Phylogenetic quantification of intra-tumour heterogeneity predicts time to relapse in high-grade serous ovarian cancer

Roland F. Schwarz[1,†] Charlotte K.Y. Ng[1,2,†]
Susanna L. Cooke[1,†] Scott Newman[1,†] Jillian Temple[1]
Anna M. Piskorz[1] Davina Gale[1] Karen Sayal[1]
Muhammed Murtaza[1] Peter J. Baldwin[3]
Nitzan Rosenfeld[1,2] Helena M. Earl[2] Evis Sala[4]
Mercedes Jimenez-Linan[3] Christine A. Parkinson[1,2]
Florian Markowetz[1,2,*] James D. Brenton[1,2,*]

Affiliations

[1] Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge, CB2 0RE, UK

[2] European Bioinformatics Institute, EMBL, Hinxton, Cambridge, CB10 1SD, UK

[3] Department of Oncology, University of Cambridge, Hutchison/MRC Research Centre, Hills Road, Cambridge, CB2 0XZ, UK

[4] Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Hills Road, Cambridge, CB2 0QQ, UK

[5] University Department of Radiology, Box 218, Level 5, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 2QQ, UK

[†] Current affiliations: RFS, European Bioinformatics Institute, Hinxton; CKYN, Memorial Sloan-Kettering Cancer Institute, New York; SLC, The Sanger Institute, Hinxton; SN, Department of Human Genetics, Emory University School of Medicine.

[*] To whom correspondence should be addressed: florian.markowetz@cruk.cam.ac.uk, james.brenton@cruk.cam.ac.uk

Abstract

Background

The major clinical challenge in the treatment of high-grade serous ovarian cancer (HGSOC) is the development of progressive resistance to platinum-based chemotherapy. We hypothesised that intra-tumor genetic heterogeneity (ITH) resulting from clonal evolution within the cancer could provide a mechanism for the development of resistance.

Methods and Findings

We performed high-resolution whole genome copy-number (CN) profiling and selected genome-wide sequencing of 138 spatially and temporally separated samples from 17 cases. We investigated the relationship between ITH and patient outcome using new methodology for evolutionary inference and phylogenetic quantification of ITH from CN profiles. We show that ITH in HGSOC is driven by ongoing clonal evolution with fully branched evolutionary trajectories that do not have clock-like evolutionary rates. We show that quantative measures of clonal expansion and temporal heterogeneity were the strongest predictor of progression-free survival. We show that clinical relapse in 2 subjects arose from clonal expansion of minor subclones present prior to therapy.

Conclusions

Quantative measures of ITH may have predictive value for outcomes after chemotherapy treatment. Rare subclonal populations can be demonstrated in tumour tissue at diagnosis in HGSOC and may undergo expansion during chemotherapy.

Introduction

Intra-tumoral genetic heterogeneity (ITH) in cancer has been investigated for almost half a century [14, 32] and recent advances in genomic technology have precisely demonstrated highly diverse genetic changes within epithelial cancers [7, 16, 20, 21, 25, 28, 29, 37, 38, 42, 43]. Multiple sampling of primary and metastatic samples in breast, pancreas and renal carcinoma has catalogued genetic divergence and confirmed metastasis-to-metastasis spread [7, 16, 28, 29, 37]. Deep sequencing of tumours has revealed the clonal compositions of individual clinical samples and shown how major and minor subclones co-exist within epithelial tumours [7, 16, 29, 31, 37].

It is hypothesized that tumors with marked ITH may be chemotherapy resistant as repeated selective sweeps following chemotherapy lead to clonal expansions of minor subclones with a fitness advantage. A tumour that exhibits sufficient ITH is thought to explore the fitness landscape widely enough to evade selection pressure from chemotherapy, and repopulate with a resistant subclone [18, 27]. However the clonal expansion hypothesis has not been widely tested in clinical settings, mainly due to difficulties in routinely acquiring genomic profiles from the same patient at different sites and time

points before and after chemotherapy. Additionally, algorithms for quantifying ITH and measuring the degree of clonal expansion have only recently been established [36].

To address whether ITH is predictive of outcome in HGSOC we performed prospective tissue collection in the CTCR-OV03 and -OV04 studies [34]. It is especially important to understand ITH in HGSOC as its clinical presentation is typically with extensive metastatic disease, involving multiple implantation sites throughout the abdomen, which enables multiple sampling from spatially-distinct sites. Neo-adjuvant chemotherapy [41] allows acquistion of samples before treatment (at biopsy) and after treatment (at surgery), in addition to relapse. In contrast to other epithelial cancers, HGSOC is highly sensitive to chemotherapy with carboplatin or cisplatin at initial treatment. The probability of a secondary response to platinum-based chemotherapy is proportional to the time interval from the initial therapy. Patients who relapse in under 6 months are clinically classified as platinum-resistant and those who relapse after 12 months are termed platinum-sensitive [24], making progression free survival a potential proxy for tumour fitness. Finally, HGSOC shows particularly high ITH on the level of genome rearrangements, leading to profound copy-number (CN) changes between clonal populations [8, 17]. Routine CN profiling techniques, such as SNP arrays, are therefore a reliable standard for collecting genomic information from the dominant clones inhabiting these spatially and temporally distinct lesions.

On this background, the OV03/04 study was established to validate the clonal expansion hypothesis in a clinical setting and investigate potential uses for routine diagnostics. Specifically, our aims were to: i) Reconstruct the evolutionary history of disease within each patient from whole-genome CN profiles. ii) Use the reconstructed evolutionary trajectories to quantify ITH in a systematic manner. In particular, estimate *temporal heterogeneity*, i.e. the amount of change in the course of treatment, and the *clonal expansion* potential of the tumour. iii) Use these indices to investigate the connection between ITH and tumour fitness, thereby exploiting the HGSOC-specific correlation between progression free survival and chemotherapy resistance [24]. iv) Validate whether relapse is a clonal expansion of a minor subclone and trace its origin in time.

To facilitate these four goals we developed MEDICC, a method for evolutionary inference and phylogenetic quantification of ITH from allele-specific copy-number profiles [36]. This is the first application of MEDICC to a large clinical study. Our results demonstrate that the clonal expansion potential and mutational rate of a tumour predict time to relapse across all patients. We provide statistical and experimental evidence that emergence of relapse is an early branching event, followed by divergent evolution and clonal expansion from a low-prevalence subclone that is present prior to therapy. We further show that HGSOC generally evolves and spreads in a tree-like branching process with mostly non-homogenous (i.e. non clock-like) rates of evolution, further supporting the idea that disease progression is shaped by selective sweeps on a cell population where most cells retain their tumorigenic potential. This is to our knowledge the first rigorous attempt to validate the clonal expansion hypothesis in a clinical setting. Our results suggest that, should sufficient samples be available, ITH can be used to predict patient outcome in HGSOC and therefore has the potential to act as a patient-specific prognostic indicator in routine clinical diagnostics.

Results

We prospectively collected temporally and spatially distinct HGSOC samples from patients undergoing neo-adjuvant platinum-based chemotherapy (Fig. 1A). CN profiles from 138 tumor samples from 17 patients were obtained with GenomeWide Affymetrix SNP6.0 arrays (Table 1) and segmented using PICNIC [19]. To exclude histological differences and different mutational profiles in key tumour suppressors as sources for any observed CN differences, we performed tagged-amplicon resequencing for genes mutated in HGSOC and a detailed histopathology review of all samples. 15/17 cases had a mutation in *TP53* consistent with HGSOC (Table 2) [1]. Of the 2 wild-type *TP53* cases, OV03-23 had strong nuclear p53 protein accumulation consistent with p53 dysfunction and OV03-04 was reclassified as a metastatic endometrioid ovarian cancer. We decided to keep it as part of this study for comparison and discuss it separately in the sections below if the results deviate from the rest of the cohort. None of the cases had a germ line mutation in *BRCA1* or *BRCA2*. OV04-20 had a somatic nonsense mutation in *BRCA2* and OV03-02 showed missense mutations of uncertain significance in *BRCA1* and *BRCA2* (Table 2). No significant differences in morphology or growth pattern were observed between metastatic sites (Fig. S1). We examined the CN profiles for evidence of known mutator phenotypes, but did not find any cases with features of tandem duplicator phenotype [26, 30].

All subsequent analyses were performed using whole genome bi-allelic integer CN profiles. 14 cases had ≥ 3 samples and were suitable for evolutionary analysis. Due to the neo-adjuvant treatment, 10 cases had samples from both biopsy prior to chemotherapy treatment and at interval debulking surgery, allowing for comparison of different time points. 2 cases included relapsed ascites samples (Table S1, for a detailed case by case description, see the Supplementary Materials).

Because ITH is the result of ongoing evolutionary change in the patient [25], reconstructing the evolutionary history of the disease within each patient is the first necessary step prior to quantifying ITH. To this end, we recently developed *MEDICC* [36], a method that employs a minimum evolution criterion to measure the genetic divergence between CN profiles, phases alleles, reconstructs evolutionary trees and ancestral genomes. The evolutionary distance between samples is estimated based on the minimum number of segmental amplifications and deletions needed to transform one genomic profile into another using optimized allele-specific assignment of major and minor CNs. Using *MEDICC* we reconstructed evolutionary trees for all 14 cases with more than 3 samples (Fig 1A). Circos plots, evolutionary trees and clinical histories for all cases can be found in the Supplementary Materials.

