



Article

# CTCF Expression is Essential for Somatic Cell Viability and Protection Against Cancer

Charles G Bailey<sup>1</sup>, Cynthia Metierre<sup>1</sup>, Yue Feng<sup>1</sup>, Kinsha Baidya<sup>1</sup>, Galina N Filippova<sup>2</sup>, Dmitri I Loukinov<sup>3</sup>, Victor V Lobanenkov<sup>3</sup> , Crystal Semaan<sup>1</sup> and John EJ Rasko<sup>1,4,\*</sup>

<sup>1</sup> Gene and Stem Cell Therapy Program Centenary Institute, The University of Sydney, Camperdown, 2050 NSW, Australia; c.bailey@centenary.org.au (C.G.B.); c.metierre@centenary.org.au (C.M.); j.feng@centenary.org.au (Y.F.); kinsha\_3@hotmail.com (K.B.); c.semaan@centenary.org.au (C.S.)

<sup>2</sup> Department of Pathology, University of Washington, 98195 Seattle, USA; gfilippo@u.washington.edu

<sup>3</sup> Laboratory of Immunogenetics, Molecular Pathology Section, National Institute of Allergy & Infectious Diseases, 20852-8152 Rockville, USA; dloukinov@niaid.nih.gov (D.I.L.); vlobanenkov@niaid.nih.gov (V.V.L.)

<sup>4</sup> Cell and Molecular Therapies, Royal Prince Alfred Hospital, Camperdown, 2050 NSW, Australia

\* Correspondence: j.rasko@centenary.org.au; Tel.: +61-2-95656116

Received: 19 October 2018; Accepted: 11 November 2018; Published: 30 November 2018



**Abstract:** CCCTC-binding factor (CTCF) is a conserved transcription factor that performs diverse roles in transcriptional regulation and chromatin architecture. Cancer genome sequencing reveals diverse acquired mutations in *CTCF*, which we have shown functions as a tumour suppressor gene. While CTCF is essential for embryonic development, little is known of its absolute requirement in somatic cells and the consequences of *CTCF* haploinsufficiency. We examined the consequences of CTCF depletion in immortalised human and mouse cells using shRNA knockdown and CRISPR/Cas9 genome editing as well as examined the growth and development of heterozygous *Ctcf* (*Ctcf*<sup>+/-</sup>) mice. We also analysed the impact of *CTCF* haploinsufficiency by examining gene expression changes in *CTCF*-altered endometrial carcinoma. Knockdown and CRISPR/Cas9-mediated editing of *CTCF* reduced the cellular growth and colony-forming ability of K562 cells. CTCF knockdown also induced cell cycle arrest and a pro-survival response to apoptotic insult. However, in p53 shRNA-immortalised *Ctcf*<sup>+/-</sup> MEFs we observed the opposite: increased cellular proliferation, colony formation, cell cycle progression, and decreased survival after apoptotic insult compared to wild-type MEFs. CRISPR/Cas9-mediated targeting in *Ctcf*<sup>+/-</sup> MEFs revealed a predominance of in-frame microdeletions in *Ctcf* in surviving clones, however protein expression could not be ablated. Examination of *CTCF* mutations in endometrial cancers showed locus-specific alterations in gene expression due to *CTCF* haploinsufficiency, in concert with downregulation of tumour suppressor genes and upregulation of estrogen-responsive genes. Depletion of CTCF expression imparts a dramatic negative effect on normal cell function. However, CTCF haploinsufficiency can have growth-promoting effects consistent with known cancer hallmarks in the presence of additional genetic hits. Our results confirm the absolute requirement for CTCF expression in somatic cells and provide definitive evidence of *CTCF*'s role as a haploinsufficient tumour suppressor gene. *CTCF* genetic alterations in endometrial cancer indicate that gene dysregulation is a likely consequence of *CTCF* loss, contributing to, but not solely driving cancer growth.

**Keywords:** CTCF; tumour suppressor gene; haploinsufficiency; zinc finger; CRISPR/Cas9; cancer; endometrial cancer; gene editing

## 1. Introduction

CTCF is a conserved multivalent transcription factor with diverse roles in transcriptional regulation and three-dimensional genome organisation such that it has been called the ‘master weaver’

protein [1]. CTCF is essential during embryonic development, as *Ctcf* null embryos are unable to implant [2]. Tissue-specific deletion of this ubiquitous factor in mice using conditional *Ctcf* alleles has highlighted the importance of CTCF availability in somatic tissues. Conditional deletion of CTCF in thymocytes can hamper T-cell differentiation and cell cycle progression, but not ablate T cell function [3]. Conditional deletion of *Ctcf* in the limb mesenchyme induces extensive apoptosis during limb development highlighting CTCF's pro-survival role [4]. Similarly, deletion of *Ctcf* specifically during early mouse brain development, led to PUMA upregulation and subsequent massive apoptosis [5]. Of relevance for our studies, *Ctcf* heterozygous mice, however, are more prone to the formation of spontaneous cancers, as well those induced by radiation and chemical means [6].

CTCF links gene regulation to genomic architecture by co-ordinating DNA looping in concert with cohesin [7–9]. Within chromosomal territories, CTCF defines boundaries between sub-megabase-scale topologically-associated domains (TADs) [10–12] in a framework that is conserved [13]. These TADs themselves can serve as large gene regulatory domains establishing specific gene expression profiles [14]. TAD organisation is CTCF site orientation-specific [13,15] and rewiring of CTCF sites can significantly perturb gene expression by affecting promoter-enhancer interactions or boundaries between euchromatin and heterochromatin [16–18]. In cancer, hypermethylation or somatic mutation of CTCF binding sites has been shown to affect chromatin boundaries. This, in turn, can induce tumour suppressor silencing [19,20]; disruption of CTCF-dependent insulation leading to aberrant TAD formation and oncogene activation [21]; and cis-activation of genes implicated in cancer [22,23].

Our previous studies first demonstrated the growth inhibitory effects of CTCF in vitro [24] and subsequently confirmed that CTCF acts as a tumour suppressor gene in vivo by suppressing tumour growth [25]. Isolated CTCF mutations have been identified in breast, prostate and Wilms' tumours [26] and acute lymphoblastic leukaemia [27]. However recent cancer genome studies have revealed the extensive somatic mutations occurring in CTCF [28]. CTCF has been classified as a significantly mutated gene owing to its high frequency of mutation and deletion in endometrial cancer [29]. CTCF mutations are detected in 35% of endometrial carcinomas exhibiting microsatellite instability (MSI), and in 20% of MSI-negative tumours [30]. One report describing 17 oncogenic signatures in cancer, defines one signature, M5, as comprising MSI-positive endometrioid cancers and some luminal A breast cancers. In this subset of endometrioid and breast cancers, CTCF mutations were identified in 40% of samples including inactivation of specific zinc fingers (ZFs) of CTCF that would lead to altered DNA binding [31]. We since revealed that CTCF genetic alterations have a pro-tumourigenic effect in endometrial cancer by altering cellular polarity and enhancing cell survival [32].

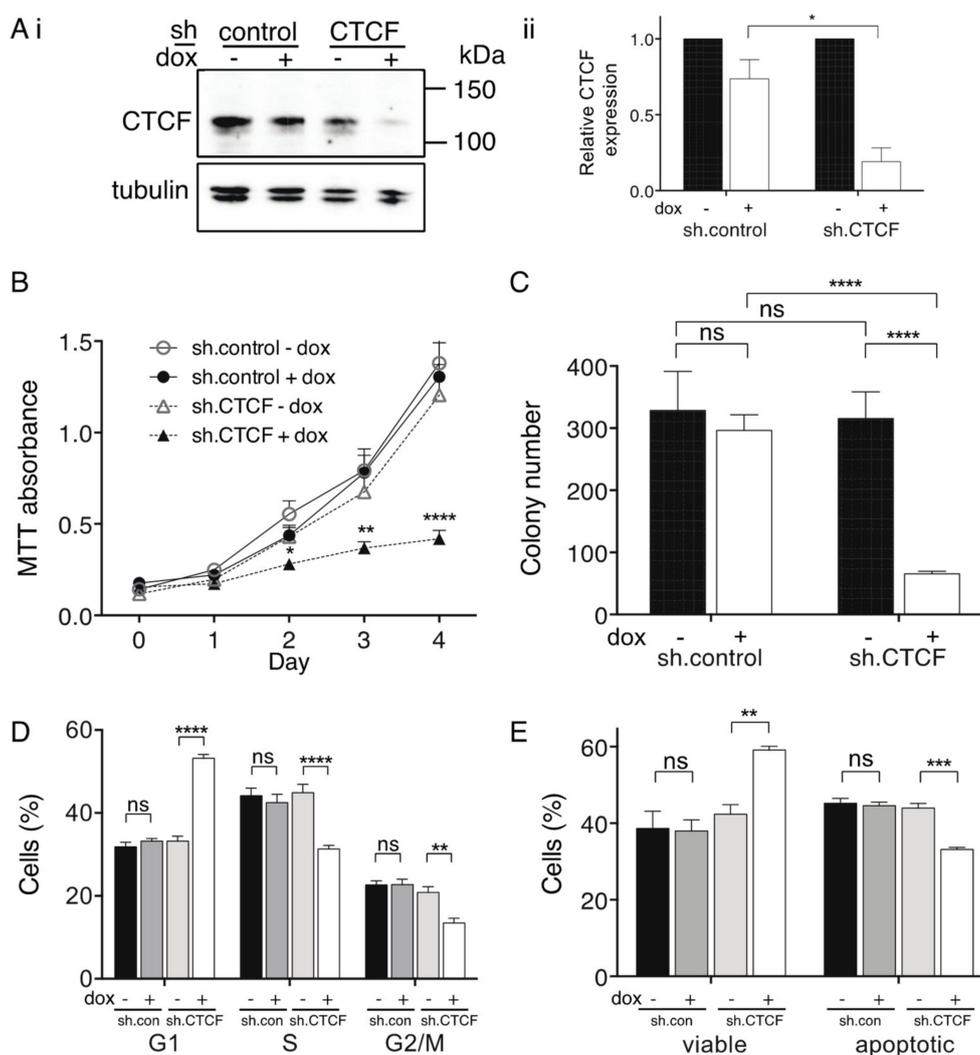
Genetic lesions in CTCF, whether heterozygous deletion, nonsense, frameshift, or even missense ZF mutations, can all result in CTCF haploinsufficiency. In endometrial cancer, CTCF mRNA transcripts expressed from alleles containing nonsense or frameshift mutations are subjected to nonsense-mediated decay [30,32]. Somatic missense mutations in residues critical for CTCF ZF binding to DNA can result in selective loss of binding to some CTCF target sites, but not all [26], indicating the functional implications of incomplete loss of CTCF binding in cancer is unclear. Loss of heterozygosity (LOH) at 16q22 can lead to CTCF haploinsufficiency and *IGF2* up-regulation in Wilms' tumours [33]. To date, modelling the full impact of CTCF haploinsufficiency on CTCF's tumour suppressor function has not been previously examined.

In this study we assessed several genetic models of CTCF haploinsufficiency to reveal in detail the impact of heterozygous loss of CTCF in somatic cells, whole mice and human endometrial cancer. Depletion of CTCF expression in K562 erythroleukaemia cells using shRNA knockdown or CRISPR/Cas9-mediated targeting of CTCF decreased cellular proliferation. In vivo, *Ctcf* heterozygosity negatively impacted the growth and gross development of mice. However, p53 shRNA-immortalised *Ctcf*<sup>+/-</sup> mouse embryonic fibroblasts (MEFs) were functionally distinct from wild-type (WT) MEFs by exhibiting increased cellular growth and other known cancer hallmarks. Importantly, we were unable to generate *Ctcf* nullizygous MEFs after CRISPR/Cas9 genome editing confirming that CTCF is

absolutely essential for somatic cell viability. Finally, we examined curated human endometrial carcinoma genomic data and observed that *CTCF* haploinsufficiency contributed to the transcriptional dysregulation of specific loci as well as inducing a unique gene signature in human cancers.

