# Supporting Information

Zhang *et al.* 10.1073/pnas.0801615105

## SI Methods

**Patients and Specimens.** Clinical characteristics were as defined (1). Optimal surgical debulking was ≤1 cm residual individual tumor nodules. Front-line chemotherapy comprised platinum, platinum-cyclophosphamide, or (after 1995) platinum-paclitaxel. Complete response to therapy was defined by normalization of physical examination, abdomino-pelvic CT scan and serum CA-125. Noncomplete response included partial response (≥50% decrease in the sum of greater tumor dimensions by CT) and no response (<50% decrease or any increase in tumor). Progression-free survival was the time between completion of chemotherapy and first recurrence (if a complete response had been achieved) or progression, defined as ≥50% tumor increase by CT scan or two rising CA-125 values. Overall survival was the interval between diagnosis and death. Observation time was the interval between diagnosis and last contact (death or last follow-up). Time was censored at the last follow-up for patients without recurrence, progression or death. Nonserous tumors comprised serous, endometrioid, clear cell, and undifferentiated tumors. All tumors were from primary sites, and were immediately snap-frozen and stored at −80°C. Specimens were acquired and processed under procedures approved by the local IRBs and were compliant with the HIPAA act.

**Cell Lines and Cell Culture.** A total of 18 ovarian cell lines were used in this study. All cancer cell lines were cultured in RPMI medium 1640 (Invitrogen) supplemented with 10% FBS (FBS, Invitrogen). Four IOSE cells, provided by N. Auersperg (2), were cultured in 1:1 media 199: MCDB 105 (Sigma) supplemented with 15% FBS.

**RNA and DNA Isolation.** Total RNA was isolated from 100–500 mg of frozen tissue or $1 \times 10^6$ cultured cells with TRIzol reagent (Invitrogen). The quality and quantity of the isolated RNA was analyzed by using a Bioanalyzer 2100 system (Agilent). Genomic DNA was isolated from frozen tumors or cultured cells by overnight digestion, phenol-chloroform extraction, and ethanol precipitation.

**TaqMan miRNA Assay.** Expression of mature miRNAs were analyzed by TaqMan miRNA Assay (Applied Biosystems) under conditions defined by the supplier (3). Briefly, single-stranded cDNA was synthesized from 5.5 ng of total RNA in 15 μl of reaction volume, using the TaqMan MicroRNA reverse transcription kit (Applied Biosystems). The reactions were incubated first at 16°C for 30 min, then at 42°C for 30 min. The reactions were inactivated by incubation at 85°C for 5 min. Each cDNA generated was amplified by quantitative PCR, using sequence specific primers from the TaqMan microRNA Assays Human Panel on an Applied Biosystems 7900HT Sequence Detection System (Applied Biosystems). The 20-μl PCR included 10 μl of 2× Universal PCR Master Mix (No AmpErase UNG), 2 μl of 10× TaqMan MicroRNA assay mix, and 1.5 μl of RT product. The reactions were incubated in a 384-well plate at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. The threshold cycle ($C_T$) is defined as the fractional cycle number at which the fluorescence passes the fixed threshold (0.2). All signals with $C_T < 37.9$ were set as undetermined.

**miRNA Microarray.** miRNA microarray was performed as described in ref. 4. Briefly, 5 μg of total RNA was reverse transcribed by using biotin end-labeled random-octamer oligonucleotide primer. Hybridization of biotin-labeled complementary DNA was performed on the Ohio State University miRNA microarray chip (OSU_CCC version 3.0), which contains 1100 miRNA probes, including 326 human miRNA genes, spotted in duplicates. Often, more than one probe exists for a given mature miRNA. Additionally, there are quadruplicate probes corresponding to most premiRNAs. The hybridized chips were washed and processed to detect biotin-containing transcripts by streptavidin-Alexa647 conjugate and scanned on an Axon 4000B microarray scanner (Axon Instruments).

**cDNA Microarray.** A two-cycle amplification protocol (Affymetrix) was used to generate sufficient cRNA for microarray analysis according to the recommendations of the manufacturer. Briefly, 10 ng of total RNA was used in the first-round synthesis of double-stranded cDNA. The RNA was reverse transcribed by using a two-cycle cDNA synthesis kit (Affymetrix) followed by amplification, using a MEGAscript T7 kit (Ambion). The cRNA generated from the first-round synthesis was cleaned by using a GeneChip sample clean-up Module IVT column (Affymetrix). The second-round double-strand cDNA amplification was carried out by using an IVT labeling kit (Affymetrix). The resulting biotin-labeled cRNA was purified by using an IVT clean-up kit (Affymetrix) and quantified by using a UV spectrophotometer ($A_{260/280}$; Beckman). An aliquot (15 μg) of cRNA was fragmented by heat and ion-mediated hydrolysis at 94°C for 35 min. Fragmented cRNA was hybridized in a hybridization oven (16 h at 45°C) to a human genome U133 Plus 2.0 GeneChip oligonucleotide array (Affymetrix) representing 47,000 transcripts and variants, including 38,500 well characterized human genes. The washing and staining of the arrays with phycoerythrin-conjugated streptavidin (Molecular Probes) was completed in Fluidics Station 450 (Affymetrix). The arrays were subsequently scanned by using a confocal laser GeneChip Scanner 3000 and GeneChip operating software (Affymetrix).

