The genetic landscape of endometrial clear cell carcinomas

Deborah F DeLair*, Kathleen A Burke, Pier Selenica, Raymond S Lim, Sasinya N Scott, Sumit Middha, Abhinita S Mohanty, Donavan T Cheng, Michael F Berger, Robert A Soslow* and Britta Weigelt*

Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

*Correspondence to: DF DeLair, Department of Pathology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. E-mail: delaird@mskcc.org
Or RA Soslow, Department of Pathology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. E-mail: soslowr@mskcc.org
Or B Weigelt, Department of Pathology, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. E-mail: weigeltb@mskcc.org

Abstract

Clear cell carcinoma of the endometrium is a rare type of endometrial cancer that is generally associated with an aggressive clinical behaviour. Here, we sought to define the repertoire of somatic genetic alterations in endometrial clear cell carcinomas (ECCs), and whether ECCs could be classified into the molecular subtypes described for endometrial endometrioid and serous carcinomas. We performed a rigorous histopathological review, immunohistochemical analysis and massively parallel sequencing targeting 300 cancer-related genes of 32 pure ECCs. Eleven (34%), seven (22%) and six (19%) ECCs showed abnormal expression patterns for p53, ARID1A, and at least one DNA mismatch repair (MMR) protein, respectively. Targeted sequencing data were obtained from 30 of the 32 ECCs included in this study, and these revealed that two ECCs (7%) were ultramutated and harboured mutations affecting the exonuclease domain of POLE. In POLE wild-type ECCs, TP53 (46%), PIK3CA (36%), PPP2R1A (36%), FBXW7 (25%), ARID1A (21%), PIK3R1 (18%) and SPOP (18%) were the genes most commonly affected by mutations; 18% and 11% harboured CCNE1 and ERBB2 amplifications, respectively, and 11% showed DAXX homozygous deletions. ECCs less frequently harboured mutations affecting CTNNB1 and PTEN but more frequently harboured PPP2R1A and TP53 mutations than non-POLE endometrioid carcinomas from The Cancer Genome Atlas (TCGA). Compared to endometrial serous carcinomas (TCGA), ECCs less frequently harboured TP53 mutations. When a surrogate model for the molecular-based TCGA classification was used, all molecular subtypes previously identified in endometrial endometrioid and serous carcinomas were present in the ECCs studied, including POLE, MMR-deficient, copy-number high (serous-like)/p53 abnormal, and copy-number low (endometrioid)/p53 wild-type, which were significantly associated with disease-free survival in univariate analysis. These findings demonstrate that ECCs constitute a histologically and genetically heterogeneous group of tumours with varying outcomes. Furthermore, our data suggest that the classification of ECCs as being generally ‘high-grade’ or ‘type II’ tumours may not be warranted.

Keywords: endometrial clear cell carcinoma; massively parallel sequencing; somatic mutations; copy number alterations; molecular classification

Received 26 April 2017; Revised 20 May 2017; Accepted 5 July 2017

No conflicts of interest were declared.

Introduction

Endometrial clear cell carcinoma (ECC) is a rare type of endometrial cancer that accounts for <3% of all endometrial cancers [1–3]. ECCs are generally morphologically similar to their ovarian counterparts, but it should be noted that the diagnosis of these lesions can be challenging, and that significant interobserver diagnostic variability has been reported [3–6].

Endometrial carcinoma has been traditionally classified into two groups on the basis of clinical, endocrine and epidemiological observations, the so-called type I and type II cancers [7,8]. Type I cancers are typically endometrioid carcinomas that tend to show a good prognosis, whereas type II cancers are generally associated with a poor prognosis and include serous and clear cell histologies [7–10]. There is burgeoning evidence to demonstrate, however, that endometrial cancer is a biologically, clinically and genetically heterogeneous disease, and that this dualistic classification may not reflect the actual heterogeneity observed [8]. A study by The Cancer Genome Atlas (TCGA) combining somatic mutations, copy-number alterations and microsatellite instability (MSI) data classified endometrial endometrioid and serous carcinomas into four molecular subtypes [11]: (1) the POLE (ultramutated) tumours, which are characterized by extremely high mutation rates and mutations in the exonuclease domain...
of POLE; (2) the MSI (hypervariable) tumours, which show very high mutation rates and few copy-number alterations; (3) the copy-number low (endometrioid) tumours, which are microsatellite stable, and harbour lower mutation frequencies and recurrent CTNNB1 mutations; and (4) the copy-number high (serous-like) group, which includes all serous carcinomas and a subset of the grade 3 endometrioid carcinomas, and are characterized by high levels of copy-number alterations, low mutation frequencies, and recurrent TP53, PPP2R1A and FBXW7 somatic mutations [11].

