


Serial genomic analysis of endometrium supports the existence of histologically indistinct endometrial cancer precursors

Mitzi Aguilar^{1†}, He Zhang^{2†}, Musi Zhang¹, Brandi Cantarell³, Subhransu S Sahoo¹, Hao-Dong Li¹, Ileana C Cuevas¹, Jayanthi Lea^{4,5}, David S Miller^{4,5}, Hao Chen¹, Wenxin Zheng^{1,4,5}, Jeffrey Gagan^{1,4}, Elena Lucas^{1,4} and Diego H Castrillon^{1,4,5*} 

¹ Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX, USA

² Quantitative Biomedical Research Center, Department of Population and Data Sciences, University of Texas Southwestern Medical Center, Dallas, TX, USA

³ Lyda Hill Department of Bioinformatics, University of Texas Southwestern Medical Center, Dallas, TX, USA

⁴ Harold C. Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, TX, USA

⁵ Department of Obstetrics and Gynecology, University of Texas Southwestern Medical Center, Dallas, TX, USA

*Correspondence to: DH Castrillon, UT Southwestern Department of Pathology, 5323 Harry Hines Boulevard, Dallas, TX 75390-9072, USA.

E-mail: diego.castrillon@utsouthwestern.edu

†These authors contributed equally to this study.

Abstract

The endometrium is unique as an accessible anatomic location that can be repeatedly biopsied and where diagnostic biopsies do not extirpate neoplastic lesions. We exploited these features to retrospectively characterize serial genomic alterations along the precancer/cancer continuum in individual women. Cases were selected based on (1) endometrial cancer diagnosis/hysterectomy and (2) preceding serial endometrial biopsies including for some patients an early biopsy before a precancer histologic diagnosis. A comprehensive panel was designed for endometrial cancer genes. Formalin-fixed, paraffin-embedded specimens for each cancer, preceding biopsies, and matched germline samples were subjected to barcoded high-throughput sequencing to identify mutations and track their origin and allelic frequency progression. In total, 92 samples from 21 patients were analyzed, providing an opportunity for new insights into early endometrial cancer progression. Definitive invasive endometrial cancers exhibited expected mutational spectra, and canonical driver mutations were detectable in preceding biopsies. Notably, ≥ 1 cancer mutations were detected prior to the histopathologic diagnosis of an endometrial precancer in the majority of patients. In 18/21 cases, ≥ 1 mutations were confirmed by abnormal protein levels or subcellular localization by immunohistochemistry, confirming genomic data and providing unique views of histologic correlates. In 19 control endometria, mutation counts were lower, with a lack of canonical endometrial cancer hotspot mutations. Our study documents the existence of endometrial lesions that are histologically indistinct but are *bona fide* endometrial cancer precursors.

© 2021 The Authors. *The Journal of Pathology* published by John Wiley & Sons, Ltd. on behalf of The Pathological Society of Great Britain and Ireland.

Keywords: endometrial cancer; endometrial precancer; atypical endometrial hyperplasia/endometrioid intraepithelial neoplasia; DNA sequencing; mutation

Received 8 November 2020; Revised 12 January 2021; Accepted 20 January 2021

No conflicts of interest were declared.

Introduction

Endometrial cancer is the most common malignancy of the female reproductive tract [1]. The majority of endometrial cancers are endometrioid and arise from histologic precursors termed atypical hyperplasia (AH) or endometrioid intraepithelial neoplasia (EIN), depending on the classification system [2–4]. In an idealized endometrial precancer–cancer sequence, estrogen stimulates endometrial gland proliferation, resulting in mild architectural abnormalities (dilated glands without hypercellularity) termed simple hyperplasia or disordered

proliferative endometrium. Simple hyperplasia is a physiologic response to estrogen and ‘fertile soil’ but is not believed to represent a true precancer (i.e. is polyclonal). In time, mutations in cancer driver genes such as *PTEN* produce a *bona fide* monoclonal neoplastic proliferation (AH/EIN) that comprises precancers with elevated cancer risk. The median time for progression of a histologically-recognizable precursor lesion to endometrial cancer is estimated at 6.7 years, although the range is wide (1–24.5 years) [5]. Additional mutations accumulate, aided by genomic instability and clonal evolution, with progression through a histologic continuum of increasing

severity, culminating in invasive endometrial adenocarcinoma capable of metastasis [6].

The Cancer Genome Atlas (TCGA) Program and other systematic cancer genome sequencing studies (MSK-IMPACT) comprehensively cataloged cancer driver mutations [7]. Over 500 endometrial cancers have been subjected to such analysis, showing that they are driven by mutations in tumor suppressors and oncogenes, without recurring chromosomal translocations [8–10]. PI3K pathway components (e.g. *PTEN*, *PIK3CA*, *KRAS*) are the most frequently mutated genes, with other cancer drivers participating in diverse cellular pathways including *CTNNB1* (encoding β -catenin), *ARID1A*, and *FBXW7* [8–14]. Endometrial cancers are categorized into four major classes with differing biological behavior, as defined by *TP53* mutations (associated with chromosomal instability), deficient mismatch repair, *POLE*-driven ultramutation, or absence of these signatures [9].

To improve cancer detection and assessment, there is interest in the study of precancers; for example, the NCI PreCancer Atlas initiative. However, for most cancer types, sampling precancers is difficult, creating logistical challenges [15]. Endometrial cancer is exceptional in that (1) the endometrium can be sampled repeatedly, (2) it is relatively accessible, and (3) the diagnostic procedure (biopsy or curettage) rarely, if ever, extirpates neoplastic lesions. Other anatomic locations (e.g. tube/ovary, pancreas, lung, brain) are less accessible and are not biopsied prior to tumor resection, or are biopsied only once. The detection of a precancer most often leads to extirpation of the lesion (e.g. resection of adenomata during colonoscopy, lumpectomy of breast carcinoma *in situ* following core biopsy, or cone resection following a biopsy diagnosis of cervical high-grade dysplasia), making longitudinal studies difficult [16]. In the uterine cervix or vulva, precancers driven by HPV often resolve spontaneously [17,18], a phenomenon undocumented in the endometrium. These factors could make the endometrium a useful model system for the systematic study of the formation and progression of early precancers [19].

With these features in mind, we developed a platform for high-throughput sequencing (HTS) of endometrial biopsies, inclusive of genes with recurrent mutations in endometrial cancers. We identified women with endometrial cancer with preceding biopsies, ideally over many years and with ≥ 1 early biopsy preceding a diagnosis of AH/EIN. Allelic frequencies were tracked over time. Class-defining mutations such as *TP53* or *POLE* were readily identified. To validate and extend HTS results, immunohistochemistry was performed for commonly mutated genes where mutant clones (even if microscopic) can be detected [*PTEN*, *ARID1A*, *CTNNB1*/ β -catenin, mismatch repair (MMR) factors *MLH1*/*MSH2*/*MSH6*/*PMS2*]. Strikingly, mutant clones could be demonstrated by immunohistochemistry for the majority of patients prior to the initial histopathologic diagnosis of endometrial precancer. Our study provides novel and unique views of endometrial precancer formation and progression, and confirms the existence of *bona fide* endometrial precancers not readily identifiable histologically.

Materials and methods

Sample selection

This was a retrospective study utilizing samples obtained during routine diagnostic and treatment procedures at the Parkland and William P. Clements University Hospitals with retrieval/analysis conducted under UTSW IRB-approved protocol STU112016-062. Original diagnoses in the surgical pathology reports were used for this study [3]. The first diagnosis of atypical hyperplasia was denoted as 'AH/EIN', as per current classification/recommendations (WHO, 2020) [20].

DNA preparation

DNA was prepared from entire tissue sections cut from a single formalin-fixed, paraffin-embedded (FFPE) tissue block. Most biopsies were embedded in a single block; otherwise, the block with the most material likely to be neoplastic based on histology was selected. No macro/microdissection was employed for endometrial biopsies/curettings. Matched germline samples were obtained from hysterectomy tissue blocks confirmed not to harbor metastasis or contaminating carcinoma (ovary/tube/cervix). Tissue sections from even the scantiest biopsy specimens provided sufficient DNA for library preparation (0.2 μ g). In most cases, this required $\leq 5 \times 4 \mu$ m sections, although some scanty specimens required as many as ten sections.

