Patterns of CTCF and ZFHX3 Mutation and Associated Outcomes in Endometrial Cancer

Christopher J. Walker, Mario A. Miranda, Matthew J. O’Hern, Joseph P. McElroy, Kevin R. Coombes, Ralf Bundschuh, David E. Cohn, David G. Mutch, Paul J. Goodfellow

Affiliations of authors: Department of Obstetrics and Gynecology, Division of Gynecology Oncology (CJW, MAM, MJO, DEC, PJG), Department of Biomedical Informatics, Center for Biostatistics, College of Medicine (JPM, KRC), The Ohio State University Comprehensive Cancer Center - James Cancer Hospital and Solove Research Institute (CTJ, MAM, MJO, JPM, KRC, DEC, PJG), Department of Physics, Department of Chemistry & Biochemistry, Department of Internal Medicine, Division of Hematology, Center for RNA Biology (RB), The Ohio State University, Columbus, OH; Department of Obstetrics and Gynecology, Division of Gynecology Oncology, Washington University, St. Louis, MO (DGM).

Correspondence to: Paul Goodfellow, PhD, Professor Obstetrics and Gynecology, 460 W. 12th Avenue, BRT 808, Columbus, OH 43210 (e-mail: paul.goodfellow@osumc.edu).

Abstract

Background: The genetic events responsible for tumor aggressiveness in endometrioid endometrial cancer (EEC) remain poorly understood. The chromosome 16q22 tumor suppressor genes CTCF and ZFHX3 are both frequently mutated in EEC, but their respective roles in outcome have not been determined.

Methods: Targeted deep sequencing of CTCF and ZFHX3 was performed for 542 EEC samples. Copy number loss (CNL) was determined using microsatellite typing of paired tumor and normal DNA and a novel Bayesian method based on variant allele frequencies of germline polymorphisms. All statistical tests were two-sided.

Results: Mutation rates for CTCF and ZFHX3 were 25.3% and 20.4%, respectively, and there was a statistically significant excess of tumors with mutation in both genes (P = .003). CNL rates were 17.4% for CTCF and 17.2% for ZFHX3, and the majority of CNLs included both CTCF and ZFHX3. Mutations were more frequent in tumors with microsatellite instability, and CNLs were more common in microsatellite-stable tumors (P < .001). Patients with ZFHX3 mutation and/or CNL had higher-grade tumors (P = .001), were older (P < .001), and tended to have more frequent lymphovascular space invasion (P = .07). These patients had reduced recurrence-free and overall survival (RFS: hazard ratio [HR] = 2.35, 95% confidence interval [CI] = 1.38 to 3.99, P = .007; OS: HR = 1.51, 95% CI = 1.11 to 2.07, P = .04).

Conclusions: Our data demonstrate there is strong selection for inactivation of both CTCF and ZFHX3 in EEC. Mutation occurs at high frequency in microsatellite-unstable tumors, whereas CNLs are common in microsatellite-stable cancers. Loss of these two tumor suppressors is a frequent event in endometrial tumorigenesis, and ZFHX3 defects are associated with poor outcome.

It is estimated that more than 50 000 cases of endometrial cancer (EC) will be diagnosed in 2015 and over half of all EC deaths can be attributed to patients with endometrioid subtype tumors (1–3). Treating cases of recurrent or advanced-stage (not confined to the uterus) disease is clinically challenging. To improve outcomes in this population, three major questions must be addressed: How do we identify early-stage patients who will recur, can we alter treatment to prevent recurrence, and how do we treat recurrent/metastatic disease? Many studies, including The Cancer Genome Atlas (TCGA) for EC, have identified...
commonly mutated genes in endometrioid endometrial carcinoma (EEC) (4–9). However, focused investigation is necessary to better understand the dysregulated molecular mechanisms responsible for tumorigenesis and disease maintenance, especially those associated with aggressive tumors.