**Array-Based Comparative Genomic Hybridization (aCGH).** BAC clones included in the 1-Mb array platform were recently described in ref. 5. Briefly, 4,134 clones from the CalTech A/B and RPCI-11 libraries were collected from both commercial and private sources and were mapped to build 34 of the human genome, using either an STS-marker (29%), end sequences (68%), or full sequences (3%). A minimum of two replicates per clone were printed on each slide. One μg of tumor and reference DNA were labeled with Cy3 or Cy5, respectively (Amersham), using the BioPrime random-primed labeling kit (Invitrogen). In parallel experiments, tumor DNA and reference DNA were labeled with the opposite dye, to account for difference in dye incorporation and provide additional data for analysis. A systematic protocol was used to analyze aCGH data for copy number alterations. For quality control purposes, clones demonstrating an adjusted foreground to background intensity ratios of <0.8 in the reference channel were removed. With dye swap data merged as input, copy number breakpoints were estimated for each sample by the Circular Binary Segmentation (CBS) algorithm, using breakpoint significance based on 10,000 permutations (6). Additional analyses and visualization of aCGH data were done by using the *CGHAnalyzer* software suite described in ref. 7.

**Tissue Microarray.** The tissue microarray was constructed as described (8, 9). Total 504 specimens were printed in the array

slides. In brief, tumors were embedded in paraffin and 5-$\mu$m sections were stained with hematoxylin–eosin to select representative regions for biopsies. Four core tissue biopsies were obtained from each specimen. The presence of tumor tissue on the arrayed samples was verified on hematoxylin–eosin stained section.

**Microarray Analysis.** The normalized microarray data were managed and analyzed by GeneSpring (Agilent), microarray software suite 4 (TM4; www.tm4.org) (10), BRB-ArrayTools version 3.6 (linus.nci.nih.gov/BRB-ArrayTools.html), and GenePattern (www.broad.mit.edu/cancer/software/genepattern) (11). Cluster and TreeView (rana.lbl.gov/EisenSoftware.htm) was used for cluster analysis and tree visualization (12).

**Immunohistochemistry (IHC).** IHC was performed by using the VECTASTAIN ABC Kit as described by the manufacturer (Vector). We used the following primary antibodies: mouse anti-human Dicer (13D6, 1:20; Abcam) and rabbit anti-human Drosha (1:200; Abcam), mouse-anti-human Ki67 (1:200; DAKO). Antibodies were incubated for 2 h at room temperature or overnight at 4°C. The immunoreaction was visualized with 3,3′-diaminobenzidine (Vector). Images were collected through Cool SNAP Pro color digital camera (Media Cybernetics) and staining index was analyzed by using Image-Pro Plus 4.1 software (Media Cybernetics).

**Northern Blot Analysis.** The 5′-digoxigenin labeled locked nuclear acid probes for miRNA detection were ordered from Exiqon. Five to 10 $\mu$g of total RNA was loaded on a 15% TBE-Urea gel (Invitrogen) and electrophoresed at 180 V for 1.5 h. The gel was stained in 1x TBE plus 1 $\mu$g/ml ethidium bromide for 40 min and photographed before electroblotting onto positively charged nylon membrane (Roche) in a TransBlot SD SemiDry electrophoretic transfer cell (Bio-Rad) at 20 V for 1.5 h. The membrane was hybridized with 100 ng/ml probe at 37°C overnight. On the following day, the membrane was washed twice in 2× SSC at room temperature (21°C) for 5 min each and twice in 0.1× SSC at 37°C for 15 min each. Digoxigenin signals were detected with DIG Northern Starter kit (Roche) following the manufacturer's instructions.

**Microarray Data Retrieval.** The public expression microarray data were retrieved from authors' web site and further analyzed by using the web-based microarray analysis software, ONCOMINE (www.oncomine.org) (13).

**Bioinformatic Analysis.** 3′UTR sequences- The human 3′UTRs were extracted from Ensembl version 45 (dataset Homo sapiens genes National Center for Biotechnology Information 36). Affymetrix probe identifiers were mapped to Ensembl gene IDs, using BioMart (www.ensembl.org/biomart/martview). For each Ensembl gene, we used the longest 3′UTR transcript. 948 up-regulated and 15,212 unchanged 3′UTRs with the length of at least 30 and <10,000 nt were obtained. The one-tailed Wilcoxon rank sum test was performed by using R, for each hexamer occurrence distributions in up-regulated versus unchanged 3′UTRs.

**Gene Ontology (GO) Analysis.** FuncAssociate (with default settings) was used to search for GO attributes in which genes containing at least one of the hexamers corresponding to the miRNA were overrepresented (http://llama.med.harvard.edu/cgi/func/funcassociate) (14). For genes that are up-regulated in late EOC, we provide the number of hexamers that correspond to miRNAs from the Chr.14 cluster in Dataset S6, in which we also provide computational predicted nonconserved targets including 5′-dominant and 3′-compensatory target sites (15) and conserved targets as predicted from the programs TargetScan (16) and PicTar (17), if available.