For normal (non-malignant) endometrial controls, hysterectomies for benign indications were selected; endometrial samples were similarly obtained from entire tissue sections, with endometrium macrodissected with a razor blade. This was done to limit subjacent myometrium to less than 10% of the overall sample. Matched germline samples were obtained from other normal tissues present in the hysterectomy (ovary, tube, or cervix). Hysterectomies for benign indications were used to isolate endometrial and matched germline DNA (supplementary material, Table S4).

DNA extraction was performed with the ReliaPrep FFPE System (Promega, Madison, WI, USA) according to the manufacturer's instructions with modifications as described. For deparaffinization, 1 ml of mineral oil was used in a thermomixer (Eppendorf, Enfield, CT, USA) at 80 °C for 30 min (1400 rpm). For sample lysis, 56 °C and 80 °C incubation steps were performed with shaking (700 rpm). For column binding, the aqueous phase was centrifuged at 2000 $\times g$ for 5 min at room temperature. Columns were washed using the same centrifugation parameters. DNA was eluted by centrifugation after 1-min incubation. DNA concentration was quantified using a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA).

Immunohistochemistry (IHC)

Fixation, sectioning, antigen retrieval, blocking, and secondary detection were performed as previously

described [21], with the following antibodies: ARID1A [#12354, rabbit monoclonal antibody (mAb); Cell Signaling Technology, Danvers, MA, USA; 1:200], PTEN (Dako #M326-7, mouse mAb; Agilent Technologies, Santa Clara, CA, USA; 1:150), p53 (Dako #DO-7, prediluted; Agilent), and β -catenin (Dako #IR702, mouse mAb). IHC for four MMR proteins was performed as previously described [22]. Antigen retrieval was performed following the manufacturers' recommendations. Most patients had undergone Lynch/MMR screening by IHC for the four markers. If screening was not performed or results were unavailable in the medical record, IHC for the four markers was performed on the hysterectomy or preceding biopsy showing carcinoma.

Methods for Design of the cancer gene panel and Bioinformatics are provided in supplementary material, Supplementary materials and methods.

Results

Study design, custom gene panel for HTS, and initial demonstration of ability to detect mutations

A hybrid capture panel was designed for cervical, endometrial, and tubo-ovarian cancers. Many genes undergo recurrent mutation at ≥ 1 Müllerian site (e.g. *FBXW7* in cervical + endometrial, *TP53* in endometrial + tubo-ovarian), so a comprehensive panel could be designed

with a relatively small number of genes ($n = 80$) and target size less than 400 kb. Published studies including TCGA and cBioPortal were mined to select relevant gene targets [8,9,23–26]. Genes with a low mutation frequency but well-documented roles in Müllerian carcinomas were included, whereas loci that remain unvalidated and of questionable biological significance (e.g. large loci such as *MUC16*) were excluded (supplementary material, Table S1).

For initial technical validation, two cases showing endometrioid adenocarcinoma in the endometrial biopsy/sample immediately preceding a hysterectomy (termed bx^0) were analyzed (cancers A and B). Cancer A bx^0 showed five mutations (variant alleles), including *KRAS*^{G12C}, *PIK3CA*^{Q546E}, and two *PTEN* mutations including one frameshift, consistent with biallelic inactivation. Cancer B bx^0 showed eight mutations, including two *ARID1A* and two *PTEN* mutations (Figure 1A). Detection of variants was set at a conservative variant allele frequency (AF) threshold of 1% (dashed gray line, Figure 1A). In this and subsequent graphs, cancer is indicated by the ∞ symbol, with AH/EIN indicated by an asterisk. The spectrum of mutations in cancers A and B was typical of endometrioid cancers, both in numbers and in the presence of canonical endometrial cancer drivers [9]. An additional AH/EIN that subsequently underwent hysterectomy was also analyzed, revealing two mutations including a *PTEN* truncating mutation (AH/EIN, Figure 1A). Thus, the HTS workflow proved capable of detecting endometrial precancer/cancer driver mutations.

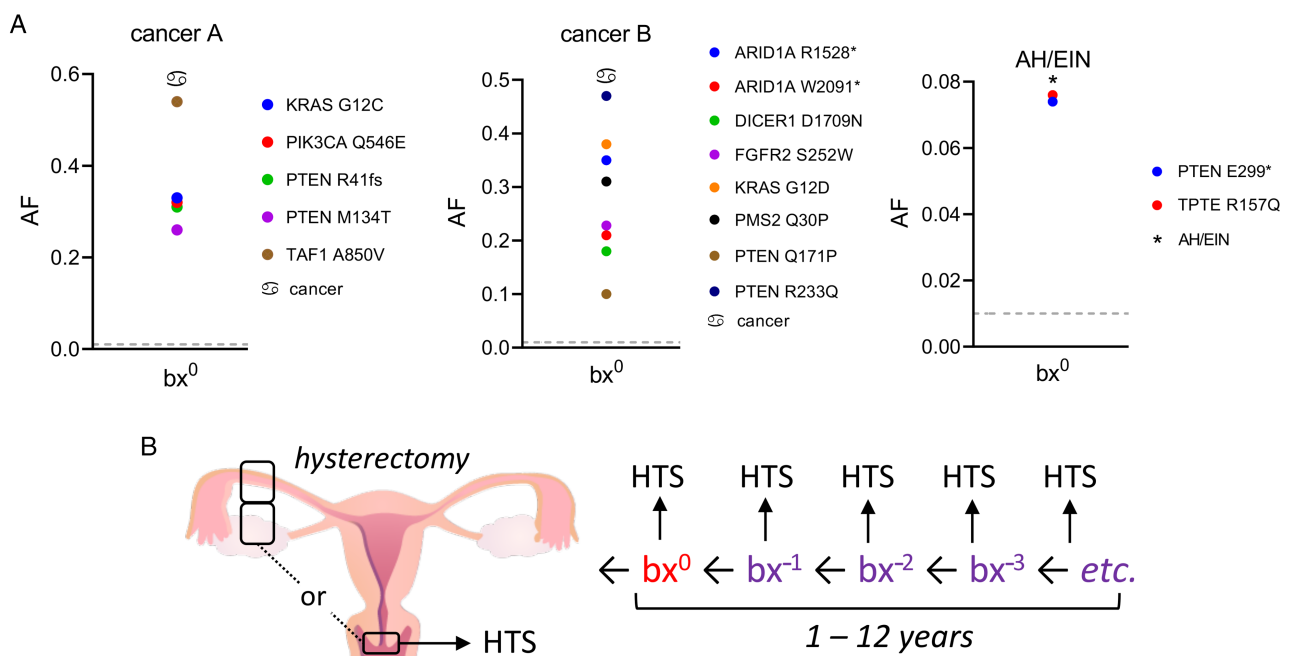


Figure 1. Custom HTS panel can detect mutations in endometrial cancers/precancers and schematic of study design. (A) Analysis of two patients with endometrial cancer and one patient with AH/EIN diagnosed on the biopsy preceding hysterectomy (termed bx^0 throughout this article) ∞ signifies a diagnosis of cancer and * signifies a diagnosis of AH/EIN. The mutation calling threshold at 0.01 AF is shown with a dashed gray line. (B) Schematic of workflows and specimen selection for cases with preceding bx . The biopsies preceding bx^0 are termed bx^{-1} , bx^{-2} , etc. The hysterectomy, which showed endometrial cancer in all cases unless otherwise stated, was used as a source of germline DNA (uninvolved and uncontaminated ovary, tube, or cervix) for more accurate mutation calling. HTS, high-throughput sequencing.

Table 1. Description of cases ($n = 21$).

