OPEN ACCESS
CellPress

Cell Reports
Report

CTCF Haploinsufficiency Destabilizes DNA Methylation and Predisposes to Cancer

Christopher J. Kemp,[1,*] James M. Moore,[1] Russell Moser,[1] Brady Bernard,[2] Matt Teater,[3] Leslie E. Smith,[1] Natalia A. Rabaia,[1] Kay E. Gurley,[1] Justin Guinney,[4] Stephanie E. Busch,[1] Rita Shaknovich,[3] Victor V. Lobanenkov,[5] Denny Liggitt,[6] Ilya Shmulevich,[2] Ari Melnick,[3] and Galina N. Filippova[1,*]

[1]Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
[2]Institute for Systems Biology, Seattle, WA 98106, USA
[3]Division of Hematology/Oncology, Weill Cornell Medical College, New York, NY 10021, USA
[4]Sage Bionetworks, 1100 Fairview Avenue, Seattle, WA 98109, USA
[5]Molecular Pathology Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, NIH, Rockville, MD 20852, USA
[6]Department of Comparative Medicine, University of Washington, Seattle, WA 98195, USA
*Correspondence: cjkemp@fhcrc.org (C.J.K.), gfilippo@fhcrc.org (G.N.F.)
http://dx.doi.org/10.1016/j.celrep.2014.04.004

SUMMARY

Epigenetic alterations, particularly in DNA methylation, are ubiquitous in cancer, yet the molecular origins and the consequences of these alterations are poorly understood. CTCF, a DNA-binding protein that regulates higher-order chromatin organization, is frequently altered by hemizygous deletion or mutation in human cancer. To date, a causal role for CTCF in cancer has not been established. Here, we show that *Ctcf* hemizygous knockout mice are markedly susceptible to spontaneous, radiation-, and chemically induced cancer in a broad range of tissues. *Ctcf*$^{+/-}$ tumors are characterized by increased aggressiveness, including invasion, metastatic dissemination, and mixed epithelial/mesenchymal differentiation. Molecular analysis of *Ctcf*$^{+/-}$ tumors indicates that *Ctcf* is haploinsufficient for tumor suppression. Tissues with hemizygous loss of *CTCF* exhibit increased variability in CpG methylation genome wide. These findings establish *CTCF* as a prominent tumor-suppressor gene and point to CTCF-mediated epigenetic stability as a major barrier to neoplastic progression.

INTRODUCTION

CTCF (CCCTC-binding factor) is a highly conserved 11 Zn finger, DNA-binding protein that utilizes different combinations of its Zn fingers to bind a large number of highly divergent target sequences throughout the genome (Kim et al., 2007; Nakahashi et al., 2013). CTCF establishes chromatin boundaries and mediates long-range chromatin interactions (Phillips and Corces, 2009). Numerous epigenetic phenomena regulated by CTCF include X chromosome inactivation, imprinting, noncoding transcription, and RNA processing (Filippova, 2008; Ong and Corces, 2014). Further, CTCF binds to target DNA sequences in a DNA-methylation-dependent manner and regulates spreading of DNA methylation (Mukhopadhyay et al., 2004; Wang et al., 2012; Zampieri et al., 2012).

Chromosomal deletion at 16q22.1 is well documented in several human cancers and is one of the most common genetic events in breast cancer, with frequencies ranging from 28% to 90%, depending on the study and molecular subtype (Filippova et al., 1998; Rakha et al., 2006). Extensive genetic and molecular analyses have implicated the involvement of several candidate tumor-suppressor genes within 16q22.1; however, with the exception of *CDH1* (Berx et al., 1996), inactivating "second hit" mutations in other genes are rare, thus hampering efforts to confirm additional candidates. As *CTCF* maps to 16q22.1, we hypothesized that it might be a haploinsufficient tumor-suppressor gene in which inactivation of just one allele would increase cancer risk (Payne and Kemp, 2005). To directly address this possibility, we examined the tumor predisposition of *Ctcf* hemizygous knockout mice.

RESULTS

Ctcf Is a Tumor-Suppressor Gene

Ctcf nullizygous embryos failed to thrive due to cell death by apoptosis, demonstrating that CTCF is indispensable for development (Moore et al., 2012). C57BL6/129 (B6/129) F_1 *Ctcf* $^{+/-}$ mice were born at the expected Mendelian frequency and displayed no overt developmental defects, excepting a slight reduction in mature B cells (data not shown). However, *Ctcf* heterozygous knockout mice were markedly predisposed to spontaneous tumor development in a broad range of tissues. By 100 weeks of age, 80% of *Ctcf* $^{+/-}$ mice succumbed to cancer, compared to only 40% of wild-type littermates (Figure 1A). *Ctcf* $^{+/-}$ mice were also three times more likely to develop multiple tumors per mouse (Table S1). The tumor spectrum associated with *Ctcf* deficiency included benign and malignant uterine tumors, histiocytic sarcomas that presented as aggressive, metastatic disease, and diploid T cell and T cell-infiltrating B cell lymphomas (Figures 1B, 1C, S1, and S2). The latter findings indicate a role for CTCF in lymphocyte maturation

