DNA Methylome Analyses Implicate Fallopian Tube Epithelia as the Origin for High-Grade Serous Ovarian Cancer

David Klinkebiel¹, Wa Zhang², Stacey N. Akers³, Kunle Odunsi³,⁴,⁵, and Adam R. Karpf²,⁶

Abstract

High-grade serous ovarian cancer (HGSC) is the most common and lethal form of epithelial ovarian cancer (EOC). Two distinct tissues have been suggested as the tissue of origin: ovarian surface epithelia (OSE) and fallopian tube epithelia (FTE). We hypothesized that the DNA methylome of HGSC should more closely resemble the methylome of its tissue of origin. To this end, we profiled HGSC (n = 10), and patient-matched OSE and FTE (n = 5) primary fresh-frozen tissues, and analyzed the DNA methylome using Illumina 450K arrays (n = 20) and Agilent Sure Select methyl-seq (n = 7). Methylomes were compared using statistical analyses of differentially methylated CpG sites (DMC) and differentially methylated regions (DMR). In addition, methylation analysis was evaluated within a variety of different genomic contexts, including CpG island shores and Homeobox (HOX) genes, due to their roles in tissue specification. Publicly available HGSC methylome data (n = 628) were interrogated to provide additional comparisons with FTE and OSE for validation. These analyses revealed that HGSC and FTE methylomes are significantly and consistently more highly conserved than are HGSC and OSE. Pearson correlations and hierarchal clustering of genes, promoters, CpG islands, CpG island shores, and HOX genes all revealed increased relatedness of HGSC and FTE methylomes. Thus, these findings reveal that the landscape of FTE more closely resembles HGSC, the most common and deadly EOC subtype.

Implications: DNA methylome analyses support the hypothesis that HGSC arise from the fallopian tube and that due to its tissue-specificity and biochemical stability, interrogation of the methylome may be a valuable approach to examine cell/tissue lineage in cancer.

Introduction

High-grade serous ovarian cancer (HGSC), the most common and lethal subtype of epithelial ovarian cancer (EOC), is frequently diagnosed at an advanced stage, where long-term survival is poor (1). Understanding the mechanisms driving initiation and progression of HGSC is critical for the development of new diagnostic and therapeutic approaches. In this context, the tissue and cellular origin of HGSC remains a critical question in the field (2–4). The first widely accepted model for the origin of HGSC implicated transformation of the ovarian surface epithelia (OSE), possibly by incessant ovulation (5, 6). This hypothesis is supported by several observations, including the existence of benign epithelial cysts (cystadenomas) in the ovary, a precursor lesion for EOC, and epidemiologic links between ovulation and ovarian cancer (3, 4). On the basis of the OSE origin model, experimental model systems for EOC have used primary and immortalized mouse and human OSE cells, and transgenic mice created by genetic manipulation of the OSE (7, 8). A more recent hypothesis for the origin of HGSC invokes transformation of fallopian tube fimбриae epithelia (FTE; refs. 9, 10). This model is supported by the identification of precursor lesions in the distal fallopian tube, including “p53 signatures,” lesions characterized by increased p53 protein expression due to TP53 mutations, and serious tubal intraepithelial carcinoma (STIC) lesions, which are characterized by increased proliferation and multiple cell layers. Cells emanating from STIC lesions are hypothesized to spread to the ovary, where they develop into invasive serous carcinoma (1, 11). The FTE model is supported by several observations: (i) identification of common FTE but not OSE precursor lesions in BRCA mutation carriers (12), (ii) reduced EOC risk in BRCA mutation carriers following bilateral tubal ligation (13), (iii) gene-expression profiling studies showing similarity of FTE and HGSC (14, 15), and (iv) development of HGSC in vivo following the engineering of specific genetic alterations (Tp53, Brcal21, and Pten) in mouse FTE (16).