CTCF is the major insulator protein in vertebrates and mediates long-range chromatin regulation (10–14). TCGA reports have shown that CTCF is specifically mutated at a high frequency in EEC tumors but not other malignancies (5,15). A recent report by Kemp et al. demonstrated that hemizygous Ctf knockout is sufficient to globally dysregulate the methylene and predispose animals to tumor formation (16). This confirmed our earlier work implicating CTCF as a haploinsufficient tumor suppressor gene (TSG) (17). CTCF maps to 16q22, a region of frequent deletion in a variety of malignancies (18–26). However, most 16q22 deletions are large and involve additional tumor suppressors including ZFHX3 (also called ATBF1), NQO1, CDH1, and EZF4 (reviewed in [27]). It remains uncertain which of these genes or combinations of genes confer selective advantage and which are coincidently deleted.

ZFHX3 is frequently mutated in EEC (5) and studies in other tumor types suggest ZFHX3 acts as a tumor suppressor through multiple mechanisms (22,28–32). ZFHX3 is involved in the regulation of the estrogen receptor (ER) pathway through a complex feedback loop (33,34), and dysregulated ER signaling is the hallmark of EEC. Mouse models have shown, that like Ctf, Zfhx3 is a haploinsufficient TSG and loss of Zfhx3 causes neoplastic growth and differential expression of genes in the progesterone/estrogen signaling pathway (35,36).

Thus we sought to more closely investigate genetic alterations of CTCF and ZFHX3 in EEC. We used targeted deep sequencing to detect mutation and copy number alterations in these two genes in a large cohort of patients with long-term follow-up. We also determined the clinicopathologic associations between CTCF and ZFHX3 defects (alone and co-occurring) with features that portend poor outcome.

Methods

Study Population

The patient population and tumor specimens investigated have been previously described (17,37,38). Our cohort is representative of a typical EEC patient population, as presented in the Supplementary Methods and Supplementary Table 1 (available online). All 542 subjects studied were treated at Washington University School of Medicine, St. Louis, MO, and provided written, informed consent to Washington University Human Research Protection Office (91–507 and 93–0828). All tumor DNAs were prepared from high neoplastic fractions by FIGO 2009 classification (37,38). Seventeen MSI-low tumors were included with 396 cases of heterozygous polymorphisms detected by deep sequencing (personal communication Kevin R. Coombes). Sixty-three germline single nucleotide variants (SNVs) were detected with 396 cases heterozygous for one or more SNV. In brief, the model considers the observed VAFs relative to the percent tumor cellularity and predicts the probability of a deletion/amplification event based on deviation from the expected VAF of .50. For our cohort, the mean and median neoplastic cellularity were 78.5% and 80.0%, respectively. The observed number k of variant reads out of a total of n reads is modeled using a binomial distribution, Binom (k, n; θ), where the ‘success’ parameter θ depends on the copy number state (Deleted, Normal, or Gained) and the fraction, v, of normal cells. This binomial model defines the likelihood function for the copy number state and normal fraction. We then place a discrete prior on the copy number state and a beta prior on the normal fraction. In this context, the posterior distribution can be computed by brute force by evaluating the likelihood at a grid of points over the interval (0,1) and applying Bayes’ formula. Microsatellite typing and the Bayesian method showed high concordance (98%) and were performed by researchers blinded to clinical data.

Normalized read depth was calculated by jointly minimizing the logarithmic normalized coverage difference for the middle two quartiles of all amplicons for each sample pair, fitting a Gaussian distribution to this data for the middle two quartiles of samples for each amplicon, and reporting z-scores relative to this Gaussian distribution for each amplicon of each sample.

Statistical Analysis

Clinicopathologic variables were compared between indicated mutational subgroups using chi-squared or Fisher’s exact tests for dichotomized variables. The continuous variables age and body mass index (BMI) were dichotomized at 60 years (42) and 30 kg/m² (43), respectively. Stage and grade were determined by FIGO 2009 classification (44), and stage was treated as a

Assessing Loss of Heterozygosity and Determining Copy Number Deletion

Fluorescence-based microsatellite typing was used to test for loss of heterozygosity (LOH) (41). Because of frequent strand slippage mutations in MSI tumors, microsatellite LOH analyses were limited to MSS tumors (n = 306). Seven highly informative dinucleotide repeats in the 16q22–23 region were evaluated (Supplementary Table 2, available online). Fragment sizing was performed using a 3730 DNA Analyzer (Applied Biosystems, Foster City, CA). Three hundred and one samples were informative for one or more marker tested.

