Whole-exome sequencing combined with functional genomics reveals novel candidate driver cancer genes in endometrial cancer

Han Liang,1,9,11 Lydia W.T. Cheung,2,9 Jie Li,2 Zhenlin Ju,1 Shuangxing Yu,2 Katherine Stemke-Hale,2 Turgut Dogruluk,3 Yiling Lu,2 Xiuping Liu,4 Chao Gu,2 Wei Guo,2 Steven E. Scherer,5 Hannah Carter,6 Shannon N. Westin,7 Mary D. Dyer,2 Roeland G.W. Verhaak,1 Fan Zhang,2 Rachel Karchin,6 Chang-Gong Liu,4 Karen H. Lu,7 Russell R. Broaddus,8 Kenneth L. Scott,3 Bryan T. Hennessy,2,10 and Gordon B. Mills2

1Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA; 2Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA; 3Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, USA; 4Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA; 5Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas 77030, USA; 6Department of Biomedical Engineering and Institute for Computational Medicine, Johns Hopkins University, Baltimore, Maryland 21205, USA; 7Department of Gynecologic Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA; 8Department of Biomedical Engineering and Institute for Computational Medicine, Johns Hopkins University, Baltimore, Maryland 21205, USA; 9Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA

Endometrial cancer is the most common gynecological malignancy, with more than 280,000 cases occurring annually worldwide. Although previous studies have identified important common somatic mutations in endometrial cancer, they have primarily focused on a small set of known cancer genes and have thus provided a limited view of the molecular basis underlying this disease. Here we have developed an integrated systems-biology approach to identifying novel cancer genes contributing to endometrial tumorigenesis. We first performed whole-exome sequencing on 13 endometrial cancers and matched normal samples, systematically identifying somatic alterations with high precision and sensitivity. We then combined bioinformatics prioritization with high-throughput screening (including both shRNA-mediated knockdown and expression of wild-type and mutant constructs) in a highly sensitive cell viability assay. Our results revealed 12 potential driver cancer genes including 10 tumor-suppressor candidates (ARID1A, INHBA, KMO, TTLL5, GRM8, AKTIP, PHKA2, TRPS1, and WNT11) and two oncogene candidates (ERBB3 and RPS6KC1). The results in the “sensor” cell line were recapitulated by siRNA-mediated knockdown in endometrial cancer cell lines. Focusing on ARID1A, we integrated mutation profiles with functional proteomics in 222 endometrial cancer samples, demonstrating that ARID1A mutations frequently co-occur with mutations in the phosphatidylinositol 3-kinase (PI3K) pathway and are associated with PI3K pathway activation. siRNA knockdown in endometrial cancer cell lines increased AKT phosphorylation supporting ARID1A as a novel regulator of PI3K pathway activity. Our study presents the first unbiased view of somatic coding mutations in endometrial cancer and provides functional evidence for diverse driver genes and mutations in this disease.

[Supplemental material is available for this article.]

Endometrial cancer is the most common gynecological malignancy and the fourth most common cancer among women in Western countries, with more than 280,000 cases occurring annually worldwide (Ferlay et al. 2010). Although the prognosis is favorable for patients identified with early-stage disease, the outcomes for patients with metastatic or recurrent endometrial cancer remain abysmal, with median survival of <1 yr. For effective cancer prevention and treatment, it is necessary to identify critical genetic changes that initiate endometrial cancer and contribute to its progression. Over the last decade, multiple groups have identified common somatic mutations of several important genes in endometrial cancer (Dedes et al. 2011). For example, PIK3CA mutations occur in ~30% of endometrial cancers, with the mutation frequency varying according to histological grade (Oda et al. 2005; Velasco et al. 2006; Catasus et al. 2008; Miyake et al. 2008; Hayes et al. 2009; Cheung et al. 2011). PTEN, a major negative regulator of the phosphatidylinositol 3-kinase (PI3K) pathway, is frequently mutated in endometrial tumors (Risinger et al. 1998; Mutter et al. 2000; Sun et al. 2000; Salvesen et al. 2004). Other cancer genes such as TP53, KRAS, AKT, and CTNNB1 are also frequently mutated in this disease (Lax et al. 2000; Schlosshauer et al. 2000; Moreno-Bueno et al. 2002; Stefansson et al. 2004; Byron et al. 2008; Jia et al. 2008; Catasus et al. 2009; Shoji et al. 2009). More recently, we have demonstrated that PI3K pathway aberrations occur in >80% of endometrial cancers, often with co-mutations of multiple pathway
Novel driver cancer genes in endometrial cancer

Results

Landscape of somatic mutations in the exomes of endometrial cancer

To gain an unbiased view of somatic alterations that contribute to the pathogenesis of endometrial cancer, we performed whole-exome sequencing of 14 endometrial tumor samples (one tumor with an unusually high mutation frequency and low estimated tumor content was excluded from further analysis) and matched normal DNA samples using the SOLiD platform. Supplemental Table 1 provides the histologic information and other clinical characteristics of these samples. On average, we obtained 185 million 50-nt single reads for tumor samples and 62 million reads for normal samples (Table 1). The Agilent SureSelect capture reagent used for enrichment targets 170,843 exons (~1.2% of the human genome); for the targeted regions, the average coverage was 94x and 42x for tumor and normal samples, respectively. We developed a computational pipeline to detect somatic mutations by comparing single-nucleotide variations in tumor and normal samples (Supplemental Figs. 1, 2; Supplemental Methods). At the depth of sequencing used and with the majority of tumors being low grade (grade 2), we identified a mean of 89.5 somatic point mutations (range 34–263) per tumor corresponding to an average of 3.7 mutations per megabase (range 0.98–12.0) (Table 1). As shown in Figure 2, the most common mutations were transitions in the CpG context (that is, G→A/C→T), which parallels several other cancer lineages (Ding et al. 2010; Chapman et al. 2011; Puente et al. 2011). In addition, the tumors contained a mean of 14.2 somatic coding small insertion/deletions (indels) (Table 1; Supplemental Table 2; Supplemental Methods).