In the mammalian embryo, extensive cytosine DNA demethylation erases the bulk of the gamete methylation pattern, which is followed by coordinated remethylation establishing cell and tissue type-specific DNA methylation in somatic cells (17). Consequently, DNA methylation participates in the establishment of an epigenetic signature that contributes to cell and tissue type-
specific chromatin organization and gene expression (18–21). Whole-genome bisulfite sequencing analysis of several human tissues in different individuals has recently revealed widespread tissue-specific DNA methylation variation in humans (22). In addition to its normal function in X-chromosome inactivation, genomic imprinting, tissue differentiation, and transposable element silencing, altered DNA methylation makes a major contribution to human diseases, including cancer (23, 24). In cancer, two general DNA methylation defects are common: gene-specific promoter hypermethylation and global DNA hypomethylation (24). Despite these oncogenic changes, tumors might be anticipated to retain tissue-specific DNA methylation that reflect their cellular and tissue origin (25, 26).

In this study, we hypothesized that the DNA methylyome provides a means to investigate cell/tissue origin in HGSC. Comparative analysis of the methylyome of HGSC, normal FTE, and normal OSE may provide insight into the tissue of origin of HGSC, due to tissue-specific methylation (18, 22, 27, 28). We interrogated the methylyomes of HGSC, FTE, and OSE using two complementary methods, and analyzed methylation in several different genomic contexts, to determine the degree of relatedness of the HGSC methylyome to FTE and OSE. Our data indicate that the HGSC methylyome is consistently and significantly more similar to FTE than OSE. Beyond the implications of this study for understanding the origin of HGSC, DNA methylation profiling may serve as a useful method for cell of origin mapping for other cancers. The stability of DNA methylation, a covalent DNA modification, is an advantageous aspect of this approach in the clinical setting.

Materials and Methods

Human tissues

We obtained fresh-frozen primary HGSC (n = 10) from patients undergoing surgical resection at Roswell Park Cancer Institute (RPCI), and patient-matched fresh-frozen normal OSE and FTE (n = 5) at RPCI, as described previously (Supplementary Table S1; ref. 29). Briefly, OSE and FTE obtained from patients without malignancy were harvested by mechanical scraping and processing of the epithelial layer of resected ovaries and the fimbriae end of fallopian tubes, immediately upon surgical removal. HGSC tissues were estimated by pathology to contain >80% neoplastic cells. All samples were collected using IRB-approved protocols at RPCI, and sample processing has been described previously (29, 30). We isolated genomic DNA using the Puregene Tissue Kit (Qiagen), which includes RNase treatment. When selecting HGSC samples, we took into account cancer-specific global DNA hypomethylation, in which repetitive DNA elements are significantly hypomethylated genome-wide (29, 31). Specifically, to eliminate potentially confounding effects due to this phenotype, we used HGSC showing similar LINE-1 DNA methylation as FTE and OSE, as determined by pyrosequencing (data not shown; refs. 29, 31). However, an independent analysis of HGSC samples displaying global DNA hypomethylation affirmed the conclusions presented here (data not shown).

DNA methylation data

New data. We performed Illumina Infinium 450K bead arrays (450K), which assessed methylation at approximately 470,000 CpG sites, at the RPCI (n = 7) and University of Utah (n = 13) Genomics Core Facilities (Supplementary Table S1). We performed Agilent SureSelect Methylation bisulfite sequencing (methyl-seq), a targeted solution hybridization method (32), which analyzed methylation at approximately 4 × 10^6 CpGs, at the University of Nebraska Medical Center (UNMC) Epigenomics Core (n = 7; Supplementary Table S1). We used either the Zymo Pico Methyl-Seq or Agilent SureSelect Methyl-Seq Kit for library preparations, and Agilent SureSelect baits to pull down the final sequences. We conducted high-throughput sequencing at the UNMC Sequencing Core, using an Illumina HiSeq 2500 Genome Analyzer. Sequencing parameters and results are shown in Supplementary Table S2. We aligned sequence tags to the human genome (hg19) using Bismark, and selected only those CpG with ≥10× coverage (33). For clarity, we refer to newly generated methylene data as “Karpf 450K data” or “Karpf Methyl-seq data.”

Public data. We used 450K data from FTE (n = 7) and primary HGSC (n = 78) from Gene Expression Omnibus (GEO) GSE65821 (Supplementary Table S3); these data have been published and are referred to as “Bowtell 450K data” (34). We only used primary tumor samples, to avoid the influence of disease recurrence and drug resistance. We also utilized Illumina Infinium 27K methylation data (27K) from primary HGSC (n = 550) using The Cancer Genome Atlas (TCGA) data portal (35); these data are referred to as “TCGA 27K data.” For analysis of TCGA data, we only utilized CpG sites that overlapped Karpf FTE and OSE 450K data (25,779 CpG sites).

