

ORIGINAL ARTICLE

CTCF genetic alterations in endometrial carcinoma are pro-tumorigenic

AD Marshall^{1,2}, CG Bailey^{1,2}, K Champ^{1,2}, M Vellozzi^{1,2}, P O'Young^{1,2}, C Metierre^{1,2}, Y Feng^{1,2}, A Thoeng^{1,2}, AM Richards³, U Schmitz^{1,2}, M Biro⁴, R Jayasinghe^{5,6}, L Ding^{5,6}, L Anderson^{2,7}, ER Mardis^{5,6} and JEJ Rasko^{1,2,8}

CTCF is a haploinsufficient tumour suppressor gene with diverse normal functions in genome structure and gene regulation. However the mechanism by which CTCF haploinsufficiency contributes to cancer development is not well understood. CTCF is frequently mutated in endometrial cancer. Here we show that most CTCF mutations effectively result in CTCF haploinsufficiency through nonsense-mediated decay of mutant transcripts, or loss-of-function missense mutation. Conversely, we identified a recurrent CTCF mutation K365T, which alters a DNA binding residue, and acts as a gain-of-function mutation enhancing cell survival. CTCF genetic deletion occurs predominantly in poor prognosis serous subtype tumours, and this genetic deletion is associated with poor overall survival. In addition, we have shown that CTCF haploinsufficiency also occurs in poor prognosis endometrial clear cell carcinomas and has some association with endometrial cancer relapse and metastasis. Using shRNA targeting CTCF to recapitulate CTCF haploinsufficiency, we have identified a novel role for CTCF in the regulation of cellular polarity of endometrial glandular epithelium. Overall, we have identified two novel pro-tumorigenic roles (promoting cell survival and altering cell polarity) for genetic alterations of CTCF in endometrial cancer.

Oncogene (2017) 36, 4100–4110; doi:10.1038/onc.2017.25; published online 20 March 2017

INTRODUCTION

Uterine malignancies represent the most prevalent gynaecological cancers in the developed world. These tumours are becoming an increasing health burden due to associations with increased body mass index, nulliparity and increased life expectancy.^{1,2} The majority (95%) of these tumours are endometrial carcinomas that originate from the endometrial glandular epithelium. The remaining 5% have a mesenchymal component such as endometrial stromal sarcoma or are mixed epithelial and stromal tumours such as adenosarcomas and carcinosarcomas.^{3,4} Within endometrial carcinoma, the majority of tumours (~80%) are endometrioid adenocarcinomas, which are generally hormone-responsive, are typified by *PTEN* deletion, and carry a good prognosis. The next most common subtypes are serous carcinoma (5–10%) and clear cell carcinoma (<5%), which are generally non-hormone-responsive and associated with a poorer prognosis. Additionally, serous tumours often carry *TP53* mutations.^{3–5} Rarer subtypes include mucinous carcinoma, squamous cell carcinoma, undifferentiated carcinoma, and mixed carcinoma which by definition contains at least two histological cell types such as endometrioid and serous carcinoma.³

Cancer genome sequencing studies focusing on large patient cohorts of specific cancers are revealing the molecular genetic landscapes of tumours. Recently, we showed that endometrioid tumours had frequent mutations in *PTEN*, *CTNNB1*, *PIK3CA*, *ARID1A*, *ARID5B* and *KRAS*.⁵ We also discovered that *CTCF*, encoding the chromatin organising protein CTCF is mutated in about one

quarter of endometrial carcinoma.⁵ This was verified in independent endometrial cancer cohorts.^{6,7} Importantly, of 127 significantly mutated genes (SMGs) in 12 different cancers, *CTCF* was identified as an SMG in endometrial cancer.⁸ *CTCF* mutations have also been identified at similar or lower frequencies in other tumour types including breast and prostate cancer, Wilms' tumour and leukaemia.^{9–12} A *CTCF* mutation has been identified at relapse in acute lymphoblastic leukaemia¹¹ indicating a potential role for *CTCF* mutation in tumour relapse. Specific mutations in the zinc finger (ZF) domain of *CTCF* result in reduced or abrogated *CTCF* DNA binding activity at certain cognate target sites, but not others.¹⁰

CTCF is an 11-ZF DNA binding protein that has been named the 'master weaver of the genome'¹³ due to its diverse functions in the regulation of chromatin structure and function including: inter- and intra- chromosomal interactions, gene regulation, nucleosome positioning, gene insulation, genetic imprinting, demarcation of lamina associated domains, X-chromosome inactivation, alternative splicing and telomere elongation (reviewed in Marshall *et al.*¹⁴). We first showed that *CTCF* can repress cell growth and colony formation suggesting a tumour suppressor function for *CTCF*.¹⁵ In addition, *CTCF* expression also confers a protective effect against apoptosis in various cell lines.^{16,17} *CTCF* has been shown to modulate the expression of various cancer-associated gene loci including notable examples: *CMYC*,¹⁸ *IGF2*,¹⁹ *TP53*,²⁰ *TERT*²¹ and *INK4/ARF*.²² Loss of *CTCF* binding via aberrant DNA methylation of critical control regions can induce epigenetic silencing of tumour suppressor loci or lead

¹Gene and Stem Cell Therapy Program, Centenary Institute, Camperdown, New South Wales, Australia; ²Sydney Medical School, University of Sydney, Sydney, New South Wales, Australia; ³Gynaecological Oncology, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia; ⁴Cell Motility and Mechanobiology, School of Medical Sciences, University of New South Wales, Sydney, New South Wales, Australia; ⁵Cancer Genomics, McDonnell Genome Institute, Washington University in St Louis, St Louis, MO, USA; ⁶Division of Oncology, Department of Medicine, Washington University in St Louis, St Louis, MO, USA; ⁷Department of Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia and ⁸Cell and Molecular Therapies, Royal Prince Alfred Hospital, Sydney, New South Wales, Australia. Correspondence: Professor JEJ Rasko, Gene and Stem Cell Therapy, Centenary Institute, Building 93, RPA Hospital, Sydney 2050, New South Wales, Australia. E-mail: j.rasko@centenary.org.au

Received 20 October 2016; revised 21 December 2016; accepted 5 January 2017; published online 20 March 2017

to activation of oncogenes.^{23,24} *Ctcf* heterozygous mice are prone to tumour development in various tissues including the uterus.²⁵ This verifies *CTCF*'s role as a haploinsufficient tumour suppressor gene.

Here we show that *CTCF* mutations identified in primary human endometrial carcinoma predominantly showed loss-of-function phenotypes due to nonsense-mediated decay of mutant transcripts or abrogation of functional activity. In addition, we also identified a recurrent mutation in *CTCF*, K365T, which showed a gain-of-function phenotype promoting increased cell survival following apoptotic insult. To recapitulate *CTCF* haploinsufficiency in *CTCF* wild-type (WT) cells we used shRNA knockdown of *CTCF*. We show that reduced *CTCF* expression reduces the proportion of spheroids showing luminal localisation of apical polarity markers F-Actin and ZO-1 in 3D endometrial spheroid culture, indicating *CTCF* plays a role in regulating the normal structure of endometrial glands. Furthermore, our copy number variation analysis of primary human endometrial cancers demonstrates that *CTCF* deletion occurs more commonly in tumours with a propensity for relapse or metastasis, and in tumours with clear cell histology. These findings provide further evidence that genetic lesions in *CTCF* promote endometrial carcinogenesis.

RESULTS

CTCF is mutated in endometrioid adenocarcinomas and related cell lines

We and others have identified over 200 somatic *CTCF* mutations occurring in primary human endometrial carcinoma samples^{5-7,26} (Figure 1a and Supplementary Table 1). The majority (71%) are inactivating mutations resulting from frameshift (45%), nonsense (20%) and splice site mutations (6%). The remaining mutations (29%) are missense mutations with the majority (43/60, 72%, $P < 0.001$ Fisher's exact test) occurring within the 11-ZF DNA binding domain of *CTCF*. All of the recurrently mutated residues in *CTCF* (S345, K352, K365, G375, R377, P378, A387, and R567) occur within the ZF domain and are predominantly clustered within ZF 4 (Figure 1b, $P = 0.0041$, Fisher's exact test). As *CTCF* is a significantly mutated gene in endometrial cancer, we sequenced *CTCF* in five endometrial cancer cell lines and identified four mutations (Figures 1a and c). These included two recurrent *CTCF* mutations: T204fs*26 in Ishikawa and R342H in both HEC1A and HEC1B and two novel mutations: G19* in HEC1B and R278C in RL95-2. Interestingly HEC1B, which is a sub-strain of HEC1A isolated from a patient with grade II endometrial cancer,²⁷ has acquired an additional somatic mutation in *CTCF*, G19*. Only the KLE endometrial cancer cell line did not contain any *CTCF* mutations. KLE has previously been shown to be diploid for Ch16²⁸ and does not contain focal deletions within Ch16q22.1.²⁹ The location of missense mutations within the ZF domains of *CTCF* is shown in Figure 1d and includes mutations affecting zinc ion binding, DNA binding and non-binding residues.

