

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Cheasley, D;Nigam, A;Zethoven, M;Hunter, S;Etemadmoghadam, D;Semple, T;Allan, P;Carey, MS;Fernandez, ML;Dawson, A;Köbel, M;Huntsman, DG;Le Page, C;Mes-Masson, AM;Provencher, D;Hacker, N;Gao, Y;Bowtell, D;deFazio, A;Gorringe, KL;Campbell, IG

Title:

Genomic analysis of low-grade serous ovarian carcinoma to identify key drivers and therapeutic vulnerabilities

Date:

2021-01-01

Citation:

Cheasley, D., Nigam, A., Zethoven, M., Hunter, S., Etemadmoghadam, D., Semple, T., Allan, P., Carey, M. S., Fernandez, M. L., Dawson, A., Köbel, M., Huntsman, D. G., Le Page, C., Mes-Masson, A. M., Provencher, D., Hacker, N., Gao, Y., Bowtell, D., deFazio, A., ... Campbell, I. G. (2021). Genomic analysis of low-grade serous ovarian carcinoma to identify key drivers and therapeutic vulnerabilities. *Journal of Pathology*, 253 (1), pp.41-54. <https://doi.org/10.1002/path.5545>.

Persistent Link:

<https://hdl.handle.net/11343/276541>

**Genomic analysis of low-grade serous ovarian carcinoma to identify key drivers and therapeutic vulnerabilities**

Dane Cheasley<sup>1,2</sup>, Abhimanyu Nigam<sup>1,2</sup>, Magnus Zethoven<sup>1,3</sup>, Sally Hunter<sup>2</sup>, Dariush Etemadmoghadam<sup>2</sup>, Timothy Semple<sup>4</sup>, Prue Allan<sup>5</sup>, Mark S. Carey<sup>6</sup>, Marta Llaurodo Fernandez<sup>6</sup>, Amy Dawson<sup>6</sup>, Martin Köbel<sup>8</sup>, David G Huntsman<sup>7</sup>, Cécile Le Page<sup>9</sup>, Anne-Marie Mes-Masson<sup>9</sup>, Diane Provencher<sup>9</sup>, Neville Hacker<sup>10</sup>, Yunkai Gao<sup>1,2</sup>, David Bowtell<sup>11</sup>, Anna deFazio<sup>12</sup>, Kylie L. Goringe<sup>+2</sup>, and Ian G. Campbell<sup>+\*1,2</sup>

<sup>1</sup>Cancer Genetics Laboratory, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia

<sup>2</sup>Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, VIC, Australia

<sup>3</sup>Bioinformatics Consulting Core, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia

<sup>4</sup>Molecular Genomics Core, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia

<sup>5</sup>Department of Clinical Pathology, Peter MacCallum Cancer Centre, and University of Melbourne, VIC, Australia

<sup>6</sup>Department of Obstetrics & Gynaecology, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1002/path.5545](https://doi.org/10.1002/path.5545)

<sup>7</sup>Department of Pathology and Laboratory Medicine, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

<sup>8</sup>Department of Pathology and Laboratory Medicine, University of Calgary, Calgary, AB, Canada

<sup>9</sup>Centre de recherche du Centre hospitalier de l'Université de Montréal (CRCHUM) and Institut du cancer de Montréal, Montreal, QC, Canada

<sup>10</sup>Prince of Wales Clinical School, University of New South Wales, NSW, Australia

<sup>11</sup>Cancer Genetics and Genomics Program, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia

<sup>12</sup> Centre for Cancer Research, The Westmead Institute for Medical Research, The University of Sydney and the Department of Gynaecological Oncology, Westmead Hospital, NSW, Australia

†These authors contributed equally

\***Correspondence to:** D Cheasley, Cancer Genetics Laboratory, Research Division, Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, Victoria, 3000, Australia. E-mail: [dane.cheasley@petermac.org](mailto:dane.cheasley@petermac.org)

**Conflict of interest:** No conflicts of interest were declared.

**Running title:** Genomic analysis of low-grade serous ovarian carcinoma

Word count: ~4400 OK



## Abstract

Low-grade serous ovarian carcinoma (LGSOC) is associated with a poor response to existing chemotherapy, highlighting the need to perform comprehensive genomic analysis and identify new therapeutic vulnerabilities. The data presented here represents the largest genetic study of LGSOC to date (n=71), analysing 127 candidate genes derived from whole exome sequencing cohorts to generate mutation and copy-number variation data. Additionally, immunohistochemistry was performed on our LGSOC cohort assessing estrogen receptor, progesterone receptor, TP53 and CDKN2A status. Targeted sequencing identified 47% of cases with mutations in key *RAS/RAF* pathway genes (*KRAS*, *BRAF*, and *NRAS*), as well as mutations in putative novel driver genes including *USP9X* (27%), *MACF1* (11%), *ARID1A* (9%), *NF2* (4%), *DOT1L* (6%) and *ASH1L* (4%). Immunohistochemistry evaluation revealed frequent estrogen/progesterone receptor positivity (85%); along with CDKN2A protein loss (10%) and CDKN2A protein overexpression (6%) which were linked to shorter disease outcomes. Indeed, 90% of LGSOC samples harboured at least one potentially actionable alteration, which in 19/71 (27%) cases were predictive of clinical benefit from a standard treatment, either in another cancer's indications or in LGSOC specifically. In addition, we validated ubiquitin-specific protease 9X (*USP9X*), which is a chromosome X linked substrate-specific deubiquitinase and tumour suppressor, as a relevant therapeutic target for LGSOC. Our comprehensive genomic study highlighted that there is an addition to a limited number of unique 'driver' aberrations that could be translated into improved therapeutic paths.

**Keywords:** Low grade serous ovarian carcinoma, genomics, mutation, copy number, somatic, cancer driver genes, ubiquitin-specific protease 9X.

## Introduction

Low-grade serous ovarian carcinomas (LGSOC) represent 3–5% of all ovarian carcinomas and differ from high-grade serous ovarian carcinomas (HGSOC) in that over half harbour activating mutations of the RAS-RAF-MAPK pathway, are *TP53* wild-type and have comparatively fewer genomic copy number (CN) alterations. Women with LGSOC often have a poor prognosis, with the majority having advanced stage disease at diagnosis that is largely unresponsive to standard ovarian cancer chemotherapeutics, resulting in a high case-fatality-rate (80–90% at 10 years) similar to HGSOC [1-4]. LGSOC affects a greater proportion of younger women than HGSOC, causing loss of many more years of life [3].

There is considerable trial data optimizing treatment for HGSOC, although much less is known about the best treatment strategies for LGSOC. It is clear now that new therapies will not come from extrapolation of other ovarian cancer subtypes and in order to make significant treatment advances, a comprehensive molecular landscape of this disease is required.

There are few genomic studies of LGSOC, but the data suggest a comparatively low somatic mutation burden especially compared to HGSOC and endometrial serous carcinomas [5,6]. LGSOC could therefore be addicted to only a limited number of ‘driver’

genes and cancer signalling pathways, and therefore targeting these genes might be efficacious.

Previous combined whole-exome sequencing (WES) (n=38) and Sanger sequencing validation (n=19) of LGSOC cases identified recurrent mutations in *KRAS* (~22%), *BRAF* (~16%) and *NRAS* (~24%), and additional driver genes including the deubiquitinase *USP9X* and the protein translational regulator *EIF1AX* [5,7,8]. Previous CN analysis showed that LGSOC harbour few CN alterations, with the exception of frequent CN loss on chromosome 1p and homozygous deletions of the *CDKN2A/2B* locus [5,9]. The aim of the current study was to validate the frequency of these genomic alterations observed in the largest cohort of LGSOC sequenced to date as a means of identifying therapeutic avenues.

## Materials and methods

### Ethics approval and consent to participate

The biobanks received ethics approval from their local review boards to collect and share samples and clinical data. All subjects gave broad written consent to future research with their samples and data, without restriction. Additionally, the collection of the COEUR repository samples and data, received local ethics approval by the 'Comité d'éthique de la recherche du CHUM' (project reference #39-27-01-2017). Tumour sequencing was approved by the Peter MacCallum Cancer Centre Human Ethics Committee under protocol #09/29.