Focusing on somatic mutations that potentially affect protein function, across the 13 tumor samples, we identified 576 unique nonsynonymous mutations and 24 nonsense mutations in 566 genes (Supplemental Table 2). The most perturbed biological pathways were integrin, angiopoietin, complement system, and PTEN signaling (Fisher exact test, \( P < 2.8 \times 10^{-3} \), false discovery rate [FDR] < 0.05, Supplemental Table 3). Based on potential biological interest, we selected 97 mutation sites for Sequenom MassARRAY validation and obtained a true-positive rate of 81%, leading to 69 genes with validated mutations and a prediction of 487 nonsynonymous mutations and 18 nonsense mutations in the 13 tumors (Supplemental Table 2; Supplemental Methods). To estimate the sensitivity of our mutation calling algorithm, we examined full-length sequences of nine genes with concurrent Sanger sequencing and Sequenom MassARRAY detection. The estimated sensitivity was 80%, with false negatives occurring primarily due to low coverage (Supplemental Methods). These results indicate that the identified mutated gene set largely characterizes the landscape of somatic coding mutations in the 13 endometrial cancers assessed.

Novel candidate driver cancer genes revealed by Ba/F3 functional assays

Several of the validated mutated genes have previously been shown to be targeted in endometrial cancers, including TP53, PTEN, CTNNB1, PIK3CA, PIK3R1, and PIK3R2 (Lax et al. 2000; Sun et al. 2001; Moreno-Bueno et al. 2002; Oda et al. 2005; Cheung et al. 2011). We recently reported their mutation profiles in a large collection of well-characterized endometrial tumors, demonstrating that aberrations, including frequent co-mutations, of the PI3K pathway occur in ~80% of endometrial cancers (Cheung et al.

Figure 1. Overview of our systems-biology approach to identifying cancer driver genes in endometrial cancer.
Table 1. Landscape of somatic mutations in the exomes of 13 endometrial tumors

| Patient ID | Tumor bases sequenced | Normal bases sequenced | Tumor exome coverage | Normal exome coverage | Callable positions | All point mutations | Nonsilent mutations | Mutation rate (per Mb) | Coding indels |
|------------|-----------------------|-----------------------|----------------------|-----------------------|-------------------|---------------------|---------------------|---------------------|--------------|
| E27        | $9.75 \times 10^8$    | $4.90 \times 10^8$    | $93.0\times$         | $55.0\times$          | $30.9 \times 10^6$ | 91                  | 49                  | 2.95                | 15           |
| E28        | $8.84 \times 10^8$    | $4.18 \times 10^8$    | $86.3\times$         | $35.8\times$          | $25.1 \times 10^6$ | 75                  | 48                  | 2.99                | 7            |
| E35        | $8.75 \times 10^8$    | $4.60 \times 10^8$    | $92.0\times$         | $47.4\times$          | $29.0 \times 10^6$ | 35                  | 19                  | 1.21                | 10           |
| E58        | $9.17 \times 10^8$    | $4.69 \times 10^8$    | $94.1\times$         | $41.7\times$          | $29.1 \times 10^6$ | 44                  | 24                  | 1.51                | 6            |
| E62        | $9.55 \times 10^8$    | $2.72 \times 10^8$    | $99.6\times$         | $18.7\times$          | $21.9 \times 10^6$ | 263                 | 146                 | 12.0                | 52           |
| E70        | $9.76 \times 10^8$    | $4.35 \times 10^8$    | $88.3\times$         | $45.2\times$          | $20.7 \times 10^6$ | 79                  | 43                  | 3.82                | 8            |
| E82        | $9.22 \times 10^8$    | $4.01 \times 10^8$    | $87.1\times$         | $44.8\times$          | $31.2 \times 10^6$ | 30                  | 13                  | 0.96                | 14           |
| E99        | $8.84 \times 10^8$    | $4.74 \times 10^8$    | $89.2\times$         | $48.5\times$          | $29.9 \times 10^6$ | 36                  | 17                  | 1.20                | 10           |
| E101       | $9.13 \times 10^8$    | $4.60 \times 10^8$    | $95.7\times$         | $52.8\times$          | $31.3 \times 10^6$ | 34                  | 18                  | 1.09                | 14           |
| E114       | $8.97 \times 10^8$    | $4.24 \times 10^8$    | $99.6\times$         | $39.0\times$          | $25.9 \times 10^6$ | 66                  | 43                  | 2.54                | 6            |
| E161       | $9.21 \times 10^8$    | $4.22 \times 10^8$    | $91.6\times$         | $43.3\times$          | $18.7 \times 10^6$ | 182                 | 81                  | 9.74                | 10           |
| E170       | $9.10 \times 10^8$    | $4.09 \times 10^8$    | $106\times$          | $39.6\times$          | $28.1 \times 10^6$ | 188                 | 98                  | 6.69                | 24           |
| E172       | $9.67 \times 10^8$    | $3.53 \times 10^8$    | $99.5\times$         | $30.2\times$          | $28.3 \times 10^6$ | 41                  | 26                  | 1.45                | 9            |

The mutation rate was calculated by dividing the total number of somatic mutations by the total number of callable nucleotide positions ($\geq 15 \times$ in tumor and $\geq 8 \times$ in matched normal samples).