DNA methylation data analysis

We used RnBeads (36) to analyze all methylene data (450K, 27K, Methyl-seq), and restricted our analysis to CpG methylation. We analyzed both differentially methylated CpG sites (DMC; ≥25% methylation change) and differentially methylated regions (DMR; contiguous regions of any length containing ≥3 CpGs and ≥25% methylation change). RnBead data included 5kb genomic tiles (n = 131,408), genes [transcriptional start site (TSS) to transcription end, n = 30,514], promoters (~1,500 to 500 bp relative to the TSS, n = 30,630) and CpG islands (n = 26,595). We additionally analyzed CpG island shores (~2,000bp of CpG island; n = 53,190; ref. 37), and enhancers (n = 32,693; Transcribed Enhancer Atlas Database; http://enhancer.binf.ku.dk/index.php). We determined the overlap of CpG sites with different genomic regions using the Bedtools intersect routine (38). We used an FDR-adjusted P value of <0.05 within RnBeads. We created hierarchal clustering and methylation heat maps using T4M microarray software Multi Experiment Viewer (MeV), based on a Pearson correlation metric and average linkage (39, 40).

To compare individual HGSC DNA methylomes with FTE and OSE, we determined DMC for each HGSC using the R software package DSS (41), by smoothing the RnBeads normalized percent methylation values for 0.5 kb units, and using a moving average algorithm. DMR were defined as regions of any length containing ≥3 CpGs and ≥25% methylation change. We quantified the number of HGSC DMC and DMR that showed a significant difference as compared to FTE or OSE, using a Wald test P value of ≤0.05.

HOX gene analysis

We downloaded the coordinates of all human HOX genes and pseudogenes (n = 333) from the Homebox database (http://
homoeodb.zoo.ox.ac.uk/), and aligned CpG sites using genomic locations from −10 kb upstream the TSS to the transcript end. We determined the overlap of CpG sites that were significantly different (FDR-adjusted P value of <0.05) between FTE and OSE with HOX gene genomic regions using the Bedtools intersect routine (38). We created hierarchical clustering and methylation heatmaps using Tm4 microarray software Multi Experiment Viewer (MeV), based on a Pearson correlation metric and average linkage (39, 40).

Genomic Data Deposit and Public Access

DNA methylation data (450K, .idat files, and methyl-seq, .fastq files) were deposited into the NCBI Gene Expression Omnibus (GEO# GSE81228).

Results

Comparison of the HGSC, FTE, and OSE methylomes

The aim of this study was to compare the degree of relatedness of the DNA methylome of primary HGSC with primary FTE and OSE, to infer the tissue lineage of HGSC. For this task, we used a set of patient-matched primary FTE and OSE tissues from patients without malignancy (n = 5). To verify the origin of these samples, we compared their 450K methylation with recently reported FTE 450K methylome data (n = 7; ref. 34). We compared CpG methylation within a variety of genomic contexts, as well as comparisons of total DMC and DMR (see Materials and Methods). This analysis verified conservation of the two independent FTE methylome data and illustrated significant divergence, in all genomic contexts, using patient-matched OSE (Supplementary Table S4).

We analyzed and compared HGSC, FTE, and OSE methylomes from Karpf 450K data. Figure 1 presents a comparison of methylation of all CpG sites (DMC) and all DMR (5 kb tiles). In both comparisons, HGSC methylation was strikingly more similar to FTE than OSE. To validate this finding, we calculated Pearson correlation coefficients for the comparisons, using all CpG sites or methylation restricted to different genomic regions. In each instance, HGSC showed greater correlation with FTE, as compared with OSE (Supplementary Fig. S1). We next used principal component analyses (PCA) to test this relationship, and again observed an increased similarity of HGSC with FTE, as compared with OSE (Supplementary Fig. S2).

