Epigenome-wide ovarian cancer analysis identifies a methylation profile differentiating clear-cell histology with epigenetic silencing of the HERG K$^+$ channel

Mine S. Cicek[1], Devin C. Koestler[6], Brooke L. Fridley[7], Kimberly R. Kalli[2], Sebastian M. Armasu[1], Melissa C. Larson[1], Chen Wang[1], Stacey J. Winham[1], Robert A. Vierkant[1], David N. Rider[1], Matthew S. Block[3], Brandy Klotzle[9], Gottfried Konecny[10], Boris J. Winterhoff[4], Habib Hamidi[10], Viji Shridhar[5], Jian-Bing Fan[9], Daniel W. Visscher[5], Janet E. Olson[1], Lynn C. Hartmann[3], Marina Bibikova[9], Jeremy Chien[8], Julie M. Cunningham[5] and Ellen L. Goode[1],*

[1]Department of Health Sciences Research, [2]Department of Medicine, [3]Department of Oncology, [4]Department of Obstetrics and Gynecology and [5]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA, [6]Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth College, Lebanon, NH 03756, USA, [7]Department of Biostatistics and [8]Department of Cancer Biology, University of Kansas Medical Center, Kansas City, KS 66160, USA, [9]Illumina, Inc., 5200 Illumina Way, San Diego, CA 92122, USA and [10]Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90404, USA

Ovarian cancer remains the leading cause of death in women with gynecologic malignancies, despite surgical advances and the development of more effective chemotherapeutics. As increasing evidence indicates that clear-cell ovarian cancer may have unique pathogenesis, further understanding of molecular features may enable us to begin to understand the underlying biology and histology-specific information for improved outcomes. To study epigenetics in clear-cell ovarian cancer, fresh frozen tumor DNA ($n = 485$) was assayed on Illumina Infinium HumanMethylation450 BeadChips. We identified a clear-cell ovarian cancer tumor methylation profile ($n = 163$) which we validated in two independent replication sets (set 1, $n = 163$; set 2, $n = 159$), highlighting 22 CpG loci associated with nine genes (*VWA1*, *FOXP1*, *FGFRL1*, *LINC00340*, *KCNH2*, *ANK1*, *ATXN2*, *NDRG21* and *SLC16A11*). Nearly all of the differentially methylated CpGs showed a propensity toward hypermethylation among clear-cell cases. Several loci methylation inversely correlated with tumor gene expression, most notably *KCNH2* (HERG, a potassium channel) ($P = 9.5 \times 10^{-7}$), indicating epigenetic silencing. In addition, a predicted methylation class mainly represented by the clear-cell cases (20 clear cell out of 23 cases) had improved survival time. Although these analyses included only 30 clear-cell carcinomas, results suggest that loss of expression of KCNH2 (HERG) by methylation could be a good prognostic marker, given that overexpression of the potassium (K$^+$) channel Eag family members promotes increased proliferation and results in poor prognosis. Validation in a bigger cohort of clear-cell tumors of the ovary is warranted.

INTRODUCTION

Clear-cell ovarian cancer cases are known to be associated with endometriosis, diagnosed at younger age and lower stage, have lower frequency of *BRCA1* and *BRCA2* mutations and higher frequency of *ARID1A* and *PIK3CA* mutations and have greater resistance for first-line platinum and taxane-based chemotherapy and poorer survival in advanced stages than

---

*To whom correspondence should be addressed. Tel: +1 5072667997; Fax: +1 5072662478; Email egoode@mayo.edu

other subtypes (1,2). Clear-cell subtype is one of the most common subtypes of ovarian cancer: high-grade serous (70%), low-grade serous (5%), endometrioid (10%), clear cell (10%) and mucinous (3%) (3). The current standard of care for the treatment is surgical resection (debulking) followed by adjuvant chemotherapy (4,5). Despite advances in surgical technologies and the development of more effective chemotherapeutics, epithelial ovarian cancer (EOC) remains the leading cause of death in women with gynecologic malignancies (6).

At present, ovarian cancer is clinically considered as a single disease, and treatment is dependent upon tumor stage and grade rather than histology. However, the increasing evidence indicates that clear-cell ovarian cancer has a different etiology, clinical characteristics and outcomes that are distinct from the other histological subtypes (1,7–9). Given our current knowledge, ovarian clear-cell carcinoma is distinct from other ovarian cancer histologies and warrants identification of molecular features to understand its biology, as further understanding of this disease may enable subtype-specific treatment options to improve patient outcomes. Thus far, efforts in distinguishing clear-cell from other ovarian cancer have been based on gene expression analysis showing gene expression patterns reflecting both morphological features and biological behavior (10).

DNA methylation, predominantly of cytosine residues adjacent to guanine bases (i.e. CpG dinucleotides), is a key epigenetic process involved in the regulation of gene expression across a variety of cellular processes, including genomic imprinting, chromosome X inactivation and chromosomal integrity (11,12). Dysregulation of DNA methylation is associated with multiple human diseases including cancer (13), and there is a growing interest in the role of DNA methylation in carcinogenesis and tumorigenesis and the use of such information for differentiating cancer subtypes at a molecular level. In an effort to identify molecular alterations, several epigenomic studies examining tumor DNA methylation have been conducted in ovarian cancer; however, these studies were not genome-wide, or they focused on only the high-grade serous subtype (14–18). Thus, more comprehensive studies involving epigenome-wide assessment of DNA methylation are needed to fully understand ovarian clear-cell cancer. Here, in order to begin to understand the underlying biology unique to clear-cell ovarian cancer, we examined the hypothesis that the ovarian cancer clear-cell subtype has a distinct tumor methylation profile from other subtypes that also is associated with epigenetic silencing. Assessing over 485 000 CpG sites in a total of 485 ovarian cancer cases including 30 clear-cell carcinomas, the present study represents the most comprehensive study to date of tumor DNA methylation in the context of ovarian cancer.

RESULTS

Methylation clustering

An ovarian cancer tumor DNA methylation profile associated with clear-cell histology was identified and validated using three roughly equally sized collections, including one training data set ($n = 163$) and two independent replication data sets

Table 1. Characteristics of study participants, n (%)

	Testing set ($n = 163$)	Replication set 1 ($n = 163$)	Replication set 2 ($n = 159$)
Site			
Ovarian	155 (95)	147 (90)	157 (99)
Fallopian tube	6 (4)	13 (8)	2 (1)
Primary peritoneal	2 (1)	3 (2)	0
Histology			
High-grade serous	115 (71)	115 (71)	114 (72)
Low-grade serous	7 (4)	7 (4)	2 (1)
Endometrioid	27 (17)	27 (17)	26 (16)
Clear cell	9 (6)	8 (5)	13 (8)
Mucinous	4 (2)	4 (2)	4 (2)
Serous, grade unknown	2 (1)	1 (1)	0 (0)
Age at diagnosis, years			
Median (range)	61 (28–85)	62 (36–86)	62 (21–88)
Stage			
Stage 1	20 (12)	22 (14)	18 (11)
Stage 2	9 (6)	9 (6)	12 (8)
Stage 3	102 (63)	105 (64)	99 (62)
Stage 4	32 (20)	27 (17)	30 (19)
Grade			
Low grade	24 (15)	28 (17)	16 (10)
High grade	138 (85)	133 (82)	143 (90)
Unknown	1	2	0
Ascites			
Yes	93 (70)	93 (66)	66 (65)
No	40 (30)	49 (34)	35 (35)
Unknown	30	21	58
Surgical debulking status			
Optimal (<1 cm)	133 (83)	141 (88)	125 (80)
Sub-optimal (≥1 cm)	27 (17)	20 (12)	31 (20)
Unknown	3	2	3
Peritoneal cytology			
Positive	126 (84)	113 (77)	108 (71)
Negative	24 (16)	33 (23)	45 (29)
Unknown	13	17	6

Numbers are n (%) unless otherwise specified.

