome-wide significance) because of our a priori hypothesis that this subset of SNPs may be associated with EOC risk. Similar to other studies that investigated associations between individual SNPs in miRNA biogenesis genes and cancer risk (11, 12, 14), we did not adjust for multiple comparisons in an effort to minimize false-negatives. However, we permuted case–control status 100,000 times to generate pointwise empirical $P$ values. Because of the heterogeneous nature of EOC, exploratory subgroup analysis was conducted to estimate genotype-specific odds ratios for serous carcinomas (the most predominant histologic subtype).

For all SNPs evaluated in stage 1, a meta-analysis was carried out in PLINK by combining results from log-additive models across studies by using the Mantel–Haenszel method for combining raw data (23). Combined ORs, 95% CIs, and $P$ values were generated by using fixed-effect and random-effect models. The $I^2$ test of heterogeneity (24) was estimated to quantify the proportion of total variation due to heterogeneity across studies. The $R$ statistical package "r-meta" was used to quantify the proportion of total variation due to heterogeneity (24) was estimated to estimate genotype-specific odds ratios for serous carcinomas (the most predominant histologic subtype).

For all SNPs evaluated in stage 1, a meta-analysis was carried out in PLINK by combining results from log-additive models across studies by using the Mantel–Haenszel method for combining raw data (23). Combined ORs, 95% CIs, and $P$ values were generated by using fixed-effect and random-effect models. The $I^2$ test of heterogeneity (24) was estimated to quantify the proportion of total variation due to heterogeneity across studies. The $R$ statistical package "r-meta" was used to quantify the proportion of total variation due to heterogeneity (24) was estimated to estimate genotype-specific odds ratios for serous carcinomas (the most predominant histologic subtype).

For all SNPs evaluated in stage 1, a meta-analysis was carried out in PLINK by combining results from log-additive models across studies by using the Mantel–Haenszel method for combining raw data (23). Combined ORs, 95% CIs, and $P$ values were generated by using fixed-effect and random-effect models. The $I^2$ test of heterogeneity (24) was estimated to quantify the proportion of total variation due to heterogeneity across studies. The $R$ statistical package "r-meta" was used to quantify the proportion of total variation due to heterogeneity (24) was estimated to estimate genotype-specific odds ratios for serous carcinomas (the most predominant histologic subtype).

In silico functional assessment of SNPs

The SNPinfo (26) and FastSNP (27) in silico tools were used to predict the functional impact of risk-associated and strongly-correlated SNPs.

Plasmid construction

To evaluate the functional significance of LIN28B rs12194974 (G>A) located 727 bases upstream from the LIN28B transcriptional start site, we generated luciferase reporter constructs by amplifying the 5′–flanking region of LIN28B spanning −1,414 bases upstream of the transcriptional start site to +1,403 bases downstream by using previously described methods (28). The amplified fragment was cloned into a pGL3 luciferase reporter vector to generate pGL3-LIN-28B-P2-IRES, which contained the wild-type allele (−727-G), and served as the template for generating the variant construct (−727-A) by using the QuikChange site-directed mutagenesis (Agilent Technologies, Inc.). The constructs were verified by DNA sequencing, using BigDye Terminator V3.1 Cycle Sequencing Kits and the 3130X Genetic Analyzer (Applied Biosystems).

Cell lines and cultures

Ovarian adenocarcinoma (OV420, OV432, CAOV3, SKOV3, OV8, A2780CP, and A2780S) and T80 human immortalized surface epithelial (HIose) cells were obtained from ATCC and the University of Pennsylvania, and were cultured at 37°C and 5% CO₂. OV8 cells were cultured in RPMI1640; A2780S and A2780CP cells were cultured in DMEM/F12; SKOV3 cells were cultured in McCoy’s 5A; and Ovca420, Ovca432, CAOV3, and T80 HIOSE cells were cultured in M199/MCD8105. All the culture media were supplemented with 10% FBS (Invitrogen Co.).

