Genes predisposed to DNA hypermethylation during acquired resistance to chemotherapy are identified in ovarian tumors by bivalent chromatin domains at initial diagnosis

Edward Curry[1], Constanze Zeller[1], Nahal Masrour[1], Darren K Patten[1], John Gallon[1], Charlotte S Wilhelm-Benartzi[2], Sadaf Ghaem-Maghami[1], David D Bowtell[3], Robert Brown[1,4]

[1] Department Surgery & Cancer, Imperial College London, London W12 0NN, UK

[2] Centre for Trials Research, Cardiff University, Cardiff, CF14 4YS, UK

[3] Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia 3002

[4] Institute for Cancer Research, Sutton, SM2 5NG, UK

Corresponding Author: Prof Robert Brown, Imperial College London, IRDB 4th Floor, Hammersmith Hospital, London W12 0NN, UK. E.mail b.brown@imperial.ac.uk. Tel +44 207 5941804, Fax +44 207 5942129

Running title: bivalent chromatin and acquired HGSOC CpG methylation

Keywords: ovarian cancer, epigenetics, histones, DNA methylation, drug resistance

Conflicts of interest: The authors declare that they have no competing interests

Word count: 5064

Figures: 3

Tables: 1

ABSTRACT

Bivalent chromatin domains containing both active H3K4me3 and repressive H3K27me3 histone marks define gene sets poised for expression or silencing in differentiating embryonic stem (ES) cells. In cancer cells aberrantly poised genes may facilitate changes in transcriptional states after exposure to anti-cancer drugs. In this study, we used ChIP-seq to characterize genome-wide positioning of H3K4me3 and H3K27me3-associated chromatin in primary high-grade serous ovarian carcinomas and in normal ovarian surface and fallopian tube tissue. Gene sets with proximal bivalent marks defined in this manner were evaluated subsequently as signatures of systematic change in DNA methylation and gene expression, comparing pairs of tissue samples taken from patients at primary presentation and relapse following chemotherapy. We found that gene sets harboring bivalent chromatin domains at their promoters in tumor tissue, but not normal epithelia, overlapped with Polycomb-repressive complex target genes as well as transcriptionally silenced genes in normal ovarian and tubal stem cells. The bivalently marked genes we identified in tumors before chemotherapy displayed increased promoter CpG methylation and reduced gene expression at relapse after chemotherapy of ovarian cancer. Overall, our results support the hypothesis that pre-existing histone modifications at genes in a poised chromatin state may lead to epigenetic silencing during acquired drug resistance.

Precis: Results suggest epigenetic targets for intervention to prevent the emergence of cancer drug resistance.

INTRODUCTION

Genetic information is packaged into human cells in the form of chromatin: double-stranded molecules of DNA wrapped around complexes of histone proteins. Covalent modifications of the histone tails can alter the density of compaction of chromatin, which can influence the accessibility of DNA to the transcription machinery that is required for gene expression. The Polycomb repressive complex (PRC2), suppresses target gene expression through catalysing the trimethylation of Lysine 27 of the histone protein H3 (H3K27me3)(1). This modified histone forms broad domains that are enriched at repressed genes(2). PRC2-null mouse Embryonic Stem (ES) cells can be derived, but do not contribute to viable organisms, so it is widely accepted that PRC2 and H3K27me3 play an important role in the differentiation of pluripotent cells. In ES cells, genomic loci bound by the core PRC2 members (SUZ12, EED and EZH2) tend to display both the repressive mark H3K27me3 and permissive mark H3K4me3 (trimethylation of Lysine 4 of the histone protein H3). These 'bivalent' promoters, simultaneously containing both active and repressive histone marks, are strongly enriched for genes whose expression either rapidly increases or rapidly decreases upon differentiation from ES cells(3). Recent physical evidence supports the co-occurrence of both H3K4me3 and H3K27me3 on the same nucleosomes, and demonstrates resolution of bivalent domains upon differentiation(4). This leads to current understanding that these marks indicate genes 'poised' for response to some differentiation or selective stimulus.

Aberrant gene silencing, leading to inactivation of tumour-suppressors and dysregulation of cell growth, has long been regarded an important feature of carcinogenesis(5). More recently, it has been shown to play an active role in the acquisition of drug resistance driven by epigenetic and genetic mechanisms (6-8). While a large body of work in the epigenetic area focuses on methylation of cytosines in DNA (CpG methylation), recent evidence also implicates histone modifications, particularly H3K27me3, as a means of aberrant gene

silencing in cancer. These two mechanisms are known to be linked(9), and genes associated with bivalent marks in ES cells are enriched in genes with promoter CpG hypermethylation in cancers(10). A further observation linking PRC2 with cancer, and particularly to cancer stem cells, is that the histone methyltransferase component of PRC2, EZH2, is overexpressed in tumour-initiating subpopulations of ovarian cancer cell lines, and inhibition of EZH2 inhibited tumour growth(11). High EZH2 expression has been associated with poor outcomes in a range of cancer types(12,13).

We previously showed that a primary high-grade serous ovarian tumour had bivalent chromatin domains enriched for loci that are bivalently marked in ES cells, but additionally at loci detected only in the tumour(14). These bivalently marked genes were preferentially silenced in tumour-initiating subpopulations of an ovarian cancer cell line and in the platinum-resistant isogenic pair of a platinum-sensitive ovarian cancer cell line. This suggests that the acquisition of stem cell associated bivalent chromatin domains could provide tumour cells a mechanism for rapid adaptation to drug exposure, with expression changes made more persistent through acquisition of DNA methylation. The involvement of these epigenetic alterations in cancer cell populations may result from adaptation of normal cellular differentiation involved in tissue maintenance and regeneration, and could link the process of tumour initiation to later acquired drug resistance(15). It is therefore relevant that a definitive stem cell populations from both ovarian and fallopian tube epithelia have recently been characterized(16).