Patient-specific CN changes cluster anatomical sites

As expected CN alterations were highly patient-specific. An evolutionary tree reconstructed using all genomic profiles across all patients irrespective of actual evolutionary relatedness correctly clusters patients (Fig. 1B). However, no clustering into resistant or sensitive subgroups was evident at this stage of the analysis (outer color bar, Fig 1A).

The majority of cases showed strong overall divergence with a median of 77 (s.d. = 46) genomic events between samples from each case and there was clear intratumoral divergence from the diploid (median of 133 events, s.d. = 70). Three cases showed less marked changes; reconstruction of events for OV03-01 and -17 had limited phylogenetic signal owing to low divergence or limited numbers of samples, and in OV03-01we found very high ITH and strong clonal expansion (see in-depth discussion below).

HGSOC is known to spread by physical shedding of cells from the ovaries. Interestingly, we often observed co-clustering of anatomical sites (spatial heterogeneity) in the patient-specific evolutionary trees (Fig. 3D and individual trees in the Supplementary Materials). For example, OV03-20 showed clear separation of omentum and small bowel mesentery samples (Fig. 3D) with right paracolic gutter and peritoneum as the respective outgroups. Genetic markers of this divergent evolution included Chr. 2q and 3p as well as amplifications on Chr. 10 (see Fig. 3C). Paired-end whole genome sequencing on samples from 6 cases confirmed divergent genetic change at higher resolution (Fig. S18). For example in OV03-20, there were three deletion breakpoints and an insertion breakpoint that were only present in the omentum sample and three deletion breakpoints that were only in the peritoneal sample (Fig. S18).

ITH is driven by ongoing clonal evolution after metastasis with changing evolutionary rates

The observed spatial heterogeneity in many patients lead us to ask whether metastases themselves are capable of spreading, or if alternatively HGSOC metastasizes from the primary site only. Because metastasis requires rapid proliferation and is closely linked to the evolutionary process, we can answer this question from the pattern of the evolutionary relationships between samples within each patient. Given a fixed number of metastases, two scenarios are possible: If only a low percentage of the primary tumor mass can metastasize, with high probability most of the metastases present must have originated from the primary tumour, leading to an evolutionary star topology (Fig. 2A). By contrast, if a high percentage of cells retain their metastatic potential after metastasis, ongoing spread (and the associated genetic change) will lead to a fully branched evolutionary tree (Fig. 2B).

We used *MEDICC* to test each patient for the null hypothesis that the pairwise distances were derived from a star topology [36]. To verify and visualize the findings we applied the neighbour-net method [6] which captures non-tree-like evolutionary signals in distance data. From 9 cases with > 3 samples, 8 showed significant branching ($p < 0.05$; chi-squared test for goodness-of-fit with Benjamini and Hochberg correction for false discovery rate). The remaining case (OV03-22) showed only a weak tree structure and the null hypothesis could not be rejected ($p = 0.22$, Fig. 2A). We additionally investigated with *MEDICC* whether the reconstructed evolutionary trees are clock-like, i.e. show a constant rate of evolution across the whole tree. After correction for multiple testing, 4 out of 14 patients (28.6%) showed significant (0.05 alpha level) non-clock like evolutionary trajectories. We conclude that in HGSOC metastases retain the ability to spread and ITH is generated through ongoing clonal evolution [21] with potentially

unknown mutator phenotypes present.

HGSOC shows genomic change during treatment

As most metastases are established before onset of treatment, the immediate next question addresses the speed of this ongoing clonal evolution and if it has a significant impact on the ability of the tumour to adapt to selective sweeps from chemotherapy. Due to the neo-adjuvant treatment we had samples available from biopsy and surgery, between which the patient receives typically 3 cycles of chemotherapy. Time between initial treatment and surgery was on median 2 months (between 1 and 4 months in our study with most patient receiving surgery after 2 months). We used this stratification of samples and the algorithms implemented in *MEDICC* to quantify the average genomic change during treatment (temporal divergence index, TD). To exclude that differences in cellularity bias estimation of evolutionary change in this time frame, we compared cellularity estimates from histopathology between the two strata and found no significant differences (Welch Two-Sample T-test, p-value = 0.702).

MEDICC measures temporal divergence by mapping genomes into a high-dimensional space (the "mutational landscape" [18, 27]) in which distances correspond to evolutionary distances between genomes. It then calculates the TD index as the distance between the robust centers of mass of biopsy and surgery samples [36], leading to a robust estimate of change during treatment. An average of 46 (s.d. = 13) new genomic events occurred between biopsy and surgery samples. For example, for OV03-10 *MEDICC* measured a profound difference between the two sample subgroups (TD index 0.66) with early and long branching of the pre-treatment omental biopsy sample away from the remaining omental samples, indicating divergent evolution (Fig. 3A and 3B). The CN events responsible for this divergence included deletions on Chr. 1p, 1q, 3p, 7q, 9q and 11p. In OV03-20, one of the three omentum samples differed in 18.1% of its genome from the omentum samples at surgery (Fig. 3C and 3D). Even though *MEDICC* detected significant genomic differences between biopsy and surgery, visual inspection of the Circos plots showed strong overall conservation, indicating that the main karyotypes for each case were established before onset of treatment (Fig. 3A and 3C and Figs. S3–S16). We concluded that HGSOC shows ongoing evolutionary change in a time frame of about 2 months which is small but significant compared to the total amount of change since onset of the disease.

HGSOC is freqently polygenomic and exhibits high potential for clonal expansion

It has previously been shown in breast cancer that clonal expansions of minor subpopulations of cells lead to so called *polygenomic* tumours, while others appear *monogenomic* [28]. These changes are potentially modulated by selection pressure from chemotherapy or other factors and might have prognostic value. *MEDICC* allows statistical quantification of the degree of clonal expansion on a continuous scale (CE index) by testing for local spatial clustering of genomes in the mutational landscape [36]. A value of CE > 1

thereby indicates detectable clonal expansion, whereas CE < 1 suggests little to no clonal expansion.

Compareable to breast cancer, we found the CE index to be variable across the cohort (median 0.73, IQR $0.65 - 1.24$) with a bimodal distribution, in which cases have either CE < 1 or CE > 1 with a clearly visible margin between the two groups (Fig. S19). This bi-partition of cases is also visible in the evolutionary trees. Cases with low CE, for example OV03-22, showed linear emission of samples and rather homogenous branch lengths (Fig. S12). By contrast, OV03-08 and OV03-17 showed strong clonal expansions (CE index 1.47 and 2.24) with multiple samples in strongly diverging subclades (Fig. 5).

Genomic change during treatment and clonal expansions predict resistance development

In order for a tumour to overcome selection pressure by chemotherapy it needs to be able to explore the mutational landscape efficiently [18]. Polygenomic tumours with a high mutation rate that have already undergone clonal expansions are likely to be at an advantage. We therefore investigated if our measures of ITH, the temporal divergence and clonal expansion indices, can predict progression-free survival (PFS), as this is the key determinant of clinical outcome and therefore the closest clinical measure of tumour fitness.

PFS was only weakly correlated with overall divergence (the total length of the evolutionary tree, cor. test $p = 0.06$). In a multivariate linear model we attempted to predict PFS from the CE and TD indices and all other clinical covariates available (age, stage, CA125 response, RECIST scores) and found only TD and CE to be significant (TD, normalized log coefficient -0.84, $p = 0.00005$; CE -0.56, $p = 0.000365$ respectively, regression $R^2 = 0.84, p = 0.000122$). CE and TD indices were strong predictors of PFS (Fig. 4) with platinum-resistance relapse occurring in cases with high clonal expansion and profound changes from biopsy to surgery. All cases with significantly increased CE indices (> 1) had less than 12 months PFS and together with TD succesfully divided the cohort into resistant and sensitive subgroups (Fig. S17). The sample size was not significantly correlated with PFS (cor. test $p = 0.25$) and did not change the regression model (coefficient specific significance $p = 0.47$). To ensure that the correlation was not dominated by the large Cook's distances of OV03-20 and OV03-13, we repeated the regression analysis excluding these two patients ($p = 0.01083$, $R^2 = 0.6$), and performed a separate rank-based analysis with all patients ($p = 0.00407$). We validated these results by using the TD and CE indices as predictors in a Cox proportional hazard model and found both coefficients to be significant (both $p < 0.05$), whereas other clinical covariates did not show significant coefficients. It is interesting to note that OV03-04, the endrometrioid cancer we kept for comparison, is also a clear outlier in the analysis.

Resistant subclones are embedded in presentation disease

That a relatively small number of changes from biopsy to surgery had a significant impact on PFS suggested that drug resistant clones arose early in evolution with subsequent

acquisition of mutations that increased their relative fitness. We were able to investigate this prediction in 2 cases with platinum-resistant relapsed disease (OV03-08 and OV03-17).