We also applied a Bayesian method to determine allelic imbalances based on the variant allele frequencies (VAFs) of heterozygous polymorphisms detected by deep sequencing (personal communication Kevin R. Coombes). Sixty-three germline single nucleotide variants (SNVs) were detected with 396 cases heterozygous for one or more SNV. In brief, the model considers the observed VAFs relative to the percent tumor cellularity and predicts the probability of a deletion/amplification event based on deviation from the expected VAF of .50. For our cohort, the mean and median neoplastic cellularity were 78.5% and 80.0%, respectively. The observed number k of variant reads out of a total of n reads is modeled using a binomial distribution, Binom (k, n; θ), where the ‘success’ parameter θ depends on the copy number state (Deleted, Normal, or Gained) and the fraction, v, of normal cells. This binomial model defines the likelihood function for the copy number state and normal fraction. We then place a discrete prior on the copy number state and a beta prior on the normal fraction. In this context, the posterior distribution can be computed by brute force by evaluating the likelihood at a grid of points over the interval (0,1) and applying Bayes’ formula. Microsatellite typing and the Bayesian method showed high concordance (98%) and were performed by researchers blinded to clinical data.

Normalized read depth was calculated by jointly minimizing the logarithmic normalized coverage difference for the middle two quartiles of all amplicons for each sample pair, fitting a Gaussian distribution to this data for the middle two quartiles of samples for each amplicon, and reporting z-scores relative to this Gaussian distribution for each amplicon of each sample.
of matched tumor and normal DNA. Sixty-five calls were confirmed as somatic mutations, 16 were novel germline polymorphisms, and two were false-positives.

The overall rates of somatic mutation in CTCF and ZFHX3 were 25.3% (87 different mutations in 136 tumors) and 20.4% (148 different mutations in 110 tumors), respectively. The vast majority of CTCF-mutated and most ZFHX3-mutated tumors had loss-of-function (LOF) mutations (Figure 1; details provided in Supplementary Table 3, available online). Twenty-nine tumors in our series have POLE mutations that have been associated with an ultramutated phenotype, and 24 (82.8%) had mutations in CTCF and/or ZFHX3. Because the POLE-mutated samples likely carry many passenger mutations, we considered this subset separately (Supplementary Table 3, available online).

CTCF and ZFHX3 mutations were more frequent in MSI tumors than MSS tumors (P < .001 for either gene, two-sided Fisher’s exact test). Among the 214 MSI tumors, 83 had CTCF mutations (38.8%) and 77 had ZFHX3 mutations (36.0%), compared with 12.8% and 4.7% for MSS tumors, respectively. The CTCF p.T204fs hotspot mutation in an A, coding repeat (17) was seen in 49 MSI tumors, and two hotspot frameshift mutations in ZFHX3, p.E763fs and p.R1893fs*35/p.D1894fs*21, were observed exclusively in MSI tumors (Figure 1). Both ZFHX3 hotspots involved repeat regions: a G, coding repeat in exon 2 was mutated in eight patients, and an AG, repeat in exon 9 (not previously reported in EC) was mutated in eight patients (Supplementary Table 3, available online). Comutation of both CTCF and ZFHX3 was seen in 33 specimens, which is statistically significantly greater than expected based on the individual mutation rate of each gene (22 expected, P = .003 two-sided Fisher’s exact, excluding samples with allelic deletion outlined below).

In addition to mutations, we detected frequent LOH/allelic imbalance in the 16q22 region. Combined microsatellite typing and determining LOH using deep sequencing data (see Methods) gave an overall allelic loss rate of 18.1%. Representative findings for LOH detected by each method are presented in Figure 1. Analysis of normalized amplicon read depth revealed that allelic imbalance was typically associated with copy number loss (CNL) (Supplementary Figure 1, available online).