At variance with endometrial endometrioid and serous carcinomas, there is a paucity of data on the genomic landscape of ECCs, partly because of the rarity of this tumour type. Candidate gene analyses of small series of ECCs have shown mutations affecting PIK3CA, ARID1A, PPP2R1A, TP53, PIK3R1, PTEN, and KRAS [4, 12, 13], and molecular similarities to both serous and endometrioid endometrial cancers have been found [14], but little is known about copy-number changes and molecular subtypes of these lesions, and whether the genetic alterations correlate with outcome.

To address this gap in our understanding of the genetics of ECCs, we subjected a series of centrally reviewed ECCs to immunohistochemical and massively parallel sequencing analysis to investigate: (1) whether ECCs harbour mutations affecting 300 key cancer genes, (2) whether ECCs show a repertoire of somatic mutations that is distinct from that of endometrial endometrioid and serous carcinomas, and (3) whether ECCs could be classified into the molecular subtypes described for endometrial endometrioid and serous carcinomas.

Materials and methods

Case selection

We selected from the files of the Department of Pathology at Memorial Sloan Kettering Cancer Center (MSKCC) all ECCs (n = 45) diagnosed between 1996 and 2013 that had both slides and blocks available. Samples were anonymized prior to analysis. This study was approved by the Institutional Review Board of the MSKCC, and patient consent was obtained where appropriate. Representative haematoyxin and eosin (H&E)-stained sections of each case diagnosed as ECC were independently reviewed by two specialized gynaecological pathologists with expertise in clear cell morphology (D.F.D. and R.A.S.) [4, 5]. For inclusion in the study, the tumours had to show typical clear cell morphology, as defined by the World Health Organization [2] and Fadare et al [3]. More specifically, carcinomas with abundant nuclear stratification, diffuse severe pleomorphism and abundant columnar nuclear shapes were excluded. Immunohistochemistry (IHC) was not used in the classification of these tumours, following Fadare et al [3].

Specific morphological features were evaluated, including the presence of classic architectural patterns, lymphocytic infiltration, mitotic index [number of mitoses per 10 high-power fields (HPFs)] and nuclear grade by use of a three-tiered scale based on nuclear pleomorphism [3] (Supplementary materials and methods). Following this review, we included 32 ECCs. Of the 45 cases reviewed, 13 (29%) were excluded; these were cases for which both reviewers diagnosed a tumour other than clear cell carcinoma, cases lacking diagnostic consensus between both pathologists, and/or mixed epithelial carcinoma with a clear cell component. Clinical information, including age, stage, location of metastases at presentation, and follow-up, was retrieved from the medical records.

Immunohistochemistry

IHC for p53, ARID1A/BAF250a and the DNA mismatch repair (MMR) proteins MSH2, MSH6, MLH1 and PMS2 was performed on all cases, as described previously [15–17]; see supplementary material, Supplementary materials and methods.

DNA extraction

Formalin-fixed paraffin-embedded tumour and normal sections were reviewed by a pathologist (D.F.D.). Five-micrometre sections from tumour samples were manually macrodissected to ensure that there were >20% neoplastic cells. Normal tissue sections, usually from a benign lymph node, were confirmed to be devoid of any neoplastic cells. Genomic DNA from tumour-matched and patient-matched normal samples was extracted with the DNeasy Blood & Tissue kit (Qiagen, Germantown, MD, USA).

Targeted capture massively parallel sequencing

Tumour and normal DNA samples were subjected to targeted massively parallel sequencing with the Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT) assay, targeting all exons of 300 key cancer genes (supplementary material, Table S1), as described previously [18, 19]. Sequence reads were aligned to the human reference genome GRCh37 using the Burrows–Wheeler Aligner (BWA; v0.7.10) [20], and local realignment, duplicate removal and base quality recalibration were performed with the Genome Analysis Toolkit (GATK; v3.1.1) [21]. Variant calling and copy-number analysis were performed as described previously [19, 22–27] (Supplementary materials and methods). Mutational hotspots were assigned according to Chang et al [28]. Cancer cell fractions of all mutations were inferred using ABSOLUTE (v1.0.6) [29], as described previously [19, 22], and the potential functional effect of each mutation was investigated as described previously [19, 29–37]; see also Supplementary materials and methods.
Comparisons of ECCs with endometrial carcinomas from the TCGA dataset

For comparisons of mutational frequencies of ECCs with those of endometrial endometrioid and serous carcinomas, the clinicopathological data and whole exome sequencing-derived mutational data from the TCGA data portal (https://tcga-data.nci.nih.gov/docs/publications/ucec_2013/; files ‘Key Clinical Data’, ‘UCEC Somatic Mutations’, and ‘Cumulative Data Freeze List’) were retrieved as described previously [19]. We restricted the comparison to the 300 genes targeted by our sequencing panel. All comparisons were performed using Fisher’s exact tests, corrected for multiple comparisons using the Benjamini–Hochberg method.