We have now generated histone modification profiles from further primary high-grade serous ovarian tumours and from normal ovarian tissue in an attempt to shed light on the role of bivalent chromatin domains in ovarian carcinogenesis and acquired drug resistance in patients. A schematic summary of our approach and our key findings are shown in Figure 1. As would be expected, there is variation in the sets of H3K27me3 and H3K4me3 marked genes from tumour to tumour. However, we show that the expanded common set of genes

bivalently marked in ovarian tumours is strongly enriched for the canonical PRC2 target genes associated with differentiation. Importantly, we use paired tumour samples taken from a cohort of patients before and after the acquisition of platinum resistance to show that these bivalently-marked PRC2 target genes are predisposed towards epigenetic silencing via gain of DNA methylation following chemotherapy, during acquired drug resistance of patients' tumours.

MATERIALS AND METHODS

Sample collection and cell lines

Four high-grade serous epithelial ovarian tumour samples from untreated patients at primary presentation, three normal ovarian surface epithelium samples and one normal fallopian tube epithelium sample were approved and obtained from the 'HTA-approved' Imperial College Healthcare NHS Trust Tissue Bank (authorized by the Wales MREC). Tumour collection and analysis for this study was approved by the West London Research Ethics Committee (reference 09/H0707/89) according to Declaration of Helsinki. Written consent was obtained from all patients included in this study who provided tumour tissue for research.

PEA1 and PEA2 cell lines (17) were obtained and used within 6 months from Department of Surgery and Cancer cell stocks. STR profiling was used to authenticate cell lines prior to freezing and to confirm that pairs of lines were related. All cell lines were mycoplasma free and checked fortnightly.

Chromatin immunoprecipitation, sequencing and PCR

All ChIP-seq sample processing and library preparation was performed as previously described(14). ChIP libraries were sequenced using an Illumina HiSeq2000 at the CSC Genomics Core Laboratory of Imperial College London. Individual 50bp reads were filtered if they contained 5 base calls with quality Phred score less than 30, or if they contained any undetermined bases. The remaining reads were mapped to the hg19 reference genome using Bowtie2(18). Reads were further filtered so that a maximum of one uniquely-mapped read was kept for each genomic co-ordinate. All reads mapping to the DAC Consensus Excluded Regions were discarded and cross-correlation for each filtered sequence library was calculated as described(19) across a range of inter-strand offsets from 1 to 200bp, at 5bp intervals. Locations of significant peaks of enriched DNA in each of the ChIP samples were identified using MACS v1.4.2(20) using a p-value cutoff of 1e-4.

IP-enrichment of candidate regions was confirmed through qPCR of IP chromatin from the high-grade serous ovarian cancer cell line PEA1. Cells were crosslinked with 1% formaldehyde and processed according to (21). Briefly, chromatin extract were sonicated using a Diagenode sonicator using 20 cycles (30 s on and 30 s off) at maximum intensity. Purified chromatin was then immunoprecipitated using 4 µg of H3K4me3 (Abcam 8580) or H3K27me3 (Abcam 6002) antibodies per ChIP. Non-immunoprecipitated chromatin was used as Input control. PCR primer sequences are given in Supplementary Table S1. IP enrichment was calculated as $\frac{input\_DNAconc}{ChIP\_DNAconc} 2^{-\Delta Ct}$, where ΔCt denotes the difference in amplification Ct values between input DNA and antibody pull-down DNA.

Assessing differential methylation and expression in clinical samples

All DNA methylation data was taken from ovarian tumour and ascites samples from the ICGC study(22) profiled on the Illumina 450k HumanMethylation arrays. 11 paired primary-relapse samples were available from which the primary sample was tumour but the paired relapse sample was from ascites, and 1 pair where tumour material from relapse was also obtained. In addition, 3 pairs of tumour and ascites samples both taken at primary presentation were available. To account for tissue-specific methylation differences between tumour cells and ascites, differential methylation was evaluated in terms of the fold-increase in methylation M-values between primary and relapse samples for each pair, divided by the average fold-increase in methylation M-values between tumour and ascites pairs. For differential gene expression, log-transformed, normalized FPKM RNA-seq counts were obtained for each gene from the ICGC ovarian cancer study(22). Statistics for differential expression were calculated using LIMMA(23) with a linear model term for each patient, a term to model tissue-specific differences and a term to model differences from primary to relapse. Empirical Bayes moderated t-statistics were obtained for the primary-vs-relapse coefficient for each gene.

Bisulphite Pyrosequencing

A total amount of 500 ng of genomic DNA was bisulphite modified using the Zymo Research EZ-DNA Methylation Kit (Cambridge Bioscience) according to the manufacturer's instructions. Sequences of pyrosequencing primer sets are provided in Supplementary Table S1. Pyrosequencing PCR was performed in duplicate for each sample in a 25 µl volume containing an end-concentration of 1 U Faststart Taq polymerase (Roche, Welwyn Garden City, UK), 1x FastStart PCR Buffer including 2 µM $MgCl_2$ (Roche), 0.05 mm dNTPs (Roche), 0.4 µM primers (each) adding 1 µl of modified DNA template using the following conditions: 95 °C for 6 min, 45 cycles of 95 °C for 30 s, 58 °C for 30 s, 72 °C for 30 s, followed by 72 °C for 5 min and terminating at 4°C. Pyrosequencing of PCR products was performed on PyroMark Q 96 MD using the PyroGold Reagent Kit (Qiagen) according to the manufacturer's instructions. The methylation percentage of CpG sites for individual genes was calculated by using the Pyro Q-CpG software (version 1.0.9), Biotage (Uppsala, Sweden).

Statistical analysis

Hypergeometric test refers to use of the hypergeometric distribution in **R** to calculate the probability of at least as great an observed overlap between two sets of genes occurring purely by chance.

Pearson Correlation Coefficients and their corresponding significance estimates were calculated in **R**.

The 'GeneSetTest' function in the Limma package(23) from Bioconductor was used to evaluate the statistical significance of a systematic shift of a set of genes towards an increase (or decrease) in measurements between two groups of samples measured in a molecular profiling experiment.

Module Maps(24) were used to create signature scores for each sample, summarizing the level of gene expression or DNA methylation across a set of genes or CpG loci, respectively.

Functional annotation of poised, hypermethylated genes was carried out using DAVID according to described protocol(25).

Assessment of statistical significance of observed increase in DNA methylation estimates from pyrosequencing of PEA1 and PEA2 cell lines was performed using the binomial distribution implemented in R, with null hypothesis based on an assumed probability of 0.5 of each locus showing higher methylation in PEA2 than PEA1.

