Targeted DNA Methylation Screen in the Mouse Mammary Genome Reveals a Parity-Induced Hypermethylation of Igf1r That Persists Long after Parturition

Tiffany A. Katz1, Serena G. Liao2, Vincent J. Palmieri1, Robert K. Dearth3, Thushangi N. Pathiraja4, Zhiguang Huo2, Patricia Shaw2, Sarah Small1, Nancy E. Davidson1, David G. Peters7, George C. Tseng2, Steffi Oesterreich1, and Adrian V. Lee1

Abstract

The most effective natural prevention against breast cancer is an early first full-term pregnancy. Understanding how the protective effect is elicited will inform the development of new prevention strategies. To better understand the role of epigenetics in long-term protection, we investigated parity-induced DNA methylation in the mammary gland. FVB mice were bred or remained nulliparous and mammary glands harvested immediately after involution (early) or 6.5 months following involution (late), allowing identification of both transient and persistent changes. Targeted DNA methylation (109 Mb of Ensemble regulatory features) analysis was performed using the SureSelectXT Mouse Methyl-seq assay and massively parallel sequencing. Two hundred sixty-nine genes were hypermethylated and 128 hypomethylated persistently at both the early and late time points. Pathway analysis of the persistently differentially methylated genes revealed Igf1r to be central to one of the top identified signaling networks, and Igf1r itself was one of the most significantly hypermethylated genes. Hypermethylation of Igf1r in the parous mammary gland was associated with a reduction of Igf1r mRNA expression. These data suggest that the IGF pathway is regulated at multiple levels during pregnancy and that its modification might be critical in the protective role of pregnancy. This supports the approach of lowering IGF action for prevention of breast cancer, a concept that is currently being tested clinically. Cancer Prev Res. 8(10); 1000–9. ©2015 AACR.

Introduction

An early first full-term pregnancy is the most effective method for breast cancer prevention. The very first report associating nulliparity with increased breast cancer risk was more than 300 years ago. Bernardino Ramazzini, the father of industrial medicine, noticed that "tumors of the breast are found more often in nuns than any other women" and speculated that this was due to a life of celibacy (1). In 1970, this phenomenon was revisited in a landmark case-control study finding that the risk of breast cancer in parous women who gave birth before the age of 20 years is half that of nulliparous women (2). These observations were supported by further studies showing that women who gave birth at or before 20 or 25 years of age have a 50% and 38% reduction in lifetime risk of breast cancer, respectively (3, 4); however, the detailed cellular and molecular mechanisms underlying this phenomenon in humans remain unclear. Studies have shown morphologic changes and molecular alterations in the postpartum breast that are likely linked to the reduced breast cancer risk (5). In order for women to be protected from breast cancer for such a long period of time postpartum (30–40 years later), there must be a permanent alteration driving these molecular and morphologic changes. We hypothesized that epigenetic alterations to the genome, which are inducible and long-lasting, mediate parity-induced protection against breast cancer.

Parity-induced protection from mammary cancer has been replicated in multiple rodent models, including full-term pregnancy and pseudopregnancy, using both carcinogen and spontaneous carcinoma models (6–8). Nulliparous animals treated with estradiol at a dose that results in circulating levels similar to pregnancy leads to decreased carcinogen-induced mammary cancer, and administering estradiol plus progesterone enhances the protective effect presumably because these conditions better mimic a pregnant state (6, 7). The molecular and cellular mechanisms of pregnancy protection are likely multifaceted involving both changes in systemic hormones and in cell differentiation, signaling, and survival in the mammary gland (9). An early full term birth (FFTB) remains the single most effective natural method to prevent breast cancer, and the discovery of mechanisms driving this phenomenon will stimulate development of new prevention approaches.
Epigenetic changes are known to control mammary growth and development. During pregnancy and lactation, DNA is hypo-methylated to allow expression of genes controlling remodeling of the gland and milk production (10). In addition, changes in chromatin structure have been shown to be induced by parity. In nulliparous women, breast epithelial nuclei are large and euchromatic, in contrast to breast tissue from parous women that display small heterochromatic nuclei with strong methylation of histones at repressive marks (11). Epigenetic alterations are both stable and inducible and thus can persistently alter gene expression and induce long-term phenotypic “memory-like” changes. Therefore, it is likely that epigenetic regulations underlie the mechanisms responsible for the protective effect of pregnancy against breast cancer. We have used a new massively parallel targeted sequencing approach to analyze differentially methylated regions (DMR) across the genome in parous and nulliparous mammary glands, with the goal of identifying pathways that are important in preventing breast cancer. The insulin-like growth factor receptor (Igf1r), as well as other members of the IGF signaling pathway, displayed increased methylation with parity and may thus be involved in protecting parous individuals from breast cancer.