Molecular classification

To classify the ECCs into the molecular subtypes described for endometrial endometrioid and serous carcinomas by TCGA, we employed a surrogate model described by Talhouk et al [38] (Supplementary materials and methods). As a second approach, hierarchical clustering was performed by the use of Ward’s algorithm and Euclidean distance, with the ECCs and all endometrial endometrioid and serous carcinomas in the TCGA dataset [11] filtered for the 300 genes targeted by our sequencing panel. The stability of the hierarchical cluster analysis was assessed with pvclust [39].

Statistical analysis

The association between molecular subtype and disease-free survival was analysed, and survival curves were calculated by use of the Kaplan–Meier method with the log-rank test. Associations between specific clinicopathological features and molecular subtypes were tested with Fisher’s exact and t-tests. P values of <0.05 were considered to be statistically significant.

Results

ECCs are phenotypically heterogeneous

After central review of 45 cases initially classified as ECCs, a final diagnosis of pure ECC was rendered in 32 cases, which were included in this study (see Materials and methods). The median age of the patients was 65 years (range 33–83 years). At presentation, 50% (16/32) of patients were of FIGO stage I, 6% (2/32) were of FIGO stage II, and 44% (14/32) were of FIGO stage IV (Table 1). The median follow-up was 29 months (range 5–105 months), and, at the end of the follow-up period, 47% (15/32) of patients had died of disease, 41% (13/32) showed no evidence of disease, 9% (3/32) were alive with disease, and 3% (1/32) had died of another cause.

The ECCs included in this study showed varying combinations of the typical morphological patterns previously described in Mullerian clear cell carcinomas (Table 1 and Figure 1), with papillary and tubulocystic combinations (19/32, 60%) and papillary, tubulocystic and solid combinations (7/32, 22%) being the most common. All ECCs predominantly showed grade 2 nuclei, and, in 50% of the cases, focal areas composed of cells with grade 3 nuclei were present (Figure 1). The median mitotic index was five mitotic figures per 10 HPFs (range 1–18).

Immunohistochemical analysis revealed that six ECCs (19%) showed abnormal expression patterns for DNA MMR proteins (Table 1 and Figure 1), with MSH6 being the most commonly altered. Of these, CC11 had a personal history of colorectal cancer and an immediate family member with a Lynch syndrome-related tumour. CC10 was not subjected to germline testing, and the remaining four patients were found not to harbour any pathogenic mutations affecting the canonical DNA MMR genes. In addition, CC05 showed MLH1 promoter methylation, whereas CC20 did not. Of the 32 ECCs analysed, 11 (34%) showed aberrant p53 expression, six showed complete loss of ARID1A expression, and one showed partial geographical loss of ARID1A. All ECCs with loss of ARID1A expression showed wild-type p53 expression patterns (Table 1). Taken together, the ECCs analysed here were found to be heterogeneous at the histological level, and in the expression of DNA MMR markers, p53 and ARID1A.

The repertoire of mutations and gene copy-number alterations of ECCs

Of the 32 ECCs included in this study, we obtained high-quality targeted massively parallel sequencing data for 30 cases, at a median depth of coverage of 453× (range 156–838×) and 255× (118–540×) for tumour and normal samples, respectively (supplementary material, Table S2). ECCs harboured a median of five non-synonymous somatic mutations (range 2–538) in the 300 genes tested. In comparison, endometrial endometrioid and serous carcinomas from the TCGA dataset harboured a median of eight (2–276) and four (1–39) non-synonymous somatic mutations in the 300 genes studied here, respectively, which is statistically significantly different (ECCs versus endometrioid, P = 0.0014; ECCs versus serous, P = 0.0421; Mann–Whitney U-test). Two of the ECCs analysed here (6%) harboured mutations affecting the exounuclease domain of POLE (CC26, V411L; CC31, P436H), characterized by a very high number of somatic mutations (supplementary material, Table S3). After removing the two POLE ECCs from the analysis, we identified TP53 (46%), PIK3CA (36%), PPP2R1A (36%), FBXW7 (25%), ARID1A (21%), PIK3RI (18%), SPOP (18%) and KRAS (14%), as the most commonly mutated genes (Figure 2; supplementary material, Figure S1 and Table S3). Many of these mutations affected hotspots, including nine of 13 (69%) TP53 mutations, six of 10 (60%) PPP2R1A mutations, six of 10 (60%) PIK3CA mutations, and four of four (100%) KRAS mutations (Figure 2), and these were...
found to be clonal (i.e. biominatorically inferred to be present in almost all cancer cells within a tumour) (supplementary material, Figure S2 and Table S3). It is noteworthy that a subset of the mutations identified in the non-POLE ECCs affected genes previously reported to be preferentially mutated in endometrioid endometrioid carcinomas, including TP53, PPP2R1A, FBXW7, and SPOP, whereas others were previously found to be preferentially mutated in endometrioid endometrioid carcinomas, including PIK3R1 and KRAS (Figure 2) [4,8,11,13,14].

