Published in final edited form as:

Hum Mutat. 2014 January ; 35(1): 63–65. doi:10.1002/humu.22463.

High frequency strand slippage mutations in *CTCF* in MSI-positive endometrial cancers

Israel Zighelboim^{1,8}, David G. Mutch¹, Amy Knapp², Li Ding^{3,4,5,6,7}, Mingchao Xie^{3,4,7}, David E. Cohn², and Paul J. Goodfellow²

¹Department of Obstetrics and Gynecology and Siteman Cancer Center Washington University School of Medicine, St. Louis, Missouri, USA

²Department of Obstetrics and Gynecology and James Comprehensive Cancer Center, Columbus, Ohio, USA

³The Genome Institute, Washington University in St. Louis, MO, USA

⁴Department of Genetics, Washington University in St. Louis, MO, USA

⁵Department of Pediatrics, Washington University in St. Louis, MO, USA

⁶Department of Medicine, Washington University in St. Louis, MO, USA

⁷Siteman Cancer Center, Washington University in St. Louis, MO, USA

Abstract

Tumors with defective mismatch repair acquire large numbers of strand slippage mutations including frameshifts in coding sequence repeats. We identified a mutational hotspot, p.T204fs, in the insulator-binding protein (*CTCF*) in MSI-positive endometrial cancers. Although *CTCF* was described as a significantly mutated gene by the endometrial cancer TCGA, the A₇ track variants leading to T204 frameshifts were not reported. Reanalysis of TCGA data using Pindel revealed frequent T204fs mutations, confirming *CTCF* is an MSI target gene and revealed the same frameshifts in tumors with intact mismatch repair. We show that T204fs transcripts are subject to nonsense-mediated decay and as such, T204fs mutations are unlikely to act as dominant negatives. The spectrum and pattern of mutations observed is consistent with *CTCF* acting as a haploinsufficient tumor suppressor.

Keywords

endometrial cancer; MSI; *CTCF*; strand slippage mutation; tumor suppressor

Endometrial cancer is the most common gynecologic malignancy in the United States. The majority of endometrial cancers are histologically classified as endometrioid adenocarcinomas (Siegel and Jemal, 2012). Endometrioid tumors frequently have defective DNA mismatch repair that leads to tumor microsatellite instability (MSI) (Duggan et al, 1994). MSI is associated with elevated mutation frequency and in particular, strand slippage mutations (frameshifts) in simple repeats. Coding sequence repeats in a number of loci have been reported (*MSH3*, *BAX*, *ATR*, *TARBP2* and others)(Swisher et al, 1998; Schwartz et al, 1999; Vassileva et al, 2002; Melo et al, 2009; Zighelboim et al, 2009). However, unlike

Correspondence should be addressed to: Paul J. Goodfellow, 460 W 12th Ave, BRT 808, Columbus OH 43210, Office: 614.685.6911, Fax: 614.688.4181, Paul.Goodfellow@osumc.edu.

⁸Present Address: Gynecologic Oncology, St. Luke's Cancer Care Associates, Assistant Clinical Professor of Obstetrics, Gynecology and Reproductive Services Temple University School of Medicine, Temple University

colon cancers with MSI, for which several MSI target genes have been described (*TGFRB2*, *BAX* and additional loci [Markowitz et al, 1995]), *bona fide* MSI mutational targets in endometrial cancer remain elusive. We identified frequent insertion/deletion mutations in CCCTC-binding insulator and transcription factor (*CTCF*; MIM# 604167) in MSI-positive endometrial cancers.

We performed exome sequencing for primary tumors from eight women with early stage MSI-positive endometrioid adenocarcinomas who subsequently recurred (Supp. Methods). Three tumors carried *CTCF* variants: one with ap.R29W substitution and two with an insertion, p.T204fs (Figure 1a). Neither variant was reported by the endometrial Cancer Genome Atlas (TCGA) (Kandoth et al, 2013).