CpG island shores are genomic regions adjacent to CpG islands that display tissue-specific DNA methylation (37), making them relevant for this study. We identified individual CpG sites within CpG island shores showing differential methylation in FTE and OSE, and used these to perform hierarchical clustering of the three sample groups. Importantly, all HGSC samples clustered more closely to FTE than to OSE (Fig. 2A). We additionally observed that one FTE sample clustered within the OSE samples (Fig. 2A; see also Figs. 2B and 3). We speculate this is due to a predominance of individual-specific DNA methylation differences in this particular patient.

A summary of all Karpf 450K methylene comparisons is provided in Table 1 (top section). χ² testing validated the significantly increased similarity of HGSC and FTE as compared with OSE, in each comparison.

**Figure 1.** Differential methylation of HGSC versus patient-matched FTE or OSE, using Karpf 450K data. Each shows an x-y plot of the mean methylation beta value of specific CpG sites or 5-kb tile regions in HGSC (y-axes) and FTE or OSE (x-axes). Sites or regions without a significant change in methylation are not shown (blank regions in middle of graphs). Sites or regions with increased methylation in HGSC are indicated on the top left of graphs, whereas sites or regions with increased methylation in FTE or OSE are indicated on the bottom right of graphs. A, DMC comparison of HGSC (n = 10) and FTE (n = 5; DMC = mean CpG methylation beta value difference ≥25%; FDR adjusted P value of <0.05). B, DMC comparison of HGSC (n = 10) and OSE (n = 5), as described in A. C, DMR comparison between HGSC (n = 10) and FTE (n = 5). DMR were 5-kb tiles (n = 131,408) with a mean beta value difference of ≥25% and an FDR-adjusted P value of <0.05. D, comparison of HGSC (n = 10) and OSE (n = 5), as described in C. The total number of DMC and DMR meeting the differential cut-off value in each comparison is indicated on the figure. All analyses were performed using RnBeads (see Materials and Methods).
The methylation patterns of PAX8, mesothelin (MSLN), and Homeobox (HOX) genes are divergent in OSE as compared with HGSC and FTE.

We examined the methylation of two genes involved FTE lineage specificity, PAX8 and MSLN (14, 16), hypothesizing that they may show divergence in OSE. As anticipated, PAX8 and MSLN showed relatively similar methylation in HGSC and FTE as compared with OSE (Supplementary Fig. S3). We next examined methylation of HOX genes, as they are involved in development and tissue differentiation, are known to be regulated by DNA methylation, and can show altered methylation and expression in HGSC (42–45). We analyzed methylation of all HOX (and HOX domain-containing) genes (n = 333 genes, n = 9011 CpGs) using Karpf 450K data. Hierarchical clustering of sample groups using HOX-associated CpGs that showed significant differential methylation between FTE and OSE (n = 179) revealed increased similarity of HGSC and FTE (Fig. 3). Specific examples of HOX genes showing divergent methylation patterns in OSE as

Figure 2.
Unsupervised hierarchal clustering of sample groups (HGSC, FTE, and OSE) using differentially methylated CpGs associated with CpG island shores. CpGs selected for analysis were: (i) located within CpG island shores, and (ii) showed a significant mean beta value difference between FTE and OSE (no beta value cutoff; FDR-adjusted P value <0.05). A, hierarchical cluster dendogram of HGSC (n = 10), FTE (n = 5), and OSE (n = 5) using Karpf 450K data. Data are comprised of 1827 CpG sites. B, hierarchical cluster dendogram of HGSC (n = 45), FTE (n = 5), and OSE (n = 5), using Bowtell 450K (HGSC) and Karpf 450K (FTE, OSE) data. Data are comprised of 1753 CpG sites. All analyses were performed using RnBeads (see Materials and Methods).
compared with FTE and HGSC are HOXB1 and HOXB7, which were hypermethylated in OSE as compared to FTE and HGSC (Fig. 2B) and HOXD3 and EMX2 (a homeodomain-containing gene), which were hypermethylated in OSE as compared to FTE or HGSC (Fig. 4C and D). We also noted that several HOX genes, including HOXA5, A9, A10, A11, B5, and D11, showed cancer-specific DNA hypermethylation, in agreement with earlier reports (Supplementary Figs. S4–S5; refs. 45–47).