Reporter gene analysis

Cells were transiently transfected with wild-type and variant LIN28B constructs. The amount of DNA in each transfection was kept constant by the addition of empty vector. After 36 hours of transfection, relative luciferase activity (RLA) was measured by a luciferase assay reagent (Promega). Transfection efficiency was normalized by cotransfection with a β-galactosidase–expressing vector; β-galactosidase activity was measured by Galacto-Light (Tropix). RLA was expressed as relative light units or fold change. All experiments were repeated in triplicate. Differences between wild-type and variant LIN28B reporter activity were tested by Student’s $t$ tests. A 2-sided $P$ value of less than 0.05 was considered statistically significant.

Quantitative RT-PCR

To validate the results from reporter assays, quantitative real-time (RT)-PCR (qRT-PCR) was used to evaluate mRNA expression of LIN28B exons 1 to 3 in the context of −727-G or −727-A promoters. Total RNA was isolated from transfected constructs in EOC cell lines A2780CP and A2780S, using Trizol reagents (Invitrogen). cDNA was prepared from total RNA samples by TaqMan RNA Reverse Transcription kit (Applied Biosystems). qRT-PCR was done by TaqMan Universal PCR Master mix (Applied Biosystems) according to the manufacturer’s protocol. PCR cycling began with template denaturation and hot start Taq polymerase activation at 95°C for 1 minute, then 40 cycles of 95°C for 5 seconds, and 60°C for 30 seconds performed on a 7900HT Fast Real-Time PCR System with data collected during each cycle at the 60°C extension step with 7900HT SDS v2.3 (Applied Biosystems). Threshold and baselines were manually determined with thresholds typically set between 0.05 and 0.1 paired with a baseline starting at 1 to 3 cycle thresholds (Ct) and finishing at 15 to 17 Ct. Relative LIN28B expression was determined by the delta–delta comparative threshold ($\Delta\Delta$Ct) method with actin. $P$ values were determined by a 2-tailed Student’s $t$ test from Ct values for normalization to total RNA.
**Table 1.** MiRNA biogenesis SNPs significantly associated with EOC risk overall and/or serous carcinoma after permutation testing among stage 1 participants