Patient	Age at hysterectomy (years)	FIGO grade	Stage	Biopsies analyzed (n)
1	29	1	1A	2
2	66	1	1A	2
3	67	1	1A	3
4	71	2	1B	2
5	33	1	1A	4
6	46	1	1A	4
7	68	1	1A	2
8	41	1	1A	2
9	38	1	1A	5
10	66	1	1A	2
11	46	1	1A	4
12	47	3	1B	2
13	44	1	1A	5
14	66	1	1A	9
15	39	1	1A	6
16	51	3	1A	2
17	53	1	1A	2
18	60	1	1A	2
19	52	1	1A	3
20	48	1	1B	6
21	55	AH / EIN		2

AH/EIN, patient with AH/EIN as most severe disease (not diagnosed with adenocarcinoma at any time).

Table 2. Additional details for cases/controls.

Cancer and AH/EIN cases
Total patients = 21
Mean age = 51.7 years
Biopsies analyzed = 71
Hysterectomies with germline control tissues = 21
Biopsies per patient = 3.4 (average)
Breakdown by grade: ACH/EIN = 1; FIGO 1 = 17; FIGO 2 = 1; FIGO 3 = 2
Non-neoplastic controls
Total patients = 19
Mean age = 45.1 years
Hysterectomies = 19
Germline control tissues = 19
Total number of study samples analyzed by HTS = 130

Variant allele frequency (AF) progression in endometrial samples from individuals

Next, we selected $n = 21$ patients with biopsies preceding bx^0 ($n = 20$ invasive adenocarcinomas and $n = 1$ EIN/AH) (Table 1). Samples preceding bx^0 were designated bx^{-1} , bx^{-2} , bx^{-3} , etc. The total number of samples analyzed including bx^0 was 71, with range 2–9 and average 3.4 per patient (Table 2). An overall study design schematic is shown in Figure 1B. Following quality control measures and application of filters, bx^0 single nucleotide variants not present in the germline sample were tabulated for each patient, and AFs in the preceding bx were graphed over time for each patient. In these graphs, ⊗ indicates the first diagnosis of endometrial cancer; an asterisk indicates the first definitive diagnosis of AH/EIN. Concurrent high-dose progestin treatment for a prior diagnosis of AH/EIN or cancer is indicated by a ⊗ symbol (Figure 2A and supplementary material, Figure S1).

AFs showed a propensity to track together across samples from the same patient. For example, in patient (pt) 1, pt2, pt3, and pt4, AFs were high at bx^0 and consistently lower at bx^{-1} (Figure 2A and supplementary material, Figure S1). However, there were many deviations from this principle. In pt5, a *CTNNB1*^{S33Y} (β -catenin) mutation was detectable in all samples but had the highest AF at bx^{-3} (Figure 2A). Also, an *NF1* mutation first appeared at bx^0 . Progestin treatment (⊗) resulted in consistent decreases of AFs. For example, pt9 (Figure 2A) was diagnosed with cancer in the first available sample (bx^{-4}), and subsequent progestin treatment lowered AFs for both *PTEN* mutations to just above the detection threshold. Then, following cessation of progestin, there was a resurgence of AFs for both *PTEN* mutations along with a new *CHD4*^{K52fs} mutation. Such results are readily explained by a marked decrease in gland numbers and epithelial:stromal cell ratios typifying progestin treatment [27,28]. In some cases, stepwise acquisition of mutations was observed. In pt13, the first sample showed only a *CTNNB1*^{D32Y} (exon 3, protein-stabilizing) mutation, followed by *PTEN*^{E43*}, and then an *ARID5B*^{L497fs} mutation in the first bx (bx^0) showing carcinoma (Figure 2A).

Ability of the HTS panel to identify class-defining or germline cancer predisposition variants

Two *POLE* mutations (both V411L) were identified among the $n = 21$ cases (pt2 and pt16, Figure 2A and supplementary material, Figure S1). These cases were associated with high numbers of mutations. Interestingly, all AFs were much lower in the bx^{-1} (4.7 and 3.4 years prior, supplementary material, Table S2), neither of which was diagnosed as AH/EIN. For pt2, *POLE*^{V411L} AF was 0.071 at bx^{-1} (diagnosed as inactive endometrium), indicating that the mutation was already present, but for pt16, *POLE*^{V411L} AF was 0.0 (undetectable) at bx^{-1} (diagnosed as proliferative endometrium, supplementary material, Table S3). These findings might suggest that *POLE* endometrial cancers progress rapidly, but definitive conclusions cannot be drawn due to the small number of cases and potential sampling issues. Pt16's carcinoma also showed deficient MMR (dMMR) (*MSH2/MSH6* loss by immunohistochemistry), and her bx^0 showed the highest number of mutations ($n = 70$) among all patients. This finding is consistent with studies showing that dMMR and *POLE* mutations sometimes co-occur and synergize by abrogating the MMR pathway that corrects mutations introduced by the mutant error-prone *pol* ϵ [29]. dMMR and *POLE* cases exhibited higher mutation counts than did non-dMMR/non-*POLE* cases (Figure 2B); ≥ 9 mutations correlated significantly with dMMR or *POLE* status ($p = 0.0069$, Fisher's exact test). The only exception (pt12) was MMR intact by immunohistochemistry (IHC) but harbored an *MSH6*^{G557A} mutation of unknown biological significance (supplementary material, Figure S1).

A germline mutation detection algorithm for clinically-actionable cancer predisposition genes per the American College of Medical Genetics and Genomics [30]