Within our published patient cohort *CTCF* was sequenced in 172 endometrioid adenocarcinomas and 60 serous carcinomas.⁵ In all, 16 missense and 25 inactivating mutations (including 17 nonsense and 8 frameshift mutations) were identified (Supplementary Table 2). All somatic *CTCF* mutations were identified within the endometrioid subtype (Fisher's exact test $P < 0.001$; Figure 2a). We determined the proportion of mutant to wild-type alleles in tumour samples following DNA and RNA sequencing. In tumour DNA, there was a similar ratio of mutant to wild-type alleles for missense mutations and inactivating mutations ($37.6 \pm 3.2\%$ and $32.8 \pm 2.3\%$, respectively, mean \pm s.e.m) consistent with monoallelic mutation of *CTCF* (Figure 2b). The proportion of expressed alleles with missense mutations was $63.1 \pm 6.3\%$ and with inactivating mutations was $7.7 \pm 3.2\%$ (mean \pm s.e.m.), indicating that most nonsense- and frameshift-mutated alleles were

eliminated by nonsense-mediated decay (NMD; Figure 2b). Consistent with nonsense-mediated decay being the mechanism by which nonsense- and frameshift-mutated alleles are degraded, all but two of these mutations obeyed the rule for termination-codon position^{30,31} in which the premature termination codon must be > 55 nt upstream of the last exon-exon junction to result in NMD (Supplementary Table 1). These data indicate that the majority of *CTCF* mutations in endometrial carcinoma result in inactivation of one allele.

CTCF haploinsufficiency occurs in serous endometrial carcinoma
We next determined if somatic copy number alterations in *CTCF* occurred in endometrial carcinoma (as estimated by GISTIC from within the TCGA portal). Loss in *CTCF* copy number occurred in 7.0% of endometrioid and 65.5% of serous subtype tumours, indicating that *CTCF* genetic deletion was more common in poorer prognosis serous subtype tumours (Fisher's exact test, $P < 0.001$; Figure 2c). When samples with somatic *CTCF* mutations were excluded, relative linear copy number values for *CTCF* correlated with *CTCF* mRNA expression ($\rho = 0.384$, Spearman's log-rank correlation). Furthermore, samples with somatic *CTCF* deletion expressed significantly less *CTCF* mRNA ($P < 0.001$, Mann-Whitney test; Figure 2d). *CTCF* copy number gain was observed in only a single serous carcinoma sample. Survival analysis restricted to patients with serous tumours showed that tumours with *CTCF* copy number loss resulted in reduced overall survival compared to *CTCF* diploid tumours (Figure 2e and Supplementary Figure 1). These indicate that genetic deletion of *CTCF*, potentially in concert with other deleted genes on chromosome 16q, results in poorer patient outcomes.

Loss- and gain-of-function *CTCF* mutations are observed in endometrial cancer

We incorporated three recurrent *CTCF* missense mutations found in endometrial carcinoma,⁵ K365T, R377H and P378L (Figures 1a and c) into lentiviral expression constructs containing HA-tagged wild-type (WT) *CTCF*. We functionally characterised WT and mutant *CTCF* in *CTCF*-haploinsufficient Ishikawa cells. All three mutants of *CTCF* were indistinguishable from WT in their nuclear localisation (Figure 3a) and were expressed at similar levels (Figure 3b). We next performed established cellular assays to measure the activity of *CTCF* including proliferation and colony-forming assays.^{15,17} Both R377H and P378L abrogated the growth inhibitory effects of WT *CTCF* ($P < 0.05$) such that they were indistinguishable from control eGFP vector (Figure 3c). The K365T mutation did not alter the anti-proliferative effect of WT *CTCF*. Furthermore, R377H and P378L showed loss-of-function by abrogating the inhibition of clonogenicity of WT *CTCF* ($P < 0.01$ and $P < 0.001$ respectively), while K365T again behaved similarly to WT *CTCF* (Figure 3d). We explored the effect of each mutation on *CTCF* function further by examining UV-induced apoptosis in Ishikawa cells. WT *CTCF* has a protective effect compared to control ($P < 0.01$; Figure 3e) as expected.¹⁷ The K365T mutation enhanced this pro-survival effect ($P < 0.001$), suggesting a gain-of-function in *CTCF* while neither R377H nor P378L showed any increase in survival compared to control.

CTCF knockdown promotes loss of endometrial cell line polarity in 3D culture

As the majority of *CTCF* mutations identified in endometrial carcinoma effectively result in inactivation of one *CTCF* allele, we modelled this using inducible shRNA knockdown of *CTCF*. In shCTCF KLE cells we reduced *CTCF* expression by about half (Figure 4a). However, we did not detect any significant change in cell proliferation after *CTCF* knockdown (Figure 4b). As with *CTCF* overexpression studies, we observed a decrease of colony

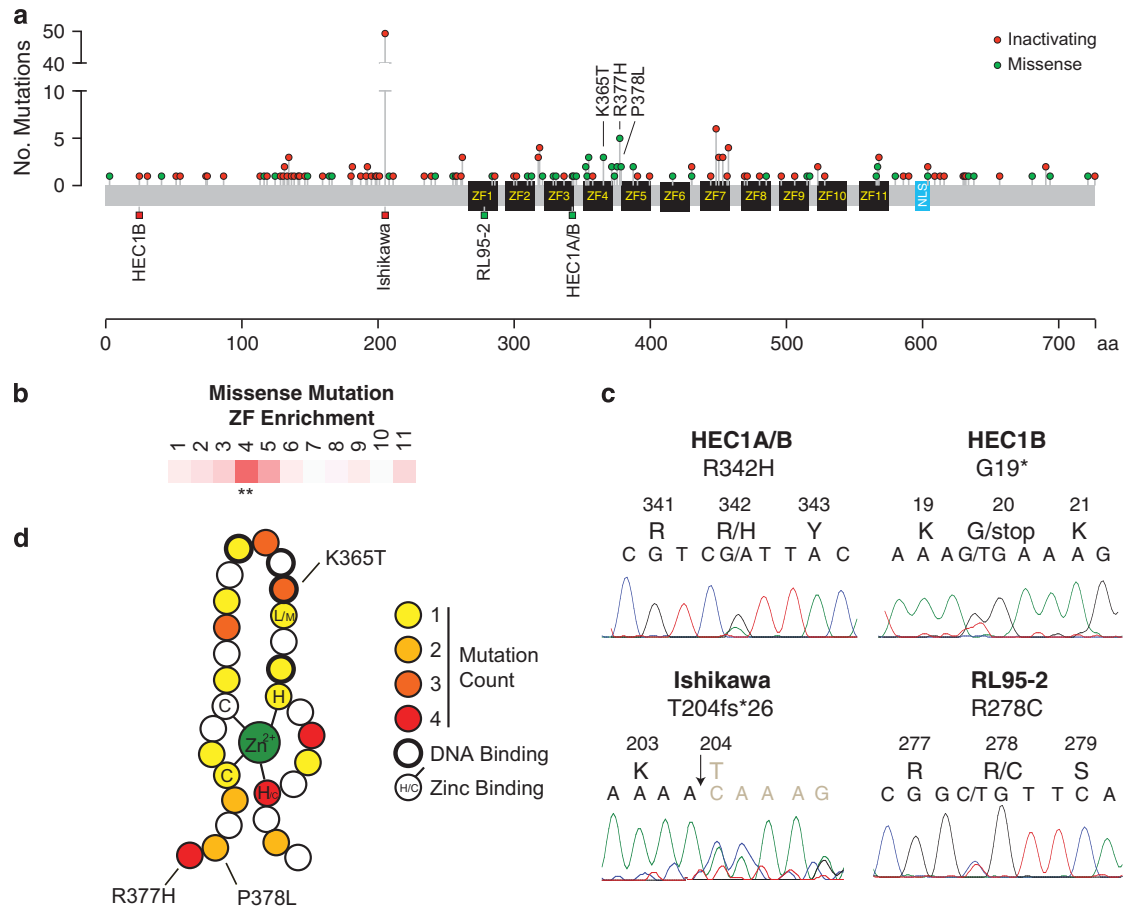


Figure 1. CTCF mutations in human endometrial carcinoma. **(a)** Summary of all CTCF inactivating (nonsense and frameshift; red) and missense (green) mutations identified in primary human endometrial carcinomas^{5–7} (round lollipops, above). Novel CTCF mutations identified in endometrial cancer cell lines (square lollipops, below). The location of ZFs (black) and nuclear localisation signal (NLS; blue). Recurrent mutations K365T, R377H and P378L functionally characterised in this study are indicated. **(b)** Heatmap showing the frequency of missense mutations in different ZF domains of CTCF ranging from 0 (white) to 15 (dark pink) residues, ***P* < 0.01 Fisher's exact test compared to expected. **(c)** Representative electropherograms of somatic mutations in endometrial cancer cell lines: R342H, G19*, T204fs*26, R278C. **(d)** The frequency of CTCF missense mutations found in endometrial carcinoma superimposed on a C2H2 ZF schematic. Note: some ZFs contain only 3 residues juxtaposed between conserved C-terminal histidine residues.

formation with shRNA knockdown of CTCF (Figure 4c). This is consistent with preliminary CRISPR genome editing studies performed in our laboratory (data not shown). However, knockdown of CTCF did not have any effect on KLE cell survival in response to UV treatment (Figure 4d). These data indicate that there is not a simple linear relationship between CTCF expression level and cell characteristics such as cell growth, colony formation and protection from apoptosis. Colony formation in particular shows that neither over- nor underexpression of CTCF is optimal for maximal colony formation.