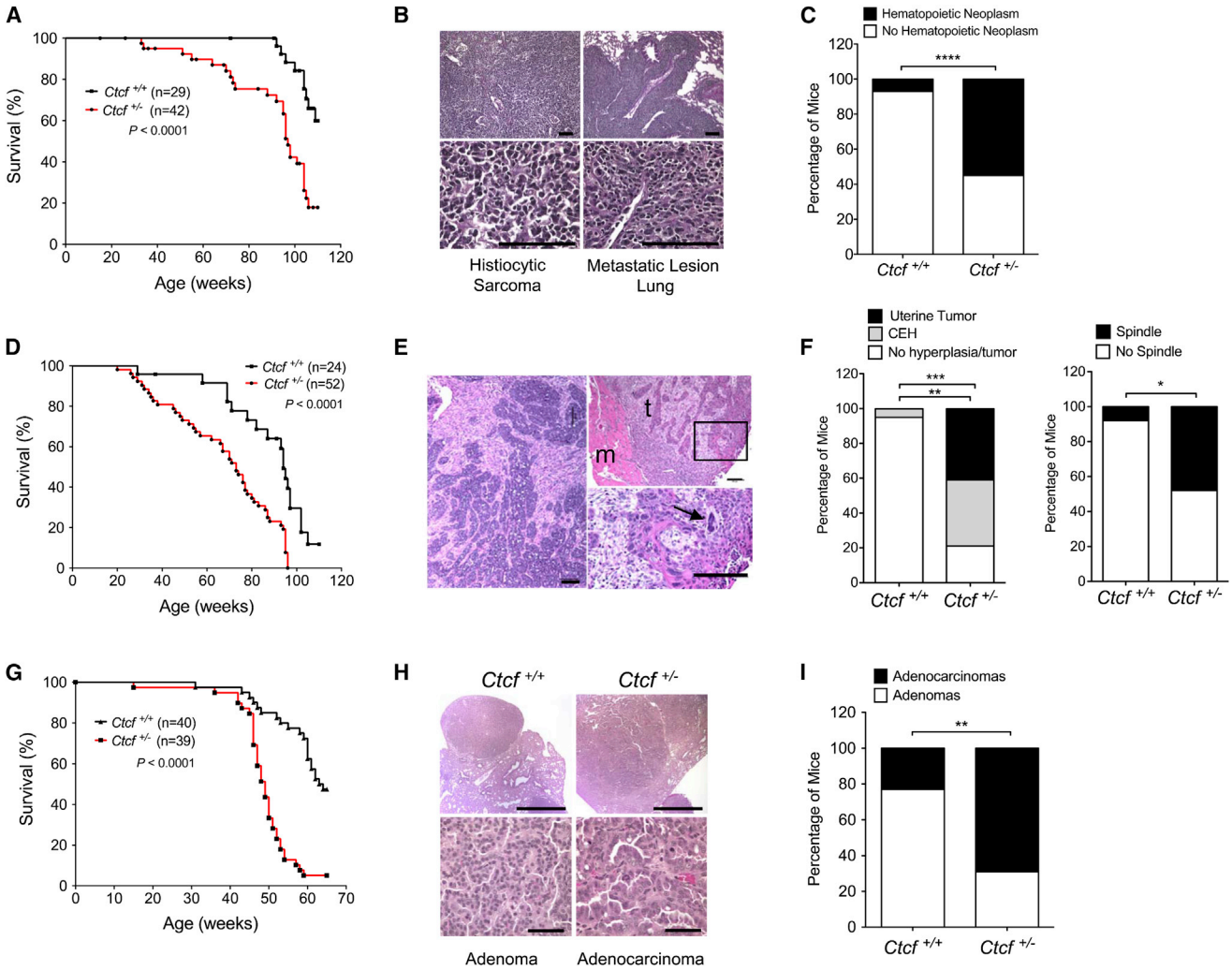


Figure 1. *Ctcf* $^{+/-}$ Mice Are Susceptible to Tumor Development

(A) Kaplan-Meier analysis of tumor-free survival of *Ctcf* $^{+/-}$ (n = 42) and *Ctcf* $^{+/+}$ (n = 29) mice, p < 0.0001.

(B) Hematoxylin and eosin (H&E) staining of spontaneous primary histiocytic sarcoma and corresponding lung metastasis from a *Ctcf* $^{+/-}$ mouse. The scale bars represent 100 μm.

(C) Frequency of hematopoietic neoplasms by *Ctcf* genotype, ****p < 0.0001.

(D) Kaplan-Meier analysis of tumor-free survival of irradiated *Ctcf* $^{+/-}$ (n = 52) and *Ctcf* $^{+/+}$ (n = 24) mice, p < 0.0001.

(E) H&E staining of DMBA-induced tumors from *Ctcf*$^{+/-}$ mice. Left, uterine endometrial adenocarcinoma with epithelial and mesenchymal components. Right, mammary gland adenocarcinoma with areas of spindle cell differentiation (t) and invasion of adjacent skeletal muscle (m); magnification of boxed region (bottom) reveals nuclear atypia (arrowhead). The scale bars represent 100 μm.

(F) Frequency of endometrial lesions (left) and mammary gland histopathology (right) from DMBA-treated mice. Spindle indicates prominent neoplastic spindle cells intermixed within the tumor, **p < 0.01, *p < 0.05.

(G) Kaplan-Meier analysis of tumor-free survival in urethane-treated *Ctcf* $^{+/-}$ (n = 39) and *Ctcf* $^{+/+}$ (n = 40) mice, p < 0.0001.

(H) H&E staining of urethane-induced lung tumors. The scale bars represent 1 mm (top) and 50 μm (bottom).

(I) Distribution of lung adenomas/adenocarcinomas in each genotype, **p = 0.002. See also Figure S1.

and lymphomagenesis, consistent with the reported block in T cell development after conditional deletion of *Ctcf* (Heath et al., 2008) and DNA-methylation-profiling studies of B cell lymphomas (De et al., 2013).

To determine if CTCF functions to suppress ionizing radiation (IR)-induced carcinogenesis, we subjected a second cohort of mice to 4 Gy IR at 2 weeks of age. Irradiated *Ctcf* $^{+/-}$ mice developed tumors earlier than their wild-type littermates (Figure 1D) and displayed a broader range of neoplasms, including thymic and splenic lymphomas, and benign; malignant; and metastatic tumors of the lung, liver, pituitary, ovary, gastrointestinal tract, bone, adrenal cortex, harderian gland, mammary gland, and thyroid gland (Table S2; Figure S1).

Ctcf $^{+/-}$ mice were also predisposed to chemically induced cancers, particularly of the uterus. Approximately 80% of

