

Reproductive genetics

Genomic alterations in ovarian endometriosis and subsequently diagnosed ovarian carcinoma

A. Linder ¹, S. Westbom-Fremer^{2,3}, C. Mateoiu⁴, A. Olsson Widjaja¹, T. Österlund ^{5,6}, S. Veerla^{2,3}, A. Ståhlberg^{5,6,7}, B. Ulfenborg ⁸, I. Hedenfalk ^{2,3,*†}, and K. Sundfeldt ^{1,9,*†}

¹Department of Obstetrics and Gynecology, Sahlgrenska Center for Cancer Research, Institute of Clinical Sciences, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden

²Division of Oncology, Department of Clinical Sciences Lund, Lund University, Lund, Sweden

³Lund University Cancer Centre (LUCC), Lund University, Lund, Sweden

⁴Department of Pathology, Sahlgrenska University Hospital, Gothenburg, Sweden



⁵Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Gothenburg, Sweden

⁶Department of Clinical Genetics and Genomics, Sahlgrenska University Hospital, Gothenburg, Sweden

⁷Department of Laboratory Medicine, Sahlgrenska Center for Cancer Research, Institute of Biomedicine, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden

⁸Department of Biology and Bioinformatics, Systems Biology Research Center, School of Bioscience, University of Skövde, Skövde, Sweden

⁹Region Västra Götaland, Sahlgrenska University Hospital, Gothenburg, Sweden

*Correspondence address. Department of Obstetrics and Gynecology, Sahlgrenska Center for Cancer Research, Institute of Clinical Sciences, Sahlgrenska Academy at University of Gothenburg, SE-405 30 Gothenburg, Sweden. E-mail: karin.sundfeldt@gu.se (K.S.)  <https://orcid.org/0000-0002-7135-3132>; Division of Oncology, Department of Clinical Sciences Lund, Lund University, SE-223 81 Lund, Sweden. E-mail: ingrid.hedenfalk@med.lu.se (I.H.)  <https://orcid.org/0000-0002-6840-3397>

†These authors contributed equally to this work.

ABSTRACT

STUDY QUESTION: Can the alleged association between ovarian endometriosis and ovarian carcinoma be substantiated by genetic analysis of endometriosis diagnosed prior to the onset of the carcinoma?

SUMMARY ANSWER: The data suggest that ovarian carcinoma does not originate from ovarian endometriosis with a cancer-like genetic profile; however, a common precursor is probable.

WHAT IS KNOWN ALREADY: Endometriosis has been implicated as a precursor of ovarian carcinoma based on epidemiologic studies and the discovery of common driver mutations in synchronous disease at the time of surgery. Endometrioid ovarian carcinoma and clear cell ovarian carcinoma are the most common endometriosis-associated ovarian carcinomas (EAOs).

STUDY DESIGN, SIZE, DURATION: The pathology biobanks of two university hospitals in Sweden were scrutinized to identify women with surgically removed endometrioma who subsequently developed ovarian carcinoma (1998–2016). Only 45 archival cases with EAO and previous endometriosis were identified and after a careful pathology review, 25 cases were excluded due to reclassification into non-EAO (n = 9) or because ovarian endometriosis could not be confirmed (n = 16). Further cases were excluded due to insufficient endometriosis tissue or poor DNA quality in either the endometriosis, carcinoma, or normal tissue (n = 9). Finally 11 cases had satisfactory DNA from all three locations and were eligible for further analysis.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Epithelial cells were collected from formalin-fixed and paraffin-embedded (FFPE) sections by laser capture microdissection (endometrioma n = 11) or macrodissection (carcinoma n = 11) and DNA was extracted. Normal tissue from FFPE sections (n = 5) or blood samples collected at cancer diagnosis (n = 6) were used as the germline controls for each included patient. Whole-exome sequencing was performed (n = 33 samples). Somatic variants (single-nucleotide variants, indels, and copy number alterations) were characterized, and mutational signatures and kataegis were assessed. Microsatellite instability and mismatch repair status were confirmed with PCR and immunohistochemistry, respectively.

MAIN RESULTS AND THE ROLE OF CHANCE: The median age for endometriosis surgery was 42 years, and 54 years for the subsequent ovarian carcinoma diagnosis. The median time between the endometriosis and ovarian carcinoma was 10 (7–30) years. The data showed that all paired samples harbored one or more shared somatic mutations. Non-silent mutations in cancer-associated genes were frequent in endometriosis; however, the same mutations were never observed in subsequent carcinomas. The degree of clonal dominance, demonstrated by variant allele frequency, showed a positive correlation with the time to cancer diagnosis (Spearman's rho 0.853, $P < 0.001$). Mutations in genes associated with immune escape were the most conserved between paired samples, and regions harboring these genes were frequently affected by copy number alterations in both sample types. Mutational burdens and mutation signatures suggested faulty DNA repair mechanisms in all cases.

LARGE SCALE DATA: Datasets are available in the supplementary tables.

LIMITATIONS, REASONS FOR CAUTION: Even though we located several thousands of surgically removed endometriomas between 1998 and 2016, only 45 paired samples were identified and even fewer, 11 cases, were eligible for sequencing. The observed high level

Received: September 18, 2023. Revised: January 25, 2024. Editorial decision: February 1, 2024.

© The Author(s) 2024. Published by Oxford University Press on behalf of European Society of Human Reproduction and Embryology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

of intra- and inter-heterogeneity in both groups (endometrioma and carcinoma) argues for further studies of the alleged genetic association.

WIDER IMPLICATIONS OF THE FINDINGS: The observation of shared somatic mutations in all paired samples supports a common cellular origin for ovarian endometriosis and ovarian carcinoma. However, contradicting previous conclusions, our data suggest that cancer-associated mutations in endometriosis years prior to the carcinoma were not directly associated with the malignant transformation. Rather, a resilient ovarian endometriosis may delay tumorigenesis. Furthermore, the data indicate that genetic alterations affecting the immune response are early and significant events.

STUDY FUNDING/COMPETING INTEREST(S): The present work has been funded by the Sjöberg Foundation (2021-01145 to K.S.; 2022-01-11:4 to A.S.), Swedish state under the agreement between the Swedish government and the county councils, the ALF-agreement (965552 to K.S.; 40615 to I.H.; 965065 to A.S.), Swedish Cancer Society (21-1848 to K.S.; 21-1684 to I.H.; 22-2080 to A.S.), BioCARE—A Strategic Research Area at Lund University (I.H. and S.W.-F.), Mrs Berta Kamprad's Cancer Foundation (FBKS-2019-28, I.H.), Cancer and Allergy Foundation (10381, I.H.), Region Västra Götaland (A.S.), Sweden's Innovation Agency (2020-04141, A.S.), Swedish Research Council (2021-01008, A.S.), Roche in collaboration with the Swedish Society of Gynecological Oncology (S.W.-F.), Assar Gabrielsson Foundation (FB19-86, C.M.), and the Lena Wäpplings Foundation (C.M.). A.S. declares stock ownership and is also a board member in Tulebovaasta, SiMSen Diagnostics, and Iscaff Pharma. A.S. has also received travel support from EMBL, Precision Medicine Forum, SLAS, and bioMCC. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Keywords: ovarian endometriosis / clear cell ovarian carcinoma / endometrioid ovarian carcinoma / exome-sequencing / mutation / copy number alteration

Introduction

Endometriosis, a benign disease that affects 10–18% of young fertile women (Giudice, 2010; Zondervan et al., 2018), may transform into cancer in the ovaries later in life. There is substantial epidemiologic evidence supporting a link between ovarian endometriosis (OE) and epithelial ovarian carcinoma (OC). Women with endometriosis have a lifetime risk of OC of 2–3% compared to a risk of 1–1.5% in the overall female population (Pearce et al., 2012; Hermens et al., 2020). There are no known predictive molecular markers or signs indicating which women with endometriosis are at risk of developing OC. If they were known, active surveillance or prophylactic surgery could be indicated to save lives.

Endometrioid ovarian carcinoma (E-OC) and clear cell ovarian carcinoma (CC-OC) are the most common endometriosis-associated ovarian carcinomas (EAOs). The genomic landscape of endometriosis and EAO has been explored in recent years. Mechanisms involved in endometriosis development include epigenetic defects, chronic inflammation, and aberrations in the regulation of estrogen and progesterone receptor pathways, with complex interactions between stromal, inflammatory, and epithelial cells (Bulun et al., 2019; Mari-Alexandre et al., 2019).

Several findings of known cancer-driving mutations in deep infiltrating endometriosis (DIE), OE, and iatrogenic endometriosis with benign histology have been reported (Anglesio et al., 2015, 2017; Lac et al., 2018; Suda et al., 2018). Frequently affected genes in endometriosis include *ARID1A*, *PIK3CA*, *KRAS*, *PTEN*, and *TP53*, which are also frequently mutated in EAO (Kuo et al., 2009; Wiegand et al., 2010; Lu et al., 2015; Murakami et al., 2017).

Endometriosis is defined as the presence of endometrial stroma and epithelium outside of the uterine corpus. The lesions are most typically located in the pelvis area, on the peritoneal surface or infiltrating nearby organs, but can be found throughout the abdominal cavity and at distant sites such as the lungs. OE, DIE, adenomyosis, and peritoneal endometriosis are suggested to be distinct subtypes. Endometriosis lesion in the vicinity of, or in continuity with, EAO is a common histopathologic finding, in which shared somatic mutations have been demonstrated (Sato et al., 2000; Anglesio et al., 2015; Er et al., 2016; Wiegand et al., 2010; Yamamoto et al., 2011). Multiple mutations in cancer-associated genes in endometrial (eutopic) tissue samples and multifocal endometriosis (ectopic) lesions have previously been described, showing that *KRAS* and *PIK3CA* mutations are shared across multiple regions of endometriosis but not

between eutopic and ectopic endometrial tissue of individual subjects (Suda et al., 2018). Conserved mutations in cancer-associated genes have also been shown between epithelial endometrial, endometriotic, and cancer cells derived from a single ancestral clone (Suda et al., 2020). Based on these findings, it is appealing to imagine a shared mutational lineage between endometrial tissue, endometriotic lesions, and possibly metachronous EAO.

Investigations on EAO cases and endometriosis prior to cancer diagnosis have not yet been conducted and may shed light on causative mechanisms. In this study, we have explored the mutational and copy number landscapes in EAO and preceding specimens of OE in paired samples from 11 patients using archival material, macro- and laser capture micro dissection, and whole-exome sequencing.

Materials and methods

Patient selection and tissue samples

Formalin-fixed and paraffin-embedded (FFPE) tissue specimens from 45 archival cases with EAO and previous endometriosis, surgically removed prior to OC diagnosis, between 1998 and 2016 were identified at the Pathology Departments of Sahlgrenska University Hospital in Gothenburg (n = 13) and Skåne University Hospital in Lund (n = 32) (Fig. 1). After pathology review (C.M./S.W.-F.), 25 cases were excluded due to reclassification into non-EAO (n = 9) or because OE could not be confirmed (n = 16). Further cases (n = 9) had either insufficient endometriosis tissue or showed poor DNA quality in either normal or endometriosis tissue (Supplementary Table S1). Eleven cases had satisfactory DNA from all three locations and were eligible for further analysis. FFPE sections rich in non-neoplastic epithelium from the fallopian tube or uterine cervix (n = 5) or blood samples collected at cancer diagnosis (n = 6) were used as the germline reference.