(set 1, $n = 163$; set 2, $n = 159$). Characteristics of the Mayo Clinic cases studied ($n = 485$) are shown in Table 1 and reflect the expected distributions of clinical features and similarity across analytical sets. A Gaussian distributed recursively partitioned mixture model (RPMM) was fit to the training data set to cluster ovarian cancer cases using the logit-transformed methylation beta values for the CpG loci most strongly associated with clear-cell histology. The optimal number of markers, M, was found to be 22 based on a nested cross-validation procedure (Supplementary Material, Figure S1); thus, the top 22 CpG loci were selected in the training set as the basis of the clear-cell-specific methylation profile (Table 2).

The clustering of the training set of samples using these 22 CpG loci resulted in 8 methylation classes, delineated as right branch R of the dendrogram (RLL, RLR, RRL, RRR) and left branch L of the dendrogram (LLL, LLR, LRL, LRR). Several of the 22 most discriminatory CpGs reside in close proximity (e.g. those on chromosomes 4, 7, 8, 14, 17 and X). Therefore,

Table 2. Median methylation of CpG loci represented in ovarian cancer clear-cell and other histologies

CpG	Chr	Position	Relationship to CpG island	Associated gene (CpG to TSS distance)	Replication set 1 ($n = 163$)			Replication set 2 ($n = 159$)		
					Clear cell	Other	P-value	Clear cell	Other	P-value
cg06444575	1	1 369 772	N. shore	*VWA1* (−1131)	0.5129	0.1115	4.0×10^{-6}	0.4170	0.1266	5.6×10^{-4}
cg15958828	3	71 598 380	Ocean	*FOXP1* (34 760)	0.4448	0.8758	2.3×10^{-6}	0.5994	0.9100	2.4×10^{-6}
cg17105609	4	1 006 237	Island	*FGFRL1* (627)	0.2680	0.0315	8.8×10^{-6}	0.0245	0.0182	0.01
cg08774969	4	1 006 240	Island	*FGFRL1* (630)	0.2417	0.0235	7.7×10^{-5}	0.0281	0.0157	2.7×10^{-3}
cg05866411	4	1 006 278	Island	*FGFRL1* (668)	0.1913	0.0454	1.3×10^{-4}	0.0405	0.0293	9.5×10^{-3}
cg00641409	6	21 666 778	Island	*LINC00340* (103)	0.1834	0.0720	4.0×10^{-5}	0.2104	0.1052	9.3×10^{-4}
cg18423852	7	150 653 001	Island	*KCNH2* (22 401)	0.7852	0.0721	2.7×10^{-6}	0.5094	0.0459	7.0×10^{-5}
cg15472092	7	150 653 079	Island	*KCNH2* (22 323)	0.7020	0.1077	2.5×10^{-6}	0.5759	0.0986	3.8×10^{-5}
cg18296036	7	150 653 115	N. shore	*KCNH2* (22 287)	0.6379	0.1098	2.7×10^{-6}	0.5045	0.1273	3.0×10^{-4}
cg13152952	8	41 570 554	Ocean	*ANK1* (84 586)	0.7901	0.0619	2.5×10^{-6}	0.4769	0.0353	4.1×10^{-5}
cg10601625	8	41 570 622	Ocean	*ANK1* (84 518)	0.6353	0.1234	2.0×10^{-6}	0.4996	0.1370	4.5×10^{-6}
cg07058694	12	111 909 145	Ocean	*ATXN2* (128 335)	0.4374	0.8757	2.4×10^{-6}	0.6712	0.8922	4.2×10^{-5}
cg01574694	14	21 493 964	Island	*NDRG2* (−29)	0.3912	0.0769	4.3×10^{-6}	0.1334	0.0206	2.5×10^{-3}
cg10963192	14	100 751 514	N. shore	None	0.2577	0.1813	2.8×10^{-5}	0.1848	0.1607	4.3×10^{-3}
cg15869718	14	100 751 572	N. shore	None	0.3435	0.0660	9.7×10^{-4}	0.1002	0.0575	1.6×10^{-4}
cg18139254	14	100 751 903	Island	None	0.0729	0.0396	6.6×10^{-4}	0.0266	0.0247	0.22
cg05940703	17	6 947 032	Island	*SLC16A11* (210)	0.1755	0.1243	0.003	0.2449	0.1338	1.7×10^{-4}
cg15639045	17	6 947 200	Island	*SLC16A11* (42)	0.1681	0.0970	1.8×10^{-3}	0.2098	0.1138	1.1×10^{-3}
cg03892045	17	6 947 288	Island	*SLC16A11* (−46)	0.5585	0.0713	9.2×10^{-6}	0.2467	0.0476	6.8×10^{-5}
cg03712541	17	6 947 294	Island	*SLC16A11* (−52)	0.4910	0.0479	9.4×10^{-5}	0.1948	0.0282	4.8×10^{-5}
cg13976265	X	3 732 500	N. shore	None	0.1860	0.0695	0.005	0.0995	0.0752	6.8×10^{-4}
cg04174838	X	3 733 239	Island	None	0.2515	0.0849	0.03	0.1301	0.1002	8.9×10^{-4}

P-value represents the Wilcoxon signed-rank test; Chr, chromosome; build 37; other histologies include serous, endometrioid and mucinous; None, not in a gene; N. Shore, 0–2 kb upstream of a CpG island; Ocean, >4 kb from a CpG island; CpG to TSS distance, distance from CpG to the transcriptional start site of the gene; Gene Refseq IDs: *VWA1*_ NM_022834, FOXP1_ NM_032682, *FGFRL1*_ NM_001004356, *KCNH2*_ NM_000238, *ANK1*_ NM_000037, *ATXN2*_ NM_002973, *NDRG2*_NM_201535, *SLC16A11*_ NM_153357; horizontal lines indicate positionally grouped CpG 'blocks'.

methylation beta values are likely to be highly correlated (19,20), representing essentially 11 'blocks' of differential methylation. All but three of these 'blocks' were located in CpG islands or shores (flanking regions 0–2 kb from CpG islands), and all but two of these 'blocks' were in a known gene. The nine genes associated with the CpGs highlighted in this analysis included *VWA1* (von Willebrand factor A domain containing 1), *FOXP1* (forkhead box P1), *FGFRL1* (fibroblast growth factor receptor-like 1), *LINC00340* (long intergenic non-protein coding RNA 340), *KCNH2* [potassium voltage-gated channel, subfamily H (eag-related), member 2], *ANK1* (ankyrin 1, erythrocytic), *ATXN2* (ataxin 2), *NDRG21* (NDRG family member 2) and *SLC16A11* [solute carrier family 16, member 11 (monocarboxylic acid transporter 11); (Table 2)].