The phylogenetic reconstruction for case OV03-17 showed that the ovarian biopsy was derived early in tumour evolution but the relapsed clone showed rapid divergence (Fig. 5). The CN profiles of the relapsed ascites sample revealed a focal deletion in the tumor suppressor gene *NF1*, but not in the ovarian and omental samples (Fig. 5). *NF1* shows recurrent somatic alteration in HGSOC [8, 35]. Because CN profiling mostly recovers the dominant clone per physical sample, we investigated subclonal population structure using digital PCR. We quantified the proportion of tumor cells containing the deletion and the tumour cellularity in each sample of those patients by counting the number of copies of DNA containing the patient-specific *TP53* mutation, wild-type *TP53* and the *NF1* breakpoint. The *NF1* deletion was detected at 5% and 26% in the pre-treatment samples (absolute cellularity 80% and 41%) and in $25 - 100\%$ of the interval debulking samples (median cellularity 49%). The histological analysis of case OV03-17 had demonstrated that it had arisen from a primary fallopian tube lesion (Fig. S1). We therefore extended the digital PCR analysis to DNA from microdissected tissues from formalin-fixed tissue blocks including the left fimbria. The *NF1* deletion was present in 1.2% of the primary invasive carcinoma in the fallopian tube and 7.9% of the biopsy from the adjacent left ovarian metastasis (Table S2).

In OV03-08, the relapsed ascites sample also diverged early with a comparatively long branch (Fig. 5) and indicating marked divergent evolution associated with deletions on Chr. 1q, 15q and 18q (Fig. S6A). In summary, the *NF1* deletion, while part of the dominant subpopulation at relapse, was already present at biopsy. As it is unlikely for this focal deletion to have arisen twice with the same size in the course of disease evolution, we conclude that the relapse was a clonal expansion of a minor subclone of presentation disease.

Discussion

Recent studies have described the extent of ITH to great detail and hypothesised that it influences patient outcome, but no rigorous attempt has been made so far to quantify this relationship. In this study, we combined a patient cohort that is sufficiently large in both number of patients and samples per patient with rigorous evolutionary methodology [36] to quantify ITH per patient and investigate its impact on patient outcome. Neoadjuvant chemotherapy, which has equivalent survival to primary cytoreductive surgery [41], thereby provided us with the unique opportunity to compare samples from spatially different sites, as well as different time points.

We have demonstrated that the metastatic process in HGSOC is tree-like and governed by ongoing clonal evolution in all metastatic sites with coincident local metastasis in tissues such as the omentum. For one patient (OV03-22), the null hypothesis of star-like evolution could not be rejected. This patient had 10 out of 17 samples rejected during data pre-processing due to low signal-to-noise ratio or poor cellularity. Low

overall divergence within the samples from this patient resulted in low confidence values at critical branching points in the tree. This may be explained by comparatively good response in this case to chemotherapy and there were no other histopathological or mutational characteristics that were different from the others in the study.

We not only show that rapid clonal expansions increase the probability of developing platinum-resistant relapse, but that on-going evolutionary change is also of critical importance. This result is in contrast to recent findings where the analysis of single nucleotide mutations did not show effects on ongoing evolutionary change [9]. However, these findings were limited to 3 cases of HGSOC and are dependent upon the depth of sequencing achieved. We believe that the analysis of CN changes in our evolutionary framework has greater robustness in HGSOC, where evolution is predominantly driven by CN change. The multivariate linear model showed that temporal heterogeneity and the degree of clonal expansion were the most important factors dictating progression-free survival. That total divergence is weakly correlated with PFS is not surprising and could also reflect different proliferation rates between tumors. However, several cases (OV03-21, OV03-17, OV03-22) that showed average total divergence had short PFS times, but instead showed exceptionally high divergence in the course of treatment or significant clonal expansions.

Genomic events that increase mutation rates give rise to mutator phenotypes [23] which in turn may provide the tumor with mechanisms to overcome severe population bottlenecks induced by chemotherapy. At present only tandem duplicator and *BRCA1/2* germ line mutations have been characterised as mutator phenotypes in HGSOC [26, 30], none of which were found in our cohort. Given that 28.6% of patients showed clear non-clock like evolutionary trajectories, this might suggest that other mutator phenotypes exist in HGSOC. However, we did not find not find a significant difference in survival between patients with and without constant evolutionary rates in the tree (t-test, p-value 0.88), so that the clinical significance of those potential unknown mutator phenotypes remains unclear.

Our phylogenetic reconstructions further allowed us to assess the robustness of the evolutionary trees, and thereby the certainty of placement of a sample in the tree. With this we addressed the question of when in the course of disease the relapse clone evolved. In both cases OV03-08 and OV03-17 we were able to determine an early branching point as the origin of relapse that shares an immediate ancestor with a pre-treatment sample. This together with the branch lengths on the adjacent samples indicates that in both cases the relapsed clone diverged before treatment and was a clonal expansion of pre-existing disease. Digital PCR of the focal *NF1* deletion conclusively shows that there was low allelic abundance of the deletion in pre-treatment samples and at the tubal primary site, but this mutation was dominant after initial chemotherapy. This is consistent with findings suggesting clonal expansion of relapse in HGSOC [11, 13].

In conclusion, using this unique dataset in combination with detailed evolutionary analysis allowed us for the first time to quantify the relationship between ITH and patient outcome and validate the clonal expansion hypothesis in a clinical setting. It provides a promising case study that shows how quantifying ITH can act as a prognostic indicator in HGSOC. In the light of these results it is difficult to overstate the importance of ITH

for resistance development. At the same time our findings strongly support previous doubts about the efficiency of personalised medicine approaches based on single sample biopsies [16].

Materials and Methods

Patients and samples

Clinical data and tissue samples for 33 patients were collected on the prospective CTCR-OV03 and CTCR-OV04 clinical studies designed to identify biomarkers of heterogeneity (REC reference numbers 05/Q0102/160 and 08/H0306/61). Imaging data for subjects from CTCR-OV03 has been previously described [34]. All patients received platinum-based chemotherapy and all but one case (OV04-30) underwent neo-adjuvant treatment. The median number of chemotherapy cycles prior to interval debulking surgery was 3 (range 3–7; $n = 16$). Patients provided written, informed consent for participation in these studies and for the use of their donated tissue and blood specimens for the laboratory studies carried out in these investigations.

Samples were stored in RNAlater and FFPE immediately after acquisition and later histologically examined and scored for cellularity by a sub-specialist gynaecological pathologist (M.J-L.). 15/33 cases were excluded from further analysis: 9 were not HG-SOC histology, 2 had no research tissue available and 4 cases had samples from only 1 timepoint or 1 metastatic site. DNA extraction was performed using the DNeasy Blood & Tissue Kit (QIAGEN) according to manufacturer's instructions.

SNP arrays

180 samples from 18 patients were profiled for copy number (CN) using Affymetrix Genome-Wide SNP 6.0 arrays (Table S1). Hybridisation of DNA to SNP 6.0 arrays was performed by AROS Applied Biotechnology (Aarhus, Denmark) following the manufacturer's protocol. Array data is available online at the NCBI Gene Expression Omnibus under accession number GSE40546. The datasets were segmented using PICNIC [19] (using the "primary" option), which further corrects for cellularity and estimates integer major and minor copy numbers. Data from 42/180 arrays were excluded because of tumour cellularity $< 50\%$ or high noise leading to removal of subject OV03-24 from further analysis. 34/42 excluded arrays were from samples obtained at interval debulking surgery following pre-operative chemotherapy suggesting that these cases had low cellularity resulting from response to treatment as has been previously observed (Table S1) [12]. 138 copy number profiles of 17 remaining patients were merged using CGHregions [40] and used for subsequent analysis.

Evolutionary inference and tree robustness

The MEDICC algorithm (Minimum Event Distance for Intra-tumour Copynumber Comparisons) was implemented in Python and C using the OpenFST finite-state transducer

library [2] and is available at `http://bitbucket.org/rfs/medicc/wiki`. The algorithm and strategy for copy number reconstruction and quantification of ITH have been described separately [36]). We used algorithms for determining the clonal expansion and temporal divergence index as described therein for all patients with more than 3 samples and where both biopsy and surgery samples were available. In 4 cases only biopsy or only surgery samples were available. We imputed missing TD index data by regressing CE against TD and prediction. Median imputation and complete removal of partially observed cases were performed as an alternative and both variants did not qualitatively alter the results.

Paired-end sequencing and breakpoint validation

DNA extracted from tumour samples and matched normal blood was processed using Illumina paired-end (PE) sample prep kit (Illumina). 41-bp (in some cases 50bp trimmed to 41bp) paired-end whole genome sequencing was performed on the Illumina Genome Analyzer IIx, where the median number of read pairs for each library was 153 million and the median sequencing depth was $\times$16.7. Sequencing data was processed using analysis pipelines as described [30]. Briefly, reads were aligned using BWA [22] and Novoalign (Novocraft Technologies, Selangor, Malaysia) and discordantly mapped read pairs were used to identify putative structural variants using a custom pipeline. PCR primers for validating structural variants were designed using Primer3 [39].

Deletion and insertion breakpoints were considered confirmed if they were detected as copy number decrease or increase on SNP array respectively by $> 50\%$ reciprocal overlap. Additionally, deletion, insertion, inter-chromosomal and inversion breakpoints were considered confirmed if both ends of a breakpoint were within 10kb of copy number breakpoints in any of the sequenced samples of the tumour.