## 2. Results

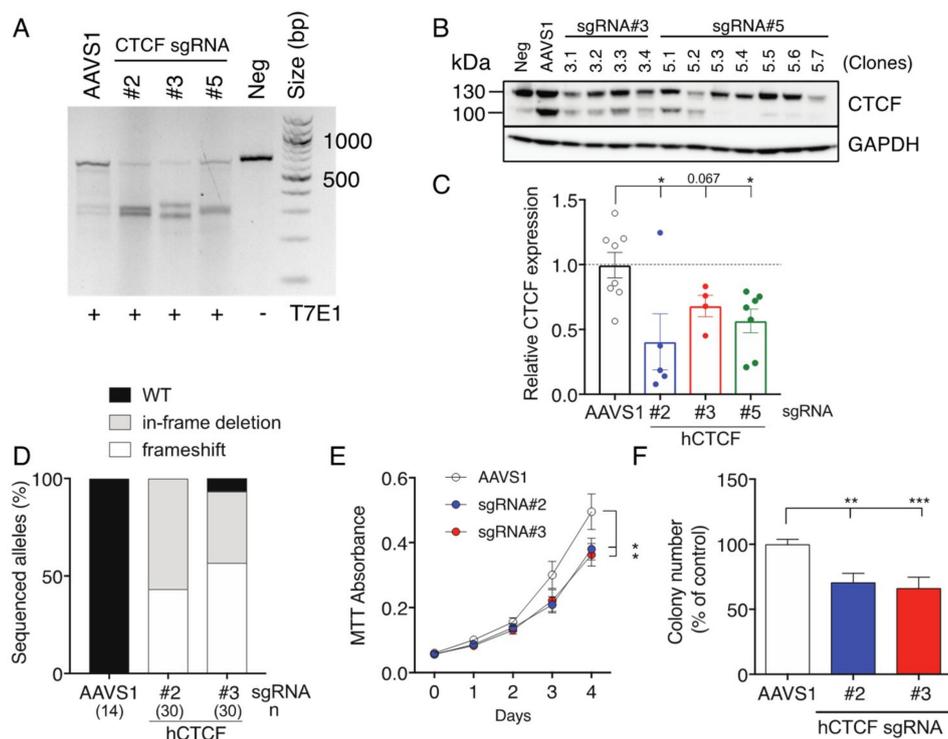
We used shRNA knockdown to model the cellular consequences of reduced *CTCF* expression in K562 cells. Western blots showed that *CTCF* protein expression was significantly knocked down by ~80% in the presence of doxycycline (dox) in sh.*CTCF* K562 cells compared to non dox-treated cells and sh.control cells (Figure 1Ai,ii). Cellular proliferation showed that *CTCF* knockdown resulted in a significant reduction of proliferation ( $p < 0.0001$ , Figure 1B). We similarly observed a significant reduction in sh.*CTCF* K562 colony number compared to non dox treatment and to sh.control (both  $p < 0.0001$ , Figure 1C). *CTCF* knockdown led to growth arrest with an increase in G1 phase ( $p < 0.0001$ ), and a concomitant reduction of cells in S ( $p < 0.0001$ ) and G2/M phases ( $p = 0.0036$ , Figure 1D). We next examined the response of sh.*CTCF* K562 cells after UV insult and observed *CTCF* knockdown in K562 cells resulted in an increase in cell viability after recovery from UV exposure ( $p = 0.004$ ) and a decrease in apoptotic (Annexin V-positive) cells ( $p = 0.0006$ , Figure 1E).



**Figure 1.** Inhibition of cell proliferation and clonogenicity following *CTCF* shRNA knockdown in K562 cells. (A) Immunoblot of *CTCF* shRNA knockdown in the presence and absence of doxycycline (dox) compared to control shRNA after 4 d: representative immunoblot (i); and relative *CTCF* expression normalised to  $\alpha$ -tubulin confirmed by ImageJ densitometric analysis (ii). Functional assays performed

after 4 d knockdown including: (B) MTT proliferation; (C) colony forming assay; (D) cell cycle analysis; and, (E) apoptotic response after recovery from UV insult. Data represent the mean  $\pm$  SEM for 3 experiments each performed in triplicate. Statistical analysis was performed using a Mann-Whitney U-test (ns = not significant, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ).

These data and our previous studies indicated that CTCF dosage is critical for its tumour suppressive functions [25,32], however CTCF haploinsufficiency has not been definitively modelled in vitro. To address this, we used CRISPR/Cas9-directed genome editing to induce genetic lesions in K562 cells (Supplementary Figure S1A), which we previously verified to contain wild type CTCF alleles using Sanger sequencing [25]. SgRNAs were designed to direct Cas9 nuclease cleavage on both strands of the critical third exon of CTCF, encoding the entire N-terminus of CTCF (Supplementary Figure S1B). All three CTCF sgRNA achieved efficient Cas9 cleavage of CTCF exon 3 (Figure 2A). A representative Western blot showing CTCF protein expression in clones isolated after CRISPR/Cas9-mediated targeting of CTCF using two sgRNAs (#3 and #5) is shown in Figure 2B. CTCF protein expression was decreased by approximately 50% in most surviving clones irrespective of the sgRNA used (Figure 2C). As each clone should contain at least one edited CTCF allele, we PCR-amplified the edited region in CTCF, cloned the PCR products and then sequenced them. We detected a mixture of CTCF alleles arising in clones including frameshifts induced by deletion or insertions near the protospacer adjacent motif (PAM) site or in-frame deletions leading to microdeletions in the CTCF protein (Figure 2D). In some clones, we observed three distinct edited CTCF alleles, consistent with K562 cells having a hypotriploid karyotype [34]. SgRNAs #2 and #3 induced 100% and ~96% gene editing efficiency respectively with a ~50:50 mixture of frameshifts and in-frame deletions (Figure 2D).

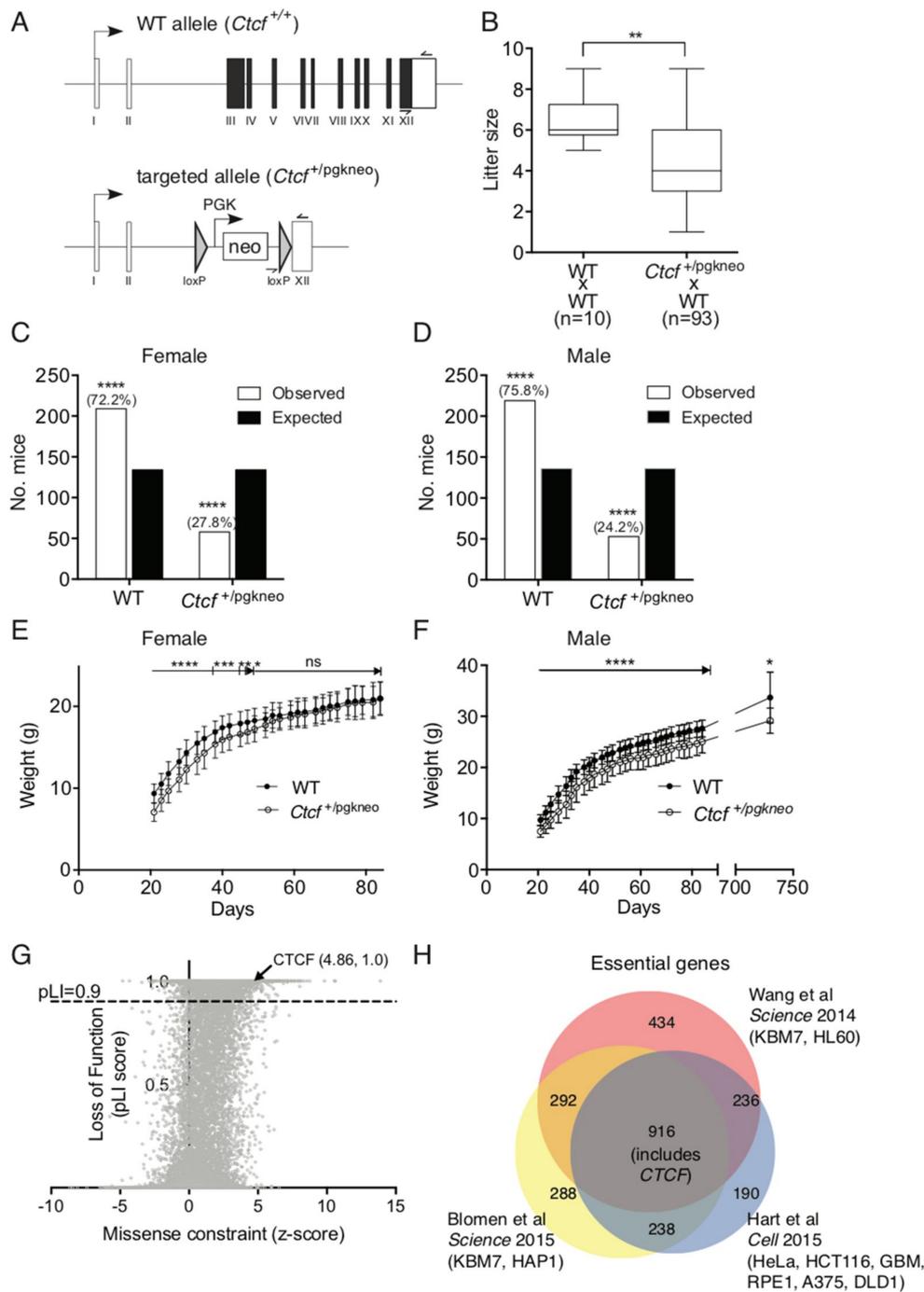


**Figure 2.** Inhibition of cell proliferation and clonogenicity following CRISPR/Cas9 targeting of CTCF in K562 cells. K562 cells were transduced with Cas9 and sgRNA-containing lentivectors (AAVS1 sgRNA = control; human CTCF exon 3 sgRNAs #2, #3, #5) and enriched for eGFP<sup>+</sup>mCherry<sup>+</sup> cells using FACS; Neg = untransduced K562 cells. (A) CTCF exon 3 PCR amplification and T7 endonuclease I (T7E1) digestion: approximate expected sizes (in bp) for digested products #2 (310, 345), #3 (296, 359) and #5 (323, 332). Analysis of CTCF protein levels in K562 clones: (B) immunoblot;

and (C) densitometric analysis of upper 130 kDa band. CTCF protein expression normalised to GAPDH expression in each sample is shown relative to untransduced K562 cells. (D) Summary of results after sequencing of *CTCF* exon 3 PCR amplicons from individual clones;  $n$  = number of clones sequences (in brackets). Functional assays performed were MTT cell proliferation (E); and clonogenicity assays (F). Quantitative data represent the mean  $\pm$  SEM for 3–4 experiments each performed in triplicate. Statistical analysis was performed using a Mann-Whitney U-test (ns = not significant, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ).