<table>
<thead>
<tr>
<th>Gene (focus)</th>
<th>SNP (major*/minor allele)</th>
<th>MAF&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Model&lt;sup&gt;c&lt;/sup&gt;</th>
<th>All subtypes (1,815 cases, 1,900 controls)</th>
<th>Serous carcinomas (1,070 cases, 1,900 controls)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OR (95% CI)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>P&lt;sup&gt;i&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>AGO2 (8q24.3)</strong></td>
<td>rs1996715&lt;sup&gt;a&lt;/sup&gt; (A/C)</td>
<td>0.48</td>
<td>R</td>
<td>0.85 (0.73–0.99)</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>rs2176397&lt;sup&gt;a&lt;/sup&gt; (G/A)</td>
<td>0.30</td>
<td>R</td>
<td>0.72 (0.57–0.92)</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>rs3864659 (A/C)</td>
<td>0.11</td>
<td>R</td>
<td>2.53 (1.38–4.65)</td>
<td>0.003</td>
</tr>
<tr>
<td><strong>DCP1 (12p13.33)</strong></td>
<td>rs10505725 (G/A)</td>
<td>0.28</td>
<td>R</td>
<td>1.24 (0.98–1.58)</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>rs2240611&lt;sup&gt;a&lt;/sup&gt; (A/G)</td>
<td>0.49</td>
<td>D</td>
<td>0.83 (0.71–0.96)</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>rs6489338&lt;sup&gt;a&lt;/sup&gt; (G/A)</td>
<td>0.46</td>
<td>R</td>
<td>1.20 (1.02–1.40)</td>
<td>0.025</td>
</tr>
<tr>
<td><strong>DICER (14q32.13)</strong></td>
<td>rs1110386&lt;sup&gt;a&lt;/sup&gt; (G/A)</td>
<td>0.22</td>
<td>R</td>
<td>1.53 (1.12–2.09)</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>rs11621737&lt;sup&gt;a&lt;/sup&gt; (G/A)</td>
<td>0.24</td>
<td>R</td>
<td>0.71 (0.53–0.95)</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>rs3353726&lt;sup&gt;a&lt;/sup&gt; (G/A)</td>
<td>0.43</td>
<td>R</td>
<td>0.85 (0.72–1.01)</td>
<td>0.059</td>
</tr>
<tr>
<td><strong>DROSHA (5p13.3)</strong></td>
<td>rs10067066&lt;sup&gt;a&lt;/sup&gt; (A/G)</td>
<td>0.07</td>
<td>R</td>
<td>2.57 (1.04–6.35)</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>rs13186629&lt;sup&gt;a&lt;/sup&gt; (A/G)</td>
<td>0.40</td>
<td>A</td>
<td>0.90 (0.82–0.99)</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>rs17404622&lt;sup&gt;a&lt;/sup&gt; (A/G)</td>
<td>0.37</td>
<td>A</td>
<td>0.90 (0.82–0.99)</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>rs2161006&lt;sup&gt;a&lt;/sup&gt; (A/G)</td>
<td>0.37</td>
<td>A</td>
<td>0.90 (0.81–0.99)</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>rs6450839&lt;sup&gt;a&lt;/sup&gt; (A/G)</td>
<td>0.30</td>
<td>D</td>
<td>1.16 (1.01–1.32)</td>
<td>0.032</td>
</tr>
<tr>
<td><strong>FMR1 (Xq27.3)</strong></td>
<td>rs25704 (G/A)</td>
<td>0.23</td>
<td>A</td>
<td>1.12 (1.01–1.25)</td>
<td>0.041</td>
</tr>
<tr>
<td><strong>LIN28 (1p36.11)</strong></td>
<td>rs11247946&lt;sup&gt;a&lt;/sup&gt; (A/G)</td>
<td>0.34</td>
<td>R</td>
<td>1.25 (1.02–1.52)</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>rs1588746&lt;sup&gt;a&lt;/sup&gt; (A/G)</td>
<td>0.19</td>
<td>D</td>
<td>0.87 (0.72–1.00)</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>rs12728900 (G/A)</td>
<td>0.26</td>
<td>A</td>
<td>0.86 (0.77–0.96)</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>rs6598964&lt;sup&gt;a&lt;/sup&gt; (G/A)</td>
<td>0.36</td>
<td>R</td>
<td>1.26 (1.04–1.53)</td>
<td>0.017</td>
</tr>
<tr>
<td><strong>LIN28B (6q16.3-q21)</strong></td>
<td>rs12194974 (G/A)</td>
<td>0.15</td>
<td>A</td>
<td>0.84 (0.74–0.97)</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>rs17065417 (A/C)</td>
<td>0.09</td>
<td>D</td>
<td>1.23 (1.04–1.46)</td>
<td>0.016</td>
</tr>
<tr>
<td><strong>PACT (2p31.2)</strong></td>
<td>rs2059691&lt;sup&gt;a&lt;/sup&gt; (G/A)</td>
<td>0.30</td>
<td>R</td>
<td>0.77 (0.61–0.97)</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>rs948588&lt;sup&gt;a&lt;/sup&gt; (G/A)</td>
<td>0.08</td>
<td>R</td>
<td>2.80 (1.22–6.40)</td>
<td>0.015</td>
</tr>
</tbody>
</table>

NOTE: Statistically significant P and EMP1 values (<0.05) are in bold type. All P values are 2-sided.

*The major allele represents the most frequently-occurring allele and serves as the reference allele during modeling.

*MAF in controls.

*Best-fitting genetic model (A, log-additive; D, dominant; R, recessive).

*OR and 95% CI adjusted for age, study site, and the first principal component for European ancestry.

*Wald test P value for testing genetic effects related to overall ovarian cancer risk before permutation testing.

*EMP1 value for testing genetic effects related to overall ovarian cancer risk.

*Empirical value for testing genetic effects related to serous carcinoma risk.

*Empirical value for testing the genetic effects related to serous carcinoma risk.

*SNPs deviating from HWE among controls (P<sub>HWE</sub> < 0.05), with P<sub>HWE</sub> values for AGO2 rs11996715, DICER rs1110386 and rs2353726, and PACT rs2059691 of 0.009, 0.007, 0.013, and 0.007, respectively.

*SNP pair in linkage disequilibrium (r<sup>2</sup> = 0.93).