### Low grade serous ovarian cancer cohort

A total of 78 LGSOCs were evaluated; 77 from the Canadian Ovarian Experimental Unified Resource (COEUR [2,10]) and 1 from the Australian Ovarian Cancer Study (AOCS). The summary of key clinical characteristics is shown in **supplementary material, Table S1**. All LGSOCs were TP53 wild-type by Immunohistochemistry (IHC) [2], which is a key diagnostic criterion.

### Library construction and massively parallel sequencing

Only cases that had  $\geq 20$  ng of tumour DNA, and in a PCR-based quality assay

had amplifiable products of  $\geq 200$  bp were selected for sequencing (n=71) using a SureSelect XT Custom Panel (Agilent Technologies, Santa Clara, CA, USA), targeting 127 genes (**supplementary material, Table S2**). Library preparation was performed on tumour DNA using the KAPA Hyper Prep Kit (Kapa Biosystems, Wilmington, MA, USA). Libraries and capture were performed using the Bravo Automated Liquid Handling Platform (Agilent Technologies). Sequencing of target-enriched libraries was performed using the NextSeq500 (Illumina, San Diego, CA, USA) generating 75 bp paired-end sequence reads. The median sequencing depth for all LGSOC samples sequenced was 354 (range 83 to 1380) with 98% having  $\geq 100$ -fold coverage. Sequencing performance metrics for all samples is shown in **supplementary material, Table S3**.

### **Somatic and germline mutation detection**

Targeted sequence variants were called using GATK UnifiedGenotyper [11], Platypus [12] and VarScan 2 [13]. Called variants were annotated using the Ensembl Variant Effect Predictor Release 78. As no samples had matched normal DNA, variants were filtered to identify high confidence somatic variants as follows: Variants were excluded if reported in Gnomad, ExAc (minus TCGA samples), 1000 Genomes or EVS at an allele frequency of  $\geq 0.0001$ ; or if detected in  $>1$  of our in-house germline ovarian exomes ( $>600$  cases), or if detected in  $>20\%$  of the LGSOC cohort (except for known hotspot mutations in *KRAS* codon 12, *BRAF V600E*, *NRAS* codon 61 or variants with a COSMIC ID). A variant was also excluded if it was detected in a problematic gene listed by Scheinen *et al* [14]. The following filters were applied to reduce common technical artefacts. Variants had to have a QUAL score  $\geq 30$ , read depth  $\geq 20$ , alternative base read

depth  $\geq 10$ , variant read proportion  $> 0.2$  after tumour purity adjustment, and called by either VarScan, Unified Genotyper and Platypus. Variants also had to be absent from the normal female B-lymphocyte DNA control run in the same sequencing batch for the normalisation baseline (NA12878, Coriell Institute, Camden, NJ, USA). The germline variant calling pipeline for hereditary breast and ovarian cancer (HBOC) genes has been described previously by our group [15,16].

### **Genome-wide copy number analysis**

Off-target sequencing reads were used to generate genome-wide CN data using CopywriteR [17] utilising the female NA12878 control run in the same sequencing batch for normalisation. Data was then imported into Nexus Copy Number™ (v8.0, BioDiscovery Inc., El Segundo, CA, USA), segmented using a FASST2 algorithm, and visualized. Thresholds were  $\log_2$  ratios of  $\pm 0.2$  for gains and losses,  $> 0.6$  for high level gains and  $< -1$  for homozygous deletions.

Details for Extraction of tumour DNA, Sequencing of hereditary breast and ovarian cancer (HBOC) genes in tumour tissue, Immunohistochemistry (IHC), Homologous recombination deficiency and fraction of the genome altered scoring, Tumour purity estimation, Analysis of loss of heterozygosity across chromosome X, and CpG island methylation analysis of *USP9X* are provided in supplementary material, Supplementary materials and methods.

## Results

### LGSOC targeted gene panel design

To identify potential LGSOC driver genes for inclusion in the targeted sequencing validation, variant files from previously published WES data on 21 LGSOC and 13 serous borderline tumours (SBTs) [7]; 9 LGSOCs and 1 mixed carcinoma [5]; and 8 LGSOC [8] were merged. From this dataset, 32 genes were identified harbouring putative somatic mutations in at least two cases and 64 that were mutated in one case but fulfilled one of the following criteria: Classified as a somatic driver in either the COSMIC or IntOGen databases (n=50); reported mutated in TCGA (n=9) [18]; showed two point mutations (n=3); was an essential splice site variant (n=3); or a truncating variant (n=5). A further 32 genes were added to the panel including commonly mutated genes in the RAF/MEK/ERK signalling pathway, genes previously reported over-expressed/mutated in LGSOC and commonly mutated in other ovarian cancer histotypes [19]. In total 127 genes were selected (**supplementary material, Table S2**) for full exon sequencing.

### Mutation landscape

Targeted sequencing was successfully performed on 71 LGSOC cases. Of the 127 genes sequenced, 84 harboured putative somatic loss-of-function (LoF) or non-synonymous mutations in one (n=38) or multiple cases (n=47), with a median mutation frequency of 3 (range 0 to 21) and mean of 3.2 ( $\pm$  2.8 SD, **Figure 1A**). All somatic mutations observed are shown in **supplementary material, Table S4** and recurrent mutations ( $\geq 2$ ) visualised in **Figure 1B**. Forty-one of the genes with somatic mutations

are classified as known somatic drivers in either the COSMIC and IntOGen databases. Additionally, 31 mutated genes (48 variants) displayed loss-of-heterozygosity (LOH) through copy-neutral LOH, and 8 mutated genes (12 variants) through copy loss of the wild-type allele.

Consistent with the frequency of reported variants in the literature [5,7,8], mutations in RAS/RAF signalling genes were common, including *KRAS* (26.7%), *BRAF* (12.6%), and *NRAS* (8.5%). Hot spot mutations in these genes were mutually exclusive of each other, consistent with previous findings [20], with 1 case (c1466) containing non-hotspot drivers for both *KRAS* (p.Gln61His) and *BRAF* (p.Gly596Arg). Variants in all RAS/RAF signalling linked genes (**supplementary material, Table S5**) were observed collectively in 57.7% (41/71) of cases. Validating previous findings, 4/6 *NRAS* Q61\* mutant cases co-occurred with missense *EIF1AX* mutations [7]. The high frequency of *USP9X* somatic point mutations (11/71, 15.5%) is consistent with the results of previous smaller studies [5,7].

Known driver genes in other cancer types and identified in our LGSOC cases included *MACF1* (11.2%), *ARID1A* (9.9%), *NF2* (4.2%), *DOT1L* (5.6%) and *ASH1L* (4.2%). Three of the four *DOT1L* point mutations co-occurred with *ASH1L* mutations (no co-occurring mutations were seen in the WES data from the 38 samples). IHC evaluation of *CDKN2A* revealed recurrent loss of expression (9.9%) and overexpression (5.6%) consistent with previous studies [5,21].

To evaluate whether key aberrations were associated with poor outcome, we tested associations with disease specific survival (DSS) (**supplementary material, Figure S1**). There were no significant differences in outcome comparing stage at diagnosis or surgical debulking status. Cases that were either *USP9X* mutation positive, *KRAS/BRAF/NRAS* mutation positive or *KRAS/BRAF/NRAS/USP9X* mutation negative showed no significant difference in DSS between the three groups. Interestingly, both CDKN2A loss and overexpression were associated with worse DSS when compared to normal CDKN2A expression (HR 3.64, 95% CI 0.84-15.8 and HR 3.43, 95% CI 0.44-26.8 respectively). CDKN2A IHC was performed and expression quantitated on an additional 13 LGSOC obtained through the AOCS [7]. This analysis when combined with the COEUR cohort showed both CDKN2A loss (n=12) and overexpression (n=6) were still associated with worse DSS when compared to normal CDKN2A expression (n=70, HR 3.30, 95% CI 1.18–9.17 and HR 3.54, 95% CI 0.71–17.60 respectively).

To determine if aberrations were associated with shorter survival, cases with survival below the median survival rate in our cohort (66 months) were compared to cases with survival above the median rate (**Table 1**), showing significant differences associated with aberrant CDKN2A (31% versus 0%, p=0.0148). When this data was combined with the extended AOCS IHC cases, aberrant CDKN2A was still significantly enriched in cases where the survival rate was below the median compared to survivors above the median (39% versus 0%, p=0.0020).