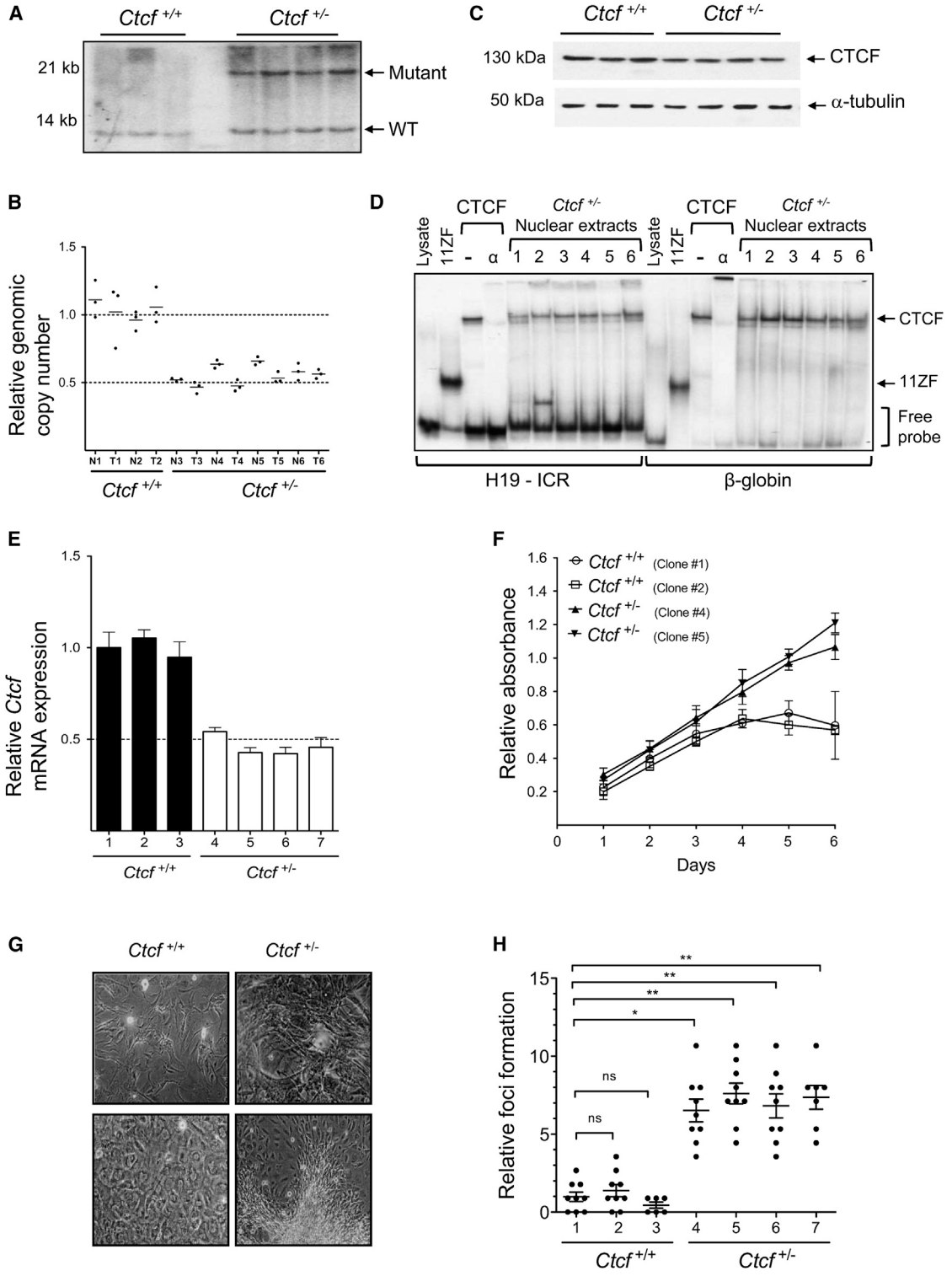


Figure 2. *Ctcf* Is Haploinsufficient for Tumor Suppression

(A) Southern blot analysis of lung tumors from *Ctcf* $^{+/+}$ (lanes 1–3) and *Ctcf* $^{+/-}$ (lanes 4–7) mice. WT, wild-type allele.

(B) qPCR analysis of genomic DNA from normal lung (N) and lung tumors (T) from each genotype.

(C) Immunoblot analysis of CTCF protein in lung tumors from *Ctcf* $^{+/+}$ (lanes 1–3) and *Ctcf* $^{+/-}$ (lanes 4–7) mice. α-tubulin served as loading control.

(legend continued on next page)

7,12-dimethylbenz[a]anthracene (DMBA)-treated *Ctcf* $^{+/-}$ female mice developed uterine lesions, ranging from cystic endometrial hyperplasias (CEH) to leiomyomas/sarcomas and highly aggressive endometrial adenocarcinomas, as compared to only a 5% incidence of CEH in wild-type mice (Figures 1E, 1F, and S1). DMBA-treated female mice also developed mammary gland tumors that, on a wild-type background, were typically adeno- or adenosquamous carcinomas. In contrast, *Ctcf*$^{+/-}$ mice exhibited a broader histopathological spectrum, ranging from adenosquamous carcinomas to more-primitive-appearing tumors with mixed luminal epithelial and mesenchymal components. Importantly, tumor cells also exhibited a pleomorphic spindle morphology consistent with myoepithelial derivation and/or epithelial-mesenchymal transition (EMT). Cells with spindle morphology showed reduced E-cadherin and increased vimentin staining, suggesting that *Ctcf* hemizygosity affects transdifferentiation of these tumors (Figures 1E, 1F, and S1). DMBA-treated *Ctcf* $^{+/-}$ female mice also developed tumors of the lung, gastrointestinal tract, skin, and ovary (Table S3). Thus, hemizygous deletion of *Ctcf* sensitizes a broad spectrum of cell lineages and tissues to spontaneous, radiation-, and chemically induced cancers, establishing CTCF as a pan-tissue tumor suppressor.

CTCF Suppresses *Kras*-Driven Lung Cancer

To address the interaction between an epigenetic regulator and a genetic driver of cancer, we next asked if reduction of *Ctcf* cooperates with mutated *Kras* in a model of urethane-induced non-small cell lung carcinoma (NSCLC). These tumors closely resemble human NSCLC in morphologic and molecular characteristics, and over 80% harbor activating mutations in the *Kras* oncogene (Gurley et al., 2014). Urethane-treated *Ctcf* $^{+/-}$ mice developed more lung tumors that were significantly larger compared to wild-type mice (Figure S1). This enhanced lung tumor burden led to earlier mortality (Figure 1G). The majority of lung tumors from wild-type mice were benign adenomas (17/22; 77%) with low mitotic activity, uniform small nuclei, and well-defined boundaries. In contrast, 69% (22/32) of *Ctcf* $^{+/-}$ lung tumors were classified as malignant adenocarcinomas with abundant mitotic activity, large and irregularly shaped nuclei, disorganized growth patterns, and frequent invasion into local parenchyma and airways (Figures 1H and 1I). Lung tumors from *Ctcf* $^{+/-}$ mice exhibited increased proliferation as measured by bromodeoxyuridine (BrdU) labeling (Figure S1), whereas apoptosis was negligible in both genotypes (not shown). Thus, reduction of *Ctcf* accelerated the development of *Kras*-driven cancer, arguing that epigenetic events regulated by CTCF play a significant role in suppressing RAS-mediated tumorigenesis.

Ctcf Is Haploinsufficient for Tumor Suppression

Some tumor-suppressor genes are recessive and require a second "hit" for abrogation of function (Payne and Kemp, 2005). However, complete loss of *Ctcf* leads to apoptotic cell death (Moore et al., 2012) and therefore is unlikely to provide a selective advantage. Southern blot analysis and quantitative PCR (qPCR) showed retention of the wild-type *Ctcf* allele in 100% (4/4) of *Ctcf*$^{+/-}$ lung tumors (Figures 2A and 2B). RT-PCR and immunoblot analysis showed full-length *Ctcf* mRNA transcript and CTCF protein were maintained in both tumors and normal tissue (Figure 2C; not shown). Sequencing of *Ctcf* cDNA from 19 representative tumors from spontaneous, irradiated, and urethane-treated *Ctcf* $^{+/-}$ mice revealed no mutations throughout the entire coding region (Table S4). No mutations or deletions in *Ctcf* were observed in five tumors from *Ctcf* wild-type mice. Gel mobility shift analysis of CTCF DNA-binding activity in nuclear extracts confirmed retention of functional CTCF in 6/6 (100%) of *Ctcf* $^{+/-}$ tumors (Figure 2D).