Tissue dissection and DNA extraction

All tissue biopsies were assessed by a specialist in gynecologic pathology (C.M./S.W.-F.) on standard hematoxylin and eosin slides. OE epithelium was collected using the PALM MicroBeam system (Zeiss, Stockholm, Sweden) by Laser Capture Microdissection. Epithelium was collected from a median of one separate biopsy of OE tissue (range, 1–5) corresponding to a median of two (range, 1–5) disconnected OE loci per specimen

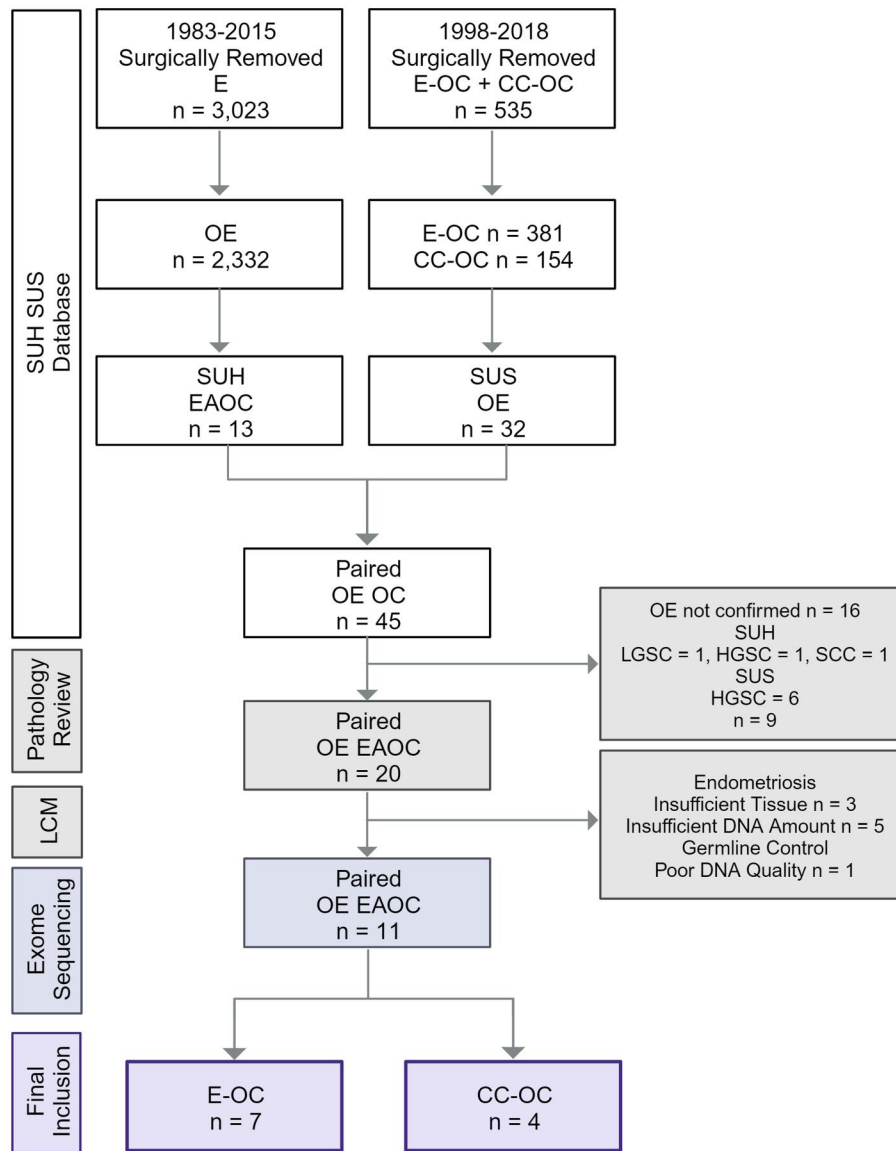


Figure 1. Patient flow chart and inclusion. CC-OC: clear cell ovarian carcinoma; E: endometriosis; EAOE: endometriosis-associated ovarian carcinoma; E-OC: endometrioid ovarian carcinoma; HGSC: high-grade serous carcinoma; LCM: laser capture microdissection; LGSC: low-grade serous carcinoma; OE: ovarian endometriosis; SCC: squamous cell carcinoma; SUH: Sahlgrenska University Hospital; SUS: Skåne University Hospital. Created with BioRender.com.

(Supplementary Fig. S1). DNA was extracted using the QIAamp DNA Micro kit (Qiagen) for Gothenburg cases and QIAamp FFPE kit (Qiagen) for Lund cases. Cell-rich carcinoma and normal tissue were macrodissected and DNA was extracted using the GeneRead DNA FFPE kit (Qiagen, Hilden, Germany) for Gothenburg cases and QIAamp FFPE kit (Qiagen, Hilden, Germany) for Lund cases. DNA from blood was extracted using QIAamp DNA Micro Kit according to the protocol (Qiagen). DNA concentration was assessed with the Qubit Fluorometer (ThermoFisher Scientific, MA, USA) and DNA quality was evaluated by TruSeq FFPE DNA Library Prep QC Kit where a ΔCt value <7.0 was required for continuation to exome sequencing.

Exome sequencing

Sample and library preparation, enrichment, and elution were performed according to the manufacturer's protocols (Human Comprehensive exome, Twist Bioscience, Carlsbad, CA). For library preparation, 50ng of genomic DNA was fragmented by enzymatic digestion, followed by end repair and dA-tailing. The

library was enriched by seven PCR cycles using twist unique dual index primers. Then, 200ng of the library was hybridized to the Twist CoreExome EF Multiplex Complete hybridization probes in a six-sample multiplex reaction. The library was amplified with seven cycles, purified, quantified, and sequenced (Nextseq500, Illumina, San Diego, CA). Image analysis and base calling were performed (Illumina Real Time Analysis, Illumina) with default parameters.

Data processing and analysis of somatic mutations

Read mapping and data processing were performed using the Sarek pipeline (<https://nf-co.re/sarek>). FreeBayes, Mutect2, and Strelka were run for calling single-nucleotide variants (SNVs). Variants were accepted if they passed all three callers, or if they passed FreeBayes and Strelka or Mutect2. For FreeBayes, a cutoff of QUAL >20 was used. Accepted mutations were filtered to remove those with an observed allele frequency of $<0.5\%$, those with sequencing depth of <5 observed reads in samples, and

those present in the normal reference sample. Variants were annotated with SnpEff, Variant Effect Predictor (VEP), and Catalogue of Somatic Mutations in Cancer (COSMIC) cancer census genes and counts (COSMIC, v95). Function and clinical implications were assessed based on Sift, PolyPhen, and ClinVar according to the VEP annotations. Previously suggested spurious genes (Lawrence et al., 2013; Tamborero et al., 2013; Shyr et al., 2014) were excluded unless identified as endometriosis-associated genes (Joseph and Mahale, 2019). The Maftools R package was used for characterization and visualization of mutational patterns and kataegis analysis (Mayakonda et al., 2018). Gene ontology and gene set enrichment were assessed with GSEA (Subramanian et al., 2005; Liberzon et al., 2015) and DAVID (Dennis et al., 2003). Mutational profiles were compared to the COSMIC single-base substitution (SBS), doublet-base substitution (DBS), and small indel (ID) mutational signatures (Alexandrov et al., 2013), by calculating cosine similarities (lsa package, version 0.73.3). Signatures with a similarity <0.5 to all samples were excluded. All samples and remaining signatures were grouped by similarity using hierarchical clustering and visualized as a heatmap (pheatmap package, version 1.0.12).

Copy number analysis

Somatic copy number alteration (CNA) analysis was performed using SAMtools (Li et al., 2009) pileup of normal and tumor sample BAM files, followed by CNA calling with VarScan2 (Koboldt et al., 2012). Calls were normalized by the ratio of uniquely mapped reads in the normal sample to uniquely mapped reads in the tumor sample and adjusted for GC content. VarScan2 results were smoothed and segmented with DNACopy package version 1.72.0 (Seshan, 2022), and recentered genome-wide. CNA segments were merged by VarScan2's mergeSegments.pl script using default parameters, and mapped to genomic annotations with the annotatr package 1.24.0 (Cavalcante and Sartor, 2017).

Immunohistochemistry

Immunohistochemical (IHC) staining was performed for MLH1 (FLEX Monoclonal mouse Anti-Human Clone ES05, Dako, Glostrup, Denmark), MSH2 (FLEX Monoclonal Mouse Anti-Human Clone FE11, Dako), MSH6 (FLEX Monoclonal Rabbit Anti-Human Clone EP49, Dako), and PMS2 (FLEX Monoclonal Rabbit Anti-Human 2 Clone EP51, Dako) using a Dako Autostainer (Autostainer Link 48) and EnVision FLEX, High pH (Link) kit. Loss of mismatch repair (MMR) protein expression was defined as granular or absence of nuclear staining for any of the MMR proteins in all epithelial cells or a distinct area of OE or tumor, with normal immunoreactivity in the internal positive control (i.e. lymphocytes and stromal cells). IHC for ARID1A (Monoclonal Rabbit Anti-Human clone EPR13501, Abcam) was performed with Ventana Benchmark Ultra. Complete loss of ARID1A expression in all tumor or OE cells was considered negative and for a subclonal pattern, evidence of complete negativity in sections of the tumor or OE was required. IHC was assessed independently by two observers (C.M., S.W.-F.). One of the EAO blocks with inadequate fixation was replaced by an omental metastasis from the same case.

Truemark microsatellite instability assay

For microsatellite instability (MSI) analysis, remnant DNA from the exome sequencing was used when available (Lund) or freshly prepared (Gothenburg). Tissue samples were macrodissected and DNA was extracted from the FFPE (QIAamp DSP DNA FFPE Tissue Kit, Qiagen) and corresponding blood (QIAamp DNA Micro Kit). PCR reactions were set up with 2 ng DNA and prepared for

electrophoresis (TrueMark™ MSI Assay, ThermoFisher Scientific). The fragment analysis reactions were diluted 1:100, or adjusted after initial assessment, before data collection (Applied Biosystems 3500 Genetic Analyzer). Data were analyzed with the TrueMark™ MSI Analysis Software (v1.0).