Results

Mutation and Allelic Deletion of CTCF and ZFHX3 in EEC

Targeted deep sequencing of CTCF and ZFHX3 was successfully completed for 538 of 542 EEC samples with an average read depth of 345X. The entire CTCF coding region was sequenced. Three regions of ZFHX3 were not analyzed because of poor coverage, which code for amino acids 765–833, 2026–2171, and 3187–3703. Overall 79.0% of the ZFHX3 coding sequence was covered. We undertook extensive technical validation of variant calls. Deep sequencing of 109 tumors that were previously Sanger sequenced for mutations in CTCF (17) correctly identified 20 different mutations in 34 tumors, and four additional nonclonal mutations were discovered (minor sequencing peaks were evident upon review of the original chromatograms). An additional 83 potential mutations in CTCF and ZFHX3 within the rest of the cohort were analyzed using Sanger sequencing.
Table 1

<table>
<thead>
<tr>
<th>1499</th>
<th>1593</th>
</tr>
</thead>
<tbody>
<tr>
<td>variant</td>
<td>total</td>
</tr>
<tr>
<td>n143837268</td>
<td>217</td>
</tr>
<tr>
<td>n6499137</td>
<td>1008</td>
</tr>
<tr>
<td>n140139907</td>
<td>5</td>
</tr>
</tbody>
</table>

![Figure 2A](https://academic.oup.com/jnci/article-abstract/107/11/djv249/2457691)

Figure 2. Representative example of allelic loss in the CTCF/ZFHX3 region. Left panel shows variant allele frequencies (VAFs) for single-nucleotide polymorphisms (SNPs) in CTCF and ZFHX3 in tumor 1593 (harboring allelic loss) and tumor 1499 (a typical two-copy tumor). Allelic imbalance is indicated by heterozygous SNPs with VAFs that deviate from the expected 0.50 (underlined). On the right of the ideogram for 16q region of interest are results for fluorescence-based microsatellite (MS) typing of paired tumor and normal DNA. MSs that revealed loss of heterozygosity (LOH) or showed no loss (WT) are given for both tumors. Uninformative markers indicated by *.-* Traces for chr16ms4 typing are shown in the right panel. Arrow is used to emphasize the 371 bp allele. LOH = loss of heterozygosity; VAF = variant allele frequency; WT = wild-type.

Most instances of CNL included both CTCF and ZFHX3 (16.6% of all tumors), with 1.5% showing loss of either gene alone (Supplementary Table 3, available online). CNL rates for CTCF and ZFHX3 were 17.4% and 17.2%, respectively. CNL was statistically significantly more common in MSS than MSI tumors, with deletion in 68 and 19 case patients, respectively (P < .001 two-sided Fisher's exact test). A single instance of homozygous deletion was identified by greatly reduced normalized read depth within the CTCF coding sequence (Figure 3A). Long-range polymerase chain reaction (PCR) revealed a 1741 bp deletion spanning intron 2–3 (1157 bp) and exon 3 (594 bp), including the initiation codon (chr16:67,643,579-67,645,330) (Figure 3B). Homozygous deletion was confirmed by quantitative PCR of genomic DNA (data not shown).

Clinicopathologic Features Associated With CTCF and ZFHX3 Mutation and Copy Number Status

The frequent codeletection and comutation of CTCF and ZFHX3 observed (23.3%) suggests selection for inactivation of both genes. Patients with defects in both genes had higher-grade tumors (FDR P = .001), had more frequent lymphovascular space invasion (LVSI; FDR P = .04), and were older (>60 years, FDR P = .04) when compared with wild-type patients (Table 1). Tumors with defects in CTCF only and ZFHX3 only were also seen at appreciable frequencies (16.5% and 9.2%, respectively) (Figure 4A). We present the relationship between clinicopathologic features and mutation and/or deletion stratified by the status of both genes (both genes defective, CTCF only defective, ZFHX3 only defective, and both genes wild-type) in Supplementary Table 4 (available online). Given the differences between MSS and MSI tumors in mutation rates, the spectra of mutations, and the frequency of CNL, we also analyzed MSS and MSI patients separately. Compared with wild-type, MSS tumors with defects in both CTCF and ZFHX3 were higher grade (FDR P < .001), and these patients were nonwhite (FDR P = .06). There were no statistically significant clinicopathologic associations in the MSI tumors (Supplementary Table 5, available online).