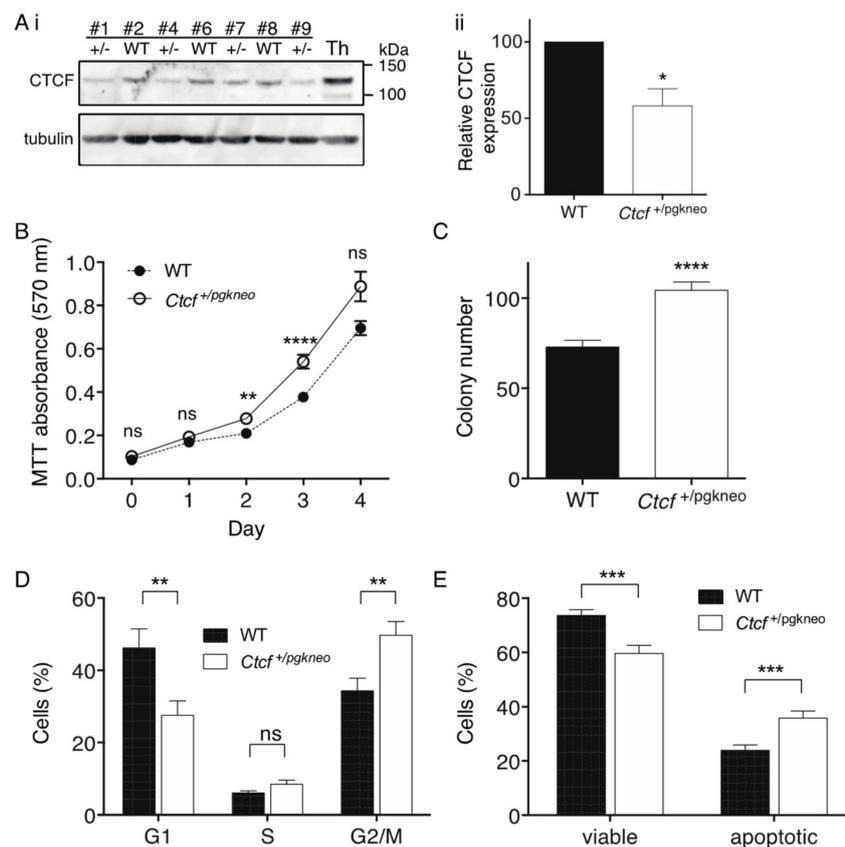
We next performed MTT cell proliferation assays on eGFP<sup>+</sup> mCherry<sup>+</sup> K562 cell pools (Supplementary Figure S1B) and showed that cells targeted using *CTCF* sgRNAs #2 and #3 exhibited reduced cellular proliferation ( $p = 0.014$  and  $p = 0.012$  respectively, Figure 2E). We also performed clonogenicity assays and confirmed that CRISPR/Cas9-directed genome editing of *CTCF* inhibited colony-forming ability by ~30 % for sgRNAs #2 and #3 (Figure 2F). Therefore, inducing genetic lesions in *CTCF* leading to haploinsufficient levels of CTCF in K562 cells had a negative impact on cellular growth.

We then examined *Ctcf* heterozygous mice to better determine what impact heterozygous deletion of the *Ctcf* locus has on post-natal growth and development. We backcrossed these *Ctcf*<sup>+/pgkneo</sup> mice (Figure 3A), which were originally described on a mixed 129SvJ:C57Bl/6J background and exhibited embryonic lethality as homozygotes [2], onto C57Bl/6J mice for at least 10 generations. Backcrossed C57Bl/6J *Ctcf*<sup>+/pgkneo</sup> mice bred with wild type C57Bl/6J (WT) mice had smaller mean litter sizes than from normal WT x WT mice (Figure 3B). This was explained by both female and male *Ctcf*<sup>+/pgkneo</sup> mice being born at sub-Mendelian ratios (~28% and ~24% respectively) compared to WT (*Ctcf*<sup>+/+</sup>) mice (both  $p < 0.0001$  *Chi*-square test; Figure 3C,D). After weaning at approximately day 21 we recorded mouse weights 3 times a week until 12 weeks of age. Female *Ctcf*<sup>+/pgkneo</sup> mice were smaller than WT littermates up to 7 weeks of age (~14% less body weight, Figure 3E), whereas male *Ctcf*<sup>+/pgkneo</sup> mice were consistently smaller than WT littermates in the first 12 weeks of age (~12% less body weight, Figure 3F). This reduced weight phenotype was maintained in male *Ctcf*<sup>+/pgkneo</sup> mice even beyond two years of age (Figure 3F). These data show *Ctcf* haploinsufficiency can significantly impact growth and development in mice. Examination of genetic variation occurring in CTCF in humans using the ExAC database [35] revealed that CTCF is extremely intolerant to genetic variation within the protein-coding region. CTCF exhibits significantly fewer nonsynonymous variants than expected ( $z$  score = 4.86) and can be classified as haploinsufficient due to intolerance to heterozygous loss-of-function variation (pLI score = 1.0; Figure 3G). Two genome-wide CRISPR screens in diploid cells [36,37] and a synthetic lethal gene trap screen in haploid cells [38] identified 916 core fitness genes essential for cell viability common to all 3 screens, including *CTCF* (Figure 3H). These data confirm CTCF as an essential gene in higher order eukaryotes.



**Figure 3.** CTCF haploinsufficiency phenotype in mice and humans. (A) Schematic of targeted inactivation of *Ctcf* in mice. Open boxes represent untranslated regions, filled boxes represent coding region. Genotyping primers used to distinguish alleles are indicated with half-arrowheads. Litter sizes (B), and Mendelian ratios (% in brackets) of female (C) and male (D) pups born from WT x *Ctcf*<sup>+/pgkneo</sup> intercrosses. Weights of pups during development following weaning at day 21 (mean ± SD): (E) for female; and, (F) aged male (>2 yo) mice. (G) Analysis of *CTCF* genetic variation in humans using the ExAC database with pLI > 0.9 indicating intolerance to heterozygous loss-of-function variation. The missense constraint is a measure of the deviation away from the observed variants in a gene versus the expected variants (high positive z-scores indicated intolerance to variation). (H) Venn diagram of essential genes identified in three independent genetic screens in human cells. Statistical analysis was performed using Fisher’s exact test or Mann-Whitney U-test (ns = not significant, \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001, \*\*\*\* *p* < 0.0001).

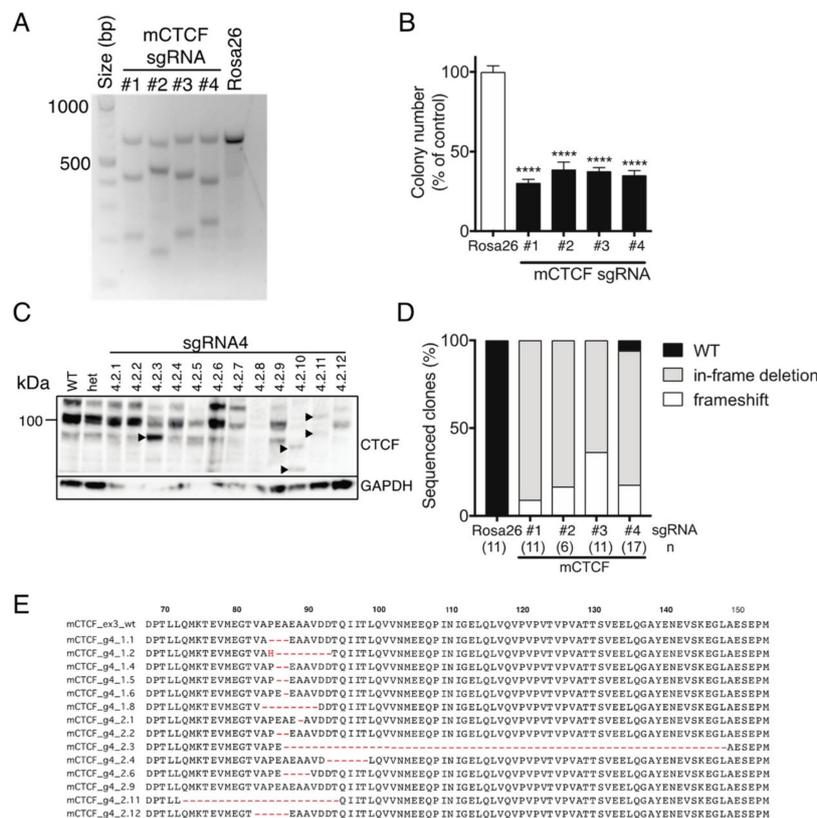
To examine the cellular consequences of *Ctcf* haploinsufficiency, we isolated mouse embryonic fibroblasts (MEFs) from a single litter containing 4 *Ctcf*<sup>+/-</sup>/*pgkneo* and 3 WT pups. These MEFs were immortalised with a retrovirus encoding stable shRNA knockdown of *p53* and then analysed by immunoblot for *Ctcf* protein expression (Figure 4Ai). Densitometric analysis of the *Ctcf*<sup>+/-</sup>/*pgkneo* and WT MEF samples confirmed *Ctcf* protein was reduced in *Ctcf*<sup>+/-</sup>/*pgkneo* MEFs to a mean of 58% of WT ( $p = 0.033$ , Figure 4Aii). We performed MTT assays and showed immortalised *Ctcf*<sup>+/-</sup>/*pgkneo* MEFs exhibited an increase in cellular proliferation compared to WT MEFs ( $p = 0.0028$  day 2,  $p < 0.0001$  day 3, Figure 4B). *Ctcf*<sup>+/-</sup>/*pgkneo* MEFs also displayed an increase in colony-forming ability compared to WT ( $p < 0.0001$ , Figure 4C). We analysed cell cycle kinetics and showed that *Ctcf*<sup>+/-</sup>/*pgkneo* MEFs exhibited a decrease in G1 phase compared to WT MEFs ( $p = 0.0072$ ) with a concomitant increase in G2/M phase ( $p = 0.0043$ ) (Figure 4D). We next examined the cellular response to UV-induced apoptosis and found that immortalised *Ctcf*<sup>+/-</sup>/*pgkneo* MEFs exhibited a decrease in viability and concomitant increase in Annexin V-positive cells compared to WT cells ( $p = 0.0002$  and  $p = 0.0005$  respectively) (Figure 4E). Immortalised *p53*-deficient *Ctcf* heterozygous MEFs exhibit pro-tumourigenic characteristics, indicating that *Ctcf* is acting as a haploinsufficient tumour suppressor gene.



**Figure 4.** Functional characterisation of immortalised *Ctcf*<sup>+/-</sup> MEFs. (A) Immunoblot of whole cell lysates isolated from WT and *Ctcf*<sup>+/-</sup>/*pgkneo* (+/-) MEFs (clone number indicated), Th=thymus (i); densitometric analysis of *Ctcf* protein normalised to the  $\beta$ -tubulin loading control (ii). Functional assays including: MTT proliferation (B); clonogenicity (C) cell cycle analysis (D); and apoptosis assay following 18 h recovery from UV irradiation (E). Data represent the mean  $\pm$  SEM for 3 experiments each performed with 4 *Ctcf*<sup>+/-</sup>/*pgkneo* and 3 WT cell lines. Statistical analysis was performed using a Mann-Whitney U-test (ns = not significant, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ).