1. Zhang L, *et al.* (2003) Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med* 348:203-213.
2. Maines-Bandiera SL, Kruk PA, and Auersperg N (1992) Simian virus 40-transformed human ovarian surface epithelial cells escape normal growth controls but retain morphogenetic responses to extracellular matrix. *Am J Obstet Gynecol* 167:729-735.
3. Chen C, *et al.* (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33:e179.
4. Liu CG, *et al.* (2004) An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc Natl Acad Sci USA* 101:9740-9744.
5. Greshock J, *et al.* (2004) 1-Mb resolution array-based comparative genomic hybridization using a BAC clone set optimized for cancer gene analysis. *Genome Res* 14:179-187.
6. Olshen AB, Venkatraman ES, Lucito R, Wigler M (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 5:557-572.
7. Margolin AA, *et al.* (2005) CGHAnalyzer: a stand-alone software package for cancer genome analysis using array-based DNA copy number data. *Bioinformatics* 21:3308-3311.
8. Kononen J, *et al.* (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 4:844-847.
9. Lassus H, Leminen A, Lundin J, Lehtovirta P, Butzow R (2003) Distinct subtypes of serous ovarian carcinoma identified by p53 determination. *Gynecol Oncol* 91:504-512.
10. Saeed AI, *et al.* (2003) TM4: A free, open-source system for microarray data management and analysis. *Biotechniques* 34:374-378.
11. Reich M, *et al.* (2006) *Nat Genet* 38:500-501.
12. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) *Proc Natl Acad Sci USA* Cluster analysis and display of genome-wide expression patterns. 95:14863-14868.
13. Rhodes DR, *et al.* (2007) Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* 9:166-180.
14. Berriz GF, King OD, Bryant B, Sander C, Roth FP (2003) Characterizing gene sets with FuncAssociate. *Bioinformatics* 19:2502-2504.
15. Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA-target recognition. *PLoS Biol* 3:e85.
16. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120:15-20.
17. Krek A, *et al.* (2005) Combinatorial microRNA target predictions. *Nat Genet* 37:495-500.