Availability of data and material

ChIP-seq data for 4 primary tumour samples and 4 normal tissue samples, along with microarray data from 4 primary tumour samples, will be made available through Gene Expression Omnibus (GSE107931). RNA-seq and DNA methylation data from ICGC ovarian cancer patients at primary presentation and relapse are available from the ICGC Data Portal (https://icgc.org/icgc/cgp/67/304/809). All processed datasets, annotation resources and data analysis code will be made available for public access on the first author's Imperial College website (http://www.imperial.ac.uk/people/e.curry).

RESULTS

Bivalent chromatin marks in primary ovarian tumours associate with gene silencing

Chromatin was extracted from 4 primary High Grade Serous Ovarian tumours, 3 healthy ovarian surface epithelium samples and 1 fallopian tube epithelium sample. Following extraction, Chromatin Immunoprecipitation (IP) in conjunction with high-throughput sequencing (ChIP-seq) was performed (as described in materials and methods) to identify genome-wide positioning of the permissive histone mark H3K4me3 and the repressive mark H3K27me3. Regions of significant enrichment of IP over input DNA control, referred to as 'peaks,' were computed for each sample and each histone mark using MACS(20). Genes with peaks in both active and repressive histone marks lying within 2kb of their TSS were called bivalently-marked. All samples showed evidence of bivalent promoters, with all tumour samples showing significantly more genes carrying bivalent marks than would be expected from random overlap of the individual active and repressive marks (see Supplementary Information). The number of bivalently marked genes in each sample is markedly correlated (Pearson Correlation Coefficient 0.856, p=0.0067) to the number of genes marked with H3K27me3 in each sample, but not H3K4me3. That is, samples with more bivalently-marked genes tend to have more H3K27me3-marked genes, but not more H3K4me3-marked genes, than samples with fewer bivalently-marked genes. This suggests that the H3K27me3 mark is more critical in creating the bivalent chromatin state, supporting the well-established links between PRC2 and repression of gene expression at bivalent promoters. It also suggests that the set of bivalently-marked genes in each profiled tissue is only a subsample of the full complement that could have been observed. This motivated us to identify regions of the genome that show a disposition toward bivalency across multiple tumours but not in the normal tissue samples.

Combining this with the set of bivalently marked genes from a previous experiment(14) gave us a list of genes with evidence for poised promoter chromatin states in multiple primary

high-grade serous ovarian tumours, but not in normal ovarian surface or fallopian epithelia (provided in Supplementary Table S2).

Previous work showed that genes with bivalent promoters were more likely to be transcriptionally silenced in the profiled tumour than genes with active promoter marks(14). Applying the same analysis strategy to gene expression microarray data obtained from the tumours that were used for ChIP-seq profiling of bivalent chromatin, the average proportion of present, marginal and absent detection calls for all probe-sets mapping to bivalently-marked genes were compared against the proportions for probe-sets mapping to genes with H3K4me3 mark only. Genes with bivalently-marked promoters were significantly less likely to be expressed at a detectable level (Chi-squared test p<2.2e-16), suggesting that the set of bivalently-marked genes are representative of a degree of functional epigenetic silencing.

Poised genes in primary tumours are enriched for regulators of cell fate
Genes marked with bivalent promoters in ES cells are considered to be key determinants of differentiation pathways, being in a poised state that retains the capacity to either increase or decrease transcription rates rapidly when required. The genes harbouring proximal bivalent chromatin domains in the ovarian tumours analysed are highly enriched for Polycomb targets with poised promoters in ES cells (Hypergeometric test p=5.8e-100, PRC2 targets obtained from (26)). Although the cell of origin of serous ovarian cancers remains contentious with both ovarian surface and fallopian tube epithelia being suggested(27,28), a recent study provided molecular characterization of a definitive ovarian and fallopian tube epithelial stem cell population in the mouse, marked by expression of Lgr5(16). The set of genes with bivalently-marked chromatin in the ovarian tumour samples was significantly enriched for human orthologs of genes that were downregulated in the Lgr5+ stem cell compartment relative to Lgr5- comparison (Hypergeometric test p= 0.006). As there was no such enrichment for orthologs of the genes upregulated in Lgr5+ cells, it would appear that repression of transcription of genes marked with bivalent chromatin in ovarian tumour cells

would keep the cells in a more 'stem-cell like' state, but maintaining the capacity to lose stem cell associated features rapidly as and when required. Table 1 displays a summary of the numbers of genes in each set.

Bivalent chromatin domains and DNA hypermethylation during acquisition of drug resistance

To examine the hypothesis that bivalent promoters in tumours at presentation may predispose genes towards hypermethylation during acquisition of drug resistance, we have integrated our ChIPseq data with RNA expression and DNA methylation data generated by the International Cancer Genome Consortium (ICGC) of high grade serous ovarian cancer, including paired samples from patients both at primary, chemo-naïve, presentation and upon relapse following platinum-based chemotherapy(22). It has previously been shown that key differentiation genes bivalently marked with promoters in ES cells are predisposed toward hypermethylation during tumour development(10,29). Consistent with this, the set of promoters we identify that have hypermethylation in primary tumours relative to normal fallopian tube tissue samples show enrichment for ES Polycomb target genes (Limma Gene Set Test p=0.01). Further, we see an enrichment of the ES Polycomb targets among the set of genes which gain promoter CpG methylation at drug-resistant relapse compared to primary chemo-naïve tumours (Hypergeometric test p=2.1e-08). These ES Polycomb target genes also show a systematic downregulation of gene expression from primary to relapse samples (Limma Gene Set Test p=6.0e-10). Genes with proximal CpG loci gaining methylation, compared to all genes, show a significant systematic downregulation of gene expression from primary to relapse samples (Limma Gene Set Test p=3.5e-05), arguing that the gain of DNA methylation is associated with the reduction in expression.

