ARTICLE

doi:10.1038/nature14410

Whole-genome characterization of chemoresistant ovarian cancer

Ann-Marie Patch[1,2]*, Elizabeth L. Christie[3]*, Dariush Etemadmoghadam[3,4,5]*, Dale W. Garsed[3]*, Joshy George[6], Sian Fereday[3], Katia Nones[1,2], Prue Cowin[3], Kathryn Alsop[3], Peter J. Bailey[1,7], Karin S. Kassahn[1,8], Felicity Newell[1], Michael C. J. Quinn[1,2], Stephen Kazakoff[1,2], Kelly Quek[1], Charlotte Wilhelm-Benartzi[9], Ed Curry[9], Huei San Leong[3], The Australian Ovarian Cancer Study Group†, Anne Hamilton[3,10,11], Linda Mileshkin[3,5], George Au-Yeung[3], Catherine Kennedy[12], Jillian Hung[12], Yoke-Eng Chiew[12], Paul Harnett[13], Michael Friedlander[14], Michael Quinn[11], Jan Pyman[11], Stephen Cordner[15], Patricia O'Brien[15], Jodie Leditschke[15], Greg Young[15], Kate Strachan[15], Paul Waring[4], Walid Azar[3], Chris Mitchell[3], Nadia Traficante[3], Joy Hendley[3], Heather Thorne[3], Mark Shackleton[3,5], David K. Miller[1], Gisela Mir Arnau[3], Richard W. Tothill[3,5], Timothy P. Holloway[3], Timothy Semple[3], Ivon Harliwong[1], Craig Nourse[1], Ehsan Nourbakhsh[1], Suzanne Manning[1], Senel Idrisoglu[1], Timothy J. C. Bruxner[1], Angelika N. Christ[1], Barsha Poudel[1], Oliver Holmes[1,2], Matthew Anderson[1], Conrad Leonard[1,2], Andrew Lonie[16], Nathan Hall[17], Scott Wood[1,2], Darrin F. Taylor[1], Qinying Xu[1,2], J. Lynn Fink[1], Nick Waddell[1], Ronny Drapkin[18], Euan Stronach[9], Hani Gabra[9], Robert Brown[9], Andrea Jewell[19], Shivashankar H. Nagaraj[1], Emma Markham[1], Peter J. Wilson[1], Jason Ellul[3], Orla McNally[12], Maria A. Doyle[3], Ravikiran Vedururu[3], Collin Stewart[20], Ernst Lengyel[19], John V. Pearson[1,2], Nicola Waddell[1,2], Anna deFazio[12]§, Sean M. Grimmond[1,7]§ & David D. L. Bowtell[3,4,5,9,21]§

Patients with high-grade serous ovarian cancer (HGSC) have experienced little improvement in overall survival, and standard treatment has not advanced beyond platinum-based combination chemotherapy, during the past 30 years. To understand the drivers of clinical phenotypes better, here we use whole-genome sequencing of tumour and germline DNA samples from 92 patients with primary refractory, resistant, sensitive and matched acquired resistant disease. We show that gene breakage commonly inactivates the tumour suppressors *RB1*, *NF1*, *RAD51B* and *PTEN* in HGSC, and contributes to acquired chemotherapy resistance. *CCNE1* amplification was common in primary resistant and refractory disease. We observed several molecular events associated with acquired resistance, including multiple independent reversions of germline *BRCA1* or *BRCA2* mutations in individual patients, loss of *BRCA1* promoter methylation, an alteration in molecular subtype, and recurrent promoter fusion associated with overexpression of the drug efflux pump MDR1.

Recurrent, chemotherapy-resistant HGSC[1] accounts for most epithelial ovarian cancer deaths. Primary disease is characterized by ubiquitous *TP53* mutation[2,3], extensive copy number change[3–5], and in about half of all HGSCs, there is evidence of mutational[3] and functional[6] inactivation of homologous recombination repair. Extensive intratumoural heterogeneity in primary HGSCs has been recently documented, as it has in other solid cancers[7–9]. However, relatively little is known of the genomic evolution of HGSC under the selective pressure of chemotherapy.

As HGSC is driven by genomic copy number change rather than recurrent point mutation[5], we undertook the first whole-genome sequence (WGS) analysis of a large HGSC cohort to examine structural variation at high resolution. We performed transcriptome, methylation, and microRNA (miRNA) expression analyses to support the WGS data, and compared resistant, refractory and sensitive primary disease. We also evaluated paired primary sensitive and relapse-resistant samples, and end-stage HGSC obtained from patients who underwent rapid autopsy for research.

Driver mutation by gene breakage

We performed whole genome sequencing (40× and 52× average coverage, germline and tumour, respectively; Supplementary Table 1), complemented by transcriptome sequencing (average 210 million reads per sample; Supplementary Table 2), and methylation, copy number and miRNA array analysis on a total of 114 tumour samples from 92 patients (Methods and Extended Data Fig. 1). Confirming previous reports[2,3], *TP53* mutations were prevalent (Fig. 1 and Supplementary Table 3). Inactivating germline or somatic mutations in genes associated with homologous recombination repair, or *BRCA1* methylation, were detected collectively in half of the primary tumours (Fig. 1 and

[1]Queensland Centre for Medical Genomics, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4067, Australia. [2]QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4006, Australia. [3]Peter MacCallum Cancer Centre, East Melbourne, Victoria 3002, Australia. [4]Department of Pathology, University of Melbourne, Parkville, Victoria 3052, Australia. [5]Sir Peter MacCallum Cancer Centre Department of Oncology, University of Melbourne, Parkville, Victoria 3052, Australia. [6]The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut 06030, USA. [7]WolfsonWohl Cancer Research Centre, Institute of Cancer Sciences, University of Glasgow, Glasgow G61 1QH, UK. [8]Technology Advancement Unit, Genetics and Molecular Pathology, SA Pathology, Adelaide, South Australia 5000, Australia. [9]Ovarian Cancer Action Research Centre, Department of Surgery and Cancer, Imperial College London, London W12 0HS, UK. [10]Department of Medicine, University of Melbourne, Parkville, Victoria 3052, Australia. [11]The Royal Women's Hospital, Parkville, Victoria 3052, Australia. [12]Centre for Cancer Research, University of Sydney at Westmead Millennium Institute, and Department of Gynaecological Oncology, Westmead Hospital, Sydney, New South Wales 2145, Australia. [13]Crown Princess Mary Cancer Centre and University of Sydney at Westmead Hospital, Westmead, Sydney, New South Wales 2145, Australia. [14]Prince of Wales Clinical School, University of New South Wales, Sydney, New South Wales 2031, Australia. [15]Victorian Institute of Forensic Medicine, Southbank, Victoria 3006, Australia. [16]Victorian Life Sciences Computation Initiative, Carlton, Victoria 3053, Australia. [17]La Trobe Institute for Molecular Science, Bundoora, Victoria 3083, Australia. [18]Dana-Farber Cancer Institute, Boston, Massachusetts 02115-5450, USA. [19]University of Chicago, Chicago, Illinois 60637, USA. [20]The University of Western Australia, Crawley, Western Australia 6009, Australia. [21]Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3052, Australia. †A list of authors and affiliations appears in the Supplementary Information.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

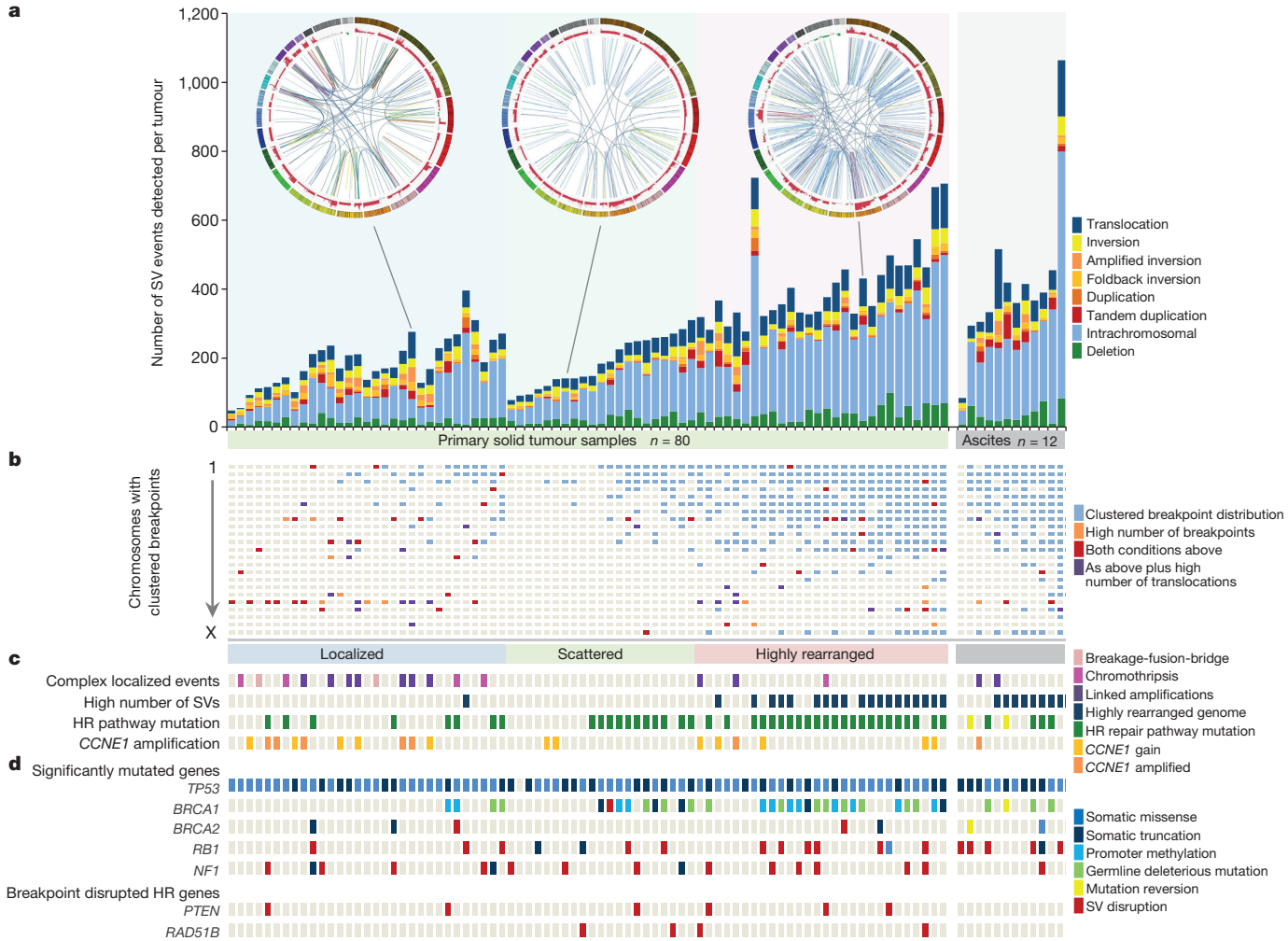


Figure 1 | Genomic features of HGSC. **a**, Structural variants (SVs) in 80 primary solid tumours and 12 ascites samples from 92 individual cases. **b**, Highly rearranged genomes have at least 7 chromosomes containing clustered breakpoint regions (blue squares), high density of breakpoints (orange), both previous conditions (red) or high density of inter-chromosomal translocations (purple). **c**, Genomes with less than 7 chromosomes with clustered breakpoints show localized complex rearrangements (chromothripsis, pink; breakage-fusion-bridge, pale pink; and amplification linked by translocations, purple) or scattered breakpoints. HR, homologous recombination. **d**, Significantly mutated driver genes (IntoGen q-value <0.01) and homologous recombination repair genes disrupted by rearrangements.

Supplementary Table 4). *CCNE1* gain/amplification (19% cases) was largely exclusive of *BRCA1/2* pathway disruption (Fig. 1).

A total of 36,561 somatic structural variants were detected in primary and recurrence samples, ranging from 48 to 1,064 per tumour (Fig. 1a and Supplementary Table 5). The frequency of detected structural variant events (Extended Data Fig. 2a) or single nucleotide variants (SNVs) (Extended Data Fig. 2b) was not related to tumour sample purity. We observed cases with highly rearranged genomes, a low number of scattered breakpoints or a high concentration of events involving one or few chromosomes (Fig. 1a, b and Supplementary Table 6). Similar to basal-like breast cancers[10], primary solid tumour samples with *BRCA1* inactivation had more structural variants ($P < 0.0001$; Extended Data Fig. 2c), involving numerous chromosomal regions (Fig. 1a–d). Tumours with *BRCA1* mutations also had a higher number of genome-wide ($P < 0.001$) and coding ($P < 0.001$) small mutations (SNVs, and insertions or deletions (indels)) compared to homologous-recombination-intact samples (Extended Data Fig. 2d, e). *BRCA1* mutation is frequently associated with a pronounced intratumoural immune response[11–13], which may relate to the greater number of mutations observed in these tumours. We identified one instance of *BRCA2* disruption by chromothripsis[14], but this mutational mechanism was infrequent, as was the occurrence of breakage-fusion-bridge cycles[15] (Fig. 1c, Extended Data Fig. 3 and Supplementary Information). Additional highly complex focal events were observed, including a pattern of amplified loci linked by high-density inter-chromosomal translocations similar to chromoplexy[16], and chromothriptic-associated production of double minute chromosomes[17,18] (Supplementary Information). Chromosome 19 commonly contained clusters of localized breakpoints, particularly samples with amplification involving 19q12 and the *CCNE1* locus (Fig. 1b, c).

Previous exome analyses concluded that apart from *TP53*, somatic point mutations in driver genes are infrequent in primary HGSC[3], an observation consistent with our findings (Fig. 1d and Supplementary Table 7). However, as exome studies have a limited ability to detect gene mutation by structural rearrangement, we considered inactivation of genes by disruption of transcriptional units (gene breakage). Although *NF1* and *RB1* were inactivated by truncating point mutations and indels in only 6% of primary samples (5 out of 80), inclusion of gene breakage raised the frequency of inactivating mutations to 20% for *NF1* and 17.5% for *RB1*, an observation supported by expression levels and transcriptome sequence data (Fig. 1d, Extended Data Fig. 4 and Supplementary Table 8). Most involved breakage without deletion, and therefore would not have been detected by copy analysis. Gene inactivation by breakage was also seen for *PTEN* and *RAD51B* (Extended Data Fig. 4).

We searched for gene fusion events and identified 822 genes involved in intergenic fusion events potentially capable of producing a fused transcript but without evidence of recurrent, biologically plausible driver events in primary HGSC (Supplementary Information). We did not detect any evidence supporting previously reported *ESRRA* and *C11orf20* (also known as *TEX40*)[19], and *CDKN2D* and *WDFY2* (ref. 20) fusions in our cohort (Supplementary Information).