2011). However, the functional roles of the remaining mutated genes in endometrial cancer are largely unknown. To identify potential driver cancer genes for further characterization, we performed bioinformatics analyses and selected 30 genes using the following criteria: (1) if the mutation in a gene is predicted to have a high impact on the protein function by CHASM (FDR of impact score $< 30\%$) (Carter et al. 2010) or MutationAssessor (high impact) (Reva et al. 2011); or (2) a gene contains multiple or recurrent mutations (Methods). Supplemental Table 4 provides details on selection of candidates for further analysis. We then evaluated the functional impact of these selected genes in murine Ba/F3 cells (Warmuth et al. 2007). This cell line is a murine bone marrow-derived pro-B-cell line that becomes IL3-independent in the presence of an oncogene or oncogenic event. Thus, the Ba/F3 cells represent a sensitive tool for measuring the effect of an introduced perturbation on cell proliferation and survival, not only for kinase genes but also non-kinase genes (Lutz et al. 2000; Warmuth et al. 2007; Yoda et al. 2010; Cheung et al. 2011).

We first performed a “loss-of-function” analysis on the 30 candidate genes using shRNA gene silencing. As shown in Figure 3A, compared with parental cells and cells transfected with empty vector or nonspecific shRNAs, inhibiting the expression of eight genes ($\text{NRHBA, KMO, PHKA2, TTL5, AKT1, IGFBP3, GRM8, and ARID1A}$) with two independent shRNAs was sufficient to promote the IL3-independent survival of Ba/F3, compatible with these genes having tumor-suppressor-like activity. In contrast, the introduction of $\text{ERBB3}$ and $\text{RPS6KC1}$ shRNAs remarkably inhibited IL3-independent cell survival, compatible with oncogene-like activity. These observations were confirmed by at least one of two additional independent shRNAs in a secondary screen (data not shown). We confirmed shRNA knockdown efficacy by Western blotting for eight proteins where appropriate antibodies were available (Fig. 3C). Nine genes with efficient shRNA knockdown had no effect in the Ba/F3 cells with four ($\text{NOS1AP, CDK17, C1orf198, and PHKG1}$) shown in Figure 3A and Supplemental Figure 3a. One of two shRNAs in the initial screen for two genes ($\text{TRPS1 and WNT11}$) promoted IL3-independent Ba/F3 survival. The activity was confirmed by two additional independent shRNAs (Fig. 3B; Supplemental Fig. 3b). Significant knockdown was not observed or antibodies were unavailable for the remaining nine genes. Taken together, the shRNA-mediated knockdown screen revealed 12 potential driver cancer genes including 10 tumor suppressors and two oncogene candidates from the 30 genes assessed in the Ba/F3 survival assay.

To complement the shRNA screen, we performed a “gain-of-function” analysis on 17 out of the 30 genes by overexpressing the wild-type or mutated gene in the Ba/F3 cells. Among the 17 genes examined, six genes were scored positive in the Ba/F3 shRNA
screen. As shown in Figure 4A, overexpression of wild-type ERBB3 significantly increased cell survival, consistent with its proposed role as an oncogene, but the P590L mutation showed no difference relative to the wild-type, suggesting that this mutation is not functional, albeit in terms of survival in the Ba/F3 assay. Compared with controls, overexpression of wild-type AKTIP and INHBA significantly reduced cell survival, consistent with their potential roles as tumor suppressors. Interestingly, their corresponding mutations (Q281K in AKTIP and R310Q in INHBA) significantly increased the Ba/F3 survival compared to the wild-type constructs \((P < 0.05)\), compared with Ba/F3 parental control. (B) Whole cell lysates were also collected for Western blotting with indicated antibodies, and ERK2 was used as the loading control.

**Figure 3.** Novel candidate driver cancer genes identified by shRNA screening in Ba/F3 viability assay. (A) Ba/F3 cells were transfected with short-hairpin RNAs (shRNA) targeting indicated genes. Empty vector (pGIPZ) and nonspecific shRNA served as the control. Cells were cultured without IL-3 for 4 wk and harvested for viability assay. Cell viability relative to Ba/F3 parental cells was shown. (*) \(P < 0.05\), compared with Ba/F3 parental control. (B) Whole cell lysates were also collected for Western blotting with indicated antibodies, and ERK2 was used as the loading control.