We next used CRISPR/Cas9-mediated targeting of *Ctcf* in monoallelic hemizygous *Ctcf*<sup>+/-</sup>/*pgkneo* MEFs to assess the impact of inducing potentially deleterious *Ctcf* genetic lesions. Four sgRNAs were designed to target the first coding exon (exon 3) of *Ctcf*, as well as a control sgRNA targeting

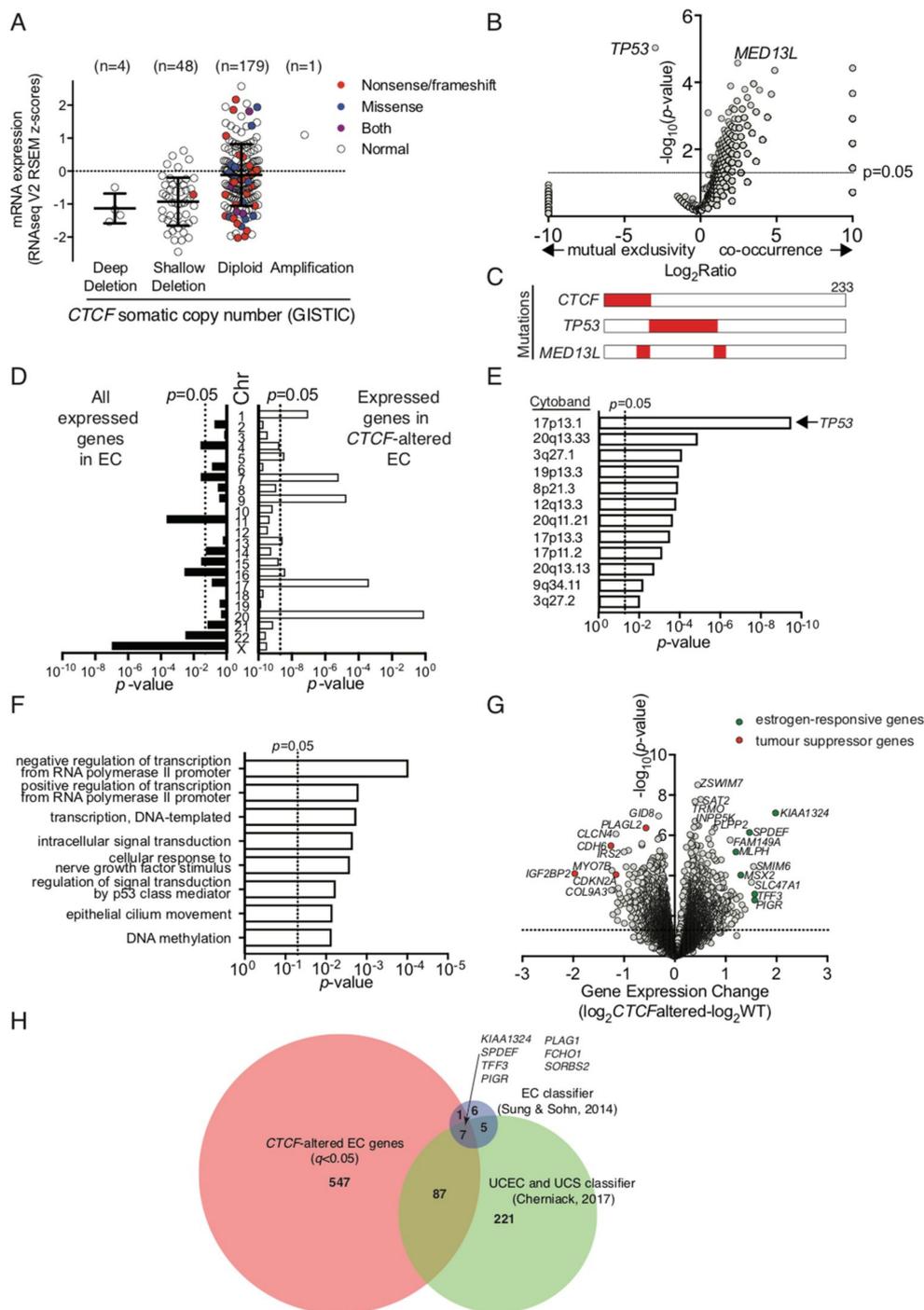
the *Rosa26* locus (Supplementary Figure S1C). Efficient Cas9-directed cleavage of *Ctcf* exon 3 was observed using each sgRNA against *Ctcf*, whereas *Rosa26* sgRNA had no detectable effect (Figure 5A). A clonogenicity assay was performed using *Ctcf*-targeted eGFP<sup>+</sup>mCherry<sup>+</sup> *Ctcf*<sup>+/pgkneo</sup> MEFs which showed that colony-forming capacity was significantly reduced to ~30–40% of control (Figure 5B). We isolated individual clones for each *Ctcf* sgRNA by FACS and then examined *Ctcf* protein expression. *Ctcf* expression in surviving clones was maintained despite attempts to inactivate the hemizygous *Ctcf* allele, however, in some clones lower molecular weight *Ctcf* species were detected, e.g., for sgRNA#4 (Figure 5C), sgRNA #1, #2, #3 (Supplementary Figure S2). These most likely result from in-frame deletions; or alternatively, frameshift mutations that occur in the first coding exon of *Ctcf*, such that alternative ATG start codons are utilised leading to N-terminal *Ctcf* protein truncations. Sequencing of CRISPR/Cas9 genome-edited surviving clones showed 44 out of 45 clones exhibited in-frame deletions or frameshift-inducing indels (Figure 5D). More than two-thirds of clones had in-frame deletions proximal to the PAM site leading to N-terminal microdeletions in *Ctcf* ranging in size from 1–62 aa (sgRNA#4; Figure 5E), some of which could be detected by Western blot (e.g., clones 4.2.3, 4.2.10, and 4.2.11; Figure 5C). These data confirm that CTCF is essential in somatic cells and that CTCF nullizygoty cannot be sustained in viable cells.



**Figure 5.** CRISPR/Cas9-directed editing of *Ctcf* in hemizygous MEFs. (A) *Ctcf*<sup>+/pgkneo</sup> MEFs transduced with Cas9 and sgRNA-containing lentivectors (mouse *Ctcf* exon 3 sgRNAs #1, #2, #3, and #4; *Rosa26* sgRNA) were FACS-enriched and subjected to T7EI digestion of *Ctcf* exon 3 amplicons amplified from isolated gDNA. Approximate expected sizes (in bp) for digested products #1 (427, 214), #2 (476, 165), #3 (428, 213), and #4 (399, 242). Clonogenicity assay (B); Western blot analysis of individual clones (from sgRNA#4, arrowheads indicate lower molecular weight *Ctcf* variants) (C); and molecular genetic analysis of individual clones; n=number of clones sequenced (in brackets) (D). (E) Examples of frequently occurring in-frame deletions in *Ctcf*<sup>+/pgkneo</sup> MEFs (from sgRNA#4). Quantitative data represent the mean  $\pm$  SEM for three experiments each performed in triplicate. Statistical analysis was performed using Mann-Whitney U-test (ns = not significant, \*\*\*\*  $p < 0.0001$ ).

To ascertain the impact of *CTCF* haploinsufficiency in the context of human cancers, we examined a uterine corpus endometrial carcinoma (UCEC) dataset from The Cancer Genome Atlas, which exhibits *CTCF* genetic alterations in 45 out of 232 patient samples (~19%) [28]. GISTIC analysis of this cohort assigned each patient sample into potential somatic copy number alterations based on relative *CTCF* expression level (Figure 6A). *CTCF* expression was decreased in a substantial proportion of endometrial cancer samples, some of which can be directly attributed to genetic deletion of *CTCF* (deep deletion). Many inactivating nonsense and frameshift mutations in *CTCF* were found in the notionally diploid population (40 out of 179, 22.3%; Figure 6A). Samples with inactivating mutations and confirmed deletions are classified herein as 'CTCF-altered'. We then analysed other gene mutations that co-occurred with or were mutually exclusive in *CTCF*-altered endometrial cancers. *TP53* mutations (66 out of 68) occurred with mutual exclusivity to *CTCF* mutations ( $p = 9.28 \times 10^{-6}$ , Figure 6B,C); whereas mutations in *MED13L*, encoding a subunit of the Mediator transcriptional co-activation complex, co-occurred with *CTCF* mutations in 13 out of 23 cancers ( $p = 2.64 \times 10^{-5}$ , Figure 6B).

We analysed RNAseq data available for these endometrial cancer samples and showed that *CTCF* gene expression was not significantly decreased in *CTCF*-altered cancers despite the presence of inactivating mutations (Supplementary Figure S3A). We next examined the chromosomal distribution of all expressed genes detected above threshold in endometrial cancers (~13,000) and found an enrichment for genes expressed on chromosomes 11, 16, 22 and X (Figure 6D). However, in *CTCF*-altered cancers there was enrichment for genes expressed on chromosomes 1, 7, 9, 17, and 20 (Figure 6C). Further analysis of enriched chromosomal regions with altered gene expression highlighted multiple loci on the short arm of chromosome 17 including 17p13.1 (which contains the *TP53* locus), and the long arm of chromosome 20 (Figure 6D). *TP53* gene expression was significantly decreased in *CTCF*-altered cancers ( $p = 0.0437$ , Supplementary Figure S3B), however, there was no significant difference after the exclusion of samples containing *TP53* mutations from the analysis (Supplementary Figure S3C). To gain further insight, gene ontology analysis of biological processes in *CTCF*-altered endometrial cancers indicated *CTCF* mutation may impact predominantly on transcriptional regulation, cell signalling pathways, such as p53, and DNA methylation (Figure 6E). Closer examination of genes that were dysregulated in *CTCF*-altered cancers showed expression of the *CTCF* paralog *CTCF*L was decreased ( $p = 0.0167$ ; Supplementary Figure S3D), the exemplar *CTCF*-regulated gene *H19* was decreased ( $p = 0.0087$ ; Supplementary Figure S3E), whilst no change was observed in *ZFH3* expression which is located adjacent to *CTCF* on chromosome 16q22 (Supplementary Figure S3F). Importantly, expression levels of the tumour suppressor genes *CDKN2A* and *PIK3CA*, which are deleted or mutated in endometrial cancer [28,39], were decreased in *CTCF*-altered samples ( $p = 0.0006$  and  $p = 0.0007$ , respectively; Figure 6G, Supplementary Figure S3G,H). Putative tumour suppressor genes *CDH6* and *IGF2BP2* were two of the most significantly fold-decreased genes ( $p = 0.0003$  and  $p = 5.71 \times 10^{-5}$  respectively; Figure 6F, Supplementary Figure S3I,J). Furthermore, the expression of estrogen-responsive genes *KIAA1324*, *MLPH*, *MSX2*, *SPDEF*, *TFF3*, and *PIGR* were all significantly up-regulated in *CTCF*-altered endometrial cancers ( $p = 0.0004$ , 0.0007, 0.0032, 0.0009, 0.0122, and 0.0028 respectively, Figure 6F, Supplementary Figure S3K–P). Lastly, differentially expressed genes in *CTCF*-altered cancers were significantly overrepresented in a 19 gene signature that classifies endometrial cancers into endometrioid and serous subtypes (8 out of 19,  $p = 1.46 \times 10^{-6}$ ) [40] as well as a 320 gene classifier that distinguishes endometrioid and serous endometrial cancers from uterine carcinosarcomas (94 out of 320,  $p = 2.96 \times 10^{-44}$ ) [39].