The semi-supervised RPMM (SS-RPMM) solution fit to the training data set based on the profile determined using the 22 CpG loci was then used to predict methylation class membership in two independent replication data sets. In the first replication set, methylation right branch (R) classes (those beginning with R) included 12 cases and were significantly enriched for clear-cell cases compared with L classes (those beginning with L), which were devoid of clear-cell cases (permutation chi-square $P < 1 \times 10^{-16}$). Further, the RRL class contained all of the clear-cell cases ($n = 8$) as well as one serous and one mucinous case (Fig. 1A). At most loci, clear-cell tumors were hypermethylated compared with other histologies; however, at the loci on chromosomes 3 and 12 (cg15958828 in *FOXP1*, and cg07058694 in *ATXN2*, respectively), hypomethylation in clear-cell tumors was observed (Table 2).

Similar results were seen in the analysis of the second replication data set, which consisted of 159 cases, including 13 of clear-cell histology. The clear-cell methylation profile again provided good prediction for clear-cell tumors; the R class included a greater proportion of clear-cell cases than the L class ($P < 7.8 \times 10^{-6}$), and the RRL class again contained the majority of the clear-cell cases (Fig. 1B). Unlike in the first replication set, some clear-cell cases were assigned to other classes, with four in the LLR cases and one in the LRR classes. The patterns of methylation by subtype were similar to those in the first replication data set (Table 2).

As expected, the correlation between methylation beta values among the 22 CpGs was high for CpGs in close proximity (Pearson's correlation coefficient >0.62), namely 'blocks' of CpGs, as mentioned earlier, on chromosomes 4, 7, 8, 14, 17 and X. As shown in Supplementary Material, Figure S2, correlation patterns were similar in both replication sets. The two loci showing clear-cell hypomethylation (cg15958828 in *FOXP1* on chromosome 3 and cg07058694 in *ATXN2* on chromosome 12) were correlated with each other (Pearson's correlation coefficient = 0.84) and negatively correlated with the other loci.

Correlation of methylation with gene expression

Promoter-associated CpGs are generally thought to be unmethylated (21); however, hypermethylation of CpG island loci may be associated with the modification of gene activation, especially gene silencing (22,23). Therefore, cis association between expression and methylation was examined ($n = 171$)
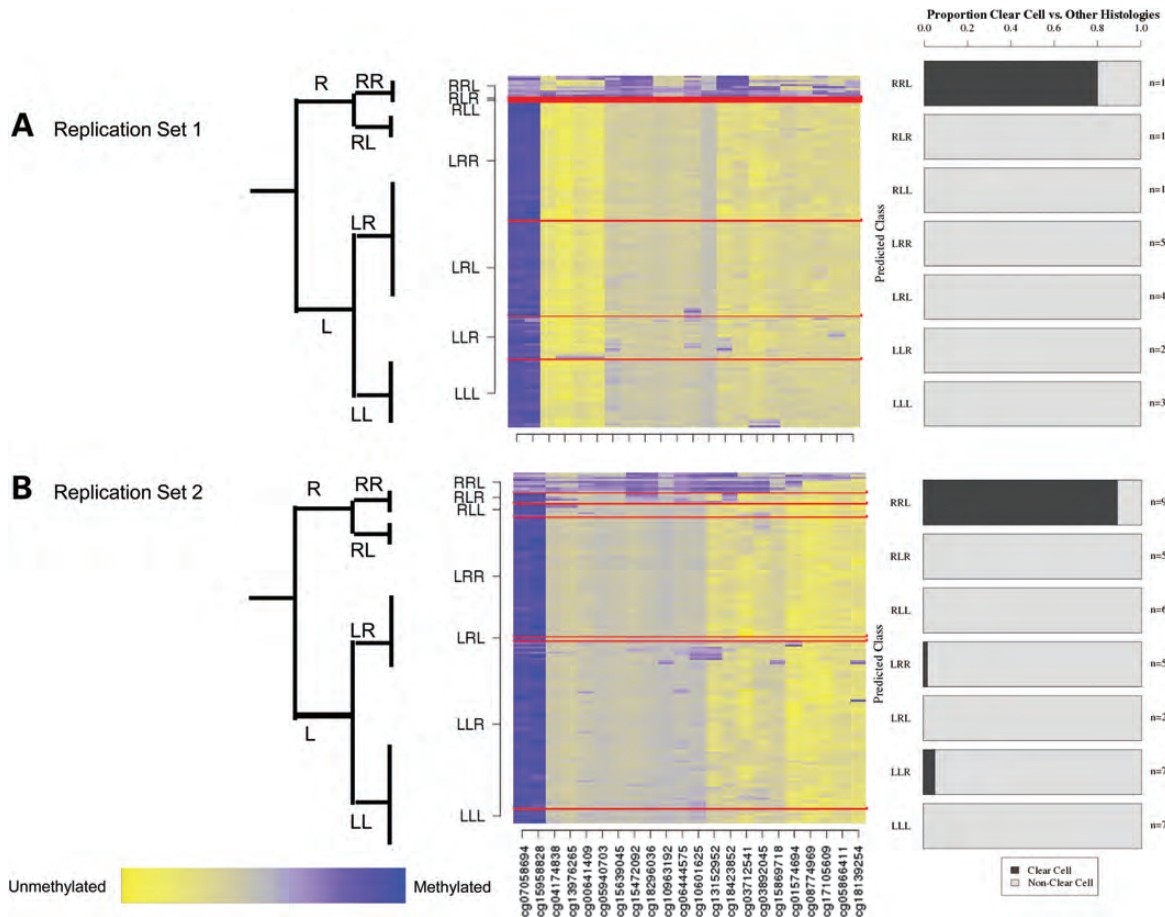
Figure 1. Clustering heat map of methylation in ovarian cancer. Results from the methylation profile analysis of the replication sets. The dendrogram on the left indicates methylation class relationships. Figures to the left are heat maps with methylation values expressed as methylated (blue) and unmethylated (yellow); rows represent subjects, which are grouped by predicted class membership (red lines), and columns represent the 22 CpG loci in hierarchical clustering order that were used to generate the methylation classes. Figures to the right are bar plots depicting the percent clear-cell/non-clear-cell cases across the predicted methylation classes, with numbers on the right representing the numbers of cases in each predicted methylation class. (**A**) Replication set 1 ($n = 163$) and (**B**) Replication set 2 ($n = 159$).

for the 22 CpG loci (Supplementary Material, Table S1), of which 59% ($n = 13$) were located in CpG island regions. A total of 18 transcripts from nine genes were investigated (*VWA1*, *FOXP1*, *FGFRL1*, *LINC00340*, *KCNH2*, *ANK1*, *ATXN2*, *NDRG2* and *SLC16A11* containing CpGs). Since many of the 22 CpGs were located in CpG islands, we observed an expected inverse relationship between methylation and expression given the well-established role of CpG island hypermethylation on repression of gene expression. The strongest relationship between hypermethylation and reduced mRNA expression was at *KCNH2* on chromosome 7 with cg15472092 in a CpG island (Spearman's correlation $= -0.37$, $P = 9.5 \times 10^{-7}$); consistent results were seen at related *KCNH2* CpGs and probes. The gene expression was reduced in clear-cell cases compared with non-clear-cell cases measured with two KCNH2 expression probes, A_23_P377882 ($P = 6.5 \times 10^{-4}$) and A_23_P168403 ($P = 0.04$). Although, methylation on chromosome 8 (cg13152952), located >4 kb from the nearest CpG island, was associated with increased expression of *ANK1* (Spearman's correlation $= 0.27$, $P = 3.4 \times 10^{-4}$),

which might be expected, as the CpG site was not in a CpG island. Modest associations were seen at *VWA1*, *NDRG2*, *SLC25A29*; at other loci, strong cis associations were generally not observed (Table 3).