Endometrial cancer cell lines such as KLE are derived from glandular epithelium, and as such, can be induced *in vitro* to grow spheroids reminiscent of endometrial glands.³² These endometrial spheroids display basal/apical polarity when stained for F-actin, a marker of apical polarity based on phalloidin staining, and the apical tight junction protein ZO-1.³³ Using this staining technique we found that spheroids generated from KLE cells after CTCF knockdown had a reduced proportion of spheroids showing apical/central F-actin and ZO-1 staining compared to control cells (Figures 4e and f). This shows that reduced CTCF gene expression disrupts endometrial cancer cell polarity, a morphological change characteristic of epithelial cancer development.³⁴ There was no consistent change observed with CTCF knockdown compared to the three relevant controls in spheroid area, spheroid

number, ZO-1 staining intensity or F-actin staining intensity (Supplementary Figure 2).

CTCF knockdown alters KLE gene expression

To identify the impact of CTCF knockdown on global gene expression in KLE cells we performed gene expression microarrays. We identified a total of 1744 genes that were differentially expressed in KLE cells after 8 days of induction of CTCF knockdown (Figure 5a and Supplementary Table 3). Gene ontology analysis of this data set revealed significant enrichment for the following terms: Process: regulation of retinoic acid receptor signaling pathway; Function: identical protein binding, protein homodimerisation activity, protein binding, and protein dimerisation activity; Component: intracellular part, cytoplasmic part (Figure 5b).

CTCF genetic deletion is associated with relapse or metastasis, and clear cell histology

We examined a small independent cohort of 23 primary endometrial malignancies as well as 8 primary tumour samples which developed local relapse and 7 primary tumour samples that presented or later developed metastasis. DNA isolated from FFPE samples was analysed using a custom Nanostring nCounter copy number variation panel of cancer-associated genes across the

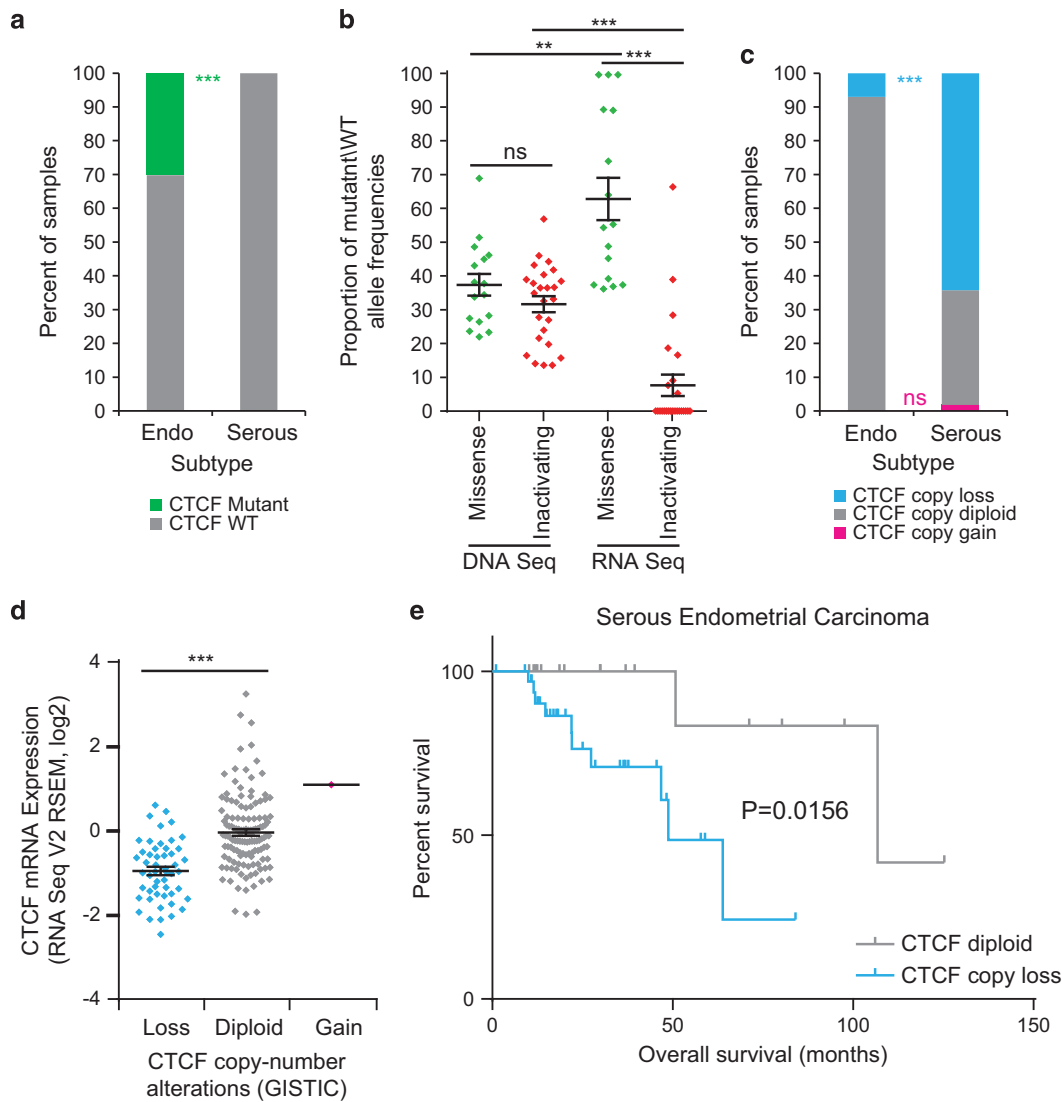


Figure 2. CTCF is mutated in endometrioid carcinoma and deleted in serous carcinoma. Analysis of mutation and copy number variation from published data.⁵ (a) Proportion of endometrioid (Endo, $n = 172$) and serous samples ($n = 60$) with mutations in CTCF; $***P < 0.001$ Fisher's exact test. (b) Proportion of missense ($n = 16$) or inactivating (nonsense and frameshift; $n = 25$) mutations in all reads obtained by DNA Seq and RNA Seq spanning the mutation site in samples; ns not significant, $**P < 0.01$, $***P < 0.001$, Mann-Whitney non-parametric test. (c) CTCF gene copy number analysis of endometrioid ($n = 172$) and serous ($n = 60$) samples using GISTIC; $***P < 0.001$ Fisher's exact test. (d) CTCF gene expression in non-CTCF mutant endometrioid and serous samples categorised for CTCF copy number using GISTIC (Loss $n = 51$, Diploid $n = 128$ and Gain $n = 1$); $***P < 0.001$ Fisher's exact test. (e) Survival analysis of serous endometrial cancer samples with ($n = 34$) or without ($n = 17$) CTCF genetic deletion; $P = 0.0156$, log-rank (Mantel Cox) test.

long arm of chromosome 16 (16q) with a particular focus on 16q22.1 where CTCF is located (Figure 6). Within the 23 primary samples only a single grade 2 endometrioid carcinoma showed CTCF deletion. Among the 8 tumours that relapsed, 3 showed CTCF deletion including a grade 2 endometrioid tumour and two clear cell carcinomas. Among tumours that developed metastases, two of seven showed CTCF deletion including a mixed grade 3 endometrioid/clear cell tumour and a grade 2 endometrioid tumour. Collectively CTCF was more likely to be deleted in endometrial tumours that developed relapse or metastasis than primary malignancies ($P = 0.0268$, Fisher's exact test). Moreover, two of three relapsed tumours exhibited clear cell histology and one of two metastatic tumours displayed mixed endometrioid/clear cell histology. This indicates a possible association between the presence of clear cell histology and CTCF deletion ($P = 0.0259$, Fisher's exact test), which warrants further investigation in a more extensive cohort. We found no evidence that CTCF deletion

occurred specifically in relapsed or metastatic tumours using matched primary and relapse or metastatic samples (four each, Supplementary figure 3).

DISCUSSION

We and others⁵⁻⁷ have shown that CTCF is mutated in about one quarter of endometrioid adenocarcinomas. The majority of these somatic CTCF mutations are inactivating mutations resulting from nonsense, frameshift or splice site mutations, which we show are rarely expressed due to nonsense-mediated decay³⁵ of mutant transcripts, consistent with other reports.⁷ The missense mutations which have been identified occurred predominantly in the 11-ZF DNA binding domain of CTCF. A number of CTCF ZF mutations previously identified in isolated cancers were shown to inhibit CTCF binding to specific DNA binding sites such as *CMYC* and *IGF2/H19* regulatory regions while binding to other sites such as