Mutation detection by resequencing

The coding sequences of *TP53*, *BRCA1*, *BRCA2*, *PTEN*, *PIK3CA*, *EGFR* and *APC* were sequenced using the TAm-Seq method using Fluidigm Access Array 48.48 platform as described previously [15] with minor modifications: pre-amplification was ommitted as high-molecular weight DNA was extracted from fresh frozen tumour specimens; 50 ng of sample DNA and multiplexed primers were used in the target-specific amplification step. Primer sequences are available upon request. Sequencing data and variant verification was performed using IGV software as described [15, 33].

Digital PCR

Digital PCR was performed using the Fluidigm Biomark microfluidic system according to the manufacturer's instructions. Primers were designed spanning the *NF1* deletion (Forward: 5′-TTTTGTTTACGAGCACAGATAACC-3′; Reverse: 5′-GAAACAGAAGATGACAGCAAAGAA-3′). Reaction mixes were prepared containing 1$\times$ TaqMan Gene Expression mastermix (Applied Biosystems), 1$\times$ EvaGreen DNA binding dye (Biotium), 1$\times$ DNA Binding Dye Sample Loading Reagent (Fluidigm), 10nM primers and template

DNA. Prior to loading into a 12.765 Fluidigm digital chip, reactions were heated to 95℃ for 1 minute and placed on ice. Reactions were thermocycled at 50℃ for 2 min; 95℃ for 10 min followed by 55 cycles of 95℃ for 15 s and 56℃ or 60℃ for 1 minute. Digital PCR was also performed on the same samples using a assay for p.R175H mutation in *TP53* as previously described [4] (Forward: 5′-CCATCTACAAGCAGTCAC-3′; Reverse: 5′-GTCACCATCGCTATCTGAG-3′; Mutant specific probe: [6FAM]-TTGTGAGGCACTG-CCCCC-[BHQ1]; Wild-type specific probe: [HEX]-TTGTGAGGCGCTGCCCCC-[BHQ1]). The proportion of the *NF1* deletion in tumour cells was calculated from the estimated counts of both the *NF1* deletion and mutant *TP53* p.R175H specific assays.

References

[1] Ahmed Ashour Ahmed, Dariush Etemadmoghadam, Jillian Temple, Andy G. Lynch, Mohamed Riad, Raghwa Sharma, Colin Stewart, Sian Fereday, Carlos Caldas, Anna Defazio, David Bowtell, and James D. Brenton. Driver mutations in tp53 are ubiquitous in high grade serous carcinoma of the ovary. *J. Pathol.*, 221 (1):49–56, May 2010. doi: 10.1002/path.2696. URL `http://dx.doi.org/10.1002/path.2696`.

[2] Cyril Allauzen, Michael Riley, Johan Schalkwyk, Wojciech Skut, and Mehryar Mohri. OpenFst: A General and Efficient Weighted Finite-State Transducer Library. *Proceedings of the Ninth International Conference on Implementation and Application of Automata, (CIAA), in Lecture Notes in Computer Science*, 4783: 11–23, 2007.

[3] Kathryn Alsop, Sian Fereday, Cliff Meldrum, Anna deFazio, Catherine Emmanuel, Joshy George, Alexander Dobrovic, Michael J. Birrer, Penelope M. Webb, Colin Stewart, Michael Friedlander, Stephen Fox, David Bowtell, and Gillian Mitchell. Brca mutation frequency and patterns of treatment response in brca mutation-positive women with ovarian cancer: a report from the australian ovarian cancer study group. *J. Clin. Oncol.*, 30(21):2654–2663, Jul 2012. doi: 10.1200/JCO.2011.39.8545. URL `http://dx.doi.org/10.1200/JCO.2011.39.8545`.

[4] K. M. Archibald, H. Kulbe, J. Kwong, P. Chakravarty, J. Temple, T. Chaplin, M. B. Flak, I. A. McNeish, S. Deen, J. D. Brenton, B. D. Young, and F. Balkwill. Sequential genetic change at the tp53 and chemokine receptor cxcr4 locus during transformation of human ovarian surface epithelium. *Oncogene*, 31(48):4987–4995, Nov 2012. doi: 10.1038/onc.2011.653. URL `http://dx.doi.org/10.1038/onc.2011.653`.

[5] Kelly L Bolton, Georgia Chenevix-Trench, Cindy Goh, Siegal Sadetzki, Susan J Ramus, Beth Y Karlan, Diether Lambrechts, Evelyn Despierre, Daniel Barrowdale, Lesley McGuffog, Sue Healey, Douglas F Easton, Olga Sinilnikova, Javier Benítez, María J García, Susan Neuhausen, Mitchell H Gail, Patricia Hartge, Susan Peock, Debra Frost, D Gareth Evans, Rosalind Eeles, Andrew K Godwin, Mary B Daly, Ava Kwong, Edmond S K Ma, Conxi Lázaro, Ignacio Blanco, Marco Montagna, Emma D'Andrea, Maria Ornella Nicoletto, Sharon E Johnatty, Susanne Krüger Kjær, Allan Jensen, Estrid Høgdall, Ellen L Goode, Brooke L Fridley, Jennifer T Loud, Mark H Greene, Phuong L Mai, Angela Chetrit, Flora Lubin, Galit Hirsh-Yechezkel, Gord Glendon, Irene L Andrulis, Amanda E Toland, Leigha Senter, Martin E Gore, Charlie Gourley, Caroline O Michie, Honglin Song, Jonathan Tyrer, Alice S Whittemore, Valerie McGuire, Weiva Sieh, Ulf Kristoffersson, Håkan Olsson, Åke Borg, Douglas A Levine, Linda Steele, Mary S Beattie, Salina Chan, Robert L Nussbaum, Kirsten B Moysich, Jenny Gross, Ilana Cass, Christine Walsh, Andrew J Li, Ronald Leuchter, Ora Gordon, Montserrat Garcia-Closas, Simon A Gayther,

Stephen J Chanock, Antonis C Antoniou, Paul D P Pharoah, EMBRACE, kConFab Investigators, and Cancer Genome Atlas Research Network. Association between brca1 and brca2 mutations and survival in women with invasive epithelial ovarian cancer. *JAMA*, 307(4):382–390, Jan 2012. doi: 10.1001/jama.2012.20.

[6] David Bryant and Vincent Moulton. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.*, 21(2):255–265, Feb 2004. doi: 10.1093/molbev/msh018. URL `http://dx.doi.org/10.1093/molbev/msh018`.

[7] Peter J. Campbell, Shinichi Yachida, Laura J. Mudie, Philip J. Stephens, Erin D. Pleasance, Lucy A. Stebbings, Laura A. Morsberger, Calli Latimer, Stuart McLaren, Meng-Lay Lin, David J. McBride, Ignacio Varela, Serena A. Nik-Zainal, Catherine Leroy, Mingming Jia, Andrew Menzies, Adam P. Butler, Jon W. Teague, Constance A. Griffin, John Burton, Harold Swerdlow, Michael A. Quail, Michael R. Stratton, Christine Iacobuzio-Donahue, and P Andrew Futreal. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature*, 467(7319): 1109–1113, Oct 2010. doi: 10.1038/nature09460. URL `http://dx.doi.org/10.1038/nature09460`.

[8] Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature*, 474(7353):609–615, Jun 2011. doi: 10.1038/nature10166. URL `http://dx.doi.org/10.1038/nature10166`.

[9] Mauro Castellarin, Katy Milne, Thomas Zeng, Kane Tse, Michael Mayo, Yongjun Zhao, John R. Webb, Peter H. Watson, Brad H. Nelson, and Robert A. Holt. Clonal evolution of high-grade serous ovarian carcinoma from primary to recurrent disease. *J. Pathol.*, Sep 2012. doi: 10.1002/path.4105. URL `http://dx.doi.org/10.1002/path.4105`.

[10] Angela Chetrit, Galit Hirsh-Yechezkel, Yehuda Ben-David, Flora Lubin, Eitan Friedman, and Siegal Sadetzki. Effect of brca1/2 mutations on long-term survival of patients with invasive ovarian cancer: the national israeli study of ovarian cancer. *J. Clin. Oncol.*, 26(1):20–25, Jan 2008. doi: 10.1200/JCO.2007.11.6905.

[11] S. L. Cooke, C K Y. Ng, N. Melnyk, M. J. Garcia, T. Hardcastle, J. Temple, S. Langdon, D. Huntsman, and J. D. Brenton. Genomic analysis of genetic heterogeneity and evolution in high-grade serous ovarian carcinoma. *Oncogene*, 29(35): 4905–4913, Sep 2010. doi: 10.1038/onc.2010.245. URL `http://dx.doi.org/10.1038/onc.2010.245`.

[12] S. L. Cooke, J. Temple, S. Macarthur, M. A. Zahra, L. T. Tan, R A F. Crawford, C K Y. Ng, M. Jimenez-Linan, E. Sala, and J. D. Brenton. Intra-tumour genetic heterogeneity and poor chemoradiotherapy response in cervical cancer. *Br. J. Cancer*, 104(2):361–368, Jan 2011. doi: 10.1038/sj.bjc.6605971. URL `http://dx.doi.org/10.1038/sj.bjc.6605971`.