We did not identify any correlation between architectural patterns and specific mutations or genes affected by mutations (data not shown); however, the presence of focal areas with grade 3 nuclei was significantly higher in non-POLE TP53-mutant ECCs than in TP53-wild-type ECCs (n = 11 versus n = 2; P = 0.002, Fisher’s exact test). Thirteen of the 28 non-POLE ECCs subjected to targeted sequencing harboured a mutation in TP53, of which 10 (77%) showed aberrant p53 expression by IHC. CC33 lacked any p53 expression (i.e. null pattern), but no TP53 somatic mutation could be identified (Table 1; supplementary material, Table S3). All six ECCs found to have an ARID1A loss-of-function mutation (i.e. frameshift or stop-gain) lacked ARID1A expression (n = 5) or showed partial geographical loss of expression (n = 1) by IHC (Table 1; supplementary material, Table S3). CC05, however, lacked ARID1A expression but was found to be ARID1A wild-type. Finally, three of the four ECCs subjected to sequencing analysis and showing loss of MSH6 expression (n = 3) or equivocal (focal, weak) MSH6 expression (n = 1) by IHC also harboured somatic MSH6 loss-of-function mutations, whereas CC11 lacked MSH6 expression but no somatic MSH6 mutation was identified. In CC05, which lacked MLH1 and PMS2 expression, no somatic genetic alterations affecting MLH1 or PMS2 were found (Table 1; supplementary material, Table S3); however, MLH1 hypermethylation was identified by clinical genetics testing (data not shown).

At the copy-number level, the most recurrent somatic mutations were amplifications of CCNE1 (18%, five of 28 non-POLE ECCs) and ERBB2 (11%, 3/28).
and homozygous deletions of DAXX on 6p21 (11%, 3/28).

These findings demonstrate that ECCs are genetically heterogeneous, and that their repertoire of somatic genetic alterations includes recurrent hotspot mutations in TP53, PIK3CA and PPP2R1A, as well as CCNE1 and ERBB2 amplifications. In addition, a subset of ECCs harboured POLE exonuclease domain mutations and showed an ultramutator phenotype.
Figure 2. Non-synonymous somatic mutations detected by targeted capture massively parallel sequencing in ECCs. Recurrent \( n \geq 2 \) non-synonymous somatic mutations (top) and recurrent \( n \geq 2 \) amplifications and homozygous deletions (bottom) identified in 28 POLE-wild-type ECCs by massively parallel sequencing targeting 300 cancer-related genes are shown. Cases are depicted in columns; genes are depicted in rows. Mutation types and gene copy-number alterations are colour-coded according to the legend. Loss of heterozygosity of the wild-type allele in association with a somatic mutation is depicted by a diagonal bar. Two additional cases, CC26 and CC31, were found to harbour somatic POLE exonuclease domain mutations (EDMs), and are shown on the right. The phenobar provides information on stage, POLE status, and ARID1A and p53 IHC, and is colour-coded according to the legend. Indel, small insertion/deletion; SNV, single-nucleotide variant.

Comparison of the mutational repertoire of ECCs with that of endometrioid and serous carcinomas

Given that ECCs showed mutations in genes previously reported to be preferentially mutated in either endometrial endometrioid or serous carcinomas (see above), we sought to compare the repertoire of mutations affecting the 300 genes analysed in the 28 non-POLE ECCs included in this study with that of endometrioid and serous carcinomas from the TCGA dataset [11]. ECCs harboured significantly fewer mutations in CTNNB1 (0% ECC versus 37% endometrioid, adjusted \( P < 0.0001 \), Fisher’s exact test) and PTEN (7% ECC versus 78% endometrioid, adjusted \( P < 0.0001 \), Fisher’s exact test) than non-POLE endometrial endometrioid carcinomas (\( n = 183 \)), but significantly more mutations in PPP2R1A (36% ECC versus 5% endometrioid, adjusted \( P < 0.0001 \), Fisher’s exact test) and TP53 (46% ECC versus 11% endometrioid, adjusted \( P = 0.0151 \), Fisher’s exact test) (supplementary material, Table S1). In contrast, following corrections for multiple comparisons, the only gene that was more frequently altered in 44 serous carcinomas than in ECCs was TP53 (89% versus 46%, adjusted \( P < 0.0001 \), Fisher’s exact test; supplementary material, Table S1).