Pathway analysis

We further evaluated robustness of our model by increasing the number of selected top CpG probes from 22 to 100 in the training set. Clustering results always showed significant consistency with our optimal model using 22 CpGs even when we increased the number of top CpGs to 100 or even more (data not shown). This overall consistency enabled us to include more probes (genes) in an exploratory pathway analysis with the goal of examining the relationships between clear-cell histology and molecular functions of genes associated with differentially methylated CpGs. We chose to use the top 100 CpGs for the pathway analysis. For the 54 genes which included these 100 top CpGs, a function group entitled 'Differentiation of Cells' appeared to be over-represented

Table 3. Correlation between clear-cell CpG loci and cis gene expression ($n = 171$)

CpG	Chr	Position	Gene	Expression probe	Expression probe genomic coordinates	Spearman's correlation	P-value
cg06444575	1	1 369 772	*VWA1*	A_23_P86089	1 376 065–1 376 124	−0.23	0.002
				A_24_P122746	1 377 713–1 377 772	−0.24	0.002
cg15958828	3	71 598 380	*FOXP1*	A_24_P362737	71 006 080–71 006 021	−0.05	0.51
				A_23_P155257	71 008 467–71 008 408	−0.10	0.20
cg17105609	4	1 006 237	*FGFRL1*	A_24_P46212	1 018 729–1 018 788	0.04	0.63
				A_23_P92349	1 019 939–1 019 998	0.13	0.08
cg08774969	4	1 006 240	*FGFRL1*	A_24_P46212	1 018 729–1 018 788	−0.15	0.06
				A_23_P92349	1 019 939–1 019 998	−0.04	0.65
cg05866411	4	1 006 278	*FGFRL1*	A_24_P46212	1 018 729–1 018 788	−0.06	0.45
				A_23_P92349	1 019 939–1 019 998	−0.05	0.51
cg00641409	6	21 666 778	*LINC00340*	A_24_P838448	22 194 466–22 194 525	0.08	0.27
				A_24_P861099	22 195 280–22 195 339	0.11	0.16
cg18423852	7	150 653 001	*KCNH2*	A_23_P168403	150 642 151–150 642 092	−0.16	0.04
				A_23_P377882	150 647 043–150 646 984	0.03	0.68
cg15472092	7	150 653 079	*KCNH2*	A_23_P168403	150 642 151–150 642 092	−0.37	9.5×10^{-7}
				A_23_P377882	150 647 043–150 646 984	−0.17	0.03
cg18296036	7	150 653 115	*KCNH2*	A_23_P168403	150 642 151–150 642 092	−0.34	5.9×10^{-6}
				A_23_P377882	150 647 043–150 646 984	−0.09	0.26
cg13152952	8	41 570 554	*ANK1*	A_24_P162483	41 512 136–41 512 077	0.18	0.02
				A_24_P162485	41 512 299–41 512 240	0.27	3.4×10^{-4}
				A_23_P216108	41 513 128–41 513 069	0.27	3.2×10^{-4}
cg10601625	8	41 570 622	*ANK1*	A_24_P162483	41 512 136–41 512 077	0.08	0.30
				A_24_P162485	41 512 299–41 512 240	0.18	0.02
				A_23_P216108	41 513 128–41 513 069	0.18	0.02
cg07058694	12	111 909 145	*ATXN2*	A_23_P389907	111 890 389–111 890 330	0.00	0.99
				A_24_P930507	111 891 962–111 891 903	0.09	0.26
cg01574694	14	21 493 964	*NDRG2*	A_23_P37205	21 485 133–21 485 074	−0.21	0.005
				A_24_P244800	21 488 089–21 487 853	−0.16	0.04
cg05940703	17	6 947 032	*SLC16A11*	A_23_P363313	6 945 224–6 945 165	0.02	0.08
cg15639045	17	6 947 200	*SLC16A11*	A_23_P363313	6 945 224–6 945 165	−0.16	0.06
cg03892045	17	6 947 288	*SLC16A11*	A_23_P363313	6 945 224–6 945 165	−0.25	9.3×10^{-4}
cg03712541	17	6 947 294	*SLC16A11*	A_23_P363313	6 945 224–6 945 165	0.01	0.45

Chr, chromosome; build 37. Relationship to a canonical CpG island: shores—0–2 kb from a CpG island; shelves—2–4 kb from a CpG island; methylation probes are annotated within a gene (CpGs that are not located in a gene were excluded from this analysis; cg10963192, cg15869718, cg18139254 on chromosome 14, and cg13976265, cg04174838 on chromosome X); horizontal lines indicate positionally grouped CpG 'blocks'.

($P = 8.6 \times 10^{-3}$), including genes *ADK*, *ATXN2*, *COL11A2*, *FOXP1*, *NCOR2*, *NES*, *NFIB*, *PCSK2*, *POU2F2*, *TGFB1*, *TGFB1I1* and *WNT4*. Network analysis identified three biological networks (i.e. a group of biologically related genes), which included 10 or more of the CpG-containing genes. These biological networks were 'Developmental Disorder, Cell Death, Renal Necrosis/Cell Death', 'Cellular Assembly and Organization, Neurological Disease, Organismal Injury and Abnormalities' and 'Cell Cycle, Cellular Development, Cell Death'. The last network included the largest number of these genes ($n = 19$; *ADK*, *ANK1*, *COL11A2*, *CYTH3*, *EXOC2*, *FGFRL1*, *FOXP1*, *KCNH2*, *NCOR2*, *NDRG2*, *NES*, *NFIB*, *PCSK2*, *POU2F2*, *PROC*, *TGFB1*, *TGFB1I1*, *TRIB1* and *WNT4*) and is displayed in Supplementary Material, Figure S3.

Association with overall survival

Survival time did not differ in clear-cell cases ($n = 30$, including 20 RRL cases) compared with other histologies ($P = 0.55$)

(Supplementary Material, Figure S4A); however, cases in the predicted RRL methylation class ($n = 23$, including 20 clear-cell cases) showed a trend in improved survival time compared with other methylation classes ($P = 0.08$; Supplementary Material, Figure S4B). Thus, reclassification of only 13 cases based on predicted methylation class accounts for this difference. The survival analysis should be validated with a replication study on independent set of cases. This exploratory analysis suggests that the clear-cell-associated methylation profile, in particular the RRL class, may, in fact, differentiate a novel clinically relevant subgroup and more precisely predict long-term survival prognosis.