Mutational signatures

Analysis of mutations across cancer genomes can reveal distinct patterns of single nucleotide substitution, many of which are associated with carcinogen exposure or mutational processes operative during tumorigenesis[21]. For most samples, either the previously described[21] age- or *BRCA*-associated signatures were dominant (Fig. 2 and Extended Data Fig. 5), with a close relationship between the age signature, patient age at diagnosis ($P < 0.0001$; Extended Data Fig. 6a) and *CCNE1* gene amplification ($P < 0.01$; Extended Data Fig. 6b). In a small number of cases, the APOBEC or mismatch repair signatures were apparent (Extended Data Fig. 5 and Supplementary Information). The BRCA mutational signature was dominant in all 34 samples with germline or somatic inactivation of *BRCA1* or *BRCA2* ($P < 0.0001$), however, this association was less consistent in the small number of patients with other homologous recombination pathway mutations (Extended Data Fig. 6c). Similarly, primary sensitive cases were found to show the highest BRCA signature contribution (Extended Data Fig. 6d). A group of 12 patients with a dominant BRCA signature, but no mutation in the homologous recombination pathway, was identified (Fig. 2). Eleven of these patients were either resistant or refractory to primary treatment, consistent with intact homologous recombination function. Therefore, although the BRCA signature sensitively detected germline or somatic *BRCA1/2* mutations, it could also be dominant in samples in which there was no evidence of mutational or clinical characteristics suggestive of a homologous recombination pathway defect.

Clinical associations

We compared patterns of structural variation, gene expression, miRNA and methylation in primary tumour samples in the context of response to primary chemotherapy. Consistent with our previous findings[22], amplification of 19q12 involving *CCNE1* was the dominant structural variant associated with primary treatment failure, while most patients with germline or somatic mutation in *BRCA1* or *BRCA2* had a favourable response to treatment (Extended Data Fig. 7a, b and Supplementary Information). Only two refractory patients (2 out of 12, 16.7%) had evidence of homologous recombination deficiency, both involving somatic methylation of *BRCA1* (Extended Data Fig. 5), consistent with the previous observation[3] that methylation is not functionally equivalent to germline mutation in mediating platinum sensitivity. Although somatic mutation frequency (SNVs and indels) was significantly lower in resistant ($P < 0.05$) and refractory ($P < 0.01$) cases compared to sensitive samples (Extended Data Fig. 7c, d), other than *CCNE1* amplification and *BRCA1/2* mutation, we did not observe notable differences in the type or frequency of driver mutations, patterns of methylation or miRNA expression in the clinical subsets (Supplementary Information). We investigated the approximately one-third of tumours with neither a discernable defect in the homologous recombination pathway, nor *CCNE1* amplification or gain. Both progression-free and overall survival in these patients were poor relative to the homologous recombination defective group, similar to *CCNE1* gained/amplified patients in both our study and The Cancer Genome Atlas (TCGA) independent data set[3] (Extended Data Fig. 7b). There was no evidence for high expression of *CCNE1* (Extended Data Fig. 7e) or over representation of other driver events in this group. We previously reported[23] and validated[3,24] four molecular subtypes of HGSC defined by distinct gene expression profiles and clinical outcome. Molecular subtype analysis of the poor-prognosis, homologous-recombination-intact/*CCNE1* wild-type tumours, showed an underrepresentation of the favourable outcome C2/immune subtype in both our series and TCGA (Extended Data Fig. 7f).

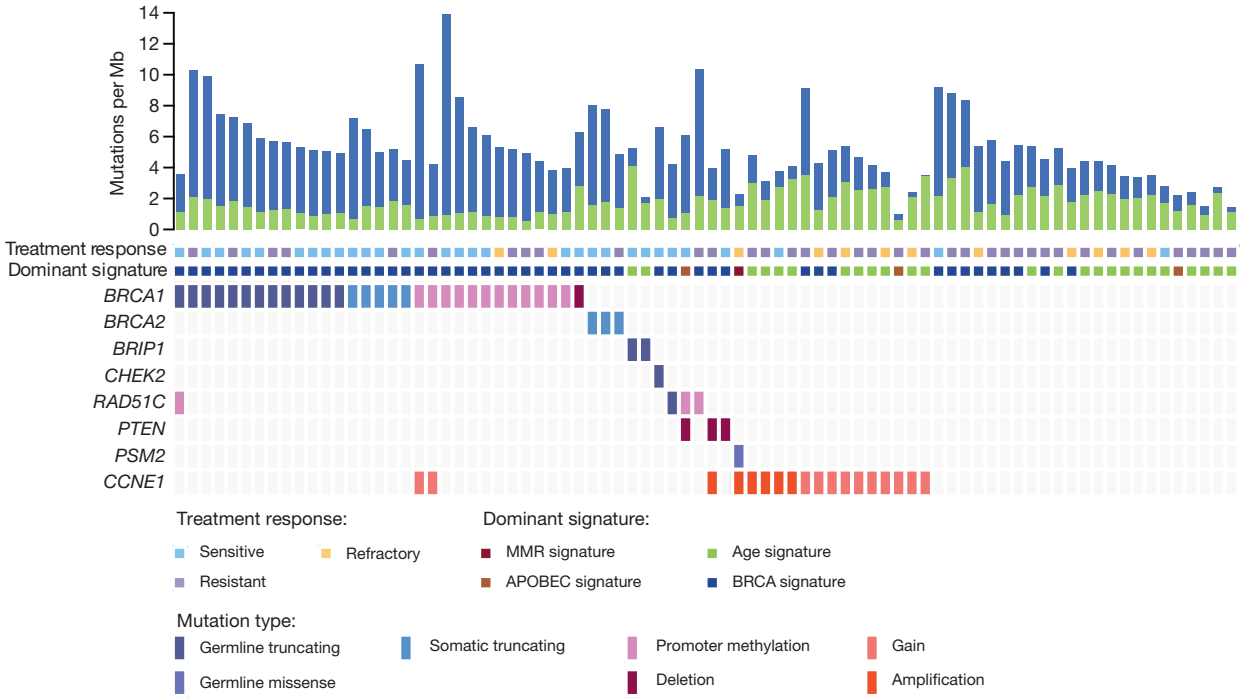


Figure 2 | Mutational signatures of primary high-grade serous ovarian cancer. Bars represent BRCA (blue) and age (green) signature mutation rates for 80 primary tumours. Coloured squares indicate primary response to chemotherapy and dominant mutational signature. Deleterious germline and somatic mutations in the homologous recombination pathway are shown, as well as cases with *CCNE1* gain or amplification, and one case with a germline mismatch repair gene (*PMS2*) mutation. Samples are grouped by mutated gene and ordered left to right by descending BRCA signature mutation rate. Mb, megabases.

Acquired resistance

Surgery or biopsy is rarely performed in recurrent HGSC and so we relied on collection of tumour cells from ascites, drained to relieve abdominal distension, to examine relapsed disease. We focused on patients who were sensitive to initial treatment but had failed subsequent therapy (Supplementary Information). Relapse samples were found to have a higher mutational burden (SNVs and indels) at the coding and non-coding level than matched primary samples (Fig. 3a). In principle, this may reflect wider representation of tumour deposits within ascites, however, few differences were observed between primary ascites and primary tumour samples (Extended Data Fig. 8a, b).

We observed a significant relationship between the number of non-silent coding SNVs unique to the relapse samples and the number of courses of platinum-based chemotherapy the patient received ($P = 0.046$; Extended Data Fig. 8c, d), suggesting that HGSC continue to evolve during treatment. To evaluate the effect of chemotherapy on the HGSC genome, we investigated mutations that were unique to the relapse samples. We observed distinct changes in the frequency of mutations in certain trinucleotide contexts in the relapse samples, with a significant increase in C(C>T)C mutations across the 15 paired cases ($P = 0.023$) (Fig. 3b). We also observed an increase in the frequency of dinucleotide substitutions in resistant (post-treatment) samples (Supplementary Information). However, we found no evidence of platinum-induced mutations specifically occurring in driver genes. We interrogated non-silent coding mutations that were unique to the relapse samples to find those that may be involved in the development of resistance. We identified 32 genes that were mutated in more than one of the paired cases; however, none was clearly linked to chemoresistance (Extended Data Fig. 8e and Supplementary Information).

We investigated other possible mechanisms of resistance, including reversion of germline mutations in *BRCA1* or *BRCA2* (ref. 25). We identified five cases with reversion mutations (Extended Data Fig. 9a), including two patients where more than one reversion mutation was identified (Extended Data Fig. 9). A notable example involved a rapid autopsy patient with a *BRCA2* germline mutation in which five independent reversion events were detected at death (Supplementary Information). This patient was non-responsive to the PARP inhibitor olaparib and to carboplatin given after her disease recurrence (Extended Data Fig. 9c). We subsequently performed deep amplicon sequencing flanking the *BRCA2* germline mutation and reversion sites in additional metastatic deposits from this patient, and identified at least a further seven high confidence reversion events (Fig. 4a, Extended Data Fig. 9d, e and Supplementary Information). We found

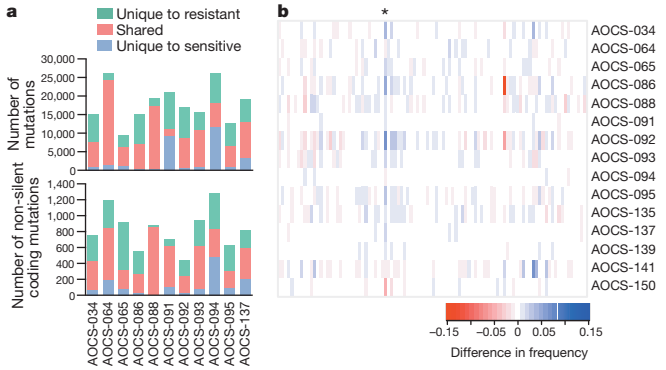


Figure 3 | Somatic mutational patterns in acquired resistant cohort. **a**, Number of mutations in paired primary and relapse samples that are unique or shared at whole genome level (top) and restricted to non-silent coding mutations (bottom). **b**, Difference in mutation frequency in post-treatment, relapse-resistant samples compared to paired samples collected earlier during disease journey. SNVs were examined in their trinucleotide sequence context. Asterisk indicates statistical significance of C(C>T)C using a test of equal or given proportions across the cohort after a Fisher's exact test (two-sided) on individual cases ($P < 0.05$).

no evidence of a deletor phenotype in the WGS data from this patient that might explain the propensity for reversion of the germline *BRCA2* allele (Supplementary Information). Evaluation of reversion allele frequency showed that several subclonal reversion events were detected in individual deposits, particularly in the abdomen where the presence of ascites may facilitate the spread of tumour cells[26] (Extended Data Fig. 9e). Reversion of mutant alleles to a wild-type sequence has been reported previously[25,27]. Minor subclones that have reverted to a wild-type sequence would be difficult to distinguish from contaminating stromal cells and therefore additional reversion events may be present at a low level.

In the second autopsy case, there was no evidence of reversion of the *BRCA1* germline allele. The primary tumour consisted of sheets of epithelial cancer cells with limited stroma and was classified as C2/immune subtype[23] by transcriptome profiling, consistent with our previous data showing enrichment of the C2/immune subtype in *BRCA1* patients[11]. However at autopsy, histological examination revealed an extensive desmoplastic stromal reaction (Fig. 4b) and a C1 subtype (Supplementary Information). Fibrotic stromal reactions are prominent in the poor outcome C1 molecular subtype of primary HGSC[23], and have been associated with poor drug uptake and primary chemoresistance in other solid cancers[28].

BRCA1 promoter methylation is an important somatic driver in ~11% of HGSCs[29]. We identified one patient in which the primary sensitive sample showed extensive promoter methylation and low *BRCA1* expression (Fig. 4c), while the relapse sample had lost *BRCA1* methylation and the gene was expressed at comparable levels to homologous-recombination-intact tumours. Comparison of global methylation patterns of primary and recurrence samples suggested a specific rather than generalized altered methylation status at relapse in this patient (Supplementary Information).

We observed that on average there were more structural variants (~1.6 times) detected in recurrence than in primary samples (Supplementary Table 5). We therefore searched for rearrangements that were unique to the post-treatment samples and that could influence chemotherapy response. Gene breakage was identified in pro-apoptotic genes *FOXO1* and *BCL2L11*, each unique to the relapse sample of a single case (Supplementary Table 8). Breakpoints in each gene were identified in a further two primary chemoresistant samples (Supplementary Information).

We also sought instances of capture and rearrangement of promoter sequences that may drive expression of a partner pro-resistance gene. Among the 114 samples, we identified 36 events in 30 unique genes. These included two recurrent promoter fusions involving *ABCB1*, which encodes the drug efflux pump MDR1. In these tumours, an intergenic deletion of ~250 kilobases (kb) juxtaposed, head-to-tail, the promoter and non-coding exon 1 of *SLC25A40* with exon 2 of *ABCB1*, resulting in a fused transcript (Fig. 4d and Extended Data Fig. 10a). There was no evidence that these events were present in matched primary, sensitive samples. Analysis of transcriptome data showed an increase in expression of *ABCB1* in both recurrent samples with the putative fusion, and a corresponding decrease in *SLC25A40*, relative to matched sensitive samples (Extended Data Fig. 10b). Indeed, the expression of *ABCB1* in these cancers was among the highest found in the cohort (Fig. 4e). Expression of a fused transcript for one of the samples where remaining material was available was confirmed by reverse transcription PCR (RT–PCR) (Extended Data Fig. 10c). Among the transcriptome data for the relapse samples we identified five additional outliers with high *ABCB1* expression. Of these, three had translocations or insertions involving the 5′ end of *ABCB1* (Extended Data Fig. 10a), which may have affected the regulation of *ABCB1*. Using primers specific for the predicted fusion, RT–PCR and Sanger sequencing, we detected a fusion product in one of these samples, suggesting the presence of an additional subclone not apparent at the coverage obtained by WGS (Supplementary Information). We evaluated an additional 51 samples collected from