Women with tumors defective in both genes and those with defects in ZFHX3 alone had reduced recurrence-free survival (Figure 4B). The hazard ratio (HR) was 2.29 for the group with both genes defective vs wild-type (95% confidence interval [CI] = 1.24 to 4.22, FDR P = .03) and was 2.32 for ZFHX3-only defective cases (95% CI = 1.02 to 5.23, FDR P = .09). OS was also statistically significantly reduced for case patients with defects in both genes (HR = 1.51, 95% CI = 1.06 to 2.15, FDR P = .07) (Supplementary Figure 2A, available online). In a multivariable model that included clinicopathologic factors statistically significantly associated with outcome (stage, grade, age, LVSI, and therapy) (38), the hazard ratio for mutation/CNL of both CTCF and ZFHX3 on RFS was 1.54 (95% CI = .79 to 3.01), but the increase in risk did not reach statistical significance (Supplementary Table 6, available online). Early-stage (stage I/II) patients with mutation/CNL of both genes recurred statistically significantly more often than those without: There were 88 early-stage patients with defects in both genes and 13 recurred (14.8%), compared with only 12 recurrences in the 190 wild-type patients (6.3%) (odds ratio = 2.56, 95% CI = 1.03 to 6.46, FDR P = .10, two-sided Fisher's exact). Kaplan-Meier analysis for RFS confirmed this finding (Supplementary Figure 2B, available online).

The Kaplan-Meier plots presented in Figure 4B show patients with defects in ZFHX3 only, and patients with both CTCF and ZFHX3 defects have worse outcome compared with the CTCF-only group and wild-type patients, which strongly implicates ZFHX3 as the factor contributing to reduced RFS. Therefore we compared cases with any ZFHX3 defect to ZFHX3 wild-type cases. Patients with ZFHX3 defects had higher-grade tumors (FDR P = .001), had more frequent LVSI (FDR P = .07), and were older (FDR P < .001) (Supplementary Table 7, available online). These women had reduced RFS and OS (RFS: HR = 2.35, 95% CI = 1.38 to 3.99, FDR P = .007; OS: HR = 1.51, 95% CI = 1.11 to 2.07, FDR P = .04) (Figure 5A and B). ZFHX3 mutation remained a statistically significant factor associated with RFS in multivariable analysis (P = .04), and only tumor grade and stage were stronger predictors of recurrence in this model (Table 2).
We used TCGA for EC dataset as a validation cohort. RFS was compared between samples with mutation/CNL in ZFHX3 (n = 45 EEC cases) and wild-type samples (n = 124 EEC cases). The RFS event rates are similar in our study and TCGA (12.6% and 13.0%). In TCGA, cohort patients with defects in ZFHX3 had reduced RFS with a hazard ratio similar to our experimental cohort, but the difference did not reach statistical significance (HR = 2.00, 95% CI = 0.84 to 4.74, P = 0.12) (Figure 5C). However, considering the event rates, hazard ratios, number of individuals, and the proportion of individuals mutated in each cohort, TCGA cohort had only 40.2% power to detect the effect, whereas our experimental cohort achieved 83.6% power to detect the effect.