**Figure 6.** The molecular genetic landscape of *CTCF*-altered endometrial cancers. Gene expression and DNA sequencing data was analysed from the TCGA endometrial carcinoma patient cohort [28]. (A) GISTIC analysis of mRNA expression indicative of somatic copy number alterations in endometrial carcinomas. Filled symbols indicate cancers with *CTCF* coding region mutations. (B) Plot of significantly co-occurring or mutually exclusive mutant genes with *CTCF*-altered cancer. (C) Schematic showing the co-occurrence between *CTCF*, *TP53* and *MED13L* mutations in endometrial cancer ( $n = 233$  patient samples). (D) Chromosomal distribution of all expressed genes in endometrial carcinoma ( $n = 13,271$ ) vs those differentially expressed ( $n = 642$ ;  $q < 0.05$ ) in *CTCF*-altered cancers (data is normalised for gene density). (E) Chromosomal location (cytoband) of genes differentially expressed in *CTCF*-altered endometrial cancers ( $q < 0.05$ ). Data for (D) and (E) were analysed using the Fisher’s exact test. (F) Biological process terms enriched in gene ontology analysis of *CTCF*-altered endometrial cancers

( $q < 0.05$ ). (G) Plot of most significantly differentially regulated genes in *CTCF*-altered cancers compared to *CTCF* WT cancers, genes of particular interest are highlighted. (H) Genes common to *CTCF*-altered differentially regulated EC genes (642 total;  $q < 0.05$ ) and two gene classifiers used to distinguish uterine corpus endometrial cancers (UCEC) from serous cancers as well as uterine carcinosarcomas (UCS) [39,40].

### 3. Discussion

Haploinsufficiency arises when only a single functional copy of a gene is inadequate for normal cell function [41]. *CTCF* was identified as one of nearly 300 haploinsufficient genes in humans based on published literature or a clear association with genetic disease [42]. Herein we empirically demonstrate that *CTCF* can be classified as haploinsufficient due to its intolerance to loss-of-function polymorphisms in humans. *CTCF* haploinsufficiency resulting from germline or de novo genetic mutations in *CTCF* (including genetic deletion, frameshift mutations or missense mutation) causes intellectual disability in humans; now classified as autosomal dominant mental retardation (MRD21; OMIM #615502) [43–45]. The impact of *CTCF* mutations on human gene expression manifested as a predominant downregulation of genes involved in the cellular response to extracellular stimuli [43] and hypermethylation of *CTCF*-binding sites [45]. The mechanisms that connect *CTCF* haploinsufficiency with cancer have yet to be elucidated.

Numerous studies over more than a decade using siRNA or shRNA knockdown of *CTCF* have only been suggestive of *CTCF*'s essential role in normal cell function. Typically, such experiments are short term, do not fully ablate *CTCF* expression and cells remain viable. Similarly, our shRNA knockdown of *CTCF* in K562 cells showed that cell proliferation and clonogenic capacity was decreased, cell survival after UV insult was increased, and cells underwent growth arrest. Paradoxically, after *CTCF* knockdown here, and in corroboration of our previous *CTCF* overexpression study [25] we observed tumour suppressive phenotypes for *CTCF* in K562 cells. These data reveal that physiological *CTCF* expression levels are critical for normal cellular function and reveal the functional importance of maintaining physiological *CTCF* expression levels.

More recently CRISPR/Cas9 mediated genome editing techniques have given tremendous insight into *CTCF*'s roles in higher-order chromatin organisation. Most studies have focussed on rewiring *CTCF*-mediated chromatin interactions such as disrupting TAD boundaries [16] and switching the orientation of *CTCF* target sequences to alter genome topology [17]. Unbiased genetic screens using CRISPR/Cas9 sgRNA libraries or synthetic lethality in haploid cells have identified *CTCF* as part of an 'essentialome' containing ~900 core fitness genes required for cell viability [36–38]. However, only two studies to date have directly focused on targeting *CTCF* in mammalian cells using CRISPR/Cas9 editing [46,47], but with disparate outcomes. *CTCF* heterozygous MCF10A clones generated by CRISPR exhibited similar proliferation rates compared to control, though cells were slower to repair double-stranded DNA breaks [46]. Silencing of *CTCF* in the RS4;11 acute lymphoblastic leukaemia cell line increased colony numbers in soft agar overlays compared to control, though *CTCF* protein appeared to be reduced to minimal levels [47].

Our strategy was to examine the essentiality of *CTCF* in somatic cells by CRISPR/Cas9 targeting of *CTCF* in K562 cells and *Ctcf* hemizygous MEFs. It was clear that despite efficient editing of *CTCF* alleles in hypotriploid K562 cells, haploinsufficient levels of *CTCF* protein expression were still maintained in surviving clones. Edited cells exhibited reduced cell proliferation and colony-forming ability consistent with our shRNA knockdown of *CTCF* in K562 cells. However, in *Ctcf* hemizygous MEFs that only contain one coding allele of *Ctcf*, we were unable to completely abolish *Ctcf* protein expression using CRISPR. Accordingly, in surviving clones we detected a high incidence of in-frame microdeletions in *Ctcf*. As these microdeletions occur in the intrinsically disordered N-terminus we do not expect them to significantly impact *Ctcf* function. Induced frameshift deletions were also likely to produce truncated *CTCF* proteins initiating from alternate in-frame ATG start codons within the N-terminus. These results

confirm that CTCF is absolutely required for somatic cell viability and that CTCF cannot be completely inactivated in cells. Interestingly, residual CTCF protein levels can be depleted to minimal amounts in the cell before viability is significantly impacted. Targeted degradation of CTCF in mouse embryonic stem (ES) cells using an auxin-inducible degron system highlighted that endogenous CTCF protein levels could be decreased by up to 99% for at least 2 d duration without a significant impact on cell proliferation or viability [48]. Acute depletion impacted on CTCF looping and insulation of TAD regions, but genomic compartmentalisation was maintained [48]. CTCF-null ES cells can progress to the blastocyst stage (E3.5) purely via retention of maternal CTCF mRNA, but exhibit peri-implantation lethality by E4.5–E5.5 [2].

We have also shown that CTCF has a dose-dependent impact on embryonic development, as even haploinsufficient levels of Ctf protein affected embryonic development in mice. We observed heterozygous *Ctf* mice being born at sub-Mendelian rates (24–28%) compared to WT littermates, which was previously suggested by a study using mice with a conditionally targeted *Ctf* allele, but was not thoroughly quantified or statistically verified [3]. Interestingly, the same *Ctf* hemizygous mice are born at normal Mendelian rates on a mixed C57Bl/6J:129SvJ background [6], indicating that the strain background is an important factor to consider in any *Ctf* genetic deficiency studies in mice. *Ctf* heterozygosity also impaired normal mouse weight gain during adult development for up to seven weeks, and which then remained constant in aged male mice. As these mice were fed a normal chow diet, we could not determine whether *Ctf* heterozygosity impairs body weight control, metabolism or nutrient signaling pathways. In future studies we will examine glucose and insulin levels in plasma after feeding-fasting cycles, body tissue composition using dual energy X-ray absorption as well as studying the impact of different chow compositions on *Ctf* heterozygous mouse development.

CTCF hemizygous mice are more susceptible than WT mice to spontaneous cancer development, as well as radiation- and chemically-induced cancers [6]. Tumours in *Ctf*<sup>+/-</sup> mice compared to WT mice also exhibit increased aggressiveness in terms of invasion, metastatic dissemination and mixed epithelial/mesenchymal differentiation, confirming CTCF as a haploinsufficient tumour suppressor [6]. Our current findings showing an increase in cell proliferation, colony forming ability and numbers of cycling cells in p53-shRNA immortalised *Ctf*<sup>+/-</sup> MEFs support these conclusions. Furthermore, CTCF depletion can increase genomic instability by hindering homologous recombination repair of DNA double-stranded breaks and cause hypersensitivity to DNA damage [46]. As a result, our observation of an increase in DNA damage after UV treatment of *Ctf*<sup>+/-</sup> MEFs is consistent with impaired DNA repair. These data may explain why CTCF haploinsufficient MEFs in the context of additional genetic hits to p53, exhibited a number of cancer hallmarks.

The TCGA UCEC cohort consisting of low- and high-grade endometrioid carcinomas and serous tumours were genetically defined into four categories [28]. The majority of CTCF somatic mutations occurred in POLE ultramutated, MSI hypermutated and copy-number low cancers, whilst copy-number high cancers with a serous-like pathology harboured TP53 mutations [28]. This was consistent with our analysis showing mutually exclusivity between CTCF and TP53 mutations in endometrial cancer. CTCF haploinsufficiency due to CTCF copy loss results in poorer survival outcomes in patients with endometrioid UCEC [6], as well as serous UCEC [32]. Our analyses provide insight into the molecular pathophysiology underlying these observations. Since CTCF is known to co-ordinate higher-order chromatin architecture to facilitate interactions between transcription regulatory sequences, our data reinforces the impact that CTCF haploinsufficient loss imparts in endometrial cancer via transcriptional regulation. CTCF haploinsufficiency results in differential regulation of genes located at specific loci, particularly on chromosomes 17 and 20, including cytoband 17p13.1 containing the TP53 locus. Whilst this is not reflected in a significant change in TP53 mRNA expression once accounting for the TP53 mutation status of patient samples, genes involved in p53-mediated signal transduction are impacted. Genes involved in DNA methylation were also differentially regulated in CTCF-altered endometrial cancers. Molecular genetic analysis of

*Ctcf*<sup>+/-</sup> mice showed DNA methylation instability compared to wild type mice [6]. Divergent CpG methylation due to *Ctcf* hemizygoty was restricted to specific loci with regions within a 2 kb window surrounding divergent CpGs exhibiting a generalised pattern of DNA hypermethylation [6]. In humans with heterozygous *CTCF* mutations exhibiting an intellectual disability, specific *CTCF* sites exhibited DNA hypermethylation [45]. This epigenetic dysregulation may offer an explanation as to why differential gene expression was observed at particular chromosomal loci in *CTCF*-altered endometrial cancers.

One possible hallmark of *CTCF*-altered endometrial cancers is the downregulation of tumour suppressor genes including *PIK3CA*, *CDKN2A*, *CDH6* and *IGF2BP2*. The tumour suppressor *PIK3CA* is ranked fifth after *CTCF* in the most frequently mutated genes in UCEC [6] whilst *CDKN2A* is downregulated in POLE, MSI, and copy-number low cancers compared to high-copy number cancers [28]. *CDH6*, which helps maintain epithelial integrity in the endometrium [49], has been shown to be a putative tumour suppressor in cholangiocarcinoma [50]. *IGF2BP2*, which was the most down-regulated gene in *CTCF*-altered endometrial cancer, was identified as a candidate tumour suppressor gene in a pan-cancer screen for homozygously deleted genes [51]. Loss of *IGF2BP2* staining, which is a feature of endometrioid cancers, but not serous cancers, has been proposed as a biomarker for distinguishing endometrial tumour pathology [52].

A second hallmark of *CTCF*-altered endometrial cancers is the upregulation of estrogen-responsive genes, which includes *KIAA1324*, *MLPH*, *MSX2*, *SPDEF*, *TFF3* and *PIGR*. *CTCF* mutations do not occur in a tumour type-specific manner, but rather they define a subset of hormone-responsive cancers [31]. *CTCF* is a negative regulator of the pioneer factor *FOXA1*, which facilitates estrogen receptor interactions with chromatin in response to estrogen [53,54]. Therefore, in *CTCF* haploinsufficient endometrial tumours, *FOXA1*/ER interactions with chromatin may increase leading to upregulation of estrogen-responsive genes. *KIAA1324*, which is a positive regulator of the autophagy pathway and may protect cells from cell death, was the most upregulated gene in *CTCF*-altered endometrial cancer [55]. *KIAA1324* is a marker of grade I endometrial cancer which decreases with increase in tumour grade and disease stage [56] and is a key member of gene signatures classifying histological subtypes [39,40]. Other estrogen-responsive genes upregulated in *CTCF*-altered cancers, namely *SPDEF*, *TFF3* and *PIGR*, are also components of these gene signatures, indicating that loss of *CTCF* could be an important factor determining endometrial cancer progression and pathology.