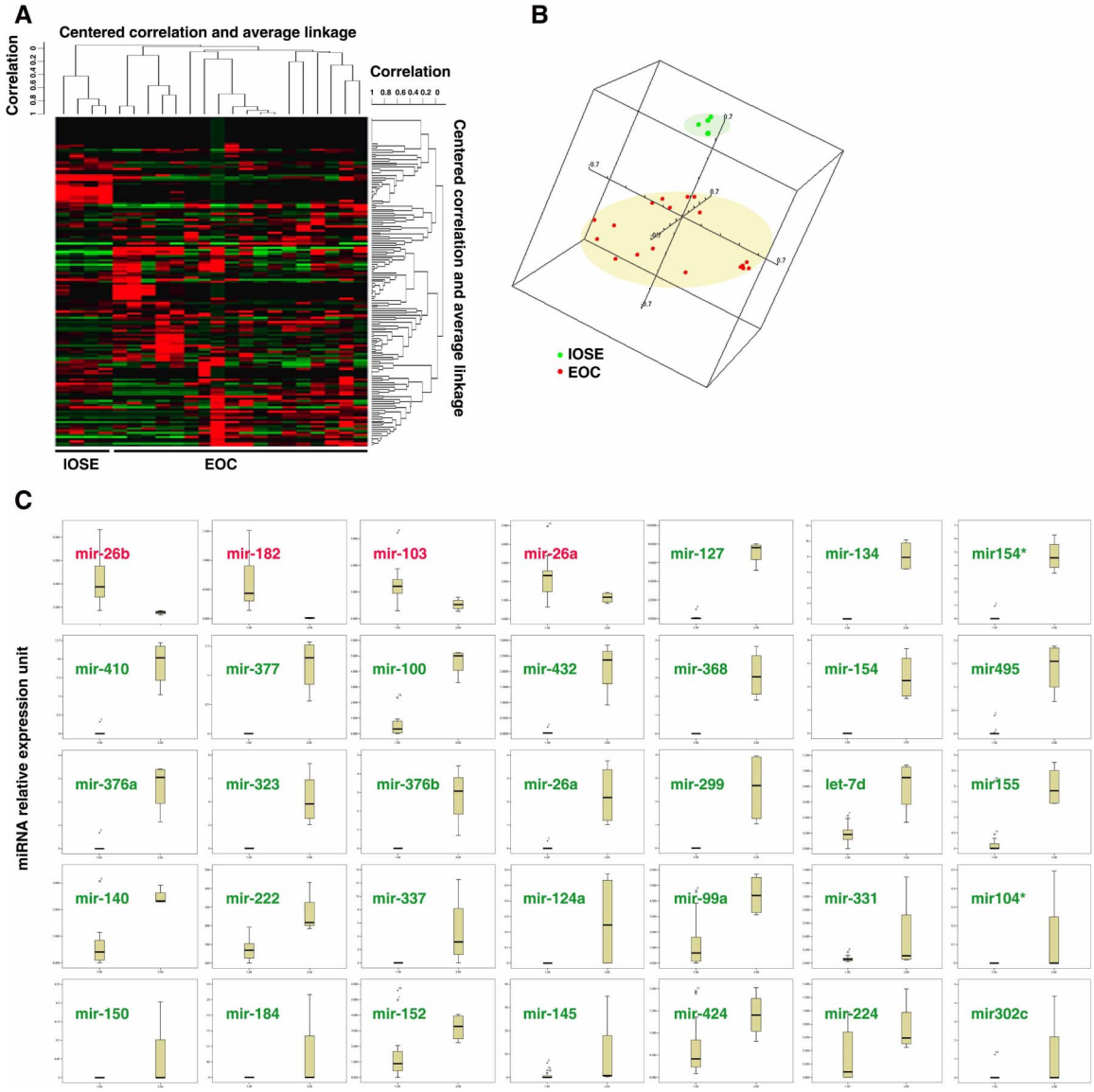



**Fig. S1.** miRNA expression profiles classify nonmalignant ovarian surface epithelium and epithelial ovarian cancer. (*A*) Unsupervised hierarchical cluster analysis of miRNA expression in 4 IOSEs and 18 EOC cell lines. (*B*) Three-dimensional scaling analysis of miRNA expression in 4 IOSEs and 18 EOC cell lines. (*C*) Relative expression of 35 miRNAs that are differentially expressed ($P < 0.05$) between IOSE and EOC cell lines. Bars represent the averaged (mean) expression of each miRNA in IOSE (green) and EOC group (red). Four miRNAs are significantly up-regulated, whereas 31 miRNAs are significantly down-regulated in EOC cell lines compared with IOSE cells.