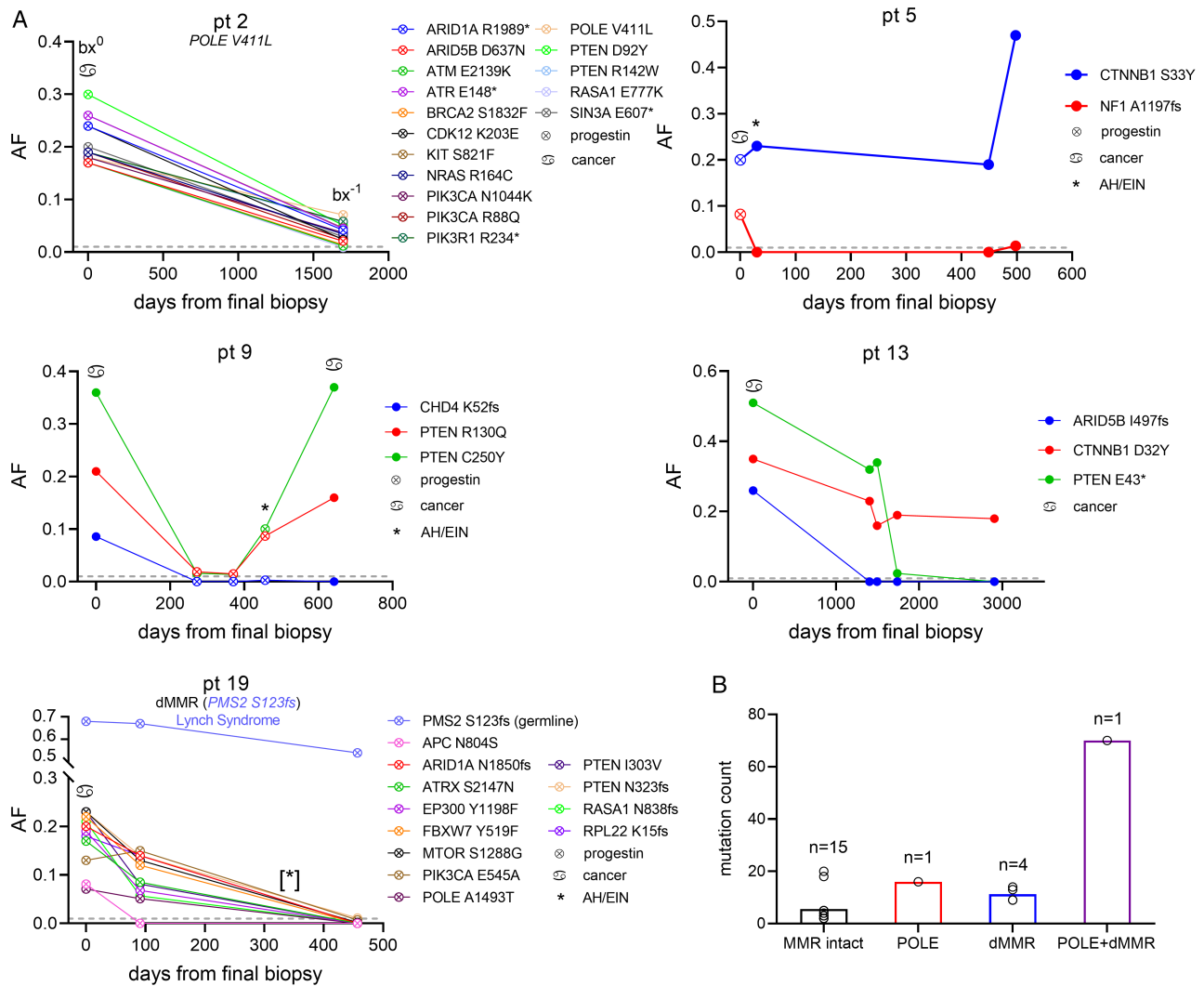


Figure 2. Variant allele frequencies in bx^0 and preceding biopsies. (A) AF graphs. Y-axis = allele frequencies per HTS data; X-axis = time scale in days relative to the final biopsy (bx^0). The leftmost data point for each graph corresponds to bx^0 , with each preceding bx^{-1} , bx^{-2} , etc. in order from left to right (bxs are labeled for pt2 only). Class-defining mutations (known ultramutating *POLE* alleles) or dMMR identified by standard IHC screening are shown below the patient number. ⊗ signifies the first diagnosis of cancer; * signifies the first diagnosis of AH/EIN. For some cases, the initial biopsy showing AH/EIN was performed at an outside institution or was unavailable, indicated by [*]. The mutation calling threshold of 0.01 AF is shown with a dashed gray line. All mutations called in the bx^0 sample are shown for all samples except for pt16 (supplementary material, Figure S1), which harbored an ultramutating *POLE V411L* allele and $n = 70$ mutations, the majority of which are not listed. (B) Mutation counts based on MMR or *POLE* status. The number of patients is shown above each bar.

identified an inherited mutation in one patient (pt19), known to have Lynch syndrome (*PMS2*^{S123fs}). As expected for a heterozygous germline variant, AF was higher than other somatically acquired mutations and showed little fluctuation (Figure 2A). Thus, the HTS panel was able to detect classifying somatic mutations such as *POLE* and germline mutations that predispose to endometrial cancer [22].

Detection of definitive cancer drivers prior to initial precancer/cancer histopathologic diagnosis

In 11/21 cases (Figure 2A and supplementary material, Figure S1), ≥ 1 recurring hotspot [31] or definitive cancer driver mutation present in the cancer [32] was identified prior to the first diagnosis of AH/EIN. For example, in

pt3, the common *PTEN*^{R130Q} hotspot mutation was detectable in bx^{-1} and bx^{-2} (AH/EIN had not been diagnosed prior to diagnosis of carcinoma in bx^0). In pt13, canonical *PTEN*^{E43*} and *CTNNB1*^{D32Y} mutations were identified in bx^{-1} where histologic features did not lead to a definitive AH/EIN diagnosis. Notably, *CTNNB1*^{D32Y} was detectable in bx^{-3} a full 8 years prior to the first definitive precancer/cancer diagnosis (Figure 2A and supplementary material, Table S2). For 7/21 cases, a diagnosis of cancer or AH/EIN had already been made prior to the earliest bx (e.g. pt9, pt20) or in the bx immediately preceding the first cancer diagnosis (e.g. pt5, pt6, pt10) (Figure 2A and supplementary material, Figure S1). Thus, although these seven cases provide insights into AF progression, they are not useful to demonstrate the ability of HTS to identify cancer drivers prior to the first AH/EIN

or carcinoma diagnosis. Taken together, however, analysis of these cases demonstrated that definitive cancer drivers were present prior to the earliest diagnosis of a histologically identifiable precancer/cancer in most patients.

Corroboration of mutations and insights into early endometrial carcinogenesis by HTS-based serial immunohistochemistry

Oncogenic mutations (gain/loss of function) of several endometrial cancer genes – *PTEN*, *CTNNB1*, *TP53*, *ARID1A*, *MLH1/MSH2/MSH6/PMS2* – disrupt protein stability or subcellular localization, rendering mutant clones detectable by IHC [33–36]. For example, inactivating mutations of the four MMR factors destabilize the proteins, and loss of protein(s) is the basis for IHC as the screening method for Lynch syndrome/dMMR [37]. *PTEN* and *ARID1A* mutations can also lead to loss of protein [34,36,38]. In contrast, inactivating *TP53* mutations lead to protein stabilization/overexpression, and exon 3 *CTNNB1* mutations prevent β -catenin degradation, resulting in overexpression or abnormal subcellular localization [39]. Detection of mutant clones by IHC is useful because it can confirm mutations identified by HTS and give morphologic insights into potential precursor lesions.

Thus, IHC for *PTEN*, *CTNNB1*, *P53*, and *ARID1A* was performed based on HTS data (i.e. if an *ARID1A* mutation was identified, *ARID1A* IHC was performed on representative bxs for that patient). IHC for MMR factors was performed on the cancer for each patient as described above. For patients demonstrating loss of any factor(s), analysis was extended to preceding biopsies. Scoring was based on ‘clonal distinctiveness’ (CD), where distinct clones different from background endometrium can be definitively identified based on IHC expression patterns, as previously described [40]. The advantage of CD as a criterion is that it is inclusive of protein loss, overexpression, or abnormal localization. Of the 21 patients with multiple bxs, 20 were analyzed with at least one marker, i.e. 20/21 (with pt10 the exception) had a mutation in *PTEN*, *CTNNB1*, *TP53*, or *ARID1A*. Of these 20 patients, CD was detected for at least one factor in 18/20 cases. None of the 18 cases had CD based solely on *MLH1/PMS2*, commonly lost due to *MLH1* promoter hypermethylation [37]. Thus, targeted IHC based on HTS yielded independent confirmation of ≥ 1 biologically significant mutation in the great majority of cases (18/21, 86%).

Several examples are presented in detail (Figure 3A and supplementary material, Figure S2), as each provides unique insights. Pt1 and pt5 highlight the ability of IHC to identify β -catenin CD and confirm *CTNNB1* mutations. Pt1 had *S37F* mutation in all samples, while pt5 had an *S33Y* (both exon 3) mutation in all four samples (bx^0 – bx^{-3}) (supplementary material, Figure S1 and Figure 2A). CD was readily detectable (Figure 3A and supplementary material, Figure S2) by nuclear localization in malignant glands in all bxs (normal endometrium

exhibits only delicate membrane localization of β -catenin throughout the endometrial cycle [40] and nuclear β -catenin has $\geq 90\%$ specificity and sensitivity for *CTNNB1* mutations [41,42]). Nuclear β -catenin was also evident in areas of squamous differentiation (sq), which is associated with *CTNNB1* mutation (pt5, supplementary material, Figure S2) [43–45]. Pt11 and pt15 (supplementary material, Figure S1) provide additional examples of β -catenin (and *PTEN*) CD in patients with corresponding mutations and further highlight β -catenin overexpression or nuclear localization as reliable criteria for establishing β -catenin CD, relative to internal control glands. Pt15 also shows the ability of HTS to reliably detect driver mutations, as both *CTNNB1*^{S37F} (exon 3) and *PTEN*^{V45fs} mutations were detected with relatively high AFs in all preceding bxs, including the earliest (bx^{-5}) (supplementary material, Table S2).