To investigate potential cancer signalling pathways involved in LGSOC pathogenesis, annotation of genes by their signalling pathways was performed utilising

Reactome Pathway Analysis [22] (**supplementary material, Table S5**). Recurrently affected pathways ( $\geq 2\%$ ) are shown in **Figure 2A**. Fifteen pathways seen in  $>5\%$  of the LGSOC cohort involved USP9X. The remaining top affected pathways included RAS signalling (22%), FGFR signalling (15%), MAPK signalling (15%), ErbB4 signalling (13%), chromatin organisation (10%) and ubiquitination (10%). Further pathway analysis was performed by grouping aberrations according to the overlapping pathways recorded in **supplementary material, Table S5** as shown in **Figure 2B**. This identified a clustering of 25 genes into 6 biological processes: chromatin regulation (which includes chromatin organization and chromatin modifying enzymes nodes), RAS/RAF/MAPK signalling (which includes RAF activation, negative regulation of MAPK pathway and oncogenic MAPK signalling nodes), FGFR signalling (which includes signalling by FGFR3 and signalling by FGFR4 nodes), WNT activation, TP53 regulation and diseases of signal transduction.

### **USP9X is a relevant therapeutic target for LGSOC**

In this cohort, *USP9X* somatic mutations (11/71, 15.5%) and CN loss (8/71, 11.2%) were collectively observed at the same frequency as *KRAS* mutations (26.7%) (**Figure 1**). The locations of the somatic mutations relative to known domains of *USP9X* are shown in **Figure 3A**. Eight LoF mutations were identified in 8 women, and 4 missense variants observed in 3 women (c988, c1323 and c17), of which 2 out of 4 variants were present within known USP9X protein domains (**Figure 3A**). Multiple computational tools (Condel, PolyPhen, SIFT, CADD and REVEL) [23-27] were used to assess the potential pathogenicity of the missense variants with two of the four being deemed pathogenic by

two or more *in silico* tools (**Table 2**). Two missense mutations were not predicted to be pathogenic, although both variants occurred in the same patient (c988) and one of the variants was present within the UBL domain which is critical for its localization at the proteasome (**Table 2**).

The cases with *USP9X* somatic mutations were further explored to determine if they had LOH. The variant allele read proportion of *USP9X* (adjusted for tumour purity) was used as a measure of allelic status, together with the allelic status of other germline single nucleotide polymorphisms (SNPs) detected across chromosome X and copy number data derived using off-target reads (**Table 2 and supplementary material, Figure S2A–K**). Eight tumours with *USP9X* mutations showed no CN loss at the locus while three showed evidence of loss of the wild-type *USP9X* allele (cases c509, c17 and c761). On this basis an additional 8 LGSOC cases without *USP9X* mutations were identified with CN LOH across the locus (**supplementary material, Figure S2L–S**).

To investigate *USP9X* expression, IHC was performed on 61 cases from the COEUR cohort that included all 11 *USP9X* mutant cancers; 6/8 *USP9X* wild-type cancers with chromosome Xq loss CN loss; and 44/51 cases with wild-type *USP9X* and no CN loss (**Figure 3B and supplementary material, Table S6**). As expected, the 2 cases with a LoF mutation and CN LOH showed null expression by IHC (**Figure 3B and supplementary material, Table S6**). *USP9X* expression was retained within the internal stroma and lymphocytes and completely absent in the tumour. The heterozygous missense cases all retained wild-type levels of expression. Interestingly, the remaining heterozygous LoF cases were also *USP9X* null. Within the sensitivity of IHC there

appears to be no reduced expression in the *USP9X* missense mutated case with copy loss of the wild-type allele, and those cases with single copy *USP9X* loss with no mutation.

To assess if *USP9X* promoter methylation fits the two-hit model of tumour-suppressor gene (TSG) inactivation, methylation of the *USP9X* CpG island region (spanning the proximal promoter into intron 1) was assessed for 6 cases where sufficient quantity and quality of DNA were available; 3 *USP9X* heterozygous mutants, 1 *USP9X* mutant with copy number LOH, and 2 *USP9X* wild-type cases. Neither the heterozygous nor the null *USP9X* tumours showed any evidence of CpG island hypermethylation (**supplementary material, Figure S3**).

Given that *USP9X* is a large gene and that matching germline DNA was unavailable, we cannot exclude the possibility that some of the missense variants are private germline variants. The LoF variants are unlikely to be germline given these are extremely rare in the general population with only three such variants reported in the entire Gnomad database (total LoF allele frequency =  $5.81 \times 10^{-06}$ ). The uncorrected allele frequencies of the missense mutations are all  $<0.3$ , which strongly suggests they are only present in the tumour. In contrast, common heterozygous germline SNPs in *USP9X* are all present at allele frequencies approximating 0.5 in these three samples. In addition, comparing the frequency of CN changes between the 8 LoF and 3 missense *USP9X* mutant carcinomas shows concordant CN events such as 8q gain and Xq loss (**supplementary material, Figure S4**), strengthening the case that these missense variants are somatic.

## Copy number (CN) analysis

Off-target sequencing reads from the sequencing panel were utilised to generate genome-wide CN data for all 71 LGSOC. Sixty five out of 71 LGSOC cases showed one or more copy number changes (**Figure 4A**). Of the 6 cases with no copy number changes, all harboured detectable somatic driver mutations which confirmed the lack of CN change was not due to low tumour DNA purity. The most frequent CN aberrations involved gains on chromosome 1q (48.4%), 8p (32.4%), 8q (32.4%), 12p (38%), 12q (32.4%); and losses on 1p (33.8%), 11p (52.1%), 16p (49.3%), 22q (32.9%) and Xq (31%). All CN gains and losses with an aggregate frequency cut off of  $\geq 25\%$ , including known cancer genes within regions are shown in **supplementary material, Table S7**.

To assess if LGSOC differed in their CN profiles depending on the driver mutation, the cases were divided into 6 sub-groups; *BRAF* mutant (**Figure 4B**, n=9); *NRAS* mutant (**Figure 4C**, n=6); *KRAS* mutant (**Figure 4D**, n=19); *USP9X* mutant negative with *USP9X* CN LOH (**Figure 4E**, n=8); *USP9X* mutant (**Figure 4F**, n=11); and those cases that were *KRAS/NRAS/BRAF/USP9X* aberrant negative with known somatic driver mutations (identified by either COSMIC or IntOGen) outside these genes (**Figure 4G**, n=19). Firstly, the frequency of CN gains on chromosome 12p was higher in *KRAS* mutant cases (11/19, 58%), and *USP9X* CN loss cases, which included 3 that were *KRAS* mutant positive (4/8, 50%), compared to the rest of the cohort (9/45, 20%) (threshold for significance is  $p < 0.05$ , greater than 25% CN frequency change between cohorts). Secondly, CN loss on chromosome Xq was higher in *USP9X* mutant cases (14/19, 73%) compared to those cases that were *USP9X* wild-type (10/52, 19%). Thirdly, CN gains on

chromosome 8q were enriched in *KRAS/BRAF/NRAS/USP9X* wild-type cases (8/19, 42%) and *USP9X* aberrant cases (15/19, 79%) compared to *KRAS/BRAF/NRAS* mutant cases (7/34, 21%).

The median HRD score of the LGSOC cohort was low at 3 (range 0 to 48), but three cases had a clinically high HRD score  $\geq 42$  (**Figure 4H**). These three cases were therefore sequenced for mutations in known *HBOC* predisposition genes and DNA-repair genes. One of those cases (c1483) with an HRD score of 43 harboured a *BRCA2* frameshift variant (p.T2722Nfs\*8) accompanied by copy neutral LOH that is likely contributing to the high HRD observed (**supplementary material, Table S4**). Also present was high-level copy amplification of *MDM2*, which has been shown to result in a loss of p53-dependent activity in *TP53* wild-type HGSOC [28]. Whilst histopathology re-revision confirmed this sample as a LGSOC, it may be on the path to becoming more HGSOC-like. Another case (c197) with an HRD score of 48 carried an *MSH3* p.A57P variant (in the heterozygous state), that is of uncertain clinical significance and would not explain the high HRD observed. The last case (c1424) with a HRD score of 42 did not carry any detectable mutations in *HBOC* or HRD genes.