Materials and Methods

Animal experiments
All animal experiments were approved by the IACUC of Baylor College of Medicine. FVB mice (Jackson Laboratories) at 55 days of age were bred to become parous or remained nulliparous (n = 10 per group at each time point). Parous animals nursed for 21 days and were allowed to undergo involution for 28 days. Animals were euthanized immediately (day 125) and 6.5 months (day 315) post-involution, respectively. Nulliparous mice were euthanized at both time points. The experimental schema is shown in Supplementary Fig. S1A. Tissues were harvested, flash-frozen in liquid nitrogen, and stored at −80°C for further analysis.

Targeted assessment of DNA methylation in DMRs
Frozen mammary glands (n = 6 per group at each time point) were cryo-homogenized and gDNA was isolated using the DNeasy Blood and Tissue reagents and protocol (Qiagen). The gDNA was sheared into 200-bp fragments using a Covaris S220 ultrasonicator. Sheared DNA underwent end-repair (XT Library Prepkit), polyA tail addition, and adapter ligation, verified by a 40-bp shift in size on an Agilent Bioanalyzer. Subsequently, the DNA was hybridized to SureSelectXT Mouse Methyl-Seq library probes overnight at 65°C and captured with streptavidin magnetic beads. The eluted DNA was bisulfite converted (Zymo EZ DNA Meth Gold) and PCR amplified. Samples were then pooled and 3 samples per lane were analyzed by 100-bp paired-end sequencing on an Illumina HiSeq 2500. The SureSelectXT Mouse Methyl-Seq design covers 109 Mb of Ensemble regulatory features (CpG shores and shelves, DNase I hypersensitive sites, transcription factor–binding sites, etc.), CpG islands, known tissue-specific DMR, and open regulatory elements.

Biostatistical analysis of DNA methylation in DMRs
Processing and quality control. The alignment reference genome used was mm9 assembly (12). The mm9 DNA reference genome was converted to a DNA methylation reference genome. Genome indexing was performed using Bismark genome preparation tools. FastQ files (23.4–146.4 million pairs of reads per sample, with an average 62 million pairs of reads) were aligned to the converted methylation reference using Bismark. The maximum number of mismatches permitted was 2. For valid paired-end alignments, the minimum insert size was 20 and the maximum insert size was 1,200. The remaining parameters from Bismark were used as default. Aligned reads outside of the targeted regions (Agilent SureSelectXT Mouse Kit) were removed. We also conducted hierarchical clustering on all samples from both time points (Supplementary Fig. S2). One sample at the early time point (NP5) was removed from further analysis due to low mapping efficiency of 9%. One sample in each time point was removed because of nonconformity in the principal component analysis (PCA; early, P12 is an outlier from all samples; late, NP4 is an outlier from all NP samples; Supplementary Fig. S3).

Differential methylation analysis. MethylKit software (a public R Package) was used to process sequencing data (13). DMR were identified by creating 120-bp windows (not overlapping) and filtering out any windows with less than 10 reads. All windows were normalized to adjust for bias across all samples. Windows that did not align in the targeted regions were discarded. The β value was calculated as methylated CpG/total CpG. Logistic regression in MethylKit was used with a sliding window analysis and a cutoff of q < 0.01 and methylation difference > 25% to identify differential methylation between parous and nulliparous samples at each time point. P values were adjusted to q values using the SLIM method to obtain false discovery rates (FDR; ref. 14). We also assigned samples to parous and nulliparous groups in random combinations at each time point and permuted this procedure 100 times. Our results are highly significant compared with randomly assigning treatment groups (P < 0.05 for all cases). A 25% difference in DNA methylation has been shown to induce a 2-fold reduction in gene expression (15). To eliminate any significance driven by outliers, windows were further filtered out by a trimmed mean difference < 0.25 (mean value with 25% observation deleted at both ends). These data are presented in the first two tabs of the Supplementary Data S1 spreadsheet.