Molecular classification of ECCs and outcome

Given the heterogeneity in the repertoire of somatic mutations identified in ECCs, and the similarities and differences with endometrial endometrioid and serous cancers at the mutational level, we sought to define whether the ECCs could be classified into the molecular subtypes. For this, we employed a surrogate model for the molecular-based TCGA classification of endometrial endometrioid and serous carcinomas as described by Talhouk et al [38]. This surrogate integrates the POLE mutation status, the IHC-based MSI status and, as a surrogate for ‘copy-number’ status, the IHC-based p53 expression status. Two of the 32 ECCs harboured a mutation in the POLE exonuclease domain (CC26 and CC31), and were classified as being of POLE subtype (Table 2). Of the remaining cases, four ECCs showed abnormal DNA MMR protein expression and were classified as being of MMR-deficient (MMR-D) subtype. Eleven ECCs showed abnormal p53 expression patterns and were classified as being of copy-number high (serous-like) subtype, called p53 abnormal (p53 abn) [38]. The remaining 15 ECCs lacking POLE mutations and showing normal DNA MMR and p53 protein expression were classified as copy-number low (endometrioid), called p53 wild-type (p53 wt) [38] (Table 2; supplementary material, Figure S3).

Nine of the eleven patients in the p53 abn group presented at advanced stage (82%), whereas only seven patients of the other subtypes presented at advanced stage (33%; \( P = 0.0233 \), Fisher’s exact test). We also noted that p53 abn ECCs showed a significantly higher rate of dissemination to the peritoneum (8/11, 73%) than...
Table 2. Molecular-based classification of endometrial clear cell carcinomas

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>POLE EDM</th>
<th>MMR IHC</th>
<th>p53 IHC</th>
<th>Molecular subtype</th>
<th>Subtype by hierarchical clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC26</td>
<td>V411 L</td>
<td>MSH6 equivocal</td>
<td>Wild-type</td>
<td>POLE</td>
<td>POLE</td>
</tr>
<tr>
<td>CC31</td>
<td>P436H</td>
<td>MSH6 loss</td>
<td>Wild-type</td>
<td>POLE</td>
<td>POLE</td>
</tr>
<tr>
<td>CC05</td>
<td>Wild-type</td>
<td>MLH1/PMS2 loss</td>
<td>Wild-type</td>
<td>MMR-D</td>
<td>CN low (endometrioid)/MSI</td>
</tr>
<tr>
<td>CC10</td>
<td>Wild-type</td>
<td>MSH6 loss</td>
<td>Wild-type</td>
<td>MMR-D</td>
<td>NA</td>
</tr>
<tr>
<td>CC11</td>
<td>Wild-type</td>
<td>MSH6 loss</td>
<td>Wild-type</td>
<td>MMR-D</td>
<td>CN low (endometrioid)/MSI</td>
</tr>
<tr>
<td>CC20</td>
<td>Wild-type</td>
<td>MLH1/PMS2/MSH6 loss</td>
<td>Wild-type</td>
<td>MMR-D</td>
<td>CN low (endometrioid)/MSI</td>
</tr>
<tr>
<td>CC01</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Null pattern</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC02</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Positive</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC07</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Positive</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC13</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Positive</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC15</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Positive</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC16</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Null pattern</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC17</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Positive</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC23</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Positive</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC24</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Positive</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC28</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Positive</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC33</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Null pattern</td>
<td>p53 abnormal</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC03</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>NA</td>
</tr>
<tr>
<td>CC04</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC06</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC08</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC12</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC14</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC18</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC19</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC21</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC22</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC25</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN low (endometrioid)/MSI</td>
</tr>
<tr>
<td>CC27</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC29</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN low (endometrioid)/MSI</td>
</tr>
<tr>
<td>CC34</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
<tr>
<td>CC35</td>
<td>Wild-type</td>
<td>Retained</td>
<td>Wild-type</td>
<td>p53 wild-type</td>
<td>CN high (serous-like)</td>
</tr>
</tbody>
</table>

CN, copy number; EDM, exonuclease domain mutation; NA, not available, as no sequencing data; *Ward’s algorithm, Euclidean distance.

the other groups (7/22, 32%, P = 0.0053, Fisher’s exact test). Of the 15 patients with p53 wt ECCs, nine (60%) presented with early-stage disease and six (40%) presented with stage IV disease. At the time of follow-up, nine patients with p53 wt ECCs had died of disease, four were alive with disease, and five had no evidence of disease (Tables 1 and 2).