To evaluate systematic changes of promoter CpG methylation following chemotherapy across genes marked with bivalent chromatin domains in primary, chemo-naïve, ovarian tumour cells, we calculated the average change in CpG methylation between matched primary and relapse samples from the same patient. Comparing the corresponding

distribution for loci in bivalently-marked promoters with that of all promoter CpGs, we see a systematic shift towards greater hypermethylation (t-test p=0.036, Supplementary Figure S1). We computed Module Map scores(24) to summarize promoter CpG methylation across entire gene sets in each sample for illustration in the heatmap in Figure 2A, with average module map scores across primary and relapse samples shown in Figure 2B. To further demonstrate that this effect was due to more hypermethylation and not less hypomethylation, we extracted only those loci with greater than 2-fold average increase in methylation in relapse samples compared to primary tumours (given in Supplementary Table S3), and found that this set was enriched for the set of promoters with bivalent chromatin domains identified by ChIPseq in primary ovarian tumours (Hypergeometric test p=0.004). Functional annotation of the list of genes that are both bivalently marked and show at least 2-fold increase in DNA methylation upon relapse suggests a link between this hypermethylation and a suppression of cellular differentiation (GO terms 'pattern specification process' and 'cell fate commitment' were enriched with adjusted p-values 7.2e-8 and 1.9e-4, respectively, a full table of enriched GO terms is provided in Supplementary Table S4). Some of the samples are observed to have clearer systematic promoter hypermethylation of these target genes than others. In fact there is considerable heterogeneity, both within the primary and the relapse tumours, in the levels of DNA methylation for even the gene set representing the most consistently differentially-methylated between the groups. This may arise from multiple alternative mechanisms for silencing being employed in the tumours, as found in (30).

In primary tumour and relapse pairs taken from the ICGC patient cohort (n=10), gene expression profiling by RNA-seq data shows a systematic down-regulation of the set of genes marked with bivalent chromatin domains in primary ovarian tumour cells (Limma Gene Set Test p=0.037). We computed Module Map scores to summarize expression levels across entire gene sets in each sample for illustration in the heatmap in Figure 2C, with average module map scores across primary and relapse samples shown in Figure 2D.

Interestingly we also see that the orthologs of Lgr5-downregulated genes in the mouse ovarian stem cell model show a strong systematic down-regulation of expression in relapsed samples relative to primary tumours (Limma Gene Set Test p<1e-12). This latter observation implies that tumour cells upon relapse have a more stem-cell-like gene expression profile and that bivalent chromatin marks may provide a mechanism for the observation that high grade serous ovarian cancer cells enhance stem cell like characteristics following exposure to chemotherapy(10) (11). The individual PRC2 complex members EZH2, EED and SUZ12 all showed higher median expression in the relapse samples relative to primary samples, but this difference was only statistically significant in the case of SUZ12 (paired t-test p=0.025). Further, we saw significant overexpression of DNMT3B (p=0.047), which suggests that transcriptional upregulation of this *de novo* DNA methyltransferase may be involved mechanistically in the observed epigenetic silencing of bivalently-marked genes.

We computed the difference in average promoter methylation and the difference in expression for each gene in each pair of primary-relapse samples. These differences were negatively correlated (Pearson correlation -0.05, product-moment test p=0.002), indicating increases in promoter methylation tending associating with decreases in expression. The set of bivalently marked genes showed on average a greater increase in methylation and greater decrease in expression between primary and relapse samples than other genes (t-test p=0.09 and p=0.08, respectively). The absolute promoter methylation levels of bivalent genes were similarly distributed to those of all other genes, but with more frequent occurrence of methylation >10% (Supplementary Figure S2). This relationship was observed in both primary and relapse samples.

To suggest key pathways affected by the epigenetic change during chemotherapy and relapse, we computed the average change in promoter methylation and gene expression across all bivalently-marked genes in pathways from the Consensus Pathway Database (31), from primary to relapse samples. A significant negative correlation was observed at the

pathway level between change in methylation and change in expression (Pearson correlation coefficient -0.2, $p<2 \times 10^{-16}$). We noted systematic increases in methylation and decreases in expression of genes in 'BMP signalling' and 'MicroRNAs in cancer' pathways (Supplementary Figure S3). miRNA signalling has previously been demonstrated to be regulated through DNA methylation (32) and implicated in drug resistance (33) in ovarian cancer. BMP signalling is known to regulate differentiation (34), and less differentiated stem cells are known to be more resistant to chemotherapy (35). These observations therefore suggest possible mechanisms relating to the acquired drug resistance observed in these relapsed tumours.

A set of candidate loci from the bivalently-marked genes showing hypermethylation on relapse were selected for independent evaluation in the PEA1/PEA2 isogenic cell line pair (17). Bivalent marks at promoters for DXL5, TRIM7, CYP21B1 and HOXD10 in the primary tumour derived cell line PEA1 were confirmed through ChIP-qPCR. Fold-enrichment from IP with antibodies targeting each of H3K4me3 and H3K27me3, relative to input DNA, is shown in Figure 3A. All selected loci show marked enrichment of the target regions through both H3K4me3 and H3K27me3 ChIP. Levels of DNA methylation for CpG sites in these promoters were assayed through bisulphite pyrosequencing, in both primary tumour derived cell line PEA1 and the cell line PEA2, which was derived from the same patient as PEA1 but after chemotherapy and subsequent relapse (17). Methylation levels (an average of 3 replicates) for these CpG sites are presented in Figure 3B, showing an increase in methylation in PEA2 relative to PEA1 for 11 of the 14 loci. Assuming under randomness that each CpG site would have a 50% chance of being measured at a higher level in PEA2 relative to PEA1, we have $p=0.02$ for seeing at least that many loci hypermethylated in PEA2 purely by chance. Of the 4 genes tested, only TRIM7 didn't feature at least 1 CpG site with >10% increase in methylation. This independent evaluation supports the systematic promoter hypermethylation of bivalently-marked genes that we established based on analysis of patient samples. To relate changes in DNA methylation to changes in gene expression, we

mapped RNA-seq reads from the PEA1 and PEA2 cell lines to these candidate genes and normalized to the number of reads mapping to GAPDH in each replicate (co-ordinates and counts given in Supplementary Table S5). Figure 3C shows 3 of the 4 genes markedly down-regulated in PEA2, and a systematic decrease across the entire bivalently marked gene set (Wilcoxon signed rank test $p<2x10^{-16}$) is observed, confirming a general trend for silencing of these genes following exposure to chemotherapy.