Meta-analysis. Once differentially methylated windows were identified, an adaptively weighted (AW) meta-analysis (16) was conducted to identify differentially methylated windows that persisted over time. The contribution of each window was marked with an AW weight (1, contribute; 0, not contribute). Persistently altered windows were defined as having weight (1, 1) with FDR < 0.05, and effect size > 10% methylation difference for both time points. We again filtered out any significant windows driven by outliers using a trimmed mean difference < 0.10. Two samples were outliers on the PCA and were therefore not included in the meta-analysis (late, P7 and P8). Windows were then annotated to genes using “Rogenome Mmuculus UCSC mm9” from bioconductor. The windows were classified as CpG island or CpG shore on the basis of CpG category database by Wu and colleagues (17). These windows were also annotated as promoter [transcription start site (TSS) ± 1,000 bp], exon, intron, and intergenic on the basis of generic parts category using TxDb.Mmuculus.UCSC.mm9.knownGene (R database) and RRange (R function) R package (18). These data are presented in the second two tabs of the Supplementary Data S1 spreadsheet.

Bisulfite sequencing of candidate genes
Genomic DNA was bisulfite-converted (Zymo EZ DNA Gold) and PCR-amplified using primers designed by Meth Primer
PCR products were cloned into pCR 4.0 by a TOPO reaction and transformed into TOP10 competent cells (Invitrogen). After transformation, cells were spread on ampicillin containing agar plates and 10 to 20 colonies were chosen for analysis. Colonies grew up overnight in ampicillin containing LB broth and DNA was isolated by Qiagen miniprep protocol. One microgram of DNA was submitted for Sanger Sequencing at The University of Pittsburgh Genomics and Proteomics Core Facility (Pittsburgh, PA). Sequences were processed using Sequencher 5.2.4 (Gene Codes) and BiQ Analyzer (Max Planck Institute for Informatics and Saarland University, Saarbrucken, Germany; ref. 19). The \( b \) value was calculated as methylated CpG/total CpGs, and data were analyzed via logistic regression in R (glm function). The gene location of the \( \text{Igf1r} \) hypermethylation was drawn with FancyGene (20) using NCBI reference sequence number NM_010513.

Quantitative real-time PCR

Frozen mammary glands were cryohomogenized and RNA was isolated. Reverse transcription was conducted using iScript and qPCR was run on a Bio-Rad cfx 384 as previously described (21), using primers in Supplementary Table S1. Data was analyzed via Student t test using Prism 5, v5.40 (GraphPad).

Data availability

The fastq files were submitted to SRA (SRP053781), which is linked to BioProject (accession no.: PRJNA273963) and BioSample (accession nos.: SAMN03324064 to SAMN03324087).

Results

Experimental design

To discover parity-induced alterations in DNA methylation of the murine mammary gland, we bred FVB/NJ mice at approximately 55 days of age and euthanized them at two time points. One group (early) was euthanized immediately after involution to identify changes directly induced by pregnancy, lactation, and involution (\( n = 8 \)). The second group (late) was euthanized 6.5 months postinvolution to identify parity-induced changes that are persistent and remain long after involution. Age-matched virgins were included at both time points (Supplementary Fig. S1A). In these studies, whole mammary glands were used. To understand

![Figure 1](https://example.com/fig1.png)

Partly-induced DNA methylation in the mouse mammary gland. A, pie charts depicting areas of the genome that are differentially methylated at the early and late time points as well as the distribution of library probes in the SureSelectXT Methyl-Seq assay. Top, the percentage of CpG islands and shores, whereas bottom shows the genomic region. B, heatmap of differentially methylated windows after hierarchical clustering in mammary glands from parous mice and age-matched nulliparous mice is depicted at each time point. Windows that had at least a coverage of 10, \( q < 0.01 \), and a difference in methylation (between nulliparous and parous) of at least 25%. We additionally filtered out windows with an absolute trimmed mean difference of \(<-0.25\). Methylation is represented as \( b \) value with green as hypomethylated and red as hypermethylated. At the early time point, 4,385 windows were differentially methylated (3,507 hypomethylated and 878 hypermethylated). At the late time point, 8,884 windows were differentially methylated (2,323 hypomethylated and 6,561 hypermethylated).
the distribution of cell types present in the gland, we conducted quantitative RT-PCR for transcripts specific for particular cell types [1, leukocyte (CD45); 2, adipose (ADIPO); 3, luminal epithelial (ESR1 and CK18); 4, myoepithelial (SMA and CK14); and 5, fibroblast (FSP)]. While some minor changes were detected, none were statistically significant, suggesting that no major changes in cell composition remain 6.5 months post-partum (Supplementary Fig. S1B). In the case of CK14, the parous samples trended higher, but the two other stromal markers were not changed, and neither epithelial marker was changed. Immediately postinvolution, adiponectin tended to be lower in the parous samples, which was expected but was not significant. We are therefore confident that the cellular components of the mammary glands are not remarkably different comparing nulliparous and parous mice.