**Discussion**

Our data show CTCF and ZFHX3 are among the most frequently altered genes in type I endometrial cancers (46): When both mutation and deletion are considered, 40% of EEC tumors have CTCF defects and 33% have abnormalities in ZFHX3. Although both CTCF and ZFHX3 were classified as statistically significantly mutated genes by TCGA, co-occurring mutations, codeletion, and the relationship between gene defects and outcome were not addressed in the TCGA endometrial cancer report (5). Our combined sequencing and CNL analysis allowed us to fully interrogate the mutational status of both genes. Although it is well established that deep sequencing data can be used to detect copy number differences, accuracy depends on read depth (47,48). Mutation detection in coding sequence repeats can be challenging with exome capture sequencing methods. For tumor types such as EEC that frequently have defective DNA mismatch repair, the difficulty in sequencing coding repeats has likely resulted in underestimation of mutational frequency (47,48). The deep read depth we achieved and use of long (425 bp) amplicons made it possible to reliably detect both allelic imbalances/CNL and insertion/deletion mutations in coding repeats.

TCGA copy number alteration data (from GISTIC) (5) confirm deletion of CTCF and ZFHX3 occurrence in EEC at modest frequency and show that most of the deletions encompass large portions of 16q. TCGA exome sequencing data demonstrate that CTCF and ZFHX3 are the only genes on chromosome 16q mutated at an appreciable frequency (>6%). Although many expressed genes map to the 16q deletion region we propose that portions of 16q are among the most frequently altered genes in type I endometrial cancers (46). Our data show CTCF and ZFHX3 are among the most frequently altered genes in type I endometrial cancers (46).
Figure 4. Patterns of CTCF and ZFHX3 defects and associated outcomes. A) Pie charts show the percentage of samples harboring mutation and/or allelic loss of CTCF only, ZFHX3 only, both genes and cases wild-type for both genes. Microsatellite-stable cases are shown on the left, and microsatellite-unstable cases are shown on the right. B) Kaplan-Meier estimates for recurrence-free survival. P values and hazard ratios for each group were calculated by Cox proportional hazard regression and false discovery rate adjusted.

Figure 5. ZFHX3 defects and association with survival. Kaplan-Meier estimates for recurrence-free (A) and overall survival (B) in our experimental cohort. C) Kaplan-Meier estimates for recurrence-free survival of endometrioid endometrial cancer patients analyzed by The Cancer Genome Atlas. P values obtained by log-rank test for univariate analyses and false discovery rate (FDR) adjusted as indicated.
Table 2. Multivariable model for outcome with ZFHX3 mutation/CNL*

<table>
<thead>
<tr>
<th>Variable</th>
<th>RFS</th>
<th>OS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 3</td>
<td>5.17 (2.25 to 11.89)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Stage III/IV</td>
<td>3.30 (1.56 to 6.99)</td>
<td>.002</td>
</tr>
<tr>
<td>Grade 2</td>
<td>2.94 (1.38 to 6.27)</td>
<td>.005</td>
</tr>
<tr>
<td>ZFHX3 mutation/CNL</td>
<td>1.82 (1.03 to 3.20)</td>
<td>.04</td>
</tr>
<tr>
<td>Presence of LVSI</td>
<td>1.48 (0.77 to 2.83)</td>
<td>.24</td>
</tr>
<tr>
<td>Use of adjuvant therapy</td>
<td>1.19 (0.56 to 2.51)</td>
<td>.65</td>
</tr>
<tr>
<td>Age &lt; 60 y</td>
<td>0.97 (0.53 to 1.77)</td>
<td>.92</td>
</tr>
</tbody>
</table>

* Model includes all clinicopathologic features listed and mutation and/or copy number loss of ZFHX3 as variables. Reference groups are as follows: for grade 2 and grade 3, grade 1; for stage, stage I/II, for ZFHX3 and CTCF defects, ZFHX3 and CTCF wild-type; for presence of LVSI, absence of LVSI; for age, < 60 years, age > 60 years; for use of adjuvant therapy, no adjuvant therapy. CI = confidence interval, CNL = copy number loss, HR = hazard ratio; LVSI = lymphovascular space invasion; OS = overall survival; RFS = recurrence-free survival.

† P values calculated using multivariable Cox proportional hazard tests. P values are two-sided.