## 4. Materials and Methods

### 4.1. Cell Lines

Human erythroid leukaemia (K562) cells were grown in RPMI 1640 medium while human embryonic kidney (HEK293T) and mouse embryonic fibroblast cells were cultured in DMEM. Basal media were supplemented with 10% FCS (*v/v*), penicillin (100 U/mL) and streptomycin (100 µg/mL). All human cell lines were authenticated by short tandem repeat profiling (Cellbank, Westmead, Australia).

### 4.2. Expression Vectors and Antibodies

*CTCF* shRNA knockdown was performed using the pFH1-UTG-*CTCF*shRNA lentivector and the corresponding control shRNA vector expressing *Arabidopsis thaliana* mir-159a [32]. This lentiviral vector contains eGFP and a doxycycline-inducible shRNA. For CRISPR/Cas9 genome editing: plasmid 52628-Bsd-T2A-H2B-mCherry was used to express single guide RNAs (sgRNAs) as a lentivector and was a kind gift from Yifei Liu (Yale Fertility Centre, New Haven, CT, USA). Sense and antisense oligonucleotides encoding the sgRNAs (Supplementary Table S1) were phosphorylated with T4 polynucleotide kinase (New England Biolabs, Ipswich, MA, USA), annealed and then cloned into 52628-Bsd-T2A-H2B-mCherry following *BspMI* digestion. Plasmid 53190-pLV-hUbC-Cas9-T2A-eGFP used for stable expression of a human-codon

optimised Cas9 nuclease [57] was obtained from Addgene (Watertown, MA, USA). For immortalisation of MEFs, pMSCVp53.1224, a retroviral vector encoding a p53 shRNA, a kind gift from Ross Dickins (Walter and Eliza Hall Institute, Melbourne, Australia), was used. Primary antibodies include: rabbit polyclonal antibody against CTCF (1:1000) [24], mouse monoclonal antibodies against CTCF (1:1000) [58],  $\alpha$ -tubulin (1:5000; sc-23948, Santa Cruz, Dallas, TX, USA) and GAPDH (1:5000; ab8245, Abcam, Cambridge, MA, USA). Secondary antibodies include: rabbit or mouse antibodies conjugated to horseradish peroxidase (Abcam, Cambridge, MA, USA; 1:5000).

#### 4.3. Retroviral and Lentiviral Transduction

Viral supernatants were produced by calcium phosphate transfection of HEK293T cells: with pJK3, pCMVtat and pL-VSV-G packaging plasmids used to produce retroviruses; and pRSV-Rev, pMDLg/p.rre and pMD2.VSV-G used to package lentiviruses. Viral supernatants collected after 24–48 h were 0.45  $\mu$ M-filtered and snap-frozen or concentrated by ultracentrifugation for 2 h at 20,000 rpm in a SW28 Beckman rotor. Viral supernatant was resuspended on ice in 10% (*v/v*) FCS/DMEM at 1/100 th of the original volume. Adherent cells ( $1\text{--}5 \times 10^5$ /well) were seeded in six-well plates before addition of fresh medium containing viral supernatant ( $\sim 5 \times 10^5$  transducing units) and Polybrene (4  $\mu$ g/mL; Sigma, Zwijndrecht, The Netherlands) and ‘spin-oculated’ for 90 min at 1500 rpm. The supernatant was replaced with medium 12 h post-transduction and fluorescent cells were purified 24 h later by fluorescence activated cell sorting (FACS; >95% purity on re-analysis) using a FACS Influx (Becton Dickinson, BD, Mountain View, CA, USA). K562 cells ( $\sim 5 \times 10^5$ /mL) in 1 mL medium with 4  $\mu$ g/mL Polybrene were placed in a 5 mL capped FACS tube and transduced with viral supernatant for 90 min by ‘spin-oculation’. The cells were resuspended and incubated at 37 °C for 4 h before removal of viral supernatant. For *in vitro* assays, cells were either plated out immediately or allowed to recover after sorting for 48–72 h in medium containing 100  $\mu$ g/mL Normocin (InvivoGen, Toulouse, France).

#### 4.4. CRISPR/Cas9 Genome Editing, Validation and Molecular Genetic Analysis

SgRNAs targeting the first coding exon of human CTCF and mouse *Ctcf* (exon 3) were designed using the Zhang lab CRISPR design tool (crispr.mit.edu). SgRNAs targeting the adeno-associated virus integration site 1 (AAVS1) and the *Rosa26* locus were used as negative control guides in human and mouse cells respectively. We used lentiviral vectors to co-express a sgRNA with mCherry, as well as a 3XFLAG-tagged Cas9 nuclease 2A-peptide linked to eGFP. Transduced cells were FACS-enriched for eGFP<sup>+</sup>mCherry<sup>+</sup> cells after 48 h from which gDNA was extracted from pools after 6 d for a T7 Endonuclease I assay to detect Cas9-directed DNA cleavage. We also isolated single eGFP<sup>+</sup>mCherry<sup>+</sup> cells by FACS into 96-well plates and expanded them before isolation of genomic DNA and whole cell lysates. Genomic DNA was isolated using the Purelink Genomic DNA Extraction kit (ThermoFisher) and PCR primers were used to amplify across the targeted region (see Supplementary Table S1). PCR amplicons were denatured and re-annealed to allow heteroduplex formation, then digested with T7 Endonuclease I (New England Biolabs, Ipswich, MA, USA) according to manufacturer’s instructions and then resolved using DNA gel electrophoresis. We PCR-amplified CTCF exon 3 from genomic DNA isolated from K562 and *Ctcf*<sup>+/-</sup> MEF clones, which had been subjected to CRISPR/Cas9-mediated gene editing, using Platinum Taq (Thermo Fisher Scientific, Waltham, MA, USA). Amplicons were ligated into pGEM-T-Easy (Promega, Madison, WI, USA) and then transformed into *E. coli*. Each clonal amplicon was then confirmed using Sanger sequencing in both directions.

#### 4.5. Isolation of Mouse Embryonic Fibroblasts

All animal experiments were performed in accordance with an approved institutional animal ethics protocol from the Royal Prince Alfred Hospital Animal Welfare Committee (SSWAHS #2013/046 approved 5 August 2013). *Ctcf*<sup>+/-</sup> mice were obtained on a mixed C57Bl/6:129SvJ

background from the Fred Hutchinson Cancer Research Centre (Seattle, WA, USA) [2]. These mice have had the complete coding region of one *Ctcf* allele replaced with a loxP-flanked cassette containing a *pgk* promoter and *neo* gene, designated *Ctcf*<sup>+/pgk<sup>neo</sup></sup> (Figure 3A). Mice homozygous for this allele (*Ctcf*<sup>pgk<sup>neo</sup>/pgk<sup>neo</sup></sup>) exhibit embryonic lethality prior to embryo implantation [2]. Mice were backcrossed at least 10 generations onto C57Bl/6 mice from the Animal Resources Centre (Perth, WA, USA) before beginning phenotyping studies. Timed matings were performed with *Ctcf*<sup>+/pgk<sup>neo</sup></sup> male mice and C57Bl/6 females and female mice were checked daily for vaginal plugs. At 13.5 dpc, pregnant females were euthanised by CO<sub>2</sub> asphyxiation. The uterine horns were removed and the foetuses released whilst immersed in PBS. Each pup was removed from its amniotic sac, decapitated and fetal liver removed. The carcasses were minced with a scalpel and then incubated in trypsin/EDTA solution (Invitrogen, Basel, Switzerland). The tissue fragments were triturated to break up clumps, and then concentrated using centrifugation to remove trypsin. Fresh trypsin was added to create a homogeneous solution of cellular material. The trypsin was inactivated in excess DMEM medium containing 10% (*v/v*) FCS and then the centrifugation step repeated. The MEFs were plated in 15 cm plates and allowed to grow for 2–3 d until there were sufficient adherent cells for cryopreservation. The remaining MEFs were transduced with MSCVp53.1224 retroviral supernatant for immortalisation. Cells were selected in 50 µg/mL hygromycin (Roche Life Science, Mannheim, Germany) to enrich for immortalised MEFs and then frozen down after 1–2 passages (P2-P3 MEFs). *Ctcf*<sup>+/pgk<sup>neo</sup></sup> mice and MEFs were genotyped according to primers listed in Supplementary Table S1.

#### 4.6. Western Blot Analysis

Protein extracts were prepared with cell lysis buffer containing 20 mM Tris-HCl (pH 8), 150 mM NaCl, 1% (*v/v*) Triton X-100, 0.1% (*v/v*) SDS, 0.5% (*w/v*) sodium deoxycholate, and EDTA-free protease inhibitor cocktail (cOmplete, Roche Life Science, Mannheim, Germany), prior to separation using denaturing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Proteins were transferred onto PVDF membranes in a semi-dry transfer apparatus before immunoblotting. Membranes were blocked in PBS/0.1% (*v/v*) Tween 20 containing 20% (*v/v*) BlokHen (AvesLab, Portland, OR, USA) or PBST containing 0.3% (*w/v*) BSA, 1% (*w/v*) polyvinylpyrrolidone and 1% (*v/v*) PEG (mw 3350). Protein expression was detected using primary antibodies followed by washing and staining with appropriate secondary antibodies conjugated to horseradish peroxidase (HRP). The HRP substrate SuperSignal<sup>®</sup> Chemiluminescent Substrate (Pierce) was detected on a Kodak ImagerStation 4000R Pro (Woodbridge, CT, USA) or BioRad Chemidoc Touch (Hercules, CA, USA). Blots were stripped with ReBlot Plus (Merck Millipore, Guyancourt, France) prior to re-probing with protein loading control antibodies. Densitometric analysis of bands from three independent blots was performed using ImageJ (National Institutes of Health, University of Wisconsin, USA).

#### 4.7. Cell Assays

Cell proliferation was assessed by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (Merck Millipore, Guyancourt, France). Adherent cells (1000/well) or suspension cells (5000/well) were plated in triplicate in a 96-well plate and proliferation was assessed over 4 d by the addition of MTT at 37 °C overnight. The reaction was quenched with isopropanol/HCl and absorbance measured at 572 nm using a Polarstar Omega plate reader (BMG Labtech, Durham, NC, USA). The clonogenic capacity of adherent cells was measured by plating 1000 cells/10 cm plate in triplicate and incubating for 8–10 d. Cells were washed with PBS, fixed with ice-cold methanol and stained with Giemsa Stain (Sigma, Zwijndrecht, The Netherlands) diluted 1:20 in triple-distilled water before scoring. The clonogenic capacity of K562 cells was measured by plating 5000 cells diluted in Iscove's Modified Dulbecco Medium (Life Technologies, Rockville, MD, USA) containing 3 mL Methocult GF H4230 (Stem Cell Technologies,

Vancouver, BC, Canada) onto 35 mm gridded plates in triplicate and incubating for 8–10 d. To assess UV-induced apoptosis, cells ( $1 \times 10^5$ /well in a 12-well plate) were plated in triplicate. The following day, medium was removed from attached cells and replaced with PBS. Plates with lids removed were placed in a Stratalinker UV Crosslinker (Stratagene, La Jolla, CA, USA) and exposed to UVC irradiation (2000  $\mu$ J for MEFs, 4000  $\mu$ J for K562 cells) and allowed to recover for 18 h. Cells were harvested and stained with anti-Annexin V-APC (BD Biosciences, San Jose, CA, USA) according to the manufacturer's protocol and with propidium iodide (PI) solution (5  $\mu$ g/mL, Sigma, Zwijndrecht, The Netherlands). Cells were analysed on a Fortessa flow cytometer (BD Biosciences, San Jose, CA, USA) with analysis performed using FlowJo 9.7.6 software (Treestar, Ashland, OR, USA). Cell viability was measured after addition of PI and then analysed by flow cytometry. The viable population represents the Annexin V<sup>-</sup>PI<sup>-</sup> cells; the apoptotic population represents the Annexin V<sup>+</sup>PI<sup>-</sup> and Annexin V<sup>+</sup>PI<sup>+</sup> cells combined.