Analysis of the *CTCF* coding region in an additional 149 primary endometrioid tumors (98 MSI-positive and 51 with normal mismatch repair previously assessed for microsatellite instability as described [Zigheboim et al, 2009; Kandoth et al, 2013]) proved that the A₇ track in exon 3 is a mutational hot spot in cancers with defective DNA mismatch repair: 25% of MSI-positive cases carried either the A insertion mutation initially identified in our exome sequencing studies, or an A deletion mutation (17 cases with p.T204fs*26 and 8 cases with p.T204fs*18, respectively) (Table 1). The overall mutation rate was 35% for MSI-positive tumors and 20% for MSI-negative tumors, with two additional MSI-positive tumors having silent mutations. The vast majority of somatic variants identified are almost certainly loss of function mutations (32 frameshifts, five stops and three splice site mutations) with only six missense mutations seen (four in MSI-positive tumor and two in MSI-negative tumors). It is noteworthy that among the 43 cases with mutations only two had a second mutation: one microsatellite stable tumor carried two frameshift mutations (p.V50fs and p.M285fs) and an MSI-positive tumor had a stop and two missense mutations (p.R457*, p.A694V and p.L722I) (Table 1). *CTCF* mutations were more common in low grade tumors than high grade tumors (37% mutation in grade 1, 28% in grade 2 and 19% in grade 3 cancers) but the difference did not reach statistical significance (p=.052 for trend). There was no association with FIGO stage, patient BMI or age at diagnosis (Supp. Table S1).

The preponderance of frameshift, splice site and stop mutations seen in endometrial cancers is consistent with *CTCF* acting as a tumor suppressor. The recurrent frameshift defects (p.T204fs*18 and p.T204fs*26) could act as dominant negatives and thus explain the absence of a “second hit” in cases with those mutations. Reverse transcriptase PCR and sequencing revealed a near complete absence of mutant transcripts in multiple primary tumors with the A insertion or deletion, consistent with A₇ insertion/deletion mutation nonsense mediated decay (Figure 1b). Primary tumors carrying stop or other frameshift mutations showed similar nonsense mediated decay of the mutant transcripts, whereas a missense change (p.H373L, c.1118A>T) was present in tumor mRNA.

The absence of mutant transcripts, coupled with the paucity of second mutations in the majority of tumors, suggests haploinsufficiency for *CTCF* in endometrioid endometrial cancers. *PTEN*, which is frequently mutated in endometrial cancer, is a one example of a haploinsufficient tumor suppressor (Di Cristofano et al, 1998; Joshi et al, 2012). Given *CTCF* is an essential gene (Merkenschlager et al, 2013) and its critical role in the global regulation of gene expression, it is possible that complete loss of *CTCF* activity is cell lethal as has been described for *DICER1*, another haploinsufficient tumor suppressor (Kumar et al, 2009).

CTCF p.T204fs mutations, present in 25% of MSI-positive endometrial cancers in our study, were not reported by the endometrial cancer TCGA (Kandoth et al, 2013). Analysis of

TCGA exome sequences for the endometrioid nonultramutators (Kandoth et al, 2013) cases using Pindel (Ye et al, 2009) did, however, reveal somatic p.T204fs mutations at a frequency comparable to that we observed using Sanger sequencing. Eleven of 63 MSI-high cases (17.5%) had an A track insertion or deletion mutation (seven and four respectively). The A₇ insertion mutation was also seen at low frequency in MSI stable tumors (four of 109, 3.7%) (Supp. Table S2). The fact that the p.T204fs*26 mutation is also seen in microsatellite stable tumors further supports the notion that the A track insertions and deletions are selected pathogenic defects. The high frequency of A₇ T204fs mutations in MSI-positive cancers (22% overall) points to *CTCF* as a *bona fide* MSI target gene in endometrial cancers and accounts for the elevated mutation rate in MSI-positive tumors compared to tumors with intact DNA mismatch repair.

There is limited overlap between the mutations we observed and those described by TCGA that reported *CTCF* is a significantly mutated gene in both MSI-positive and microsatellite stable tumors: only five of the 24 different somatic mutations we observed were seen in the TCGA cohort (including p.T204fs*26 and p.T204fs*18 identified by Pindel). The large number of novel *CTCF* mutations speaks further to the importance of *CTCF* in endometrial tumorigenesis.