Methyl-seq confirms the relatedness of the HGSC and FTE methylomes

As a complementary approach to 450K, we used methyl-seq to analyze a subgroup of tumor and normal samples, as described in Materials and Methods (Supplementary Table S1). Data from 450K and methyl-seq were well correlated (Supplementary Table S5). The compiled results of methyl-seq are reported in Table 1. As observed with Karpf 450K data, χ² testing of methyl-seq revealed significantly increased similarity of HGSC and FTE as compared with HGSC and OSE, in each genomic context.

Independent cohorts of HGSC methylation data validate the relatedness of the HGSC and FTE methylomes

We used two published primary HGSC datasets, Bowtell 450K data (n = 78; ref.34) and TCGA 27K data (n = 550; ref. 35), and compared these with our patient-matched primary FTE and OSE 450K data. Given the low genomic coverage of Illumina 27K, we confined our analysis of methylation within different genomic contexts to Bowtell 450K data. Pearson correlation and PCA analyses both indicated increased similarity of HGSC to FTE as compared with OSE (Supplementary Figs. S6–S7). Hierarchical clustering of CpG island shore methylation data also revealed increased relatedness of HGSC with FTE as compared with OSE (Fig. 2B). χ² analysis revealed significantly greater similarity of HGSC with FTE as compared with OSE, in all genomic contexts (Table 1). Despite the reduced coverage of 27K data, it still illustrated the similarity of HGSC to FTE as compared with OSE (Supplementary Fig. S8; Table 1).

Individual HGSC sample methyolome analysis validated the similarity of HGSC and FTE, as compared with OSE

The above data report group-based comparative methylome analyses. To determine the extent to which individual HGSC samples are related to FTE and OSE, we used Karpf 450K and Bowtell 450K data to classify individual samples according to DMC or total DMR. For total DMC, 76 of 78 (99%) HGSC showed increased similarity to FTE, whereas, for DMR, 77 of 78 (99%) HGSC showed increased similarity to FTE (Fig. 5A and B). These data affirm the conclusions drawn from group-wise analyses.

Table 1. Summary of HGSC versus patient-matched FTE and OSE DNA methylome comparisons

<table>
<thead>
<tr>
<th>Sample group comparison</th>
<th>All Sites</th>
<th>CpG islands</th>
<th>CpG island shores</th>
<th>HOX genes</th>
<th>Enhancers</th>
<th>Skb Tiles</th>
<th>CpG islands</th>
<th>Promoters</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karpf 450K, HGSC (n = 10), patient-matched FTE and OSE (n = 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HGSC vs. FTE</td>
<td>1,017</td>
<td>388</td>
<td>417</td>
<td>222</td>
<td>29</td>
<td>3</td>
<td>32</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>HGSC vs. OSE</td>
<td>19,102</td>
<td>3,625</td>
<td>6,096</td>
<td>1,825</td>
<td>530</td>
<td>741</td>
<td>472</td>
<td>223</td>
<td>41</td>
</tr>
<tr>
<td>Fold difference</td>
<td>18.8 (****)</td>
<td>9.3 (****)</td>
<td>14.6 (****)</td>
<td>8.2 (****)</td>
<td>18.3 (****)</td>
<td>247 (****)</td>
<td>14.8 (****)</td>
<td>223 (****)</td>
<td>41 (****)</td>
</tr>
<tr>
<td>Karpf Methyl-seq, HGSC (n = 3), patient-matched FTE and OSE (n = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HGSC vs. FTE</td>
<td>125,586</td>
<td>40,298</td>
<td>50,322</td>
<td>20,850</td>
<td>3,019</td>
<td>2,649</td>
<td>714</td>
<td>530</td>
<td>325</td>
</tr>
<tr>
<td>HGSC vs. OSE</td>
<td>161,561</td>
<td>47,886</td>
<td>62,991</td>
<td>23,550</td>
<td>4,061</td>
<td>4,580</td>
<td>1,543</td>
<td>93</td>
<td>510</td>
</tr>
<tr>
<td>Fold difference</td>
<td>1.3 (****)</td>
<td>1.2 (****)</td>
<td>1.3 (****)</td>
<td>1.1 (****)</td>
<td>1.3 (****)</td>
<td>1.7 (****)</td>
<td>2.2 (****)</td>
<td>1.7 (****)</td>
<td>1.6 (****)</td>
</tr>
<tr>
<td>Bowtell 450K, HGSC (n = 78), patient-matched FTE and OSE (n = 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HGSC vs. FTE</td>
<td>21,646</td>
<td>3,130</td>
<td>7,209</td>
<td>1,732</td>
<td>455</td>
<td>1,085</td>
<td>460</td>
<td>379</td>
<td>190</td>
</tr>
<tr>
<td>HGSC vs. OSE</td>
<td>39,439</td>
<td>5,461</td>
<td>12,158</td>
<td>2,711</td>
<td>991</td>
<td>1,879</td>
<td>882</td>
<td>641</td>
<td>302</td>
</tr>
<tr>
<td>Fold difference</td>
<td>1.8 (****)</td>
<td>1.7 (****)</td>
<td>1.7 (****)</td>
<td>1.6 (****)</td>
<td>2.2 (****)</td>
<td>1.7 (****)</td>
<td>1.9 (****)</td>
<td>1.7 (****)</td>
<td>1.6 (****)</td>
</tr>
<tr>
<td>TCGA 27K, HGSC (n = 550), patient-matched FTE and OSE (n = 5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HGSC vs. FTE</td>
<td>1,260</td>
<td>134</td>
<td>523</td>
<td>55</td>
<td>9</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>HGSC vs. OSE</td>
<td>1,592</td>
<td>180</td>
<td>583</td>
<td>77</td>
<td>8</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Fold difference</td>
<td>1.2 (****)</td>
<td>1.3 (****)</td>
<td>1.1 (ns)</td>
<td>1.4 (ns)</td>
<td>0.9 (ns)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Abbreviations: N/A, insufficient coverage to conduct DMR measurements; ns, not significant.