#### 4.8. Cell Cycle Analysis

For cell cycle analysis by DNA content, cells were washed with PBS prior to fixation in ice-cold 70% ethanol and stored overnight at 4 °C. Post-fixation, cells were washed twice with PBS to remove all traces of ethanol prior to staining with a solution containing PI (20  $\mu$ g/mL), 0.1% (*v/v*) Triton X-100 and RNase-A (200  $\mu$ g/mL). Cells were incubated for 15 min in the dark at room temperature and analysed on a Canto-II flow cytometer (BD Biosciences, San Jose, CA, USA). Cell cycle analysis was undertaken using FlowJo 9.7.6 cell cycle modelling software (Treestar, Ashland, OR, USA) by applying the Dean-Jett-Fox algorithm.

#### 4.9. Bioinformatics Analysis

Gene expression and somatic mutation data from the uterine corpus endometrial carcinoma dataset [28] was downloaded from cBioPortal. Of the 500 samples described, 240 contain matched sequencing and copy number alteration data. Statistical tests already conducted on these data were also downloaded, including Student's test (*p*) and Benjamini-Hochberg adjusted *p*-values (*q*). *CTCF*-altered cancers included those with somatic mutations (missense, nonsense, frameshift) and deep deletions (*n* = 45). Normal *CTCF* included shallow deletions and non-mutant samples (*n* = 178). The *CTCF*-altered gene signature used in subsequent analysis includes all differentially expressed genes (*q* < 0.05).

## 5. Conclusions

We examined *CTCF* essentiality and haploinsufficiency in somatic cells and mice using various molecular genetic techniques and models. Despite achieving efficient genome editing of *CTCF* using CRISPR the inability to obtain complete ablation of *CTCF* expression reinforces its requirement. In all cases, cellular fitness in *CTCF*-targeted cells was compromised leading to surviving cells compensating with reduced *CTCF* protein expression or truncated *CTCF* protein variants. Consequently cell proliferation, colony-forming ability and cycling cells were reduced. However, in the presence of additional genetic hits, such as in p53, *CTCF* haploinsufficient cells exhibited known cancer hallmarks, namely increased proliferation and reduced cell cycle control. In human endometrial cancer datasets, we identified a unique gene signature in *CTCF* haploinsufficient cancers arising from differential gene expression at specific loci. Downregulation of tumour suppressor genes and upregulation of estrogen-responsive genes may be a molecular feature of *CTCF*-altered endometrial cancers. Our study clearly demonstrates that *CTCF* is a haploinsufficient tumour suppressor gene that is essential for somatic cell viability and protects against cancer. As the master of weaver of the genome *CTCF* plays an essential role in chromatin organisation, the full impact of *CTCF* haploinsufficiency on three-dimensional chromatin architecture remains to be elucidated.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/19/12/3832/s1>.

**Author Contributions:** C.G.B. conceived the study, performed mouse experiments, designed the CRISPR/Cas9 targeting strategy, analysed data and wrote the manuscript. J.R. supervised research and governance and provided scientific guidance; C.M. and Y.F. performed cellular biology experiments; K.B. genotyped mice and performed phenotypic characterisation of mice; C.S. assisted in drafting the manuscript; and G.F., V.V.L., and D.I.L. provided reagents and discussion.

**Funding:** Financial support was provided by Tour de Cure (Scott Canner Research Fellowship) to C.G.B. and for research grants to C.G.B. and J.E.J.R.; Tour de Rocks project support to C.G.B.; National Health and Medical Research Council grants (#507776 and #1128748 to J.E.J.R.); Cancer Council NSW project grants (RG11-12, RG14-09) to J.E.J.R. and C.G.B.; and support grants from Cure The Future Foundation and anonymous foundation.

**Acknowledgments:** The authors wish to thank the Centenary Institute Advanced Cytometry Facility for FACS and the Centenary Institute animal house staff for animal husbandry.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

CTCF	CCCTC-binding factor
CRISPR	Clustered regularly interspaced short palindromic repeats
MSI	Microsatellite instability
MEF	mouse embryonic fibroblast
FACS	fluorescence-activated cell sorting

## References

- Phillips, J.E.; Corces, V.G. CTCF: Master weaver of the genome. *Cell* **2009**, *137*, 1194–1211. [[CrossRef](#)] [[PubMed](#)]
- Moore, J.M.; Rabaia, N.A.; Smith, L.E.; Fagerlie, S.; Gurley, K.; Loukinov, D.; Disteche, C.M.; Collins, S.J.; Kemp, C.J.; Lobanenkov, V.V.; et al. Loss of maternal CTCF is associated with peri-implantation lethality of *Ctcf* null embryos. *PLoS ONE* **2012**, *7*, e34915. [[CrossRef](#)] [[PubMed](#)]
- Heath, H.; De Almeida, C.R.; Sleutels, F.; Dingjan, G.; Van de Nobelen, S.; Jonkers, I.; Ling, K-W.; Gribnau, J.; Renkawitz, R.; Grosveld, F.; et al. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *EMBO J.* **2008**, *27*, 2839–2850. [[CrossRef](#)] [[PubMed](#)]
- Soshnikova, N.; Montavon, T.; Leleu, M.; Galjart, N.; Duboule, D. Functional Analysis of CTCF During Mammalian Limb Development. *Dev. Cell* **2010**, *19*, 819–830. [[CrossRef](#)] [[PubMed](#)]
- Watson, L.A.; Wang, X.; Elbert, A.; Kernohan, K.D.; Galjart, N.; Bérubé, N.G. Dual Effect of CTCF Loss on Neuroprogenitor Differentiation and Survival. *J. Neurosci.* **2014**, *34*, 2860–2870. [[CrossRef](#)] [[PubMed](#)]
- Kemp, C.J.; Moore, J.M.; Moser, R.; Bernard, B.; Teater, M.; Smith, L.E.; Rabaia, N.A.; Gurley, K.E.; Guinney, J.; Busch, S.E.; et al. CTCF Haploinsufficiency Destabilizes DNA Methylation and Predisposes to Cancer. *Cell Rep.* **2014**, *7*, 1020–1029. [[CrossRef](#)] [[PubMed](#)]
- Merkenschlager, M.; Odom, D.T. CTCF and Cohesin: Linking Gene Regulatory Elements with Their Targets. *Cell* **2013**, *152*, 1285–1297. [[CrossRef](#)] [[PubMed](#)]
- Sofueva, S.; Yaffe, E.; Chan, W.C.; Georgopoulou, D.; Vietri, R.M.; Mira-Bontenbal, H.; Pollard, S.M.; Schroth, G.P.; Tanay, A.; Hadjur, S. Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* **2013**, *32*, 3119–3129. [[CrossRef](#)] [[PubMed](#)]
- Zuin, J.; Dixon, J.R.; Van der Reijden, M.I.; Ye, Z.; Kolovos, P.; Brouwer, R.W.; Van de Corput, M.C.P.; Van de Werken, H.J.G.; Knoch, T.A.; Van IJcken, W.F.J.; et al. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 996–1001. [[CrossRef](#)] [[PubMed](#)]
- Dixon, J.R.; Selvaraj, S.; Yue, F.; Kim, A.; Li, Y.; Shen, Y.; Hu, M.; Liu, J.S.; Ren, B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **2012**, *485*, 376–380. [[CrossRef](#)] [[PubMed](#)]
- Nora, E.P.; Lajoie, B.R.; Schulz, E.G.; Giorgetti, L.; Okamoto, I.; Servant, N.; Piolot, T.; Van Berkum, N.L.; Meisig, J.; Sedat, J.; et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **2012**, *485*, 381–385. [[CrossRef](#)] [[PubMed](#)]

12. Seitan, V.; Faure, A.; Zhan, Y.; McCord, R.; Lajoie, B.; Ing-Simmons, E.; Lenhard, B.; Giorgetti, L.; Heard, E.; Fisher, A.G.; et al. Cohesin-based chromatin interactions enable regulated gene expression within pre-existing architectural compartments. *Genome Res.* **2013**, *23*, 2066–2077. [[CrossRef](#)] [[PubMed](#)]
13. Vietri Rudan, M.; Barrington, C.; Henderson, S.; Ernst, C.; Odom, D.T.; Tanay, A.; Hadjur, S. Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell Rep.* **2015**, *10*, 1297–1309. [[CrossRef](#)] [[PubMed](#)]
14. Symmons, O.; Uslu, V.V.; Tsujimura, T.; Ruf, S.; Nassari, S.; Schwarzer, W.; Ettwiller, L.; Spitz, F. Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* **2014**, *24*, 390–400. [[CrossRef](#)] [[PubMed](#)]
15. Tang, Z.; Luo, O.J.; Li, X.; Zheng, M.; Zhu, J.J.; Szalaj, P.; Trzaskoma, P.; Magalska, A.; Wlodarczyk, J.; Rusczycki, B.; et al. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell* **2015**, *163*, 1611–1627. [[CrossRef](#)] [[PubMed](#)]
16. Lupiáñez, D.G.; Kraft, K.; Heinrich, V.; Krawitz, P.; Brancati, F.; Klopocki, E.; Horn, D.; Kayserili, H.; Opitz, J.M.; Laxova, R.; et al. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell* **2015**, *161*, 1012–1025. [[CrossRef](#)] [[PubMed](#)]
17. Guo, Y.; Xu, Q.; Canzio, D.; Shou, J.; Li, J.; Gorkin, D.U.; Jung, I.; Wu, H.; Zhai, Y.; Tang, Y.; et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* **2015**, *162*, 900–910. [[CrossRef](#)] [[PubMed](#)]
18. Narendra, V.; Rocha, P.P.; An, D.; Raviram, R.; Skok, J.A.; Mazzoni, E.O.; Reinberg, D. Transcription. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **2015**, *347*, 1017–1021. [[CrossRef](#)] [[PubMed](#)]
19. Witcher, M.; Emerson, B.M. 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. [[CrossRef](#)] [[PubMed](#)]
20. Liu, Q.; Thoms, J.A.; Nunez, A.C.; Huang, Y.; Knezevic, K.; Packham, D.; Poulos, R.C.; Williams, R.; Beck, D.; Hawkins, N.J.; et al. Disruption of a -35kb enhancer impairs CTCF binding and MLH1 expression in colorectal cells. *Clin. Cancer Res.* **2018**, *24*, 4602–4611. [[CrossRef](#)] [[PubMed](#)]
21. Flavahan, W.A.; Drier, Y.; Liau, B.B.; Gillespie, S.M.; Venteicher, A.S.; Stemmer-Rachamimov, A.O.; Suvà, M.L.; Bernstein, B.E. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **2016**, *529*, 110–114. [[CrossRef](#)] [[PubMed](#)]
22. Umer, H.M.; Cavalli, M.; Dabrowski, M.J.; Diamanti, K.; Kruczyk, M.; Pan, G.; Komorowski, J.; Wadelius, C. A Significant Regulatory Mutation Burden at a High Affinity Position of the Ctf Motif in Gastrointestinal Cancers. *Hum. Mutat.* **2016**, *37*, 904–913. [[CrossRef](#)] [[PubMed](#)]
23. Guo, Y.A.; Chang, M.M.; Huang, W.; Ooi, W.F.; Xing, M.; Tan, P.; Skanderup, A.J. Mutation hotspots at CTCF binding sites coupled to chromosomal instability in gastrointestinal cancers. *Nat. Commun.* **2018**, *9*, 1520. [[CrossRef](#)] [[PubMed](#)]
24. 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. Cell growth inhibition by the multifunctional multivalent zinc-finger factor CTCF. *Cancer Res.* **2001**, *61*, 6002–6007. [[PubMed](#)]
25. Tiffen, J.C.; Bailey, C.G.; Marshall, A.D.; Metierre, C.; Feng, Y.; Wang, Q.; Watson, S.L.; Holst, J.; Rasko, J.E.J. The cancer-testis antigen BORIS phenocopies the tumor suppressor CTCF in normal and neoplastic cells. *Int. J. Cancer* **2013**, *133*, 1603–1613. [[CrossRef](#)] [[PubMed](#)]
26. 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. Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter its DNA-binding specificity. *Cancer Res.* **2002**, *62*, 48–52. [[PubMed](#)]
27. Mullighan, C.G.; Zhang, J.; Kasper, L.H.; Lerach, S.; Payne-Turner, D.; Phillips, L.A.; Heatley, S.L.; Holmfeldt, H.; Collins-Underwood, J.R.; Ma, J.; et al. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* **2011**, *471*, 235–239. [[CrossRef](#)] [[PubMed](#)]
28. 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. Integrated genomic characterization of endometrial carcinoma. *Nature* **2013**, *497*, 67–73. [[PubMed](#)]
29. Kandoth, C.; McLellan, M.D.; Vandin, F.; Ye, K.; Niu, B.; Lu, C.; Xie, M.; Zhang, Q.; McMichael, J.F.; Wyczalkowski, M.A.; et al. Mutational landscape and significance across 12 major cancer types. *Nature* **2013**, *502*, 333–339. [[CrossRef](#)] [[PubMed](#)]