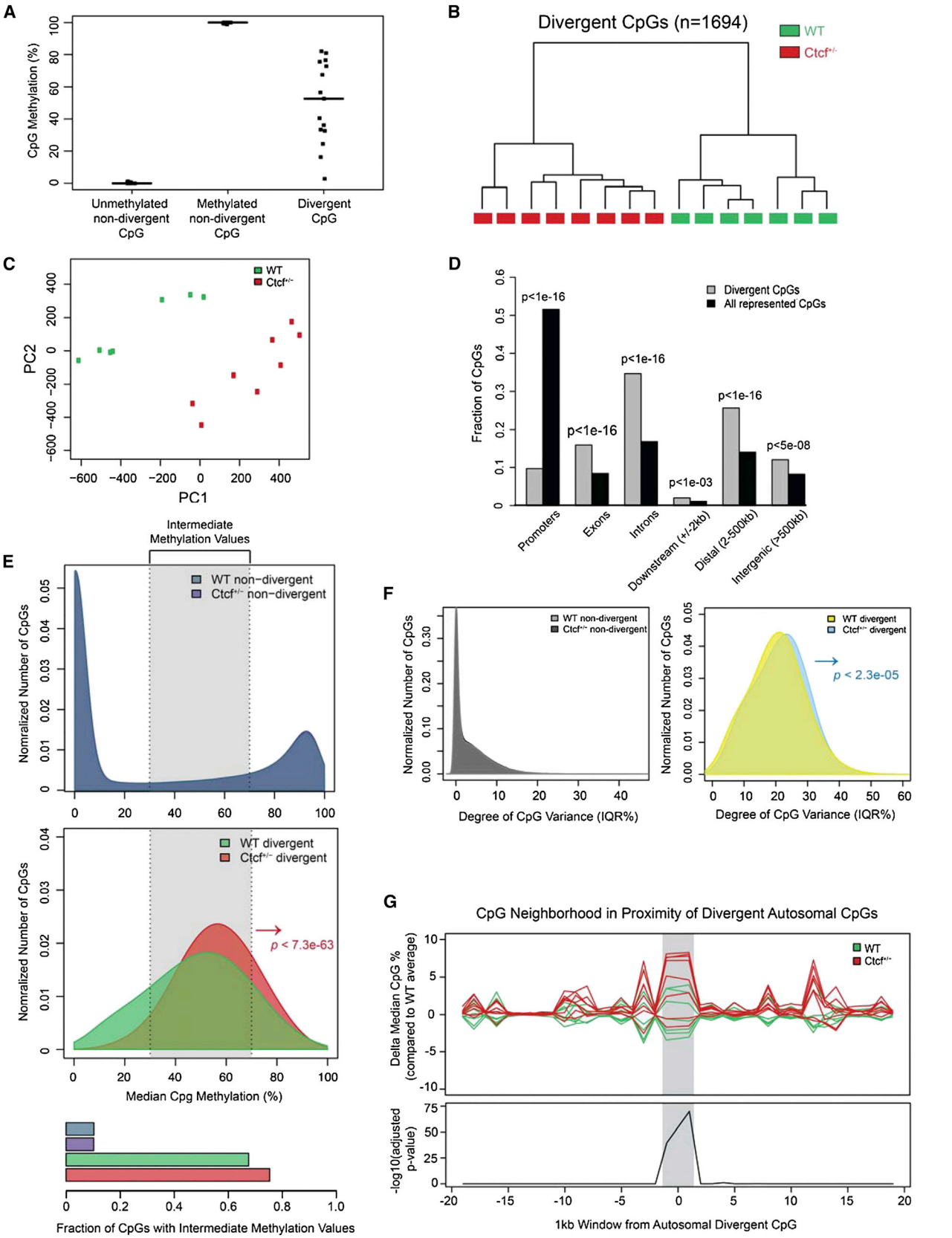
As ectopic expression of CTCF inhibits cell growth (Rasko et al., 2001), we next asked if *Ctcf* had a gene-dosage-dependent effect on cell proliferation. Mouse-embryo-derived fibroblasts (MEFs) from wild-type mice stopped proliferating at confluence, forming a uniform monolayer with a flattened morphology. In contrast, *Ctcf* $^{+/-}$ MEFs that showed a 50% reduction in *Ctcf* mRNA transcript levels (Figure 2E) continued to proliferate after reaching confluence, piling up and forming foci indicating loss of contact inhibition (Figures 2F–2H).

Collectively, the increased tumor predisposition of *Ctcf* $^{+/-}$ mice; the retention of the wild-type *Ctcf* allele, mRNA, and protein expression in tumors; and the cell autonomous loss of contact inhibition of *Ctcf* $^{+/-}$ cells establish this gene as a haploinsufficient tumor suppressor, where loss of a single allele significantly increases cancer risk.

Ctcf Hemizygosity Causes DNA Methylation Instability In Vivo

CTCF can affect cytosine methylation both locally, through binding to chromatin boundaries, and distally, through its long-range effects on 3D chromatin architecture (Zampieri et al., 2012; Wang et al., 2012). Aberrant methylation at CTCF-binding sites in several tumor-suppressor genes and imprinted loci has been reported in human cancer (Filippova, 2008). To determine if hemizygous loss of *Ctcf* affects DNA methylation in vivo, we examined the methylation status of known CTCF-binding sites at the promoters of three tumor-suppressor genes (*p16*^{Ink4a}, *p19*^{Arf}, and *Mlh1*), as well as at the imprinting control regions (ICRs) of three imprinted loci (*Igf2/H19*, *KvDMR1*, and *Rasgrf1*). Bisulfite conversion and sequencing analysis of wild-type and *Ctcf*$^{+/-}$ MEFs, as well as paired normal lung and lung tumor

(D) Gel shift analysis of nuclear extracts from *Ctcf* $^{+/-}$ lung tumors show CTCF binding at both the H19/Igf2 ICR and β-globin insulator FII loci. Positions of protein-DNA complexes with 11ZF CTCF DNA-binding domain (11ZF) or full-length CTCF protein (CTCF) are indicated. α-CTCF antibody (α) was used to supershift CTCF-DNA complexes.
(E) qRT-PCR analysis of Ctcf mRNA in *Ctcf* $^{+/+}$ (n = 3) and *Ctcf* $^{+/-}$ (n = 4) MEFs, mean ± SEM.
(F) Proliferation of *Ctcf* $^{+/+}$ and *Ctcf* $^{+/-}$ MEFs. Assays performed in triplicate for two clones, each genotype; mean ± SD.
(G) Foci formation in MEFs cultured from *Ctcf* $^{+/-}$ compared to *Ctcf* $^{+/+}$ mice.
(H) Relative MEF foci formation (# foci/10 cm plate) from *Ctcf* $^{+/+}$ (n = 3) and *Ctcf* $^{+/-}$ (n = 4) mice. Each dot represents an independent experiment plotted as mean ± SEM relative to the mean wild-type value; *p < 0.05, **p < 0.01.



(legend on next page)

tissue samples from three wild-type and four *Ctcf*$^{+/-}$ mice revealed that, of the six loci examined, one, the *Igf2/H19* ICR, showed differences in DNA methylation. The maternally inherited *Igf2/H19* ICR allele is normally unmethylated and bound by CTCF, whereas the paternal ICR is methylated and does not bind CTCF (Bell and Felsenfeld, 2000). Both normal lung and MEFs from wild-type mice showed the expected pattern of 50% methylated CpGs at two CTCF-binding sites analyzed within the *Igf2/H19* ICR. In contrast, *Ctcf* $^{+/-}$ lung tissue and MEFs exhibited between 80% and 90% CpG methylation at these sites (Figure S3).

To explore the effect of hemizygous deletion of *Ctcf* on DNA methylation genome wide, we profiled noncancerous murine lung tissue. We examined 1.75 M CpGs by enhanced reduced representation bisulfite sequencing (ERRBS), a method based on bisulfite conversion of DNA followed by deep sequencing of MspI fragments with size-based enrichment for genes and regulatory regions (Akalin et al., 2012). We first performed an unsupervised analysis based on the CpGs with the greatest degree of divergence among the 15 lung samples examined (n = 1,694 CpGs; interquartile range > 25%) (Figure 3A; Table S5). This unbiased analysis yielded two main clusters, corresponding precisely to the seven and eight samples from *Ctcf* $^{+/+}$ and *Ctcf* $^{+/-}$ mice, respectively (Figures 3B, 3C, and S3C). This suggests that the most-prominent differences in DNA methylation between these tissues are directly attributed to *Ctcf* hemizygosity. Outside of these divergent CpGs, the overall CpG methylation patterns were remarkably stable and consistent between genotypes (Figure S3D), implying that CTCF regulation of DNA methylation occurs primarily at specific loci. The divergent CpGs were nonrandomly distributed in the genome with preferential localization within introns, exons, and intergenic regions and depletion from promoter regions (Figure 3D).