Targeted bisulfite sequencing of DMRs
gDNA isolated from mammary tissues was used to conduct a targeted analysis of DMRs across the genome. DNA methylation was assessed using SureSelectXT Mouse Methyl-Seq Technology. On average, 39 million pairs of reads per sample were aligned to the reference with a mapping efficiency of 50% to 68.1% and an average 71x sequencing depth was achieved (the coverage depth was 113x before alignment; Supplementary Fig. S3). The sequencing results were processed as described in the Materials and Methods, and quality control was conducted by Pearson correlation and PCA (Supplementary Fig. S3).

Differentially methylated genes in early and late parous mammary glands
To identify differentially methylated genes, logistic regression was used to determine significantly hyper/hypomethylated windows, which were then annotated to genes. In the early time point, comparing parous and nulliparous mammary glands, 4,385 windows representing 1,880 genes were identified as differentially methylated, with 3,507 windows (1,535 genes) hypomethylated, and 878 windows (446 genes) hypermethylated (Supplementary Data S1). Additionally, 101 genes were both hyper and hypomethylated. Of the 4,385 differentially methylated windows, 2% were located in CpG islands (CpGi) and 17% in CpGi shores as annotated by UCSC genome browser (Fig. 1A). For those windows in genomic regions, 9% were located in promoter regions, 9% in exons, 48% in introns, and 35% in intergenic regions (Fig. 1A). At

Table 1. Mammary-specific differentially methylated genes

<table>
<thead>
<tr>
<th>Genes</th>
<th>SureSelectXT coverage</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOC100271722</td>
<td>False</td>
<td>Zhang et al., 2013</td>
</tr>
<tr>
<td>LOC150381</td>
<td>False</td>
<td>Zhang et al., 2013</td>
</tr>
<tr>
<td>MIRLET7</td>
<td>True</td>
<td>Zhang et al., 2013</td>
</tr>
<tr>
<td>MIRLET7AS</td>
<td>False</td>
<td>Zhang et al., 2013</td>
</tr>
<tr>
<td>MIRLET7B</td>
<td>False</td>
<td>Zhang et al., 2013</td>
</tr>
<tr>
<td>MIRLET7BGH</td>
<td>False</td>
<td>Zhang et al., 2013</td>
</tr>
<tr>
<td>ALX4</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>FEV</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>HLF</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>HOXA11</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>LYL1</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>NEUROG</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>PAX9</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>GATA5</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>MGD</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>SOX10</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>SREBF1</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>ST18</td>
<td>True</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>TP73</td>
<td>False</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>TRIM20</td>
<td>False</td>
<td>Avraham et al., 2014</td>
</tr>
<tr>
<td>ZNF436</td>
<td>False</td>
<td>Avraham et al., 2014</td>
</tr>
</tbody>
</table>

Figure 2.
Confirmation of Methyl-Seq DNA methylation by bisulfite sequencing. A, β values for the top differentially methylated genes with the highest coverage in the Methyl-Seq assay are depicted. B, DNA methylation after bisulfite sequencing is represented as β values. Logistic regression was used to compare nulliparous and parous samples for each gene (***, P < 0.001; **, P < 0.01).
the late time point, a total of 8,884 windows (2,888 genes), including 6,561 windows (2,260 genes), were hypermethylated and 2,323 windows (872 genes) were hypomethylated. Two hundred forty-four genes were both hyper- and hypomethylated. The distribution of the methylated sites in CpG islands and shores was identical in both early and late samples (Fig. 1A). In the SureSelectXT Mouse Methyl-Seq library, 7% of the probes were located in CpG islands and 11% in shores, whereas we detected only 2% in CpG islands and 16% in shores. In addition, 35% of the library probes are located in introns, 12% in exons, 8% in promoters, and 45% in intergenic regions. Our data are substantially enriched for intronic regions and CpG shores, with fewer differentially methylated windows in intergenic regions and CpG islands. This difference is highly significant by χ² analysis (genomic region: early, P = 4.75E−51; late, P = 4.91E−148; CpG/shores: early, P = 1.19E−63; late, P = 3.15E−124). It is possible that DNA methylation in introns and CpG shores plays a greater role than previously thought.