30. Zigelboim, I.; Mutch, D.G.; Knapp, A.; Ding, L.; Xie, M.; Cohn, D.E.; Goodfellow, P.J. High Frequency Strand Slippage Mutations in CTCF in MSI-Positive Endometrial Cancers. *Hum. Mutat.* **2013**, *35*, 63–65. [[CrossRef](#)] [[PubMed](#)]
31. Ciriello, G.; Miller, M.L.; Aksoy, B.A.; Senbabaoglu, Y.; Schultz, N.; Sander, C. Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* **2013**, *45*, 1127–1133. [[CrossRef](#)] [[PubMed](#)]
32. Marshall, A.D.; Bailey, C.G.; Champ, K.; Vellozzi, M.; O'Young, P.; Metierre, C.; Feng, Y.; Thoeng, A.; Richards, A.M.; Schmitz, U. CTCF genetic alterations in endometrial carcinoma are pro-tumorigenic. *Oncogene* **2017**, *36*, 4100–4110. [[CrossRef](#)] [[PubMed](#)]
33. Mummert, S.K.; Lobanenkova, V.A.; Feinberg, A.P. Association of chromosome arm 16q loss with loss of imprinting of insulin-like growth factor-II in Wilms tumor. *Genes Chromosomes Cancer* **2005**, *43*, 155–161. [[CrossRef](#)] [[PubMed](#)]
34. Naumann, S.; Reutzler, D.; Speicher, M.; Decker, H.J. Complete karyotype characterization of the K562 cell line by combined application of G-banding, multiplex-fluorescence in situ hybridization, fluorescence in situ hybridization, and comparative genomic hybridization. *Leuk. Res.* **2001**, *25*, 313–322. [[CrossRef](#)]
35. Lek, M.; Karczewski, K.J.; Minikel, E.V.; Samocha, K.E.; Banks, E.; Fennell, T.; Ware, J.S.; Hill, A.J.; Cummings, B.B.; Tukiainen, T.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [[CrossRef](#)] [[PubMed](#)]
36. Hart, T.; Chandrashekar, M.; Aregger, M.; Steinhart, Z.; Brown, K.R.; MacLeod, G.; Mis, M.; Zimmermann, M.; Fradet-Turcotte, A.; Sun, S.; et al. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **2015**, *163*, 1515–1526. [[CrossRef](#)] [[PubMed](#)]
37. Wang, T.; Wei, J.J.; Sabatini, D.M.; Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **2014**, *343*, 80–84. [[CrossRef](#)] [[PubMed](#)]
38. Blomen, V.A.; Májek, P.; Jae, L.T.; Bigenzahn, J.W.; Nieuwenhuis, J.; Staring, J.; Sacco, R.; Van Diemen, F.R.; Olk, N.; Stukalov, A.; et al. Gene essentiality and synthetic lethality in haploid human cells. *Science* **2015**, *350*, 1092–1096. [[CrossRef](#)] [[PubMed](#)]
39. Cherniack, A.D.; Shen, H.; Walter, V.; Stewart, C.; Murray, B.A.; Bowlby, R.; Sacco, R.; Van Diemen, F.R.; Olk, N.; Stukalov, A.; et al. Integrated Molecular Characterization of Uterine Carcinosarcoma. *Cancer Cell* **2017**, *31*, 411–423. [[CrossRef](#)] [[PubMed](#)]
40. Sung, C.O.; Sohn, I. The expression pattern of 19 genes predicts the histology of endometrial carcinoma. *Sci. Rep.* **2014**, *4*, 5174. [[CrossRef](#)] [[PubMed](#)]
41. Huang, N.; Lee, I.; Marcotte, E.M.; Hurles, M.E. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* **2010**, *6*, e1001154.
42. Dang, V.T.; Kassahn, K.S.; Marcos, A.E.; Ragan, M.A. Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *Eur. J. Hum. Genet.* **2008**, *16*, 1350–1357. [[CrossRef](#)] [[PubMed](#)]
43. Gregor, A.; Oti, M.; Kouwenhoven, E.N.; Hoyer, J.; Sticht, H.; Ekici, A.B.; Kjaergaard, S.; Rauch, A.; Stunnenberg, H.G.; Uebe, S.; et al. De novo mutations in the genome organizer CTCF cause intellectual disability. *Am. J. Hum. Genet.* **2013**, *93*, 124–131. [[CrossRef](#)] [[PubMed](#)]
44. Bastaki, F.; Nair, P.; Mohamed, M.; Malik, E.M.; Helmi, M.; Al-Ali, M.T.; Rezzak Hamzeh, A. Identification of a novel CTCF mutation responsible for syndromic intellectual disability—A case report. *BMC Med. Genet.* **2017**, *18*, 68. [[CrossRef](#)] [[PubMed](#)]
45. Hori, I.; Kawamura, R.; Nakabayashi, K.; Watanabe, H.; Higashimoto, K.; Tomikawa, J.; Ieda, D.; Ohashi, K.; Negishi, Y.; Hattori, A.; et al. CTCF deletion syndrome: Clinical features and epigenetic delineation. *J. Med. Genet.* **2017**, *54*, 836–842. [[CrossRef](#)] [[PubMed](#)]
46. Hilmi, K.; Jangal, M.; Marques, M.; Zhao, T.; Saad, A.; Zhang, C.; Luo, V.M.; Syme, A.; Rejon, C.; Yu, Z.; et al. CTCF facilitates DNA double-strand break repair by enhancing homologous recombination repair. *Sci. Adv.* **2017**, *3*, e1601898. [[CrossRef](#)] [[PubMed](#)]
47. Ding, L.W.; Sun, Q.Y.; Tan, K.T.; Chien, W.; Mayakonda, A.; Yeoh, A.E.J.; Kawamata, N.; Nagata, Y.; Xiao, J-F.; Loh, X-Y.; et al. Mutational Landscape of Pediatric Acute Lymphoblastic Leukemia. *Cancer Res.* **2017**, *77*, 390–400. [[CrossRef](#)] [[PubMed](#)]
48. Nora, E.P.; Goloborodko, A.; Valton, A.L.; Gibcus, J.H.; Ueberohrn, A.; Abdennur, N.; Dekker, J.; Mirny, L.A.; Bruneau, B.G. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **2017**, *169*, 930–944.e22. [[CrossRef](#)] [[PubMed](#)]

49. Singh, H.; Aplin, J.D. Endometrial apical glycoproteomic analysis reveals roles for cadherin 6, desmoglein-2 and plexin b2 in epithelial integrity. *Mol. Hum. Reprod.* **2015**, *21*, 81–94. [[CrossRef](#)] [[PubMed](#)]
50. Goepfert, B.; Ernst, C.; Baer, C.; Roessler, S.; Renner, M.; Mehrabi, A.; Hafezi, M.; Pathil, A.; Warth, A.; Stenzinger, A.; et al. Cadherin-6 is a putative tumor suppressor and target of epigenetically dysregulated miR-429 in cholangiocarcinoma. *Epigenetics* **2016**, *11*, 780–790. [[CrossRef](#)] [[PubMed](#)]
51. Cheng, J.; Demeulemeester, J.; Wedge, D.C.; Vollan, H.K.M.; Pitt, J.J.; Russnes, H.G.; Pandey, B.P.; Nilsen, G.; Nord, S.; Bignell, G.R.; et al. Pan-cancer analysis of homozygous deletions in primary tumours uncovers rare tumour suppressors. *Nat. Commun.* **2017**, *8*, 1221. [[CrossRef](#)] [[PubMed](#)]
52. Zhang, L.; Liu, Y.; Hao, S.; Woda, B.A.; Lu, D. IMP2 expression distinguishes endometrioid from serous endometrial adenocarcinomas. *Am. J. Surg. Pathol.* **2011**, *35*, 868–872. [[CrossRef](#)] [[PubMed](#)]
53. Hurtado, A.; Holmes, K.A.; Ross-Innes, C.S.; Schmidt, D.; Carroll, J.S. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.* **2011**, *43*, 27–33. [[CrossRef](#)] [[PubMed](#)]
54. Zhang, Y.; Liang, J.; Li, Y.; Xuan, C.; Wang, F.; Wang, D.; Shi, L.; Zhang, D.; Shang, Y. CTCF acts upstream of FOXA1 and demarcates the genomic response to estrogen. *J. Biol. Chem.* **2010**, *285*, 28604–28613. [[CrossRef](#)] [[PubMed](#)]
55. Deng, L.; Feng, J.; Broaddus, R.R. The novel estrogen-induced gene EIG121 regulates autophagy and promotes cell survival under stress. *Cell Death Dis.* **2010**, *1*, e32. [[CrossRef](#)] [[PubMed](#)]
56. Deng, L.; Broaddus, R.R.; McCampbell, A.; Shipley, G.L.; Loose, D.S.; Stancel, G.M.; Pickar, J.H.; Davies, P.J.A. Identification of a novel estrogen-regulated gene, EIG121, induced by hormone replacement therapy and differentially expressed in type I and type II endometrial cancer. *Clin. Cancer Res.* **2005**, *11*, 8258–8264. [[CrossRef](#)] [[PubMed](#)]
57. Kabadi, A.M.; Ousterout, D.G.; Hilton, I.B.; Gersbach, C.A. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Res.* **2014**, *42*, e147. [[CrossRef](#)] [[PubMed](#)]
58. Pugacheva, E.M.; Tiwari, V.K.; Abdullaev, Z.; Vostrov, A.A.; Flanagan, P.T.; Quitschke, W.W.; Loukinov, D.I.; Ohlsson, R.; Lobanenko, V.V. Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding and pre-emptive choices of X chromosome inactivation. *Hum. Mol. Genet.* **2005**, *14*, 953–965. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).