The divergent CpGs exhibited intermediate methylation levels (ranging from 30% to 70%), contrasting with the classical bimodal distribution among the nonvariable CpGs (Figure 3E). This indicates a high degree of intrasample heterogeneity among these divergent CpGs, which might represent hot spots for epigenetic hypervariability in lung tissue. Divergent CpGs from *Ctcf*$^{+/-}$ lungs were significantly shifted toward cytosine hypermethylation (Figure 3E, middle panel) and showed increased variance in methylation (Figure 3F), further evidence that *Ctcf* hemizygosity destabilized the regulation of DNA methylation. Whereas DNA methylation changes between genotypes occurred primarily at the set of divergent CpGs, a modest overall gain in genome-wide DNA methylation was also observed in *Ctcf*$^{+/-}$ lung (Figure S3E).

Although the methylation differences between *Ctcf* genotypes were striking, these changes would unlikely have biological relevance if they were limited to individual CpGs. Hence, to test whether CTCF also affects regions surrounding the divergent CpGs, the median methylation values within each lung sample, relative to the average wild-type methylation values, were determined for all CpGs occurring within 1 kb windows extending up- and downstream of the divergent CpGs. We found significant hypermethylation in *Ctcf*$^{+/-}$ samples that extended to ~2 kb regions surrounding divergent CpGs (Figure 3G). Interestingly, several cancer-associated genes such as *Trp53*, *Dnmt3A*, *RunX1*, *Alk*, *Card11*, *Kit*, *Mpo*, *Pou2F2*, *Spen*, *ZfhX3*, and *Arid1A*, including *Ctcf* itself, contain divergent CpGs either within or in close proximity to the gene. *Ctcf*$^{+/-}$ lung cytosine methylation patterns therefore reflect discrete regional increases in methylation diversity, with a significant tendency of these regions to shift toward a hypermethylated state, suggesting a role for CTCF in maintaining stability of cytosine methylation patterning in the genome.

CTCF Mutation and Copy Number Variation in Human Breast and Endometrial Cancer

The functions of tumor-suppressor genes are typically well conserved between mice and humans. Our mouse model results, together with the frequent 16q22.1 deletions reported in breast and other cancers (Rakha et al., 2006), predict that CTCF is a tumor suppressor in human cancer. Indeed, analysis of The Cancer Genome Atlas (TCGA) data revealed reduced gene copy number of *CTCF* in 276/484 (57%) of breast tumors (Figure 4A). Tumors with reduced *CTCF* DNA copy number showed significantly reduced levels of CTCF mRNA (Figure 4C). In addition, point mutations in *CTCF* were identified in 21 out of 772 tumors with 17 of these mapping to the protein sequence (Figure 4E). *CTCF* mutations predominantly occurred in the estrogen-receptor-positive luminal A subtype, where the frequency was 4% (Cancer Genome Atlas Network, 2012). Although complete loss of CTCF function is incompatible with cell survival, seven breast tumors showed evidence of both *CTCF* copy number reduction and mutation (Figure 4A; Table S6). One had a truncating "loss of function" mutation (E21X), possibly indicating two separate subclonal events. The other six mutations were either silent or missense mutations and clustered in the CTCF Zn finger (ZF) DNA-binding domain, particularly in ZF1 (Figures 4E, 4G, and S4). Previous analysis of similar mutations (R283C and H284N/P/Y) that disrupt either DNA sequence recognition or zinc coordination showed that these mutations selectively impaired CTCF binding to some, but not all, DNA target sites,

Figure 3. Aberrant DNA Methylation in Noncancerous Lung Tissue from *Ctcf* $^{+/-}$ Mice

(A) Examples of nondivergent and divergent CpGs within 15 lung samples.

(B) Hierarchical clustering of divergent autosomal CpG methylation values.

(C) Principle component analysis plot for divergent autosomal CpG methylation values.

(D) Genomic distribution of divergent autosomal CpGs versus all represented CpGs.

(E) Density plots of median CpG methylation values according to *Ctcf* genotype; nondivergent autosomal CpGs (top) and divergent autosomal CpGs (middle), $p < 7.3 \times 10^{-63}$). Divergent CpGs tend to have intermediate methylation values (bottom).

(F) Density plots of CpG methylation variance according to *Ctcf* genotype; nondivergent autosomal CpGs (left) and divergent autosomal CpGs (right), $p < 2.3 \times 10^{-5}$).

(G) CpG methylation values within 1 kb intervals up- and downstream of divergent CpGs. The differences of median CpG methylation values for each sample relative to the median of all seven wild-type CpG methylation values and associated p values are plotted. Shaded area indicates region with significant difference between genotypes, $p < 1 \times 10^{-25}$, ANOVA.

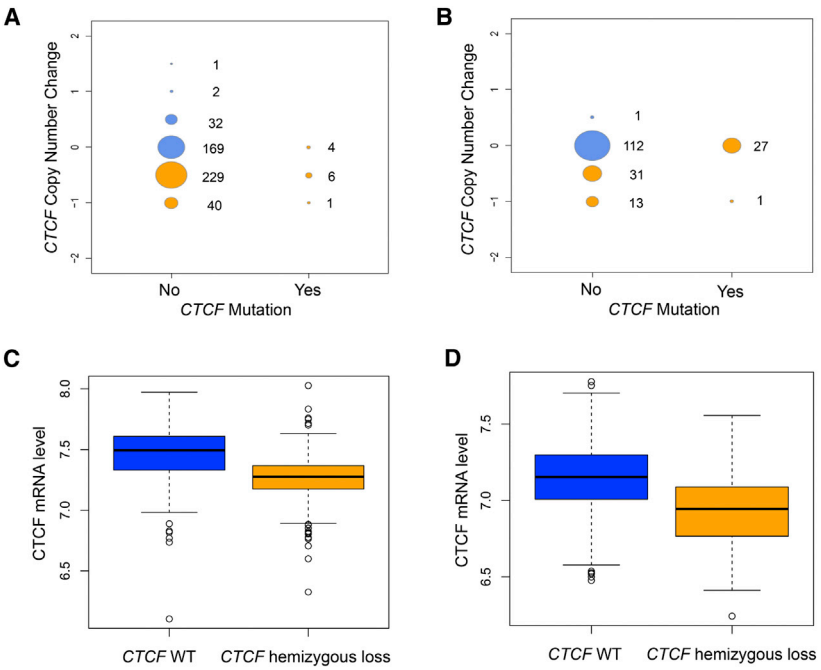


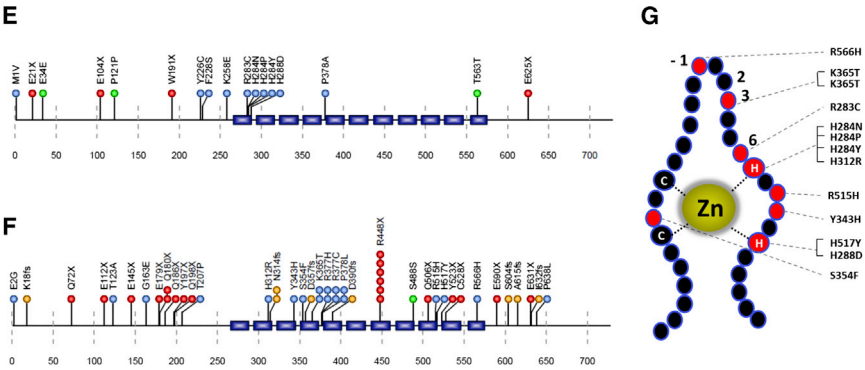
Figure 4. *CTCF* Hemizygous Deletion and Mutation in Human Breast and Endometrial Cancer

(A and B) Size plots for breast invasive carcinoma (A) and uterine corpus endometrioid carcinoma (B), indicating the total number of samples with either copy number change or mutation (yellow) or samples diploid for *CTCF* (blue). Copy number values are presented in discrete increments of 0.5.