*FDR < 0.05, methylated difference ≥25%.
*FDR < 0.05, mean methylation difference ≥25%, ≥3 CpGs per region.
*Increase in HGSC versus OSE as compared with HGSC versus FTE.
*χ²-P value: (****) < 0.0001; (****) < 0.001; (****) < 0.01; (****) < 0.05.
Discussion

Using the DNA methylome as a classifier, our data reveal that HGSC more closely resembles FTE than OSE. This relationship was conserved regardless of the HGSC sample population, methylome analysis method, or genomic location, including all CpGs, CpG islands, genes, promoters, and all DMR. In addition, CpG island shores and enhancers maintained this relationship, which is notable based on previous studies indicating the importance of tissue-specific differential methylation in CpG island shores, and the well-established role of enhancers in driving tissue specification (37, 48). Thus, our study provides support for the model that HGSC originates in the fimбриae end of the fallopian tube. We also note that use of the proper normal control is an essential aspect of molecular studies of HGSC, and our data indicate that FTE should be used for this purpose.

Within the FTE, there are two major cell types, secretory and ciliated cells. In addition, FTE harbors a minor population of basal cells, which includes the stem cell niche (3). Secretory cells are proliferative and help regenerate the epithelium during cell turnover, whereas ciliated cells are terminally differentiated (11). Determination of the DNA methylomes of individual FTE cell types will provide further insight into the cell lineage of HGSC, and how the methylome becomes altered in cancer.

Figure 4.
Methylation of specific HOX genes in HGSC, FTE, and OSE. DNA methylation of HOXB1 (A), HOXB7 (B), HOXD3 (C), and EMX2 (D) in HGSC (n = 10) and patient-matched FTE (n = 5), and OSE (n = 5), as determined using Karpf 450K data. Top, chromosome, gene, and CpG island locations (UCSC genome browser), and map of CpG sites analyzed by 450K. Bottom, CpG methylation data for sample groups; red boxes indicate regions that contain differently methylated CpG sites (FDR-adjusted P value < 0.05) between FTE or OSE as compared with HGSC. Broken arrows indicate the TSS.