[13] Susanna L. Cooke and James D. Brenton. Evolution of platinum resistance in high-grade serous ovarian cancer. *Lancet Oncol.*, 12(12):1169–1174, Nov 2011. doi: 10.1016/S1470-2045(11)70123-1. URL http://dx.doi.org/10.1016/S1470-2045(11)70123-1.

[14] D. L. Dexter, H. M. Kowalski, B. A. Blazar, Z. Fligiel, R. Vogel, and G. H. Heppner. Heterogeneity of tumor cells from a single mouse mammary tumor. *Cancer Res.*, 38(10):3174–3181, Oct 1978.

[15] Tim Forshew, Muhammed Murtaza, Christine Parkinson, Davina Gale, Dana W Y. Tsui, Fiona Kaper, Sarah-Jane Dawson, Anna M. Piskorz, Mercedes Jimenez-Linan, David Bentley, James Hadfield, Andrew P. May, Carlos Caldas, James D. Brenton, and Nitzan Rosenfeld. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma dna. *Sci. Transl. Med.*, 4(136): 136ra68, May 2012. doi: 10.1126/scitranslmed.3003726. URL http://dx.doi.org/10.1126/scitranslmed.3003726.

[16] Marco Gerlinger, Andrew J. Rowan, Stuart Horswell, James Larkin, David Endesfelder, Eva Gronroos, Pierre Martinez, Nicholas Matthews, Aengus Stewart, Patrick Tarpey, Ignacio Varela, Benjamin Phillimore, Sharmin Begum, Neil Q. McDonald, Adam Butler, David Jones, Keiran Raine, Calli Latimer, Claudio R. Santos, Mahrokh Nohadani, Aron C. Eklund, Bradley Spencer-Dene, Graham Clark, Lisa Pickering, Gordon Stamp, Martin Gore, Zoltan Szallasi, Julian Downward, P Andrew Futreal, and Charles Swanton. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.*, 366(10):883–892, Mar 2012. doi: 10.1056/NEJMoa1113205. URL http://dx.doi.org/10.1056/NEJMoa1113205.

[17] Kylie L Gorringe, Sharoni Jacobs, Ella R Thompson, Anita Sridhar, Wen Qiu, David Y H Choong, and Ian G Campbell. High-resolution single nucleotide polymorphism array analysis of epithelial ovarian cancer reveals numerous microdeletions and amplifications. *Clin. Cancer Res.*, 13(16):4731–4739, Aug 2007.

[18] Mel Greaves and Carlo C Maley. Clonal evolution in cancer. *Nature*, 481(7381): 306–13, Jan 2012. doi: 10.1038/nature10762.

[19] Chris D. Greenman, Graham Bignell, Adam Butler, Sarah Edkins, Jon Hinton, Dave Beare, Sajani Swamy, Thomas Santarius, Lina Chen, Sara Widaa, P Andy Futreal, and Michael R. Stratton. PICNIC: an algorithm to predict absolute allelic copy number variation with microarray cancer data. *Biostatistics*, 11(1):164–175, Jan 2010. doi: 10.1093/biostatistics/kxp045. URL http://dx.doi.org/10.1093/biostatistics/kxp045.

[20] L. Khalique, A. Ayhan, M. E. Weale, I. J. Jacobs, S. J. Ramus, and S. A. Gayther. Genetic intra-tumour heterogeneity in epithelial ovarian cancer and its implications for molecular diagnosis of tumours. *J. Pathol.*, 211(3):286–295, Feb 2007. doi: 10.1002/path.2112. URL http://dx.doi.org/10.1002/path.2112.

[21] Lalarukh Khalique, Ayse Ayhan, John C. Whittaker, Naveena Singh, Ian J. Jacobs, Simon A. Gayther, and Susan J. Ramus. The clonal evolution of metastases from primary serous epithelial ovarian cancers. *Int. J. Cancer*, 124(7):1579–1586, Apr 2009. doi: 10.1002/ijc.24148. URL `http://dx.doi.org/10.1002/ijc.24148`.

[22] Heng Li and Richard Durbin. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, 25(14):1754–1760, Jul 2009. doi: 10.1093/bioinformatics/btp324. URL `http://dx.doi.org/10.1093/bioinformatics/btp324`.

[23] L. A. Loeb. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.*, 51(12):3075–3079, Jun 1991.

[24] M Markman, R Rothman, T Hakes, B Reichman, W Hoskins, S Rubin, W Jones, L Almadrones, and J L Lewis, Jr. Second-line platinum therapy in patients with ovarian cancer previously treated with cisplatin. *J. Clin. Oncol.*, 9(3):389–393, Mar 1991.

[25] Andriy Marusyk, Vanessa Almendro, and Kornelia Polyak. Intra-tumour heterogeneity: a looking glass for cancer? *Nat. Rev. Cancer*, 12(5):323–334, 2012. doi: 10.1038/nrc3261. URL `http://dx.doi.org/10.1038/nrc3261`.

[26] David J. McBride, Dariush Etemadmoghadam, Susanna L. Cooke, Kathryn Alsop, Joshy George, Adam Butler, Juok Cho, Danushka Galappaththige, Chris Greenman, Karen D. Howarth, King W. Lau, Charlotte K. Ng, Keiran Raine, Jon Teague, David C. Wedge, Australian Ovarian Cancer Study Group, Xavier Caubit, Michael R. Stratton, James D. Brenton, Peter J. Campbell, P Andrew Futreal, and David Dl Bowtell. Tandem duplication of chromosomal segments is common in ovarian and breast cancer genomes. *J. Pathol.*, 227(4):446–455, Aug 2012. doi: 10.1002/path.4042. URL `http://dx.doi.org/10.1002/path.4042`.

[27] Lauren M F. Merlo, John W. Pepper, Brian J. Reid, and Carlo C. Maley. Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer*, 6(12):924–935, Dec 2006. doi: 10.1038/nrc2013. URL `http://dx.doi.org/10.1038/nrc2013`.

[28] Nicholas Navin, Alexander Krasnitz, Linda Rodgers, Kerry Cook, Jennifer Meth, Jude Kendall, Michael Riggs, Yvonne Eberling, Jennifer Troge, Vladimir Grubor, Dan Levy, Pr Lundin, Susanne Mnr, Anders Zetterberg, James Hicks, and Michael Wigler. Inferring tumor progression from genomic heterogeneity. *Genome Res.*, 20 (1):68–80, Jan 2010. doi: 10.1101/gr.099622.109. URL `http://dx.doi.org/10.1101/gr.099622.109`.

[29] Nicholas Navin, Jude Kendall, Jennifer Troge, Peter Andrews, Linda Rodgers, Jeanne McIndoo, Kerry Cook, Asya Stepansky, Dan Levy, Diane Esposito, Lakshmi Muthuswamy, Alex Krasnitz, W Richard McCombie, James Hicks, and Michael Wigler. Tumour evolution inferred by single-cell sequencing. *Nature*, 472(7341):

90–94, Apr 2011. doi: 10.1038/nature09807. URL http://dx.doi.org/10.1038/nature09807.

[30] Charlotte K Y. Ng, Susanna L. Cooke, Kevin Howe, Scott Newman, Jian Xian, Jillian Temple, Elizabeth M. Batty, Jessica C M. Pole, Simon P. Langdon, Paul A W. Edwards, and James D. Brenton. The role of tandem duplicator phenotype in tumour evolution in high-grade serous ovarian cancer. *J. Pathol.*, 226(5):703–712, Apr 2012. doi: 10.1002/path.3980. URL http://dx.doi.org/10.1002/path.3980.

[31] Serena Nik-Zainal, Peter Van Loo, David C. Wedge, Ludmil B. Alexandrov, Christopher D. Greenman, King Wai Lau, Keiran Raine, David Jones, John Marshall, Manasa Ramakrishna, Adam Shlien, Susanna L. Cooke, Jonathan Hinton, Andrew Menzies, Lucy A. Stebbings, Catherine Leroy, Mingming Jia, Richard Rance, Laura J. Mudie, Stephen J. Gamble, Philip J. Stephens, Stuart McLaren, Patrick S. Tarpey, Elli Papaemmanuil, Helen R. Davies, Ignacio Varela, David J. McBride, Graham R. Bignell, Kenric Leung, Adam P. Butler, Jon W. Teague, Sancha Martin, Goran Jnsson, Odette Mariani, Sandrine Boyault, Penelope Miron, Aquila Fatima, Anita Langerd, Samuel A J R. Aparicio, Andrew Tutt, Anieta M. Sieuwerts, ke Borg, Gilles Thomas, Anne Vincent Salomon, Andrea L. Richardson, Anne-Lise Brresen-Dale, P Andrew Futreal, Michael R. Stratton, Peter J. Campbell, and Breast Cancer Working Group of the International Cancer Genome Consortium . The life history of 21 breast cancers. *Cell*, 149(5):994–1007, May 2012.

[32] P. C. Nowell. The clonal evolution of tumor cell populations. *Science*, 194(4260): 23–28, Oct 1976.