DISCUSSION

We previously proposed a model in which overexpression of EZH2 in cancer cells, particularly those with lower proliferation rate and greater tumorigenic capacity, leads to aberrant marking of promoters with bivalent chromatin domains, predisposing these loci to epigenetic silencing via DNA methylation during the course of tumour evolution towards an acquired drug resistant state(15). We previously presented evidence of bivalent chromatin domains in a primary ovarian tumour, showed that these were enriched for Polycomb targets associated with developmental regulation, and linked the bivalently marked genes to transcriptional silencing in tumour-initiating cell enriched subpopulations and chemotherapy-resistant models of ovarian cancer cell lines(14). However, access to paired tissue samples from patients prior to and after exposure to chemotherapy in our present study allowed us to demonstrate systematic epigenetic silencing of bivalently marked gene sets occurring *in patients* in the course of treatment. During this time, the cancer cells have been exposed to platinum-based chemotherapy and have both regrown a tumour mass and altered their drug sensitivity so that the same cytotoxic chemotherapy is less effective at therapeutic doses. While genetic change such as reversion of Brca1 mutations are important in drug resistance and treatment failure at relapse, epigenetic change such as DNA methylation also has a key role (6,36). We observe bivalent genes defined from primary HGSOC tumours compared against normal ovarian and tubal epithelia, and to canonical Polycomb targets defined in ES cells, which gives assurance that the gene sets analysed are not overly dependent on data from individual subjects. Given that we see lower expression of Lgr5-downregulated orthologs in relapse tissue, and functional annotation of the genes found to be bivalently-marked and hypermethylated upon release points to a signature associated with cell fate specification, we believe that epigenetic reprogramming associated with drug resistant relapse could reflect a more stem-like phenotype through suppression of differentiation. Alternatively, it may be that stem cells and tumour cells share common mechanisms of poised chromatin states. Thus cancer cells may also be using the regulation of poised

17

chromatin states to flexibly respond to selective pressures allowing increased plasticity. It can be seen that there is variability in the extent of hypermethylation of these gene sets, with some patients showing greater increases in methylation between primary and relapse than others. This cohort is too small to consider whether there are clinical implications of the extent of hypermethylation of bivalently marked genes, but this would be interesting to investigate should methylation profiles be obtained from more paired samples.

With available data we cannot differentiate between poised states leading to the acquisition of silencing during chemotherapy, as opposed to bivalently marked loci already being fully silenced in (drug-resistant) clones within the primary tumour sample which survived and grew out following initial treatment to form the relapse. The only way to address this definitively would be to use single-cell genetic and epigenetic profiling of all tumour samples, which remains technically infeasible at this time. It also remains to be demonstrated whether or not any of the epigenetic silencing events we have observed are directly contributing to the functional acquisition of drug resistance in high grade serous ovarian cancer cells, although our set of genes with bivalent promoters in tumours included FLNC and MDK, previously identified key drivers of epigenetically-mediated acquisition of drug resistance in ovarian tumour cell line models(6). Proof of a direct causal link would require further experimental evidence. Such experiments would require emergence of epigenetic editing technology (e.g. (37)) to induce the presence or absence of bivalent marks at PRC2 target gene promoters, then follow an appropriate model through acquisition of drug resistance, monitoring changes to DNA methylation and gene expression. As such technology is still in development, experimental proof of causality is beyond the scope of this study.

Importantly, we have shown that the histone modification profile from a primary tumour at initial clinical presentation is informative of the DNA methylation and gene expression state of drug-resistant tumour at relapse many months after treatment. This suggests that drug-resistant disease could theoretically be targeted before it develops, through a personalized

medicine strategy that utilizes epigenetic profiles obtained from a primary tumour sample to predict the most likely drivers of subsequent resistance. One can imagine a situation in which the bivalent chromatin profile of a tumour is mapped and used to select, from a list of available therapeutics targeting genes or pathways known to cause chemotherapy resistance, those with targets most likely to become altered during disease relapse. Alternatively, it may be the case that the epigenetic silencing events are acquired during chemotherapy and only in a poised state at presentation, a scenario that might be implied by the observation of overexpression of the PRC2 complex member SUZ12 and the *de novo* DNA methyltransferase DNMT3B on drug-resistant relapse. The emergence of chemotherapy resistance is facilitated by the ability of cancer cells to alter their epigenetic landscape, particularly via controlling bivalent chromatin domains and DNA methylation. Treating patients with drugs that target these processes, either along with chemotherapy or in a maintenance setting during remission, could prevent the cancer cells from adapting their epigenetic landscape to acquire chemotherapy resistance.

We have shown here that ovarian tumour cells harbour the bivalent chromatin domains that are characteristic of pluripotent cells held in an undifferentiated state poised to undergo rapid phenotypic change on lineage commitment. Specifically, we see poised chromatin states leading to epigenetic silencing of genes involved in stem cell differentiation, implying this disruption of normal differentiation processes as a potential mechanism by which tumour cells could adapt to the presence of cytotoxic chemotherapy and that common mechanisms give rise to drug resistant populations as observed during normal cell differentiation. Promoters kept in a poised state by the presence of these domains would enable cells not immediately killed by cytotoxic chemotherapy to adapt to the changes in their environment brought about by this treatment, and repopulate a tumour with cells re-programmed to suppress cellular mechanisms of drug sensitivity and treatment response. While we observe these similarities of ovarian cancer cells to ovarian stem cells, we do not have data to assess the extent to which cancer cells derive from stem cells. It would be possible for these

cells to use alternative mechanisms of establishing and maintaining bivalent chromatin domains, and that these could be located in different positions throughout the genome. Lineage tracing experiments to investigate tumour formation from Lgr5-expressing cells would be of interest to us but beyond the scope of this study. In this context we highlight the observation that mouse Lgr5+ stem cells of the ovary and fallopian tube overexpress Ezh2 relative to the Lgr5- differentiated ovarian and tubal epithelia(16). Tumour-initiating subpopulations of human cancer cell lines with stem cell-like properties also overexpress EZH2 relative to non-tumour initiating subpopulations(11). If both the normal adult stem cell compartment of the ovary and tumour initiating cells of ovarian cancers overexpress the enzyme responsible for deposition of the H3K27me3 histone mark, which is a key step in formation of bivalent chromatin domains, it may suggest that this mechanism is already present in the normal physiology of the ovary. This, in turn, would indicate that high grade serous ovarian cancers are physiologically disposed towards this capacity for epigenetically driven acquired drug resistance. We speculate that therapeutics targeting aberrant PRC2 activity in tumour cells during remission, such as histone methyltransferase inhibitors, could prevent acquisition of resistance by DNA methylation to widely-used cytotoxic chemotherapies and offer a means of preventing drug resistance emergence.