In pt2 and pt3 (Figure 3 and supplementary material, Figure S1), *ARID1A* and *PTEN* mutations were IHC-confirmed in bx^0 . Protein loss (CD) for these markers can be readily detected in mutant clones because endometrial stroma serves as an internal positive control (Figure 3A). In pt3 bx^{-1} , 2.3 years prior to bx^0 , distinct *ARID1A*-negative glands were focally present, whereas the carcinoma at bx^0 was diffusely *ARID1A*-negative. *PTEN* CD was not confirmed prior to bx^0 , even though *PTEN*^{R130Q} was detectable at AF > 0.01. This could be sampling-related. Alternatively, perhaps at bx^{-1} only a heterozygous *PTEN* mutation was present. In any case, the findings are interesting because they show the ability of HTS-directed IHC to detect mutant clones earlier than routine histologic evaluation. bx^{-1} was diagnosed as endometrial polyp, while bx^{-2} was diagnosed as disordered proliferative endometrium, and the CD glands were not histologically distinct. These results also show that *ARID1A* mutations and clones can be early driver events in endometrial carcinogenesis, preceding by years the earliest histopathological diagnosis of AH/EIN.

In pt6 (dMMR for *MLH1/PMS2* by IHC), *MLH1* and *PMS2* CD were present in bx^0 and bx^{-2} (supplementary material, Figure S2). In all cases and bxs in this study, there was concordance with superimposable patterns of *MLH1* and *PMS2* CD; only *MLH1* is shown for brevity. *MLH1* CD was not detectable in bx^{-3} (–6.3 years from bx^0), arguing that *MLH1* hypermethylation was the initial event, with *ARID1A* mutation occurring subsequently. Consistent with this idea, the *ARID1A*^{P1175fs} mutation resulted from a deletion of a single C nucleotide (*c.3524delC*), likely due to strand slippage induced by dMMR, as this specific *ARID1A* mutation has been previously documented in dMMR-colorectal adenocarcinomas [46].

In pt18, four mutations were identified in bx^0 , diagnosed as atypical hyperplasia bordering on adenocarcinoma, with adenocarcinoma diagnosed in the hysterectomy. None of these mutations were detectable in bx^{-1} , taken only 197 days prior to bx^0 (supplementary material, Table S2 and Figure S1). One of the four mutations was a *PTEN*^{A120fs} frameshift mutation, and *PTEN*

CD characterized the entire adenocarcinoma (supplementary material, Figure S2). Concordant with the HTS results, however, no PTEN CD was detectable in bx^{-1} , likely due to insufficient sampling in bx^{-1} , noted to be scant in the original diagnosis.

Finally, pt21 is a striking example of HTS-directed IHC for visualization of mutant clones providing insights into early endometrial carcinogenesis. bx^0 showed *ARID1A* and *PTEN* mutations, and bx^0 IHC showed CD for both *ARID1A* and *PTEN* with superimposable patterns in serial sections (supplementary material, Figure S2). However, bx^{-1} , taken at -5.4 years, showed distinct *PTEN* loss in three adjacent glands that were morphologically unremarkable but can be

surmised to be clonal due to their close physical proximity. These same glands retained *ARID1A*, arguing that *PTEN* was the initial mutation, followed by *ARID1A*. These findings help to rationalize the early reports by Mutter of *PTEN*-null glands in morphologically normal endometria, and argue that they are likely to represent significant but very early precursor lesions [33,47]. The findings also demonstrate the ability of HTS and HTS-based CD to identify relevant mutations prior to histologically identifiable AH/EIN (bx^{-1} was diagnosed as minimally disordered PE, no hyperplasia or malignancy; supplementary material, Table S2). HTS-based CD analysis of the remaining patients is shown in supplementary material, Figure S2. We then tabulated the

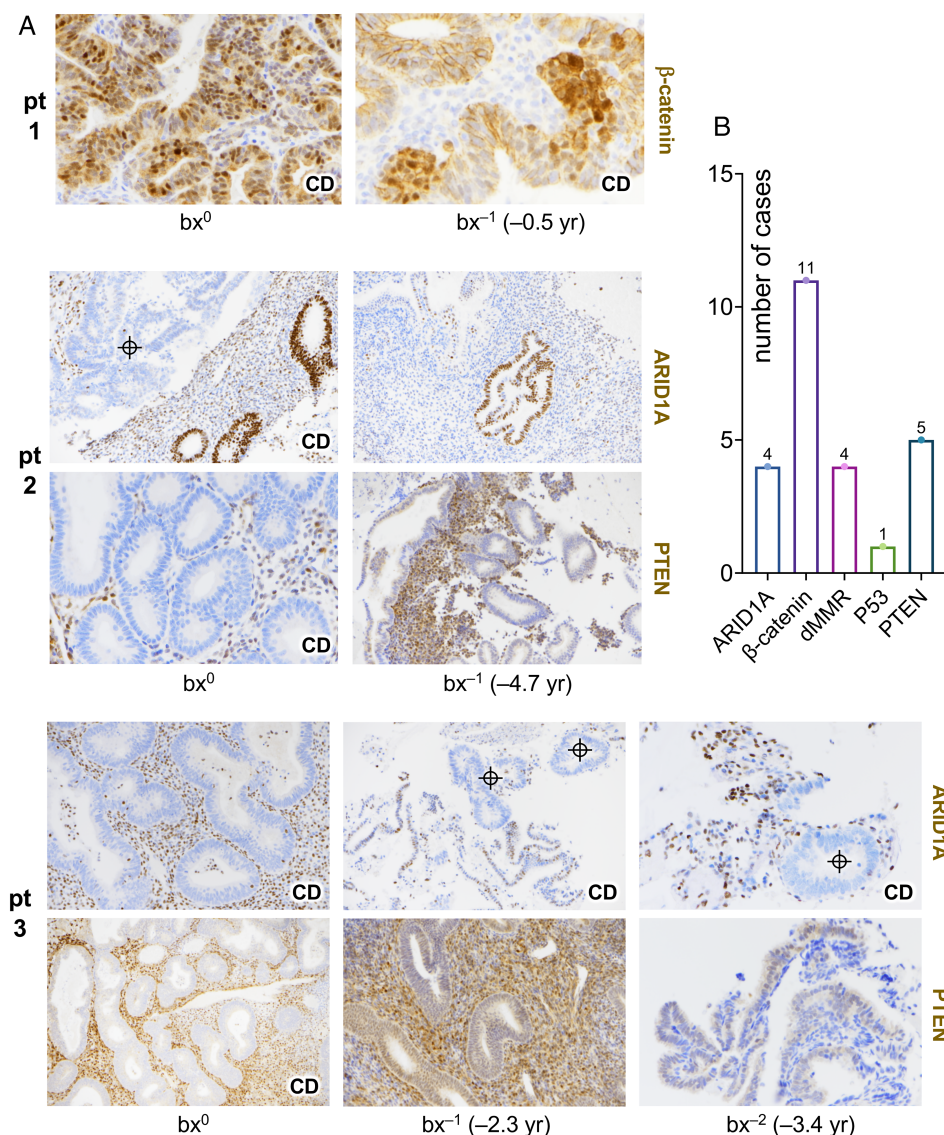


Figure 3. Mutations identified by HTS can be confirmed by immunohistochemistry: *ARID1A*, *CTNNB1* (β -catenin), *TP53* (p53), and *PTEN*. (A) Patients 1–3 are shown; see supplementary material, Figure S2 for all other cases. IHC for a relevant MMR factor (e.g. MLH1) is also shown for cases where dMMR was identified during four-factor MMR screening for further insights into precursor progression. IHC was performed on up to four samples per patient including bx^0 , spanning significant histopathological transitions including (when available) a bx preceding the first diagnosis of cancer/EIN. CD = clonal distinctiveness, signifying direct confirmation of an underlying cancer-driving mutation by IHC \oplus highlights CD glands (where only focal glands showed CD). nl = representative normal glands used to assess CD; sq = squamous differentiation. (B) Number of patients for which marker was among the first to detect CD. The total number is >21 (number of patients in this study) because for some patients there was >1 marker detecting CD in the first biopsy. dMMR, CD by MLH1, MSH2, MSH6, or PMS2.