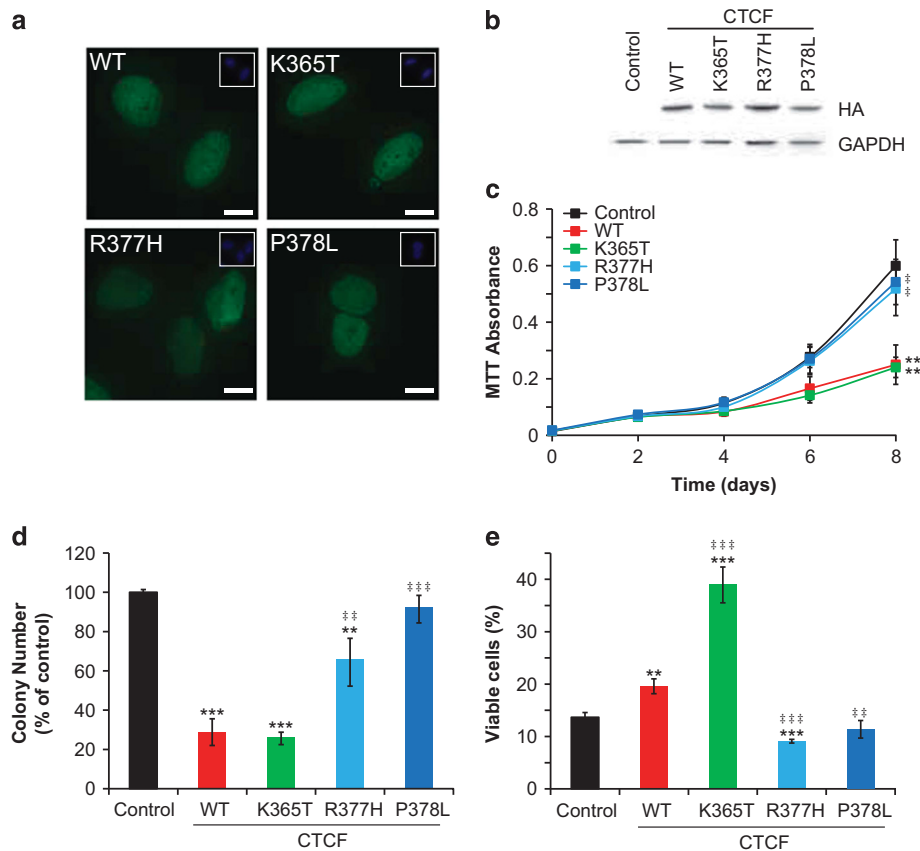


Figure 3. CTCF mutations alter normal CTCF function. **(a)** Subcellular localisation and **(b)** Western blot expression level of HA-tagged wild-type (WT), K365T, R377H and P378L CTCF in Ishikawa cells as detected by the anti-HA antibody; scale bar = 10 μ m; inset shows DAPI staining of nuclei. Functional analysis of Ishikawa cells expressing CTCF WT, K365T, R377H and P378L, or control eGFP vector: MTT proliferation assay **(c)**; colony formation assay **(d)**; and apoptosis assay following recovery from UV exposure **(e)**; * indicates a significant difference from control; † indicates a significant difference from WT CTCF; * or † $P < 0.05$, ** or †† $P < 0.01$, *** or ††† $P < 0.001$ Student's *T*-test. Data are mean \pm s.e.m of 3–4 independent experiments.

the β -globin and *APP β* regulatory regions was unaffected.¹⁰ In endometrial carcinoma these missense mutations appear to cluster within ZFs 4–5 of CTCF. ZFs 4–7 are responsible for binding to the 20 bp core binding motif of CTCF which constitutes ~80% of all CTCF target sites across the genome.³⁶

Recurrent CTCF missense mutations in endometrial cancer occur within inter-ZF regions (R377H and P378L), and within conserved residues critical for zinc co-ordination or residues that directly contact the minor groove of DNA (K365T). R377H and P378L are loss-of-function mutations which abrogate the tumour suppressive effects of CTCF. These inter-ZF mutations may alter the conserved DNA binding domain structure and prevent normal CTCF function. Conversely, K365T did not disrupt the tumour suppressive effect of CTCF on cellular growth but enhanced CTCF's response to UV-induced apoptosis indicating a pro-survival gain-of-function. As this residue directly contacts DNA, this mutation may alter DNA binding specificity. The exact nature or consequences of this altered binding has yet to be determined.

Despite our demonstration of a gain-of-function mutation in CTCF, the majority of CTCF mutations in endometrial cancer result in loss-of-function. We and others have found that the mechanism of CTCF gene inactivation differs between cancer subtypes. Microsatellite instability (MSI)-positive endometrioid tumours are prone to strand slippage mutations in CTCF resulting in the recurrent T204fs*18 and T204fs*26 mutations.⁷ We found that CTCF is genetically deleted in a further 22% of all endometrial carcinomas, and that the highest rates of CTCF genetic deletion occur in endometrial serous carcinomas (65.5%). Consequently,

these result in a significant reduction in CTCF gene expression, indicating that haploinsufficient deletion of CTCF may be important in endometrial carcinoma pathogenesis and/or progression. In addition, other genes within Chr16q are often concurrently deleted with CTCF. The haploinsufficient loss of these additional genes may synergise with CTCF haploinsufficiency in promoting more aggressive behaviour in endometrial carcinomas. Thus the mechanism of CTCF inactivation could contribute to tumour subtype and prognosis.

CTCF has previously been suggested as a target of 16q22.1 deletions observed in various cancer types.^{6,9,37,38} CTCF is located within the minimally deleted region of 16q22.1,³⁷ however other candidate genes such as E-cadherin (*CDH1*) and *ZFH3* have also been suggested.^{6,38} E-cadherin is required for normal development and function of the endometrium,³⁹ while loss of E-cadherin staining is associated with endometrial carcinoma development and poor prognosis.^{40,41} Likewise, somatic *ZFH3* mutations occur in about 20% of endometrial cancers and are associated with poor patient outcomes.⁶ We have determined that haploinsufficient deletions of 16q, which usually contain CTCF, are associated with serous, and now potentially clear cell histology and poorer prognosis. This provides the first evidence that CTCF inactivation in conjunction with inactivation of *CDH1* and/or *ZFH3* and/or other candidate genes on 16q contributes to the development of poorer prognosis tumour subtypes. Any additive or synergistic effect between CTCF haploinsufficiency and that of other candidate genes within 16q has yet to be investigated.

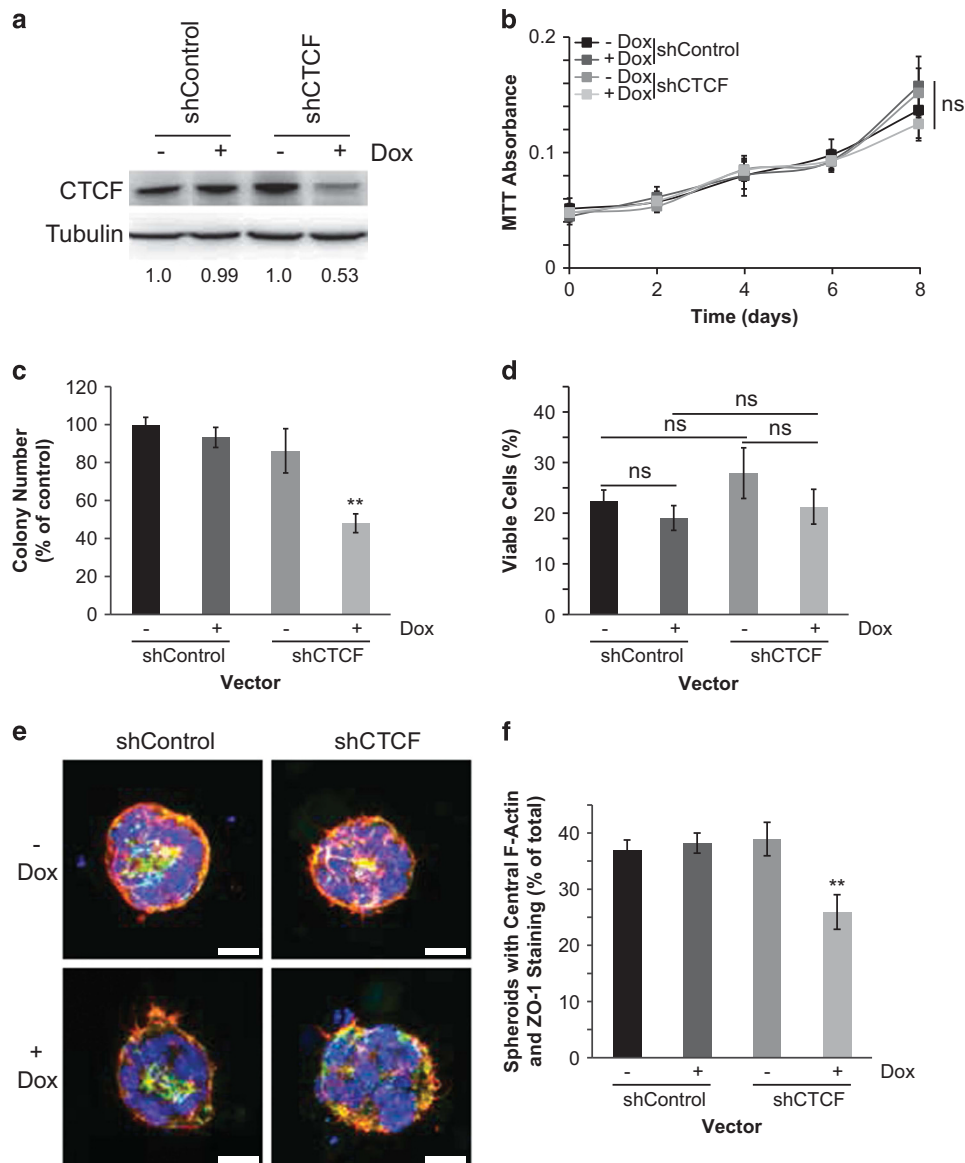


Figure 4. CTCF knockdown alters clonogenicity and endometrial cancer spheroid polarity. **(a)** Immunoblot of KLE cells transduced with inducible shControl or shCTCF shRNA vectors in KLE cells with or without doxycycline (Dox) induction. Densitometric ratio of CTCF expression compared to non-Dox induced samples is shown. Functional analysis of KLE cells expressing either shControl or shCTCF with or without Dox induction: MTT assay **(b)**; colony formation assay **(c)**; and apoptosis assay following recovery from UV insult **(d)**. **(e)** Maximum intensity projections of confocal images of representative KLE endometrial cancer spheroids expressing shControl or shCTCF stained with rabbit anti-ZO-1 (green), F-actin (rhodamine phalloidin, red) and DAPI (blue); scale bar = 10 μ m. **(f)** Quantitation of spheroids polarisation; ns not significant, ** $P < 0.01$ Student's *t*-test. Data are mean \pm s.e.m of 3–4 independent experiments.