In conclusion, we have shown that high grade serous ovarian cancer at initial presentation have bivalently marked chromatin domains, containing active H3K4me3 and repressive H3K27me3 histone marks, which define gene sets that are more likely to have increased promoter CpG methylation and reduced gene expression at patient relapse with resistant disease. This provides the first evidence from patient samples that epigenetic silencing during acquired drug resistance could be underpinned by pre-existing histone modifications associated with genes being in a poised state.

ACKNOWLEDGEMENTS

REFERENCES

1. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P*, et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science **2002**;298:1039-43
2. Pauler FM, Sloane MA, Huang R, Regha K, Koerner MV, Tamir I*, et al.* H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. Genome Research **2009**;19:221-33
3. Fisher CL, Fisher AG. Chromatin states in pluripotent, differentiated, and reprogrammed cells. Current Opinion in Genetics & Development **2011**;21:140-6
4. Voigt P, LeRoy G, Drury WJ, Zee BM, Son J, Beck DB*, et al.* Asymmetrically modified nucleosomes. Cell **2012**;151:181-93
5. Baylin SB, Jones PA. A decade of exploring the cancer epigenome — biological and translational implications. Nature Reviews Cancer **2011**;11:726-34
6. Zeller C, Dai W, Steele NL, Siddiq A, Walley AJ, Wilhelm-Benartzi CSM*, et al.* Candidate DNA methylation drivers of acquired cisplatin resistance in ovarian cancer identified by methylome and expression profiling. Oncogene **2012**;31:4567-76
7. Huang S, Benavente S, Armstrong EA, Li C, Wheeler DL, Harari PM. p53 modulates acquired resistance to EGFR inhibitors and radiation. Cancer Research **2011**;71:7071-9
8. Zhang S, Huang W-C, Li P, Guo H, Poh S-B, Brady SW*, et al.* Combating trastuzumab resistance by targeting SRC, a common node downstream of multiple resistance pathways. Nature Medicine **2011**;17:461-9
9. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nature Reviews Genetics **2009**;10:295-304
10. Easwaran H, Johnstone SE, Van Neste L, Ohm J, Mosbruger T, Wang Q*, et al.* A DNA hypermethylation module for the stem/progenitor cell signature of cancer. Genome Research **2012**;22:837-49
11. Rizzo S, Hersey JM, Mellor P, Dai W, Santos-Silva A, Liber Diea. Ovarian cancer stem cell-like side populations are enriched following chemotherapy and overexpress EZH2. Molecular Cancer Therapeutics **2011**;10:325-35
12. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG*, et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature **2002**;419:624-9
13. Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA*, et al.* EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proceedings of the National Academy of Sciences **2003**;100:11606-11
14. Chapman-Rothe N, Curry E, Zeller C, Liber D, Stronach E, Gabra H*, et al.* Chromatin H3K27me3/H3K4me3 histone marks define gene sets in high grade serous ovarian cancer that distinguish malignant, tumour sustaining and chemo-resistant ovarian tumour cells. Oncogene **2012**;32:4586-92
15. Brown R, Curry E, Magnani L, Wilhelm-Benartzi CS, Borley J. Poised epigenetic states and acquired drug resistance in cancer. Nature Reviews Cancer **2014**;14:747-53
16. Ng A, Tan S, Singh G, Rizk P, Swathi Y, Tan TZ*, et al.* Lgr5 marks stem/progenitor cells in ovary and tubal epithelia. Nature Cell Biology **2014**;16:745-57
17. Langdon SP, Lawrie SS, Hay FG, Hawkes MM, McDonald A, Hayward IP*, et al.* Characterization and properties of nine human ovarian adenocarcinoma cell lines. Cancer Research **1988**;48:6166-72
18. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature Methods **2012**;9:357-9

19.   Carroll TS, Liang Z, Salama R, Stark R, de Santiago I. Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data. Frontiers in Genetics **2014**;5

20.   Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. Nature Protocols **2012**;7:1728-40

21.   Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, Odom DT. ChIP-seq: using high-throughput sequencing to discover protein–DNA interactions. Methods **2009**;48:240-8

22.   Patch A-M, Christie EL, Etemadmoghadam D, Garsed DW, George J, Fereday S, *et al.* Whole–genome characterization of chemoresistant ovarian cancer. Nature **2015**;521:489-94

23.   Smyth GK. Limma: linear models for microarray data. Bioinformatics and Computational Biology Solutions using R and Bioconductor: Springer; 2005. p 397-420.

24.   Segal E, Friedman N, Koller D, Regev A. A module map showing conditional activity of expression modules in cancer. Nature Genetics **2004**;36:1090-8

25.   Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols **2008**;4:44-57

26.   Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. Cell **2006**;125:301-13

27.   Auersperg N. Ovarian surface epithelium as a source of ovarian cancers: unwarranted speculation or evidence-based hypothesis? Gynecologic Oncology **2013**;130:246-51

28.   Vang R, Shih IM, Kurman RJ. Fallopian tube precursors of ovarian low-and high-grade serous neoplasms. Histopathology **2013**;62:44-58

29.   Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, *et al.* A stem cell–like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nature Genetics **2007**;39:237-42

30.   Teschendorff AE, Zheng SC, Feber A, Yang Z, Beck S, Widschwendter M. The multi-omic landscape of transcription factor inactivation in cancer. Genome Medicine **2016**;8:89

31.   Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Research **2011**;39:D712-D717

32.   Iorio MV, Visone R, Di Leva G, Donati V, Petrocca F, Casalini P, *et al.* MicroRNA signatures in human ovarian cancer. Cancer Research **2007**;67:8699-707

33.   van Jaarsveld MT, Helleman J, Berns EM, Wiemer EA. MicroRNAs in ovarian cancer biology and therapy resistance. The International Journal of Biochemistry & Cell Biology **2010**;42:1282-90