(C and D) Relative CTCF mRNA levels in breast (n = 856; $p < 10 \times 10^{-16}$) (C) and endometrial (n = 362; $p < 7.15 \times 10^{-12}$) (D) tumors are plotted according to *CTCF* DNA copy number.

(E and F) Somatic mutations within the *CTCF* protein coding sequence are plotted based on amino acid position (Uniprot Identifier P49711) for breast carcinoma (E) and endometrioid carcinoma (F). Synonymous (green), missense (blue), frameshift InDels (gold), and nonsense (red) mutations are shown. Blue rectangles indicate 11 Zn finger domains.

(G) Rendition of a typical C2-H2 type zinc finger (ZF) showing composite of missense mutations from endometrial and breast cancers. Amino acids at positions −1, 2, 3, and 6 that contact DNA directly and histidine (H) and cysteine (C) residues that coordinate Zn are indicated.



CTCF Mutational Status Correlates with Whole-Genome Methylation Patterns in Human Tumors

As our mouse model demonstrated that a 50% reduction in *Ctcf* gene dosage altered DNA methylation patterns genome wide, we next asked if DNA methylation was altered in human tumors with *CTCF* hemizygous deletions or mutations. We analyzed Illumina Infinium DNA methylation data generated by the TCGA, a platform which assays methylation status of 27,000 CpGs (Kandoth et al., 2013). Endometrial adenocarcinomas with *CTCF* copy number aberrations (CNA) or point mutations exhibited significantly distinct methylation patterns, with a subset of CpGs showing either an increase or decrease in methylation compared to *CTCF* diploid tumors (Figures 5A and 5B; Table S7). *CTCF* CNA tumors tended to have a greater number of CpG methylation differences as compared to tumors with *CTCF* point mutations. This could reflect a broader destabilization associated with loss of additional genes on chromosome 16q in the CNA tumors, as, for example, seen for CpGs located on chromosome 16 itself and/or a more restrictive phenotype associated with individual *CTCF* point mutations. In either case, these hyper- and hypomethylated CpGs were distributed across the genome. Interestingly, of the 16 CpGs assayed at the *H19* locus, 12 showed a significant increase in methylation in *CTCF* CNA tumors (Table S7), a finding that is consistent with the increased methylation seen at the *Igf2/H19* ICR in murine *Ctcf* $^{+/-}$ tissues (Figure S3B). Likewise, in luminal A breast cancer, a subset of CpGs showed

consistent with the multivalent nature of this 11 Zn finger protein (Filippova et al., 2002; Nakahashi et al., 2013).

Hemizygous deletions of *CTCF* were also observed in uterine endometrial cancer (45/185; 24%), which again correlated with significantly reduced levels of CTCF mRNA (Figures 4B and 4D). Point mutations were also frequent (53/248; 21%), primarily in the type 1 endometrioid subtype (Kandoth et al., 2013). Of these tumor-specific mutations, 48/53 mapped to the CTCF protein sequence and included missense mutations localized to the ZF domain and predicted to alter DNA sequence recognition, as well as truncating and frameshift mutations predicted to delete some or all of the ZFs (Figures 4F, 4G, and S4). Only one tumor with a missense mutation (R377C in the ZF domain) showed evidence of both copy number loss and mutation (Table S6). Overall, *CTCF* is ranked as the 4th- and 16th-most significantly mutated gene in endometrial and breast cancer, respectively, comparable to well-known cancer genes such as *PTEN*, *TP53*, *PIK3CA*, and *FBXW7* (Figure S4).