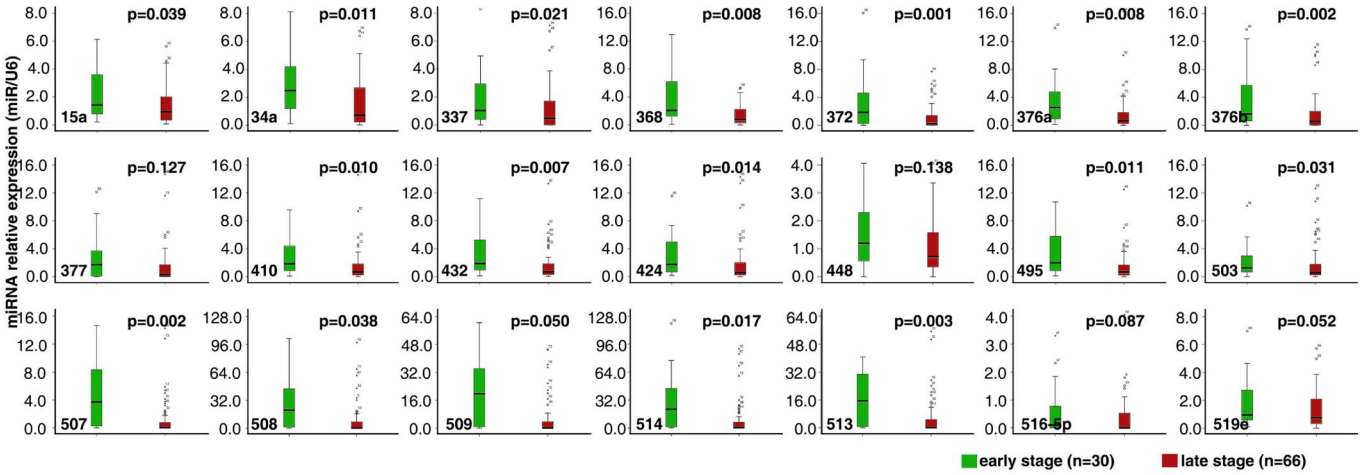



**Fig. S2.** Real-time RT-PCR validation of miRNA microarray results.

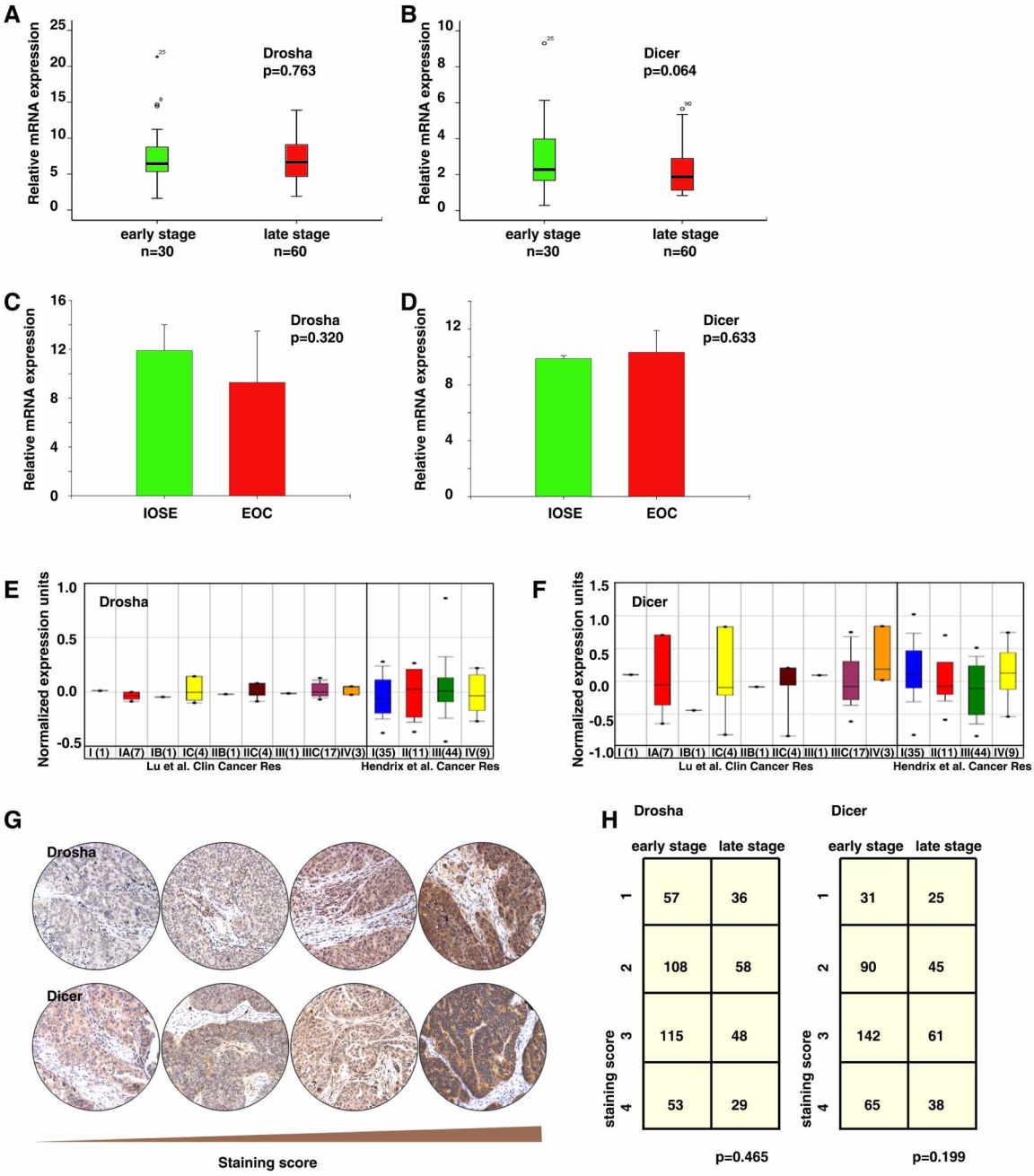



**Fig. S3.** Drosha and Dicer are not deregulated in ovarian cancer. (*A* and *B*). Relative mRNA expression of Drosha (*A*) and Dicer (*B*) in early- and late-stage EOC. (*C* and *D*). Relative mRNA expression of Drosha (*C*) and Dicer (*D*) in IOSE and EOC cell lines. (*E* and *F*). Drosha (*C*) and Dicer (*D*) mRNA expression at different stages of EOC in public microarray datasets. (*G*) Immunohistochemistry staining of Drosha and Dicer in tissue arrays containing 504 EOC specimens. (*H*) Summary of $\chi^2$ tests of Drosha and Dicer expression between early- and late-stage EOC.

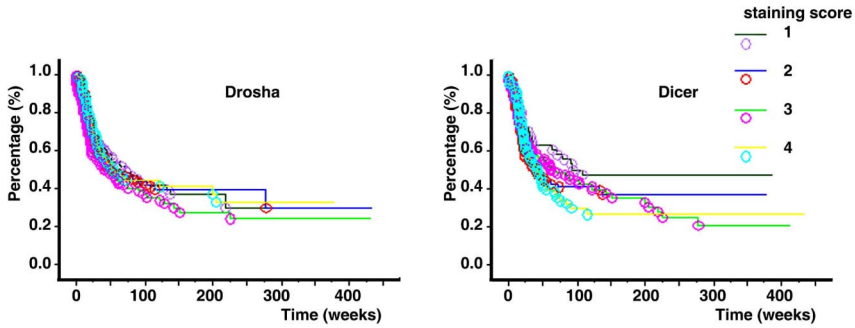



**Fig. S4.** Drosha and Dicer expression levels do not affect clinical outcome in EOC. Overall survival distribution of patients with advanced stage EOC and different expression of Drosha or Dicer protein based on tissue array analysis (all $P > 0.05$). *P* values were derived with the use of the log-rank statistic.