In order to investigate molecular subtypes within LGSOC, unsupervised clustering was performed using somatic mutation and CN data. This delineated 8 distinct clusters (**Figure 5**): Cluster 1 – high FGA, *KRAS* mutant, chromosome 1p loss and 1q, 12p/q gains; 17q and 18q gains; Cluster 2 – poor outcome, high FGA, *KRAS* and *USP9X* aberrant, chromosome 1p and Xq loss, with 1q, 3q, 7q, 8p/q, 12p/q, and 17q gains; Cluster 3 – *USP9X* mutant, chromosome 1p loss and 8p/q gains; Cluster 4 – *BRAF*

mutant and chromosome 7q gain; Cluster 5 – PR positive, *KRAS* mutant, chromosome 1p loss and 7q, 8q and 12p gains; Cluster 6 – poor outcome, PR positive, *NRAS* mutant, chromosome 1p loss, with 1q and 7q gains; Cluster 7 – poor outcome, high FGA, suboptimal surgery, *KRAS/BRAF/NRAS/USP9X* wild-type, somatic driver positive, chromosome 1p loss, with 1q and 8p/q gains; and Cluster 8 – poor outcome, PR positive, *KRAS/BRAF/NRAS/USP9X* wild-type, somatic driver positive.

A high FGA (here defined as  $\geq 10\%$ ) was not significantly associated with worse DSS when compared to FGA cases  $< 10\%$  (HR 1.75, 95% confidence interval (CI) 0.91-3.35; **supplementary material, Figure S1**). Grouping FGA based on quintiles found no significant difference comparing patients in quintile 1 to quintile 5 (HR 1.09, 95% CI 0.39-3.02; **supplementary material, Figure S1**).

#### **Assessing the clinical value of genomic findings.**

The clinical utility of individual mutant genes, CN alterations and ER/PR status detected in each sample was systematically evaluated. Each feature was curated into tiers of clinical actionability according to the OncoKB knowledge base of oncogenic effects and treatment implications [29] where levels 1-4 represent the level of evidence that the biomarker is predictive of response to a drug or predictive of drug resistance (R1 to R2) (**Figure 6A**). In addition to the OncoKB scoring, each LGSOC was scored for mutations in known oncogenes/tumour suppressor genes or genes not identified as either a tumour suppressor or oncogene (the number of mutations for these categories is shown for each LGSOC case), where compelling targeted therapeutic strategies have

yet to be developed (**Figure 6A**). The percentage of total aberrations observed for each level of treatment evidence is shown (**Figure 6B**).

## Discussion

In keeping with previous findings, mutually exclusive mutations in *KRAS/BRAF/NRAS* dominated the mutation landscape (46.5%) [20]. This predominance has led to clinical trials evaluating *MAPK/ERK* kinase inhibitor (MEKi) activity as a potential therapeutic. The MEKi selumetinib has shown low activity in patients with LGSOC, and response did not correlate with *KRAS/BRAF* mutation status [30]. Indeed, clinical response to the *BRAF* inhibitor dabrafenib has been demonstrated thus far in two women with *BRAF*<sup>V600E</sup> mutated LGSOC [31]. Even though the above trials are small, the underlying dogma of targeted treatments is that response will be strongly correlated with mutation status. However, response in a small subset of patients likely suggests that in the *RAS/RAF* mutation carriers there are other pathway alterations which can bypass the dependency on *RAS/RAF*. Indeed, 52% of the *RAS/RAF* mutated cases in our cohort harboured somatic alterations not involved in *RAS/RAF* signalling.

Outside of *RAS/RAF* signalling, mutations in *MACF1*, linked to activated WNT and MAPK signalling [32,33], were observed in 11.2% of cases. Given *MACF1* contains ~102 exons and spans over 270 kbp we cannot be certain the variants observed are not private germline variants. However, 5 out the 8 missense variants did show copy neutral LOH of the wild-type allele, strengthening this as a likely somatic driver.

The observation of co-occurring somatic mutations in *ASH1L* and *DOT1L* supports the idea that perturbation of genes involved in chromatin organisation and histone methylation defines a biologically distinct subset of LGSOC. While the number of

cases is small, data from all cancer genomes in the cBioPortal also show a significant tendency towards somatic co-occurrence ( $\log_2$  odds ratio = 1.648,  $p < 0.001$ ,  $q < 0.001$ ) [34,35]. DOT1L is the sole methyltransferase responsible for all three forms of H3K79 methylation, and somatic mutations are found in HGSOC, whereby *DOT1L* depletion in HGSOC cell lines promoted cell invasion and cancer stem-like cell properties [36]. ASH1L is a H3K4 methyltransferase, and along with DOT1L are coupled with the “on” state of transcription [37]. Mutations in *ARID1A* were also collectively observed 8.5% of cases, with only 1 case overlapping with a co-occurring *ASH1L/DOT1L* mutant cancer. LGSOC patients with mutations of histone methylation modifiers may benefit from epigenetic modifiers that are currently being explored in *ARID1A* mutant ovarian clear cell carcinoma trials [38].

Since *CDKN2A* aberrations are enriched in LGSOC cases with shorter survival, targeting *CDKN2A* represents a promising avenue for therapeutic intervention to improve outcomes for these patients. Two studies have demonstrated exceptional response to Palbociclib in a patient with refractory uterine leiomyosarcomas [39] and metastatic collecting duct carcinoma harbouring *CDKN2A* deletion [40]. However, complete/partial recession was not seen in other *CDKN2A*-null tumours, such as melanoma [41]. Both loss and overexpression have been discovered in several carcinomas, including LGSOC, and are both linked to shorter survival outcome [42], which was observed in our cohort. Mechanisms to target *CDKN2A*-overexpressing cancer cells are currently under investigation employing an inducible suicide gene regulated by the p16 promoter [43] and warrants investigation in LGSOC-derived cell lines [44].

Ubiquitin-specific protease 9X (USP9X) is an X-linked deubiquitinase that plays a major role in tissue homeostasis, and dysregulation is observed in multiple cancer-types [45]. Within our dataset 7/12 mutations observed were frameshift variants with null expression, indicating that *USP9X* represents a classical “two-hit” TSG. Additionally, 2/4 missense variants clustered within known USP9X functional domains and 2/4 were predicted to be pathogenic using *in silico* tools. While three cases showed biallelic genetic loss of *USP9X*, a further 8 *USP9X* mutant cases retained the wild-type allele. We showed that in these heterozygous cases the retained wild-type allele was not expressed in the tumour and this was not attributable to promoter hyper-methylation. This is consistent with studies showing that *USP9X* frequently escapes X-inactivation across multiple female tissues [46], including solid cancers [47]. Overall, these data indicate that the wild-type *USP9X* allele in heterozygous mutant cases is silenced by other mechanisms and provides clear evidence that USP9X functions as a classic TSG requiring bi-allelic inactivation.

The CN profiles of the remaining LGSOC cases without *USP9X* mutations were investigated and identified 8 cases with CN loss at the locus that were genomically similar to *USP9X* mutant cases regardless of expression changes, suggesting that the frequency of LGSOC with *USP9X* aberrations may be higher than observed. Additionally, a case of *USP9X* complex structural rearrangement has been previously detected in LGSOC [7]; unfortunately we could not detect similar structural perturbation due to our sequencing platform. *USP9X* appears to be an important driver in LGSOC progression as mutations were identified in only 2.6% of SBTs [5] and not seen in other ovarian

carcinoma subtypes [48,49]. A major challenge remains in understanding the molecular context of USP9X-regulated processes in LGSOC, given USP9X governs multiple signalling pathways with varied cellular responses [reviewed in 45].

Hierarchical clustering revealed 2 of 8 subgroups are enriched for poor outcomes and are *KRAS/BRAF/NRAS/USP9X* wild-type that either have a high FGA (cluster 7) or are largely progesterone receptor positive (Cluster 8). Furthermore, 2 of 3 clusters with a high FGA (clusters 2 and 7) were enriched for cases that died from LGSOC, although a high FGA was not significantly associated with disease-specific survival, likely due to the small number of cases in this analysis. Within these two clusters the shared CN changes were 1p loss with 1q and 8p/q gains, and whether individual cancer genes or gene-dosage alterations of all cancer genes within these CN regions contribute towards LGSOC disease remains unknown.