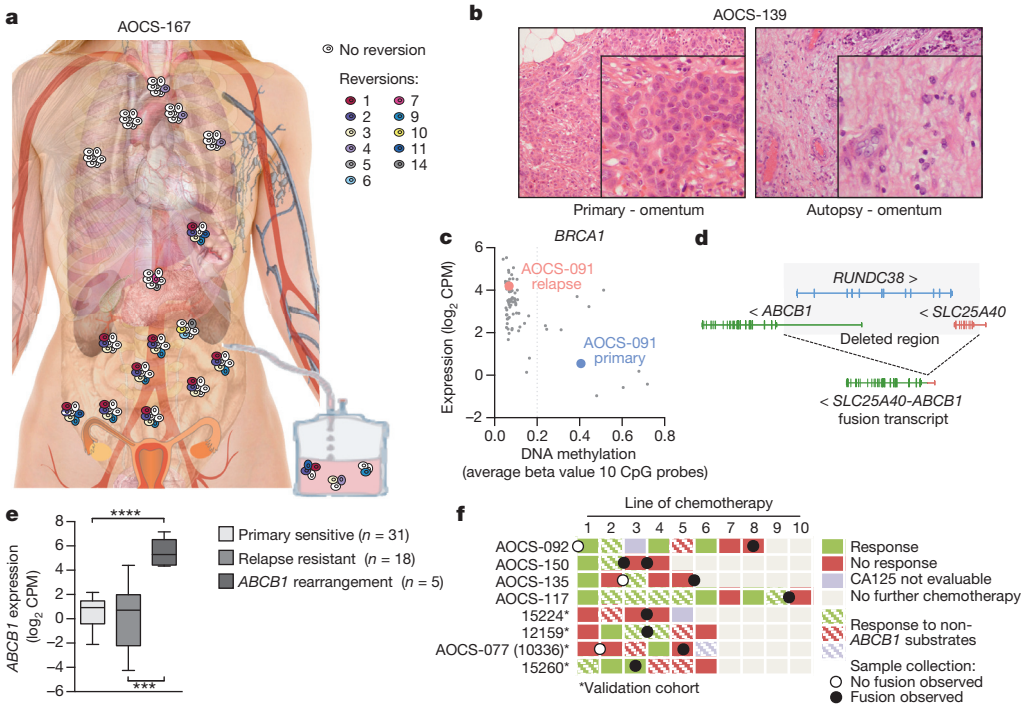


Figure 4 | Molecular changes associated with acquired resistance. **a**, In total, 11 high-confidence reversion events were identified in 16 out of 18 samples collected at autopsy in a *BRCA2* carrier. Additional large deletion event identified by WGS not shown. **b**, Haematoxylin-and-eosin-stained primary (left) and autopsy (right) tumour sections showing epithelial sheets and extensive stromal reaction, respectively. Original magnification, ×200. **c**, *BRCA1* promoter methylation and low expression in AOCS-091 primary sensitive sample compared to hypomethylation and increased expression in matched resistant relapse sample. Grey dots represent all other primary tumours ($n = 80$). **d**, Schematic of deletion and *SLC25A40-ABCB1* predicted transcript. **e**, Box plots of *ABCB1* expression; lines indicate median and whiskers show range (***$P < 0.001$, ****$P < 0.0001$, two-sided *t*-test). **f**, Treatment response in patients with *SLC25A40-ABCB1* fusion. Primary treatment (line 1) includes both surgery and chemotherapy, and CA125 response cannot distinguish between effects of the two treatments.

women with recurrent HGSC and found a further 4 independent fusion events (Extended Data Fig. 10d), all involving samples with above median *ABCB1* expression (Extended Data Fig. 10e). MDR1 (encoded by *ABCB1*) is an efflux pump for various chemotherapeutic agents used in the treatment of ovarian cancer including paclitaxel, etoposide and doxorubicin. All patients with the fusion had been exposed, and their tumours failed to respond, to one or more MDR1 substrates (Fig. 4f).

Discussion

We performed whole-genome analysis of a disease that is characterized molecularly by extensive structural genomic variation and clinically by initial chemo-sensitivity, followed by the frequent emergence of chemotherapy resistance. Despite our finding that gene breakage substantially increases the frequency of recurrent mutations in *NF1* and *RB1*, few genes other than *TP53* are recurrently mutated in HGSC at a high level. The lack of driver protein-coding fusions in primary tumour samples indicates that these are less common than previously suggested[19,20].

While the prevailing view is that primary HGSC lacks recurrent actionable point mutations, the situation in relapse disease is unknown. We did not detect recurrent point mutations that are currently actionable in relapse samples, suggesting that, at best, only low frequency events are likely to be uncovered using personalized genomic evaluation of patients with recurrent HGSC. A recent analysis of DNA changes in *Caenorhabditis elegans* exposed to platinum recorded changes in pyrimidine-rich regions[30], similar to those we observed here in post-treatment samples. Although there seems to be a chemotherapy imprint on the tumour genome that includes non-silent coding changes, these did not result in mutation of known driver or resistance genes in our series.

We observed a range of molecular changes associated with acquired chemoresistance. Reversion of *BRCA1* and *BRCA2* germline alleles has been reported previously[25,31,32], and we extend these findings with a report of an apparent reversion of a methylated *BRCA1* allele in one patient. Analogous to convergent mutation of *SETD2* in renal cancer[33], we found multiple reversion events in two patients, including one at death with 12 high confidence independent *BRCA2* reversion events, underscoring a daunting potential for evolution of resistance in HGSC genomes. We note that this patient did not have debulking surgery and therefore may have had a large tumour pool from which reverted clones could arise under selection. It is of interest to determine whether the use of neoadjuvant chemotherapy[34] in mutation carriers increases the risk of emergence of reverted germline alleles. A surprising feature of this case was that multiple reversion events were found in individual tumour deposits, whereas distinct foci with separate, largely monoclonal events may have been expected. Furthermore, clones were shared between metastatic sites, indicating metastasis to metastasis seeding and possible cooperative interactions between clones. The multiple reversions seen in one end-stage patient contrasted with no *BRCA1* reversion in another patient, where an extensive desmoplastic stromal reaction was observed at autopsy. Tumour desmoplasia has been associated with poor drug uptake and chemoresistance in pancreatic cancer[28], and it is possible that the stromal reaction in this patient allowed tumour survival despite apparent retention of a homologous-recombination-defective phenotype.

The *ABCB1* gene, encoding the multidrug-resistant protein 1 (MDR1), mediates rapid efflux of many chemotherapeutic agents including paclitaxel[35], which forms part of standard primary chemotherapy in HGSC. We found up-regulation of this transporter apparently through promoter fusion and translocation involving the 5′ region of the gene, in approximately 8% of HGSC recurrence samples.

Inhibition of MDR1 activity has not proven to be an effective strategy so far, owing to a considerable increase in paclitaxel toxicity[36,37]. Identification of patients with this novel fusion event may allow clinicians to prioritize treatment with chemotherapy that is not a substrate of MDR1, more targeted use of MDR1 inhibitors, and use of new PARP inhibitors that are poor MDR1 substrates[38].

Collectively, our findings underscore the heterogeneity and apparent adaptability of the HGSC genome under the selective pressure of chemotherapy, and indicate that overcoming resistance to conventional chemotherapy will require a diversity of approaches.

Note added in proof: The C > T mutations in pyrimidine-rich regions we observed in post-treatment samples closely resembled those associated with temozolomide treatment of glioblastoma patients[39].

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 December 2014; accepted 16 March 2015.

1. Vaughan, S. *et al.* Rethinking ovarian cancer: recommendations for improving outcomes. *Nature Rev. Cancer* **11,** 719–725 (2011).
2. Ahmed, A. A. *et al.* Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J. Pathol.* **221,** 49–56 (2010).
3. The Cancer Genome Atlas Research Network. Integrated genomic analysis of ovarian cancer. *Nature* **474,** 609–615 (2011).
4. Gorringe, K. L. *et al.* High-resolution single nucleotide polymorphism array analysis of epithelial ovarian cancer reveals numerous microdeletions and amplifications. *Clin. Cancer Res.* **13,** 4731–4739 (2007).
5. Ciriello, G. *et al.* Emerging landscape of oncogenic signatures across human cancers. *Nature Genet.* **45,** 1127–1133 (2013).
6. Mukhopadhyay, A. *et al.* Development of a functional assay for homologous recombination status in primary cultures of epithelial ovarian tumor and correlation with sensitivity to poly(ADP-ribose) polymerase inhibitors. *Clin. Cancer Res.* **16,** 2344–2351 (2010).
7. Bashashati, A. *et al.* Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. *J. Pathol.* **231,** 21–34 (2013).
8. Castellarin, M. *et al.* Clonal evolution of high-grade serous ovarian carcinoma from primary to recurrent disease. *J. Pathol.* **229,** 515–524 (2013).
9. Cowin, P. A. *et al.* LRP1B deletion in high-grade serous ovarian cancers is associated with acquired chemotherapy resistance to liposomal doxorubicin. *Cancer Res.* **72,** 4060–4073 (2012).
10. Popova, T. *et al.* Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res.* **72,** 5454–5462 (2012).
11. George, J. *et al.* Nonequivalent gene expression and copy number alterations in high-grade serous ovarian cancers with BRCA1 and BRCA2 mutations. *Clin. Cancer Res.* **19,** 3474–3484 (2013).
12. Clarke, B. *et al.* Intraepithelial T cells and prognosis in ovarian carcinoma: novel associations with stage, tumor type, and BRCA1 loss. *Mod. Pathol.* **22,** 393–402 (2009).
13. Soslow, R. A. *et al.* Morphologic patterns associated with BRCA1 and BRCA2 genotype in ovarian carcinoma. *Mod. Pathol.* **25,** 625–636 (2012).
14. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144,** 27–40 (2011).
15. Kinsella, M. & Bafna, V. Combinatorics of the breakage-fusion-bridge mechanism. *J. Comput. Biol.* **19,** 662–678 (2012).
16. Baca, S. C. *et al.* Punctuated evolution of prostate cancer genomes. *Cell* **153,** 666–677 (2013).
17. Rausch, T. *et al.* Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. *Cell* **148,** 59–71 (2012).
18. Nones, K. *et al.* Genome-wide DNA methylation patterns in pancreatic ductal adenocarcinoma reveal epigenetic deregulation of SLIT-ROBO, ITGA2 and MET signaling. *Int. J. Cancer* **135,** 1110–1118 (2014).
19. Salzman, J. *et al.* ESRRA-C11orf20 is a recurrent gene fusion in serous ovarian carcinoma. *PLoS Biol.* **9,** e1001156 (2011).
20. Kannan, K. *et al.* CDKN2D-WDFY2 is a cancer-specific fusion gene recurrent in high-grade serous ovarian carcinoma. *PLoS Genet.* **10,** e1004216 (2014).
21. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500,** 415–421 (2013).
22. Etemadmoghadam, D. *et al.* Integrated genome-wide DNA copy number and expression analysis identifies distinct mechanisms of primary chemoresistance in ovarian carcinomas. *Clin. Cancer Res.* **15,** 1417–1427 (2009).
23. Tothill, R. W. *et al.* Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin. Cancer Res.* **14,** 5198–5208 (2008).
24. Verhaak, R. G. *et al.* Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. *J. Clin. Invest.* **123,** 517–525 (2013).
25. Norquist, B. *et al.* Secondary somatic mutations restoring *BRCA1/2* predict chemotherapy resistance in hereditary ovarian carcinomas. *J. Clin. Oncol.* **29,** 3008–3015 (2011).
26. Lengyel, E. Ovarian cancer development and metastasis. *Am. J. Pathol.* **177,** 1053–1064 (2010).
27. Swisher, E. M. *et al.* Secondary *BRCA1* mutations in *BRCA1*-mutated ovarian carcinomas with platinum resistance. *Cancer Res.* **68,** 2581–2586 (2008).
28. Olive, K. P. *et al.* Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science* **324,** 1457–1461 (2009).
29. Geisler, J. P., Hatterman-Zogg, M. A., Rathe, J. A. & Buller, R. E. Frequency of BRCA1 dysfunction in ovarian cancer. *J. Natl. Cancer Inst.* **94,** 61–67 (2002).
30. Meier, B. *et al. C. elegans* whole-genome sequencing reveals mutational signatures related to carcinogens and DNA repair deficiency. *Genome Res.* **24,** 1624–1636 (2014).
31. Edwards, S. L. *et al.* Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* **451,** 1111–1115 (2008).
32. Sakai, W. *et al.* Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature* **451,** 1116–1120 (2008).
33. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366,** 883–892 (2012).
34. Vergote, I. *et al.* Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer. *N. Engl. J. Med.* **363,** 943–953 (2010).
35. Murray, S., Briasoulis, E., Linardou, H., Bafaloukos, D. & Papadimitriou, C. Taxane resistance in breast cancer: mechanisms, predictive biomarkers and circumvention strategies. *Cancer Treat. Rev.* **38,** 890–903 (2012).
36. Binkhathlan, Z. & Lavasanifar, A. P-glycoprotein inhibition as a therapeutic approach for overcoming multidrug resistance in cancer: current status and future perspectives. *Curr. Cancer Drug Targets* **13,** 326–346 (2013).
37. Lhommé, C. *et al.* Phase III study of valspodar (PSC 833) combined with paclitaxel and carboplatin compared with paclitaxel and carboplatin alone in patients with stage IV or suboptimally debulked stage III epithelial ovarian cancer or primary peritoneal cancer. *J. Clin. Oncol.* **26,** 2674–2682 (2008).
38. Jaspers, J. E. *et al.* Loss of 53BP1 causes PARP inhibitor resistance in Brca1-mutated mouse mammary tumors. *Cancer Discov.* **3,** 68–81 (2013).
39. Johnson, B. E. *et al.* Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* **343,** 189–193 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements The AOCS gratefully acknowledge the cooperation of the participating institutions in Australia, and also acknowledge the contribution of the study nurses, research assistants and all clinical and scientific collaborators including L. Galletta, C. Emmanuel, L. Bowes and J. Hallo. The authors acknowledge assistance from C. Anderson and D. Gwynne. The CASCADE investigators would like to thank the CASCADE Management Committee, all staff at the Victorian Institute of Forensic Medicine, D. Stevens and Tobin Brothers Funerals. The investigators would like to thank the Australia New Zealand Gynaecological Oncology Group (ANZGOG) and the women, and their families, who participated in these research programs. This work was supported by the National Health and Medical Research Council of Australia (NHMRC ID631701), Worldwide Cancer Research (09-0676) and Cancer Australia (1004673). The Australian Ovarian Cancer Study was supported by the US 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 National Health and Medical Research Council of Australia (NHMRC; ID400413, ID400281). The AOCS gratefully acknowledges additional support from S. Boldeman, the Agar family, Ovarian Cancer Australia and Ovarian Cancer Action (UK). The Gynaecological Oncology Biobank at Westmead, a member of the Australasian Biospecimen Network-Oncology group, was supported by grants from the NHMRC (ID 310670, ID 628903) and the Cancer Institute of New South Wales. The CASCADE study was supported by the Peter MacCallum Cancer Centre Foundation, and in kind by the Victorian Institute of Forensic Medicine and Tobin Brothers Funerals.