Statistical analysis

Statistical analysis was performed using the IBM SPSS Statistics version 28. Groups were compared with the Mann-Whitney U or Kruskal-Wallis test for two or more groups, respectively. Correlation was evaluated with Spearman's correlation. Two-sided tests were used and $P \leq 0.05$ was considered statistically significant. Two OC samples (CC-OC-04, E-OC-10) were excluded from statistics and calculations when the data were presented on OC-group level. Data in the text are presented as median with minimum and maximum values in parenthesis. A permutation test was performed to assess whether common missense mutations between paired samples were expected by chance. Permutations were performed for 1000 iterations, where mutations were randomly seeded in the exomes of 10 *in silico*-paired samples. The number of mutations per sample was assumed to follow a negative binomial distribution fitted to the mutation data in this study (except for E-OC-10 that harbored a much larger number of mutations). To make the permuted data comparable to the real data, mutations in each *in silico* sample were seeded in the exons affected by mutations in the corresponding sample.

Ethical approval

The study was approved by the ethics committee of the Sahlgrenska University Hospital (Dnr 201-15 and T522-17).

Results

Patient data

Eleven women with surgically removed OE who later developed OC were included in the present study. Seven developed E-OC and four developed CC-OC. The median age at diagnosis of OE and OC was 42 and 54 years, respectively (Supplementary Table S2). There was no difference in age distribution between the patients at the time of OE (Mann-Whitney, $P=0.315$) or OC diagnosis (Mann-Whitney, $P=0.649$) considering the two histotypes. The median time between OE and E-OC was 10 years (range, 4–27) and CC-OC 9 years (range, 9–30). There was no correlation between age at OE and time to cancer diagnosis (Spearman's rho -0.321 , $P=0.482$). Four patients had bilateral OE, and of them, one patient developed E-OC in both ovaries. Three patients had OE and OC on opposite sides, of which two underwent unilateral salpingo-oophorectomy at the time of OE surgery (E-OC and CC-OC), and the third patient developed an ultramutated E-OC. Laterality of the OE was unknown for four patients (Supplementary Table S2).

Whole-exome analysis of OE

Whole-exome sequencing of the 11 OE samples and their matched germline control was performed for the identification of endometriosis-associated somatic mutations. The median sequencing depth was 20 reads per identified variant (range, 5–726). A total of 4248 somatic mutations were detected with a median of 204 mutations per sample (range, 78–1453) or 22.6 (range, 5.5–281.8) somatic mutations per sequenced Mb (mut/mB) (Supplementary Fig. S2). The median fraction of non-silent mutations per sample was 15% (3.5–26), while 15% (10–40) and 69% (33–87) were silent or located outside transcribed region, respectively (Fig. 2A). In all, 763 non-silent mutations were identified in

624 genes (Supplementary Table S3). This corresponded to a median of 2.8 non-silent mut/Mb per sample (range, 0.2–73.9) with a median variant allele frequency (VAF) of 0.209 (range, 0.18–0.38). There was no correlation between age at OE diagnosis and the total mut/Mb (Spearman's $\rho = -0.409$, $P = 0.212$) or non-silent mut/Mb

in cancer-associated genes (Spearman's $\rho = -0.482$, $P = 0.133$). Median VAF was ≥ 0.2 in 89.5% of the variants (Fig. 2B). For endometriosis that later developed into E-OC and CC-OC, the median VAF was 0.29 (range, 0.27–0.44) and 0.30 (range, 0.27–0.32), respectively. The median VAF showed a positive correlation with time to OC diagnosis (Spearman's $\rho = 0.615$, $P < 0.05$). The relationship was emphasized when including the VAFs for non-silent mutations only (Fig. 2C, Spearman's $\rho = 0.853$, $P < 0.001$). The time between diagnoses showed a trend for positive correlation with cancer-associated mut/Mb that did not reach statistical significance (Spearman's $\rho = 0.550$, $P = 0.079$), and a corresponding sub-analysis of non-silent mutations showed a similar result (Spearman's $\rho = 0.495$, $P = 0.121$). There was a positive correlation between mut/Mb and number of individual biopsies with OE locus per specimen (Spearman's $\rho = 0.641$, $P < 0.05$), but not considering the number of spatially separate loci (Spearman's $\rho = 0.344$, $P = 0.301$). Hypermutated regions, kataegis, were observed in OE-03 and OE-05 (Supplementary Fig. S3A and B). They carried mutations possibly affecting apolipoprotein B mRNA editing enzyme catalytic polypeptide (APOBEC) genes, with an intronic variant in OE-03 (APOBEC3F c.567-50G>A) and an intergenic variant in OE-05 (RP11-444J21.2-APOBEC1 n.7781293T>C). These two samples also demonstrated a significantly high median VAF (median 0.36 and 0.44) (Fig. 2B).

Non-silent mutations in OE

Twenty-nine (29) genes were identified as affected by non-silent mutations in more than one patient (Fig. 3A). Eight of the samples displayed one or more non-silent mutations in a cancer-associated gene (Fig. 3B). In all, 136 mutations, corresponding to 37 cancer-associated genes were observed (Supplementary Table S3). This included a recurrent mutation in KRAS, represented by two likely pathogenic mutations, c.35G>C and c.35G>A. Of the 37 mutated cancer-associated genes, 17 (46%) were associated with maintenance of genome integrity such as DNA damage, methylation and chromosomal rearrangement (Fig. 3B), and six samples displayed mutations likely to contribute to genomic instability (Supplementary Table S4). Five samples harbored mutations affecting the MAPK signaling pathway, including mutations in MOS and CHUK (OE-03), FGFR4 (OE-05), KRAS (OE-06, OE-07), CACNA1B (OE-07), MAP3K13, and PPP3CA (OE-09) (Supplementary Table S3). A comparison with previously published exome-sequencing data of endometriosis, without any known association with cancer, showed that 52% of the mutated genes in the present study were mutated in OE (Li et al., 2014), while only three mutated genes were shared with endometriosis from other sites (Anglesio et al., 2017) (Fig. 3C). Of the cancer-associated genes identified in the current study, 68% (25/37) were also mutated in OE (Fig. 3D). A sub-analysis comparison to genes that were found mutated in $>10\%$ (2/17) of the samples in the external OE cohort showed that 47% (17/36) of the cancer-associated genes in our dataset also harbored a mutation in the dataset from OE in women without any known cancer association (Li et al., 2014). The corresponding comparison with data from non-OE showed that only two of the cancer-associated genes overlapped with OE: one from each data set.

Whole-exome analysis of the subsequent OC

Whole-exome sequencing of the 11 OC samples and their matched germline control was performed for identification of cancer-associated somatic mutations. The median read depth was 21 (range, 5–633). A total of 33 232 somatic mutations, corresponding to 6909 affected genes were detected (Supplementary Fig. S4). The median number of detected mutations per OC

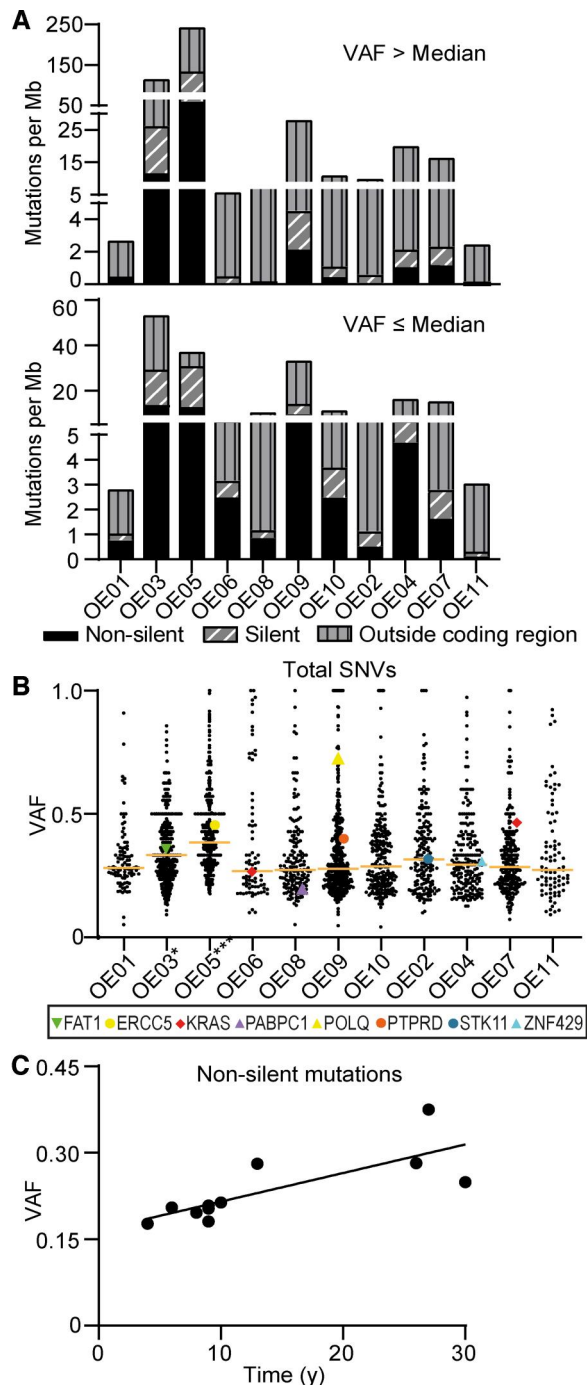


Figure 2. Evaluation of the mutational landscape of ovarian endometriosis by whole-exome sequencing. (A) Distribution of SNVs per Mb, separated by impact. Data were dichotomized by median VAF (≤ 0.353). (B) All mutations detected per specimen, horizontal line marks median VAF. Symbols annotate the non-silent SNV in the cancer-associated gene with the highest VAF. The VAF of OE-03 was higher than that of all samples except OE-05, and OE-05 had the significantly highest VAF (Kruskal-Wallis, $*P < 0.05$, $**P < 0.001$). (C) Correlation between median VAF of non-silent mutations, and time between OE and OC diagnoses. Mb: megabase; OE: ovarian endometriosis; SNV: single-nucleotide variant; VAF: variant allele frequency; Y: year.