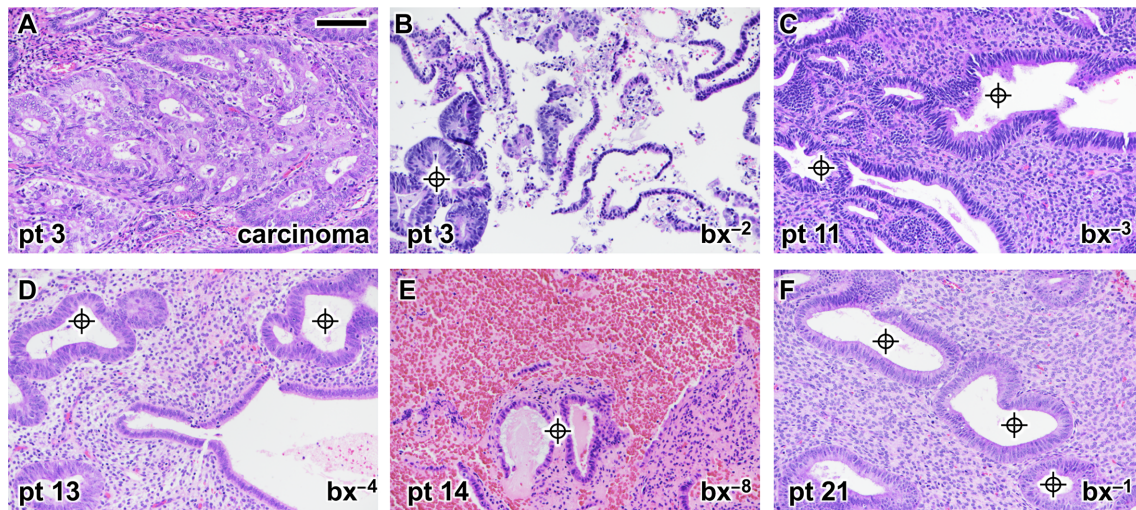


Figure 4. Histology of early biopsies with IHC-confirmed HTS mutations but lacking diagnostic features of EIN/AH. All panels are at the same magnification (bar = 100 μ m); images were taken with 20 \times objective. CD glands (per IHC step sections) are shown with \oplus symbol. (A) Representative endometrial cancer (pt 3, FIGO grade 1) for general comparison, showing crowded glands with no intervening stroma, diagnostic of adenocarcinoma. (B) Pt3. Most fragments are superficial and atrophic; rare \oplus glands in retrospect have distinct histology but such glands are too few and changes too minimal for diagnosis of AH/EIN. (C) Pt11. \oplus glands have some crowding and architectural abnormalities that fall short of AH/EIN criteria. (D) Pt13. \oplus glands have larger cells but minimal architectural irregularities. (E) Pt14. Specimen very scant and highly-fragmented limiting diagnosis. (F) Pt21. \oplus glands show mild architectural abnormalities and are not histologically very distinct from neighboring glands.

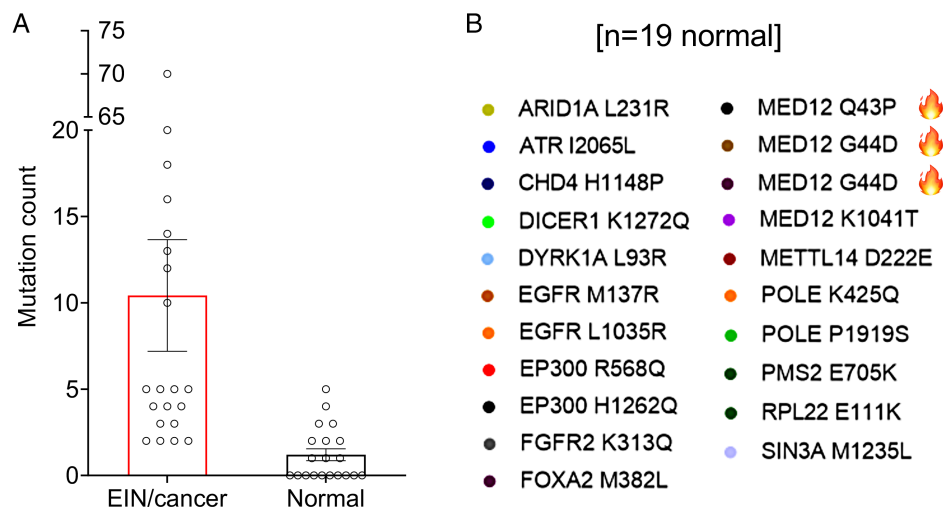


Figure 5. Analysis of non-malignant (normal) endometrial samples. Samples were subjected to similar workflow and identical mutation calling protocols. (A) Overall mutation counts in bx⁰ samples; scatter plots showing mean \pm SEM. (B) Aggregate list of mutations from the $n = 19$ normal samples analyzed; known tumor hotspot or definitive driver mutations are indicated with a 🔥 symbol. All three are *MED12* mutations (Q43P, G44D) seen in leiomyomata.

markers that detected the initial CD. Notably, all four markers and dMMR were sometimes the sentinel CD events, with β -catenin being the most common and P53 being the least common (Figure 3B). This shows that β -catenin mutations are often early events, in agreement with recent studies that β -catenin is a useful biomarker for the detection/confirmation of AH/EIN [40,42].

Re-evaluation of early biopsies with IHC-confirmed HTS mutations revealed subtle gland architectural abnormalities in most cases. In retrospect, some CD glands were histologically distinctive (i.e. different from background normal, non-CD glands) in terms of architecture and/or cell size (Figure 4). However, in such

cases, the changes were subtle and did not meet histologic criteria for AH/EIN. In other biopsies that were scant or highly fragmented (limiting reliable histologic evaluation), CD glands were not readily distinguishable (Figure 4). Overall, these results highlight the ability of HTS to detect mutations in biopsies that are histologically sub-diagnostic for endometrial precancer for a variety of reasons.

HTS analysis of normal endometrium

We then analyzed 19 cases of histologically unremarkable endometrium from women of similar age ranges.

Hysterectomies for benign indications were used to isolate endometrial and matched germline DNA. Nine of the 19 cases showed no mutations, with the remainder showing some called variants (Figure 5). However, none were categorical hotspot or predicted oncogenic mutations [32] other than *MED12*^{G44D} and *MED12*^{Q43P}, common driver mutations in benign leiomyomata, which do not occur in endometrial adenocarcinomas. The presence of *MED12* mutations correlated with the presence of submucosal leiomyomata. As expected, the number of mutations was higher (by several-fold) in EIN/AH cancer cases than in normal endometria [48].

Discussion

This study helps to establish endometrium as a model system for precancer evolution, particularly of very early precancers. Starting with women who underwent hysterectomy for a diagnosis of endometrial cancer, we identified a subset of patients with preceding serial biopsies over several years. Pathology laboratories store FFPE blocks for ≥ 10 years, and such archives are invaluable resources for diverse investigations of cancer progression. This study also showed the feasibility of identifying driver mutations in an endometrial cancer, and then tracking variant allele frequencies retrospectively in preceding biopsies.

A notable finding was the detection of mutations (present in the invasive cancer) before the diagnosis of AH/EIN. The diagnosis of AH/EIN is based on histopathologic criteria and is subjective [33,40,49,50]. The reliance in this study on original histopathologic diagnoses eliminated the possibility of retrospective observer bias. However, the prerequisite for a long preceding history of endometrial samples might introduce unforeseen bias. The unrecognized early precursor lesions harboring canonical endometrial cancer driver mutations were classified as simple hyperplasias (SH); disordered proliferative endometrium (DPE); exhibited some other abnormal histologic features not diagnostic of DPE, SH, or AH/EIN; or were noted to be scant. This strongly suggests that some proportion of biopsies routinely diagnosed as 'disordered proliferative endometrium' or 'non-atypical hyperplasia' in fact harbor molecularly-significant precursor lesions.

There are currently two distinct classification schemes for endometrial precancers (AH versus EIN systems), grounded on divergent histologic criteria. The biopsies in this study were diagnosed with the AH system. While the advantages of either system are debated [49,51], we found that premalignant lesions with definitive mutations are not always histologically recognizable as either AH or EIN and thus would not be identified by either system [15,52,53]. In any case, the lack of standardization of histopathologic criteria and their application remain significant issues, and the poor inter-observer agreement for either system poses continuing challenges for patient management and cancer prevention [51].