There is a growing number of genes involved in the regulation of chromatin structure for which somatic and constitutional defects have been associated with human malignancies and developmental abnormalities (van Bokhoven et al, 2011; Kandoth et al, 2013). *ARID1B* is one such example. *ARID1B* is frequently mutated in endometrial cancers and constitutional mutations have been associated with syndromic mental retardation, including Coffin-Siris syndrome (Hoyer et al, 2012; Tsurusaki et al, 2012). Recently *de novo* frameshift *CTCF* mutations were associated with intellectual disability and it is hypothesized that *CTCF* deficiency affects expression of enhancer-regulated genes (Gregor et al, 2013). The functional consequences of *CTCF* defects in endometrial cancer are yet to be determined and the analysis of missense mutations reported here and by the TCGA may be particularly revealing. Absence of *CTCF* binding in a chromatin boundary region upstream of the *CDKN2A* tumor suppressor is associated with p16 silencing (Witcher et al, 2009). Perturbed/reduced *CTCF* activity associated with missense mutations and/or haploinsufficiency could impact expression of a broad range of genes with known or novel roles in endometrial tumorigenesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Matt O'Hern for assistance with mutation testing and Alexis Chassen for help with manuscript preparation. The study was approved by the Human Studies Committee at Washington University in St. Louis, MO and Human Subjects Research Committee at the Ohio State University in Columbus, OH. We are grateful to the patients who volunteered to participate in the study in order to make this work possible.

Grant Sponsor

This work is supported by the National Institutes of Health (R21 CA155674 to P.J.G and U01 HG006517 to L.D.).

References

1. Di Cristofano A, Pesce B, Cordon-Cardo C, et al. *Pten* is essential for embryonic development and tumour suppression. *Nature Genetics*. 1998 Aug;19:348–355. [PubMed: 9697695]

2. Duggan BD, Felix JC, Muderspach LI, et al. Microsatellite instability in sporadic endometrial carcinoma. *Journal of the National Cancer Institute*. 1994; 86(16):1216–1221. [PubMed: 8040889]
3. Goodfellow PJ, Buttin BM, Herzog TJ, et al. Prevalence of defective DNA mismatch repair and MSH6 mutation in an unselected series of endometrial cancers. *Proc Natl Acad Sci U S A*. 2003; 100(10):5908–5913. [PubMed: 12732731]
4. Gregor A, Oti M, Kouwenhoven EN, et al. De novo mutations in the genome organizer CTCF cause intellectual disability. *Am J Hum Genet*. 2013; 93(1):124–131. [PubMed: 23746550]
5. Hill DA, Ivanovich J, Priest JR, et al. DICER1 mutations in familial pleuropulmonary blastoma. *Science*. 2009; 325(5943):965. [PubMed: 19556464]
6. Hoyer J, Ekici AB, Ende S, et al. Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am J Hum Genet*. 2012; 90(3):565–572. [PubMed: 22405089]
7. Joshi A, Wang H, Jiang G, et al. Endometrial tumorigenesis in Pten(+/-) mice is independent of coexistence of estrogen and estrogen receptor alpha. *Am J Pathol*. 2012; 180(6):2536–2547. [PubMed: 22503752]
8. Kandoth C, Schultz N, Cherniack AD, et al. Integrated genomic characterization of endometrial carcinoma. *Nature*. 2013; 497(7447):67–73. [PubMed: 23636398]
9. Kumar MS, Pester RE, Chen CY, et al. Dicer1 functions as a haploinsufficient tumor suppressor. *Genes Dev*. 2009; 23(23):2700–2704. [PubMed: 19903759]
10. Markowitz S, Wang J, Myeroff L, et al. Inactivation of the type II TGF- β receptor in colon cancer cells with microsatellite instability. *Science*. 1995; 268:1336–1338. [PubMed: 7761852]
11. Melo SA, Ropero S, Moutinho C, et al. A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. *Nat Genet*. 2009; 41(3):365–370. [PubMed: 19219043]
12. Merkenschlager M, Odom DT. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell*. 2013; 152(6):1285–1297. [PubMed: 23498937]
13. Schwartz S Jr, Yamamoto H, Navarro M, et al. Frameshift mutations at mononucleotide repeats in caspase-5 and other target genes in endometrial and gastrointestinal cancer of the microsatellite mutator phenotype. *Cancer Res*. 1999; 59(12):2995–3002. [PubMed: 10383166]
14. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin*. 2012; 62(1):10–29. [PubMed: 22237781]
15. Swisher EM, Mutch DG, Herzog TJ, et al. Analysis of MSH3 in endometrial cancers with defective DNA mismatch repair. *J Soc Gynecol Investig*. 1998; 5(4):210–266.
16. Tsurusaki Y, Okamoto N, Ohashi H, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet*. 2012; 44(4):376–378. [PubMed: 22426308]
17. van Bokhoven H. Genetic and epigenetic networks in intellectual disabilities. *Annu Rev Genet*. 2011; 45:81–104. [PubMed: 21910631]
18. Vassileva V, Millar A, Briollais L, et al. Genes involved in DNA repair are mutational targets in endometrial cancers with microsatellite instability. *Cancer Research*. 2002; 62:4095–4099. [PubMed: 12124347]
19. Witcher M, Emerson BM. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell*. 2009; 34(3):271–284. [PubMed: 19450526]
20. Ye K, Schulz MH, Long Q, et al. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*. 2009; 25(21):2865–2871. [PubMed: 19561018]
21. Zighelboim I, Schmidt AP, Gao F, et al. ATR mutation in endometrioid endometrial cancer is associated with poor clinical outcomes. *J Clin Oncol*. 2009; 27(19):3091–3096. [PubMed: 19470935]