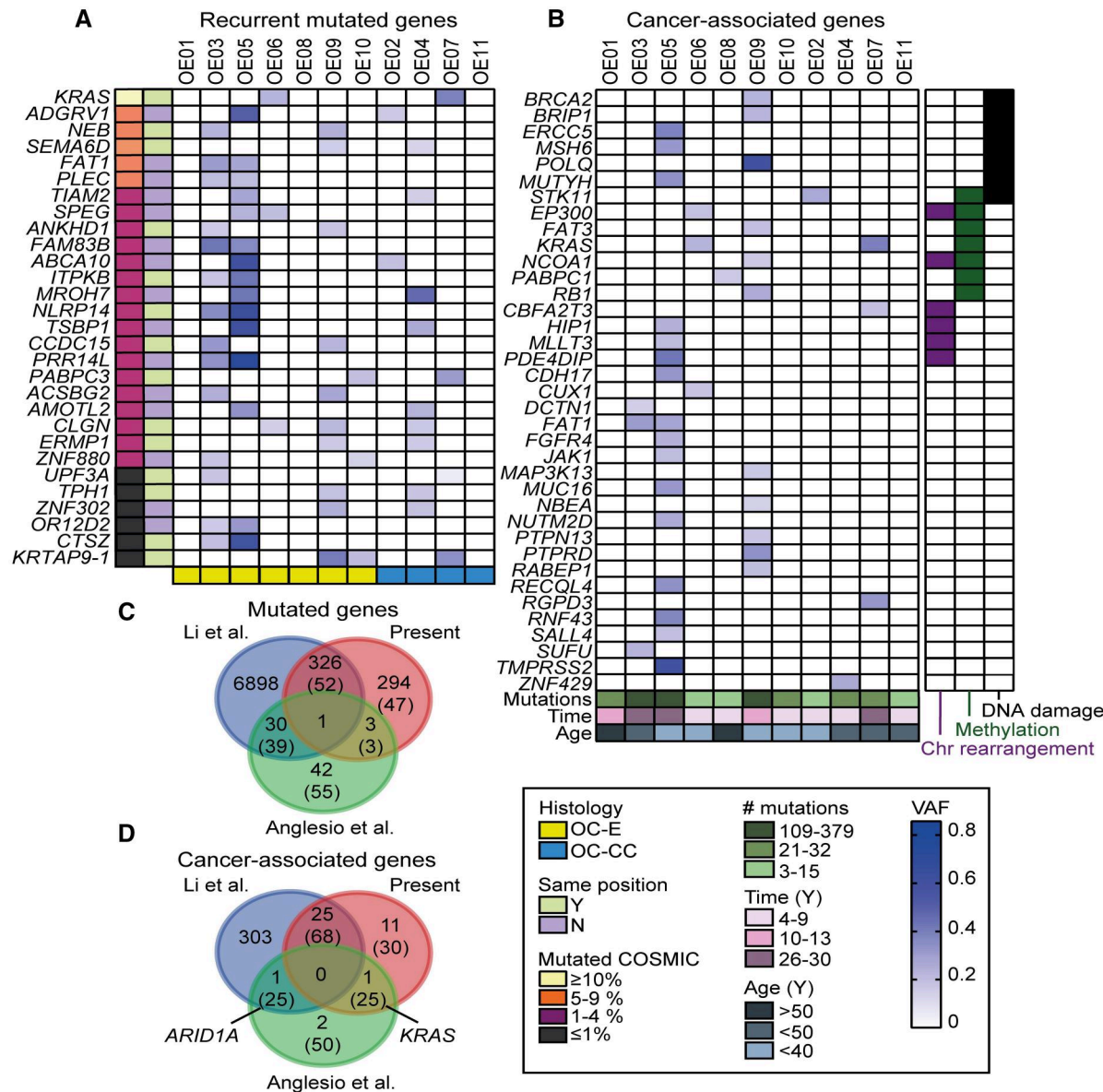


Figure 3. Non-silent mutations in ovarian endometriosis. (A) Genes harboring a mutation in more than one sample. Mutated genes assessed per sample and displayed with the histology of subsequent malignancy. In cases where the same basepair is altered in more than one specimen, this is marked in the figure. Additionally displayed is the reported mutation-frequency in COSMIC (left-hand side-panel). Color and density indicate the VAF of each somatic mutation. Reported spurious genes are excluded. (B) Mutated cancer-associated genes per sample. Patients ordered by histology of subsequent malignancy, OC-E (first seven columns) OC-CC (last four columns). Marked in the figure are the number of mutations according to low, medium, and high. Also shown is time between the two diagnoses and age at diagnosis. Overlapping functional annotations from GSEA are displayed (right-hand side-panel). (C) Venn diagram, comparison of the current dataset with previously published data. OE (Li et al.) and endometriosis in other loci (Anglesio et al.). (D) Equivalent comparison of cancer-associated genes. The number of overlapping genes is displayed together with the percentage within parenthesis. Percentage was calculated from the cohort with the least mutated genes. Chr: chromosome; COSMIC: catalogue of somatic mutations in cancer; N: No; OC-CC: clear cell ovarian carcinoma; OC-E: endometrioid ovarian carcinoma; OE: ovarian endometriosis; VAF: variant allele frequency; Y: Yes; (Y): time in years.

sample was 501 (range, 6–27 722). Two OCs were excluded from the statistical analysis and calculations; E-OC-10 had an excessive number of mutations, 27 722, of which 9666 were non-silent, and CC-OC-04 had an insufficient coverage of 1944 Mb at cutoff and only six mutations were identified, of which two were non-silent. After this exclusion, the range was 45–2097 mutations per sample, corresponding to a median of 32.9 mut/Mb (range, 3.95–328.9). Considering all mutations, 95.8% had a VAF ≥ 0.2 , with a median VAF per sample of 0.294 (range, 0.22–0.44) (Table 1) and a median VAF for non-silent mutations of 0.269 (range, 0.18–0.42). All OCs except E-OC-03 displayed an above-median VAF mutation in a cancer-associated gene (Fig. 4A). In all, 9996 non-silent

mutations were identified in 6561 genes (Supplementary Table S4). A median of 20% (range, 5–36%) of the mutations per sample were identified as non-silent (Fig. 4B), corresponding to a median number of non-silent mutations of 63 (range, 2–9666), or range of 3–295, after exclusion as above. The latter corresponded to a median of 4.8 mutations per Mb (range, 0.3–119.9).

Non-silent mutations in OC

There were 52 genes identified as affected by non-silent mutations in cancer-associated genes (Fig. 4C). The most frequently mutated cancer-associated gene was TP53, which was mutated in five OC samples (three E-OC and two CC-OC). All TP53

mutations were predicted to be pathogenic. The genes *KRAS*, *PIK3CA*, *AFF3*, and *SPEN* were mutated in three carcinomas. Two tumors, both of the CC-OC histotype, had deleterious mutations in *ERBB3*. Two tumors, one of each OC histotype, had a predicted pathogenic mutation in *PIK3CA* hotspots (c.1624G>A and c.3140A>G). Three E-OC specimens harbored mutations in *KRAS*, of which two displayed a hotspot mutation in c.35G>A, while the hypermutated E-OC-10 harbored an atypical mutation, c.39C>T. Stop-gain mutations were detected in *PTEN* and *ARID1A* in one E-OC and one CC-OC, respectively. The *ARID1A* stop-gain mutation was confirmed with IHC, while the remaining specimens were positive for *ARID1A* (Supplementary Fig. S5). The hypermutated E-OC-10 sample carried non-silent mutations in 331 cancer-associated genes, including a deleterious missense mutation in the exonuclease domain of *POLE*; in all, it was analogous with the ultramutated genotype of endometrial carcinoma. This tumor also harbored mutations in several genes associated with maintenance of genome integrity (Supplementary Table S5). Additionally, E-OC-10 showed evidence of kataegis (Supplementary Fig. S3C) and had multiple mutations in kataegis genes (seven in *APOBEC* genes and three in *REV1*), one of which was a missense variant (*APOBEC3B* c.31C>T).

Shared mutations in paired endometriosis and OC

A total of 308 shared mutations were detected in paired OE and OC samples corresponding to a median of 19 (range, 0–134) mutations per sample. All samples displayed at least one shared SNV between paired OE and OC, apart from Case 04 (Supplementary Table S6). The median VAF per sample was 0.40 (range, 0.28–0.61) and 0.39 (range, 0.16–0.70) in the OE and corresponding OC. We detected 33 shared non-silent mutations in 31 genes (Fig. 5).

Eight cases shared 1–11 non-silent variants per case. Of the variants, 13 showed a higher VAF in OC versus OE, 13 had <0.05 difference in VAFs, and seven had a lower VAF in OC compared to OE. Except for one stop codon variant in *NEB*, all variants were of missense type. Of the common variants, 51% (17/33) were predicted tolerated and benign. Seven variants were assessed as being deleterious and possibly damaging, of these, two cases had higher VAF in OC than in OE: *DLGAP3* and *NUFIP1* (Case 01), and *W12-3308P17.2* (Case 09). Six cases (55%) shared mutations associated with effects on the immune response. Mutations in *PIBF1* (Case 01), *HLA-B* (Case 02), and *HLA-DQA2* (Case 08) were predicted as damaging or deleterious, while the remaining mutations were tolerable or benign, or no prediction annotation was made. The permutation test showed that the probability of observing one or more common missense mutations by chance was <5%.

SBS, DBS, and indel signatures

Evaluation of the cosine similarity between the mutational profile of each sample and COSMIC SBS and DBS signatures suggested that DNA repair mechanisms were frequently represented by signatures with ≥ 0.5 similarity (Fig. 6). Signatures attributed to mismatch repair deficiency (MMRd) (SBS6, 15, 26, 44, and DBS7), were observed in nine OE and seven OC samples, respectively. IHC staining for MMRd confirmed aberrations in two OCs (OC-E-03 and OC-E-10) (Table 1, Supplementary Fig. S6). IHC for OC-CC-11 was inconclusive due to fixation artefacts, while the remaining samples showed no MMR aberrations. Analysis of MSI suggested that OE-02 displayed high instability, while data from OE-04 and OE-05 were inconclusive, and the remaining OEs did not display MSI. Four OCs displayed MSI, two with high (OC-E-05 and OC-E-10) and two with low MSI (OC-CC-02 and OC-E-08). The

Table 1. Data summary for all samples.

Group	Sample	SNVs				CNAs				IHC		
		Mut/Mb		VAF (median)		Total		Cancer-associated genes		ARID1A	Results (MMRd)	MSI
		Total	Non-silent	Total	Non-silent	Loss	Gain	Loss	Gain			
Ovarian endometriosis	OE-01	5.6	1.2	0.29	0.28	4138	602	146	19	P	MMRp	MSS
	OE-02	18.0	0.6	0.32	0.21	128	0	2	0	P	MMRp	MSI-High
	OE-03	169.3	25.6	0.36	0.28	776	203	22	4	P	MMRp	MSS
	OE-04	36.9	5.8	0.3	0.2	6	8	0	0	P	MMRp	No call
	OE-05	281.8	73.5	0.44	0.38	1252	19	43	0	P	MMRp	No call
	OE-06	13.0	2.5	0.27	0.21	6	14	0	0	P	MMRp	MSS
	OE-07	32.1	2.8	0.3	0.25	164	1677	5	37	P	MMRp	MSS
	OE-08	19.3	1.0	0.27	0.18	17	0	1	0	P	MMRp	MSS
	OE-09	61.7	11.8	0.28	0.21	38	0	0	0	P	MMRp	MSS
	OE-10	22.6	3.0	0.29	0.2	28	0	1	0	P	MMRp	MSS
	OE-11	5.5	0.2	0.27	0.18	77	679	2	19	P	IR	MSS
Ovarian carcinoma	OC-E-01	58.0	4.5	0.3	0.27	3145	6631	107	198	P	MMRp	MSS
	OC-CC-02	45.2	7.4	0.36	0.39	1251	414	43	13	P	MMRp	MSI-low
	OC-E-03	3.95	0.3	0.18	0.18	3981	2218	110	70	P	MMRd ^a	MSS
	OC-CC-04	3086.4*	1028.8*	0.42	0.42	0	10 468	0	329	P	MMRp	MSS
	OC-E-05	328.9	119.9	0.27	0.25	3402	2140	104	161	P	MMRp	MSI-high
	OC-E-06	94.1	4.8	0.22	0.2	555	1075	25	39	P	MMRp	MSS
	OC-CC-07	10.3	1.9	0.34	0.32	2963	2809	92	91	P	MMRp	MSS
	OC-E-08	22.0	4.9	0.47	0.34	11	4831	0	150	P	MMRp	MSI-low
	OC-E-09	32.9	7.7	0.26	0.23	1477	4114	55	143	P	MMRp	MSS
	OC-E-10	1013.6**	353.4**	0.42	0.41	100	80	5	2	P	MMRd ^b	MSI-high
	OC-CC-11	6.7	1.2	0.26	0.24	4038	1476	189	49	N	IR	MSS

CNA: copy number alterations; IHC: immunohistochemistry; IR: inconclusive results; MMRd: mismatch repair deficiency; MMRp: mismatch repair proficient; MSI: microsatellite instable; MSS: microsatellite stable; Mut/Mb: mutations per megabase; N: negative; P: positive; SNV: single-nucleotide variants; VAF: variant allele frequency.