**Regulation of PI3K pathway activation by ARID1A**

ARID1A (AT-rich interactive domain 1A; also known as BAF250) is a key member of the SWI/SNF chromatin-modeling complex, and the gene has recently been reported to be frequently mutated in a wide variety of cancer types (Jones et al. 2010; Wiegand et al. 2010, 2011; Guan et al. 2011b; Gu et al. 2011; Wang et al. 2011; Mamoa et al. 2012). While previous studies and our present results suggest its potential role as a tumor suppressor, the molecular mechanism underlying its functional role in cancer is largely unknown. To explore how ARID1A contributes to the pathogenesis of endometrial cancer, we first resequenced this gene in the same cohort of tumor samples \((n = 222)\) as in our previous study (Cheung et al. 2011) and observed a (nonsilent) mutation frequency of 41.6\% (Fig. 6A; Supplemental Fig. 5). Strikingly, ARID1A mutations co-occur with mutations in PTEN (Fisher exact test, \(P < 1.2 \times 10^{-5}\), Bonferroni corrected \(P < 1.0 \times 10^{-4}\) and PIK3CA \((P < 2.3 \times 10^{-3}\), Bonferroni corrected \(P < 1.8 \times 10^{-2}\) (Fig. 6A; Supplemental Table 5), as well as with overall PI3K pathway aberration \((P < 1.1 \times 10^{-3}\), approach. As shown in Figure 5 and Supplemental Figure 4, decreased PTEN protein and increased Akt phosphorylation was accompanied by significant increased cell viability in PTEN siRNA-transfected cells carrying a wild-type PTEN gene and expressing PTEN protein, but not in PTEN protein-negative cell lines, supporting the utility of the approach. Importantly, KLE, in which all genes assessed were wild-type, exhibited the same response to the corresponding shRNA as Ba/F3 except for IGFBP3 siRNA (Fig. 5B). The responses of EFE184 (all genes tested are wild-type except ARID1A with undetectable protein expression), SK-UT-2 (ARID1A is mutated with detectable protein; mutation status of others is unknown), and SNG-II (ARID1A is mutated with undetectable protein; the mutation status of the others is unknown) were more variable (Supplemental Fig. 4). ARID1A siRNA only demonstrated altered viability of KLE and SK-UT-2 where the ARID1A protein is expressed (Fig. 5B; Supplemental Fig. 4). Importantly, AKTIP siRNA significantly increased viability in all cell lines, consistent with its potential tumor-suppressor role suggested in the Ba/F3 system. The siRNAs for WNT11 and INHBA increased viability in three cell lines, whereas IGFBP3 siRNA increased cell viability in two cell lines. In contrast, ERBB3 and RPS6KC1 siRNAs inhibited cell viability in all cell lines at least at one time point. siRNAs targeting NOS1AP, PHKG1, and C1orf198, which had no effect in Ba/F3 survival, did not alter cell viability in all of the endometrial cell lines we examined, providing additional support for the validity of the Ba/F3 model and the pipeline of characterization of mutations. Collectively, the effect of silencing candidate driver genes in endometrial cancer cell lines largely recapitulated the results obtained in Ba/F3 cells.

**Effect of candidate driver cancer genes by siRNA knockdown in endometrial cancer cell lines**

To extend our observations in Ba/F3 cells to human endometrial cancer cells, we introduced siRNAs targeting the 12 candidate driver cancer genes into four human endometrial cancer cell lines and assessed effects on cell viability. We included an siRNA targeting the established PTEN tumor-suppressor gene to validate the
Bonferroni corrected $P < 8.6 \times 10^{-3}$). Since coordinate PI3K pathway mutations are common in endometrial cancer relative to other cancer lineages, the concordance of aberrations in the PI3K pathway and ARID1A could represent coordinate targeting of the PI3K pathway or a mutually exclusive function for ARID1A and the PI3K pathway. We thus examined the effect of ARID1A mutations on protein and phosphoprotein levels of core members in the PI3K pathway and ARID1A could represent coordinate targeting of the PI3K pathway contributing to the tumor-suppressor activity of ARID1A.

Discussion

Our study represents the first unbiased view of somatic mutations in endometrial cancer, which strongly complements previous gene- and pathway-focused studies. More importantly, our results provide substantial functional evidence for a diversity of novel candidate drivers, suggesting key insights into the pathogenesis of endometrial cancer with implications for the development of targeted therapy.

Next-generation sequencing technology has facilitated characterization of the full spectrum of aberrations in cancer genomes in a cost-effective and timely manner (Mardis 2011). However, there is still a great gap between creating a catalog of mutations and alterations and identifying a short list of “actionable” elements (Chin et al. 2011). Here we developed a systems-biology approach to filling this gap: We combined computation-prediction-based prioritization with functional screening in a highly sensitive cell viability assay. This approach allows the identification of a large number of candidate driver cancer genes in an efficient way. Focusing on a candidate driver gene of high interest, we further examined the underlying molecular mechanisms through an integrative analysis of mutation profile and protein expression on a large, well-characterized sample set and “hypothesis-driven” functional studies in endometrial cancer cell lines. These methods and the related experimental systems can be readily applied to other cancer types.