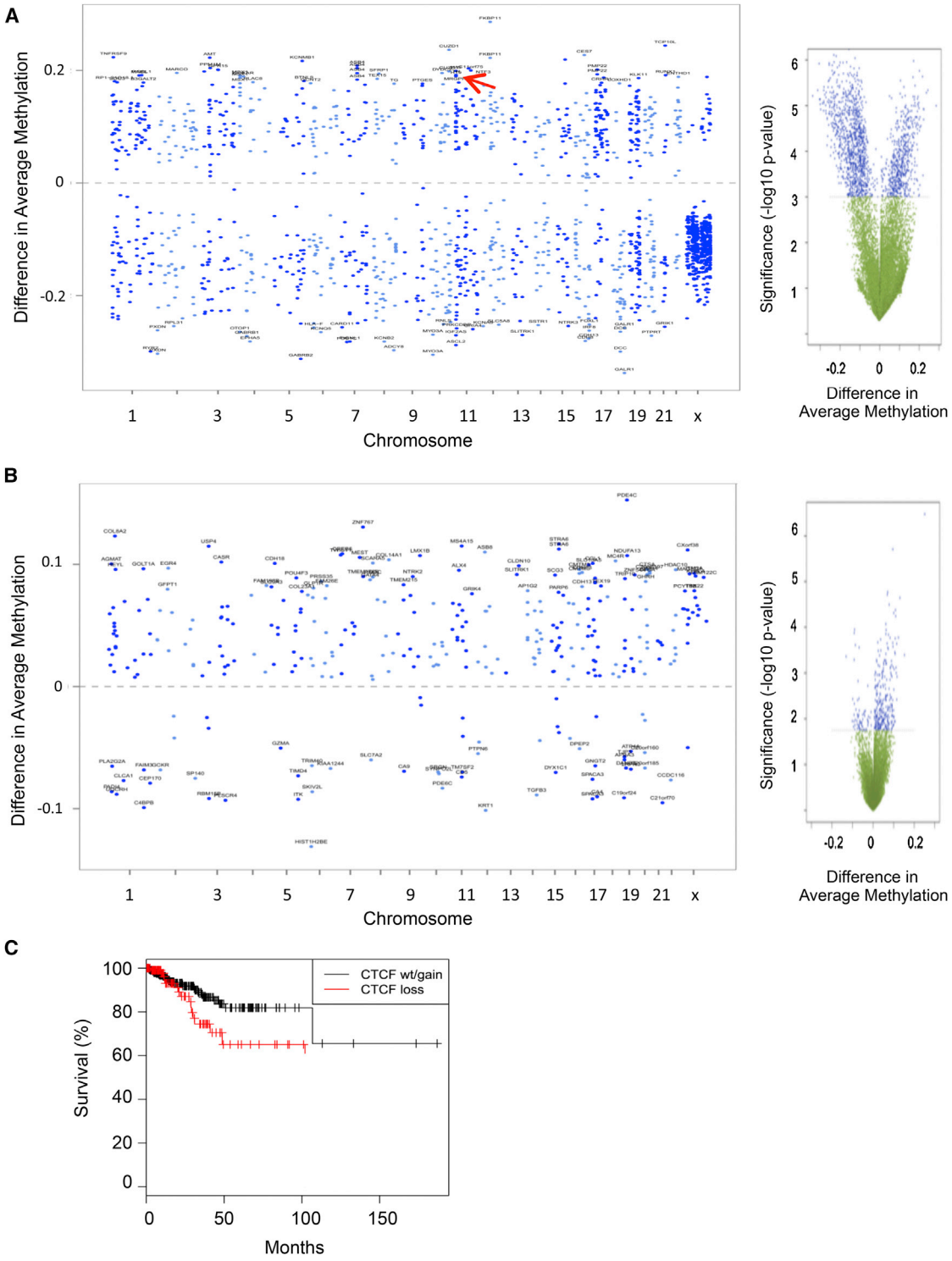


Figure 5. *CTCF* Status Correlates with Genome-wide DNA Methylation Patterns and Patient Survival in Endometrial Cancer

(A) Chromosomal plot of significantly differentially methylated CpGs between *CTCF* CNA (n = 45) versus *CTCF* diploid (n = 114) endometrial tumors, $p < 0.001$. Positive and negative differences indicate methylation probes with increased or decreased methylation in the *CTCF* CNA tumors, respectively. The H19 locus is indicated with a red arrow. Right, volcano plot of permutation-based significance (blue, $p < 0.001$) as a function of differences in average DNA methylation between *CTCF* CNA versus wild-type endometrial tumors (beta value differentials).

(legend continued on next page)

significantly altered methylation in *CTCF* CNA or mutant tumors compared to *CTCF* wild-type tumors (Figure S5; Tables S7).

The accelerated cancer-associated mortality and increased aggressiveness of tumors in *Ctcf*$^{+/-}$ mice prompted us to ask if *CTCF* copy number in human tumors was associated with survival. In the TCGA breast cancer cohort, which included all subtypes, we observed no significant association with overall survival (not shown), whereas in endometrial cancer, reduction of *CTCF* copy number was associated with poor survival (Figure 5C). Although a careful analysis of different tumor subtypes with and without *CTCF* mutation or copy number variation will be required, these data suggest that, in addition to a distinct epigenetic profile, *CTCF* hemizygous loss in human tumors might also correlate with a distinct clinical outcome.

DISCUSSION

Here, we have shown that the chromatin organizer CTCF is a haploinsufficient tumor suppressor in vivo, where loss of just one allele enhances both tumor formation and malignant progression. *Ctcf* $^{+/-}$ mice were predisposed to spontaneous, ionizing radiation and chemically induced tumors of epithelial, mesenchymal, and hematopoietic origin, indicating a broad role for CTCF in tumor suppression. Furthermore, tumors from *Ctcf* hemizygous mice were more aggressive, with frequent local invasion, metastatic dissemination, features of EMT, and mixed lineage differentiation indicating reduction of *Ctcf* enhances malignant progression.

Our functional studies in mice, when viewed together with frequent *CTCF* hemizygous deletions or point mutations found in human tumors, implicate *CTCF* as a major tumor-suppressor gene in human cancer. Indeed, analysis of 4,742 tumors across 21 cancer types revealed that *CTCF* was one of the most significantly mutated genes (Lawrence et al., 2014).

Estrogen signaling is a significant risk factor for both estrogen-receptor-positive luminal A breast cancer and type 1 endometrial cancer. The particularly high frequency of *CTCF* aberrations in these tumors, together with the susceptibility of *Ctcf* $^{+/-}$ mice to endometrial tumors and pleomorphic breast tumors, suggests a prominent role for CTCF in hormone-driven cancers. Given that CTCF is a negative regulator of FOXA1-chromatin interactions that are required for estrogen receptor activity (Hurtado et al., 2011), it is possible that *CTCF* disruption contributes to tumor development by enhancing the promoting effects of estrogen.

Our study also has implications for understanding the origins of DNA methylation alterations in cancer. That hemizygous loss of *Ctcf* destabilized DNA methylation at epigenetically variable CpGs in normal tissue and enhanced cancer progression suggests that epigenetic instability may both precede and accelerate the evolution of cancer. Indeed, human endometrial and breast tumors with genetic disruption of *CTCF* exhibited a distinct pattern of DNA methylation relative to *CTCF* intact tumors. Whether these methylation changes are a direct consequence of *CTCF* disruption remains unclear; however, the impact of *Ctcf* hemizygosity on DNA methylation profiles that we observed in mice suggests this as a strong possibility. Overall, our data support a model wherein *CTCF* hemizygous deletion or mutation leads to epigenetic instability that in turn enhances phenotypic plasticity and thereby accelerates the emergence, adaptation, and evolution of neoplastic lesions. Our findings further suggest that human tumors with *CTCF* disruption might manifest as discrete epigenetic subtypes with clinically distinct outcomes and potentially unique therapeutic opportunities.

EXPERIMENTAL PROCEDURES

Tumor Induction and Analysis

All experiments with mice were approved by the Fred Hutchinson Cancer Research Center animal care and use committee and performed according to Institutional Animal Care and Use Committee regulations. Experimental mice were generated by crossing 129/sv *Ctcf* $^{+/-}$ mice to C57BL6 *Ctcf* $^{+/+}$ mice. Isogenic C57BL6/129 F_1 *Ctcf* $^{+/-}$ and *Ctcf* $^{+/+}$ offspring were genotyped (Moore et al., 2012), maintained on standard lab chow and water ad libitum, and housed in a 12 hr light-dark cycle. A second cohort was generated and mice were exposed at 2 weeks of age to 400 cGy ionizing radiation (^{137}Cesium at 500 cGy/min). A third cohort was given 1 mg of DMBA dissolved in 200 μl of sesame oil by gavage once a week for 6 weeks. A fourth was injected at 14 to 15 days of age with a single dose of urethane (1 mg/g body weight, intraperitoneally). Eight urethane-treated mice of each genotype were injected with BrdU (100 mg/kg; Sigma) 1 hr before sacrifice. For all experiments, mice were observed on a daily basis through 24 months and euthanized when moribund. Mice were necropsied, and all grossly visible tumors and surrounding normal tissues were fixed in formalin or frozen.

Statistical Analysis

Tumor development and patient survival data were analyzed using Kaplan Meier survival plots with log rank test for significance. The mean tumor multiplicity and tumor incidence at each anatomic site were compared between the genotypes using the Mann-Whitney nonparametric test and Fisher's exact test, respectively. Significant differences in foci formation were assessed via a Kruskal Wallis test with Dennett's posttest for multiple comparisons. CTCF mRNA expression was compared to DNA copy number using Wilcoxon rank sum test. Data are presented as either mean ± SD or mean ± SEM as indicated, and asterisks depicted in figures represent statistical significance.

Molecular Biology Protocols

See Supplemental Experimental Procedures for detailed protocols on immunohistochemistry, immunophenotyping and cell cycle analysis of lymphomas, Southern blotting, quantitative RT-PCR (qRT-PCR), DNA sequencing, electrophoretic mobility shift assay, bisulfite sequencing analysis of DNA methylation, and experiments on MEFs.