Recently, mammally-specific differentially methylated regions have been identified (15, 22). Twenty-one breast-specific differentially methylated genes were identified in these 2 publications and 13 are represented in the SureSelectXT Methyl-Seq assay. Although these were identified in human samples, we found 2 genes to be significantly differentially methylated with parity in our mouse study, Mgmt and Hlf. Mgmt is hypermethylated with parity at the late time point and hypomethylated at the early time point, and Hlf is hypomethylated at the late time point (Table 1). The significantly differentially methylated windows are presented by heatmap (Fig. 1B). Windows shown have a significance of at least q < 0.01, and a methylation difference between nulliparous and parous greater than 25%.

Validation of the targeted methyl sequencing of DMRs
To first validate our assay using an orthogonal approach, we used bisulfite sequencing on the same DNA that was used in the Methyl-Seq assay. Genes with the largest difference in β value and the highest coverage were chosen for validation, specifically Ska91, Psip2, St6gal1, Irf2, Ece1, and Lck (Supplementary Data S1 and Fig. 2A). Psip2 and St6gal1 were confirmed as hypomethylated, and Ska91 and Irf2 were confirmed as hypermethylated by bisulfite Sanger sequencing, whereas a nonsignificant trend was noted for Ece1 and Lck (Fig. 2B and Supplementary Fig. S4). As an additional control, we further interrogated the Methyl-Seq data by investigating regions of the genome known to be maternally imprinted (Grb10, Cdken1c, Igf1r/Airn, Hor2, Xist, and H19/Igf1) and found that as expected, approximately 50% of the CpGs were methylated in each of these regions (Supplementary Fig. S5).

Persistent DNA methylation changes in the parous mammary gland
To determine which genes were persistently differentially methylated in the parous mammary gland 6 months postinvolution (late), we performed meta-analysis using 493,473 windows from the early time point and 473,714 windows from the late time point (432,946 windows overlapped). This analysis identified 624 genes (660 windows) and 322 genes (330 windows) that were hyper- and hypomethylated in the early group and remained altered in the late group. After filtering by trimmed means > 0.10, 276 windows (269 genes) were hypermethylated, and 131 windows (128 genes) were hypomethylated (Fig. 3 and Supplementary Data S1). A supervised hierarchical clustering analysis was also performed and is displayed in Supplementary Fig. S6. Ingenuity Pathway Analysis (INGenuity.com) on the persistently hypermethylated and hypomethylated, and both sets of altered genes simultaneously showed several signaling pathways to be altered, including P70S6K, Rho GTPases, protein kinase A, and others. The full data set from the pathway analysis is presented in the Supplementary Data S1 spreadsheet. Among the top persistently hypermethylated genes identified in the Methyl-Seq assay (Table 2), which was also central to one of the top signaling networks in the pathway analysis (Supplementary Fig. S7) was Igf1r.

Igf1r is persistently methylated in the parous mammary gland
Genes that are persistently altered at both the early and late time points are of great biologic significance and serve to address our
hypothosis that long-term epigenetic alterations underlie the protective effect of pregnancy. We have previously reported that parity results in a significant decrease in circulating growth hormone (GH) in rats that altered GH/IGF signaling in the mammary gland and others have shown alterations on the GH/IGF axis in humans (23, 24). Interestingly, in the current study, we found that Igf1r was seventh among the top hypermethylated genes after meta-analysis, indicating it is methylated in the mammary gland post-partum and remains methylated long after involution (Table 2). Because the GH/IGF pathway has previously been implicated in the protective effect of pregnancy and it is among the top persistently hypermethylated genes we further investigated this family. This hypermethylation occurs in the largest intron of the Igf1r gene between exons 2 and 3 (Fig. 4A). This intron also contains other epigenetic modifications, including histone 3 lysine 4 mono and trimethylation, as displayed in the UCSC Genome Browser, indicating that this region is likely involved in gene expression regulation (Supplementary Fig. S8A).

We confirmed these findings by bisulfite sequencing and found a significant increase in Igf1r methylation at the same above-mentioned intron in parous animals 6 months postinvolutions (Fig. 4B and C and Supplementary Fig. S8B). This increase in intron methylation was associated with a decrease in Igf1r mRNA expression in the mammary gland of parous animals (Fig. 4D). We interrogated other IGF pathway genes in the Methyl-Seq data analysis and found several to have regions that displayed significantly altered DNA methylation in the Methyl-Seq analysis (Table 3). The 5 genes with the highest ratio (number of significantly altered windows/the total number of windows), Irs1, Igf1, Igfbp4, Pthr, and Stat5b, all display increased methylation with parity in the Methyl-Seq assay (Fig. 5A), and almost all showed a trend toward decreased mRNA expression in the parous mammary gland (Fig. 5B). The differentially methylated windows in the top 3 IGF pathway genes, Irs1, Igf1, and Igfbp4, overlap with other epigenetic modifications, indicating that they are likely involved in gene regulation (Supplementary Fig. S9).