[33] James T Robinson, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S Lander, Gad Getz, and Jill P Mesirov. Integrative genomics viewer. *Nat. Biotechnol.*, 29(1):24–26, Jan 2011. doi: 10.1038/nbt.1754.

[34] Evis Sala, Masako Y Kataoka, Andrew N Priest, Andrew B Gill, Mary A McLean, Ilse Joubert, Martin J Graves, Robin A F Crawford, Mercedes Jimenez-Linan, Helena M Earl, Charlotte Hodgkin, John R Griffiths, David J Lomas, and James D Brenton. Advanced ovarian cancer: multiparametric mr imaging demonstrates response- and metastasis-specific effects. *Radiology*, 263(1):149–159, Apr 2012. doi: 10.1148/radiol.11110175.

[35] Navneet Sangha, Rong Wu, Rork Kuick, Scott Powers, David Mu, Diane Fiander, Kit Yuen, Hidetaka Katabuchi, Hironori Tashiro, Eric R Fearon, and Kathleen R Cho. Neurofibromin 1 (nf1) defects are common in human ovarian serous carcinomas and co-occur with tp53 mutations. *Neoplasia*, 10(12):1362–1372, Dec 2008.

[36] Roland F Schwarz, Anne Trinh, Botond Sipos, James D Brenton, Nick Goldman, and Florian Markowetz. Phylogenetic quantification of intra-tumour heterogeneity. *Genome Biology (submitted)*, 2013.

[37] Sohrab P. Shah, Ryan D. Morin, Jaswinder Khattra, Leah Prentice, Trevor Pugh, Angela Burleigh, Allen Delaney, Karen Gelmon, Ryan Guliany, Janine Senz, Christian Steidl, Robert A. Holt, Steven Jones, Mark Sun, Gillian Leung, Richard Moore, Tesa Severson, Greg A. Taylor, Andrew E. Teschendorff, Kane Tse, Gulisa Turashvili, Richard Varhol, Ren L. Warren, Peter Watson, Yongjun Zhao, Carlos Caldas, David Huntsman, Martin Hirst, Marco A. Marra, and Samuel Aparicio. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature*, 461(7265):809–813, Oct 2009. doi: 10.1038/nature08489. URL `http://dx.doi.org/10.1038/nature08489`.

[38] Sohrab P. Shah, Andrew Roth, Rodrigo Goya, Arusha Oloumi, Gavin Ha, Yongjun Zhao, Gulisa Turashvili, Jiarui Ding, Kane Tse, Gholamreza Haffari, Ali Bashashati, Leah M. Prentice, Jaswinder Khattra, Angela Burleigh, Damian Yap, Virginie Bernard, Andrew McPherson, Karey Shumansky, Anamaria Crisan, Ryan Giuliany, Alireza Heravi-Moussavi, Jamie Rosner, Daniel Lai, Inanc Birol, Richard Varhol, Angela Tam, Noreen Dhalla, Thomas Zeng, Kevin Ma, Simon K. Chan, Malachi Griffith, Annie Moradian, S-W Grace Cheng, Gregg B. Morin, Peter Watson, Karen Gelmon, Stephen Chia, Suet-Feung Chin, Christina Curtis, Oscar M. Rueda, Paul D. Pharoah, Sambasivarao Damaraju, John Mackey, Kelly Hoon, Timothy Harkins, Vasisht Tadigotla, Mahvash Sigaroudinia, Philippe Gascard, Thea Tlsty, Joseph F. Costello, Irmtraud M. Meyer, Connie J. Eaves, Wyeth W. Wasserman, Steven Jones, David Huntsman, Martin Hirst, Carlos Caldas, Marco A. Marra, and Samuel Aparicio. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*, 486(7403):395–399, Jun 2012. doi: 10.1038/nature10933. URL `http://dx.doi.org/10.1038/nature10933`.

[39] Andreas Untergasser, Ioana Cutcutache, Triinu Koressaar, Jian Ye, Brant C. Faircloth, Maido Remm, and Steven G. Rozen. Primer3–new capabilities and interfaces. *Nucleic Acids Res.*, Jun 2012. doi: 10.1093/nar/gks596. URL `http://dx.doi.org/10.1093/nar/gks596`.

[40] Mark A. van de Wiel and Wessel N van Wieringen. Cghregions: dimension reduction for array cgh data with minimal information loss. *Cancer Inform.*, 3:55–63, 2007.

[41] Ignace Vergote, Claes G Tropé, Frédéric Amant, Gunnar B Kristensen, Tom Ehlen, Nick Johnson, René H M Verheijen, Maria E L van der Burg, Angel J Lacave, Pierluigi Benedetti Panici, Gemma G Kenter, Antonio Casado, Cesar Mendiola, Corneel Coens, Leen Verleye, Gavin C E Stuart, Sergio Pecorelli, Nick S Reed, European Organization for Research and Treatment of Cancer-Gynaecological Cancer Group, and NCIC Clinical Trials Group. Neoadjuvant chemotherapy or primary surgery in stage iiic or iv ovarian cancer. *N. Engl. J. Med.*, 363(10):943–953, Sep 2010. doi: 10.1056/NEJMoa0908806.

[42] Joost S. Vermaat, Isaac J. Nijman, Marco J. Koudijs, Frank L. Gerritse, Stefan J. Scherer, Michal Mokry, Wijnand M. Roessingh, Nico Lansu, Ewart de Bruijn, Richard van Hillegersberg, Paul J. van Diest, Edwin Cuppen, and Emile E. Voest.

Primary colorectal cancers and their subsequent hepatic metastases are genetically different: implications for selection of patients for targeted treatment. *Clin. Cancer Res.*, 18(3):688–699, Feb 2012. doi: 10.1158/1078-0432.CCR-11-1965. URL http://dx.doi.org/10.1158/1078-0432.CCR-11-1965.

[43] Xiaochong Wu, Paul A. Northcott, Adrian Dubuc, Adam J. Dupuy, David J H. Shih, Hendrik Witt, Sidney Croul, Eric Bouffet, Daniel W. Fults, Charles G. Eberhart, Livia Garzia, Timothy Van Meter, David Zagzag, Nada Jabado, Jeremy Schwartzentruber, Jacek Majewski, Todd E. Scheetz, Stefan M. Pfister, Andrey Korshunov, Xiao-Nan Li, Stephen W. Scherer, Yoon-Jae Cho, Keiko Akagi, Tobey J. MacDonald, Jan Koster, Martin G. McCabe, Aaron L. Sarver, V Peter Collins, William A. Weiss, David A. Largaespada, Lara S. Collier, and Michael D. Taylor. Clonal selection drives genetic divergence of metastatic medulloblastoma. *Nature*, 482(7386):529–533, Feb 2012. doi: 10.1038/nature10825. URL http://dx.doi.org/10.1038/nature10825.

Acknowledgements

We thank H. Biggs, C. Hodgkin, L. Jones and the Addenbrooke's gynaecological cancer multidisciplinary team for expert assistance in sample collection and study recruitment. **Funding:** We acknowledge the support of Cancer Research UK, the University of Cambridge, National Institute for Health Research Cambridge Biomedical Research Centre, Cambridge Experimental Cancer Medicine Centre and Hutchison Whampoa Limited. C.P. was supported in part by the Academy of Medical Sciences, Wellcome Trust, British Heart Foundation and Arthritis Research UK. **Author contributions:** R.F.S. with support from F.M. designed the analysis strategy, executed all data analyses and designed and implemented the computational methods. E.S., C.A.P., K.S., P.J.B., H.M.E. and J.D.B. designed the clinical study and collected samples and clinical data. C.K.Y.N., S.L.C., S.N., J.T., A.M.P., D.G., M.M., N.R. performed genomic assays and analysis. M.J.-L. performed pathological analysis. R.F.S., F.M., and J.D.B. wrote the manuscript with assistance from C.K.Y.N and other authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Array data is available online at the NCBI Gene Expression Omnibus under accession number GSE40546.