*SNP pair in linkage disequilibrium (0.97 ≥ r<sup>2</sup> ≥ 0.99).
### Results

Despite frequency-matching on age-group, cases were slightly older (60.0 ± 11.6 years) than controls (56.8 ± 12.0 years). As compared with controls, cases tended to have lower education levels, were less likely to have used oral contraceptives, and had longer duration of hormone replacement therapy use (data not shown). The distributions of selected clinical and pathologic characteristics of stage 1 cases are summarized by study site in Supplementary Table S2. Most EOC were of serous histology, were diagnosed at stage III, and were of high grade.

In stage 1, 32 SNPs from 10 of the 18 genes were associated with EOC overall or with serous histology (n = 1,070; Supplementary Table S3). Twenty-three SNPs from 9 genes retained statistical significance (EMP1 < 0.05) after permutation testing (Table 1). SNPs within *DCP1* (rs2240610 and rs489338) and *DROSHA* (rs13186629, rs17404622, and rs2161006) are in LD (r^2 > 0.80). Of 4 SNPs in Table 1 that deviated from HWE, all are common (MAF > 0.10) and normal clustering was observed, suggesting that genotyping errors were unlikely. Most statistically significant associations were based on a recessive or dominant model and/or the minor allele was inversely associated with risk. The most notable association (EMP1 = 0.004) with EOC overall was observed with *AGO2* rs2176397 [AA genotype versus the AG/GG genotype, OR (95% CI) = 0.72 (0.57–0.92)]. Subgroup analysis of cases with serous histology revealed trends in association of similar direction and magnitude as the overall analysis (Table 1).

The joint meta-analysis included data from a total of 3,987 EOC cases (2,135 serous carcinomas) and 4,952 controls (Supplementary Table S1). Point estimates and confidence intervals were similar under fixed-effects and random-effects models (data not shown), suggesting minimal between-study heterogeneity. Under a fixed-effects model, the association for rs12728900 was 0.86 (0.78–0.96), for rs13186629, rs2161006, and rs17404622—3 SNPs in high LD (0.97 ≤ r^2 ≤ 0.99) that flank the 3' UTR of the gene and map to a 17 kb region of LD within 5p13.3 (Supplementary Fig. S1A). rs13186629 showed the strongest evidence of association in this region, with a summary OR of 1.09 (95% CI: 1.02–1.16, P = 0.008, I^2 = 0%). The next strongest association was with

#### Table 1: SNPs and Clinical Characteristics of Stage 1 Cases

<table>
<thead>
<tr>
<th>Gene</th>
<th>Site</th>
<th>OR (95% CI)</th>
<th>n</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs13186629</td>
<td>Stage I</td>
<td>1.11 (1.01–1.22)</td>
<td>3,700</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
<td>1.6</td>
</tr>
<tr>
<td>rs13186629</td>
<td>UK</td>
<td>1.07 (0.98–1.17)</td>
<td>4,122</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
<td>1.6</td>
</tr>
<tr>
<td>rs13186629</td>
<td>POL</td>
<td>1.14 (0.89–1.46)</td>
<td>774</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
<td>1.6</td>
</tr>
<tr>
<td>rs13186629</td>
<td>NEC</td>
<td>0.0</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td>8,596</td>
<td>1.09 (1.02–1.16)</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
<td>1.6</td>
<td>1.7</td>
</tr>
</tbody>
</table>

---

Figure 1. Forest plots for the 6 SNPs associated with EOC risk in the joint meta-analysis. Boxes denote OR point estimates, their areas being proportional to the inverse variance weight of the estimate. Horizontal lines represent 95% CIs. The diamond represents the summary OR computed under a fixed-effects model. The vertical line is at the null value (OR = 1.0).
**LIN28B Polymorphisms and Ovarian Cancer Risk**

LIN28B rs12194974 (summary OR = 0.90, 95% CI: 0.82–0.98, \( P = 0.015, \hat{P} = 21.8\%\)), an SNP flanking the 3′-UTR of LIN28B (Supplementary Fig. S1B). We also observed statistically significant associations for rs25704, an SNP in the 3′-UTR of the X-linked FMR1 gene (summary OR = 1.12, 95% CI: 1.02–1.24, \( P = 0.020, \hat{P} = 0\%\); Supplementary Fig. S1C), and for intronic SNP LIN28 rs12728900 (summary OR = 0.93, 95% CI: 0.86–0.99, \( P = 0.031, \hat{P} = 24.3\%\); Supplementary Fig. S1D).