Discussion

The most significant modifiable factor affecting a woman’s risk for developing breast cancer is an early age at FFB. Although the protective effect of an early FFB was identified decades ago, the mechanisms underlying this effect remain to be elucidated. Age at FFB is increasing in women in the United States and likely represents a contributing factor to the increase in breast cancer incidence (25). Understanding how an early FFB prevents breast cancer may identify promising new avenues to develop novel breast cancer prevention strategies. In this study, we have characterized epigenetic changes following pregnancy and found that Igf1r and other IGF family members are hypermethylated and downregulated. This study provides support to the concept of inhibiting the GH/IGF pathway for breast cancer prevention.

Epigenetic changes are known to play integral roles in the mammary gland (10). Because epigenetic modifications are both inducible and stable, they can persistently alter gene expression in a long-lasting "memory-like" fashion, likely playing a role in the protective effect of pregnancy against breast cancer, as speculated by Jerry and colleagues (26). Chosh et al. investigated DNA methylation in breast tissue of 30 women (19 parous, 16 nulliparous; ref. 27). Using a genome-wide methyl-binding protein pull-down to isolate all methylated DNA (a nontargeted approach), their analysis identified FOXA1 as hypermethylated and possibly silenced with parity (27). They speculated that this could attenuate the effects of estrogen receptor (ERα) leading to a reduced susceptibility to breast cancer. They also found an IGF family member altered in their analysis. The IGF acid labile subunit (IGFALS) was hypomethylated in their analysis. The IGFAI is responsible for carrying IGF1 in the circulation, and altering the levels of IGFAI may reduce IGF1 activity in target tissues. In addition, histone modifications are likely important. Trimethylation of lysine 27 on histone 3 (H3K27me3) was identified in a genome-wide histone methylation screen to increase during pregnancy in rodents (28). Our *in silico* data show that histone modifications, including but not limited to H3K27me3...
and H3K4, overlap with parity-induced DNA methylation changes in IGF pathway genes (Supplementary Fig. S9).

Parity-induced genome-wide gene expression changes have been investigated in both rodents and women. The first study to characterize changes in expression across the genome identified a gene signature indicative of parity in multiple rat strains and mice (29). In women, a genome-wide change in expression is associated with parity and parous women who never develop breast cancer display a different signature than parous women who do develop breast cancer (30), and an 11-gene signature can identify parity in women (31). Of note, in our data, at the early time point, the vast majority of differentially methylated genes was hypomethylated, whereas 6 months after involution, there is a dramatic shift to hypermethylation. This hypomethylation at the early time point could signify an immense induction of expression of genes involved in the remodeling process of involution which is multifaceted and would likely produce a complex gene signature.

Our unbiased screen identified Igf1r as one of the most hypermethylated and persistently epigenetically altered genes in parous mammary gland. The Igf1r and other members of the GH/IGF pathway are critical for mammary development and also play a role in pregnancy, lactation, and involution (32–35). The Igf1r has been shown to be important during lactation in the mouse mammary gland as overexpressing a dominant/negative IGF1R reduced alveolar development and milk protein production (36). Consistent with hypermethylation of the Igf1r intron, we found that Igf1r mRNA was decreased in parous mammary gland. Several other IGF family members also showed epigenetic alterations and gene expression changes. Other genome-wide expression studies have identified components of the GH/IGF pathway to be downregulated in the parous mammary gland (29, 37). IGF1 was found to be significantly decreased in parous mammary gland, whereas IGFBP5 was increased (29). IGF1 was also found to be decreased in the breast of parous women (37). In addition to mammary-specific changes, we have previously demonstrated that circulating GH is decreased in parous animals leading to a reduction in protein activation of key IGF/GH pathway signaling molecules in the mammary gland (23). These data, in conjunction with previous gene expression studies, and our data reported here strongly support the concept that downregulation of the GH/IGF axis (both endocrine and intrinsic to the mammary gland) may be involved in the protective effect of pregnancy against breast cancer.