CTCF has been established as a tumour suppressor gene due to its ability to suppress cancer cell growth.^{15,17,25} Along with *PTEN*, *TP53*, *PIK3CA*, *CTNNB1* and *ARID1A*, *CTCF* has been classified as a significantly mutated gene in endometrial cancer.⁸ *Ctcf* haploinsufficient mice, which are tumour prone, also develop endometrial cancers.²⁵ These observations firmly implicate *CTCF* mutation and haploinsufficiency in endometrial carcinoma pathogenesis. However the cellular consequences of *CTCF* mutation and haploinsufficiency had not been investigated. Here we show that at least one recurrent mutation in *CTCF*, K365T, exhibits a pro-survival gain-of-function phenotype which could contribute to tumourigenesis. To date, this gain-of-function mutation, K365T has been identified in two cases of endometrioid adenocarcinoma⁵ and one serous carcinoma case.²⁶ This K365T mutation is the only *CTCF* mutation to be identified in a worse prognosis serous subtype tumour.

Conversely the majority of endometrial cancer-associated genetic changes to *CTCF* result in inactivation of one allele of *CTCF* leading to haploinsufficiency. Here we show that these can deregulate endometrial cell polarity, a phenotype which is associated with malignancy. Thus, loss-of-function missense mutations, inactivating mutations and genetic copy number loss are different mechanisms which achieve reduced *CTCF* gene dosage and/or expression and/or *CTCF* protein activity. To date, no homozygous deletion resulting in complete inactivation of *CTCF* has been reported in tumours. A single case has been reported showing homozygous deletion of *CTCF* exon 3⁶ which encodes the entirety of the N-terminus of *CTCF*. RNA sequencing was not performed to determine if alternative transcripts of *CTCF* were expressed. *CTCF* is absolutely required for development as *Ctcf* null mouse embryos fail to develop into blastocysts.⁴² Thus we expect

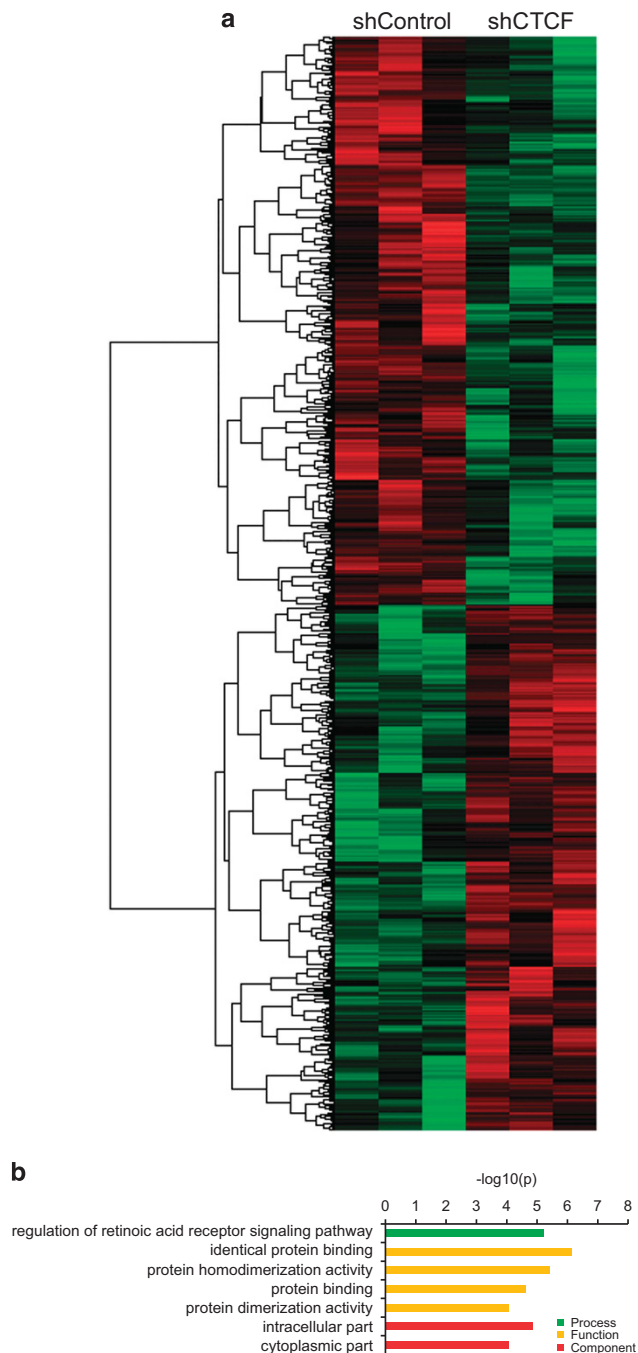


Figure 5. Knockdown of CTCF results in gene deregulation in KLE cells. **(a)** Heat map of significantly (adjusted P -value < 0.05) differentially expressed genes in shControl (+Dox) samples compared to shCTCF (+Dox) samples. Gene names are provided in Supplementary Table 3. **(b)** Gene ontology analysis of differentially expressed genes showing terms enriched with P -value $< 10^{-3}$ and an FDR Q -value of < 0.1 .

complete loss of *CTCF* in tumours would be detrimental to cell viability and tumorigenesis, similar to when enforced deletion of both alleles of *DICER1*, a known haploinsufficient tumour suppressor gene, leads to inhibition of tumorigenesis.⁴³

CTCF haploinsufficiency also destabilises DNA methylation at epigenetically variable CpGs in normal tissues.²⁵ It is not clear what specific effects this global epigenetic instability would have on endometrioid cells, but such deregulation may accelerate carcinogenesis, and is indeed recognised as an

enabling factor in cancer.⁴⁴ Paradoxically we show that knockdown of CTCF can have both pro-tumourigenic (deregulation of cellular polarity) and anti-tumourigenic (inhibition of clonogenicity) activities. Also, either CTCF overexpression or knockdown inhibit colony formation in endometrioid cancer cells. This indicates that CTCF must be tightly regulated to ensure normal cell function.

The endometrium of a reproductive human female (from puberty to menopause, excluding pregnancy) is a dynamic tissue undergoing monthly cycles of proliferation, secretion and menstruation. This cycling is accompanied by dramatic morphological changes in the endometrial glands. However, despite these changes, the glandular epithelial cells maintain normal apical-basal polarity through specific subcellular expression of polarity proteins.⁴⁵ Loss of cellular polarity allows cells to override normal contact inhibition signals and is thus a hallmark of cancer.⁴⁴ This is consistent with the finding of Kemp *et al.*²⁵ who showed that CTCF haploinsufficiency did not alter the proliferation of MEFs under subconfluent conditions but showed loss of contact inhibition at confluency. CTCF knockdown in the KLE endometrial cancer cell line reduces the proportion of endometrial cancer spheroids showing central luminal staining of apical polarity markers. Loss of cellular polarity is thought to be an early event in endometrial cancer development, as it is evident in endometrial atypical hyperplasia which progresses to malignant carcinoma in about a third of patients.^{46,47} Moreover, loss of cellular polarity is an important aspect of epithelial to mesenchymal transition (EMT) known to be involved in tumour progression.⁴⁸ Interestingly, in mice, *Ctcf*^{+/-} tumours exhibit increased aggressiveness, including invasion, metastasis and mixed epithelial and mesenchymal morphology.²⁵ This is the first time that CTCF has been implicated in the regulation of cellular polarity, and could indicate an important role for CTCF haploinsufficiency in the development of endometrial and other cancers.

Here we showed that knockdown of CTCF resulted in the significant deregulation of gene expression in KLE cells. Using gene process ontology analysis we found significant enrichment for genes associated with the retinoic acid receptor signalling pathway. Interestingly retinoic acid receptor signalling has been associated with differentiation, decidualisation and cancer formation in the endometrium,^{49–51} and is also associated with the regulation of glandular epithelium polarity.⁵² This may offer a gene regulatory mechanism by which CTCF haploinsufficiency contributes to endometrial cancer development.

In conclusion, we have shown that CTCF mutations and genetic deletions predominantly result in loss-of-function and more rarely pro-survival gain-of-function phenotypes. *CTCF* mutation is restricted to the endometrioid subtype of endometrial carcinoma, while genetic deletion of Ch16q including the *CTCF* locus, occurs predominantly in worse prognosis subtypes including serous carcinoma, and our preliminary studies indicate clear cell carcinoma. Inactivation of one allele of *CTCF* (*CTCF* haploinsufficiency) effectively reduces the amount of functional CTCF in endometrial cancer cells. CTCF knockdown to recapitulate haploinsufficiency resulted in deregulation of endometrial cancer spheroid polarity which likely contributes to endometrial cancer development and progression. Tumours with *CTCF* haploinsufficiency can now be shown to exhibit four different hallmarks of cancer: sustained proliferative signalling, resisting cell death, evading growth suppressors and epigenetic deregulation of the genome.⁴⁴ Our study provides further evidence of CTCF's role as a haploinsufficient tumour suppressor gene. Future studies to explore the global consequences of *CTCF* haploinsufficiency on three-dimensional genomic architecture and the transcriptional landscape in cancer are now warranted.