Figures

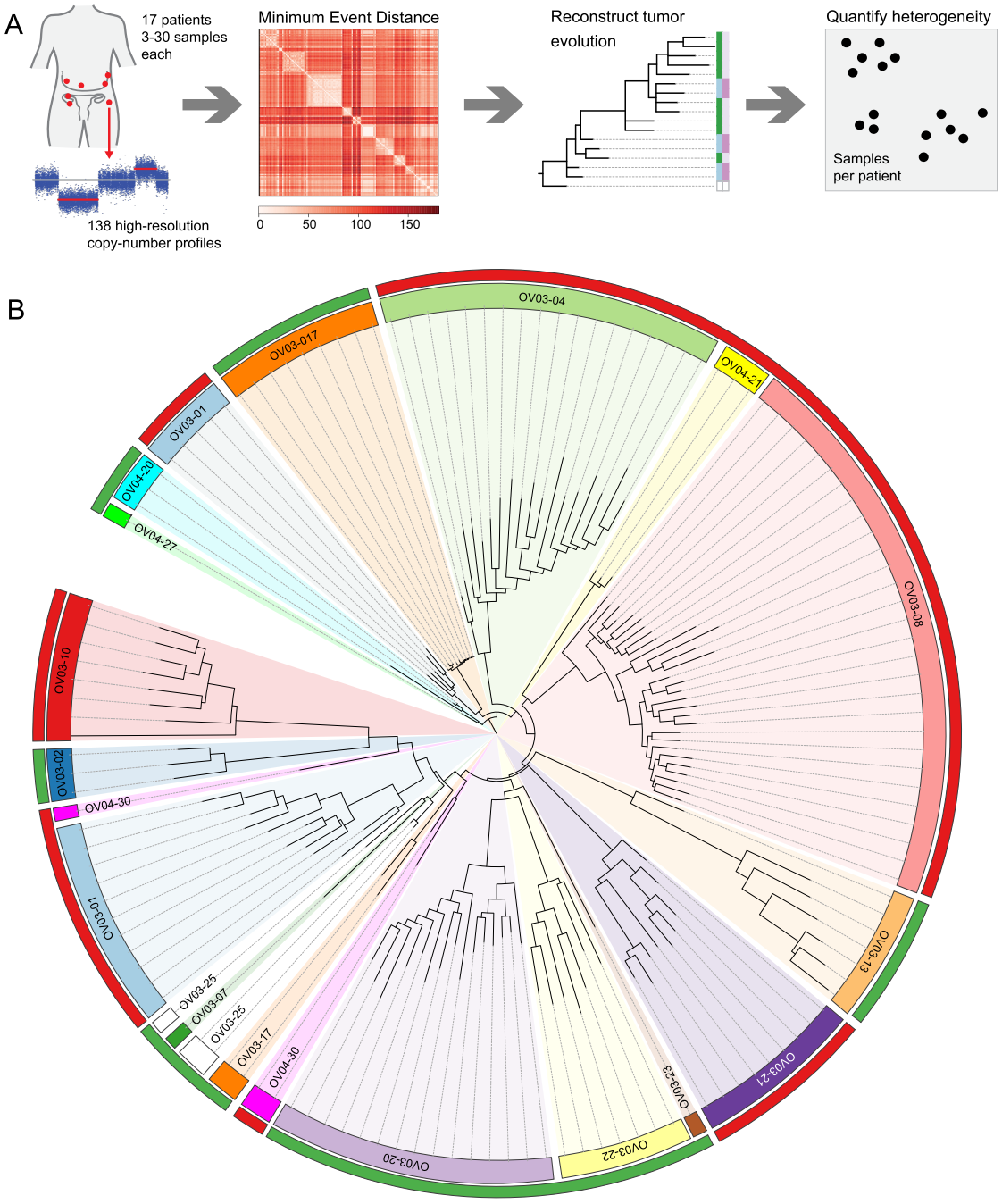

Figure 1: **A. Numerical quantification of intra-tumoral genetic heterogeneity (ITH) by evolutionary comparisons.** Copy-number (CN) profiles from 138 metastatic sites were obtained for 17 patients with HGSOC. The "Minimum Event Distance for Intra-tumor Copy number Comparisons" (MEDICC) algorithm was used to compute minimum event distances between profiles and to reconstruct the evolutionary history for each sample, enabling numerical quantification of both spatial and temporal heterogeneity and clonal expansion for each case. **B. HGSOC exhibits significant patient-specific intra-tumor genetic heterogeneity.** Neighbour-joining tree of all samples based on total copy number. Samples from each case (colored inner circular bar) cluster into clades. The outer circular bar indicates survival. Red, progression-free survival (PFS) < 12 months; green, PFS > 12 months. No immediate clustering of patients into sensitive and resistant subgroups is visible.

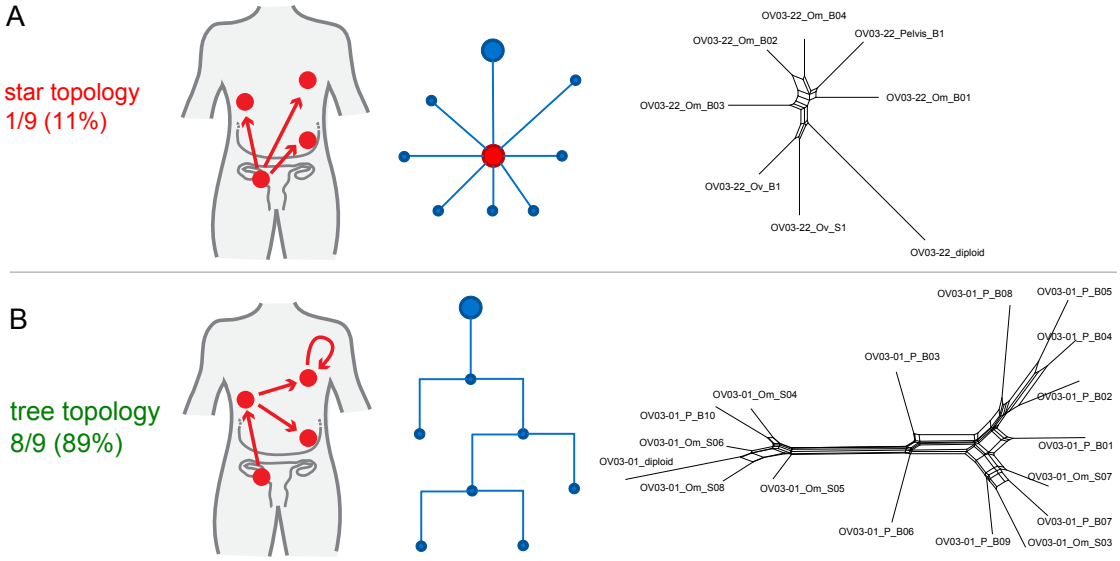

Figure 2: **A. Radial pattern of metastatic spread leads to a star topology.** If all metastases are derived from the primary lesion (left) the evolutionary relationships have star-like topology (middle). A neighbor-net representation of the distances shows deviation from a tree structure (right). **B. Branched metastatic spread leads to a tree topology**. If metastases create new metastases (left), the evolutionary history is tree-like (middle). Neighbor-net representation of the distance matrix then shows the tree-like structure (right). The number and proportion of cases classified to star or tree topology are shown.

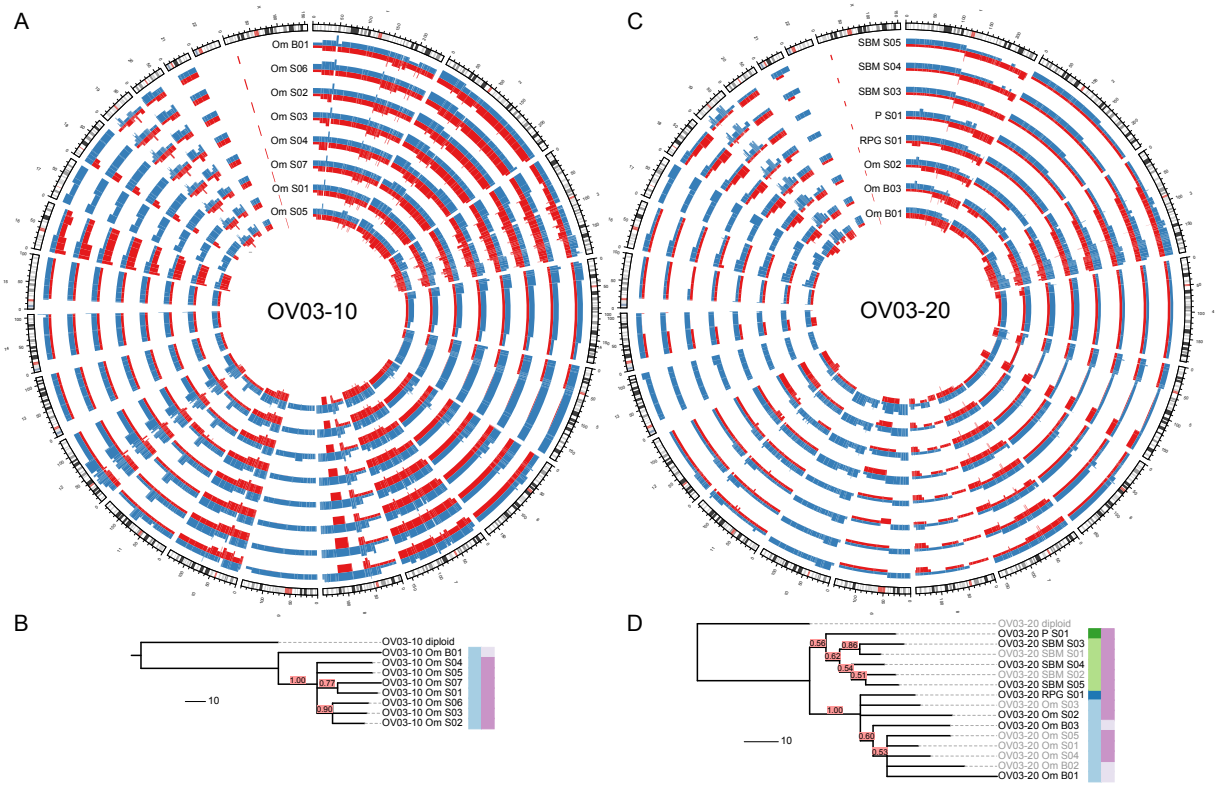

Figure 3: **Phylogenies and genome architecture for OV03-10 and OV03-20:**A representative subset of the allele-specific genomic CN profiles of each patient is shown. Seperate alleles are indicated in red and blue. The black sample names in the tree indicate the selected samples. Confidence values for each split are printed in red boxes. The color coded bars on the right of the phylogenies indicate different sites (left column) and different sampling times (right column). Branch lengths are to scale in *events* as determined by MEDICC and as indicated by the separate black unit branch.