Epigenetic modifications have been linked to breast cancer risk in multiple studies. At least three dietary compounds (folate, choline, and sulforaphane) can alter the epigenome and reduce cancer risk (38–44). Folate, a water-soluble vitamin, had been shown to alter DNA methylation by modulating both DNA methyltransferases and methyl-binding proteins (38). In women, folate or choline (a methyl donor food) intake is inversely correlated with breast cancer risk in multiple studies (39, 42, 43, 45). Sulforaphane is a modulator of histone acetylation and inhibits growth in multiple breast cancer cells and also reduces susceptibility to 7,12-dimethylbenz(a)anthracene (DMBA)-induced tumorigenesis in animals (40, 41, 44). If pregnancy is indeed modulating disease risk through epigenetic mechanisms that reduce tumorigenic pathways such as the IGE/GH pathway, it might be possible that modulating these epigenetic events in young women by dietary or therapeutic interventions will ultimately reduce risk of breast cancer.

The current landscape of evidence-based prevention therapies include inhibitors of estrogen regulation, namely, tamoxifen, raloxifene, anastrozole, and exemestane (46–49). Interestingly, parity has been associated with a reduction in ER/progesterone receptor (PR)þ breast cancers indicating a role for hormonal preventative therapies (50, 51). Unfortunately, while these therapies are effective, they also have side effects and thus these prevention therapies are only recommended to women of high risk. One major avenue of investigation for chemoprevention, currently under clinical investigation in high-risk women, is inhibition of the IGF pathway (53, 54). For example, a 10-day treatment of a somatostatin analog (SOM230, Signifor, Novartis) decreased proliferation and increased apoptosis in ductal carcinoma in situ. The treatment also reduced IGF1R, pAkt, and pERK (55).
Our study used a novel DNA methylation screen by conducting next-generation sequencing on bisulfite-converted DNA, which has been enriched for areas of the genome expected to be differentially methylated. This targeted approach enhanced our analysis by providing more than 100× coverage of specific regions important in DNA methylation. Although highly novel and

Table 3. IGF pathway members also display altered DNA methylation

<table>
<thead>
<tr>
<th>Gene</th>
<th>Total windows</th>
<th>Significant windows</th>
<th>Ratio</th>
<th>Hyper</th>
<th>Hypo</th>
<th>Total windows</th>
<th>Significant windows</th>
<th>Ratio</th>
<th>Hyper</th>
<th>Hypo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irs1</td>
<td>36</td>
<td>2</td>
<td>0.0556</td>
<td>0</td>
<td>2</td>
<td>35</td>
<td>19</td>
<td>0.5429</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>lgr1</td>
<td>26</td>
<td>3</td>
<td>0.1154</td>
<td>0</td>
<td>3</td>
<td>25</td>
<td>13</td>
<td>0.5200</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Igpbp4</td>
<td>28</td>
<td>5</td>
<td>0.1786</td>
<td>1</td>
<td>4</td>
<td>20</td>
<td>9</td>
<td>0.4500</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Stat5b</td>
<td>81</td>
<td>7</td>
<td>0.0864</td>
<td>3</td>
<td>4</td>
<td>85</td>
<td>36</td>
<td>0.4255</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Pthr</td>
<td>29</td>
<td>3</td>
<td>0.1034</td>
<td>0</td>
<td>3</td>
<td>29</td>
<td>12</td>
<td>0.4138</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Shc</td>
<td>21</td>
<td>1</td>
<td>0.0476</td>
<td>0</td>
<td>1</td>
<td>21</td>
<td>8</td>
<td>0.3810</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Igpbp3</td>
<td>11</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>0.3750</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Igpzr</td>
<td>32</td>
<td>4</td>
<td>0.1250</td>
<td>1</td>
<td>3</td>
<td>28</td>
<td>10</td>
<td>0.3571</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Igpbp6</td>
<td>3</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0.3333</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Ghr</td>
<td>45</td>
<td>5</td>
<td>0.1111</td>
<td>0</td>
<td>5</td>
<td>45</td>
<td>14</td>
<td>0.3111</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>Akt1</td>
<td>22</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>6</td>
<td>0.3000</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Akt2</td>
<td>29</td>
<td>2</td>
<td>0.0690</td>
<td>0</td>
<td>2</td>
<td>29</td>
<td>8</td>
<td>0.2759</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Irs2</td>
<td>46</td>
<td>10</td>
<td>0.2174</td>
<td>0</td>
<td>10</td>
<td>45</td>
<td>12</td>
<td>0.2667</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Sos</td>
<td>27</td>
<td>1</td>
<td>0.0370</td>
<td>1</td>
<td>0</td>
<td>34</td>
<td>9</td>
<td>0.2647</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Stat5a</td>
<td>16</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>6</td>
<td>0.2609</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Src</td>
<td>50</td>
<td>1</td>
<td>0.0200</td>
<td>1</td>
<td>0</td>
<td>48</td>
<td>10</td>
<td>0.2083</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Stat3</td>
<td>51</td>
<td>6</td>
<td>0.1766</td>
<td>3</td>
<td>3</td>
<td>54</td>
<td>10</td>
<td>0.1852</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Erk</td>
<td>12</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>2</td>
<td>0.1538</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Igpzr</td>
<td>14</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>1</td>
<td>0.0833</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Irs4</td>
<td>19</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>1</td>
<td>0.0500</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Igpbp5</td>
<td>3</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PDK</td>
<td>2</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Igpbp1</td>
<td>2</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0.0000</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