Author Contributions Project supervision: J.V.P., Nicola.W., A.D.F., S.M.G., D.D.L.B. Study design: A.M.P., E.L.C., D.E., D.W.G., S.F., P.C., Nicola.W., A.D.F., S.M.G., D.D.L.B. Sample acquisition: S.F., K.A., A.O.C.S., A.H., L.M., O.M.N., C.K., J.H., Y.E.C., P.H., M.F., M.Q., J.P., S.C., P.O.B., J.L., P.W., N.T., H.T., M.S., R.D., E.S., H.G., A.J., O.M.N., E.L., A.D.F., D.D.L.B. Sample preparation: E.L.C., D.E., D.W.G., P.C., Y.E.C., P.H., C.M., J.H. Data acquisition: E.L.C., D.E., D.W.G., S.F., D.K.M., G.M.A., T.P.H., T.S., I.H., C.N., E.N., S.M., S.I., T.J.C.B., A.N.C. Performed patient autopsy: G.Y., K.S. Sequence data management, alignment and mutation identification: A.M.P., K.N., F.N., S.K., O.H., M.A., C.L., S.W., Q.X., J.L.F., Nick.W., S.H.N., P.J.W., J.V.P., S.M.G. Genome informatics, software tool development: A.M.P., P.J.B., K.S.K., F.N., S.K., B.P., O.H., M.A., C.L., S.W., D.F.T., Q.X., J.L.F., Nick.W., J.V.P., Nicola.W. Data analysis: A.M.P., E.L.C., D.E., D.W.G., J.G., K.N., P.J.B., K.S.K., M.C.JQ., K.Q., C.W.B., E.C., H.S.L., G.A.Y., W.A., R.W.T., A.L., N.H., R.B., J.E., M.D., R.V., C.S., J.V.P., Nicola.W., S.M.G. Perform molecular/verification analysis: E.L.C., K.N., S.M., E.M. Wrote the manuscript: A.M.P., E.L.C., D.E., D.W.G., Nicola.W., A.D.F., S.M.G., D.D.L.B.

Author Information The whole genome and transcriptome sequencing data have been deposited in the European Genome-phenome Archive (EGA) repository under the accession code EGAD00001000877. Genotyping, methylation and miRNA data sets have been submitted into the Gene Expression Omnibus (GEO) accession GSE65821. A complete list of the AOCS Study Group can be found at http://www.aocstudy.org. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.D.L.B. (david.bowtell@petermac.org).

METHODS

Patient cohort description. The study population consisted of women diagnosed with epithelial ovarian, primary peritoneal or fallopian tube cancer between 1992 and 2012. The women were treated at hospitals across Australia and were recruited through the Australian Ovarian Cancer Study (AOCS)[23] or through the Gynaecological Oncology Biobank at Westmead Hospital in Sydney. Four primary refractory cases were obtained from the Hammersmith Hospital Imperial College (London, UK) and the University of Chicago (Chicago, USA). Ethics board approval was obtained at all institutions for patient recruitment, sample collection and research studies. Written informed consent was obtained from all participants in this study.

All patients were diagnosed with serous carcinoma of high-grade (grade 2 or grade 3) and advanced stage (FIGO stage III or IV, International Federation of Gynaecology and Obstetrics) and received platinum-based chemotherapy as part of primary treatment. In total, 81 out of 92 (88%) patients also received a taxane at first line treatment; 86 out of 92 (93%) patients received a taxane during the course of their clinical care.

From previously published studies we estimated that with 92 unique patients, we would expect to achieve 90% power for 90% of genes to detect mutations that occur at a frequency of ~10% above the background rate for ovarian cancer (assuming a somatic mutation frequency of <2 per Mb)[40].

Additionally, as normal controls, 7 fallopian tube samples were obtained from Dana-Farber Cancer Institute (Boston, USA). Five samples were fallopian tube secretory epithelial cells isolated from dissociated fallopian tube fimbriae and expanded in short-term culture in WIT medium (Stemgent), two samples were fallopian tube epithelium collected from dissociated fallopian tube fimbriae without expansion *in vitro*.

The validation cohort of relapse ascites samples were collected from patients recruited through the AOCS with high-grade and advanced stage disease. In total, 50 out of 51 (98%) patients received platinum-based chemotherapy as part of primary treatment; 46 out of 51 (90.2%) patients also received a taxane at first line treatment; 47 out of 51 (92.2%) received a taxane during the course of treatment.

Clinical definitions. Progression-free survival was the time interval between histological or cytological diagnosis and disease progression based on GCIG (Gynaecological Cancer Inter Group) CA125 criteria[41], imaging or clinical evaluation. Overall survival was defined as the interval between histological or cytological diagnosis and death from any cause. A response to second and subsequent lines of chemotherapy was defined as at least a 50% decrease in CA125 from an increased pre-treatment level, confirmed and maintained for at least 28 days[41]. A complete response to second and subsequent lines of chemotherapy was defined as normalization of CA125 from an increased pre-treatment level, confirmed and maintained for at least 28 days.

Inclusion criteria for treatment response groups. *Primary resistant and refractory*: This group included patients with evidence of disease progression within ~6 months from the end of primary treatment. An additional level of stringency was added by only including patients that demonstrated no response to subsequent lines of chemotherapy. A subset were designated as refractory, which included patients with disease progression while on primary treatment, or within one month of the end of primary treatment[42], evidenced by persistently elevated CA125 or clinical progression.

Primary sensitive: Patients with no evidence of disease progression within 6 months of the end of primary treatment or patients that demonstrated responses, either normalization of CA125 or a 50% decrease in CA125 (from an increased pre-treatment level), to multiple lines of platinum-based chemotherapy.

Acquired resistance: Patients that failed to respond to chemotherapy for relapsed disease having previously demonstrated sensitivity to earlier lines of chemotherapy. Initial sensitivity was inferred if patients with macroscopic residual disease after initial surgery had no progression within 6 months from the end of primary treatment and/or had a complete or 50% response to chemotherapy for second and subsequent lines of treatment, before the development of resistance (that is, no response to later lines of chemotherapy) and collection of relapse sample.

Nucleic acid isolation. DNA was isolated from peripheral lymphocytes or lymphoblastoid cell lines using the salting out method for normal DNA. For tumour DNA and RNA isolation, frozen tissue sections were cyrosectioned, with a section for haematoxylin and eosin staining being taken before and after serial sectioning for assessment of tumour content. For samples containing >70% tumour, 4 × 50 μm and 12 × 100 μm whole frozen tissue sections were used for DNA and RNA extractions, respectively. For samples containing <70% tumour, needle dissection of tumour cells was performed on up to 100 × 10 μm frozen sections for DNA and RNA extractions. Tumour cells were isolated from ascites using Dynabeads Epithelial Enrich (Invitrogen/Life Technologies), followed by DNA and RNA extractions. DNA extractions were performed using the DNeasy blood and tissue kit (QIAGEN). DNA was quantified using the Qubit dsDNA BR assay. RNA was isolated using the mirVana miRNA Isolation kit (Ambion/Life Technologies). RNA quality was assessed using the RNA 6000 Nano kit on the BioAnalyzer (Agilent) and the NanoDrop was used to assess quality and determine RNA quantity.

SNP arrays and copy number analysis. Tumour and matched normal DNA was assayed with the Omni 2.5-8, V1.0 and V1.1 IlluminaBeadChips as per manufacturer's instructions (Illumina). Single nucleotide polymorphism (SNP) arrays were scanned on an iScan (Illumina), data was processed using the Genotyping module (v.1.9.4) in GenomeStudio v.2011.1 (Illumina) to calculate B-allele frequencies (BAF) and logR ratios. Genome alteration print[43] was used to call somatic regions of copy number change after low-quality probes assessed in the matched normal sample with GenCall (GC) score of <0.7 were removed. Commonly affected genes in regions with significant ($q < 0.05$) gain and loss were determined using GISTIC v2.0 (ref. 44) and genes within regions of copy number change were annotated according to the Ensembl v70 gene model.

Whole-genome library and sequence generation. Tumour cellularity was assessed using SNP array data and qPure[45] indicating a median tumour content of 85.3% (range 49.7–99.8). Matching germline DNA from blood mononuclear cells or lymphoblastoid cell lines were analysed for all patients. Sequence libraries were generated from 1 μg of genomic DNA using the standard library preparation technique as in the protocol in the TruSeq DNA PCR-Free Sample Preparation Guide for the TruSeq DNA Sample Preparation Kit (FC-121-2001, Illumina). Sequencing was carried out by the Illumina Genome Network to a minimum average of 30-fold base coverage for germline samples and 40-fold coverage for tumour samples.

WGS data processing and quality control. Each lane of sequencing data underwent alignment to the Genome Reference Consortium human genome assembly (GRCh37) using multi-threaded BWA[46] 0.6.2-mt resulting in sorted lane level files in sequence alignment/mapping (SAM) format then compressed and converted to binary file (BAM) created by Samtools[47] 0.1.19. Whole sample level merged BAMs, one each for matched germline and tumour samples were produced by in-house tools and optical duplicate reads marked using Picard MarkDuplicates 1.97 http://picard.sourceforge.net. Quality assessment and coverage estimation was carried out by qProfiler and qCoverage (http://sourceforge.net/projects/adamajava). To test for the presence of sample or data swaps all sequence data were assessed for concordance at approximately 1.4 million polymorphic genomic positions including the genotyping array data by qSignature.

Transcriptome sequence data generation. Between 500 ng and 2 μg of total RNA (RNA integrity number >7, except for autopsy samples) was used for library preparation using the TruSeq RNA Sample Preparation v2 kit (Illumina) as per the manufacturer's instructions using the Low-Throughput protocol. All libraries were sequenced as paired-end 100-bp on a HiSeq2000 instrument (Illumina).

Transcriptome sequence data processing and quality control. Each barcode-separated lane of sequencing data underwent three streams of processing. To assess quality each was aligned to the Genome Reference Consortium human genome assembly (GRCh37) using multi-threaded BWA[46] 0.6.2-mt. RNA-seQC[48] was used to investigate RNA sequencing quality and concordance to all other sequencing was assessed at approximately 1.4 million polymorphic genomic positions by qSignature (http://sourceforge.net/projects/adamajava). Estimation of gene and transcript abundance was carried out using RSEM[49] for the Ensembl v70 gene and transcript model. The 'expected_count' values were transformed to $\log_2$-counts per million ($\log_2$CPM) and the mean–variance relationship estimated to compute appropriate observational-level weights using the methods available in R-package Limma (http://genomebiology.com/2014/15/2/R29). A quantile normalization step was applied to normalize between samples before downstream analysis using the methods available in R-package Limma. A further gapped alignment and analysis with Mapsplice[50] was also carried out in order to identify insertions and deletions as well as differential splicing patterns and gene fusion transcripts.

MicroRNA expression analysis. MicroRNA expression levels were examined using the nCounterHuman v2.1 miRNA expression analysis kit (NanoString Technologies), with 100 ng total RNA as input for hybridization to reporter/capture probes. Purification was performed on the nCounter Prep Station and detection using the nCounter Digital Analyzer (NanoString Technologies), as per the manufacturer's instructions.

The nSolver Analysis Software (NanoString Technologies) was used to extract the raw counts for all samples that passed internal QC metrics, before normalization using the housekeeping genes to control for sample input.

Methylation. Methylation levels across the genome were examined on the Infinium Human Methylation 450 Bead Chip (Illumina), the samples were processed at the Australian Genome Research Facility. The data were background corrected, normalized to internal controls and quality control was performed at

the probe and sample level. A total of 572 probes failed in greater than 25% of the samples owing to their detection *P* values being greater than 0.05, and these were removed from the data[51]. COMBAT was used to remove batch effects[52].

Somatic SNV variant detection. High confidence somatic substitutions were identified as the intersect of the post filtered output of the mutation detection tools qSNP[53] and the Unified Genotyper[54] (GATK). In addition, supporting evidence for variants was sought in matched independent and/or orthogonal data sets such as long mate pair sequencing data produced on Life Technologies SOLiD platform and HiSeq RNA transcriptome sequencing. Annotation of variants was carried out against Ensembl v.70 and consequence prediction using Ensembl perl API modules. Two methods, Mutsig[55] and IntOGen[56], were used to identify significant, recurrently mutated genes in the primary tumour data set ($n = 80$).

Somatic short insertion and deletion variant detection. Pindel[57] was used to identify indels of 1–50 bp in length and variant filtering, annotation and prioritization was carried out as previously described[18]. Manual review of both germline and somatic variants in genes highlighted in this study was carried out using Integrative Genomics Viewer (IGV)[58]. Gene model annotation of variants was carried out against Ensembl v.70 and consequence prediction using Ensembl perl API modules for variant effect prediction across all transcripts.

Structural variant detection. Structural variants of length 50 bp or greater were identified using qSV[18] (http://sourceforge.net/projects/adamajava), which discovers somatic structural rearrangements from tumour and matched normal sequencing using three methodologies: discordantly mapped read pairs; clusters of soft clipping; and a split contig alignment generated from localized *de novo* assembly of unmapped plus aberrantly mapped reads. Only high confidence somatic structural variants were used in the analysis defined as events that were identified by more than one of the detection methodologies or that had at least 10 supporting reads, and where no evidence of the event was identified in the matched germline sample including low quality data.

Breakpoint annotation and consequence predictions were performed against the Ensembl v70 gene model to calculate the consequences for flanking genes or fragments of genes. Breakpoints were also integrated with the segmented copy number data to inform characterization: intrachromosomal breakpoints flanking a copy number segment of loss were annotated as deletions; those linking segments of similar DNA copy number were annotated as intrachromosomal translocations; duplications and inversions associated with increases in copy number enabled characterization of tandem duplications and amplified or foldback inversions.

Somatic variant verification. In-house developed software examined genomic positions of variants in the matched data sets scoring supporting read evidence and assessing coverage in a similar manner to methods used in a recent genomic landscape paper for cervical carcinoma[59]. In addition for structural variants, SNP array data provided orthogonal evidence in support for structural variant events that cause a copy number shift when boundaries of copy number change were identified in close proximity to breakpoints.

Mutational signatures and detection of kataegis. Mutational signatures were deciphered using the framework developed by previously[21] (http://www.mathworks.com/matlabcentral/fileexchange/38724). Loci of localized hyper-mutation, termed kataegis, were identified by calculating the inter-mutation distances between the genomic position of sorted, high-confidence somatic substitution variant calls for each sample. Inter-mutation distances were then segmented using piecewise constant fitting to find regions of constant inter-mutation distance[60]. Parameters used for piecewise constant fitting were c525 and kmin52 (ref. 60). Putative regions of kataegis were identified as those segments containing six or more consecutive mutations with an average inter-mutation distance of less than or equal to 1,000 bp.

Detection of breakpoint clustering. Chromosomes containing highly significant non-random distribution of breakpoints were identified as recently described[61] with a stringent threshold of $P < 0.00001$. Chromosomes with remarkably high numbers of rearrangement events were identified as outliers defined as a breakpoint per megabase rate exceeding 1.5 times the length of the inter-quartile range from the seventy-fifth percentile for each sample with a minimum threshold of 35 breakpoints per chromosome. Chromosomes identified to have both characteristics were inspected against criteria for chromothripsis[61] or the breakage-fusion-bridge cumulative rearrangement model[15]. Similarly, chromosomes with high numbers of translocations were identified with a minimum threshold of 10 translocation breakpoints per chromosome following observations of localized events.