To predict the possible functional relevance of risk-associated and strongly correlated SNPs (\( \hat{r}^2 > 0.80\)) listed in Supplementary Table S4, we used the SNPinfo (26) and FastSNP (27) *in silico* tools; results suggest that 29 of these 78 SNPs are predicted to influence gene expression. Of the SNPs with evidence for statistical significance in the replication stage, the DROSHA tagSNPs have no known function, rs12194974 is predicted by SNPinfo to reside in a transcription factor binding site in the highly conserved LIN28B promoter, FMR1 rs25704 is located in a putative miRNA binding site according to SNPinfo, and LIN28 rs12728900 is predicted by FastSNP to be an intronic enhancer. Of particular interest is the LIN28B rs12194974 SNP because transcription factors (TF) predicted to bind there include PPARγ (peroxisome proliferator-activated receptor gamma), YY1 (yin yang 1), and PAX2 (paired box gene 2), each of which have been implicated in ovarian carcinogenesis (29–31). On the basis of evidence that LIN28B/LIN28 overexpression negatively regulates biogenesis of let-7, a known tumor-suppressor miRNA family frequently downregulated in EOC and other malignancies (8, 28, 32), we hypothesized that variation in the promoter may affect how TFs activate or repress LIN28B activity, ultimately affecting the level or timing of LIN28B and let-7 expression.

We explored the functional significance of rs12194974 by constructing LIN28B G>A reporters, which contain the first 3 exons of LIN28B following IRES and the luciferase gene (Fig. 2A). Luciferase reporter assay was used to measure promoter activities of wild-type (−727-G) and variant (−727-A) constructs in 8 cell lines. EOC cells transfected with −727-A had 30% to 84% lower RLA than the common −727-G reporter (\( P < 0.05\); Fig. 2B). Real-time RT-PCR of exons 1 to 3 of LIN28B confirmed 24% to 29% reduced relative mRNA LIN28B expression of −727-A compared with −727-G (Fig. 2C). These results suggest that the A allele has lower promoter activity which may lead to downregulation of LIN28B.

**Discussion**

This study examined whether inherited variants in miRNA biogenesis genes contribute to EOC risk by integrating multiple lines of evidence, including statistical, *in silico*, biological, and functional data. Consistent with results from Liang and colleagues (15), we identified SNPs within or near DROSHA as being associated with EOC risk, but we did not identify SNPs in genes evaluated in both studies (AGO1, DGC8, GEMINA, RAN, and XPO5) as risk-associated in our much larger study population. Our findings are noteworthy because several candidate genes highlighted here (LIN28, LIN28B, DROSHA, Dicer, and...
AG02) have been implicated in ovarian carcinogenesis (6–9, 33, 34). Importantly, associations identified with LIN28/LIN28B have biologic relevance because these genes are over-expressed in advanced stage EOC and may contribute to tumorigenesis by repressing let-7 and upregulating let-7 target oncogenes KRAS and MYC (8, 28). Indeed, a study of 211 EOC tumors showed that high LIN28B expression, low let-7a, and high IGF-II were correlated with a poorer prognosis compared with tumors with low LIN28B expression (9), consistent with a role for LIN28B in tumor growth (35).

Our meta-analytic results, which were largely influenced by stage 1, provided evidence of a reduced risk of EOC with the A allele of rs12194974, a variant that we showed experimentally reduces LIN28B transcriptional activity and LIN28B mRNA expression. Because of the low frequency of the rs12194974 A allele, we have not been able to identify human cancer cell lines that harbor this variant, limiting functional analyses. As such, future studies involving knock-in of the rs12194974 A allele into the LIN28B gene may be warranted to evaluate LIN28B expression. Taken together, SNPs in LIN28B may serve as biomarkers of risk, and given the role of LIN28B in stem cell renewal, disrupting its function in tumors where it is over-expressed may promote let-7 expression and therapeutic benefit (8, 9). Interestingly, polymorphisms within LIN28B may also influence the reproductive lifespan and human growth, as SNPs residing in the same LD block as rs12194974 (Supplementary Fig. S1B) have been associated with age at menarche, timing of puberty, height, and finger-length ratio (36–40).