NOTE: The total number of 120-bp windows available for each gene was compared with the number of windows that were significantly differentially methylated with parity to obtain the ratio.

Figure 5. IGF pathway members are hypermethylated with parity. A, Methyl-Seq analysis of IGF pathway members at the late time point. Logistic regression was implemented to compare nulliparous and parous. B, mRNA expression of the IGF pathway members analyzed by quantitative RT-PCR.
innovative, there were some limitations in our design. For example, the mouse estrous cycles were not synchronized, and therefore the mice may have been in different stages of the cycle at euthanasia. This is unlikely to skew our results, as DNA methylation is relatively stable, and the estrus stage of the cycle, when the greatest changes would likely occur, is relatively short; therefore, the majority of glands were not likely harvested during estrus. In addition, we analyzed the entire mammary gland, including stroma and epithelium; therefore, it is unclear in which compartments these epigenetic changes occur.

In conclusion, we have shown that members of the GH/IGF pathway undergo DNA methylation in response to pregnancy, which may contribute to a reduction in their expression. This pathway has been implicated in breast cancer and also been shown to be reduced with parity in multiple previous studies. These data together indicate that the GH/IGF pathway maybe critical in pregnancy protection from breast cancer, and prevention strategies targeting this pathway should undergo further research.

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

Authors’ Contributions

Conception and design: T.A. Katz, V.J. Palmieri, S. Oesterreich, A.V. Lee


Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): T.A. Katz, V.J. Palmieri, R.K. Dearth, T.N. Pathiraja, P. Shaw, S. Small, D.G. Peters, A.V. Lee


Writing, review, and/or revision of the manuscript: T.A. Katz, S.G. Liao, R.K. Dearth, Z. Huo, N.E. Davidson, G.C. Tseng, S. Oesterreich, A.V. Lee

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): T.A. Katz, S.G. Liao, V.J. Palmieri, Z. Huo, P. Shaw

Study supervision: T.A. Katz, S. Oesterreich, A.V. Lee

Disclaimer
The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the University of Pittsburgh.

Acknowledgments
The authors thank Drs. Kevin Poon and Alex Siebold with Agilent Technologies for their outstanding support and early access to the SureSelectXT Mouse Methyl-Seq kit, as well as Ora L. Britton for mouse husbandry and experimentation.

Grant Support
This project used the UPCI Tissue and Research Pathology Services shared facilities that are supported in part by award P30CA047904. This work was supported in part by the National Cancer Institute of the NIH under award number R01CA94118 [A. V. Lee] and T32CA186873 (T.A. Katz). A.V. Lee is a recipient of a Scientific Advisory Council award from Susan G. Komen for the Cure and a Hillman Foundation Fellow. The authors acknowledge support from the University of Pittsburgh Cancer Institute, UPAC, and by the University of Pittsburgh Center for Simulation and Modeling through the high performance computing resources provided.

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 May 1, 2015; revised July 21, 2015; accepted August 5, 2015; published OnlineFirst August 19, 2015.

References
5. Katz et al.
Targeted DNA Methylation Screen in the Mouse Mammary Genome Reveals a Parity-Induced Hypermethylation of \textit{Igf1r} That Persists Long after Parturition


Access the most recent version of this article at: doi:10.1158/1940-6207.CAPR-15-0178

Access the most recent supplemental material at:
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2015/08/19/1940-6207.CAPR-15-0178.DC1
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2015/09/19/1940-6207.CAPR-15-0178.DC2

This article cites 54 articles, 23 of which you can access for free at:
http://cancerpreventionresearch.aacrjournals.org/content/8/10/1000.full.html#ref-list-1

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

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

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