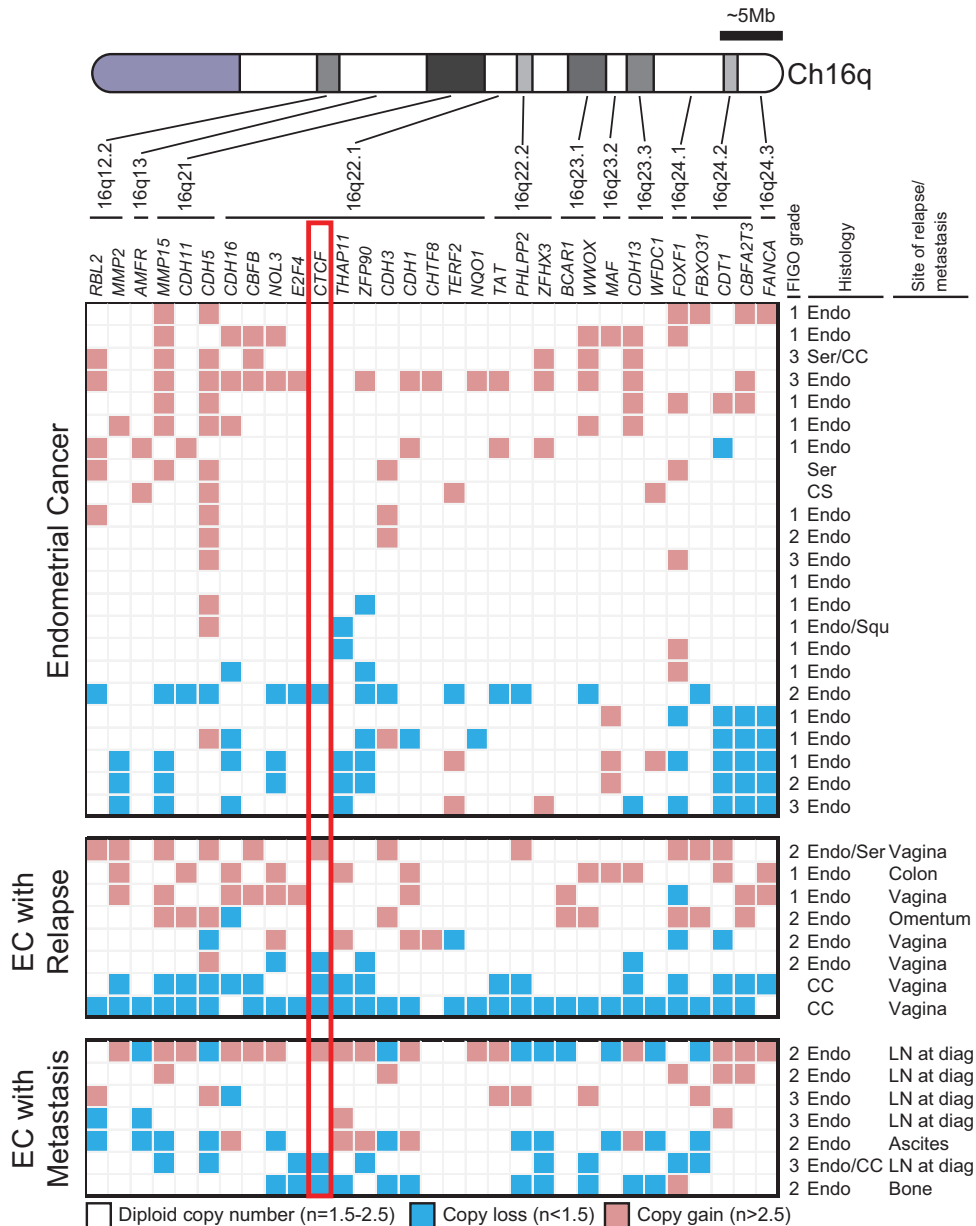


Figure 6. CTCF genetic deletion is associated with relapse and metastasis, and clear cell histology. A custom Nanostring nCounter Copy Number Variation (CNV) Code Set was designed to span cancer-associated genes in the long arm of chromosome 16 with particular focus on genes within 16q22.1 containing the *CTCF* locus (average of three probes, indicated by red border). Chromosome band and gene name are labelled above. An endometrial cancer cohort ($n = 23$, top), endometrial carcinomas (EC) which relapsed ($n = 8$, middle), and endometrial carcinomas (EC) which metastasised ($n = 7$, bottom) were analysed. CNV counts were normalised to the average of seven normal or benign endometrium samples which were all considered diploid and any variation depicted on the heatmap. Federation Internationale de Gynecologie et d'Obstetrique (FIGO) tumour grading scores are shown. Endometrial cancer histologies included endometrioid (Endo), Serous (Ser), Clear Cell (CC), Carcinosarcoma (CS) and Squamous Carcinoma (Squ). Where mixed histologies were seen multiple subtypes are listed. The sites of local relapse or metastasis are listed, and if detected at diagnosis (at diag); (LN) lymph node.

MATERIALS AND METHODS

Cell culture

Ishikawa, HEC1A, HEC1B, RL95-2 and KLE endometrial cancer cell lines were a kind gift from Associate Professor Deborah Marsh of the Kolling Institute, Sydney. Prior to commencing this study, these cell lines were authenticated by short tandem repeat profiling in 2014 (CellBank, Westmead, NSW, Australia). All cells were cultured in DMEM:F12 medium (Life Technologies, Carlsbad, CA, USA) containing 10% (v/v) FBS and penicillin-streptomycin solution at 37 °C and 5% CO₂ with or without 1 µg/ml doxycycline (Dox) as indicated. We used established methods for 3D spheroid culture.^{32,33} Briefly, 96-well imaging plates (Beckton Dickinson, Franklin Lakes, NJ, USA) were coated in Matrigel (Corning, Tewksbury, MA, USA) and allowed to harden.

KLE cells (5×10^3 /well) were plated in DMEM:F12 medium containing penicillin-streptomycin, 1 mM HEPES, 5 ng/ml EGF, 1:100 dilution of Insulin-Transferrin-Selenium (ITS) supplement (Life Technologies) and 3% (v/v) Matrigel for 10 days with media changes every 2–3 days with or without 1 µg/ml Dox.

CTCF Sequencing

Genomic DNA was extracted from Ishikawa, HEC1A, HEC1B, RL95-2 and KLE endometrial cancer cell lines using the Purelink genomic DNA isolation kit (Life Technologies). PCR amplicons spanning *CTCF* exons 3–11 including splice sites were amplified with established primers¹⁷ using Phusion DNA Polymerase (NEB, Ipswich, MA, USA). Each PCR product was Sanger

sequenced in forward and reverse directions. All mutations were confirmed by cloning each PCR amplicon into pCR-Blunt-TOPO (Life Technologies) and re-sequencing.

Generation of lentiviral expression

Point mutations (K365T, R377H and P378L) were introduced into pCCLtet eGFP-2A-HA-hCTCF¹⁷ using splice overlap extension PCR. Sequences for shRNAs targeting human CTCF (shCTCF; 5'-tcccCGAAAGCAGCATTCC TATAttcaagagaTATAGGAATGCTGCTTTCCGcttttc-3') or encoding an *Arabidopsis thaliana* microRNA 159a (shControl; 5'-tcccTTTGATTGAAGG GAGCTCtcaagagaGAGCTCCCTCAATCCAAActtttc-3') were cloned into *BbsI/XhoI* sites in the Dox-inducible vector pFH1tUTG.⁵³ Lentiviral supernatants for each vector were produced by calcium phosphate transfection of HEK293T cells using Tat-independent packaging plasmids as previously described.¹⁷ Lentiviral particles were then concentrated by ultracentrifugation for 2 h at 20,000 rpm, 4 °C in a SW28 Beckman rotor. Viral pellets were resuspended on ice in complete Dulbecco's Modified Eagle Medium (DMEM) in 1% of the original volume. Cells were transduced with lentivirus in the presence of Polybrene (8 µg/ml) by spinoculation for 1.5 h. The transduction media on cells was then replaced with fresh media. Approximately 48 h post transduction, GFP-positive cells were sorted on a FACSAria IIu (Becton Dickinson, Franklin Lakes, NJ, USA) and either plated immediately for cell biology assays (CTCF mutants) or cultured in the absence of Dox (shCTCF) until experiments were performed.

Immunoblotting, immunofluorescence, imaging and quantitation

For Western blotting, cell lysates were electrophoresed on NuPAGE Novex 4-12% Bis Tris Protein Gels (Life Technologies) and transferred to nitrocellulose membrane using the iBlot System (Life Technologies). Membranes were blocked in either 5% (v/v) skim milk in phosphate buffered saline 0.1% (v/v) Tween-20 (PBST) or 1% (w/v) polyethylene glycol MW3350, 1% (w/v) polyvinylpyrrolidone, 0.3% (w/v) bovine serum albumin in PBST for 1 h at room temperature then incubated with 1:1000 mouse anti-HA (Covance, Princeton, NJ, USA), 1:1000 rabbit anti-CTCF¹⁵ or 1:1000 mouse anti-tubulin (Santa Cruz, Dallas, TX, USA) in blocking buffer. Blots were washed three times each for five minutes in PBST. Membranes were incubated with anti-mouse/rabbit HRP (1:5000) in blocking buffer for 1 h at room temperature. Membranes were treated with chemiluminescent reagent and exposed on an Image Station 4000R Pro (Kodak, Rochester, NY, USA). Densitometry analysis was performed using Image J software (NIH, Bethesda, MD, USA).