The three cases with a high HRD represent a small group with either somatic or germline DNA repair defects. HRD scores  $\geq 42$  can predict the likelihood of response to neo-adjuvant PARP inhibition, which has now extended to HGSOC cases yielding improvements in survival [50]. It is possible that NGS-based assays that detect HRD in LGSOC, beyond *BRCA1/2* mutational status, may predict clinical benefit but further research is required.

Overall, 90% of LGSOC harboured at least one potentially actionable alteration, of which in 19/71 cases were predictive of clinical benefit from a standard-of-care treatment, either in another cancer indication (11/71) or in LGSOC specifically (9/10,

*BRAF* mutant cases). LGSOCs that were ER receptor and/or PR receptor positive were the highest recurrent actionable target in our cohort, making up ~85% of cases where compelling clinical evidence supports these biomarkers as indicative of response in another indication. Importantly though, in a retrospective analysis of LGSOC the objective response rate to endocrine therapy was only 9% [51], and in a phase-II prospective trial, the response rate was 14% [52]. Combination therapy strategies should be considered in these cases as *MAPK* activation can crosstalk with ER $\alpha$  and experimentally drive endocrine resistance in ovarian cancer models [53].

While our understanding of LGSOC has expanded significantly over the past decade, the experience with targeted therapy for these rare histologic subtypes is still limited. With multiple clinical trials targeting specific molecular targets on-going, women diagnosed with late-stage LGSOC would greatly benefit from trials that performed DNA sequencing to identify the most suitable targeted therapy. Such trials should be targeted to either single or combination aberrations with pre-clinical efficacy, contrasting with the current practice of administering platinum-based chemotherapy, which provides little benefit and considerable patient toxicity. There is tremendous potential for progress in treating LGSOC by leveraging our genomic finding and translating this understanding into improved therapeutic paths.

## Acknowledgements

The authors would like to acknowledge the Bioinformatics and Molecular Genomics Core Facilities of the Peter MacCallum Cancer Centre, which are supported by the Australian Cancer Research Foundation. This study also used resources provided by the Canadian Ovarian Cancer Research Consortium's COEUR biobank funded by the Terry Fox Research Institute (grant #2009-15) and supervised by the CHUM. The Consortium acknowledges contributions to its COEUR biobank from institutions across Canada (for a full list see [http://www.tfri.ca/en/research/translational-research/coeur/coeur\\_biobanks.aspx](http://www.tfri.ca/en/research/translational-research/coeur/coeur_biobanks.aspx)).

This work was supported by the Australian Gynaecological Cancer Foundation and Way In Network (TP825458, held by DC), the National Health and Medical Research Council (APP1092856 & APP1041975, held by IC), Cancer Australia (APP1142697, held by AdeF, KG, DE) and the Terry Fox Research Institute (#2009–15). The Australian Ovarian Cancer Study was supported by the U.S. Army Medical Research and Materiel Command under DAMD17-01-1-0729, The Cancer Council Victoria, Queensland Cancer Fund, The Cancer Council New South Wales, The Cancer Council South Australia, The Cancer Foundation of Western Australia, The Cancer Council Tasmania and the NHMRC (ID400413 and ID400281). The Australian Ovarian Cancer Study gratefully acknowledges additional support from S. Boldeman, the Agar family, Ovarian Cancer Australia and Ovarian Cancer Action. The funders had no role in the design and execution of this study.

**Author contributions statement**

MK and PA performed pathology review of cases. CLP prepared tissue samples, performed nucleic acid extraction and PCR. DGH, CLP, AMMM, and DP provided clinical information. DC, SH, DE, MSC, MLF, AdeF, AD, NH, DB, KLG and IGC contributed toward the design of sequencing panel. DC, MZ, SH and DE performed bioinformatics analyses. DC, AN and TS performed sequencing library preparations. DC, KLG and IGC coordinated the study. DC, AN and YG analysed the data. DC, KLG and IGC designed the study and were involved at all stages. DC prepared the figures and drafted the manuscript, which was then extensively edited by KLG, IGC and AN. All remaining authors contributed to the final draft of the manuscript.

**Data availability statement**

The datasets analysed in the current study are available from the corresponding author, with requests for access subject to review by the COEUR Study Committee.

## References

1. Ali RH, Kalloger SE, Santos JL, *et al.* Stage II to IV low-grade serous carcinoma of the ovary is associated with a poor prognosis: a clinicopathologic study of 32 patients from a population-based tumor registry. *Int J Gynecol Pathol* 2013; **32**: 529-535.
2. Le Page C, Rahimi K, Kobel M, *et al.* Characteristics and outcome of the COEUR Canadian validation cohort for ovarian cancer biomarkers. *BMC Cancer* 2018; **18**: 347.
3. Chen M, Jin Y, Bi Y, *et al.* A survival analysis comparing women with ovarian low-grade serous carcinoma to those with high-grade histology. *Onco Targets Ther* 2014; **7**: 1891-1899.
4. Schmeler KM, Sun CC, Bodurka DC, *et al.* Neoadjuvant chemotherapy for low-grade serous carcinoma of the ovary or peritoneum. *Gynecol Oncol* 2008; **108**: 510-514.
5. Hunter SM, Anglesio MS, Ryland GL, *et al.* Molecular profiling of low grade serous ovarian tumours identifies novel candidate driver genes. *Oncotarget* 2015; **6**: 37663-37677.
6. Iijima M, Banno K, Okawa R, *et al.* Genome-wide analysis of gynecologic cancer: The Cancer Genome Atlas in ovarian and endometrial cancer. *Oncol Lett* 2017; **13**: 1063-1070.
7. Etemadmoghadam D, Azar WJ, Lei Y, *et al.* EIF1AX and NRAS Mutations Co-occur and Cooperate in Low-Grade Serous Ovarian Carcinomas. *Cancer Res* 2017; **77**: 4268-4278.
8. Jones S, Wang T-L, Kurman RJ, *et al.* Low-grade serous carcinomas of the ovary contain very few point mutations. *The Journal of pathology* 2012; **226**: 413-420.
9. Emmanuel C, Chiew YE, George J, *et al.* Genomic classification of serous ovarian cancer with adjacent borderline differentiates RAS pathway and TP53-mutant tumors and identifies NRAS as an oncogenic driver. *Clin Cancer Res* 2014; **20**: 6618-6630.
10. Le Page C, Kobel M, de Ladurantaye M, *et al.* Specimen quality evaluation in Canadian biobanks participating in the COEUR repository. *Biopreserv Biobank* 2013; **11**: 83-93.
11. Van der Auwera GA, Carneiro MO, Hartl C, *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013; **43**: 11 10 11-33.
12. Rimmer A, Phan H, Mathieson I, *et al.* Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet* 2014; **46**: 912-918.
13. Koboldt DC, Zhang Q, Larson DE, *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012; **22**: 568-576.
14. Scheinin I, Sie D, Bengtsson H, *et al.* DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res* 2014; **24**: 2022-2032.
15. Li N, Rowley SM, Thompson ER, *et al.* Evaluating the breast cancer predisposition role of rare variants in genes associated with low-penetrance breast cancer risk SNPs. *Breast Cancer Res* 2018; **20**: 3.