Molecular subtyping. Tumour molecular subtypes were classified using previously described expression profiles[23] as the learning set. Expression of genes between each high-grade serous molecular subtype was compared to identify subtype specific gene signatures. All genes differentially expressed at a FDR < 5% were selected as subtype specific genes. The union of all subtype specific gene sets were used as the features of the classifier. The normalized data from the learning set and transformed data from the RNA-seq data (described in detail in the previous section on transcriptome data processing) were further scaled independently so that every gene in the two sets had zero mean and unit variance. Finally, the transformed gene expression profiles were used to identify the molecular subtypes of samples in the cohort using *k*-nearest neighbour classification algorithm. The accuracy of the classification procedure was estimated based on the samples of known molecular subtypes, as defined by previous microarray experiments[23]. These samples were removed from the training set and the classification procedure repeated. Of the 19 samples with known molecular subtypes in the cohort, 17 were classified correctly (89.47% accuracy).

Verification and validation of *SLC25A40-ABCB1* fusion transcript. A cohort of 51 relapse ascites samples was used to examine frequency of *SLC25A40-ABCB1* fusion. RNA from ascites samples was reverse transcribed into cDNA using random primers (Promega) and M-MLV reverse transcriptase (Promega). Quantitative reverse transcription PCR was performed to examine *SLC25A40* and *ABCB1* expression, with *GAPDH* and *ACTB* for normalization, using the following primers: *ABCB1*_forward 5′-GAGAGATCCTCACCAAGCGG-3′, *ABCB1*_reverse 5′-CGAGCCTGGTAGTCAATGCT-3′, *SLC25A40*_forward 5′-AGGGAACATTGGATGCATTTTT-3′, *SLC25A40*_reverse 5′-AGCACTTAATTGATCATAGCAGGT-3′.

The $\Delta\Delta C_t$ method was used to calculate expression levels in ascites compared to expression in the SKOV3 cell line[9].

Testing and verification of the presence of the *SLC25A40-ABCB1* fusion transcript was performed using nested PCR with primers to exon 1 of *SLC25A40* and exon 3 of *ABCB1*: outer_forward 5′-CGGCTCTGTGTTGACCAAAC-3′, outer_reverse 5′-TCTTTGCTCCTCCATTGCGG-3′; internal_forward 5′-CCCGTCACCAGGGTTATTCC-3′, internal_reverse 5′-CCCCTTCAAGATCCATTCCGA-3′.

PCR conditions were as follows: 98°C for 30 s, 30 cycles of 98°C for 10 s, 60°C for 30 s, 72°C for 10 s, and 72°C for 10 min. PCR product from first PCR was purified using QIAquick PCR purification kit (QIAGEN), PCR product from second PCR was run on a 2% agarose gel, bands were excised and purified using Wizard SV Gel and PCR Clean-up System (Promega) before Sanger sequencing using internal primers.

Examination of *BRCA2* reversion events in AOCS-167. Primers to flank the germline and reversion events identified by WGS were used to amplify the regions of interest for next-generation sequencing of additional tumour samples from AOCS-167. Each of the gene-specific primers were generated with a common sequence (CS) on the 5′ end. The paired end sequence (PE-CS) primers add the Illumina adaptor sequences and a sample specific 10-bp barcode (BC) to the libraries. *BRCA2*_1F 5′-TCATGATGAAAGTCTGAAGAAAAATGA-3′; *BRCA2*_1R 5′-GAGCAAGACTCCACCTCAAA-3′; *BRCA2*_2F 5′-AATTTTTGCAGAATGTGAAAAGCT-3′; *BRCA2*_2R 5′-ACTAAACAGAGGACTTACCATGACT-3′; *BRCA2*_3F 5′-ACTGTTTCTATGAGAAAGGTTGTGA-3′; *BRCA2*_3R 5′-TCCAATGTGGTCTTTGCAGC-3′; *BRCA2*_4F 5′-AGCACATTCTACATAAACTGTTTCT-3′; *BRCA2*_4R 5′-TGACTTTCCAATGTGGTCTTTGC-3′; CS_F 5′-ACACTGACGACATGGTTCTACA-3′; CS_R 5′-TACGGTAGCAGAGACTTGGTCT-3′; PE-CS_F 5′-AATGATACGGCGACCACCGAGATCT-3′; PE-CS_R 5′-CAAGCAGAAGACGGCATACGAGAT-BC-3′.

The first PCR used the CS-gene-specific primers and Phusion High-Fidelity DNA Polymerase (Thermo Scientific) with PCR conditions as follows: 98°C for 30 s, 30 cycles of 98°C for 10 s, 60°C for 30 s, 72°C for 10 s, and 72°C for 10 min. QIAquick PCR purification (QIAGEN) was used and the product was used as a template for a second PCR using PE-CS primers with a unique barcode for each sample and PCR conditions as follows: 98°C for 10 min, 15 cycles of 98°C for 10 s, 60°C for 30 s, 72°C for 30 s, and 72°C for 10 min. Libraries were purified using AMPure XP Beads and analysed on Agilent BioAnalyser High Sensitivity DNA chip. All 18 libraries were run on a single Illumina MiSeq flowcell.

Sequencing data was adaptively trimmed (equivalent to the trimming performed by BWA with $q = 30$) then aligned to the amplicon sequences using an in-house bioinformatics tool (Peter MacCallum Cancer Centre Bioinformatics Core, unpublished). This tool utilizes the known primer sequences to align each read to its amplicon of origin, and where there is overlap in the forward and reverse a consensus sequence is generated using the highest quality base. Fastqc analysis was performed on the fastq files (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), the average base sequence quality was above 35 for all samples. mpileup files were generated from the aligned sequence using SAMtools 0.1.18 (ref. 47) and VarScan v2.3 (ref. 62) was subsequently used to calculate the read depth at each position. The mean read depth ranged from 8,101–21,019 per sample. Indel variant calling was performed using IndelGenotyperV2 (GATK v1.0.4905)[54], with the minimum fraction of reads with the indel set to 0.001 and a minimum count of reads supporting the indel set to 1. All variants were manually reviewed in IGV.

For verification mutation specific PCR primers were designed with 1–5 bp spanning the deletion, and non-mutation specific primers used were *BRCA2*_2F and *BRCA2*_4R.

*BRCA2*_rev1_1R 5′-TGTGTTTTCACTGTCA-3′; *BRCA2*_rev1_2R 5′-TGTGTTTTCACTGTCACA-3′; *BRCA2*_rev2_1R 5′-GGACTTACCATGACTTC-3′; *BRCA2*_rev2_2R 5′-GGACTTACCATGACTTCTC-3′; *BRCA2*_REV3_1F 5′-TTATACTTTAACAGGAA-3′; *BRCA2*_rev4_1F 5′-AAACATCAGGGAATGT-3′; *BRCA2*_rev5_1R 5′-GTGTTTTCACTGTCTT-3′; *BRCA2*_rev5_2R 5′-GTGTTTTCACTGTCTTCA-3′; *BRCA2*_rev6_1R 5′-ACCATGACTTGCAGCG-3′; *BRCA2*_rev6_2R 5′-ACCATGACTTGCAGCGAT-3′; *BRCA2*_rev7_1R 5′-TACCATGACTTGCAGT-3′; *BRCA2*_rev7_2R 5′-TACCATGACTTGCAGTG-3′; *BRCA2*_rev8_1R 5′-GACTTGCAGCTTCTCTTG-3′; *BRCA2*_rev9_1R 5′-CCATGACTTGCAGCTC-3′; *BRCA2*_rev9_2R 5′-CCATGACTTGCAGCTCTC-3′; *BRCA2*_rev10_1F 5′-GTTTTATACTTTAACAGC-3′; *BRCA2*_rev10_2F 5′-GTTTTATACTTTAACAGCTG-3′; *BRCA2*_rev11_1F 5′-CTTTAACAGGATTTGGAAT-3′; *BRCA2*_rev12_1F 5′-ACAGGATTTGGAAAAC-3′; *BRCA2*_rev12_2F 5′-ACAGGATTTGGAAAACAT-3′; *BRCA2*_rev13_1F 5′-CAGGATTTGGAAAAAT-3′; *BRCA2*_rev14_1F 5′-TTGGAAAAACATCAGC-3′; *BRCA2*_rev14_2F 5′-TTGGAAAAACATCAGCTG-3′; *BRCA2*_rev15_1F 5′-GAAAAACATCAGGGAAG-3′; *BRCA2*_rev15_2F 5′-GAAAAACATCAGGGAAGTA-3′.

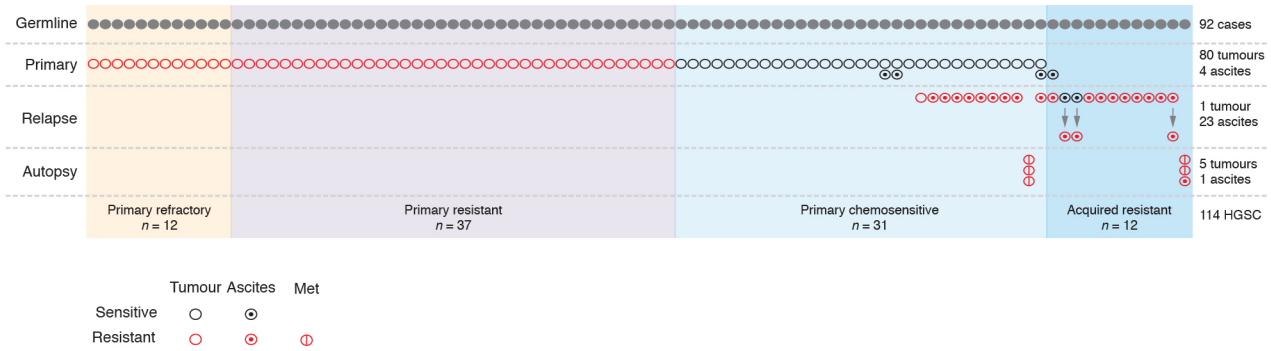
Touchdown PCR was used for mutation-specific PCR to improve sensitivity and specificity, PCR conditions were as follows: 98°C for 5 min, 10 cycles of 98°C for 10 s, 65°C for 30 s decreasing 1°C per cycle, 72°C for 10 s, 15 cycles of 98°C for 10 s, 55°C for 30 s, 72°C for 10 s and 72°C for 10 min. PCR products were run on 2% agarose gel for visual verification of presence of PCR product of correct size.

Code availability. All in house developed tools are available for download from http://sourceforge.net/projects/adamajava.

Statistics. No statistical methods were used to predetermine sample size, and the experiments were not randomized. Further details of statistics are provided in the Supplementary Information.

40. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505,** 495–501 (2014).
41. Rustin, G. J. *et al.* Definitions for response and progression in ovarian cancer clinical trials incorporating RECIST 1.1 and CA 125 agreed by the Gynecological Cancer Intergroup (GCIG). *Int. J. Gynecol. Cancer* **21,** 419–423 (2011).
42. Friedlander, M. L. *et al.* Clinical trials of palliative chemotherapy in platinum-resistant or -refractory ovarian cancer: time to think differently? *J. Clin. Oncol.* **31,** 2362 (2013).
43. Popova, T. *et al.* Genome Alteration Print (GAP): a tool to visualize and mine complex cancer genomic profiles obtained by SNP arrays. *Genome Biol.* **10,** R128 (2009).
44. Mermel, C. H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12,** R41 (2011).
45. Song, S. *et al.* qpure: A tool to estimate tumor cellularity from genome-wide single-nucleotide polymorphism profiles. *PLoS ONE* **7,** e45835 (2012).
46. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
47. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
48. DeLuca, D. S. *et al.* RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28,** 1530–1532 (2012).
49. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12,** 323 (2011).
50. Wang, K. *et al.* MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res.* **38,** e178 (2010).
51. Wilhelm-Benartzi, C. S. *et al.* Review of processing and analysis methods for DNA methylation array data. *Br. J. Cancer* **109,** 1394–1402 (2013).
52. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8,** 118–127 (2007).
53. Kassahn, K. S. *et al.* Somatic point mutation calling in low cellularity tumors. *PLoS ONE* **8,** e74380 (2013).
54. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20,** 1297–1303 (2010).
55. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499,** 214–218 (2013).
56. Gundem, G. *et al.* IntOGen: integration and data mining of multidimensional oncogenomic data. *Nature Methods* **7,** 92–93 (2010).
57. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25,** 2865–2871 (2009).
58. Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnol.* **29,** 24–26 (2011).
59. Ojesina, A. I. *et al.* Landscape of genomic alterations in cervical carcinomas. *Nature* **506,** 371–375 (2014).
60. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149,** 979–993 (2012).
61. Korbel, J. O. & Campbell, P. J. Criteria for inference of chromothripsis in cancer genomes. *Cell* **152,** 1226–1236 (2013).
62. Koboldt, D. C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22,** 568–576 (2012).