ERRBS

Genome-wide methylation analysis was performed on noncancerous murine lung tissue at approximately 48 weeks after urethane treatment, at which time no histological differences were seen in normal lungs between *Ctcf* genotypes. Five hundred nanograms of DNA from eight *Ctcf* $^{+/-}$ mice (four male, four female) and seven wild-type mice (three male, four female) was processed by the ERRBS protocol (Akalin et al., 2012) as described in Supplemental Experimental Procedures.

(B) Chromosomal plot of significantly differentially methylated CpGs ($p < 0.018$) between *CTCF* mutant ($n = 45$) versus *CTCF* diploid ($n = 114$) endometrial tumors. Note, COL14A1 at chr8 with a differential methylation value of 0.25 was deleted for clarity (see Table S7). Right, volcano plot of permutation-based significance (blue, $p < 0.018$) as a function of differences in average DNA methylation between *CTCF* mutant versus wild-type tumors (beta value differentials).
(C) Kaplan Meier survival plot of patients with endometrial cancer ($n = 492$) stratified by *CTCF* copy number ($p < 0.05$; log rank test).

TCGA Analysis

Mutation annotation format files containing somatic mutation calls for breast-invasive carcinoma and uterine corpus endometrioid carcinoma were downloaded from the TCGA Data Coordination Center. The GISTIC2 algorithm was used to identify samples with significant somatic gains or losses of *CTCF*. Methylation beta values were calculated as the fraction (ranging from 0 to 1) of methylated beads from the Illumina Infinium DNA methylation platform, taken from Level 3 TCGA data as described in Supplemental Experimental Procedures.

ACCESSION NUMBERS

Data are deposited in the National Center for Biotechnology Information's Gene Expression Omnibus (GEO) and are accessible through GEO accession number GSE48975.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.04.004.

ACKNOWLEDGMENTS

This work was supported by NIH grant numbers CA68360, ES007033, DA030326, U24CA143835, U01 CA141550, U01 CA176303, and PHS NRSA T32 GM007270 from NIGMS. We thank S. Collins, C. Grandori, S. Henikoff, S. Tapscott, and F. Yang for comments on the manuscript, and the William Guy Forbeck Research Foundation.

Received: December 27, 2013
Revised: March 5, 2014
Accepted: April 3, 2014
Published: May 1, 2014

REFERENCES

Akalin, A., Garrett-Bakelman, F.E., Kormaksson, M., Busuttil, J., Zhang, L., Khrebtukova, I., Milne, T.A., Huang, Y., Biswas, D., Hess, J.L., et al. (2012). Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. PLoS Genet. *8*, e1002781.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482–485.

Berx, G., Cleton-Jansen, A.M., Strumane, K., de Leeuw, W.J., Nollet, F., van Roy, F., and Cornelisse, C. (1996). E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. Oncogene *13*, 1919–1925.

Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70.

De, S., Shaknovich, R., Riester, M., Elemento, O., Geng, H., Kormaksson, M., Jiang, Y., Woolcock, B., Johnson, N., Polo, J.M., et al. (2013). Aberration in DNA methylation in B-cell lymphomas has a complex origin and increases with disease severity. PLoS Genet. *9*, e1003137.

Filippova, G.N. (2008). Genetics and epigenetics of the multifunctional protein CTCF. Curr. Top. Dev. Biol. *80*, 337–360.

Filippova, G.N., Lindblom, A., Meincke, L.J., Klenova, E.M., Neiman, P.E., Collins, S.J., Doggett, N.A., and Lobanenkov, V.V. (1998). 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 *22*, 26–36.

Filippova, G.N., Qi, C.F., Ulmer, J.E., Moore, J.M., Ward, M.D., Hu, Y.J., Loukinov, D.I., Pugacheva, E.M., Klenova, E.M., Grundy, P.E., et al. (2002). Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter tts DNA-binding specificity. Cancer Res. *62*, 48–52.

Gurley, K.E., Moser, A.R., and Kemp, C.J. (2014). Induction of lung tumors in mice with urethane. In Mouse Models of Cancer, C. Abate-Shen, K. Politi, L.A. Chodosh, and K.P. Olive, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory), pp. 63–65.

Heath, H., Ribeiro de Almeida, C., Sleutels, F., Dingjan, G., van de Nobelen, S., Jonkers, I., Ling, K.W., Gribnau, J., Renkawitz, R., Grosveld, F., et al. (2008). CTCF regulates cell cycle progression of alphabeta T cells in the thymus. EMBO J. *27*, 2839–2850.

Hurtado, A., Holmes, K.A., Ross-Innes, C.S., Schmidt, D., and Carroll, J.S. (2011). FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat. Genet. *43*, 27–33.

Kandoth, C., Schultz, N., Cherniack, A.D., Akbani, R., Liu, Y., Shen, H., Robertson, A.G., Pashtan, I., Shen, R., Benz, C.C., et al.; Cancer Genome Atlas Research Network (2013). Integrated genomic characterization of endometrial carcinoma. Nature *497*, 67–73.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231–1245.

Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature *505*, 495–501.

Moore, J.M., Rabaia, N.A., Smith, L.E., Fagerlie, S., Gurley, K.E., Loukinov, D., Disteche, C.M., Collins, S.J., Kemp, C.J., Lobanenkov, V.V., and Filippova, G.N. (2012). Loss of maternal CTCF is associated with peri-implantation lethality of *Ctcf* null embryos. PLoS ONE *7*, e34915.

Mukhopadhyay, R., Yu, W., Whitehead, J., Xu, J., Lezcano, M., Pack, S., Kanduri, C., Kanduri, M., Ginjala, V., Vostrov, A., et al. (2004). The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. Genome Res. *14*, 1594–1602.

Nakahashi, H., Kwon, K.R., Resch, W., Vian, L., Dose, M., Stavreva, D., Hakim, O., Pruett, N., Nelson, S., Yamane, A., et al. (2013). A genome-wide map of CTCF multivalency redefines the CTCF code. Cell Rep. *3*, 1678–1689.

Ong, C.T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nat. Rev. Genet. *15*, 234–246.

Payne, S.R., and Kemp, C.J. (2005). Tumor suppressor genetics. Carcinogenesis *26*, 2031–2045.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194–1211.

Rakha, E.A., Green, A.R., Powe, D.G., Roylance, R., and Ellis, I.O. (2006). Chromosome 16 tumor-suppressor genes in breast cancer. Genes Chromosomes Cancer *45*, 527–535.

Rasko, J.E., Klenova, E.M., Leon, J., Filippova, G.N., Loukinov, D.I., Vatolin, S., Robinson, A.F., Hu, Y.J., Ulmer, J., Ward, M.D., et al. (2001). Cell growth inhibition by the multifunctional multivalent zinc-finger factor CTCF. Cancer Res. *61*, 6002–6007.

Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R., et al. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome Res. *22*, 1680–1688.

Zampieri, M., Guastafierro, T., Calabrese, R., Ciccarone, F., Bacalini, M.G., Reale, A., Perilli, M., Passananti, C., and Caiafa, P. (2012). ADP-ribose polymers localized on Ctcf-Parp1-Dnmt1 complex prevent methylation of Ctcf target sites. Biochem. J. *441*, 645–652.