^a Subclonal MSH2/6.

^b Negative MSH6.

* Low coverage.

** Hypermutated.

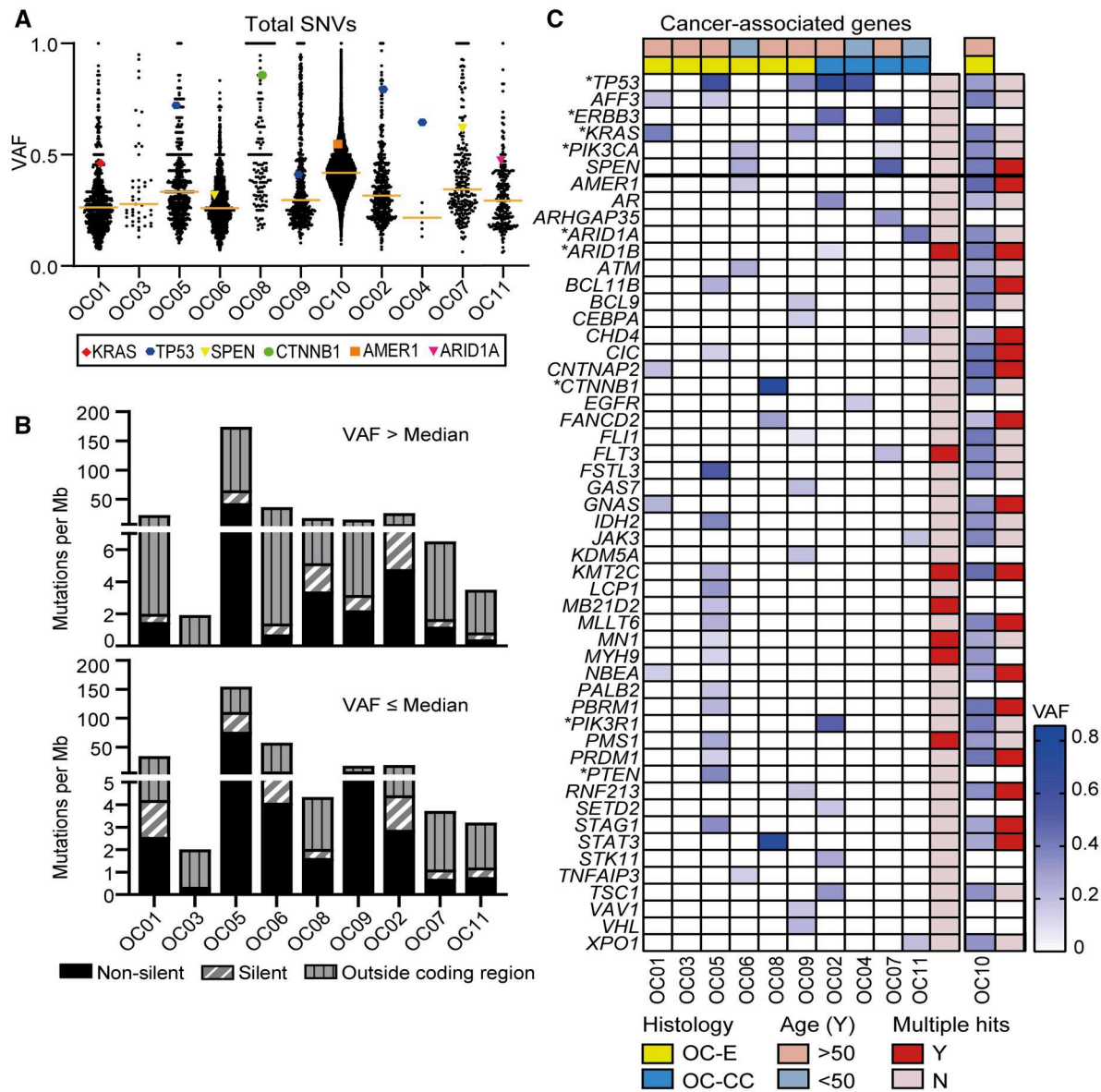


Figure 4. Characterization of mutations in endometriosis-associated ovarian carcinoma. (A) All mutations detected per specimen; the horizontal line marks the median VAF. Symbols annotate the non-silent SNV in the cancer-associated gene with the highest VAF. (B) Distribution of SNVs per Mb, separated by impact and location. Data were dichotomized by median VAF (≤ 0.402). (C) Heatmap of the 52 cancer-associated genes mutated in the 11 OC specimens. Color densities indicate the VAF of each somatic mutation. The hypermutated E-OC-10 was considered separately, and subsequently included for a complete data set. Horizontal demarcation indicates the six genes mutated in more than one sample, not including the contribution of E-OC-10. The asterisk (*) indicates genes expected to be affected by mutations in EAOC. EAOC: endometriosis-associated ovarian carcinomas; N: No; OC-CC: clear cell ovarian carcinoma; OC-E: endometrioid ovarian carcinoma; SNV: single-nucleotide variant; VAF: variant allele frequency; Y: Yes; (Y): Time in years.

signature associated with homologous repair deficiency (SBS3) was observed in eight OE and OC samples, respectively. The most frequent signature, suggestive of deficiency in transcription-coupled nucleotide excision repair (TC-NER) (SBS5) was seen in all samples, except OC-CC-04 and OC-E-10. Indel signatures (IDs), ID1 and ID2, suggestive of DNA slippage during replication, were the most frequently noted, in 77% (17/22) and 59% (13/22) of the samples. The OC-E-10, harboring missense mutations in *POLE*, *POLD1*, *MSH2*, *MSH6*, and *PMS2*, showed high similarity to additional signatures coupled to *POLE* or *POLD1* mutations (SBS20 and 14) and MMRd (SBS44 and DBS10).

CNAs in endometriosis and OCs

When comparing paired samples, OC showed more CNAs than OE, with a median of 256 (range, 17–521) events per tumor and

135 (range, 62–379) events per OE specimen. All samples in the present cohort displayed CNAs affecting multiple genes with a median of 128 (range, 14–4740), and 5542 (range, 181–10 468) genes affected in OE and OC, respectively (Table 1, Supplementary Table S7). Overall, the OEs displayed more chromosomal loss than gain of genetic material (Supplementary Fig. S7). Generally, chromosomal gains consisted of smaller chromosomal regions and partial gene sequences. Eight OEs displayed a CNA affecting one or more cancer-associated genes (Table 1), median 49.5 (range, 0–165), where a predicted impact could be estimated (i.e. deletion or amplification of transcribed regions or whole gene/s, respectively). OE-01 displayed the highest number of CNAs, including loss of a large fraction of chromosome X (ChrXp22.33-q28, LogR -0.89) (Fig. 7A), affecting numerous cancer-associated genes (145) where the entire or majority of the

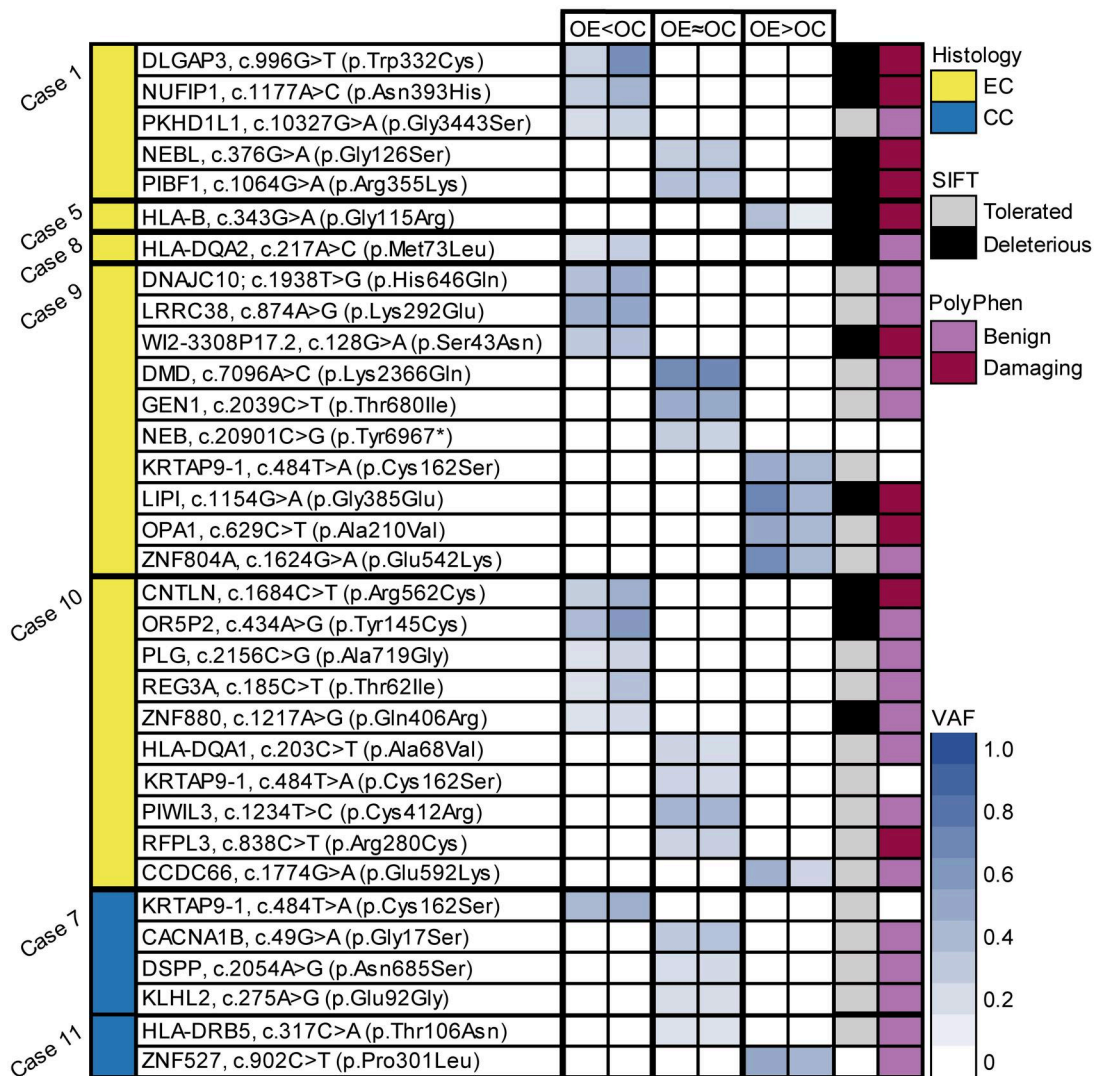


Figure 5. Shared non-silent mutations in paired endometriosis and carcinomas. Data are displayed by differences in variant allele frequency between paired samples of OE and OC. Data were trichotomized as increased VAF in OC (OE < OC), ≤ 0.05 difference (OE \approx OC), or decreased in OC (OE > OC). Affected genes are displayed with base pair and amino acid alterations. Variants are predicted as tolerated (gray) or deleterious (black) in SIFT and as benign (pink) or possibly/probably damaging (purple) in PolyPhen. CC: clear cell ovarian carcinoma; EC: endometrioid ovarian carcinoma; OE: ovarian endometriosis; VAF: variant allele frequency.