16. Rowley SM, Mascarenhas L, Devereux L, *et al.* Population-based genetic testing of asymptomatic women for breast and ovarian cancer susceptibility. *Genet Med* 2018.
17. Kuilman T, Velds A, Kemper K, *et al.* CopywriteR: DNA copy number detection from off-target sequence data. *Genome Biol* 2015; **16**: 49.
18. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011; **474**: 609-615.
19. Emmanuel C, Gava N, Kennedy C, *et al.* Comparison of expression profiles in ovarian epithelium in vivo and ovarian cancer identifies novel candidate genes involved in disease pathogenesis. *PLoS One* 2011; **6**: e17617.
20. Nakamura K, Nakayama K, Ishibashi T, *et al.* KRAS/BRAF Analysis in Ovarian Low-Grade Serous Carcinoma Having Synchronous All Pathological Precursor Regions. *International journal of molecular sciences* 2016; **17**: 625.
21. Rambau PF, Vierkant RA, Intermaggio MP, *et al.* Association of p16 expression with prognosis varies across ovarian carcinoma histotypes: an Ovarian Tumor Tissue Analysis consortium study. *The Journal of Pathology: Clinical Research* 2018; **4**: 250-261.
22. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst* 2016; **12**: 477-479.
23. González-Pérez A, López-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet* 2011; **88**: 440-449.
24. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics* 2013; **Chapter 7**: Unit7.20.
25. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols* 2009; **4**: 1073-1081.
26. Kircher M, Witten DM, Jain P, *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nature genetics* 2014; **46**: 310-315.
27. Ioannidis NM, Rothstein JH, Pejaver V, *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet* 2016; **99**: 877-885.
28. Ahmed AA, Etemadmoghadam D, Temple J, *et al.* Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *The Journal of pathology* 2010; **221**: 49-56.
29. Chakravarty D, Gao J, Phillips SM, *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol* 2017; **2017**.
30. Farley J, Brady WE, Vathipadiekal V, *et al.* Selumetinib in women with recurrent low-grade serous carcinoma of the ovary or peritoneum: an open-label, single-arm, phase 2 study. *Lancet Oncol* 2013; **14**: 134-140.
31. Moujaber T, Etemadmoghadam D, Kennedy CJ, *et al.* BRAF Mutations in Low-Grade Serous Ovarian Cancer and Response to BRAF Inhibition. *JCO Precision Oncology* 2018: 1-14.
32. Afghani N, Mehta T, Wang J, *et al.* Microtubule actin cross-linking factor 1, a novel target in glioblastoma. *Int J Oncol* 2017; **50**: 310-316.

33. Jain P, Silva A, Han HJ, *et al.* Overcoming resistance to single-agent therapy for oncogenic BRAF gene fusions via combinatorial targeting of MAPK and PI3K/mTOR signaling pathways. *Oncotarget* 2017; **8**: 84697-84713.
34. Cerami E, Gao J, Dogrusoz U, *et al.* The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* 2012; **2**: 401-404.
35. Gao J, Aksoy BA, Dogrusoz U, *et al.* Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Science Signaling* 2013; **6**: p11-p11.
36. Gao J, Aksoy BA, Dogrusoz U, *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013; **6**: p11.
37. Steger DJ, Lefterova MI, Ying L, *et al.* DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Molecular and cellular biology* 2008; **28**: 2825-2839.
38. Berns K, Caumanns JJ, Hijmans EM, *et al.* ARID1A mutation sensitizes most ovarian clear cell carcinomas to BET inhibitors. *Oncogene* 2018; **37**: 4611-4625.
39. Elvin JA, Gay LM, Ort R, *et al.* Clinical Benefit in Response to Palbociclib Treatment in Refractory Uterine Leiomyosarcomas with a Common CDKN2A Alteration. *Oncologist* 2017; **22**: 416-421.
40. Pal SK, Ali SM, Ross J, *et al.* Exceptional Response to Palbociclib in Metastatic Collecting Duct Carcinoma Bearing a CDKN2A Homozygous Deletion. *JCO Precision Oncology* 2017: 1-5.
41. Tang B, Sheng X, Kong Y, *et al.* Palbociclib for treatment of metastatic melanoma with copy number variations of CDK4 pathway: case report. *Chinese Clinical Oncology* 2018; **7**.
42. Rambau PF, Vierkant RA, Intermaggio MP, *et al.* Association of p16 expression with prognosis varies across ovarian carcinoma histotypes: an Ovarian Tumor Tissue Analysis consortium study. *J Pathol Clin Res* 2018; **4**: 250-261.
43. Kohli J, Campisi J, Demaria M. A novel suicide gene therapy for the treatment of p16(Ink4a)-overexpressing tumors. *Oncotarget* 2018; **9**: 7274-7281.
44. Fernandez ML, Dawson A, Hoenisch J, *et al.* Markers of MEK inhibitor resistance in low-grade serous ovarian cancer: EGFR is a potential therapeutic target. *Cancer cell international* 2019; **19**: 10-10.
45. Murtaza M, Jolly LA, Gecz J, *et al.* La FAM fatale: USP9X in development and disease. *Cell Mol Life Sci* 2015; **72**: 2075-2089.
46. Berletch JB, Ma W, Yang F, *et al.* Escape from X inactivation varies in mouse tissues. *PLoS genetics* 2015; **11**: e1005079-e1005079.
47. The EPC, Dunham I, Kundaje A, *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; **489**: 57.
48. Murakami R, Matsumura N, Brown JB, *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.
49. Cheasley D, Wakefield MJ, Ryland GL, *et al.* The molecular origin and taxonomy of mucinous ovarian carcinoma. *Nat Commun* 2019; **10**: 3935.

50. Morgan RD, Clamp AR, Evans DGR, *et al.* PARP inhibitors in platinum-sensitive high-grade serous ovarian cancer. *Cancer Chemother Pharmacol* 2018; **81**: 647-658.
51. Gershenson DM, Bodurka DC, Coleman RL, *et al.* Hormonal Maintenance Therapy for Women With Low-Grade Serous Cancer of the Ovary or Peritoneum. *J Clin Oncol* 2017; **35**: 1103-1111.
52. Tang M, O'Connell RL, Amant F, *et al.* PARAGON: A Phase II study of anastrozole in patients with estrogen receptor-positive recurrent/metastatic low-grade ovarian cancers and serous borderline ovarian tumors. *Gynecologic Oncology* 2019; **154**: 531-538.
53. Hou JY, Rodriguez-Gabin A, Samaraweera L, *et al.* Exploiting MEK inhibitor-mediated activation of ER $\alpha$  for therapeutic intervention in ER-positive ovarian carcinoma. *PloS one* 2013; **8**: e54103-e54103.
54. Gershenson DM, Bodurka DC, Lu KH, *et al.* Impact of Age and Primary Disease Site on Outcome in Women With Low-Grade Serous Carcinoma of the Ovary or Peritoneum: Results of a Large Single-Institution Registry of a Rare Tumor. *J Clin Oncol* 2015; **33**: 2675-2682.
55. Cheasley D, Li N, Rowley SM, *et al.* Molecular comparison of interval and screen-detected breast cancers. *J Pathol* 2019.
56. Köbel M, Rahimi K, Rambau PF, *et al.* An Immunohistochemical Algorithm for Ovarian Carcinoma Typing. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* 2016; **35**: 430-441.
57. Ayabe A, Miyazaki K, Shimizu D, *et al.* The LAST Guidelines in Clinical Practice: Implementing Recommendations for p16 Use. *American Journal of Clinical Pathology* 2015; **144**: 844-849.
58. Marquard AM, Eklund AC, Joshi T, *et al.* Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. *Biomark Res* 2015; **3**: 9.
59. Birkbak NJ, Wang ZC, Kim JY, *et al.* Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. *Cancer Discov* 2012; **2**: 366-375.
60. Popova T, Manie E, Rieunier G, *et al.* Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res* 2012; **72**: 5454-5462.
61. Abkevich V, Timms KM, Hennessy BT, *et al.* Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br J Cancer* 2012; **107**: 1776-1782.
62. Burrell RA, McClelland SE, Endesfelder D, *et al.* Replication stress links structural and numerical cancer chromosomal instability. *Nature* 2013; **494**: 492-496.
63. Riester M, Singh AP, Brannon AR, *et al.* PureCN: copy number calling and SNV classification using targeted short read sequencing. *Source Code for Biology and Medicine* 2016; **11**: 13.
64. Thompson ER, Rowley SM, Li N, *et al.* Panel Testing for Familial Breast Cancer: Calibrating the Tension Between Research and Clinical Care. *J Clin Oncol* 2016; **34**: 1455-1459.
65. Kent WJ, Sugnet CW, Furey TS, *et al.* The human genome browser at UCSC. *Genome research* 2002; **12**: 996-1006.

References 55 – 65 are cited only in the supplementary material.