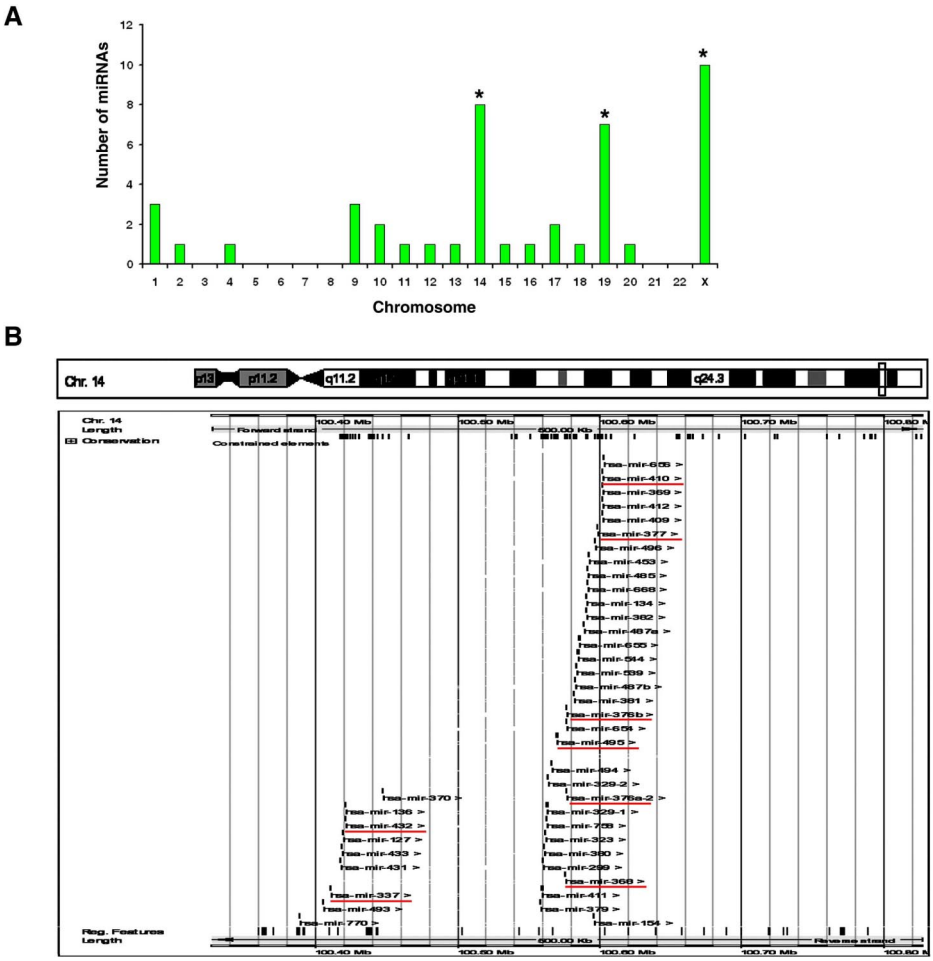



**Fig. S5.** Genomic distribution of down-regulated miRNAs in late-stage EOC. (*A*) Genomic distribution of miRNAs that were down-regulated in advanced stage EOC: 25 of 44 miRNAs (56.8%) are distributed at three chromosomes within large miRNA clusters (marked as stars, 8 located in chromosome 14; 7 located in chromosome 19; and 10 located in chromosome X). (*B*) Genomic organization of the miRNA cluster at the *Dlk1-Gtl2* domain in chromosome 14. Down-regulated miRNAs are underlined in red.