Endometrial cancers are often categorized as having either an MSI or chromosomal instability (CIN) mutator phenotype (G1–53). The genetic inactivation of CTCF and ZFHX3 we observed in cancers with normal DNA mismatch repair (MSI tumors with CIN) and those with MSI implicates strong selection in EEC. MSI tumors frequently have chromosomal deletions involving the region containing CTCF and ZFHX3, whereas MSI tumors often harbor SNV/indels in both genes, including hotspot mutations in coding sequence repeats. Examination of the VAFs of these hotspot mutations showed that the vast majority are clonal (variant frequency > 3%) and are, as such, unlikely to reflect mutational noise secondary to defective mismatch repair (data not shown). Three-fourths of CTCF mutations are LOF (frameshift, nonsense, and splice site) and, as we have previously demonstrated, subjected to nonsense-mediated decay (17). Examination of the CTCF missense mutations shows that they cluster around the zinc finger domains, and condel (54) predictions indicate most are deleterious to protein function (Figure 1; Supplemental Table 3, available online). Although the evidence for LOF of ZFHX3 is not as compelling, half of all mutations are LOF, and half of the missense mutations are predicted to be deleterious (Supplementary Table 3, available online). Thus we propose that allelic loss and mutations in EEC both lead to haploinsufficiency for CTCF and ZFHX3.

Recent studies have challenged the Knudsen two-hit model (55) for tumor suppressors and suggest that hemizygous deletions play critical roles in shaping the cancer genome (56,57). Our previous report on CTCF defects in EC strongly pointed to its function as a haploin-sufficient TSG, which has been confirmed in vivo (16,17). Knock-out mouse modeling has also implicated ZFHX3 as a haploinsufficient TSG (35). Here, we found that only 9.6% of CTCF-mutated tumors and 19.1% of ZFHX3-mutated tumors harbor a second hit, further supporting haploinsufficiency. Although one specimen in our cohort has homozygous deletion of CTCF, the alternative initiation codon for isoform 2 is retained, which in principle could compensate for the loss of isoform 1.

Importantly, our data show that women with tumors defective in ZFHX3 alone (no evidence for mutation in CTCF) have statistically significantly reduced RFS indistinguishable from that seen in cases with defects in both genes (Figure 48). Although CTCF and ZFHX3 are commonly inactivated together, our findings point to ZFHX3 mutation as the key factor associated with poor outcome.

The correlative nature of the work described here is an obvious limitation of the study and inherent to many biomarker and molecular discovery efforts. In vitro and in vivo studies to functionalize the genetic inactivation of these two genes are the necessary next steps for better understanding the roles ZFHX3 and CTCF play in endometrial tumorigenesis.

Overall, the frequency and spectra of mutations and deletions we identified in the closely linked tumor suppressors CTCF and ZFHX3 suggests strong selection for inactivation of both genes in EEC. The statistically significantly reduced RFS for women with defects in ZFHX3 highlights its underappreciated importance in EEC, specifically in biologic/clinical aggressiveness.

Funding

This work was supported by The National Institutes of Health (R21 CA155674 to PQ), the National Cancer Institute (P30 CA16058) supporting the Nucleic Acid and Biostatistics shared resources at the Ohio State University Comprehensive Cancer Center, and the Pelotonia Fellowship Program (CJW).

Notes

The sponsors had no role in the design of the study, the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication.

We would like to thank Alexis Chassen for manuscript editing and data curation, Sandya Liyanarachchi for assistance with data analysis, and Nima Esmaeil Mokaram for assistance with computational methods. We would like to acknowledge Pearily Yan and the Ohio State University Genomics Core Facility, and Tea Meulia and the Ohio State University Molecular and Cellular Imagine Center, a CFAES/OARDX core facility, in Wooster, OH.

We are very grateful to all of the patients who contributed specimens to this study and all of the attending physicians and staff at the Washington University School of Medicine Division of Gynecologic Oncology.

All authors declare they have no conflict of interest.

References