Although this study was not designed to establish clinical utility, our findings suggest that HTS evaluation of endometrial precancers could be diagnostically useful in the interpretation of endometrial biopsies with indeterminate histologic features (such as crowded glands, DPE, or SH). HTS analysis might (1) help to establish a lesion as a *bona fide* precursor signifying an increased cancer risk; (2) identify class-defining mutations earlier in clinical progression, which could further guide management; and (3) diagnose hereditary cancer syndromes earlier, which would trigger earlier surveillance and enhance clinical management. Although specific criteria would need to be developed, including the utility of or need for HTS-based IHC, our data suggest that the detection of a single known hotspot or definitive cancer driver (e.g. *KRAS*, *PTEN* loss of function, *CTNNB1* exon 3 mutation, etc.) in the context of suspicious histopathologic features could have sufficient sensitivity and specificity. Missense variants of unknown functional consequence would likely have no diagnostic value, as they appear in normal aging endometria. However, even with technical refinements, sensitivity will be limited by sampling issues or technical factors.

Recent studies of aging normal endometrium found that most microdissected single endometrial glands harbor mutations, consistent with the inexorable age-related accumulation of mutations due to normal DNA replication errors, estimated at ~ 1 per cell division. However, the vast majority of such age-related mutations are limited to one gland, and thus might be undetectable in analyses of entire (non-microdissected) samplings because the AF would be well below call thresholds [48,54,55]. We have not further investigated if the mutations that we identified in normal endometria were age-related mutations or sequencing errors; however, two mutations were in *CHD4* and *FOXA2* (Figure 5B), among the small number of genes that undergo age-related mutation/positive selection in endometrium [48]. Other recent studies have confirmed the presence of age-associated mutations in endometria and detected such mutations in non-microdissected tissues as well as by HTS-directed CD of *PTEN* and *ARID1A* [54]. The incidence of mutant clones in normal aging endometria or endometriosis and their relationship to cancer remains an intriguing subject under investigation [55–57].

In summary, this research, which serially analyzed endometrial biopsies from individual women, demonstrates that mutations in endometrial cancers/precancers are present in earlier endometria without definitive histologic features of AH/EIN, highlighting the existence of 'pre-AH/EIN' lesions not readily diagnosable histologically but that harbor clonal genetic aberrations detectable by HTS and verifiable by IHC. We found that premalignant lesions with definitive mutations that eventually gave rise to invasive cancers were not always histologically recognized at the time of biopsy [15,52,53]. Our study also showed that HTS evaluation of endometrial precancers can provide unique insights into precancer initiation and progression.

Acknowledgements

Funding to DHC was provided by the NIH/NCI (R01CA19691) and the Cancer Prevention Research Institute of Texas (RP190207). We gratefully acknowledge the support of the UTSW Tissue Management Facility through the UTSW Simmons Comprehensive Cancer Center Support Grant P30 CA142543 and the UT Southwestern Bioinformatics Core Facility, funded by the Cancer Prevention and Research Institute of Texas (RP150596), for bioinformatics pipeline support.

Author contributions statement

DHC conceived the study and secured funding. MA, MZ, BC, SS, HL, IC, JL, DM and JG designed experiments. HZ performed bioinformatic analyses. DHC, HC, WZ and EL designed immunohistochemical methods, analyzed photomicrographs, and provided histopathological interpretations. All the authors interpreted data. DHC and MA wrote the first draft of the manuscript. All the authors revised the manuscript critically for important intellectual content and approved the submitted final version.

Data availability statement

Sequence data have been deposited in GenBank (SUB7696261).

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020; **70**: 7–30.
- Kurman RJ, Carcangiu ML, Herrington CS, *et al.*, (eds). *WHO Classification of Tumours of Female Reproductive Organs* (4th edn). International Agency for Research on Cancer: Lyon, 2014.
- Kurman RJ, Kaminski PF, Norris HJ. The behavior of endometrial hyperplasia. A long-term study of “untreated” hyperplasia in 170 patients. *Cancer* 1985; **56**: 403–412.
- Huvila J, Pors J, Thompson EF, *et al.* Endometrial carcinoma: molecular subtypes, precursors and the role of pathology in early diagnosis. *J Pathol* 2020. <https://doi.org/10.1002/path.5608> [Epub ahead of print].
- Lacey JV Jr, Ioffe OB, Ronnett BM, *et al.* Endometrial carcinoma risk among women diagnosed with endometrial hyperplasia: the 34-year experience in a large health plan. *Br J Cancer* 2008; **98**: 45–53.
- Mutter GL, Zaino RJ, Baak JP, *et al.* Benign endometrial hyperplasia sequence and endometrial intraepithelial neoplasia. *Int J Gynecol Pathol* 2007; **26**: 103–114.
- Hong B, Le Gallo M, Bell DW. The mutational landscape of endometrial cancer. *Curr Opin Genet Dev* 2015; **30**: 25–31.
- Cherniack AD, Shen H, Walter V, *et al.* Integrated molecular characterization of uterine carcinosarcoma. *Cancer Cell* 2017; **31**: 411–423.
- The Cancer Genome Atlas Research Network, Kandoth C, Schultz N, *et al.* Integrated genomic characterization of endometrial carcinoma. *Nature* 2013; **497**: 67–73.
- Zehir A, Benayed R, Shah RH, *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* 2017; **23**: 703–713.
- Cuevas IC, Sahoo SS, Kumar A, *et al.* Fbxw7 is a driver of uterine carcinosarcoma by promoting epithelial–mesenchymal transition. *Proc Natl Acad Sci U S A* 2019; **116**: 25880–25890.
- Le Gallo M, O’Hara AJ, Rudd ML, *et al.* Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat Genet* 2012; **44**: 1310–1315.
- Wiegand KC, Lee AF, Al-Agha OM, *et al.* Loss of BAF250a (*ARID1A*) is frequent in high-grade endometrial carcinomas. *J Pathol* 2011; **224**: 328–333.
- Wang Y, Hoang L, Ji JX, *et al.* SWI/SNF complex mutations in gynecologic cancers: molecular mechanisms and models. *Annu Rev Pathol* 2020; **15**: 467–492.
- Srivastava S, Ghosh S, Kagan J, *et al.* The making of a PreCancer Atlas: promises, challenges, and opportunities. *Trends Cancer* 2018; **4**: 523–536.
- Bond JH. Interference with the adenoma–carcinoma sequence. *Eur J Cancer* 1995; **31A**: 1115–1117.
- Nasiell K, Nasiell M, Vačlavinková V. Behavior of moderate cervical dysplasia during long-term follow-up. *Obstet Gynecol* 1983; **61**: 609–614.
- Stephenson RD, Denehy TR. Rapid spontaneous regression of acute-onset vulvar intraepithelial neoplasia 3 in young women: a case series. *J Low Genit Tract Dis* 2012; **16**: 56–58.
- Li L, Yue P, Song Q, *et al.* Genome-wide mutation analysis in precancerous lesions of endometrial carcinoma. *J Pathol* 2021; **253**: 119–128.
- WHO Classification of Tumours Editorial Board. *Female Genital Tumours* (5th edn). International Agency for Research on Cancer: Lyon, 2020; 250–251.
- Nakada Y, Stewart TG, Pena CG, *et al.* The LKB1 tumor suppressor as a biomarker in mouse and human tissues. *PLoS One* 2013; **8**: e73449.
- Lucas E, Chen H, Molberg K, *et al.* Mismatch repair protein expression in endometrioid intraepithelial neoplasia/atypical hyperplasia: should we screen for Lynch syndrome in precancerous lesions? *Int J Gynecol Pathol* 2019; **38**: 533–542.
- Gao J, Aksoy BA, Dogrusoz U, *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013; **6**: p11.
- Bell DW, Ellenson LH. Molecular genetics of endometrial carcinoma. *Annu Rev Pathol* 2019; **14**: 339–367.
- Berger AC, Korkut A, Kanchi RS, *et al.* A comprehensive pan-cancer molecular study of gynecologic and breast cancers. *Cancer Cell* 2018; **33**: 690–705.e9.
- The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011; **474**: 609–615.
- Wheeler DT, Bristow RE, Kurman RJ. Histologic alterations in endometrial hyperplasia and well-differentiated carcinoma treated with progestins. *Am J Surg Pathol* 2007; **31**: 988–998.
- Zaino RJ, Brady WE, Todd W, *et al.* Histologic effects of medroxyprogesterone acetate on endometrioid endometrial adenocarcinoma: a Gynecologic Oncology Group study. *Int J Gynecol Pathol* 2014; **33**: 543–553.
- Shlien A, Campbell BB, de Borja R, *et al.* Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermethylated cancers. *Nat Genet* 2015; **47**: 257–262.
- Kalia SS, Adelman K, Bale SJ, *et al.* Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med* 2017; **19**: 249–255.
- Chang MT, Asthana S, Gao SP, *et al.* Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nat Biotechnol* 2016; **34**: 155–163.
- Chakravarty D, Gao J, Phillips S, *et al.* OncoKB: a precision oncology knowledge base. *JCO Precis Oncol* 2017; **2017**: PO.17.00011.