gene displayed loss. Corresponding chromosomal loss was seen in the subsequent OC (ChrXp22.33-q28, LogR -0.44 to -1.16). The most telomeric region, with the most discrepant LogR, harbored two cancer-associated genes, *FLNA* and *RPL10*. Overlapping loss of chromosome X was also seen in Case 05, where both OE and OC displayed loss of multiple cancer-associated genes including *FLNA* and *RPL10*. Both patients developed E-OC. Two OEs that later developed into CC-OC displayed large gains on chromosome 16, 16p13.3-p13.13 and 16p13.3-p13.11 (Fig. 7B) in OE-07 and OE-11, respectively. This region contained six cancer-associated genes that were amplified in both specimens. For Case 07, overlap was found in OE and the subsequent CC-OC (16p13.3-p13.13), affecting eight cancer-associated genes. Data showed that, except for Cases 08 and 10, all samples shared common regions with CNAs, with a median of 10 affected genes per sample (0–540 and 0–2179 genes with gains and losses, respectively). Two pseudogenes, *DUX4L1/3*, displayed gain in three samples, and in total 137 genes displayed gain in ≥ 2 cases, including five cancer-associated genes (*CAMTA1*, *PRDM16*, *RPL22*, *SKI*, *TNFRSF1*). Functional annotation suggested enrichment of tumor necrosis

factor signaling (Cases 07 and 11). Case 07 also displayed enrichment of amplification of genes associated with negative regulation of HLA class I molecules. Shared chromosomal loss in ≥ 2 samples affected 683 genes, including 31 cancer-associated genes, where loss of *BGN* was found in three cases. Four OCs and two OEs displayed CNAs on Chromosome 6 affecting immune response genes. Loss of 6p21.33 (Fig. 7C) was seen in three samples (OE-03, OE-05, OC-E-06), where the smallest region of overlap contained four genes (*HLA-B*, *HLA-C*, *MICA*, *MICB*). OC-E-09 displayed chromosomal gain affecting an adjacent transcribed region of *HLA-G* (201 bp, LogR 1.01). Gain of 6p22.1-p21.33 was seen in OC-CC-02; this region contained the major HLA class I genes and two additional cancer-associated genes (*POU5F1*, *TRIM27*). OC-E-03 showed deletion of 6p21.32, containing a part of *HLA-DPA1* (13 965 bp, LogR -0.82).

Discussion

In the present study, we assessed the genomic landscapes of paired OE and EAOC; 11 samples from women with OE who later

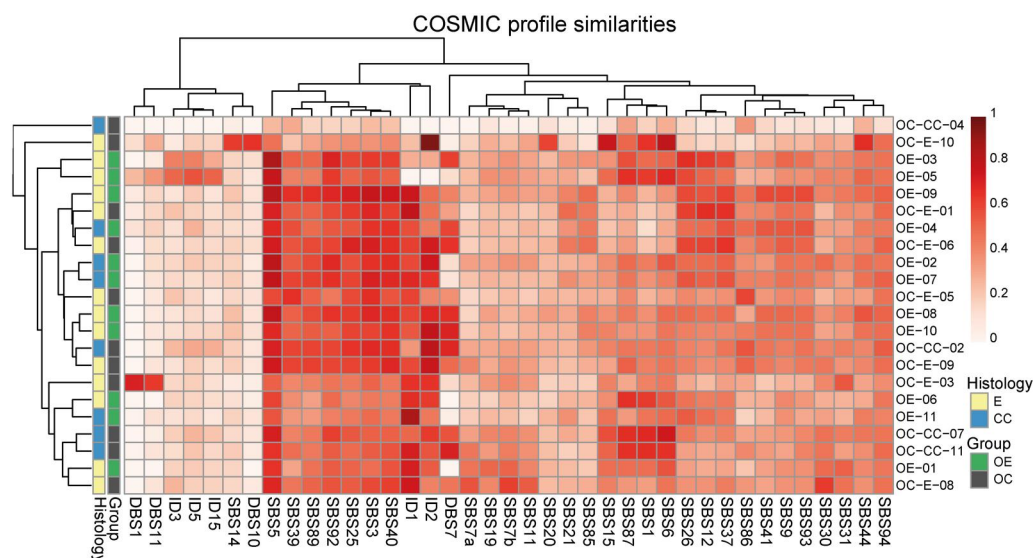


Figure 6. Profile similarities of SBS, DBS, and ID COSMIC signatures for all samples. Mutation profiles with a minimum of one sample with a signature similarity of ≥ 0.5 are included in the figure. The color density indicates the level of similarity (cosine similarity range 0.5–0.95). CC: clear cell; COSMIC: catalogue of somatic mutations in cancer; DBS: doublet-base substitution; E: endometrioid; ID: indel; OC: ovarian carcinoma; OE: ovarian endometriosis; SBS: single-base substitution.

developed either endometrioid or clear cell OC were analyzed. To our knowledge, this is the first study that has explored the mutation and copy number profiles of endometriosis with a confirmed subsequent malignant association.

We here identified conserved mutations between OE and paired OC, indicating a potential common ancestry. However, although mutations in cancer-driver genes have been suggested as the origin of malignant transformation of OE (Suda et al., 2020), no shared cancer-associated mutations were observed. Thus, although mutations in conventional cancer-driver genes were confirmed to occur frequently in OE, our original hypothesis could not be supported.

Noteworthy, the median VAFs of the shared mutations were higher than the overall median VAFs, suggesting a positive selection for clones harboring these mutations (Suda et al., 2018). Conserved mutations in genes associated with an immune response were the most prevalent, with affected genes in 5 of the 11 cases. Additionally, CNA analyses showed that aberrations involving immunological pathways were conserved. In total, seven cases exhibited genetic aberrations suggestive of an altered immune response. The single most frequently mutated type of genes was human leukocyte antigen (HLA) genes; four cases harbored a shared non-silent mutation in an HLA gene, three in class II, and one in HLA class I genes. Furthermore, the present data showed CNAs in four carcinomas and two endometriomas, mainly affecting HLA class I genes, emphasizing the importance of these mechanisms in the development of endometriosis and EAO. In line, chromosomal instability in the HLA super-locus on 6p21 has previously been described as a common alteration in EAO (Micci et al., 2014), and aberrant HLA expression has been described for both endometriosis (Semino et al., 1995; Matalliotakis et al., 2001) and OC (Hazini et al., 2021).

Additionally, the most frequently observed aberration in paired samples was chromosomal loss of BGN (encoding biglycan). Low BGN expression has been associated with decreased immunogenicity and poor patient outcome in KRAS-driven malignancies (Subbarayan et al., 2022). Furthermore, a deleterious mutation in *PIBF1* (encoding progesterone immunomodulatory

binding Factor 1) was seen in paired samples. This protein regulates the immune response during normal pregnancy, and has been suggested to facilitate tumor growth by suppressing the local antitumor response (Szekeres-Bartho and Polgar, 2010). Taken together, these genetic aberrations suggests that adaption to inflammation is an early and crucial event, in line with known mechanisms by which tumors evade the immune system (Jhunjhunwala et al., 2021). Furthermore, this finding suggests that immune-associated treatments may be beneficial not only for the treatment of endometriosis but may also have cancer-preventive effects.

Considering the chronic inflammation accompanying endometriosis (Zondervan et al., 2018), and the well-described association between inflammation and the genomic instability associated with tumor development (Coussens and Werb, 2002), it is reasonable to assume that the inflammatory environment contributes to the increased incidence of EAO. Based on other types of cancer and their association with inflammation (Balkwill and Mantovani, 2001), it could be expected that chronic inflammation would increase the incidence of cancer to an even higher extent than what is seen in association with endometriosis (Pearce, et al., 2012; Hermens et al., 2020). Furthermore, chronic inflammation is not exclusive to OE, but universal to endometriosis, and mutations in cancer-driver genes have been shown for different types of endometriosis, although no association with EAO has been described (Anglesio et al., 2015, 2017). Interestingly, the present data show a strong positive correlation between VAF and time between the two diagnoses. This implies that a dominant OE may delay tumor establishment; in other words, protect against an emerging malignancy. This is in line with the previously shown cancer-preventive properties of non-malignant epithelium in association with an accumulation of cancer-associated mutations (Colom et al., 2021). Moreover, the data show that there was no correlation between age at OE diagnosis and mutational burden, suggesting that mutations are selected, unlike mutations in normal endometrium that passively accumulate over time (Moore et al., 2020). This is in line with clonal evolution of cancer (Greaves and Maley, 2012), and