Gene symbol	Overall survival (median survival = 66 months)						P	Cosmic somatic driver	InTOgen somatic driver	Tumour suppressor gene	Oncogene
	DOD below the median (n=26)			Survived above the median (n=16)							
	Lof	MS	%	Lof	MS	%					
<i>SYNE1</i>	-	5	19%	-	-	0%	0.0648				
<i>ABL1</i>	-	3	12%	-	-	0%	0.1636	Y	Y		Y
<i>CSMD1</i>	-	3	12%	-	-	0%	0.1636			Y	
<i>FRAS1</i>	-	3	12%	-	-	0%	0.1636				
<i>KRAS</i>	-	5	19%	-	6	38%	0.1963	Y	Y		Y
<i>DNAH10</i>	-	1	4%	-	2	13%	0.2961				
<i>PTPRZ1</i>	-	1	4%	-	2	13%	0.2961				
<i>ARID1A</i>	2	1	12%	1	-	6%	0.5753	Y		Y	
<i>BRAF</i>	-	3	12%	-	1	6%	0.5753	Y	Y		Y
<i>EIF1AX</i>	-	3	12%	-	1	6%	0.5753	Y	Y	Y	
<i>NRAS</i>	-	3	12%	-	1	6%	0.5753	Y	Y		Y
<i>PLEC</i>	-	3	12%	-	1	6%	0.5753				
<i>USP9X</i>	3	1	15%	2	1	19%	0.7789	Y		Y	Y
<i>ASH1L</i>	2	-	8%	-	1	6%	0.8618			Y	
<i>BPTF</i>	-	2	8%	-	1	6%	0.8618				
<i>DOT1L</i>	-	2	8%	-	1	6%	0.8618			Y	Y
<i>MACF1</i>	-	3	12%	-	2	13%	0.9264			Y	Y
	<b>Loss</b>	<b>O/E</b>	<b>%</b>	<b>Loss</b>	<b>O/E</b>	<b>%</b>	<b>P</b>				
<i>CDKN2A</i>	5	3	31%	-	-	0%	<b>0.0148</b>	Y	Y	Y	

Plus extended AOCS cohort	n=36			n=18			<i>P</i>				
	Loss	O/E	%	Loss	O/E	%					
<i>CDKN2A</i>	9	5	39%	-	-	0%	<b>0.002</b>	Y	Y	Y	

**Table 1. Somatic aberration profile comparing long-term and short-term survivors.** The median survival rate for our LGSOC cohort was 66 months. Cases that died of disease below the median were compared to cases that survived above the median. A two-tailed p-value was calculated. CDKN2A status was analysed via immunohistochemistry performed on COEUR tissue microarrays and an extended cohort TMA published by the Australian Ovarian Cancer Study (AOCS). Y, yes; N, no; DOD, died of disease; LoF, loss-of-function variant; MS, missense variant.

Sample ID	<i>USP9X</i> nucleotide position	<i>USP9X</i> protein change	Mutation type	Read depth	Tumour purity estimate	Detected <i>USP9X</i> allele frequency	Allele frequency adjusted for tumour purity	<i>USP9X</i> CN loss	Predicted allelic status <i>USP9X</i>	Mutation within domain	<i>USP9X</i> IHC	Condel	PolyPhen	SIFT	REVEL	CADD
c1632	c.187G>T	p.E63*	Stop-gained	340	~100%	0.5	0.5	No	Heterozygous		Null					
c509	c.1153C>T	p.R385*	Stop-gained	228	40%	0.33	0.83	Yes	Loss of wild type		Null					
c17	c.1355T>G	p.L452R	Missense	330	50%	0.4	0.8	Yes	Loss of wild type		Wild type	N	N	Y	Y	N
c1279	c.1981_1982insT	p.R662*	Frameshift	202	30%	0.13	0.43	No	Heterozygous		Null					
c988	c.2518G>A	p.V840I	Missense	871	30%	0.11	0.37	No	Heterozygous	UBL	Wild type	N	N	N	N	N
c1094	c.3679delG	p.D1227fs*12	Frameshift	216	50%	0.2	0.4	No	Heterozygous		Null					
c761	c.3761_3762insA	p.Q1255Tfs*5	Frameshift	283	60%	0.55	0.92	Yes	Loss of wild type		Null					
c132	c.3867delG	p.E1290fs*2	Frameshift	863	40%	0.19	0.48	No	Heterozygous		Null					
c1323	c.4755T>G	p.I1585M	Missense	260	60%	0.28	0.47	No	Heterozygous	Catalytic CYS	Wild type	Y	Y	N	N	N
c1503	c.5013delC	p.F1671Ffs*17	Frameshift	441	60%	0.3	0.5	No	Heterozygous	Catalytic	Null					

c1531	c.5458C>T	p.R1820*	Stop-gained	1238	92%	0.23	0.25	No	Heterozygous	Catalytic	Null					
c988	c.7675A>G	p.S2559 G	Missense	831	30%	0.16	0.53	No	Heterozygous		Wild type	N	N	N	N	N

**Table 2. USP9X mutation information.** Mutations were scored for copy loss, allelic status, concordant immunohistochemistry expression and predicted to be pathogenic, damaging/deleterious, or tolerated/neutral/benign utilizing five *in-silico* prediction tools. CN, copy number; IHC, immunohistochemistry; Y, Yes; N, No.

## Legends to Figures

**Figure 1. Somatic aberration profile of the low-grade serous ovarian carcinoma cohort.** (A) The total number of coding variants (frameshift/truncating, essential splice and missense) is shown for each of the 71 LGSOCs screened. (B) A matrix of recurrently ( $\geq 2$ ) aberrant (mutated, CN/IHC amplified and loss) genes are shown. Each row represents the gene screened and each column represents a different LGSOC patient sample. The total number and percentage of somatic aberrations in the samples screened are also listed. In addition to variant calling, CDKN2A status was analysed via immunohistochemistry performed on COEUR tissue microarrays. Age at diagnosis  $\leq 35$  years was selected as it is significantly associated with worse outcome [54]. Half shaded squares represent samples with both a mutation and CN gain in that gene. ER, estrogen receptor; PR, progesterone receptor; FGA, fraction of the genome altered.

**Figure 2. Gene annotation of associated cancer signalling pathways.** Pathway annotation was performed utilising the Reactome Pathway Analysis R package [22], (A) recurrently affected pathways ( $\geq 2$ ) are shown as a percentage of total variants observed (223 variants). Additionally, the 15 pathways in which USP9X has been implicated in are highlighted in red. (B) A network of functional interactions is shown. Nodes (orange) represent cancer signalling pathways that functionally interact (grey line) with proteins found aberrant in our cohort (grey dot). The size of the node represents the number of interacting proteins. The six biological processes are highlighted within the blue dotted areas

**Figure 3. USP9X acts as a tumour suppressor gene in the context of low-grade serous ovarian carcinomas.** (A) The position of stop-gained (star), frameshift (circle) and missense mutations (arrow) are mapped to amino acid residues of USP9X. The structure of the human USP9X comprises a ubiquitin-like (UBL) domain, a catalytic ubiquitin specific protease domain (containing two short conserved cysteine and histidine catalytic motifs) and four nuclear localization sequence (NLS) motifs. (B) Immunohistochemical evaluation of USP9X expression was performed on COEUR LGSOC cancers (see supplementary material, Table S6), and representative images are shown for one sample with a USP9X stop-gained mutation + CN loss of wild type allele, two samples with a USP9X heterozygous frameshift mutation, two samples with USP9X CN loss and one sample that was USP9X wild-type. CN, copy number.

**Figure 4. Copy number analysis of low-grade serous ovarian carcinomas.** The frequency of global copy number changes was scored for (A) all 71 LGSOCs compared to LGSOCs that are (B) BRAF mutant; (C) NRAS mutant; (D) KRAS mutant; (E) USP9X mutant; (F) USP9X mutant negative with USP9X copy loss; (G) and cases that are KRAS/NRAS/BRAF/USP9X negative and somatic driver positive. Gains and losses that were statistically significantly enriched in the mutational subgroup are highlighted by black dot boxes. Copy gains are in blue and losses in red. Copy number profiles were used to generate (H) a homologous recombination efficiency sum score. Arrowhead, LGSOC with MSH3 p.A 57P mutation, Arrow, LGSOC with BRCA2 p.T2722Nfs\*8 & MDM2 high copy gain.