Through the combination of gene silencing and overexpression strategies in Ba/F3 and endometrial cancer cell lines, we revealed 12 potential driver cancer genes in endometrial cancer. Among these genes, ARID1A has attracted wide interest recently. Frequent mutations throughout the gene sequence including multiple truncation mutations have been reported in ovarian (Jones et al. 2010; Wiegand et al. 2010), gynecological (Guan et al. 2011b), bladder (Gui et al. 2011), gastric (Wang et al. 2011), breast (Mamo et al. 2012), and endometrial cancers (Guan et al. 2011a; Wiegand et al. 2011), suggesting a role as a tumor suppressor. ARID1A encodes BAF250a, a nuclear protein and a key component of the SWI/SNF chromatin-remodeling complex that functions as a regulator of gene expression and chromatin dynamics (Wu et al. 2009). However, to date, the mechanisms by which loss of ARID1A function contributes to cancer pathophysiology remains poorly understood. ARID1A has been suggested to suppress cell proliferation of ovarian and endometrial cancer cells through physically interacting with p30 to coordinately regulate the transcription levels in three cell lines (KLE, ESS1, and MFE280) expressing wild-type ARID1A (Fig. 6D). In contrast, up-regulation of AKT phosphorylation was not observed in the EFE184 cell line in which ARID1A protein is not present (Fig. 6D). These results are consistent with inhibition of the PI3K pathway contributing to the tumor-suppressor activity of ARID1A.
growth dependent on the cellular context. For example, INHBA substantially inhibited tumor growth and angiogenesis in in vivo gastric cancer and neuroblastoma models (Schramm et al. 2005; Kaneda et al. 2011). Meanwhile, INHBA mRNA was up-regulated in lung cancer and may result in the promotion of cell proliferation (Seder et al. 2009). In endometrial cancer cell lines, the role of INHBA is controversial (Di Simone et al. 2002; Tanaka et al. 2003). It is possible that the contribution of INHBA to the tumorigenesis is determined by the relative expression levels of other receptor subunits and their interacting partners. WNT11 is a key member of the Wingless-type (Wnt) signaling pathway whose deregulation has been implicated in endometrioid endometrial tumors as evidence by beta-catenin mutations and aberrant nuclear accumulation (Ikeda et al. 2000; Moreno-Bueno et al. 2002). The canonical Wnt cascade is mediated by nuclear beta-catenin binding to T-cell factor transcription factors to activate genes relevant to tumorigenesis, while noncanonical Wnt signaling is beta-catenin independent (Bejo et al. 2005). Interestingly, WNT11 is down-regulated in hepatocellular carcinoma, and it can modulate both canonical and noncanonical Wnt pathways to execute its tumor-suppressor actions (Toyama et al. 2010). Thus, the role of WNT11 in Wnt signaling in endometrial cancer warrants further investigation.

ERBB3 and RPS6KC1 are potential oncogenes identified in our study. ERBB3 (HER3) belongs to the epidermal growth factor receptor family and has been implicated in cancer, previously. ERBB2, a partner of ERBB3, as well as ERBB3 itself is often ampliﬁed and overexpressed in breast, ovarian, prostate, and lung cancers (for review, see Sihanandam and Anderson 2008). Overexpression of ERBB3 promotes tumorigenesis through multiple mechanisms including cell cycle progression, stimulation of cell migration, and invasion primarily via activation of the PI3K pathway (Sihanandam and Anderson 2008). ERBB3 is highly expressed in endometrial cancer, but the functional role remains unclear (Srinivasan et al. 1999; Ejskjaer et al. 2007). RPS6KC1 (encoding ribosomal protein S6 kinase polypeptide 1) is another potential oncogene. RPS6KC1 mutations have been previously found in breast, ovary, and lung cancers (Davies et al. 2005; Stephens et al. 2005; Cancer Genome Atlas Research Network 2011). Two independent studies show that RPS6KC1 does not have kinase activity (Hayashi et al. 2002; Liu et al. 2005), but, instead, likely functions as an adaptor molecule to recruit binding partners (sphingosine kinase-1 and peroxiredoxin-3) to early endosomes (Hayashi et al. 2002; Liu et al. 2005). These trafficking pathways within the endosomal system play a fundamental role in regulating protein degradation, recycling, secretion, and compartmentalization; and, indeed, defective vesicular trafficking is a hallmark of malignant transformation (Mosesson et al. 2008).

Our study has several limitations. First, somatic alterations in noncoding regions, genome rearrangements, and copy
number variations would have been missed. Second, the number of samples sequenced is relatively small, thus sequencing of additional samples will be necessary to establish a comprehensive genomic landscape for endometrial cancer. Finally, additional functional and clinical studies will be required before the insights from the genomic analysis can be translated into effective personalized therapy for endometrial cancer.

Figure 6. Mutational and functional analysis of ARID1A on the activation of the PI3K pathway. (A) Co-mutation patterns of ARID1A and key genes related to the PI3K pathway. (Upper panel) Mutation diagram in the full set of endometrial tumor samples (n = 222). Each column represents a tumor, and each row corresponds to a single gene. (Lower panel) Mutation or co-mutation frequencies are expressed as a percentage of all the samples, and the co-mutation frequencies from random expectation are shown in parentheses for comparison. (Dark red) Genes with statistically significant co-mutations, accompanied with Bonferroni-corrected P-values. (B) The functional effect of ARID1A mutation on protein expression of the PI3K pathway. Each arrow represents a protein marker with significant differential expression between ARID1A wild-type and mutated samples: (red arrows) phosphorylated proteins and (green arrows) total proteins with \( P < 0.05 \) (two-sided t-test, \( FDR < 0.1 \)); (orange arrows) phosphorylated proteins with marginal significance \( P < 0.07 \) (\( FDR < 0.13 \)). (Solid red) Activated genes; (solid gray) genes without available protein expression data. (C) The functional effect of ARID1A mutation on the phosphorylation of AKT and p70S6K in tumor samples in which both PTEN and PIK3CA genes are wild-type, and also PTEN expression is retained (n = 47). P-values were calculated based on a two-sided t-test. (Boxes) The distribution of individual values from the lower 25th percentile to the upper 75th percentile; (solid line in the middle) median values; (lower and upper whisker) fifth and 95th percentiles; (small circles) outlier data points. (D) Four endometrial cancer cell lines were transfected with 20 nM ARID1A siRNA or nonspecific siRNA and harvested after 72 h for Western blotting with the indicated antibodies. Numerical values below each lane of the immunoblots represent the quantification of the relative protein level by densitometry.
Methods

Sample collection
All studies of 222 patients diagnosed at the University of Texas MD Anderson Cancer Center (Houston, TX, USA) from 1998 to 2009 were approved by the Institutional Review Board. Tumor content (≥80%); histological classification, grade, and stage were reviewed by two independent pathologists. Genomic DNA from frozen tumor resections was extracted by the MD Anderson Bioextraction core using the QIAamp DNA Mini Kit (QIAGEN); normal DNA was extracted from peripheral blood leukocytes using the QIAamp Blood kit (QIAGEN).