As an exploratory analysis, we assessed whether the molecular subtypes identified with the surrogate model were associated with outcome. We observed that the molecular subtypes as defined by the surrogate model were significantly associated with disease-free survival (P = 0.0183) in univariate analysis (Figure 3). Patients with ECCs of POLE or MMR-D subtype had a favourable outcome (no events) as compared to those with ECCs of p53 wt or p53 abn subtype (Figure 3).

As a hypothesis-generating exploratory aim, we assessed whether unsupervised hierarchical clustering of the mutations identified in the 300 genes studied in the ECCs and all TCGA endometrial endometrioid and serous carcinomas (n = 244) would allow a classification of ECCs on the basis of their mutational profiles. This cluster analysis revealed three stable clusters: one enriched for endometrial carcinomas of the POLE subtype, one enriched for copy-number high (serous-like) cancers, and one encompassing the majority of copy-number low (endometrioid) and MSI (hypermutated) cancers (Figure 4; supplementary material, Figures S4 and S5). The ECCs classified as being of POLE, MMR-D and p53 abn molecular subtypes based on the surrogate IHC assay described above also clustered with the respective TCGA endometrial cancers classified as POLE, copy-number low (endometrioid)/MSI and copy-number high (serous-like) in the hierarchical cluster analysis, respectively (Table 2). In contrast, only two of the 14 ECCs classified as of p53 wt subtype by the surrogate assay clustered with the copy-number low (endometrioid)/MSI endometrial cancers from TCGA; rather, 12 of 14 of these cases clustered with the copy-number high (serous-like) endometrial cancers on the basis of their mutational profile. We noted that a subset of p53 wt ECCs harboured mutations in genes preferentially mutated in serous carcinomas [11], including SPOP, FBXW7, and/or PPP2R1A (Figure 2), which may drive the clustering. In addition, we observed that a subset of the surrogate assay-defined p53 wt ECCs showed aberrant copy-number profiles (supplementary material, Figure S6). In univariate analysis, the three subgroups identified on the basis of mutational profiles with...
Figure 3. Disease-free survival analyses of ECC patients stratified according to clinical features and molecular subtypes. (A–D) Kaplan–Meier disease-free survival curves for ECC patients stratified according to (A) stage of disease (stage I versus stage III/IV), (B) age (<65 years versus ≥65 years), (C) molecular subtypes defined with a surrogate model [38] {POLE, MMR-D, p53 wild-type [copy-number low (endometrioid)] and p53 abnormal [copy-number high (serous-like)]}, and (D) hierarchical clustering [POLE, copy-number high (serous-like) enriched, copy-number low (endometrioid)/MSI (hypermutated) enriched]. P values of the log-rank test are shown. CN, copy-number.

Figure 4. Hierarchical clustering of ECCs from this study and endometrioid and serous carcinomas from TCGA on the basis of somatic mutations identified in 300 cancer genes. Hierarchical cluster analysis of mutations identified in the 300 cancer genes included in our targeted massively parallel sequencing assay using Euclidean distance metric and Ward’s algorithm, including all ECCs from the current study, and all endometrial endometrioid and serous carcinomas from TCGA. Three stable clusters were identified (Figure S5): a POLE cluster, a cluster enriched for endometrial carcinomas of copy-number high (serous-like) subtype, and a cluster enriched for endometrial cancers of copy-number low (endometrioid) and MSI (hypermutated) subtypes. The tumour type and the molecular subtype of the endometrial endometrioid and serous carcinomas as defined by TCGA are presented in the phenobar below the heatmap, colour-coded according to the legend. The majority of ECCs clustered with the serous carcinomas/copy-number high (serous-like) tumours; however, ECCs were also found in the POLE and the copy-number low (endometrioid)/MSI (hypermutated)-enriched clusters.
hierarchical clustering were also significantly associated with outcome ($P = 0.0465$). These data provide further evidence that ECCs are heterogeneous at the genetic level, and that all molecular subtypes identified in endometrial endometrioid and serous carcinomas can be found in ECCs.

**Discussion**

Here, we have demonstrated that ECCs, which constitute a rare type of endometrial cancer, show a heterogeneous repertoire of somatic genetic alterations, affecting cancer genes previously found to be altered preferentially either in endometrioid or in serous endometrial carcinomas [8,11]. Furthermore, we observed that all four molecular subtypes identified in endometrial endometrioid and serous carcinomas in the TCGA dataset are represented in ECCs.