Figure 5.
Individual HGSC sample methylation differences as compared with patient-matched FTE or OSE group averages. DNA methylation data from Karpf 450K for HGSC (n = 10), FTE (n = 5), and OSE (n = 5), and Bowtell 450K HGSC data (n = 78) were used. A, DMC between Karpf and Bowtell HGSC samples and FTE or OSE. B, DMR between Karpf and Bowtell HGSC samples and FTE or OSE. In both, a Wald test P value of < 0.05 was used.
HOX proteins play key roles in cell and tissue identity and HOX genes can be regulated by DNA methylation (42–44). We therefore analyzed HOX genes and observed that overall HOX methylation in HGSC more closely resembled FTE as compared with OSE. In particular, we show that HOXB1, HOXB7, HOXD3, and EMX2 illustrate a methylation signature that was conserved between HGSC and FTE, but divergent in OSE.

Although overall our data implicate strong conservation between the HGSC and FTE methyleomes as compared with OSE, we noted in one comparison (of CpG island shore methylation) that a small subset of HGSC clustered more closely to OSE (Fig. 2B). Therefore, we cannot formally exclude that a subpopulation of HGSC has greater similarity to, and therefore potentially arises from, OSE. In addition, we cannot exclude that the robust relative conservation of the HGSC and FTE methyleomes observed overall, rather than implicating FTE as the cell/tissue origin, could result from a transdifferentiation event, for example, Müllerian metaplasia (4). A final limitation of our study is that it used clinically advanced HGSC. Methyleome studies of early-stage tumors or HGSC precursor lesions will provide important additional insight into the cellular origin of HGSC.

Beyond the relevance of this study for addressing the tissue of origin of HGSC, our data more generally suggest that DNA methyleome analyses may serve as a useful method for mapping cellular and tissue origins in other cancers. The stability of DNA methylation patterns may serve as a useful method for mapping cellular and tissue origins in other cancers. The stability of DNA methylation, as both a covalent chemical modification and a stable (replication-coupled) epigenetic mark, is advantageous when considering methyleome profiling in the clinical setting. DNA methylation can be measured in biological fluids, fresh-frozen tissue samples, and formalin-fixed paraffin-embedded archival materials (49). Moreover, emerging methyleome data implicate DNA methylation in the processes of tissue differentiation and the maintenance of tissue specificity (18, 19, 22, 27, 28). These two traits (biochemical stability and tissue specificity) support the utility and relevance of DNA methyleome profiling to investigate the cellular origins of cancer.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: D. Klinkebiel, A.R. Karpf
Development of methodology: D. Klinkebiel
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D. Klinkebiel, W. Zhang, S. Akers, K. Odunsi
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D. Klinkebiel, A.R. Karpf
Writing, review, and/or revision of the manuscript: D. Klinkebiel, S. Akers, K. Odunsi, A.R. Karpf
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S. Akers, A.R. Karpf
Study supervision: A.R. Karpf

Acknowledgments
We thank Ann-Marie Patch and David Bowtell for helpful assistance and data sharing, and Shashikant Lele for support. We thank the University of Buffalo, Center of Excellence in Genomics and Bioinformatics, RPCI Bioinformatics and Genomics Core, University of Utah Genomics Core, UNMC Epigenomics Core, UNMC Bioinformatics Core, and the UNMC DNA Sequencing and Microarray Core (supported by NICMS 8P20GM103427 and P20GM103471) for assistance.

Grant Support
This work was supported by The Otis Glee Medical Research Foundation (A.R. Karpf), The Betty J. and Charles D. McKinsey Ovarian Cancer Research Fund (to A.R. Karpf), The Fred & Pamela Buffett Cancer Center (NIH P30 CA036727), and NIH (to A.R. Karpf), and NIH T32CA108456 (to S.N. Akers). K. Odunsi acknowledges support from NIH P30 CA036727, NIH R01CA158318, and the Roswell Park Alliance Foundation.

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

Received March 18, 2016; revised May 9, 2016; accepted May 16, 2016; published OnlineFirst June 3, 2016.

www.aacrjournals.org
DNA Methyleome of Ovarian Cancer and Tissues of Origin

References

4. Dubéau L, Drapkin R. Coming into focus: the nonovarian origins of HGSC. Methyleome studies of early-stage tumors or HGSC precursor lesions will provide important additional insight into the cellular origin of HGSC.

Published OnlineFirst June 3, 2016; DOI: 10.1158/1541-7786.MCR-16-0097
Mol Cancer Res; 14(9) September 2016
OF7

Downloaded from mcr.aacrjournals.org on October 27, 2017. © 2016 American Association for Cancer Research.

""