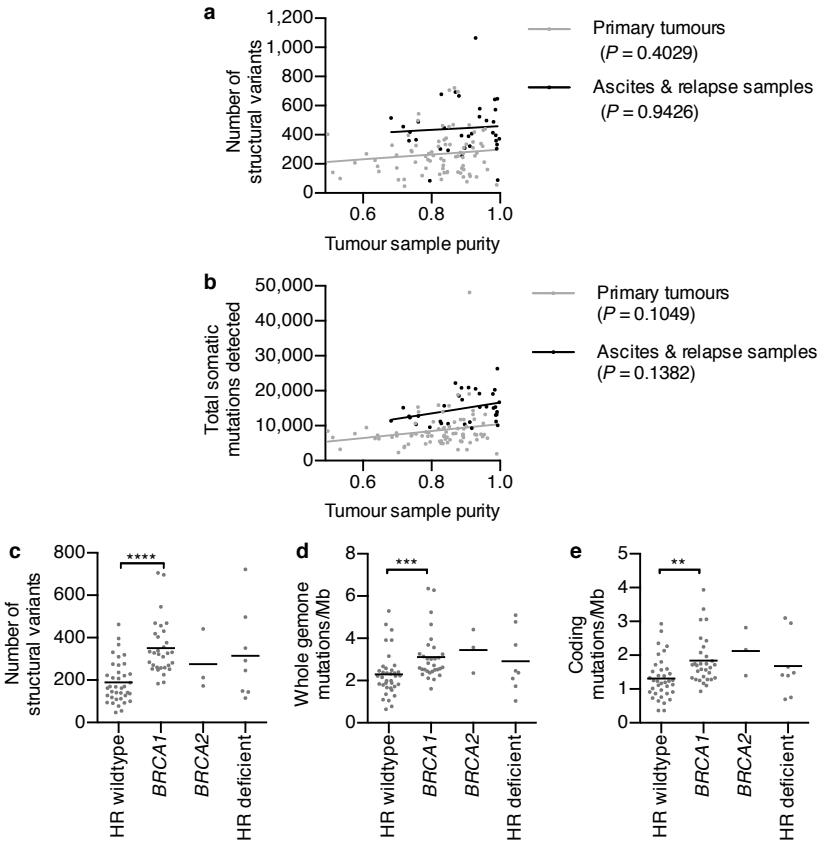
a



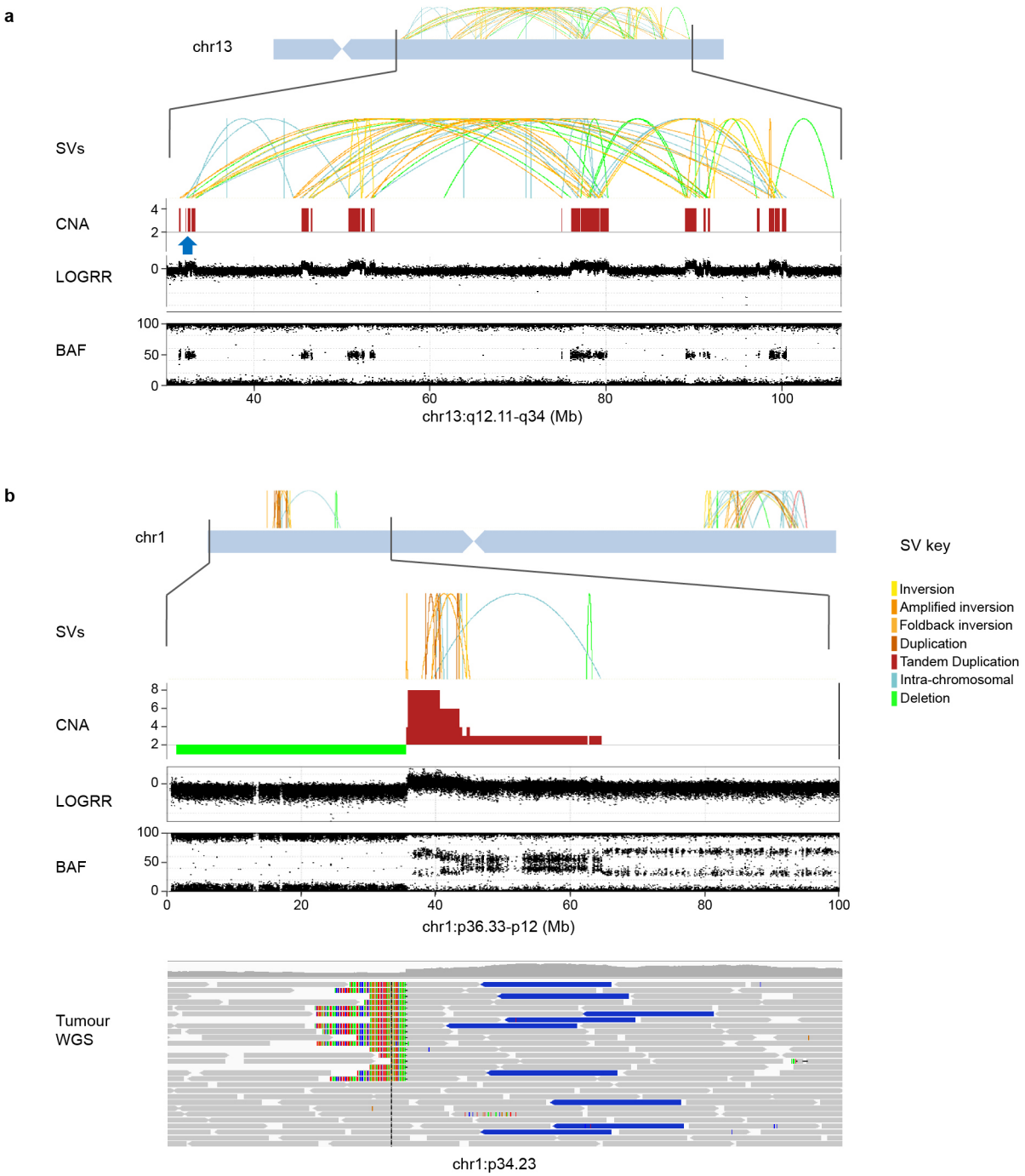
b

	Primary refractory	Primary resistant	Primary chemosensitive	*P* value	Acquired resistant (no primary tumour)	All cases
Age						
Mean	61.7	61.3	57.6	0.188[a]	57.7	59.7
Range	(52.5 - 74.7)	(44.5 - 78.4)	(39.2 - 73)		(45.5 - 70.8)	(39.2 - 78.4)
Primary site						
Ovary	8 (67%)	27 (73%)	28 (90%)	0.115[b]	10 (83%)	73 (79%)
Peritoneum	4 (33%)	9 (24%)	2 (6%)		2 (17%)	17 (18%)
Fallopian tube	0 (0%)	1 (3%)	1 (3%)		0 (0%)	2 (2%)
Grade						
2	2 (17%)	4 (11%)	5 (16%)	0.745[b]	1 (8%)	12 (13%)
3	10 (83%)	33 (89%)	26 (84%)		9 (75%)	78 (85%)
Ungraded*	0 (0%)	0 (0%)	0 (0%)		2 (17%)	2 (2%)
Stage						
III	12 (100%)	31 (84%)	25 (81%)	0.300[b]	10 (83%)	78 (85%)
IV	0 (0%)	6 (16%)	6 (19%)		2 (17%)	14 (15%)
Residual disease						
No macroscopic	1 (8%)	2 (5%)	1 (3%)	0.876[b]	1 (8%)	5 (5%)
Macroscopic ≤ 1 cm	6 (50%)	23 (62%)	18 (58%)		2 (17%)	49 (53%)
Macroscopic >1 cm	5 (42%)	12 (32%)	12 (39%)		4 (33%)	33 (36%)
Macroscopic (size unknown)	0 (0%)	0 (0%)	0 (0%)		5 (42%)	5 (5%)
Progression free survival						
Number of events	12 (100%)	37 (100%)	24 (77%)	<0.0001[c]	12 (100%)	85 (92%)
Median months^	4.98	9.24	17.72		21.21	10.39
95% CI of median	(4.14 - 7.04)	(7.56 - 10.13)	(13.28 - 21.96)		(13.35 - 26.01)	(9.67 - 11.47)
Overall survival						
Number of events	11 (92%)	36 (97%)	19 (61%)	<0.0001[c]	10 (83%)	76 (83%)
Median months^	12.92	20.22	55.86		46.04	27.11
95% CI of median	(6.41 - 34.68)	(16.31 - 22.49)	(45.96 - 69.07)		(31.76 - 64.83)	(21.50 - 34.49)
Total cases	**12**	**37**	**31**	**80**	**12**	**92**

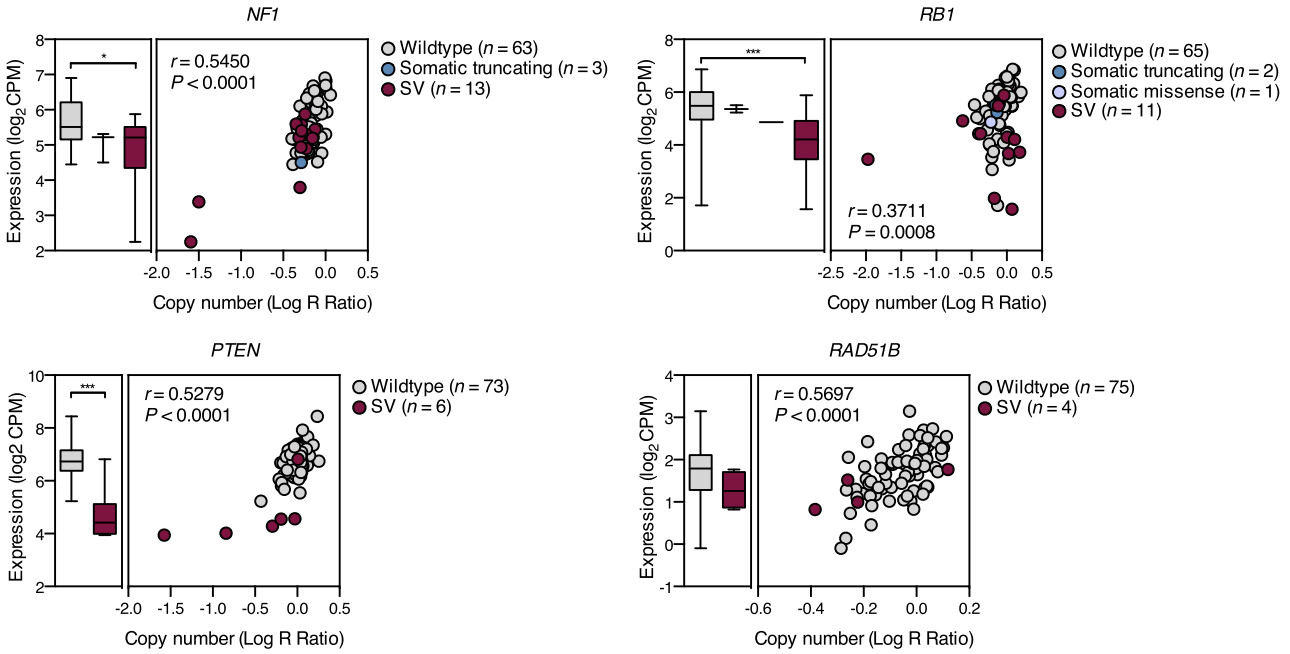
Extended Data Figure 1 | Patient cohort. **a**, Summary of whole-genome sequenced patients ($n = 92$) and samples ($n = 114$). **b**, Clinical characteristics of patients by clinical response group and acquired resistant cases without matching primary tumour material. *Primary tumour ungraded; diagnosis from ascites or pleural fluid. ^Time to progression or death measured from diagnosis. [a]Kruskal–Wallis, [b]Fisher or [c]log-rank test *P* values comparing primary tumour clinical groups reported. Median follow up time of cohort was 97.3 months.



Extended Data Figure 2 | Structural variants and somatic mutations. **a**, **b**, Scatter plot shows that the number of structural variants (**a**) and SNVs (**b**) detected per sample was not dependent on the purity of primary tumours ($n = 80$, grey) and ascites/relapse samples ($n = 34$, black). Spearman correlation P values indicated. **c**, Number of structural variants detected in primary tumours ($n = 80$) grouped by homologous recombination (HR) mutation status: HR wild-type, *BRCA1/2* altered, and HR deficient (lines indicate mean; $^{****}P < 0.0001$, Kolmogorov–Smirnov test). **d**, **e**, Whole-genome (**d**) and coding mutations (**e**) per megabase for samples ($n = 79$) stratified by *BRCA1*, *BRCA2* or homologous recombination pathway mutation status (lines indicate mean; $^{**}P < 0.01$, $^{***}P < 0.001$, Kolmogorov–Smirnov test). Sample AOCS-166 with germline mismatch repair mutation excluded from analysis (see Supplementary Information section 4.3).



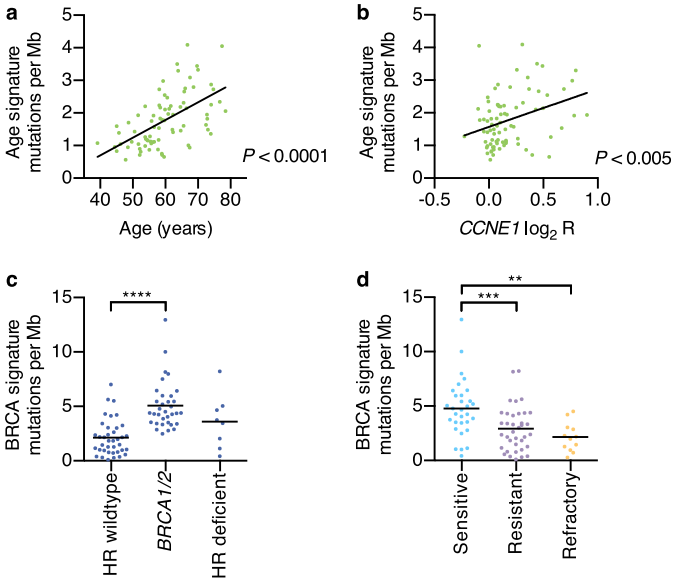
Extended Data Figure 3 | Chromothripsis and breakage-fusion-bridge. **a**, Chromothripsis affecting chromosome 13, including the *BRCA2* locus (arrow), for a primary chemotherapy-resistant tumour. A high number of breakpoints (structural variants) are observed with oscillations of copy number (CNA and LOGRR) indicating regions of retained heterozygosity (BAF) for a single haplotype. **b**, Breakage-fusion-bridge amplification is observed on chromosome 1 for a primary chemotherapy-resistant tumour. A cluster of breakpoints (structural variants), mostly inversions, on the distal p arm are associated with blocks of amplification (CNA). IGV review of the tumour WGS confirms that the telomere region has been lost.