We confirm previous studies showing that TP53/p53 is the most commonly altered gene in ECCs [4,12–14]. On the basis of immunohistochemical analysis of 21 ECCs with oestrogen receptor, progesterone receptor, p53 and Ki67, Lax et al [40] concluded that there were three types of ECC: typical ECC, serous-like ECC, and endometrioid-like ECC. Hoang et al [13] suggested that a subset of ECCs with typical clear cell morphology may be biologically and clinically related to serous cancers. In addition, Fadare et al [3] found that 34% of ECCs harboured aberrant p53 expression by IHC, which was associated with significantly lower progression-free survival in univariate analysis. In our dataset, 62% of the TP53-mutant ECCs studied harboured concomitant mutations in PPP2R1A (six of 13 non-POLE TP53 mutant) or SPOP (two of 13 non-POLE TP53 mutant), akin to serous/copy-number high (serous-like) carcinomas [11]. Furthermore, we observed that 34% of ECCs were classified as being of the p53 abn subtype using a surrogate model, and were associated with poor disease-free survival in univariate analysis and showed a higher rate of peritoneal metastases than the other subtypes. Our findings provide additional evidence that the subgroup of ECCs harbouring TP53 mutations may be similar to endometrial serous carcinoma not only in biological behaviour but also at the genetic level.

Importantly, however, not all ECCs are similar to endometrial serous cancers in terms of their molecular profile and clinical behaviour. In fact, the entire spectrum of molecular subtypes previously described for endometrial endometrioid and serous carcinomas were identified in the ECCs studied here, which also include POLE, MMR-D and p53 wt/copy-number low (endometrioid) cancers. Two ECCs were found to harbour POLE exonuclease domain mutations (6%) associated with an extremely high mutational burden (371 and 727 somatic mutations in the 300 genes studied), and both showed abnormal MSH6 expression patterns. In addition, 13% (4/32) of the non-POLE ECCs were classified as MMR-D associated with an increased number of somatic mutations [median of 11 mutations (range 8–43) versus non-MMR-D, median of four mutations (range 2–25), $P = 0.012$, Mann–Whitney $U$-test]. Akin to the observations in endometrial endometrioid carcinomas harbouring POLE exonuclease domain mutations [11,41–43], both patients with POLE-mutant ECCs are currently alive without evidence of disease. Although a significant association between lost/equivocal expression of MSH6 and POLE exonuclease domain mutations was observed here ($P < 0.05$, Fisher’s exact test), only two ECCs concurrently showed both alterations (Table 2). Therefore, these findings should be perceived as hypothesis-generating, and warrant further studies to define the frequency and the molecular basis of the association between alterations of MSH6 and POLE exonuclease domain mutations. It should be noted that universal Lynch syndrome testing has been recommended for all newly diagnosed endometrial cancers [44]. Importantly, endometrial carcinomas showing a clear cell component have been found to be overrepresented among tumours with DNA MMR abnormalities [45].

MMR-deficient and POLE-mutant endometrial endometrioid cancers have been reported to show distinctive histological features [15,43]. A detailed histological review revealed no histological differences between ECCs with and without abnormal DNA MMR protein expression or POLE mutations other than differences in the pattern of immune infiltrate; 78% (25/32) were found to have a lymphoplasmacytic inflammatory response, which is a typical histological feature of ECCs. Importantly, however, 19% (6/32) of cases were found to harbour prominent peritumoural and/or intratumoural infiltrating lymphocytes. Of these six ECCs, four showed abnormal DNA MMR protein expression, and one also harboured a POLE exonuclease domain mutation.

The surrogate model for molecular subtype classification employed here was developed with endometrioid and serous/mixed carcinomas [38], and, although the associations with outcome in our ECCs were statistically significant, improvements to this model may be possible. Different types of TP53 mutation (e.g. missense, truncation, and frameshift) have been shown to affect the assessment of p53 by IHC [46–48], and not all TP53 mutants identified by massively parallel sequencing showed abnormal IHC patterns. In the model employed, p53 IHC is used as a surrogate for ‘copy-number’ status. We observed, however, that a subset of p53/TP53-wild-type ECCs had aberrant gene copy-number profiles (supplementary material, Figure S6), and on the basis of their mutational profile, most clustered with copy-number high (serous-like) endometrioid and serous carcinomas, suggesting that alterations in genes other than TP53 may lead to a ‘serous-like’ genetic make-up. Similarly, although the immunohistochemical analysis of DNA MMR proteins performed has high sensitivity and specificity for MSI [49], the correlation is not perfect.
The stratification of ECCs based on their genetic make-up may not only identify subsets with distinct outcomes but, in the era of precision medicine, may also help to guide treatment decision-making in the future. Eleven of the 30 ECCs subjected to targeted sequencing analysis (37%) harboured mutations affecting PIK3CA, PIK3R1, and/or PTEN, suggesting that, as with endometrioid and serous carcinomas, targeting of the phosphoinositide 3-kinase (PI3K) pathway in a subset of ECCs may constitute a therapeutic strategy [11,50]. Furthermore, we observed that three ECCs (10%) harboured ERBB2 gene amplifications and one (CC12) an ERBB2 hotspot mutation (S310F), providing evidence to suggest that ERBB2 may be a therapeutic target in a subset of ECCs. ARID1A, a member of the SWI/SNF complex, has been previously reported to be important in the pathogenesis of gynaecological clear cell carcinomas, and has been found to be mutated in 46–57% and 13% of ovarian and endometrial clear cell carcinomas, respectively [13,51,52]. Of the ECCs studied here, 22% lacked ARID1A expression, and all but one of these had an underlying ARID1A frameshift mutation. It has recently been suggested that pharmacological inhibition of EZH2 may represent a novel treatment strategy for cancers harbouring ARID1A mutations [53]. In addition, there is evidence indicating that ARID1A-mutated cancers may also be sensitive to targeting of the PI3K–AKT pathway or DNA damage response [54], and a clinical trial combining olaparib with the AKT inhibitor AZD5363 in ARID1A-mutant advanced solid tumours is currently recruiting participants (ClinicalTrials.gov: NCT02576444). Finally, immunotherapy has recently been added to the repertoire of possible treatments in patients with hypermutated and ultramutated cancers [55].