Cells or spheroids were rinsed in PBS and fixed in 10% (v/v) neutral buffered formalin (Fronine, Riverstone, NSW, Australia) for 10 min at room temperature (RT). Samples were washed 3 times for 5 min in phosphate buffered saline (PBS) and then permeabilised with 0.5% (v/v) Triton X-100 in PBS for 10 min at RT. Samples were then washed in PBS as above. Samples were blocked in 20% (v/v) BlokHen (Aves Labs, Tigard, OR, USA) in PBS, 0.02% (v/v) Triton X-100 for 30 min at room temperature and incubated with primary antibodies as follows: mouse anti-HA (1:1,000; Covance) or rabbit anti-ZO-1 (1:100; Cell Signaling, Boston, MA, USA) in blocking buffer for 1.5 h at RT. Samples were then washed 3 times for 5 min in PBS 0.1% (v/v) Triton X-100. Secondary antibodies: anti-mouse Alexa Fluor 594 (1:5,000; Life Technologies) or anti-rabbit Alexa Fluor 488 (1:5,000; Life Technologies) and rhodamine phalloidin to bind F-actin (1:200, Life Technologies) were incubated in blocking buffer for 1 h at RT. Samples were washed 3 times for 5 min in PBS 0.1% Triton X-100. For spheroids the first wash contained 0.2 µg/ml DAPI (Life Technologies) to counterstain nuclei. Cells were mounted in Prolong Gold plus DAPI (Life Technologies), and spheroids were covered with PBS. Imaging of cells was performed on the DeltaVision Personal Deconvolution Microscope (GE Healthcare, Chicago, IL, USA) and images deconvolved using SoftWoRx software (GE Healthcare). Imaging of spheroids for scoring was performed on a BD pathway (Becton Dickinson) and representative images were taken on a Confocal SP5 (Leica, Wetzlar, Germany). Spheroids were scored for central staining of both F-Actin and ZO-1 by a researcher blinded to sample identity. There were an average of 105 ± 6 (mean ± s.e.m) spheroids per image, one image per well, and duplicate wells for each experiment. Four replicate experiments were analysed.

Cell biology assays

All cell biology assays were performed on GFP-positive cells as previously described.¹⁷ Briefly, MTT cell proliferation assays were performed using 200-250 cells/well plated in a 96-well plate in triplicate and proliferation

was assessed at 0, 2, 4, 6 and 8 d by the addition of 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, Merck, Kenilworth, NJ, USA) overnight. The reaction was quenched with isopropanol/HCl and the optical density at 570 and 630 nm measured using a POLARstar Omega plate reader (BMG Labtech, Ortenberg, Germany). Clonogenicity assays were performed using 1000 cells/10 cm dish in triplicate and incubated for 10-20 d. Cells were rinsed in PBS and fixed with ice-cold methanol, then stained with Giemsa solution diluted 1:20 in H₂O. The colonies were scored by a researcher blinded to sample identity. For apoptosis experiments, 1 × 10⁵ cells were plated in 24-well plates and incubated overnight. The medium was replaced with PBS and cells were UV irradiated (Ishikawa 1000 µJ, KLE 4000 µJ) in a Stratallinker (Stratagene, San Diego, CA, USA). After 18 h recovery, both adherent and non-adherent cells were stained for flow cytometry using Annexin V-APC (Becton Dickinson) and DAPI (Sigma, Croydon, UK). Viable cells were defined as those cells that were negative for both stains.

Gene expression array analysis

Triplicate samples of KLE cells expressing either shControl or shCTCF knockdown lentivectors were treated with doxycycline for 8 days prior to isolation of RNA. Gene expression analysis was performed using the GeneChIP PrimeView Human Gene Expression Array (Affymetrix, Santa Clara, CA, USA). Analysis was performed using the Bioconductor packages *affy* and *limma*. Array data were normalised using the Robust Multi-array Average (RMA) approach. We applied an intensity based filter to remove probe sets that were not expressed. Differentially expressed genes were identified by fitting a linear model to the data and applying an empirical Bayes smoothing to the standard errors. *P*-values were Benjamini-Hochberg adjusted. Gene ontology analysis was performed using GOrilla (<http://cbl-gorilla.cs.technion.ac.il>)⁵⁴ using a background of expressed genes, a *P*-value threshold of < 10⁻³, and FDR *Q*-value threshold of < 0.1.

Genetic analysis of primary human samples

Archived formalin-fixed, paraffin-embedded (FFPE) primary human endometrial cancer samples and associated clinical data (which was extracted from patient case notes and the secure password-protected hospital database) were deidentified and obtained from the Royal Prince Alfred Hospital in accordance with Sydney Local Health District-approved human ethics protocol (X12-0380). DNA from samples was isolated using the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. DNA concentrations were determined by spectrophotometry (Nanodrop) and 500 ng of DNA was prepared for use on a custom nCounter Copy Number Variation Panel (Nanostring, Seattle, WA, USA) with probes measuring the copy number of cancer-associated genes along the long arm of chromosome 16. Copy number values were normalised to seven normal or benign endometrium FFPE samples isolated from the same location which were considered diploid.

ABBREVIATIONS

CTCF, CCCTC-binding factor; Dox, Doxycycline; FBS, Fetal bovine serum; FFPE, Formalin-fixed paraffin-embedded; HRP, Horseradish peroxidase; SMG, Significantly mutated gene; WT, Wild type; ZF, Zinc finger.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

Financial support was provided by Tour de Cure (Scott Canner Research Fellowship) to CGB and for research grants to ADM, CGB and JEJR; Cancer Council NSW project grants (RG14-09) to JEJR and CGB; (RG14-05) to JEJR; support grants from Cure The Future Foundation and an anonymous foundation; National Institute of General Medical Sciences Cell and Molecular Biology training grant (GM 007067) to RJ. We thank the Centenary Institute Advanced Cytometry Facility for FACS and imaging. The authors acknowledge the University of Sydney HPC service at The University of Sydney for providing HPC, visualisation, database, and/or grid resources that have contributed to the research results reported within this paper. URL: http://sydney.edu.au/research_support/.

AUTHOR CONTRIBUTIONS

CTCF gene sequencing: ADM, CGB, KC, YF and CM. Analysis of previously published CTCF data in endometrial cancer: ADM, CGB, RJ, LD and EM. Generation of lentiviral expression vectors: CGB, CM and PO. Immunofluorescence and imaging: ADM, KC and MV. Cell biology assays: ADM, CGB, KC, MV, MB and PO. Analysis of independent endometrial cancer cohort: ADM, AR and LA. Conception, Supervision and Experimental Design: ADM, CGB and JEJR. Drafting of the article: ADM. Critical revision of manuscript: CGB and JEJR.