34.   Winnier G, Blessing M, Labosky PA, Hogan B. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes & Development **1995**;9:2105-16

35.   Alvero AB, Chen R, Fu H-H, Montagna M, Schwartz PE, Rutherford T, *et al.* Molecular phenotyping of human ovarian cancer stem cells unravels the mechanisms for repair and chemoresistance. Cell Cycle **2009**;8:158-66

36.   Swisher EM, Sakai W, Karlan BY, Wurz K, Urban N, Taniguchi T. Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance. Cancer Research **2008**;68:2581-6

37.   Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nature Biotechnology **2015**;33:510-7

TABLES

Table 1: Summary of numbers of genes with bivalent marks in ovarian tumour and normal tissues

Tissue Source	Bivalently Marked Genes	ES Polycomb Targets[a] (out of 1567)	Ovarian Lgr5 Targets[b] (out of 203)
Ovarian Tumour	747	217	11
Normal (OSE, FT)	207	23	3

a. Previously described Polycomb targets in ES cells(26)

b. Orthologs of genes downregulated in Lgr5+ mouse ovary cells(16).

FIGURE LEGENDS

Figure 1 – Graphical Summary of Analytical Approach and Key Findings

ChIP-seq for H3K27me3 and H3K4me3 (in comparison to input DNA) was used to obtain genome-wide bivalent mark profiles in 4 primary ovarian tumours and 4 normal tissue samples, supplemented by the profile from a primary ovarian tumour obtained from our previous study. Genes with both H3K27me3 and H3K4me3 marks within 2kb of the TSS in any of the tumours but none of the normal tissues were taken forward as representative of poised genes in primary ovarian cancers. These genes systematically gained methylation and reduced expression in ovarian cancer patient samples obtained at relapse following cytotoxic chemotherapy treatment, relative to their primary tumour samples. The set of genes which were both poised in primary tumours and epigenetically silenced upon relapse were strongly enriched for regulators of differentiation, including the classical ES cell Polycomb targets.

Figure 2 – Module Maps Summarizing DNA Methylation and Gene Expression Signatures

Heatmaps show module map scores for gene sets in terms of DNA methylation (A) and gene expression (C) in individual primary tumour and relapse ascites samples from the ICGC high-grade serous ovarian cancer cohort. DNA methylation module map scores $\log_2$-transformed in (A). Set of genes with bivalent chromatin domains in ovarian tumours is labelled 'tumour.bivalent', set of polycomb targets in ES cells is labelled 'prc2', set of orthologs of genes repressed in Lgr5+ mouse ovary stem cells is labelled 'lgr5.targets', sets of genes up- and down-regulated in expression in relapse ascites relative to primary tumours are labelled 'relapse.hi' and 'relapse.lo' respectively, and sets of genes hyper- and hypo-methylated in relapse ascites relative to primary tumours are labelled 'relapse.hyper' and 'relapse.hypo' respectively. Colour bar along top of the heatmaps indicates primary samples with blue and relapse samples with pink. Average module map scores across all primary

24

samples (blue) and all relapse samples (pink) are shown in bar charts (B) for DNA methylation and (D) for gene expression. These show that methylation tends to increase in relapse for the genes with bivalent marks in primary ovarian tumours, for the polycomb targets in ES cells and for genes decreasing expression on relapse. Methylation tends to decrease in relapse for the genes increasing expression on relapse. Gene expression shows the opposite trend, decreasing for the genes with bivalent marks in primary ovarian tumours and for the polycomb targets in ES cells.

Figure 3 – Evaluation of bivalent chromatin marks and hypermethylation in PEA1/PEA2 ovarian cancer cell line model