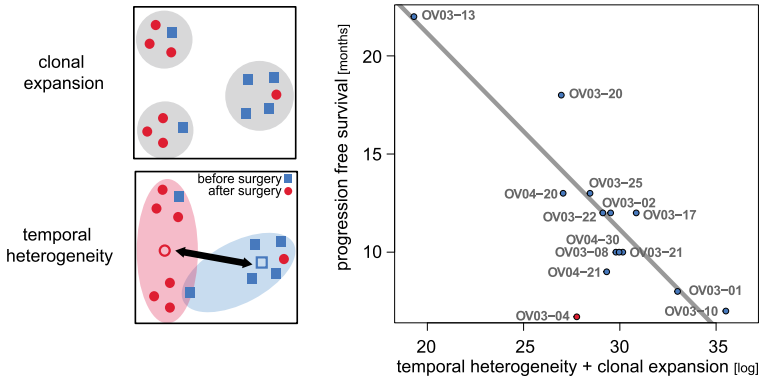

Figure 4: **Patient survival is determined by temporal divergence and degree of clonal expansion.** The degree of clonal expansion was quantified as the deviation from non-randomness in the spatial distribution of samples in the first two principal components (gray circles, top left). Temporal heterogeneity was determined as the distance between the robust centroids of biopsy (blue) and surgery samples (red, bottom left). Linear regression shows relationship between temporal heterogeneity, the degree of clonal expansion and progression free survival for all patients except OV03-04 (right, $R^2 = 0.84, p = 0.000122$). Patient OV03-04, indicated in red, was affected by endometrioid cancer, not HGSOC (see results) and shows as a clear outlier in the analysis.

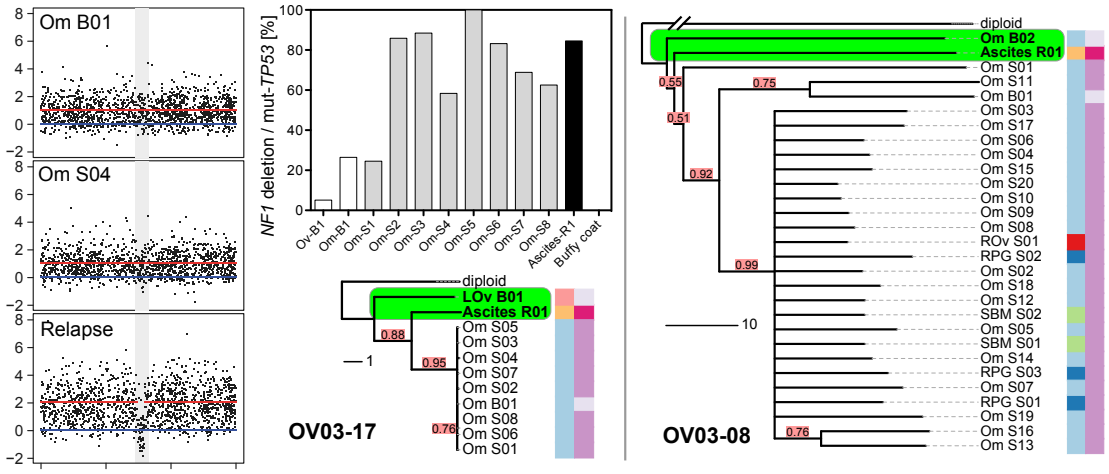

Figure 5: **Relapse is an early diverged clonal expansion of a low-prevalence subclone of presentation disease.** In both relapse cases the relapse was placed next to a biopsy sample and close to the diploid as an early branched event (see trees). The long branches indicate strong divergence. Additionally, in case 17, the focal *NF1* deletion which was dominant in the relapse and only there visible on the aCGH (left), was confirmed to be present in all samples by digital PCR (left top), but only as a minor subclone at biopsy.

Tables

Case	Age	Stage	Resp.	Tx	CA125	PFS	#samples	TD index	CE index	ST pval
OV03-01	47	4	PR	Neo	-93	8	16/20	4.7825	1.2605	0.000
OV03-02	62	3C	PR	Neo	-92	12	3/5	NA	0.7106	0.670
OV03-04	69	4	SD	Neo	-92	6	18/20	NA	1.2433	0.000
OV03-07	48	3	PR	Neo	-91	20	1/3	NA	NA	NA
OV03-08	63	4	PR	Neo	-76	10	29/29	3.8437	1.4706	0.000
OV03-10	59	4	SD	Neo	-80	7	8/8	6.6617	0.7299	0.001
OV03-13	61	3C	PR	Neo	-43	22	7/8	3.0334	0.6837	0.000
OV03-17	51	3C	PR	Neo	-24	12	11/14	3.4609	2.2358	0.000
OV03-20	71	3	SD	Neo	-100	18	15/16	4.5374	0.6494	0.000
OV03-21	60	3	PR	Neo	-87	10	9/11	4.7720	0.8686	0.000
OV03-22	58	3C	PR	Neo	-98	12	7/17	5.7657	0.4834	0.282
OV03-23	60	3	SD	Neo	-88	>29	1/4	NA	NA	NA
OV03-25	57	3C	PR	Neo	-28	>12	3/3	NA	0.6215	0.484
OV04-20	63	3C	PR	Neo	NA	>12	3/4	4.6729	0.6083	0.740
OV04-21	54	3	SD	Neo	NA	9	3/5	NA	0.7413	0.740
OV04-27	58	3	PR	Neo	NA	>12	1/3	NA	NA	NA
OV04-30	60	3C	*SD	PS	NA	10	3/4	NA	0.8591	0.640

Table 1: Summary of samples from the OV03/04 clinical studies. The table shows patients recruited into the studies and the number of samples available before and after quality control. Tx: Treatment (Neo: neo-adjuvant, PS: primary surgery), Resp: Response according to RECIST evaluation (PR: Partial response, SD: Stable disease, PD: progressive disease, *SD is not comparable to SD as treatment modalities are different), CA125: CA125 tumour marker reduction (pc), PFS: progression-free survival (months), TD index: temporal divergence index, CE index: degree of clonal expansion, ST p-val: p-value of test for star topology (BH corrected). Patients with < 3 samples could not be evaluated for TD and CE indices. Patients with CE but no TD index did not have both biopsy and surgery samples available.

	Effect	Gene	Protein change	cDNA change	RefSeq ID
OV03-01	ms	TP53	p.Y234C	c.A701G	NM_000546
OV03-02	ns	TP53	p.Y234X	c.C702A	NM_000546
OV03-02	ms	BRCA1	p.Y179C	c.A536G	NM_007294
OV03-02	ms	BRCA1	p.N550H	c.A1648C	NM_007294
OV03-02	ms	BRCA1	p.F486L	c.T1456C	NM_007294
OV03-02	ms	BRCA2	p.E1110V	c.A3329T	NM_000059
OV03-04	ND	TP53			
OV03-07	ms	TP53	p.H214R	c.A641G	NM_000546
OV03-08	ms	TP53	p.C141R	c.T421C	NM_000546
OV03-10	fs	TP53	p.P153fs	c.459_469del11	NM_000546
OV03-13	ms	TP53	p.R273C	c.C817T	NM_000546
OV03-13	ms	APC	p.S2596A	c.T7786G	NM_000038
OV03-17	ms	TP53	p.R175H	c.G524A	NM_000546
OV03-17	silent	BRCA2	p.G1552G	c.T4656C	NM_000059
OV03-20	fs	TP53	p.I195fs	c.583_584dupA	NM_000546
OV03-21	ms	TP53	p.S215G	c.A643G	NM_000546
OV03-21	ms	APC	p.D1714N	c.G5140A	NM_000038
OV03-22	ns	TP53	p.R306X	c.C916T	NM_000546
OV03-23	ND	TP53			
OV03-25	ms	TP53	p.Y236S	c.A707C	NM_000546
OV03-25	silent	BRCA2	p.V465V	c.A1395C	NM_000059
OV04-20	ms	TP53	p.V216L	c.G646T	NM_000546
OV04-20	ns	BRCA2	p.L2732X*	c.T8195A	NM_000059
OV04-21	ms	TP53	p.C135R	c.T403C	NM_000546
OV04-27	ms	TP53	p.C275Y	c.G824A	NM_000546
OV04-30	ms	TP53	p.R273H	c.G818A	NM_000546

Table 2: Mutations detected in samples from CTCR-OV03 and CTCR-OV04 subjects using TAm-Seq. OVO4-20 had a deleterious somatic nonsense mutation (p.L2732X*) in *BRCA2*. This mutation was not detected in two independent germline DNA samples from OV04-20. All other BRCA1/2 mutations shown are not pathogenic or are of no/unknown clinical importance according to BIC and LOVD-IARC databases. ND, no mutation detected; Ms, missense; ns, nonsense; fs, frameshift.