REFERENCES

- Haidopoulos D, Simou M, Akrivos N, Rodolakis A, Vlachos G, Fotiou S *et al*. Risk factors in women 40 years of age and younger with endometrial carcinoma. *Acta Obstet Gynecol Scand* 2010; **89**: 1326–1330.
- Ukrainseva SV, Arbeev KG, Yashin AI. Epidemiology of hormone-associated cancers as a reflection of age. *Adv Exp Med Biol* 2008; **630**: 57–71.
- Kurman RJ, Carcangiu ML, Herrington CS, Young RHE. World Health Organization Classification of Tumours of Female Reproductive Organs, 4th edn, IARC: Lyon, France, 2014, p 128.
- Llaurado M, Ruiz A, Majem B, Ertekin T, Colas E, Pedrola N *et al*. Molecular bases of endometrial cancer: new roles for new actors in the diagnosis and the therapy of the disease. *Mol Cell Endocrinol* 2012; **358**: 244–255.
- Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, Shen H *et al*. Integrated genomic characterization of endometrial carcinoma. *Nature* 2013; **497**: 67–73.
- Walker CJ, Miranda MA, O'Hern MJ, McElroy JP, Coombes KR, Bundschuh R *et al*. Patterns of CTCF and ZFX3 mutation and associated outcomes in endometrial cancer. *J Natl Cancer Inst* 2015; **107**: djv249.
- Zigelboim I, Mutch DG, Knapp A, Ding L, Xie M, Cohn DE *et al*. High frequency strand slippage mutations in CTCF in MSI-positive endometrial cancers. *Hum Mutat* 2014; **35**: 63–65.
- Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C *et al*. Mutational landscape and significance across 12 major cancer types. *Nature* 2013; **502**: 333–339.
- Aulmann S, Blaker H, Penzel R, Rieker RJ, Otto HF, Sinn HP. CTCF gene mutations in invasive ductal breast cancer. *Breast Cancer Res Treat* 2003; **80**: 347–352.
- Filippova GN, Qi CF, Ulmer JE, Moore JM, Ward MD, Hu YJ *et al*. Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alters its DNA-binding specificity. *Cancer Res* 2002; **62**: 48–52.
- Mullighan CG, Zhang J, Kasper LH, Lerach S, Payne-Turner D, Phillips LA *et al*. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* 2011; **471**: 235–239.
- Yoshida K, Toki T, Okuno Y, Kanezaki R, Shiraishi Y, Sato-Otsubo A *et al*. The landscape of somatic mutations in Down syndrome-related myeloid disorders. *Nat Genet* 2013; **45**: 1293–1299.
- Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell* 2009; **137**: 1194–1211.
- Marshall AD, Bailey CG, Rasko JE. CTCF and BORIS in genome regulation and cancer. *Curr Opin Genet Dev* 2014; **24**: 8–15.
- Rasko JE, Klenova EM, Leon J, Filippova GN, Loukinov DI, Vatolin S *et al*. Cell growth inhibition by the multifunctional multivalent zinc-finger factor CTCF. *Cancer Res* 2001; **61**: 6002–6007.
- Li T, Lu L. Functional role of CCCTC binding factor (CTCF) in stress-induced apoptosis. *Exp Cell Res* 2007; **313**: 3057–3065.
- Tiffen JC, Bailey CG, Marshall AD, Metierre C, Feng Y, Wang Q *et al*. The cancer-testis antigen BORIS phenocopies the tumor suppressor CTCF in normal and neoplastic cells. *Int J Cancer* 2013; **133**: 1603–1613.
- Lobanenkov VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, Polotskaja AV *et al*. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene* 1990; **5**: 1743–1753.
- Watanabe N, Haruta M, Soejima H, Fukushi D, Yokomori K, Nakadate H *et al*. Duplication of the paternal IGF2 allele in trisomy 11 and elevated expression levels of IGF2 mRNA in congenital mesoblastic nephroma of the cellular or mixed type. *Genes Chromosomes Cancer* 2007; **46**: 929–935.
- Soto-Reyes E, Recillas-Targa F. Epigenetic regulation of the human p53 gene promoter by the CTCF transcription factor in transformed cell lines. *Oncogene* 2010; **29**: 2217–2227.
- Renaud S, Loukinov D, Bosman FT, Lobanenkov V, Benhattar J. CTCF binds the proximal exonic region of hTERT and inhibits its transcription. *Nucleic Acids Res* 2005; **33**: 6850–6860.
- Hirosue A, Ishihara K, Tokunaga K, Watanabe T, Saitoh N, Nakamoto M *et al*. Quantitative assessment of higher-order chromatin structure of the INK4/ARF locus in human senescent cells. *Aging Cell* 2012; **11**: 553–556.
- Lai AY, Fatemi M, Dhasarathy A, Malone C, Sobol SE, Geigerman C *et al*. DNA methylation prevents CTCF-mediated silencing of the oncogene BCL6 in B cell lymphomas. *J Exp Med* 2010; **207**: 1939–1950.
- Witcher M, Emerson BM. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell* 2009; **34**: 271–284.
- Kemp CJ, Moore JM, Moser R, Bernard B, Teater M, Smith LE *et al*. CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep* 2014; **7**: 1020–1029.
- Le Gallo M, O'Hara AJ, Rudd ML, Urick ME, Hansen NF, O'Neil NJ *et al*. Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat Genet* 2012; **44**: 1310–1315.
- Kuramoto H, Tamura S, Notake Y. Establishment of a cell line of human endometrial adenocarcinoma in vitro. *Am J Obstet Gynecol* 1972; **114**: 1012–1019.
- Richardson GS, Dickersin GR, Atkins L, MacLaughlin DT, Raam S, Merk LP *et al*. KLE: a cell line with defective estrogen receptor derived from undifferentiated endometrial cancer. *Gynecol Oncol* 1984; **17**: 213–230.
- Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S *et al*. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012; **483**: 603–607.
- Jung H, Lee D, Lee J, Park D, Kim YJ, Park WY *et al*. Intron retention is a widespread mechanism of tumor-suppressor inactivation. *Nat Genet* 2015; **47**: 1242–1248.
- Nagy E, Maquat LE. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem Sci* 1998; **23**: 198–199.
- Marshall AD, van Geldermalsen M, Otte NJ, Anderson LA, Lum T, Vellozzi MA *et al*. LAT1 is a putative therapeutic target in endometrioid endometrial carcinoma. *Int J Cancer* 2016; **139**: 2529–2539.
- Eritja N, Llobet D, Domingo M, Santacana M, Yeramian A, Matias-Guiu X *et al*. A novel three-dimensional culture system of polarized epithelial cells to study endometrial carcinogenesis. *Am J Pathol* 2010; **176**: 2722–2731.
- Royer C, Lu X. Epithelial cell polarity: a major gatekeeper against cancer? *Cell Death Differ* 2011; **18**: 1470–1477.
- Wong JJ, Ritchie W, Ebner OA, Selbach M, Wong JW, Huang Y *et al*. Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* 2013; **154**: 583–595.
- Nakahashi H, Kwon KR, Resch W, Vian L, Dose M, Stavreva D *et al*. A Genome-wide Map of CTCF Multivalency Redefines the CTCF Code. *Cell Rep* 2013; **3**: 1678–1689.
- Filippova GN, Lindblom A, Meincke LJ, Klenova EM, Neiman PE, Collins SJ *et al*. A widely expressed transcription factor with multiple DNA sequence specificity, CTCF, is localized at chromosome segment 16q22.1 within one of the smallest regions of overlap for common deletions in breast and prostate cancers. *Genes Chromosomes Cancer* 1998; **22**: 26–36.
- Rakha EA, Armour JA, Pinder SE, Paish CE, Ellis IO. High-resolution analysis of 16q22.1 in breast carcinoma using DNA amplifiable probes (multiplex amplifiable probe hybridization technique) and immunohistochemistry. *Int J Cancer* 2005; **114**: 720–729.
- Reardon SN, King ML, MacLean JA 2nd, Mann JL, DeMayo FJ, Lydon JP *et al*. CDH1 is essential for endometrial differentiation, gland development, and adult function in the mouse uterus. *Biol Reprod* 2012; **86**: 141.
- Carico E, Atlante M, Giarnieri E, Raffa S, Bucci B, Giovagnoli MR *et al*. E-cadherin and alpha-catenin expression in normal, hyperplastic and neoplastic endometrium. *Anticancer Res* 2010; **30**: 4993–4997.
- Gonzalez-Rodilla I, Aller L, Llorca J, Munoz AB, Verna V, Estevez J *et al*. The E-Cadherin expression vs. tumor cell proliferation paradox in endometrial cancer. *Anticancer Res* 2013; **33**: 5091–5095.
- Moore JM, Rabaia NA, Smith LE, Fagerlie S, Gurley K, Loukinov D *et al*. Loss of maternal CTCF is associated with peri-implantation lethality of Ctf null embryos. *PLoS One* 2012; **7**: e34915.
- Kumar MS, Pester RE, Chen CY, Lane K, Chin C, Lu J *et al*. Dicer1 functions as a haploinsufficient tumor suppressor. *Genes Dev* 2009; **23**: 2700–2704.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; **144**: 646–674.
- Jiménez-Ayala M, J-AP B. Cytology of the normal endometrium – cycling and postmenopausal. *Endometrial Adenocarcinoma: Prevention and Early Diagnosis*. Karger: Basel, Switzerland, 2008; pp 32–39.
- Allison KH, Reed SD, Voigt LF, Jordan CD, Newton KM, Garcia RL. Diagnosing endometrial hyperplasia: why is it so difficult to agree? *Am J Surg Pathol* 2008; **32**: 691–698.
- Lacey Jr JV, Chia VM. Endometrial hyperplasia and the risk of progression to carcinoma. *Maturitas* 2009; **63**: 39–44.

- 48 Moreno-Bueno G, Portillo F, Cano A. Transcriptional regulation of cell polarity in EMT and cancer. *Oncogene* 2008; **27**: 6958–6969.
- 49 Brar AK, Kessler CA, Meyer AJ, Cedars MI, Jikihara H. Retinoic acid suppresses in-vitro decidualization of human endometrial stromal cells. *Mol Hum Reprod* 1996; **2**: 185–193.
- 50 Cheng YH, Utsunomiya H, Pavone ME, Yin P, Bulun SE. Retinoic acid inhibits endometrial cancer cell growth via multiple genomic mechanisms. *J Mol Endocrinol* 2011; **46**: 139–153.
- 51 Li R, Saito T, Tanaka R, Satohisa S, Adachi K, Horie M *et al*. Hypermethylation in promoter region of retinoic acid receptor-beta gene and immunohistochemical findings on retinoic acid receptors in carcinogenesis of endometrium. *Cancer Lett* 2005; **219**: 33–40.
- 52 Seewaldt VL, Caldwell LE, Johnson BS, Swisshelm K, Collins SJ, Tsai S. Inhibition of retinoic acid receptor function in normal human mammary epithelial cells results in increased cellular proliferation and inhibits the formation of a polarized epithelium in vitro. *Exp Cell Res* 1997; **236**: 16–28.
- 53 Herold MJ, van den Brandt J, Seibler J, Reichardt HM. Inducible and reversible gene silencing by stable integration of an shRNA-encoding lentivirus in transgenic rats. *Proc Natl Acad Sci USA* 2008; **105**: 18507–18512.
- 54 Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 2009; **10**: 48.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies this paper on the Oncogene website (<http://www.nature.com/onc>)