Major strengths of our study include the large sample size, the quality control measures put into place to generate high quality genotype data, and the procedures used to ensure a genetically homogeneous study population of women with European ancestry (i.e., strict inclusion/exclusion criteria and assessment of population substructure). Furthermore, this represents the first report involving miRNA biogenesis SNPs and cancer risk to follow-up on observed associations by conducting a joint meta-analysis and incorporating functional data. As we did not adjust for multiple comparisons, our findings should be considered suggestive that inherited variants in miRNA biogenesis genes, particularly LIN28B, affect EOC risk. Future studies may involve fine mapping of the highlighted regions of interest, using data from the 1000 Genomes Project. Efforts are also underway to replicate the most promising associations reported herein in the OCAC, and to evaluate variants that may impact miRNA primary transcripts or miRNA–target interactions. Knowledge of inherited variation in miRNA-related genes may help to identify high-risk populations and aid in the development of diagnostic, prognostic, and therapeutic strategies for EOC and likely other cancers.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank all of the individuals who participated in this research along with all of the researchers, clinicians, and staff who have contributed to the participating studies.

Grant Support

The genotyping, bioinformatic, and biostatistical data analysis for stage 1 was supported by R01-CA-114343 and R01-CA114343-S1 and the Moffitt Cancer Center/University of South Florida Anna Valentine Fund and Cancer Center Support Grant P30-CA0762-92. The MAY study is supported by R01-CA-122443 and P50-CA-136939 and funding from the Mayo Foundation. The NCO study is supported by R01-CA-76166. The TBO study is supported by R01-CA-106414, the American Cancer Society (CTIG-00-196-01-CCE), and the Advanced Cancer Detection Center Grant, Department of Defense (DAMD-17-98-1-8659). The TOR study is supported by grants from the Canadian Cancer Society and the NIH (R01-CA-63682 and R01-CA-63678). The Mayo Clinic Genotyping Shared Resource is supported by the National Cancer Institute (P30-CA-15083).

Of studies that participated in stage 2, the NEC study is supported by grants CA-54419 and P50 CA16509. The POL study was supported by the Intramural Research Program of the NIH, National Cancer Institute, Division of Cancer Epidemiology and Genetics, and the Center for Cancer Research. The SEA study is funded by a program grant from Cancer Research U.K. The UKO study is supported by funding from Cancer Research U.K., the Eve Appeal, and the OAK Foundation. Some of this work was undertaken at UCLH/UCL, who received some funding from the Department of Health's NIHR Biomedical Research Centre funding scheme. U.K. genotyping and data analysis was supported by a project grant from Cancer Research U.K. U.K. studies also make use of data generated by the Wellcome Trust Case–Control consortium. A list of investigators who contributed to the generation of data is available at www.wtccc.org.uk.

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 November 17, 2010; revised March 17, 2011; accepted March 21, 2011; published OnlineFirst April 11, 2011.

References

3. Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to tumorigenesis by repressing let-7 and upregulating let-7 target oncogenes KRAS and MYC (8, 28). Indeed, a study of 211 EOC tumors showed that high LIN28B expression, low let-7a, and high IGF-II were correlated with a poorer prognosis compared with tumors with low LIN28B expression (9), consistent with a role for LIN28B in tumor growth (35).

www.aacrjournals.org Cancer Res; 71(11) June 1, 2011 3903

Published OnlineFirst April 11, 2011; DOI: 10.1158/0008-5472.CAN-10-4167
LIN28B Polymorphisms Influence Susceptibility to Epithelial Ovarian Cancer

Jennifer Permuth-Wey, Donghwa Kim, Ya-Yu Tsai, et al.

Cancer Res 2011;71:3896-3903. Published OnlineFirst April 11, 2011.

Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-10-4167

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2011/04/13/0008-5472.CAN-10-4167.DC1

Cited articles
This article cites 38 articles, 11 of which you can access for free at:
http://cancerres.aacrjournals.org/content/71/11/3896.full.html#ref-list-1

Citing articles
This article has been cited by 14 HighWire-hosted articles. Access the articles at:
/content/71/11/3896.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.