DISCUSSION

Ovarian cancer is the leading cause of death in women with gynecologic malignancies; its current treatment scheme is solely dependent on stage and grade of the disease. The different histological subtypes which exhibit distinct clinical and genetic features warrant further research in an effort to

understand the biology of ovarian cancer (3,24). In this epigenetic study, we investigated genome-wide DNA methylation and histologic subtypes using fresh frozen DNA from 485 ovarian cancer tumors, using Illumina HumanMethylation450 BeadChips (over 440 000 CpG sites following exclusions). We identified eight ovarian cancer methylation classes based on the semi-supervised RPMM analyses, one of which, termed RRL, was enriched for clear-cell cases. This is the most comprehensive ovarian cancer tumor epigenetic study to date. There are a number of studies in the literature reporting methylation in ovarian cancer and ovarian cancer cell lines (15,16,25–27), and methylation differences in ovarian cancer versus normal tissue (28–30), but sample sizes were often small and comparison across subtypes was rarely done (15), especially with a high-coverage methylation array. The TCGA reported 168 genes as epigenetically silenced (promoter hypermethylation associated with decreased gene expression) in high-grade serous ovarian cancer samples compared with fallopian tube controls (17); as expected, none of these were among our best 100 clear-cell differentiating CpGs. That the TCGA only examined high-grade serous ovarian tumors and examined lower density CpGs (average two CpGs per promoter) on the Illumina Infinium Human-Methylation27 BeadChip demonstrated the need for more comprehensive assessment of ovarian cancer methylation.

Human promoters containing a CpG island tend to remain nucleosome-free as well as methylation-free. However, methylation of these regions is now firmly established as an important mechanism for gene inactivation. CpG island hypermethylation of tumor-suppressor genes has been described in almost every tumor type. In contrast, exons demonstrate a high degree of methylation and nucleosome occupancy (31). Importantly, exonic CpG islands are more susceptible to *de novo* methylation than promoter islands and methylation seeded in exonic regions may spread to other islands, including promoter regions (31). Furthermore, recent work has demonstrated that exonic DNA methylation seems to function together with exonic nucleosomes and H3K36me3 for the proper splicing of transcripts with different expression levels (31).

Because there is evidence in the literature showing importance of SWI/SNF (SWItch/Sucrose NonFermentable) genes in clear-cell tumors (1), we examined the difference in the methylation in 10 of these genes between clear-cell and non-clear-cell cases. In total, there were 264 CpGs that passed our QC, which were associated with *MTOR*, *ARID1A*, *PIK3CA*, *PTEN*, *CHD4*, *TP53*, *HNF1B*, *PPM1D*, *PPP2R1A* and *EP300*. We first implemented a series of locus-by-locus analyses for examining the association between the methylation of these 264 CpGs and clear-cell histology. As before, these models were controlled for age, stage, grade, ascites and peritoneal cytology. Our results showed that for both the training set and replication set 1, there did not appear to be any strong relationship between the methylation of these 264 CpGs and clear-cell histology, as evidenced by their uniform distribution of *P*-values. For the combined analysis, there was a slight trend toward non-uniformly distributed *P*-values, with heavier weight toward lower *P*-values; however, the strength of association for these 264 CpGs was considerably less than the top 100 CpGs identified in our previous analysis. Along these lines, among these 264 CpGs,

none appeared in our list of the top 1000 CpGs that associate with clear-cell histology. Thus, our agnostic examination identified probes that are much more strongly associated with clear-cell histology than any of the probes associated with the SWI/SNF genes.

In addition to molecular profiling of clear-cell ovarian cancer cases, we classified the 54 genes which included the top 100 CpGs differentially methylated in clear cell using Ingenuity Pathway Analysis (IPA). These genes were correlated in the human interactome by evidencing the functional networks. Clear-cell-specific differentially methylated genes were found to be involved in various molecular and cellular functions and most specifically, 19 genes including *KCNH2* in 'Cell Cycle, Cellular Development, Cell Death'. *KCNH2* methylation and expression are the most consistently associated, suggesting that investigation of its role may shed light on the biology of clear-cell ovarian cancer. The *KCNH2* gene region was also evaluated as an example to show that even when considering a broader region (± 2 Mb), the same methylation signal associated with clear-cell histology was detected. This gene encodes an ether-a-go-go potassium channel in cardiac cells (32). KCNH2, also known as HERG, belongs to the Eag family, which have been implicated in signaling pathways leading to cell proliferation or apoptosis (programmed cell death) (33). Studies on ovarian and endometrial cancer cell lines have shown preferential expression of HERG in tumor cell lines, and high expression was correlated with cell proliferation (34,35). Furthermore, K^+ channel blockers, 4-AP and TEA, have been reported to significantly inhibit the proliferation of HERG-positive SKOV-3 cells (36). EAG (Kv10.1) overexpression in ovarian cancer was associated with poor prognosis, identifying it as a therapeutic target. KCNH2 (Kv11.1, HERG) is also reported to be overexpressed in ovarian cancer. Of the 25 clear-cell cancers identified in that report, 17 had lower levels of Eag expression. Their study did not specifically highlight the expression levels of HERG in clear-cell tumors. However, based on our data, we speculate that loss of expression of KCNH2 (HERG) by methylation could be a good prognostic marker, given that overexpression of the Eag family members promoted increased proliferation and results in poor prognosis.

A number of factors strengthen our study. The number of tumors assessed ($n = 485$), the use of fresh frozen tumor material, the inclusion of technical validation and the validation of the profile in two additional independent ovarian tumor case collections give us confidence in our results. Even though the number of clear-cell cases was small ($n = 30$), we were able to validate the clear-cell-specific methylation profile in two replication sets. It was also advantageous to have consistent pathology review and sample handling for all samples. Compared with conventional unsupervised analytical methods, the approach we used, SS-RPMM, also provided some strengths. We incorporated phenotype information into the learning of clustering patterns, and prevented data over-fitting through nested cross-validations. This enabled automatic selection of optimal cluster number, *M*, which is usually difficult to determine for non-parametric methods, such as *K*-mean and hierarchical clustering. Since logit-transformed methylation beta values are approximately Gaussian-distributed, model-based RPMM is also

expected to be more robust to outliers and powerful to detect underlying data structures than distribution-free clustering approaches. Nonetheless, we note that a limitation of SS-RPMM and all clustering methods is that they are mainly driven by data measurement; as a result, they could not incorporate biological context, such as interactions among genes and coordination of molecular functions. Also, optimality of the numerical model is conditional on statistical assumptions, some of which may not be valid for genomic data.

In conclusion, this study provides a comprehensive analysis of ovarian tumor DNA methylation and demonstrates that methylation in clear-cell ovarian cancer is distinct from other ovarian cancer histologies. Ovarian clear-cell carcinoma is a distinct histopathologic subtype of EOC with an incidence of 5–10% of all ovarian malignancies, and the small sample size (6.2%) in our study is within the expected range of incidence. However, our findings should be interpreted with caution, and validation in a bigger cohort of clear-cell tumors of the ovary is warranted.

MATERIALS AND METHODS

Eligible cases were women aged ≥ 20 years who were ascertained between 1992 and 2009 at the Mayo Clinic within 1 year of diagnosis with pathologically confirmed primary invasive epithelial ovarian, fallopian tube or primary peritoneal cancer. Borderline and non-EOC cases were excluded. Participants were split into three approximately equally sized collections representing a training set ($n = 163$) and two replication sets ($n = 163$, $n = 159$; Table 1). Tumors were snap-frozen immediately following surgery and stored at $-80°C$. The clinical diagnoses were confirmed by re-review by a gynecologic pathologist who verified histology and tumor grade. A gynecologic pathologist (D.W.V.) reviewed each tissue to ensure 70% tumor content prior to DNA extraction. Cases gave informed consent; all protocols were approved by the Mayo Clinic Institutional Review Board.