Exome sequencing and detection of somatic alterations
Whole-exome libraries (captured by Agilent SureSelect All Exon Kit) were constructed and sequenced on a SOLiD V3.0 Genome Analyzer using 50-nt single reads. Sequencing reads were processed, mapped to the human genome (hg19), and analyzed using Bioscope (version 1.21) combined with in-house scripts. Supplemental Figure 1 shows the overall computational pipeline for detecting somatic mutations. Mutation validation was performed by Sequenom MassARRAY genotyping, and whole-gene ressequencing on the large sample cohort was performed by Sanger sequencing. A complete description is provided in the Supplemental Methods.

Mutational analysis
Mutation data of ARID1A in 222 endometrial tumors were obtained as previously described (Cheung et al. 2011). A two-sided Fisher exact test was used to assess whether (nonsilent) mutations in ARID1A tend to co-occur with mutations in other genes (or PI3K pathway aberrations) in the tumor samples. PI3K pathway aberrations were defined as PTEN loss or mutations in any pathway genes (including PTEN, PI3CA, PIK3R1, PIK3R2, and AKT1). A Bonferroni correction was used to control multiple testing.

RPPA data analysis
To examine the effect of ARID1A mutations on the PI3K pathway, high-throughput RPPA data for 24 PI3K-pathway-related proteins in the 222 samples were obtained from our previous study (Cheung et al. 2011). RPPA data were processed and normalized as previously described (Park et al. 2010). A two-sided t-test was used to test whether the expression of a given protein in ARID1A wild-type samples was significantly different from that in ARID1A mutated samples, and a false discovery rate (Benjamini and Hochberg 1995) was used to control multiple testing.

Bioinformatics analysis
Pathway enrichment analysis was performed with Ingenuity Pathways Analysis (IPA) (version 8.0). The status of PTEN was determined as in our previous study (Cheung et al. 2011). A two-sided Fisher exact test was used to assess whether ARID1A mutations were correlated with the status of PTEN (loss or retained). To select mutated genes for shRNA studies, functional effects of somatic mutations were predicted with two programs: (1) CHASM (Carter et al. 2010), mutations with FDR < 30% were selected; (2) MutationAssessor (Reva et al. 2011), mutations with high functional effect were selected.

Ba/F3 viability assays
The interleukin-3 (IL-3)–dependent prolymphoid cell line Ba/F3 was maintained in RPMI1640 medium containing 10% fetal bovine serum supplemented with 5 ng/mL IL-3 at 37°C in a 5% CO2 atmosphere. Constructs containing wild-type genes or their corresponding mutants cloned into pLeni6.3 (Invitrogen) and shRNAs in pGPZIP vector (Open Biosystems) were transfected into Ba/F3 cells using the Neon electroporation system according to the manufacturer’s instructions (Invitrogen). At 96 h post-transfection, the cells were resuspended in medium without IL-3. Cells (5 × 10^3) were plated in 96-well plates and cultured for 4 wk. Cell viability was evaluated using Cell Titer Blue (Promega) for mitochondrial dehydrogenase activity. Statistical analysis was performed using ANOVA followed by Tukey’s post hoc test (GraphPad Software). P < 0.05 was considered statistically significant.

siRNA transfection in endometrial cancer cell lines
Endometrial cell lines EFE-184, ESS-1, and MFE280 were purchased from DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). KLE was provided by Dr. Russell Broadus (MD Anderson Cancer Center, Houston, TX), and SK-UT-2 was kindly provided by Dr. Bo R. Rueda (Massachusetts General Hospital, Boston). All cell lines were cultured in media according to the suppliers’ instructions at 37°C in a 5% CO2 atmosphere. Cells were transfected with 20 nM siRNAs (Dharmacon) using Lipofectamine RNAiMAX reagent (Invitrogen) according to the manufacturers’ instructions. The cells were harvested at time points as indicated in the figure legends and processed for either cell viability assay using Cell Titer Blue (Promega) or Western blot analysis.

Western blotting
Whole-cell lysates were extracted with RIPA lysis buffer (1% NP-40, 5 mM EDTA, 1 mM sodium orthovanadate, 1% phenylmethylsulphonyl fluoride, and complete protease inhibitor cocktail). The protein concentrations were determined with DC Protein Assay Reagent (Bio-Rad Laboratories). Cell lysates (25 μg) were loaded onto SDS/PAGE and transferred to a Hybond-ECL nitrocellulose membrane (Amersham Biosciences). The membrane was blocked with 5% nonfat milk and incubated with primary antibody overnight at 4°C. Protein expression was visualized with an ECL plus kit (Amersham Biosciences).

Data access
The sequencing data from this study can be accessed at the European Genome-phenome Archive (EGA) (http://www.ebi.ac.uk/ega/) under accession number EGAS00001000318.