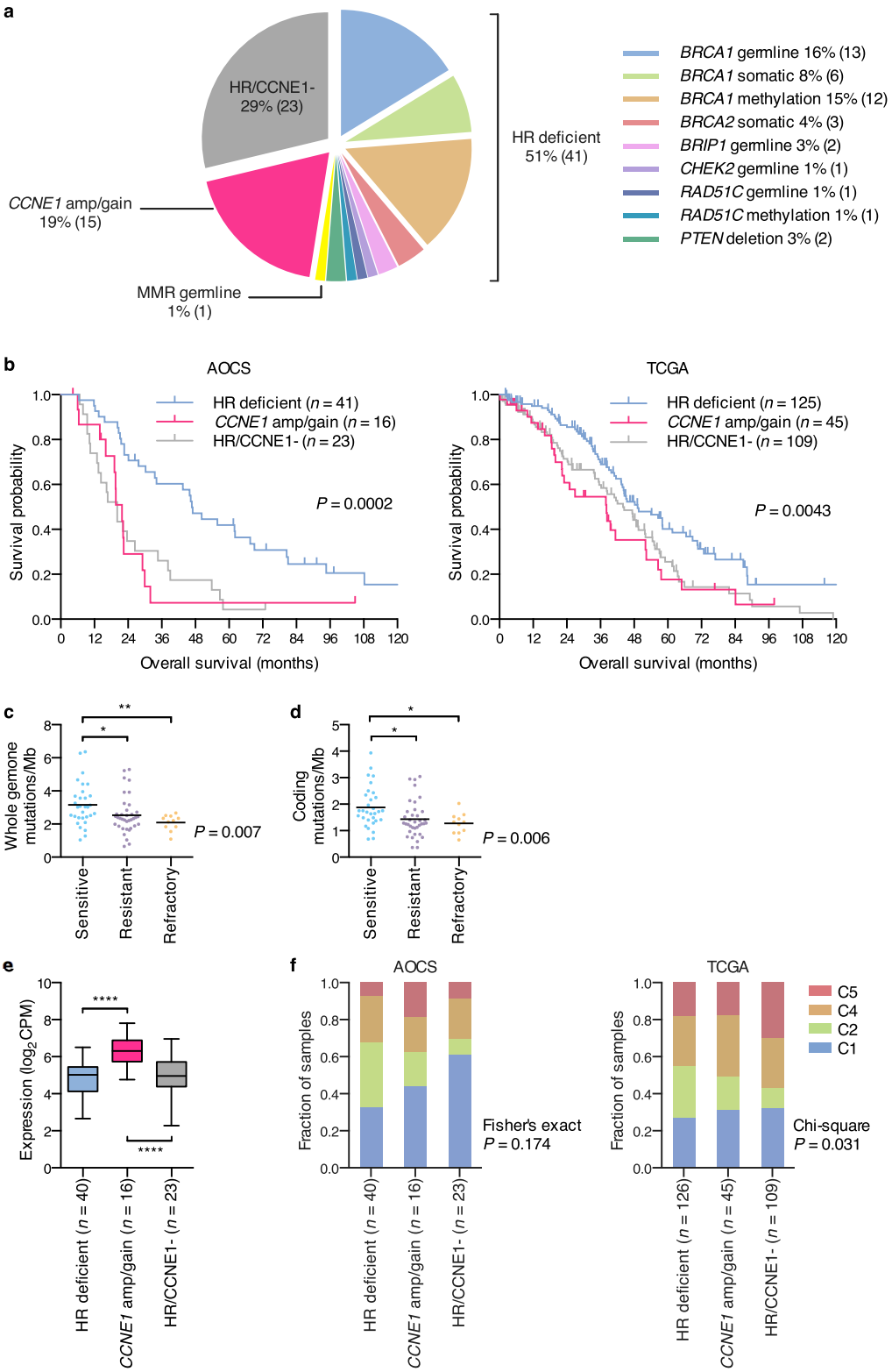
Extended Data Figure 4 | Expression of genes altered by structural variation. Scatter graphs show expression of *RB1*, *NF1*, *PTEN* and *RAD51B* plotted against copy number in primary tumours ($n = 79$; Spearman correlation analysis). Boxplots summarize expression by mutation type; lines indicate median and whiskers show range (two-tailed Mann-Whitney test; $*P < 0.05$, $***P < 0.001$). Samples with somatic interrupting structural variants and truncating mutations are indicated. Not all structural variants are associated with DNA loss and translocations can be observed to produce high expression counts for non-functional transcripts. Expression values lower than the median value for tumours and controls was observed for the majority of samples with structural variant events in *RB1*, *NF1*, *PTEN* and *RAD51B*.

Study ID	Age	Response group	HR pathway status	HR gene mutation	*CCNE1* status	*CCNE1* $\log_2$ R (CBS segment)	MMR signature mut/Mb	APOBEC signature mut/Mb	Age signature mut/Mb	BRCA signature mut/Mb	Dominant signature
AOCS-064	67.5	Sensitive	HRD	*BRCA1 meth*		-0.0	2.1	1.7	0.9	13.1	4
AOCS-094	59.0	Sensitive	HRD	*BRCA1 meth*	Gain	0.40	3.8	3.0	0.7	10.0	4
AOCS-106	64.6	Resistant	HRD	*RAD51C meth*		0.11	1.7	1.1	2.2	8.2	4
AOCS-108	61.2	Resistant	HRD	*GL BRCA1*		-0.0	2.3	2.3	2.1	8.2	4
AOCS-139	62.4	Sensitive	HRD	*GL BRCA1*		0.11	1.6	1.3	2.0	8.0	4
AOCS-091	39.2	Sensitive	HRD	*BRCA1 meth*		0.16	1.6	0.8	1.1	7.5	4
AOCS-090	53.8	Sensitive	WT			0.09	2.1	1.1	2.2	7.0	4
AOCS-171	52.1	Sensitive	HRD	*SM BRCA1*		0.04	1.8	1.5	0.7	6.5	4
AOCS-149	56.6	Sensitive	HRD	*SM BRCA2*		0.23	2.4	1.8	1.6	6.4	4
AOCS-145	62.2	Sensitive	HRD	*GL BRCA1*		0.15	1.3	1.1	1.5	6.0	4
AOCS-122	59.0	Sensitive	HRD	*SM BRCA2*		-0.0	0.9	1.0	1.8	6.0	4
AOCS-116	63.6	Resistant	WT		Gain	0.44	2.6	1.4	3.5	5.6	4
AOCS-112	70.0	Resistant	WT			0.26	1.2	1.1	3.3	5.5	4
AOCS-058	58.4	Resistant	HRD	*GL BRCA1*		0.07	0.6	1.2	1.8	5.5	4
AOCS-125	54.6	Sensitive	HRD	*BRCA1 meth*		0.16	1.2	0.6	1.2	5.4	4
AOCS-095	50.6	Sensitive	HRD	*GL BRCA1*		-0.0	1.0	0.9	1.4	5.4	4
AOCS-093	57.7	Sensitive	HRD	*BRCA1 meth*		0.04	1.3	1.4	0.9	5.2	4
AOCS-147	55.7	Sensitive	HRD	*RAD51C meth; SM PTEN del*		0.07	1.2	6.6	1.1	5.0	2
AOCS-152	73.0	Sensitive	HRD	*SM BRCA1*		0.04	0.9	1.5	1.5	5.0	4
AOCS-088	56.5	Sensitive	HRD	*GL BRCA1*		-0.0	0.7	0.7	1.1	4.8	4
AOCS-170	72.8	Sensitive	HRD	*GL CHEK2*		0.17	2.3	1.6	1.9	4.6	4
AOCS-158	52.5	Refractory	HRD	*BRCA1 meth*		0.02	0.7	0.9	0.8	4.5	4
AOCS-114	55.3	Resistant	HRD	*GL BRCA1*		0.32	0.6	0.8	1.3	4.4	4
AOCS-128	45.1	Resistant	HRD	*BRCA1 meth*		0.02	0.7	1.9	0.8	4.4	4
AOCS-107	47.9	Resistant	HRD	*BRCA1 meth*		0.01	1.4	0.8	0.6	4.4	4
AOCS-113	77.4	Resistant	WT			-0.1	1.5	1.2	4.1	4.3	4
AOCS-105	73.9	Resistant	HRD	*GL BRCA1*		0.05	0.9	0.7	1.4	4.3	4
AOCS-065	46.3	Sensitive	HRD	*GL BRCA1*		0.01	0.8	0.4	1.0	4.3	4
AOCS-146	52.1	Sensitive	HRD	*GL BRCA1*		0.10	0.8	1.0	0.9	4.3	4
AOCS-153	57.9	Refractory	WT			0.03	0.9	0.8	1.2	4.2	4
AOCS-056	72.5	Resistant	WT			0.03	0.7	0.6	1.6	4.1	4
AOCS-131	67.1	Sensitive	HRD	*GL BRCA1*		0.14	0.4	3.1	1.0	4.0	4
AOCS-034	51.8	Sensitive	HRD	*GL BRCA1*		0.27	1.1	1.4	1.0	3.9	4
AOCS-148	45.7	Sensitive	HRD	*SM PTEN del*		0.10	0.7	0.9	1.4	3.8	4
AOCS-092	68.7	Sensitive	HRD	*SM BRCA1 del*		0.20	1.5	0.8	2.8	3.5	4
AOCS-086	64.7	Sensitive	HRD	*SM BRCA1*		0.01	0.9	0.9	1.5	3.5	4
AOCS-126	58.9	Sensitive	HRD	*GL RAD51C*		-0.0	0.7	0.9	0.8	3.4	4
AOCS-104	48.5	Resistant	WT			0.04	0.8	0.9	0.9	3.4	4
AOCS-063	62.3	Resistant	HRD	*SM BRCA2*		-0.0	0.8	0.9	1.4	3.4	4
AOCS-079	72.8	Resistant	HRD	*SM BRCA1*		0.02	0.7	0.9	1.8	3.4	4
AOCS-164	55.1	Resistant	HRD	*BRCA1 meth*	Gain	0.34	0.8	1.3	0.9	3.3	4
AOCS-057	56.1	Resistant	HRD	*BRCA1 meth*		0.13	1.0	0.6	1.1	3.3	4
AOCS-096	61.6	Resistant	WT			0.14	1.6	0.5	2.2	3.3	4
AOCS-004	53.8	Refractory	WT		Gain	0.57	1.7	0.6	1.3	3.0	4
AOCS-078	59.9	Resistant	WT		Gain	0.43	1.6	2.1	2.1	3.0	4
AOCS-130	65.6	Sensitive	HRD	*SM BRCA1*		0.07	0.7	1.7	1.6	2.9	4
AOCS-001	53.8	Refractory	HRD	*BRCA1 meth*		-0.0	0.7	1.4	1.0	2.9	4
AOCS-144	54.4	Sensitive	HRD	*BRCA1 meth*		0.08	0.4	1.3	1.2	2.8	4
AOCS-160	66.1	Resistant	WT			0.06	0.5	0.6	2.8	2.6	3
AOCS-143	49.7	Sensitive	HRD	*GL BRCA1; RAD51C meth*		0.16	0.5	0.2	1.1	2.5	4
AOCS-083	58.4	Resistant	WT			0.25	1.3	0.7	2.2	2.4	4
AOCS-165	74.0	Resistant	WT			0.08	0.4	0.5	2.9	2.4	3
AOCS-080	67.6	Refractory	WT		Gain	0.36	0.8	0.7	3.1	2.3	3
AOCS-168	59.8	Refractory	WT			-0.1	0.6	1.6	1.8	2.2	4
AOCS-081	59.8	Resistant	WT		Gain	0.54	1.3	0.6	2.5	2.2	3
AOCS-076	58.1	Resistant	WT			0.08	1.0	0.6	2.2	2.2	3
AOCS-115	56.4	Resistant	HRD	*SM PTEN del*	Amplified	0.78	0.6	0.3	1.9	2.0	4
AOCS-161	60.6	Refractory	WT			-0.0	0.8	1.2	2.5	2.0	3
AOCS-163	65.6	Resistant	WT			0.31	0.5	0.6	2.3	1.8	3
AOCS-061	59.3	Resistant	WT		Amplified	0.77	0.7	0.5	3.0	1.8	3
AOCS-111	65.8	Resistant	WT		Gain	0.46	0.8	0.2	2.7	1.5	3
AOCS-084	60.9	Refractory	WT			0.15	1.0	0.6	2.0	1.5	3
AOCS-159	78.4	Resistant	WT			0.01	0.7	0.6	2.1	1.3	3
AOCS-169	74.1	Refractory	WT			0.08	0.5	0.4	2.3	1.3	3
AOCS-005	51.6	Resistant	WT		Amplified	0.90	0.9	1.6	1.9	1.2	3
AOCS-133	66.8	Sensitive	HRD	*GL BRIP1*		0.31	0.6	0.4	4.1	1.1	3
AOCS-123	55.8	Sensitive	WT			-0.0	0.5	0.2	1.8	1.1	3
AOCS-124	58.1	Sensitive	WT		Amplified	0.69	0.7	0.4	2.7	1.0	3
AOCS-075	54.0	Resistant	WT			0.27	0.4	2.3	1.2	1.0	2
AOCS-162	74.7	Refractory	WT		Gain	0.37	0.7	0.5	2.7	1.0	3
AOCS-109	63.9	Resistant	WT		Amplified	0.80	0.7	0.5	3.3	0.8	3
AOCS-157	45.1	Resistant	WT			0.20	0.3	0.2	1.6	0.8	3
AOCS-166	58.0	Refractory	WT	*MMR mutation*	Amplified	0.67	34.4	1.1	1.5	0.7	1
AOCS-085	44.5	Resistant	WT			0.05	0.4	0.2	1.0	0.6	3
AOCS-132	48.7	Sensitive	HRD	*GL BRIP1*		0.10	0.3	0.2	1.7	0.4	3
AOCS-059	76.7	Resistant	WT			0.00	0.6	0.4	2.3	0.4	3
AOCS-055	59.6	Resistant	WT		Gain	0.55	1.2	12.7	0.7	0.4	2
AOCS-097	55.5	Resistant	WT			0.07	0.2	0.1	1.1	0.3	3
AOCS-077	66.5	Refractory	WT		Gain	0.50	0.3	0.3	2.1	0.3	3
AOCS-060	69.7	Resistant	WT		Gain	0.49	0.3	0.6	3.4	0.1	3

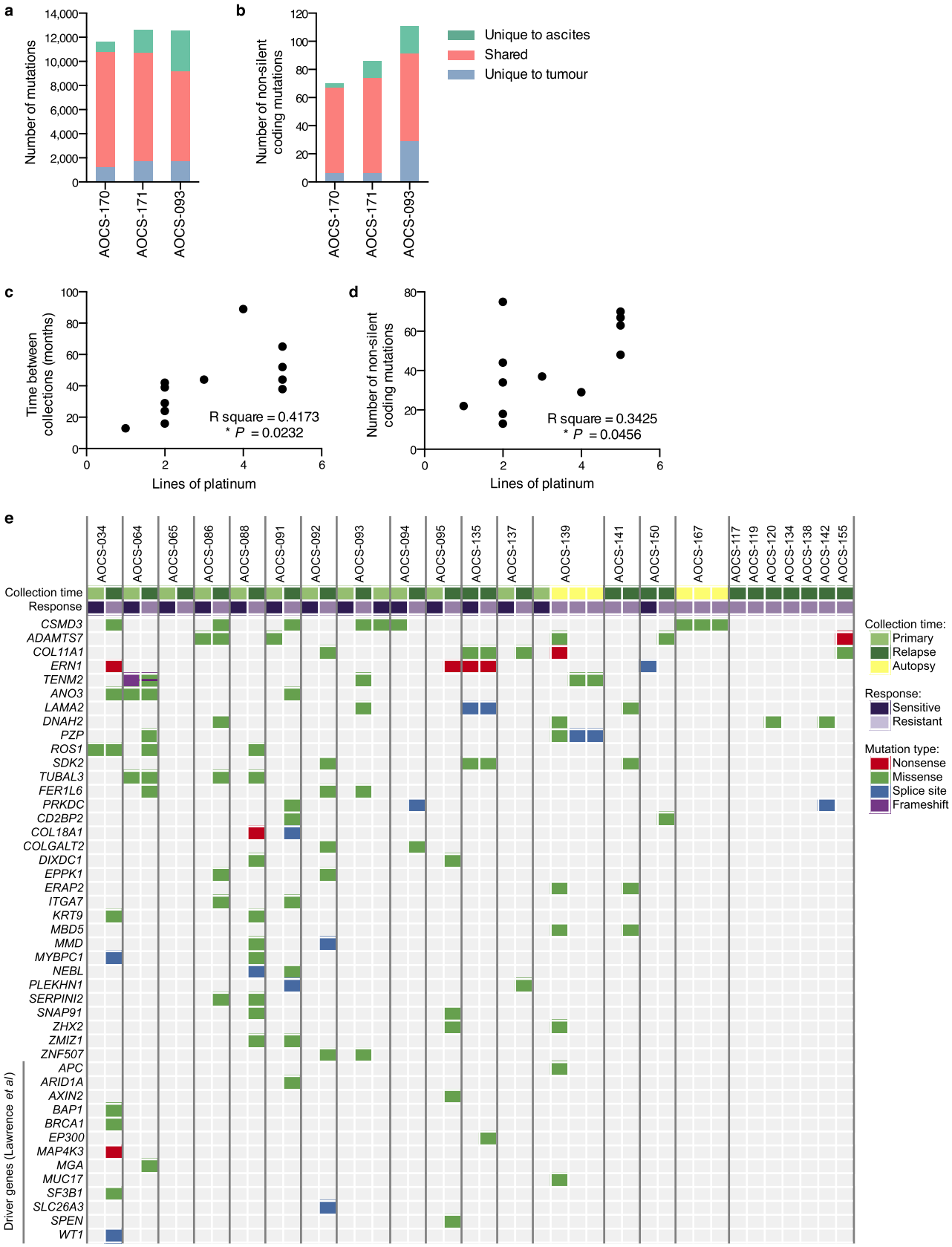
Extended Data Figure 5 | Molecular characteristics of primary tumours. Clinical and molecular characteristics of primary tumours sorted by BRCA signature contribution ($n = 80$).