Methylation arrays

Samples from the training set and the first replication set were assayed and underwent combined quality control procedures at the Mayo Clinic Molecular Genome Facility (Rochester, MN, USA); the second replication set was analyzed at Illumina Corporation (San Diego, CA, USA). Tumor DNA (1 μg at Mayo Clinic, and 250 ng at Illumina) was bisulfite-modified (BSM) using the Zymo EZ96 DNA Methylation Kit (Zymo Research, Orange, CA, USA) according to the manufacturer's protocol. Methylation analyses used the Illumina Infinium HumanMethylation450 BeadChip, which interrogates more than 485 000 CpG loci across the genome (37,38), following the manufacturer's protocol using 250 ng (Mayo Clinic) or 160 ng (Illumina) of BSM DNA. At the Mayo Clinic, BSM-negative and -positive controls were also included ($n = 9$ and $n = 16$, respectively), as were Centre d'Etudes du Polymorphisme Humain (CEPH) DNA replicates ($n = 9$), and samples were run in two batches. Briefly, fragmented DNA was hybridized to BeadChips, which were then processed through a primer extension and an immunohistochemistry

staining protocol to allow detection of a single-base extension reaction. Finally, BeadChips were coated and then imaged on an Illumina iScan. Analysis included control probes for assessing sample-independent and sample-dependent performance. The methylation level of each CpG locus was calculated in GenomeStudio® Methylation module (v.1.9.0). The methylation status of the target CpG sites was determined by comparing the ratio of fluorescent signal from the methylated allele with the sum from the fluorescent signal from both methylated and unmethylated alleles. The two replication sets had 32 samples in common and were used to validate performance of the methylation assays at two different institutions. Intraclass correlation of beta values for each sample that was run at both laboratories was estimated. Each intraclass correlation coefficient was >0.85 (data not shown). One set of the duplicate samples was taken out from replication set 2 for further analysis.

Technical validation

Nine CpGs with a broad spectrum of percent methylation (range 0.23–0.78) and relatively high variability (standard deviation 0.13–0.36) were also assessed using bisulfite Pyrosequencing. Ninety-five samples were tested, including 35 samples that had been run on BeadChips at the Mayo Clinic, 24 samples that had been run on BeadChips at Illumina, 30 samples which had been inadvertently processed on BeadChips at both laboratories (removed from the second replication set for analyses), four BSM-negative controls and two BSM-positive controls (CpGenome™ Universal Methylated DNA; Millipore Corporation, Billerica, MA, USA). Primers (Supplementary Material, Table S1) were designed using the Pyrosequencing Assay Design Software. DNA (20–30 ng) was PCR-amplified using primers, one of which was biotinylated. Briefly, the incorporated biotinylated amplicon was immobilized on streptavidin-coated beads used to purify and render the denatured, single-stranded and biotinylated PCR product. The single-stranded DNA was purified using the Pyrosequencing vacuum workstation. The single-stranded product was annealed to 0.3 μM of the sequencing primer complementary to the single-stranded template and placed at 85°C for 2 min and then cooled to room temperature for 5 min. Pyrosequencing reactions were performed on Biotage PyroMark MD, and data were analyzed using the PyroMD Software. Percent methylation was quantified as methylated C-to-unmethylated C ratio using the Pyro Q-CpG software, which analyzes 96 samples and provides automatic quality control for each sample for the completion of bisulfite conversion and estimates of non-converted DNA. Results suggested that BeadChip data were satisfactory; the median Pearson correlation of methylation values between BeadChip and Pyrosequencing assays was 0.84 for Mayo Clinic and 0.87 for Illumina (Supplementary Material, Table S1).

Methylation data pre-processing

Using the Illumina GenomeStudio software, DNA methylation values from BeadChip assays were scored as beta values, ranging from 0 (unmethylated) to 1 (methylated). Probes which had a reference SNP id or were positioned at a single-nucleotide polymorphism (dbSNP build 137), which had high

beta values in BSM-negative controls (beyond 4 standard deviations of mean) or which were detected in <70% of samples were excluded. A total of 440 649 and 448 543 Illumina probes passed quality control in training set/replication set 1 and replication set 2, respectively (Supplementary Material, Table S2). The intraclass correlation for beta values among CEPH replicates at training and replication set 1 was 0.99 and that for tumor duplicate samples was >0.99. We removed samples that failed or were from cases who had received neoadjuvant chemotherapy. Within each batch, based on an observed plate effect, a correction was applied by fitting a linear model with fixed plate effect for each logit-transformed CpG locus with the unstandardized residual saved. The logit-transformed locus mean was added back onto the residual before back-transforming the value to a 0–1 scale. To adjust for batch effects, we used the COMBAT methodology, which has recently been shown to perform effectively and efficiently compared with other batch-adjustment methodologies (39); briefly, this involved a location (mean) and scale (variance) adjustment using an empirical Bayes framework (40).

Methylation clustering

The scheme of our analysis strategy aimed at identifying and validating novel epigenetic biomarkers of clear-cell ovarian cancer is depicted in Supplementary Material, Figure S5. We used SS-RPMM to determine methylation profiles that are associated with clear-cell histology (41). Using the training set, linear models were fit to each of the CpG loci separately to examine the association between logit-transformed methylation (42) and clear-cell histology. CpG loci were then ranked based on their strength of association with clear-cell histology, and the top M CpG loci were selected for subsequent clustering analysis; M, the optimal number of CpG loci, was selected using a nested cross-validation procedure (Supplementary Material, Figure S1).

Clustering of cases was achieved using the RPMM (43), a hierarchical model-based method for clustering that has been extensively used for the clustering of array-based methylation data (41,44–46). Briefly, RPMM begins by comparing the model goodness-of-fit between one-class and two-class mixture models. If the two-class model fits the data better, these classes are further split into two new classes and compared with the previous split in terms of model goodness-of-fit. Recursion continues until the algorithm arrives at the most parsimonious representation of the data, resulting in an estimate of the number of clusters, as well as the class membership probabilities for each subject. Because of the hierarchical nature of RPMM, classes are referred to in terms of their position on a tree dendrogram. Thus, the class labels for the two-class RPMM solution would be referred to as R and L, representing the right (R) and left (L) branches of the tree, respectively. Successive branches of the tree follow the same nomenclature, wherein the label for each child node is preceded by the label for the respective parent node. Based on the RPMM fit to the training data, a naive Bayes classifier, a probabilistic classifier, was used to predict methylation class membership for the observations in each of the two independent replication set. Subjects were then assigned to the class with the highest membership probability. All models were adjusted for subject age, stage, grade, peritoneal cytology and ascites, as these potential confounding factors were found to be associated with histology.

The association between the predicted methylation classes in the replication sets and clear-cell histology was assessed using permutation chi-square tests and unconditional logistic regression models. In addition, for the top M (22) CpGs, the difference in methylation between clear-cell and non-clear-cell histology was tested in the independent replication sets using a non-parametric Wilcoxon rank-sum test. In addition, an exploratory analysis examined the clinical outcome of the RRL methylation class (consisting primarily, but not exclusively, of clear-cell cases). Specifically, we examined survival in clear-cell cases versus others and in RRL-assigned cases versus others by plotting Kaplan–Meier curves using all data ($n = 485$), accounting for left truncation and right censoring at 10 years of follow-up. The follow-up for all cases was right-censored at 10 years in an attempt to minimize the number of non-ovarian deaths in our analysis.