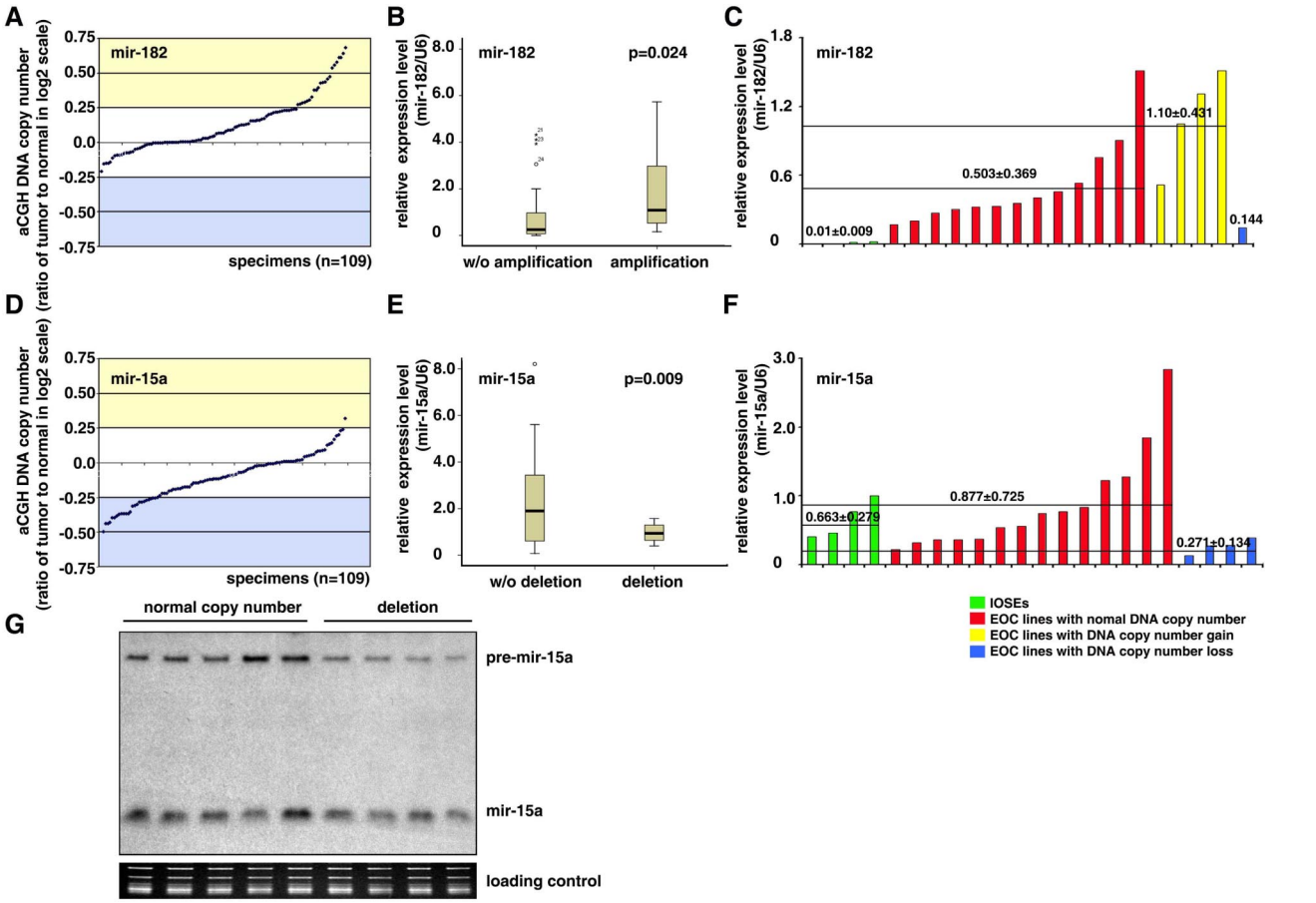



**Fig. S6.** DNA copy number deletions contribute to the down-regulation of miRNAs. (*A*) The genomic locus harboring *mir-182* exhibits significant DNA copy number gain in a set of 109 EOC. (*B*) Mature *miR-182* expression in EOC tumor specimens with or without DNA amplification involving the locus at which *mir-182* is located. (*C*) Mature *miR-182* expression in IOSE and EOC cell lines with or without the DNA amplification involving the locus at which *mir-182* is located. (*D*) The genomic locus harboring *mir-15a* exhibits significant DNA copy number loss in a set of 109 EOC. (*E*) Mature *miR-15a* expression in EOC tumor specimens with or without DNA copy number loss involving the locus at which *mir-15a* is located. (*F*) Mature *miR-15a* expression in IOSE, and EOC cell lines with or without the DNA deletion involving the locus at which *mir-15a* is located. (*G*) Northern blot analysis of pre-*miR-15a* and mature *miR-15a* in EOC cell lines with or without DNA deletion.

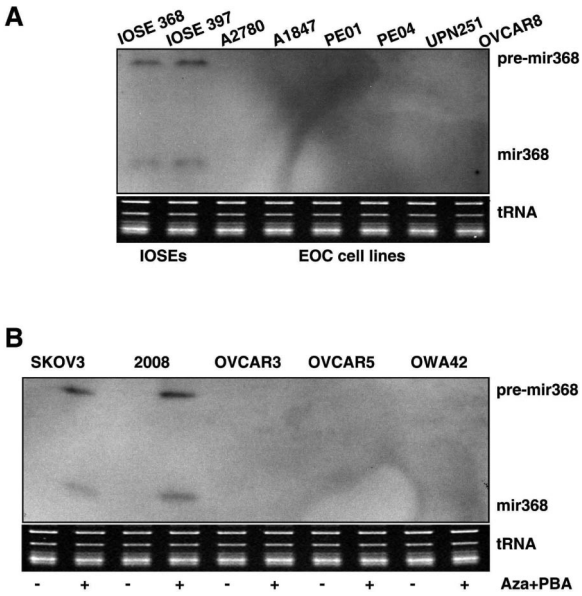
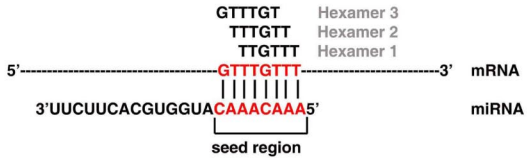



**Fig. S7.** Epigenetic alterations silence miRNA expression in ovarian cancer. (*A*) Northern blot analysis of *miR-386* expression in 2 IOSE cells and 6 EOC cell lines. (*B*) Northern blot analysis of *miR-368* expression in EOC cells treated with 5-Aza-CdR and PBA for 6 days.

A




B

| hexamer | -lnP | number of UTRs (UP) | hsa-miR | start position |
|---|---|---|---|---|
| TTTGTT | Inf | 566 | hsa-miR-495 | 2 |
| TTGTTT | 36.74 | 556 | hsa-miR-495 | 1 |
| AATTTA | 26.66 | 432 | hsa-miR-513-3p | 1 |
| GTTTGT | 24.31 | 400 | hsa-miR-495 | 3 |
| AAATTT | 23.86 | 500 | hsa-miR-513-3p | 2 |
| TGTGTT | 23.79 | 475 | hsa-miR-362-3p | 1 |
| GTTATA | 17.11 | 235 | hsa-miR-410 | 3 |
| TATATT | 16.15 | 431 | hsa-miR-410 | 1 |
| GTTGAA | 14.06 | 294 | hsa-miR-95 | 1 |
| TTATAT | 13.94 | 395 | hsa-miR-410 | 2 |
| GAAATT | 12.67 | 383 | hsa-miR-513-3p | 3 |
| TTTCAA | 9.72 | 417 | hsa-miR-488 | 1 |
| TATAGG | 8.78 | 179 | hsa-miR-337-3p | 3 |
| GCATTA | 8.51 | 215 | hsa-miR-365 | 1 |
| ACTTTG | 8.15 | 400 | hsa-miR-519d | 1 |
| AGTGAT | 7.89 | 291 | hsa-miR-34b | 3 |
| GTGATT | 7.53 | 298 | hsa-miR-34b | 2 |
| TATGAT | 7.52 | 248 | hsa-miR-376b | 1 |
| TATGAT | 7.52 | 248 | hsa-miR-376a | 1 |