This study has several limitations. First, given that all ECCs were formalin-fixed and paraffin-embedded, and that the tumour and normal tissue material was limited, we subjected the ECCs to targeted sequencing focusing on cancer genes rather than to whole exome sequencing, and could not define the prevalence of somatic genetic alterations not surveyed with the methods employed in this study. Third, the tumour cell content of a subset of samples was low (i.e. <50%), which may have affected the mutation and/or copy-number analysis. The sequencing depth for these samples was high, however (median of 453x for tumours), and the frequency of mutations affecting known genes, including TP53 and ARID1A, was similar to that previously described. Fourth, given the low number of somatic mutations in ECCs in the 300 genes analysed, with the exception of the tumours harbouring POLE mutations, mutational signatures could not be determined. Further studies are warranted to define the mutational signatures in ECCs.

Despite the limitations, our data demonstrate that ECCs constitute a heterogeneous group of tumours in terms of histology, somatic genetic alterations, and clinical behaviour. Akin to endometrial endometrioid cancers, all molecular subtypes were represented in the ECCs studied, including POLE and MMR-deficient tumours with an excellent prognosis, and copy-number low (endometrioid)/p53 wt and copy-number high (serous-like)/p53 abn tumours with a poor prognosis. On the basis of these findings, the classification of all ECCs as ‘high-grade’ or ‘type II’ tumours according to Bokhman [7] may not be warranted.

Acknowledgements

We thank Jorge S. Reis-Filho for critically reading the manuscript. The research reported in this paper was supported in part by the Farmer Family Foundation and a Cancer Center Support Grant of the National Institutes of Health/National Cancer Institute (Grant No. P30CA008748). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This work was presented in part as a platform presentation at the 104th United States and Canadian Academy of Pathology (USCAP) Annual Meeting, Boston, MA, USA, 2015.

Author contributions statement

The authors contributed in the following way: DFD, RAS: conceived the study; DFD, RAS: reviewed and selected the cases; KAB, PS, RSL, SNS, SM, ASM, DTC, MFB: performed bioinformatics analysis; DFD, KAB, BW: analysed and interpreted the data; DFD, BW: wrote the first draft of the manuscript. All authors edited and approved the final draft of the manuscript.

References


**SUPPLEMENTARY MATERIAL ONLINE**

**Supplementary materials and methods**

Figure S1. Likely pathogenic non-synonymous somatic mutations detected by targeted capture massively parallel sequencing in endometrial clear cell carcinomas

Figure S2. Cancer cell fractions of non-synonymous somatic mutations detected by targeted capture massively parallel sequencing in endometrial clear cell carcinomas

Figure S3. Histological features and copy number profiles of molecular subtypes of endometrial clear cell carcinomas

Figure S4. Hierarchical clustering of endometrial clear cell carcinomas from this study and endometrioid and serous carcinomas from TCGA using somatic mutations identified in 300 cancer genes

Figure S5. Assessment of the stability of the hierarchical cluster analysis

Figure S6. Copy number profiles of endometrial clear cell carcinomas classified as of copy-number low (endometrioid)/p53 wild-type subtype using a surrogate model

Table S1. List of 300 genes included in the targeted capture massively parallel sequencing assay and frequency of mutations affecting these genes in endometrial clear cell carcinomas from this study and endometrial endometrioid and endometrial serous carcinomas from The Cancer Genome Atlas (TCGA) dataset

Table S2. Sequencing statistics of endometrial clear cell carcinomas subjected to targeted massively parallel sequencing

Table S3. Somatic mutations identified in endometrial clear cell carcinomas by targeted massively parallel sequencing