Preliminary analyses were unadjusted for potential confounders; analyses were carried out using R version 2.13 (www.cran.r-project.org).

Gene expression analyses

For the top M (22) CpGs, analyses of mRNA expression at relevant genes were also performed. Tumor RNA was isolated from 171 fresh frozen samples using the Qiagen RNEasy protocol and quantitated using a Nanodrop Spectrophotomer (Agilent Technologies, Santa Clara, CA, USA). An amount of 750 ng of high-quality total RNA with RNA integrity numbers >8.0 was labeled with cyanine 5-CTP or cyanine 3-CTP, using the Low RNA Input Fluorescent Linear Amplification Kit (Agilent Technologies), purified on RNeasy Mini columns (Qiagen) and hybridized to Agilent whole human genome $4 \times 44K$ expression arrays (using a mixed reference of 107 tumor samples including 63% serous, 5% endometrioid, 3% mucinous, 3% clear cell). Slides were scanned using the Agilent 2565BA Scanner, and data were exported by the Agilent Feature Extraction Software (version 7.5.1) and transferred into the Rosetta Resolver software (Rosetta Inpharmatics LLC, WA, USA) for further analysis. Gene expression data in the form of $\log_2$ ratios of signal from individual tumor to signal from the reference mix were used for analysis. The intensity ratios between the tumor sample and mixed reference calculated for each sequence were computed according to the Agilent error model. For each CpG locus of interest within a gene, we assessed the association of mRNA expression at the corresponding gene with DNA methylation beta values using Spearman's correlations. We tested the difference in methylation between clear-cell and non-clear-cell cases, using a non-parametric Wilcoxon rank-sum test.

Pathway analysis

The relationships between the genes of interest were investigated using knowledge base IPA (Ingenuity® Systems, www.ingenuity.com). The robustness of our clustering allowed us to increase the number of CpG loci from optimal M (22) to 100 that best discriminated of clear-cell histology. This was

achieved by ranking the CpGs based on their strength of association with clear-cell histology and selecting the top 100 on the basis of their *P*-value. Although our SS-RPMM analysis used the top 22 CpGs for the generation of methylation profiles, the top 100 CpGs were selected for our pathway analysis, which are associated with 54 unique genes, which provided us with a larger pool of genes for identifying biologically informative pathways. It is also important to note that the methylation profiles ascertained using the top 22 CpGs versus the top 100 were highly consistent, giving us confidence in using the 54 unique genes associated with the top 100 CpGs. As a result, these 54 genes were used in our pathway analysis, allowing us to gain further insight into the molecular functions of genes associated with differentially methylated CpGs. The aim was to examine whether any biological pathway was over-represented among these genes using a broad functional enrichment analysis with Fisher's exact testing, and to identify and visualize specific networks modeling the molecular relationships of these genes.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

FUNDING

REFERENCES

1. Del Carmen, M.G., Birrer, M. and Schorge, J.O. (2012) Clear cell carcinoma of the ovary: a review of the literature. *Gynecol. Oncol.*, **126**, 481–490.

2. Pearce, C.L., Templeman, C., Rossing, M.A., Lee, A., Near, A.M., Webb, P.M., Nagle, C.M., Doherty, J.A., Cushing-Haugen, K.L., Wicklund, K.G. *et al.* (2012) Association between endometriosis and risk of histological subtypes of ovarian cancer: a pooled analysis of case-control studies. *Lancet Oncol.*, **13**, 385–394.

3. Prat, J. (2012) Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features. *Virchows Arch.*, **460**, 237–249.

4. Ozols, R.F. (2006) Systemic therapy for ovarian cancer: current status and new treatments. *Semin. Oncol.*, **33**, S3–S11.

5. Romero, I. and Bast, R.C. Jr (2012) Minireview: human ovarian cancer: biology, current management, and paths to personalizing therapy. *Endocrinology*, **153**, 1593–1602.

6. Siegel, R., Naishadham, D. and Jemal, A. (2012) Cancer statistics, 2012. *CA Cancer J. Clin.*, **62**, 10–29.

7. Bolton, K.L., Ganda, C., Berchuck, A., Pharoah, P.D. and Gayther, S.A. (2012) Role of common genetic variants in ovarian cancer susceptibility and outcome: progress to date from the ovarian cancer association consortium (OCAC). *J. Intern. Med.*, **271**, 366–378.

8. White, K.L., Schildkraut, J.M., Palmieri, R.T., Iversen, E.S. Jr, Berchuck, A., Vierkant, R.A., Rider, D.N., Charbonneau, B., Cicek, M.S., Sutphen, R. *et al.* (2012) Ovarian cancer risk associated with inherited inflammation-related variants. *Cancer Res.*, **72**, 1064–1069.

9. Skirnisdottir, I., Seidal, T., Karlsson, M.G. and Sorbe, B. (2005) Clinical and biological characteristics of clear cell carcinomas of the ovary in FIGO stages I-II. *Int. J. Oncol.*, **26**, 177–183.

10. Schwartz, D.R., Kardia, S.L., Shedden, K.A., Kuick, R., Michailidis, G., Taylor, J.M., Misek, D.E., Wu, R., Zhai, Y., Darrah, D.M. *et al.* (2002) Gene expression in ovarian cancer reflects both morphology and biological behavior, distinguishing clear cell from other poor-prognosis ovarian carcinomas. *Cancer Res.*, **62**, 4722–4729.

11. Miranda, T.B. and Jones, P.A. (2007) DNA methylation: the nuts and bolts of repression. *J. Cell. Physiol.*, **213**, 384–390.

12. Lande-Diner, L., Zhang, J., Ben-Porath, I., Amariglio, N., Keshet, I., Hecht, M., Azuara, V., Fisher, A.G., Rechavi, G. and Cedar, H. (2007) Role of DNA methylation in stable gene repression. *J. Biol. Chem.*, **282**, 12194–12200.

13. Jones, P.A. and Baylin, S.B. (2007) The epigenomics of cancer. *Cell*, **128**, 683–692.

14. Barton, C.A., Hacker, N.F., Clark, S.J. and O'Brien, P.M. (2008) DNA methylation changes in ovarian cancer: implications for early diagnosis, prognosis and treatment. *Gynecol. Oncol.*, **109**, 129–139.

15. Houshdaran, S., Hawley, S., Palmer, C., Campan, M., Olsen, M.N., Ventura, A.P., Knudsen, B.S., Drescher, C.W., Urban, N.D., Brown, P.O. *et al.* (2010) DNA methylation profiles of ovarian epithelial carcinoma tumors and cell lines. *PLoS One*, **5**, e9359.

16. Campan, M., Moffitt, M., Houshdaran, S., Shen, H., Widschwendter, M., Daxenbichler, G., Long, T., Marth, C., Laird-Offringa, I.A., Press, M.F. *et al.* (2011) Genome-scale screen for DNA methylation-based detection markers for ovarian cancer. *PLoS One*, **6**, e28141.

17. Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature*, **474**, 609–615.

18. Dai, W., Teodoridis, J.M., Zeller, C., Graham, J., Hersey, J., Flanagan, J.M., Stronach, E., Millan, D.W., Siddiqui, N., Paul, J. *et al.* (2011) Systematic CpG islands methylation profiling of genes in the Wnt pathway in epithelial ovarian cancer identifies biomarkers of progression-free survival. *Clin. Cancer Res.*, **17**, 4052–4062.