33. Nucci MR, Castrillon DH, Bai H, et al. Biomarkers in diagnostic obstetric and gynecologic pathology: a review. *Adv Anat Pathol* 2003; **10**: 55–68.
34. Monte NM, Webster KA, Neuberger D, et al. Joint loss of PAX2 and PTEN expression in endometrial precancers and cancer. *Cancer Res* 2010; **70**: 6225–6232.
35. Singh N, Piskorz AM, Bosse T, et al. p53 immunohistochemistry is an accurate surrogate for TP53 mutational analysis in endometrial carcinoma biopsies. *J Pathol* 2020; **250**: 336–345.
36. Hecht JL, Pinkus JL, Pinkus GS. Enhanced detection of atypical hyperplasia in endometrial polyps by PTEN expression. *Appl Immunohistochem Mol Morphol* 2004; **12**: 36–39.
37. Stelloo E, Jansen AML, Osse EM, et al. Practical guidance for mismatch repair-deficiency testing in endometrial cancer. *Ann Oncol* 2017; **28**: 96–102.
38. Mao TL, Ardighieri L, Ayhan A, et al. Loss of ARID1A expression correlates with stages of tumor progression in uterine endometrioid carcinoma. *Am J Surg Pathol* 2013; **37**: 1342–1348.
39. Gao C, Wang Y, Broadus R, et al. Exon 3 mutations of CTNNB1 drive tumorigenesis: a review. *Oncotarget* 2018; **9**: 5492–5508.
40. Strickland AL, Rivera G, Lucas E, et al. PI3K pathway effectors pAKT and FOXO1 as novel markers of endometrioid intraepithelial neoplasia. *Int J Gynecol Pathol* 2019; **38**: 503–513.
41. Kim G, Kurnit KC, Djordjevic B, et al. Nuclear beta-catenin localization and mutation of the CTNNB1 gene: a context-dependent association. *Mod Pathol* 2018; **31**: 1553–1559.
42. Costigan DC, Dong F, Nucci MR, et al. Clinicopathologic and immunohistochemical correlates of CTNNB1 mutated endometrial endometrioid carcinoma. *Int J Gynecol Pathol* 2020; **39**: 119–127.
43. Saegusa M, Okayasu I. Frequent nuclear beta-catenin accumulation and associated mutations in endometrioid-type endometrial and ovarian carcinomas with squamous differentiation. *J Pathol* 2001; **194**: 59–67.
44. Brachtel EF, Sánchez-Estevéz C, Moreno-Bueno G, et al. Distinct molecular alterations in complex endometrial hyperplasia (CEH) with and without immature squamous metaplasia (squamous morules). *Am J Surg Pathol* 2005; **29**: 1322–1329.
45. Lin MC, Lomo L, Baak JP, et al. Squamous morules are functionally inert elements of premalignant endometrial neoplasia. *Mod Pathol* 2009; **22**: 167–174.
46. Cajuso T, Hänninen UA, Kondelin J, et al. Exome sequencing reveals frequent inactivating mutations in ARID1A, ARID1B, ARID2 and ARID4A in microsatellite unstable colorectal cancer. *Int J Cancer* 2014; **135**: 611–623.
47. Mutter GL. Histopathology of genetically defined endometrial precancers. *Int J Gynecol Pathol* 2000; **19**: 301–309.
48. Moore L, Leongamornlert D, Coorens THH, et al. The mutational landscape of normal human endometrial epithelium. *Nature* 2020; **580**: 640–646.
49. Ordi J, Bergeron C, Hardisson D, et al. Reproducibility of current classifications of endometrial endometrioid glandular proliferations: further evidence supporting a simplified classification. *Histopathology* 2014; **64**: 284–292.
50. Skov BG, Broholm H, Engel U, et al. Comparison of the reproducibility of the WHO classifications of 1975 and 1994 of endometrial hyperplasia. *Int J Gynecol Pathol* 1997; **16**: 33–37.
51. Sanderson PA, Critchley HO, Williams AR, et al. New concepts for an old problem: the diagnosis of endometrial hyperplasia. *Hum Reprod Update* 2017; **23**: 232–254.
52. Jarboe EA, Mutter GL. Endometrial intraepithelial neoplasia. *Semin Diagn Pathol* 2010; **27**: 215–225.
53. Wong S, Hui P, Buza N. Frequent loss of mutation-specific mismatch repair protein expression in nonneoplastic endometrium of Lynch syndrome patients. *Mod Pathol* 2020; **33**: 1172–1181.
54. Lac V, Nazeran TM, Tessier-Cloutier B, et al. Oncogenic mutations in histologically normal endometrium: the new normal? *J Pathol* 2019; **249**: 173–181.
55. Suda K, Nakaoka H, Yoshihara K, et al. Clonal expansion and diversification of cancer-associated mutations in endometriosis and normal endometrium. *Cell Rep* 2018; **24**: 1777–1789.
56. Anglesio MS, Papadopoulos N, Ayhan A, et al. Cancer-associated mutations in endometriosis without cancer. *N Engl J Med* 2017; **376**: 1835–1848.
57. Noe M, Ayhan A, Wang TL, et al. Independent development of endometrial epithelium and stroma within the same endometriosis. *J Pathol* 2018; **245**: 265–269.
58. Schneider VA, Graves-Lindsay T, Howe K, et al. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res* 2017; **27**: 849–864.
59. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009; **25**: 1754–1760.
60. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
61. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 2012; **6**: 80–92.
62. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015; **4**: 7.
63. Cunha KS, Oliveira NS, Fausto AK, et al. Hybridization capture-based next-generation sequencing to evaluate coding sequence and deep intronic mutations in the NF1 gene. *Genes (Basel)* 2016; **7**: 133.
64. The Cancer Genome Atlas Research Network, Albert Einstein College of Medicine, Analytical Biological Services, et al. Integrated genomic and molecular characterization of cervical cancer. *Nature* 2017; **543**: 378–384.
65. Heidenreich B, Kumar R. TERT promoter mutations in telomere biology. *Mutat Res* 2017; **771**: 15–31.

References 58–65 are cited only in the supplementary material.

SUPPLEMENTARY MATERIAL ONLINE

Supplementary materials and methods

Supplementary figure legends

Figure S1. Variant allele frequencies in bx⁰ and preceding biopsies (all cases not shown in supplementary material, Figure S2A)

Figure S2. HTS-driven validation of mutations that can be confirmed by immunohistochemistry: ARID1A, CTNNB1 (β-catenin), TP53 (p53), and PTEN

Table S1. List of genes on custom capture panel for gynecologic malignancies including endometrial cancer

Table S2. Detailed timeline with summary of clinical diagnoses

Table S3. Confirmation of HTS data and detection of mutant clones by IHC

Table S4. Description of controls