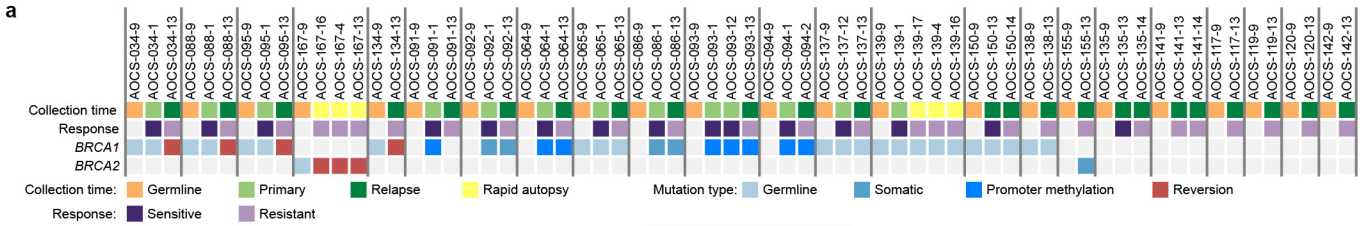
Extended Data Figure 6 | Mutational signature associations. **a**, **b**, Age mutational signature association with age at diagnosis (**a**) and *CCNE1* copy number (**b**) (Spearman correlation *P* value reported, $n = 80$). **c**, **d**, BRCA mutational signature association with *BRCA1/2* and other homologous recombination pathway mutations (**c**) and primary treatment response (**d**) (lines indicate mean; Kolmogorov–Smirnov $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$; $n = 80$).



Extended Data Figure 7 | Molecular drivers and clinical associations. **a**, Percentage (*n*) of primary tumours (total $n = 80$) affected by homologous recombination pathway mutations and *CCNE1* copy number gains. One driver mutation counted for samples with more than one change, ranking mutations in *BRCA1/2*, followed by other germline, somatic, amplification, deletion and methylation events respectively. **b**, Association of driver mutation subgroup with overall survival in AOCS and TCGA cohorts (Kaplan–Meier analysis, *P* value calculated by Mantel–Cox log-rank test). **c**, **d**, Whole genome (**c**) and coding mutations (**d**) per megabase for samples stratified by primary clinical response group (lines indicate mean). Kruskal–Wallis test *P* value reported (*$P < 0.05$, **$P < 0.01$, Kolmogorov–Smirnov test). **e**, Boxplots summarize *CCNE1* expression in different driver mutation subgroups (****$P < 0.0001$, unpaired two-tailed *t*-test). Middle bar, median; whiskers, data range. **f**, Proportions of gene expression molecular subtypes between driver mutation subgroups. HR/CCNE1− subgroup has no detected homologous recombination pathway mutations or *CCNE1* copy number changes.



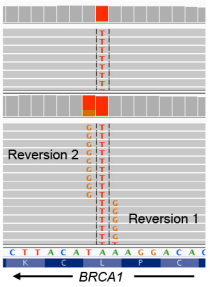
Extended Data Figure 8 | Analysis of acquired resistance cases. **a**, **b**, Matched primary ascites share most variants with primary tumour samples across the whole genome (**a**) and for non-silent coding mutations (**b**). **c**, **d**, Significant correlations are observed between the time between collection of the primary and relapse samples and the number of lines of platinum treatment the patient received (Pearson correlation, $P = 0.0232$, two-tailed) (**c**), and between the number of non-silent coding mutations unique to the relapse samples and the number of lines of platinum treatment the patient received (Pearson correlation, $P = 0.0456$, two-tailed) (**d**). **e**, Genes with recurrent non-silent coding mutations that are unique to the second sample in the acquired resistance cohort or that are driver genes reported previously[55].

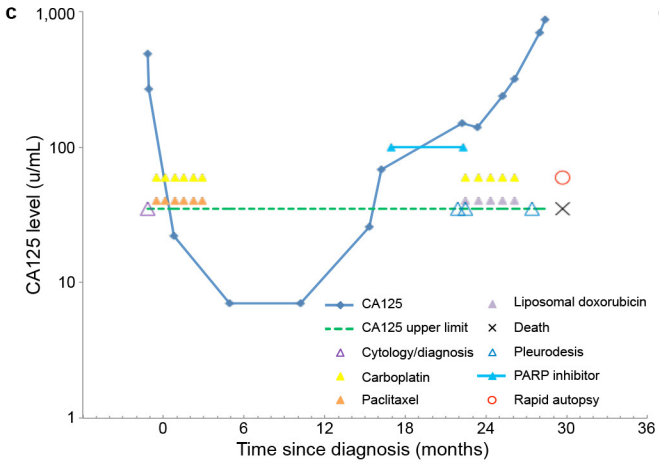


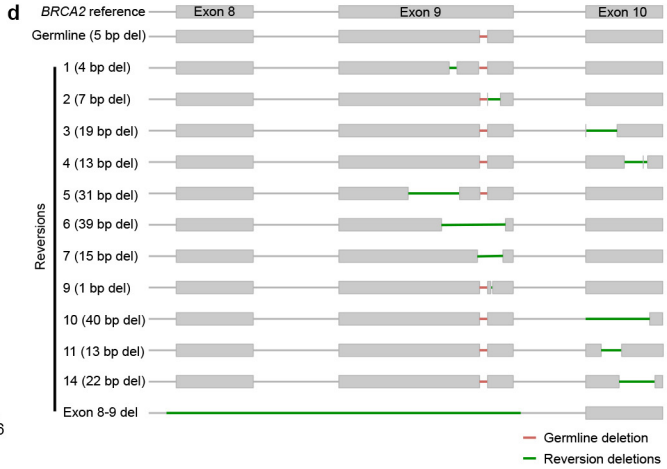
b

AOCS-034	Position	Mutation†	Codon‡	Amino acid	DNA variant allele ratio	
					Primary	Relapse
BRCA1 reference			TTA	Leucine		
Germline mutation	17:41258497	A>T	TAA	Stop	0.90	0.93
Reversion 1	17:41258498	A>G	CAA	Glutamine	0.00	0.32
Reversion 2	17:41258496	T>G	TAC	Tyrosine	0.00	0.09

† Mutation identified on plus strand; ‡ Codon encoded on minus strand





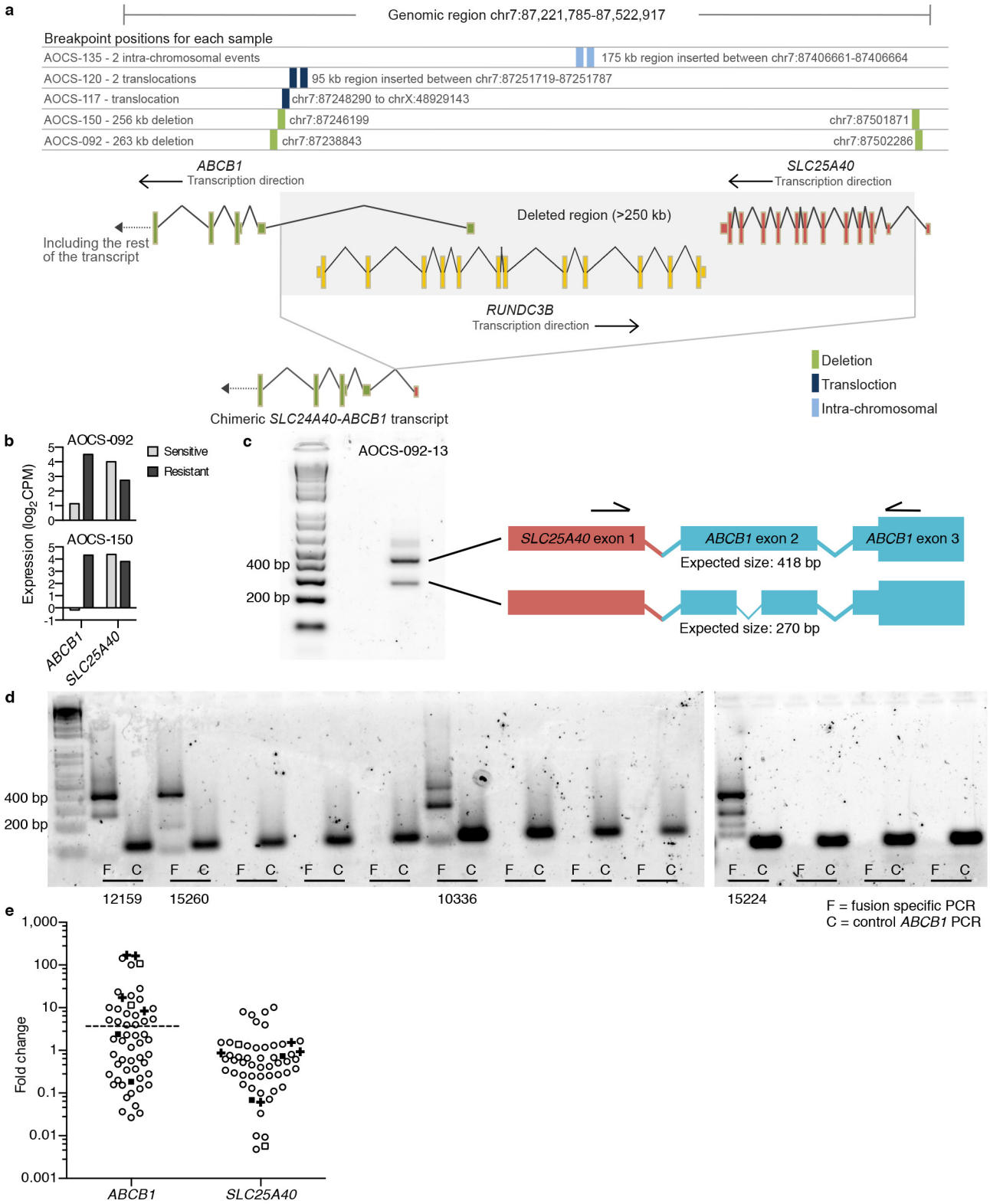


e

AOCS-167 relapse samples	Germline mutation 5 bp del	Reversions 1 4 bp del	2 7 bp del	3 19 bp del	4 13 bp del	5 31 bp del	6* 39 bp del	7* 15 bp del	8*^ 6 bp del	9 1 bp del	10 40 bp del	11 13 bp del	12^ 1 bp del	13^ 25 bp del	14 22 bp del	15^ 10 bp del
Small bowel mesentry	0.722	0.003	0.165	0.033						0.006						
Small bowel wall	0.883	0.246	0.263	0.003												
Large bowel wall	0.772	0.034	0.239	0.002												
Greater omentum	0.822	0.040 †	0.133 †	0.114 †						0.008				0.001		
Caecum	0.811	0.048	0.236	0.059						0.021						
Transverse colon	0.785	0.037	0.067	0.008						0.041						
Descending colon	0.768	0.093	0.040	0.044						0.023						
Appendix	0.689	0.032	0.0001 ‡	0.070						0.010		0.002				
Para-aortic LN	0.880					0.759		0.058								
Sub-carinal	0.651				0.469											
Omental mass	0.743						0.003		0.001		0.001				0.088	0.001
Left diaphragm	0.715	0.011	0.128	0.017						0.011		0.008	0.001			
Right pleural	0.572															
Left pleural	0.916				0.986											
Left hilum LN	0.773				0.950											
Right hilum LN	0.500															
Pericardial LN	0.909				0.927 †											
Ascites	0.727	0.033 †	0.127 †	0.131 †	0.002					0.011		0.003				

* Encompasses deletion; † Confirms observations in WGS; ‡ Not called but observed in IGV; ^ Low confidence event

Extended Data Figure 9 | *BRCA1/2* reversion events. **a**, *BRCA1* and *BRCA2* mutations are indicated by sample for acquired resistance cases ($n = 23$). Of the ten cases with germline mutations, five show secondary somatic mutations in the relapse-resistant samples that are proposed to restore gene function (that is, reversion). **b**, Two independent point mutations revert the *BRCA1* nonsense mutation in AOCS-034. **c**, CA125 serum marker profile for case AOCS-167. **d**, Schematic of 12 high confidence reversion events identified in AOCS-167. **e**, Allele frequency of fifteen *BRCA2* reversion events identified in deep amplicon sequencing across 18 relapse samples from case AOCS-167. Reversions observed in single deposits at low allele frequencies (≤0.001) were considered low confidence events.



Extended Data Figure 10 | *SLC25A40-ABCB1* fusion transcript. **a**, Schematic showing the locations of structural variants identified upstream or internal to *ABCB1* from WGS. **b**, *ABCB1* and *SLC25A40* expression levels in sensitive and resistant samples of AOCS-092 and AOCS-150. **c**, RT–PCR verification of *SLC25A40-ABCB1* fusion transcript expression in AOCS-092 and schematic of expected RT–PCR products. **d**, RT–PCR results from a validation cohort of relapse ascites identifying four additional samples with the fusion transcript ($n = 51$). **e**, *ABCB1* and *SLC25A40* expression levels compared to the SKOV3 cell line by qRT–PCR in the validation cohort and selected cases. Relapse samples with the fusion are indicated by open squares and the matched primary samples indicated by closed squares; validation cases with the fusion are indicated by '+'; the median plus median absolute deviation expression level is indicated by the dotted line ($n = 56$).