A.

B.

Figure 1. *CTCF* mutations in primary endometrial cancers

A. Novel *CTCF* variants in MSI-positive endometrial cancers. Matched normal and tumor DNAs showing somatic mutations. The p.R29W and p.T204fs*26 mutations (circled) were identified by exome sequencing and validated by Sanger sequencing. High fidelity Taq polymerase PCR was used for exon 3 p.T204fs mutation analysis. Arrows indicate mutations. Methods for mutation detection are included in the Supp. Methods.

B. Nonsense mediated decay of *CTCF* transcripts. Stop mutation (p.S238*) present in tumor DNA is absent from corresponding tumor cDNA (left panel). Missense mutation (p.H373L) is not subject to decay and is present in tumor cDNA (middle panel). Nonsense mediated decay of the A₇ track insertion mutation p.T204fs (right panel) is further evidenced by the dramatic reduction of the T allele at rs6499137 (3' UTR). Arrows indicate mutation or SNP of interest. Primers and conditions are presented in the Supp. Methods.

Table 1

CTCF mutations identified in primary endometrioid endometrial cancers

Mutation		Tumor MMR status	Mutations observed*
	c.610insA	MSI+	17
	c.610delA	MSI+	
	c.1342 C>T	MSS and MSI+	8
	c.626_627delAG	MSI+	2
	c.950_951delCA	MSI+	
	c.1056 G>T	MSI+	Seen in one tumor only
	c.1025 G>A	MSI+	
	c.713 C>A	MSI+	
	c.1161 C>T	MSI+	
	c.1578 C>T	MSI+	
	c.1702-3T>G	MSI+	
All three mutations	c.1369 C>T	MSI+	
	c.2081 C>T	MSI+	
	c.2164 C>A	MSI+	
	c.405delG	MSS	
	c.1437delC	MSS	
	c.1248 T>G	MSS	
	c.1118 A>T	MSS	
	c.1407 C>A	MSS	
	c.422_426AAGTG>T (complex)	MSS	
	c.791+2T>G	MSS	
	c.1702-1G>C	MSS	
Both mutations in	c.149delT	MSS	
	c.853_854AT>G	MSS	

* p.R448*, p.G216_splice c.791+2T>G, and p.N568_splice c.1702-1G>C were reported for endometrial cancer TCGA series (Kandoth, 2013).

Two additional silent mutations identified in two MSI+ tumors (p.S526S, c.1578 C>T and p.T518T, c.1554 C>T).

MSS: microsatellite stable

Note: Mutations listed here for 149 primary tumors do not include p.R29W and 2 instances of p.T204fs*26 mutations identified in initial exome sequencing study.

Mutations are referenced to NM_006565.3

Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence (NM_006565.3). The initiation codon is codon 1.

Variants have been submitted to COSMIC (cancer.sanger.ac.uk/)