C

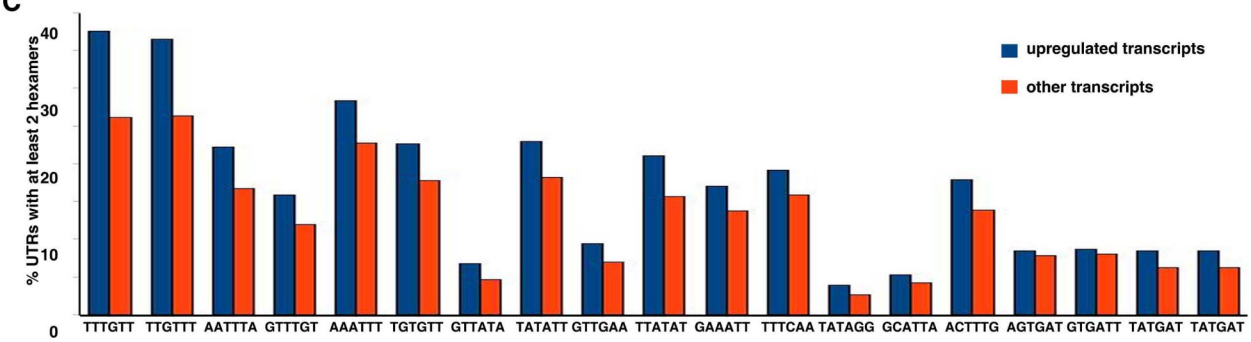



**Fig. S8.** miRNA deregulation affects mRNA transcripts. (*A*) Illustration of the seed region of miRNA and its targeting mRNA 3′UTR. (*B*) List of significant hexamers. (*C*) Percentage of 3′UTRs of up-regulated (blue bars) and unchanged genes (red bars) with at least two occurrences.

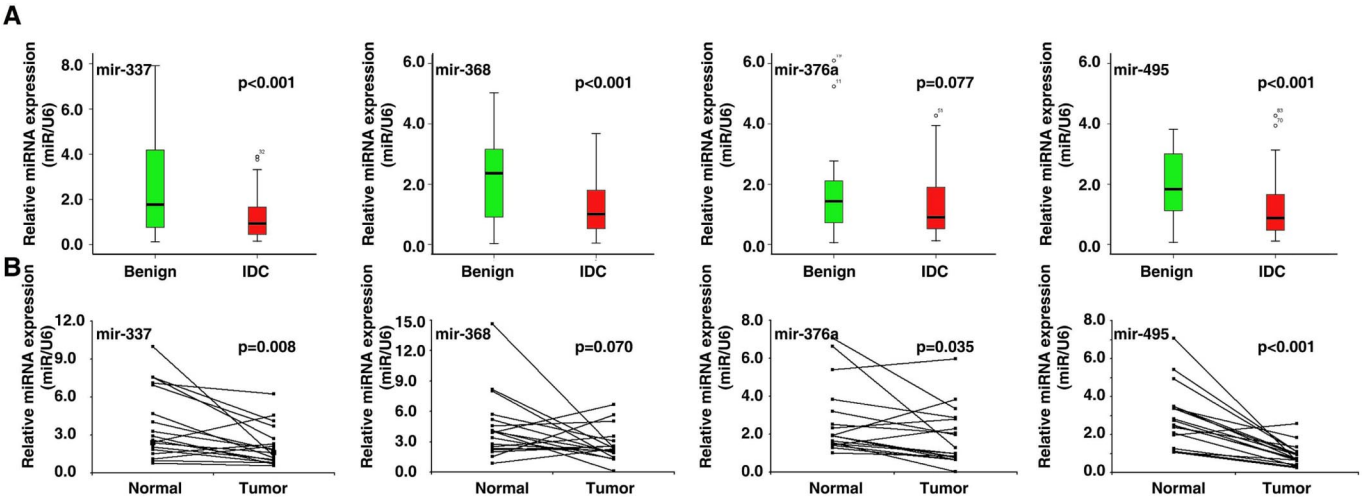



**Fig. S9.** miRNA cluster at the *Dlk1-Gtl2* domain is altered in epithelial cancers. (*A*) miRNAs from the cluster located at the *Dlk1-Gtl2* domain are down-regulated in invasive ductal carcinoma (IDC, $n = 73$) compared with benign breast lesions ($n = 23$). (*B*) miRNAs from this cluster are down-regulated in colon cancer compared with adjacent normal colon tissues ($n = 18$).

## Other Supporting Information Files

Dataset S1
Dataset S2
Dataset S3
Dataset S4
Dataset S5
Dataset S6
Dataset S7