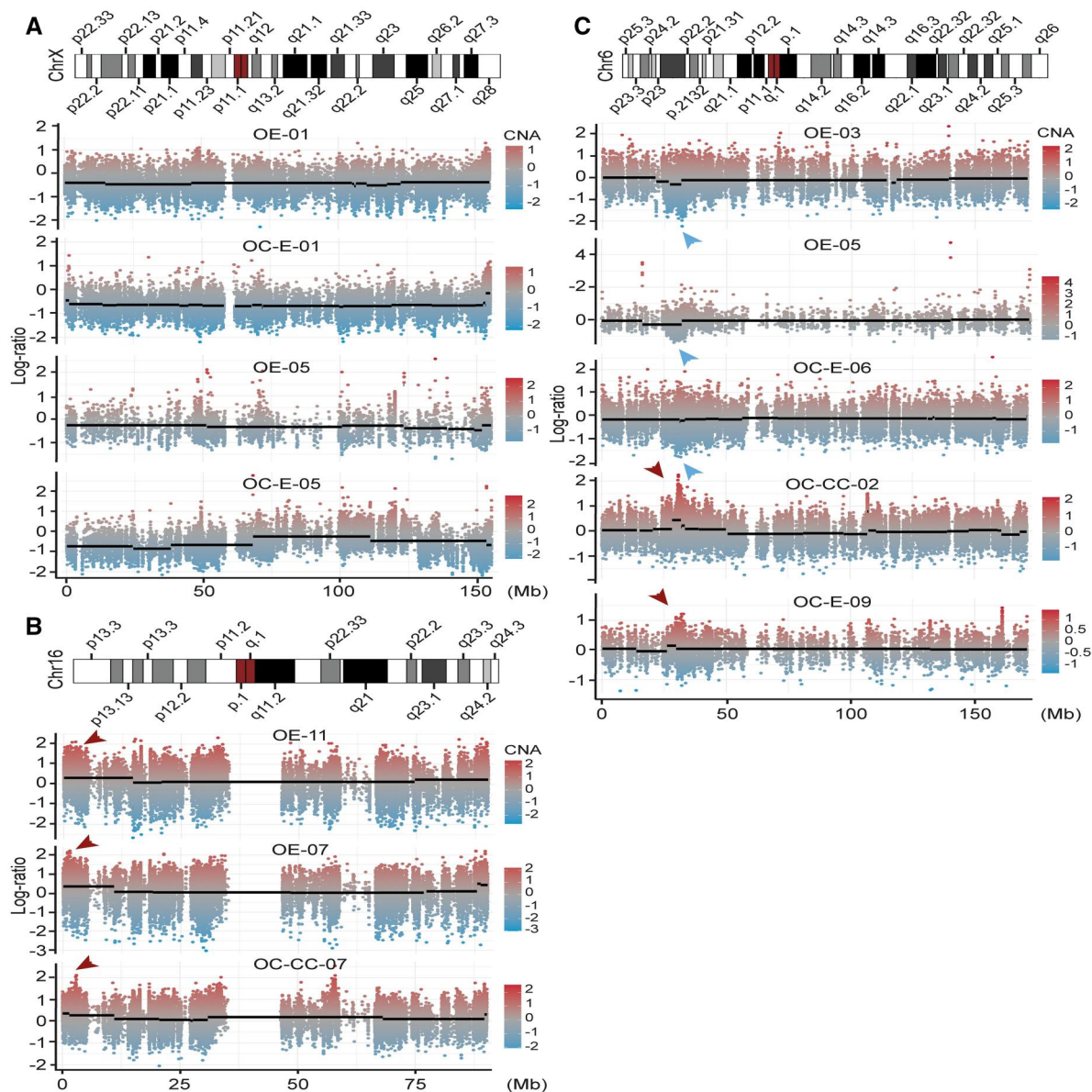


Figure 7. Visualization of recurrent large copy number alterations. (A) Extensive loss of chr X in two OE and paired OC with endometrioid histotype. (B) Gain of chr16 in OE that later developed OC of clear cell histotype, a corresponding gain was detected in paired OC for one of the cases. (C) Two OEs and four OCs displayed CNAs affecting chr 6, including loss and gain affecting HLA genes of class I and II. The y-axis represents the normalized median logR. chr: chromosome; CNA: copy number alteration; Mb: megabase; OC-CC: clear cell ovarian carcinoma; OC-E: endometrioid ovarian carcinoma; OE: ovarian endometriosis.

supports previous observations that clonal expansion of OE involves the accumulation of mutations in cancer-driver genes (Suda et al., 2018). However, the absence of conserved cancer-associated mutations between OE and OC indicates that the mutations associated with the malignant transformation are later events that occur in an originally less-potent precursor. There is sparse data on the genomic configuration of multiple and/or bilateral endometriotic lesions in the same individual. One case report described multiple shared mutations in cancer-associated genes in synchronous samples from the endometrium, different endometriosis lesions, and CC-OC (Suda et al., 2020) and in another cohort, the same group found shared KRAS or PIK3CA mutations between multiple unilateral OE lesions in six individuals, and between bilateral OE lesions in a single case (Suda et al., 2018). These findings suggest an inherent potential for progression to cancer in all OE lesions within these individuals. Hence, despite the discrepancy in laterality between OE and

OC in some of the cases in our study, the pair-wise comparisons made are relevant and interesting in relation to clonal evolution in cancer progression.

KRAS was the most frequently mutated cancer-associated gene in the current study. All affected specimens displayed the same hotspot mutation (c.35G) except the hypermutated OC-10, which carried a rarer mutation in an adjacent codon. It has been suggested that selection for specific c.35G mutations drives different tumor types (Ostrow et al., 2016). The most well-described KRAS mutation in OE (p.G12V, c. G35T) (Suda et al., 2018) was observed in one of the OE samples. This specific mutation has previously been linked to inflammation in OE (Yachida et al., 2021a). Of note, expected cancer-associated genes such as PIK3CA, ARID1A, PPP2R1A, and FBXW7 (Anglesio et al., 2015, 2017; Suda et al., 2018) did not harbor any non-silent mutations in OE. Moreover, the number of OE specimens with observed cancer-associated mutations in the present data was higher than what

has previously been reported (Anglesio et al., 2017; Suda et al., 2018). Comparative analyses suggested that OEs share several commonly affected genes (Li et al., 2014), including a large fraction (68%) of cancer-associated genes, compared to only 1% in data from endometriosis at other locations (Anglesio et al., 2017), suggesting that this genetic profile is associated with the OE subtype. A limitation of the present study was the small cohort, thus further studies are warranted to determine whether a specific genetic profile of OE contributes to the development of EAOC.

Considering OC, the mutation frequency of the conventional EAOC-associated genes differed from what has previously been described. The most frequently reported mutated gene in EAOC, ARID1A, described in 50% and 30% of CC-OC and E-OC tumors (Yachida et al., 2021), was only found to be mutated in one CC-OC. The effect of the mutation was confirmed by IHC as the single case displaying aberrant ARID1A protein expression. Thus, the frequency of aberrant ARID1A protein expression was also considerably lower than expected (Heinze et al., 2022). PIK3CA, with a described mutation frequency of 41% and 53% in E-OC and CC-OC (Yachida et al., 2021b), respectively, was found mutated in E-OC only. In contrast, the most frequently mutated gene in the tumor specimens was TP53: mutated in three E-OC and two CC-OC specimens, respectively. This suggests that in the present cohort TP53 mutations were overrepresented, particularly in CC-OC.

The data showed that both OE and OC displayed mutational profiles associated with deficiency in genes controlling genomic integrity. Analysis of cosine similarity to mutational signatures suggested that most of the specimens displayed a profile associated with deficiency in homologous repair as well as defective DNA mismatch repair. Although these findings could only be confirmed for two OCs using IHC, the MSI results suggested that at least one OE and four OCs displayed instabilities consistent with MMR deficiency. Furthermore, seven of the OEs and four of the OCs displayed genetic aberrations such as non-silent mutations or CNAs in genes associated with maintenance of genomic integrity or repair mechanisms. This included aberrations shared between paired samples such as mutations in the homologous repair gene *GEN1* (Lee et al., 2015) and deletion of *FLNA*, associated with the BRCA damage response pathway (Velkova et al., 2010). Additionally, the ultramutated OC and two of the OEs displayed signs of cancer-like kataegis (Alexandrov et al., 2013). However, only the OC specimen displayed the nucleotide substitution patterns typically associated with APOBEC-related kataegis. The OEs, on the other hand, displayed significantly high VAFs and prolonged time between diagnoses, suggesting an exceedingly strong clone.

Taken together, these data suggest that an innate defective DNA damage response and overall genomic instability, while central to the malignant transformation, also account for the resilience of OE.

The study has several limitations. The aim was to study possible causality in surgically resected OE prior to OC diagnosis in the same patient. We acknowledge that even though this is the first study of its kind, the sample size was small, the sequenced endometriosis was resected, sometimes on the contralateral side of subsequent OC, and no in-between surgery data was known, i.e. recurrence of endometrioma. However, 8 of the 11 OCs were at stages I and II at the time of diagnosis and second surgery, which might suggest that these patients were more closely surveilled since OCs are commonly diagnosed at stages III and IV.

Taken together, the present study suggests that, as a robust endometriosis lesion prevails, the accumulation of specific

cancer-driver mutations may ultimately lead to the development of a malignancy. The time required, and probability, for this to occur are associated with the genotype of the original endometriotic lesion and the environment where the precursor is located. The positive correlation between the number of distinct OE subpopulations and mutational burden, and the overall low number of recurrent mutations in the cohort, suggest intra- and inter-patient heterogeneity. This study provides a novel comprehensive understanding of the genetic diversity of OE associated with metachronous carcinoma. The complexity presented here prompts further investigative studies for an in-depth understanding of the risk of malignant transformation associated with OE.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Data availability

The data underlying this article are available in the article and in its online [supplementary material](#).

Acknowledgements

We thank Birgitta Weijdegård, Maria Yhr, and Anna Ebbesson for excellent technical work, and Eva Rambech for MSI data evaluation. The authors would like to acknowledge the SciLifeLabs, Bioinformatics and Data Centre (BDC), Gothenburg University, and Center for Translational Genomics (CTG), Lund University for providing expertise and service with sequencing.

Authors' roles

Conceptualization: K.S., C.M., A.S., and T.Ö.; methodology: K.S., C.M., I.H., A.L., T.Ö., A.S., and B.U.; investigation: A.L., S.W.-F., C.M., B.U., S.V., and T.Ö.; validation: S.W.-F. and A.O.W.; visualization: A.L., B.U., A.O.W., S.W.-F., and S.V.; supervision: K.S., I.H., and A.L.; writing the original draft: A.L., S.W.-F., C.M., B.U., and K.S.; critical reviewing of data and draft: K.S., I.H., A.L., A.S., T.Ö., S.W.-F., A.O.W., S.V., and B.U.; and funding: K.S., I.H., A.S., S.W.-F., and C.M.. All authors read and approved the final manuscript.

Funding

The present work has been funded by the Sjöberg Foundation (2021-01145 to K.S.; 2022-01-11:4 to A.S.), Swedish state under the agreement between the Swedish government and the county councils, the ALF-agreement (965552 to K.S.; 40615 to I.H.; 965065 to A.S.), Swedish Cancer Society (21-1848 to K.S.; 21-1684 to I.H.; 22-2080 to A.S.), BioCARE—a strategic research area at Lund University (I.H. and S.W.-F.), Mrs Berta Kamprad's Cancer Foundation (FBKS-2019-28, I.H.), Cancer and Allergy Foundation (10381, I.H.), Region Västra Götaland (A.S.), Sweden's Innovation Agency (2020-04141, A.S.), Swedish Research Council (2021-01008, A.S.), Roche in collaboration with the Swedish Society of Gynecological Oncology (S.W.-F.), Assar Gabrielsson Foundation (FB19-86, C.M.), and the Lena Wäpplings Foundation (C.M.).