19. Ehrich, M., Turner, J., Gibbs, P., Lipton, L., Giovanneti, M., Cantor, C. and van den Boom, D. (2008) Cytosine methylation profiling of cancer cell lines. *Proc. Natl Acad. Sci. USA*, **105**, 4844–4849.

20. Nautiyal, S., Carlton, V.E., Lu, Y., Ireland, J.S., Flaucher, D., Moorhead, M., Gray, J.W., Spellman, P., Mindrinos, M., Berg, P. *et al.* (2010) High-throughput method for analyzing methylation of CpGs in targeted genomic regions. *Proc. Natl Acad. Sci. USA*, **107**, 12587–12592.

21. Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.*, **39**, 457–466.

22. Baylin, S.B. and Jones, P.A. (2011) A decade of exploring the cancer epigenome – biological and translational implications. *Nat. Rev. Cancer*, **11**, 726–734.

23. Du, X., Han, L., Guo, A.Y. and Zhao, Z. (2012) Features of methylation and gene expression in the promoter-associated CpG islands using human methylome data. *Comp. Funct. Genom.*, **2012**, 598987.

24. McCluggage, W.G. (2011) Morphological subtypes of ovarian carcinoma: a review with emphasis on new developments and pathogenesis. *Pathology*, **43**, 420–432.

25. Wu, Q., Lothe, R.A., Ahlquist, T., Silins, I., Trope, C.G., Micci, F., Nesland, J.M., Suo, Z. and Lind, G.E. (2007) DNA methylation profiling of ovarian carcinomas and their in vitro models identifies HOXA9, HOXB5, SCGB3A1, and CRABP1 as novel targets. *Mol. Cancer*, **6**, 45.

26. Bol, G.M., Suijkerbuijk, K.P., Bart, J., Vooijs, M., van der Wall, E. and van Diest, P.J. (2010) Methylation profiles of hereditary and sporadic ovarian cancer. *Histopathology*, **57**, 363–370.

27. Gloss, B.S., Patterson, K.I., Barton, C.A., Gonzalez, M., Scurry, J.P., Hacker, N.F., Sutherland, R.L., O'Brien, P.M. and Clark, S.J. (2012) Integrative genome-wide expression and promoter DNA methylation profiling identifies a potential novel panel of ovarian cancer epigenetic biomarkers. *Cancer Lett.*, **318**, 76–85.

28. Yang, H.J., Liu, V.W., Wang, Y., Tsang, P.C. and Ngan, H.Y. (2006) Differential DNA methylation profiles in gynecological cancers and correlation with clinico-pathological data. *BMC Cancer*, **6**, 212.

29. Shih, I.-M., Chen, L., Wang, C.C., Gu, J., Davidson, B., Cope, L., Kurman, R.J., Xuan, J. and Wang, T.-L. (2010) Distinct DNA methylation profiles in ovarian serous neoplasms and their implications in ovarian carcinogenesis. *Am. J. Obstet. Gynecol.*, **203**, 584.e1–584.e22.

30. Michaelson-Cohen, R., Keshet, I., Straussman, R., Hecht, M., Cedar, H. and Beller, U. (2011) Genome-wide de novo methylation in epithelial ovarian cancer. *Int. J. Gynecol. Cancer*, **21**, 269–279.

31. Choi, J.K. (2010) Contrasting chromatin organization of CpG islands and exons in the human genome. *Genome Biol.*, **11**, R70.

32. Jehle, J., Schweizer, P.A., Katus, H.A. and Thomas, D. (2011) Novel roles for hERG K+ channels in cell proliferation and apoptosis. *Cell Death Dis.*, **2**, e193.

33. Lang, F., Foller, M., Lang, K.S., Lang, P.A., Ritter, M., Gulbins, E., Vereninov, A. and Huber, S.M. (2005) Ion channels in cell proliferation and apoptotic cell death. *J. Membr. Biol.*, **205**, 147–157.

34. Cherubini, A., Taddei, G.L., Crociani, O., Paglierani, M., Buccoliero, A.M., Fontana, L., Noci, I., Borri, P., Borrani, E., Giachi, M. *et al.* (2000) HERG potassium channels are more frequently expressed in human endometrial cancer as compared to non-cancerous endometrium. *Br. J. Cancer*, **83**, 1722–1729.

35. Asher, V., Warren, A., Shaw, R., Sowter, H., Bali, A. and Khan, R. (2011) The role of Eag and HERG channels in cell proliferation and apoptotic cell death in SK-OV-3 ovarian cancer cell line. *Cancer Cell Int.*, **11**, 6.

36. Asher, V., Khan, R., Warren, A., Shaw, R., Schalkwyk, G.V., Bali, A. and Sowter, H.M. (2010) The Eag potassium channel as a new prognostic marker in ovarian cancer. *Diagn. Pathol.*, **5**, 78.

37. Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L. *et al.* (2011) High density DNA methylation array with single CpG site resolution. *Genomics*, **98**, 288–295.

38. Sandoval, J., Heyn, H.A., Moran, S., Serra-Musach, J., Pujana, M.A., Bibikova, M. and Esteller, M. (2011) Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics*, **6**, 692–702.

39. Chen, C., Grennan, K., Badner, J., Zhang, D., Gershon, E., Jin, L. and Liu, C. (2011) Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. *PLoS One*, **6**, e17238.

40. Johnson, W.E., Li, C. and Rabinovic, A. (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, **8**, 118–127.

41. Koestler, D.C., Marsit, C.J., Christensen, B.C., Karagas, M.R., Bueno, R., Sugarbaker, D.J., Kelsey, K.T. and Houseman, E.A. (2010) Semi-supervised recursively partitioned mixture models for identifying cancer subtypes. *Bioinformatics*, **26**, 2578–2585.

42. Kuan, P.F., Wang, S., Zhou, X. and Chu, H. (2010) A statistical framework for Illumina DNA methylation arrays. *Bioinformatics*, **26**, 2849–2855.

43. Houseman, E.A., Christensen, B.C., Yeh, R.F., Marsit, C.J., Karagas, M.R., Wrensch, M., Nelson, H.H., Wiemels, J., Zheng, S., Wiencke, J.K. *et al.* (2008) Model-based clustering of DNA methylation array data: a recursive-partitioning algorithm for high-dimensional data arising as a mixture of beta distributions. *BMC Bioinformatics*, **9**, 365.

44. Koestler, D.C., Marsit, C.J., Christensen, B.C., Accomando, W., Langevin, S.M., Houseman, E.A., Nelson, H.H., Karagas, M.R., Wiencke, J.K. and Kelsey, K.T. (2012) Peripheral blood immune cell methylation profiles are associated with nonhematopoietic cancers. *Cancer Epidemiol. Biomarkers Prev.*, **21**, 1293–1302.

45. Christensen, B.C., Smith, A.A., Zheng, S., Koestler, D.C., Houseman, E.A., Marsit, C.J., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Wrensch, M.R. *et al.* (2011) DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma. *J. Natl. Cancer Inst.*, **103**, 143–153.

46. Langevin, S.M., Koestler, D.C., Christensen, B.C., Butler, R.A., Wiencke, J.K., Nelson, H.H., Houseman, E.A., Marsit, C.J. and Kelsey, K.T. (2012) Peripheral blood DNA methylation profiles are indicative of head and neck squamous cell carcinoma: an epigenome-wide association study. *Epigenetics*, **7**, 291–299.