Acknowledgments
This work was supported by a Stand Up to Cancer Dream Team Translational Research Grant, a Program of the Entertainment Industry Foundation (SU2C-AAAR-DT0209) to G.B.M.; a Uterine Cancer SPORE grant (NIH/NCI P50 CA098258) to R.R.B, K.H.L., and G.B.M.; U01 CA168394 to G.B.M.; a Career Development Award from the Conquer Cancer Foundation of the American Society of Clinical Oncology to B.T.H.; NIH/NCI R21CA152432 to G.B.M.; and G.B.M.; grants from the G.S. Cancer Biology Training Program of the Gulf Coast Consortia (CPRIT Grant No. RP101489) to L.W.T.C.; grants from the G.S. Industry Foundation (SU2C-AACR-DT0209) to G.B.M.; a Uterine Translational Research Grant, a Program of the Entertainment Industry Foundation (SU2C-AAAR-DT0209) to G.B.M.; and G.B.M.; a Career Development Award from the Conquer Cancer Foundation of the American Society of Clinical Oncology to B.T.H.; NIH/NCI R21CA152432 to R.K.; a training fellowship from the Keck Center Computational Cancer Biology Training Program of the Gulf Coast Consortia (CPRIT Grant No. RP101489) to L.W.T.C.; grants from the G.S. Hogan Gastrointestinal Research Fund and the Lorraine Dell Program in Bioinformatics for Personalization of Cancer Medicine to H.L.; and an Institutional Core grant (NIH/NCI P30 CA016672).
Author contributions: H.L., G.B.M., and B.T.H. designed the research. L.W.T.C., J.L., S.Y., K.S.H., T.D., Y.L., C.G., W.G., S.E.S., and K.L.S. performed the experimental work, including Sequencing, Sanger sequencing, Ba/F3 viability assays, and Western blots; X.L. and C.G.L. performed library preparation and SOLID sequencing; H.L., L.W.T.C., J.L., H.C., S.N.W., M.D., R.G.W.V., F.Z., R.K., and G.B.M. performed data analyses; K.H.L., R.R.B., and B.T.H. provided patient samples; H.L., L.W.T.C., J.L., and G.B.M. produced the text and figures; H.L. and G.B.M. provided leadership for the project. All authors contributed to the final manuscript.

References

Bejsovec A. 2005. Wnt pathway activation: New relations and locations. Cell 120: 11–14.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate—a practical and powerful approach to multiple testing. J Stat Soc Ser B Methodol 57: 289–300.

Byron SA, Gartside MG, Wellens CL, Mallon MA, Keenan JB, Powell MA, Goodfellow PJ, Pollock PM. 2008. Inhibition of activated fibroblast growth factor receptor 2 in endometrial cancer cells induces cell death despite PTEN abrogation. Cancer Res 68: 6902–6907.

Cancer Genome Atlas Research Network. 2008. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455: 1061–1068.

Cancer Genome Atlas Research Network. 2011. Integrated genomic analyses of ovarian carcinoma. Nature 474: 609–615.

Carter H, Samayoa J, Hruban RH, Karchin R. 2010. Prioritization of driver mutations in pancreatic cancer using cancer-specific high-throughput annotation of somatic mutations (CHASM). Cancer Biol Ther 10: 582–587.

Catasus I, Gallardo A, Cuatreroscas M, Prat J. 2008. PIK3CA mutations in the kinase domain (exon 20) of uterine endometrial adenosarcomas are associated with adverse prognostic parameters. Mod Pathol 21: 131–139.

Catasus I, Gallardo A, Cuatreroscas M, Prat J. 2009. Concomitant PIK3- Akt and PTEN alterations in endometrial carcinomas are associated with poor prognosis. Mod Pathol 22: 522–529.

Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Shinchel AC, Harvieve CL, Brunet P, Ahmann GJ, Adli M, et al. 2011. Initial genome sequencing and analysis of multiple myeloma. Nature 471: 467–472.

Cheung LW, Hennessey BT, Li J, Yu S, Myers AP, Djordjevic B, Lu Y, Stemke-Hale K, Zhang F, Ju Z, et al. 2011. High frequency of PIK3R1 and PIK3R2 mutations in endometrial cancer elucidates a novel mechanism for regulation of PTEN protein stability. Cancer Discov 1: 170–185.

Chin L, Hahn WC, Getz G, Meyerson M. 2011. Making sense of cancer genome and functional analysis of CpG island methylation in human cancers. J Cell Physiol 226: 1564–1572.

Davies H, Hunter C, Smith R, Stephens P, Greenman C, Bignell G, Teague J, Butler A, Edkins S, Stevens C, et al. 2005. Somatic mutations of the protein kinase family in human lung cancer. Cancer Res 65: 7591–7597.

Dedes KJ, Wetterskog D, Ashworth A, Reis-Filho JS. 2011. Identification and characterization of novel phosphoglycerate kinase-1-binding protein. J Biol Chem 287: 33319–33324.

Hayes MP, Douglas W, Ellenson LH. 2009. Molecular alterations of EGRF and PIK3CA in uterine serous carcinoma. Gynecol Oncol 113: 370–373.

Hennessey BT, Smith DR, Ram PN, Ly U, Mills GB. 2005. Exploiting the PI3K/ AKT pathway for cancer drug discovery. Nat Rev Drug Discov 4: 988–1004.

Hollander MC, Blumenthal GM, Dennis PA. 2011. PTEN loss in the continuum of common cancers, rare syndromes and mouse models. Nat Rev Cancer 11: 289–301.