(A) Fold-enrichment from H3K4me3 and H3K27me3 immunoprecipitation of PEA1 cell line DNA at promoter regions of candidate genes, based on ChIP-qPCR quantification. Each IP shown is an average of two independent chromatin extractions, pull-down and amplification. Values shown reflect concentration of target region DNA relative to all DNA in the sample, normalized to no-IP input DNA. (B) Quantification of promoter CpG methylation levels in PEA1 and PEA2 cell lines by bisulphite pyrosequencing. Values shown represent proportion of sequenced DNA fragments with the corresponding CpG site methylated, out of all sequenced DNA fragments mapping to that CpG site. Each assay reflects the average of 3 independent DNA extractions. (C) Quantification of gene expression levels in PEA1 and PEA2 cell lines from RNA-seq. Values shown represent number of reads mapped to each gene relative to number of reads mapping to GAPDH, in each of 3 replicates.
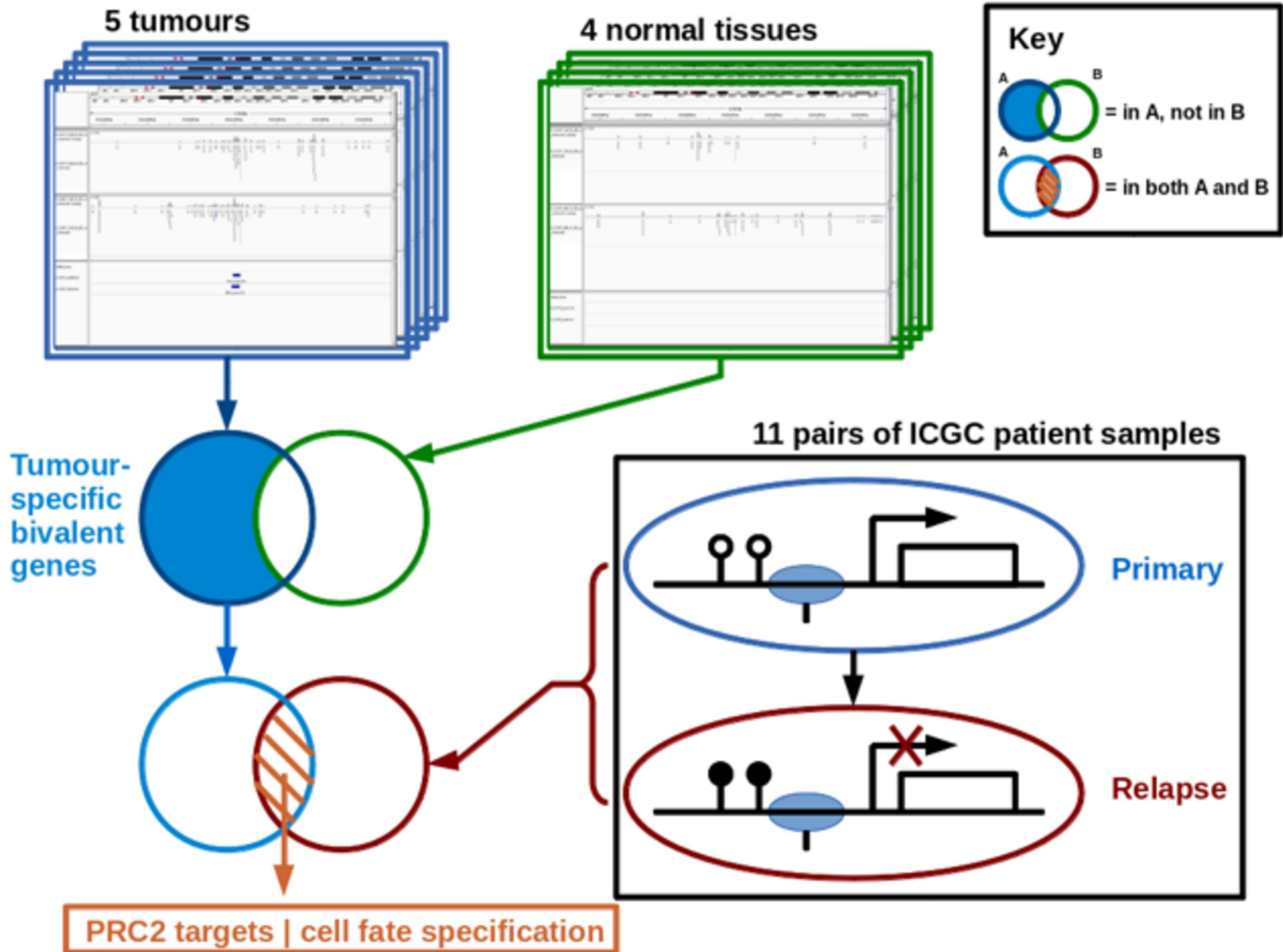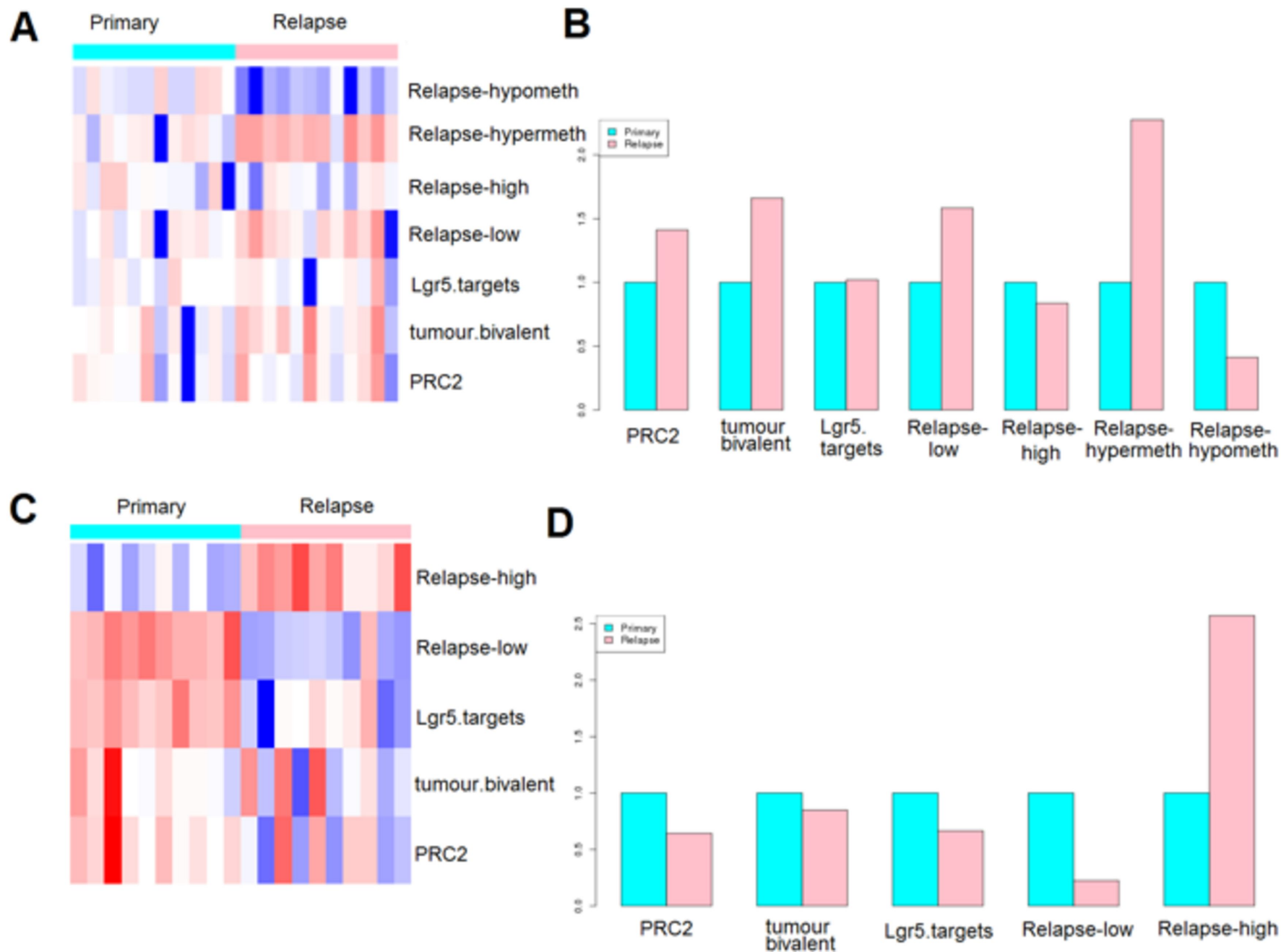
Figure 1

Figure 2

Figure 3