**Figure 5. Unsupervised hierarchical cluster analysis of low-grade serous ovarian carcinomas.** Cluster analysis was performed on clinical data, *KRAS*, *BRAF*, *NRAS* and *USP9X* aberrations, and regions of CN gain and loss seen in  $\geq 25\%$  of cases, where known cancer genes fall within the region. 8 clusters were identified and labelled C1 – C8. DOD, died of disease; CN, copy number; PR, progesterone receptor; FGA, fraction of the genome altered; HRD, homologous recombination deficiency.

**Figure 6. Levels of treatment evidence assigned to genomic aberrations observed in the low-grade serous ovarian carcinoma cohort. (A)** Alterations were annotated based on their clinical actionability according to OncoKB, and LGSOC samples were assigned to the level of the most actionable alteration. Levels of evidence varied according to whether mutations are FDA-recognized biomarkers (level 1), predict response to standard-of-care therapies (level 2) or predict response to investigational agents in clinical trials (level 3). Levels 2 and 3 were subdivided according to whether the evidence existed for LGSOC (2A, 3A) or a different tumour type (2B, 3B). Samples were additionally analysed for genes considered as oncogenic and TSG but not actionable, or genes not identified as either an oncogenic or TSG. The distribution of the highest level of actionability across all patients is displayed. **(B)** Pie chart showing the percentage of total aberrations detected based on treatment level. Key adapted from [29]

**Supplementary materials and methods**

**Figure S1.** Combined genomic and clinicopathological analysis

**Figure S2.** Analysis of loss of heterozygosity and allelic imbalance across chromosome X, including USP9X

**Figure S3.** Bisulphite sequencing of LGSOC tumours within the *USP9X* promoter region

**Figure S4.** Copy number analysis of USP9X mutant tumours

**Table S1.** Clinical characteristics of LGSOC cohort

**Table S2.** Inclusion criteria for genes on the targeted LGSOC SureSelect panel

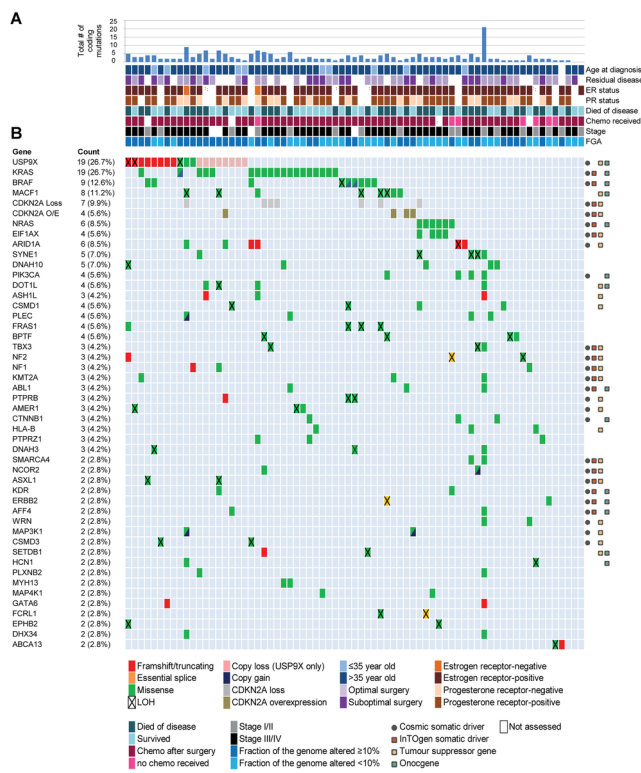
**Table S3.** Sequencing metrics

**Table S4.** Tumour sequencing metrics

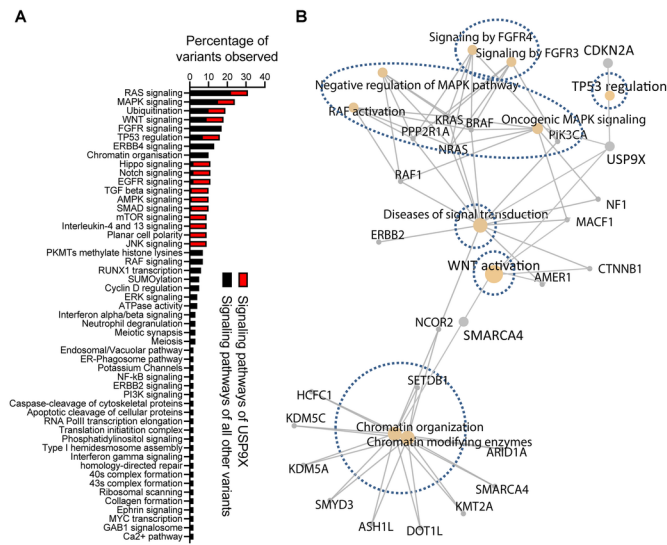
**Table S5.** Variants observed and associated cancer signalling pathways

**Table S6.** USP9X immunohistochemistry and expression evaluation on the COEUR tissue microarray.

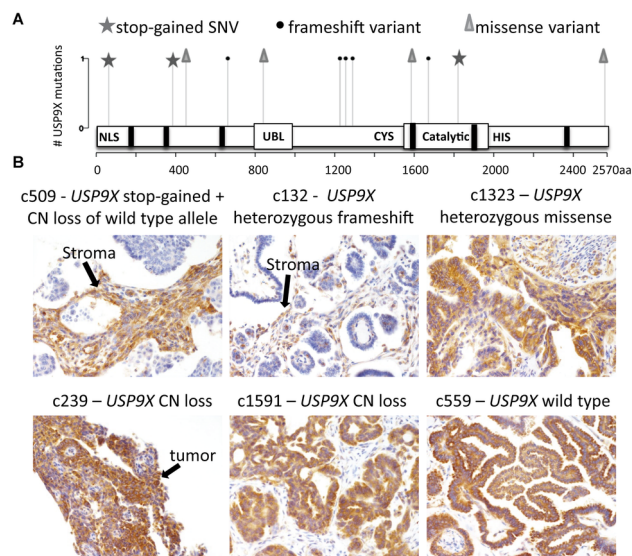
**Table S7.** Significant CN gains and loss across the LGSOC cohort



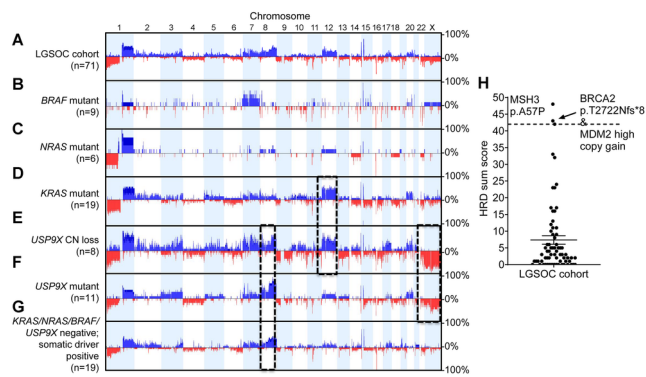
PATH\_5545\_Figure\_1\_finalRP.tif



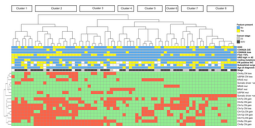
PATH\_5545\_Figure\_2\_finalRP.tif



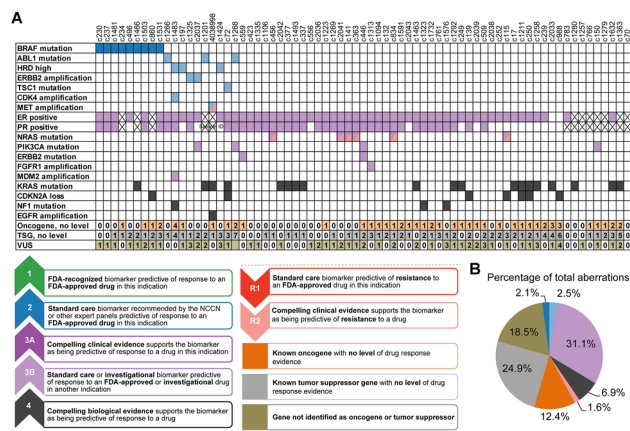
PATH\_5545\_Figure\_3\_finalRP.tif



PATH\_5545\_Figure\_4\_finalRP.tif



PATH\_5545\_Figure\_5\_finalRP.tif



PATH\_5545\_Figure\_6\_finalRP.tif