Ikeda T, Yoshinaga K, Semba S, Kondo E, Horii A. 2000. Mutation and loss of expression of ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers. Cancer Res 60: 24: 265–266.

Gual Y, Gao G, Huang Y, Hu X, Tang A, Gao S, Wu R, Chen C, Li X, Zhou L, et al. 2011. Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. Nat Genet 43: 875–878.

Hayashi S, Okada T, Igarashi N, Fujita T, Jalhanseer S, Nakamura S. 2002. Identification and characterization of novel phosphoglycerate kinase-1-binding protein. J Biol Chem 277: 33319–33324.

Hao A, Tatsuno M, Kudo K, Kudoh K, Sakai K, Nanai T, Fujita Y, et al. 2011. Activin A inhibits vascular endothelial cell growth and suppresses tumour angiogenesis in gastric cancer. Br J Cancer 105: 1210–1217.

Kanaeda H, Aoto T, Matsumoto K, Kudo K, Aoki M, Kudoh K, Sakai K, Sagano T, et al. 2011. Activin A inhibits vascular endothelial cell growth and suppresses tumour angiogenesis in gastric cancer. Br J Cancer 105: 1210–1217.

Kato N, Nakayama K, Kihara S, Hori R, Nakamura Y, Hashimoto J, et al. 1999. Molecular alterations of PTEN in the human colorectal cancer cell line AS1. Gastroenterology 117: 123–133.

Kato N, Ishida Y, Kuroda K, Shimizu Y, et al. 2000. Altered PTEN expression as a diagnostic marker for the earliest endometrial precursor lesion. Cancer Discov 1: 1564–1572.

Kato N, Nakayama K, Kihara S, Hori R, Nakamura Y, Hashimoto J, et al. 1999. Molecular alterations of PTEN in the human colorectal cancer cell line AS1. Gastroenterology 117: 123–133.

Kato N, Ishida Y, Kuroda K, Shimizu Y, et al. 2000. Altered PTEN expression as a diagnostic marker for the earliest endometrial precursor lesion. Cancer Discov 1: 1564–1572.

Kato N, Nakayama K, Kihara S, Hori R, Nakamura Y, Hashimoto J, et al. 1999. Molecular alterations of PTEN in the human colorectal cancer cell line AS1. Gastroenterology 117: 123–133.

Kato N, Ishida Y, Kuroda K, Shimizu Y, et al. 2000. Altered PTEN expression as a diagnostic marker for the earliest endometrial precursor lesion. Cancer Discov 1: 1564–1572.

Kato N, Nakayama K, Kihara S, Hori R, Nakamura Y, Hashimoto J, et al. 1999. Molecular alterations of PTEN in the human colorectal cancer cell line AS1. Gastroenterology 117: 123–133.

Kato N, Ishida Y, Kuroda K, Shimizu Y, et al. 2000. Altered PTEN expression as a diagnostic marker for the earliest endometrial precursor lesion. Cancer Discov 1: 1564–1572.

Kato N, Nakayama K, Kihara S, Hori R, Nakamura Y, Hashimoto J, et al. 1999. Molecular alterations of PTEN in the human colorectal cancer cell line AS1. Gastroenterology 117: 123–133.

Kato N, Ishida Y, Kuroda K, Shimizu Y, et al. 2000. Altered PTEN expression as a diagnostic marker for the earliest endometrial precursor lesion. Cancer Discov 1: 1564–1572.
Risbridger GP, Schmitt JJ, Robertson DM. 2001. Activins and inhibins in endocrine and other tumors. Endocr Rev 22: 836–858.

Risinger JI, Hayes K, Maxwell GL, Carney ME, Dodge RK, Barrett JC, Berchuck A. 1998. PTEN mutation in endometrial cancers is associated with favorable clinical and pathologic characteristics. Clin Cancer Res 4: 3005–3010.

Salvesen HB, Stefansson I, Kretzschmar EL, Gruber P, MacDonald ND, Ryan A, Jacobs IJ, Akslen LA, Das S. 2004. Significance of PTEN alterations in endometrial carcinoma: A population-based study of mutations, promoter methylation and PTEN protein expression. Int J Oncol 25: 1615–1623.

Schlosshauer PW, Pirog EC, Levine RL, Ellenson LH. 2000. Mutational analysis of the CTNNB1 and APC genes in uterine endometrioid carcinoma. Modern Pathol 13: 1066–1071.

Schramm A, von Schuetz V, Christiansen H, Havers W, Papoutsi M, Wilting J, Schweigerer L. 2005. High activin A-expression in human neuroblastoma: Suppression of malignant potential and correlation with favourable clinical outcome. Oncogene 24: 680–687.

Seder CW, Hartojo W, Lin L, Silvers AL, Wang ZW, Silvers AL, Silvers AL, Silvers AL, Silvers AL. 1999. Expression of the c-erbB-3/HER-3 and c-erbB-4/HER-4 growth factor receptors and their ligands, neuregulin-1a, neuregulin-1b, and betacellulin, in normal endometrium and endometrial cancer. Clin Cancer Res 5: 2877–2883.

Stefansson HM, Salvesen HB, Akslen LA. 2004. Prognostic impact of alterations in P-cadherin expression and related cell adhesion markers in endometrial cancer. J Clin Oncol 22: 1242–1252.

Received January 11, 2012; accepted in revised form July 2, 2012.