Conflict of interest

A.S. declares stock ownership and is also a board member in Tulebovaasta, SiMSen Diagnostics, and Iscaff Pharma. A.S. has

also received travel support from EMBL, Precision Medicine Forum, SLAS, and bioMCC. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL et al.; ICGC PedBrain. Signatures of mutational processes in human cancer. *Nature* 2013;**500**:415–421.
- Anglesio MS, Bashashati A, Wang YK, Senz J, Ha G, Yang W, Aniba MR, Prentice LM, Farahani H, Li Chang H et al. Multifocal endometriotic lesions associated with cancer are clonal and carry a high mutation burden. *J Pathol* 2015;**236**:201–209.
- Anglesio MS, Papadopoulos N, Ayhan A, Nazeran TM, Noe M, Horlings HM, Lum A, Jones S, Senz J, Seckin T et al. Cancer-associated mutations in endometriosis without cancer. *N Engl J Med* 2017;**376**:1835–1848.
- Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? *Lancet* 2001;**357**:539–545.
- Bulun SE, Wan Y, Matei D. Epithelial mutations in endometriosis: link to ovarian cancer. *Endocrinology* 2019;**160**:626–638.
- Cavalcante RG, Sartor MA. annotatr: genomic regions in context. *Bioinformatics* 2017;**33**:2381–2383.
- Colom B, Herms A, Hall MWJ, Dentro SC, King C, Sood RK, Alcolea MP, Piedrafita G, Fernandez-Antoran D, Ong SH et al. Mutant clones in normal epithelium outcompete and eliminate emerging tumours. *Nature* 2021;**598**:510–514.
- Coussens LM, Werb Z. Inflammation and cancer. *Nature* 2002;**420**:860–867.
- Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;**4**:P3.
- Er TK, Su YF, Wu CC, Chen CC, Wang J, Hsieh TH, Herreros-Villanueva M, Chen WT, Chen YT, Liu TC et al. Targeted next-generation sequencing for molecular diagnosis of endometriosis-associated ovarian cancer. *J Mol Med (Berl)* 2016;**94**:835–847.
- Giudice LC. Clinical practice. Endometriosis. *N Engl J Med* 2010;**362**:2389–2398.
- Greaves M, Maley CC. Clonal evolution in cancer. *Nature* 2012;**481**:306–313.
- Hazini A, Fisher K, Seymour L. Deregulation of HLA-I in cancer and its central importance for immunotherapy. *J Immunother Cancer* 2021;**9**:1–17.
- Heinze K, Nazeran TM, Lee S, Kramer P, Cairns ES, Chiu DS, Leung SCY, Kang EY, Meagher NS, Kennedy CJ et al. Validated biomarker assays confirm that ARID1A loss is confounded with MMR deficiency, CD8(+) TIL infiltration, and provides no independent prognostic value in endometriosis-associated ovarian carcinomas. *J Pathol* 2022;**256**:388–401.
- Hermens M, van Altena AM, Nieboer TE, Schoot BC, van Vliet H, Siebers AG, Bekkers RLM. Incidence of endometrioid and clear-cell ovarian cancer in histological proven endometriosis: the ENOCA population-based cohort study. *Am J Obstet Gynecol* 2020;**223**:107.e1–107.e11.
- Jhunjhunwala S, Hammer C, Delamarre L. Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. *Nat Rev Cancer* 2021;**21**:298–312.
- Joseph S, Mahale SD. Endometriosis knowledgebase: a gene-based resource on endometriosis. *Database (Oxford)* 2019;**2019**:1–9.
- Koboldt DC, Zhang QY, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;**22**:568–576.
- Kuo KT, Mao TL, Jones S, Veras E, Ayhan A, Wang TL, Glas R, Slamon D, Velculescu VE, Kuman RJ et al. Frequent activating mutations of PIK3CA in ovarian clear cell carcinoma. *Am J Pathol* 2009;**174**:1597–1601.
- Lac V, Verhoef L, Aguirre-Hernandez R, Nazeran TM, Tessier-Cloutier B, Praetorius T, Orr NL, Noga H, Lum A, Khattri J et al. Iatrogenic endometriosis harbors somatic cancer-driver mutations. *Hum Reprod* 2018;**34**:69–78.
- Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013;**499**:214–218.
- Lee SH, Princz LN, Klugel MF, Habermann B, Pfander B, Biertumpfel C. Human Holliday junction resolvase GEN1 uses a chromodomain for efficient DNA recognition and cleavage. *Elife* 2015;**4**:1–24.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;**25**:2078–2079.
- Li X, Zhang Y, Zhao L, Wang L, Wu Z, Mei Q, Nie J, Li X, Li Y, Fu X et al. Whole-exome sequencing of endometriosis identifies frequent alterations in genes involved in cell adhesion and chromatin-remodeling complexes. *Hum Mol Genet* 2014;**23**:6008–6021.
- Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;**1**:417–425.
- Lu Y, Cuellar-Partida G, Painter JN, Nyholt DR, Morris AP, Fasching PA, Hein A, Burghaus S, Beckmann MW, Lambrechts D et al.; International Endogene Consortium (IEC). Shared genetics underlying epidemiological association between endometriosis and ovarian cancer. *Hum Mol Genet* 2015;**24**:5955–5964.
- Mari-Alexandre J, Carcelen AP, Agababyan C, Moreno-Manuel A, Garcia-Oms J, Calabuig-Farinas S, Gilabert-Estelles J. Interplay between microRNAs and oxidative stress in ovarian conditions with a focus on ovarian cancer and endometriosis. *Int J Mol Sci* 2019;**20**:1–24.
- Matalliotakis IM, Athanassakis I, Goumenou AG, Neonaki MA, Koumantakis EE, Vassiliadis S, Koumantakis EE. The possible anti-inflammatory role of circulating human leukocyte antigen levels in women with endometriosis after treatment with danazol and leuporelin acetate depot. *Mediators Inflamm* 2001;**10**:75–80.
- Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018;**28**:1747–1756.
- Micci F, Haugom L, Abeler VM, Davidson B, Trope CG, Heim S. Genomic profile of ovarian carcinomas. *BMC Cancer* 2014;**14**:315.
- Moore L, Leongamornlert D, Coorens THH, Sanders MA, Ellis P, Dentro SC, Dawson KJ, Butler T, Rahbari R, Mitchell TJ et al. The mutational landscape of normal human endometrial epithelium. *Nature* 2020;**580**:640–646.
- Murakami R, Matsumura N, Brown JB, Higasa K, Tsutsumi T, Kamada M, Abou-Taleb H, Hosoe Y, Kitamura S, Yamaguchi K et al. Exome sequencing landscape analysis in ovarian clear cell carcinoma shed light on key chromosomal regions and mutation gene networks. *Am J Pathol* 2017;**187**:2246–2258.
- Ostrow SL, Simon E, Prinz E, Bick T, Shentzer T, Nagawkar SS, Sabo E, Ben-Izhak O, Hershberg R, Hershkovitz D. Variation in KRAS driver substitution distributions between tumor types is

- determined by both mutation and natural selection. *Sci Rep* 2016;**6**:21927.
- Pearce CL, Templeman C, Rossing MA, Lee A, Near AM, Webb PM, Nagle CM, Doherty JA, Cushing-Haugen KL, Wicklund KG et al.; Ovarian Cancer Association Consortium. Association between endometriosis and risk of histological subtypes of ovarian cancer: a pooled analysis of case-control studies. *Lancet Oncol* 2012;**13**:385–394.
- Sato N, Tsunoda H, Nishida M, Morishita Y, Takimoto Y, Kubo T, Noguchi M. Loss of heterozygosity on 10q23.3 and mutation of the tumor suppressor gene PTEN in benign endometrial cyst of the ovary: possible sequence progression from benign endometrial cyst to endometrioid carcinoma and clear cell carcinoma of the ovary. *Cancer Res* 2000;**60**:7052–7056.
- Semino C, Semino A, Pietra G, Mingari MC, Barocci S, Venturini PL, Ragni N, Melioli G. Role of major histocompatibility complex class I expression and natural killer-like T cells in the genetic control of endometriosis. *Fertil Steril* 1995;**64**:909–916.
- Seshan VE. DNACopy: DNA copy number data analysis. R package version 1.70.0, 2022.
- Shyr C, Tarailo-Graovac M, Gottlieb M, Lee JJ, van Karnebeek C, Wasserman WW. FLAGS, frequently mutated genes in public exomes. *BMC Med Genomics* 2014;**7**:64.
- Subbarayan K, Massa C, Leisz S, Steven A, Bethmann D, Biehl K, Wickenhauser C, Seliger B. Biglycan as a potential regulator of tumorigenicity and immunogenicity in K-RAS-transformed cells. *Oncoimmunology* 2022;**11**:2069214.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;**102**:15545–15550.
- Suda K, Cruz Diaz LA, Yoshihara K, Nakaoka H, Yachida N, Motoyama T, Inoue I, Enomoto T. Clonal lineage from normal endometrium to ovarian clear cell carcinoma through ovarian endometriosis. *Cancer Sci* 2020;**111**:3000–3009.
- Suda K, Nakaoka H, Yoshihara K, Ishiguro T, Tamura R, Mori Y, Yamawaki K, Adachi S, Takahashi T, Kase H et al. Clonal expansion and diversification of cancer-associated mutations in endometriosis and normal endometrium. *Cell Rep* 2018;**24**:1777–1789.
- Szekeres-Bartho J, Polgar B. PIBF: the double edged sword. Pregnancy and tumor. *Am J Reprod Immunol* 2010;**64**:77–86.
- Tamborero D, Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Kandath C, Reimand J, Lawrence MS, Getz G, Bader GD, Ding L et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci Rep* 2013;**3**:2650.
- Velkova A, Carvalho MA, Johnson JO, Tavtigian SV, Monteiro ANA. Identification of filamin A as a BRCA1-interacting protein required for efficient DNA repair. *Cell Cycle* 2010;**9**:1421–1433.
- Wiegand KC, Shah SP, Al-Agha OM, Zhao Y, Tse K, Zeng T, Senz J, McConechy MK, Anglesio MS, Kalloger SE et al. ARID1A mutations in endometriosis-associated ovarian carcinomas. *N Engl J Med* 2010;**363**:1532–1543.
- Yachida N, Yoshihara K, Suda K, Nakaoka H, Ueda H, Sugino K, Yamaguchi M, Mori Y, Yamawaki K, Tamura R et al. Biological significance of KRAS mutant allele expression in ovarian endometriosis. *Cancer Sci* 2021a;**112**:2020–2032.
- Yachida N, Yoshihara K, Yamaguchi M, Suda K, Tamura R, Enomoto T. How does endometriosis lead to ovarian cancer? The molecular mechanism of endometriosis-associated ovarian cancer development. *Cancers (Basel)* 2021b;**13**:1–14.
- Yamamoto S, Tsuda H, Miyai K, Takano M, Tamai S, Matsubara O. Gene amplification and protein overexpression of MET are common events in ovarian clear-cell adenocarcinoma: their roles in tumor progression and prognostication of the patient. *Mod Pathol* 2011;**24**:1146–1155.
- Zondervan KT, Becker CM, Koga K, Missmer SA, Taylor RN, Vigano P. Endometriosis. *Nat Rev Dis Primers* 2018;**4**:9.

© The Author(s) 2024. Published by Oxford University Press on behalf of European Society of Human Reproduction and Embryology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Human Reproduction, 2024, 39, 1141–1154
<https://doi.org/10.1093/humrep/deae043>
Original Article