

Interplay between base excision repair protein XRCC1 and ALDH2 predicts overall survival in lung and liver cancer patients

Xin Chen^{1,2,3,*}, Arnaud J. Legrand^{1,*}, Siobhan Cunniffe¹, Samuel Hume¹, Mattia Poletto¹, Bruno Vaz¹, Kristijan Ramadan¹, Dengfu Yao^{2,†}, Grigory L. Dianov^{1,4,5,†}

¹CRUK & MRC Oxford Institute for Radiation Oncology, University of Oxford, Department of Oncology, Old Road Campus Research Building, OX3 7DQ Oxford, UK

²Research Centre of Clinical Medicine, Affiliated Hospital of Nantong University, Jiangsu, China

³School of Life Science, Nantong University, Nantong, China

⁴Institute of Cytology and Genetics, Russian Academy of Sciences, Lavrentyeva 10, Novosibirsk 630090, Russian Federation

⁵ Novosibirsk State University, Novosibirsk 63000, Russian Federation

*These authors contributed equally to this work

†Co-corresponding authors

Address correspondence to:

Grigory Dianov, CRUK & MRC Oxford Institute for Radiation Oncology,
University of Oxford, Department of Oncology,
Old Road Campus Research Building,
OX3 7DQ Oxford, UK
tel. +441865617325, e-mail: grigory.dianov@oncology.ox.ac.uk

Dengfu Yao,
Research Centre of Clinical Medicine,
Affiliated Hospital of Nantong University,
Jiangsu, China
e-mail: yaodf@ahmnc.com

Running Title: Interplay between XRCC1 and ALDH2 in lung and liver cancers

Abstract

Background To deliver efficacious personalised cancer treatment, it is essential to characterise the cellular metabolism as well as the genetic stability of individual tumours. In this study, we describe a new axis between DNA repair and detoxification of aldehyde derivatives with important implications for patient prognosis and treatment.

Methods Western blot and qPCR analyses were performed in relevant non-transformed and cancer cell lines from lung and liver tissue origin in combination with bioinformatics data mining of The Cancer Genome Atlas database from lung and hepatocellular cancer patients.

Results Using both biochemical and bioinformatics approaches, we revealed an association between the levels of expression of the aldehyde detoxifying enzyme aldehyde dehydrogenase 2 (ALDH2) and the key DNA base excision repair protein XRCC1. Across cancer types, we found that if one of the corresponding genes exhibits a low expression level, the level of the other gene is increased. Surprisingly, we found that low ALDH2 expression levels associated with high XRCC1 expression levels are indicative for a poor overall survival, particularly in lung and liver cancer patients. In addition, we show that Mithramycin A, a XRCC1 expression inhibitor, efficiently kills cancer cells expressing low levels of ALDH2.

Conclusions This study suggests that lung and liver cancers require efficient single-strand break repair for their growth in order to benefit from a low aldehyde detoxification metabolism. We also propose that the ratio of XRCC1 and ALDH2 levels might be a useful prognosis tool in these cancer types.

Keywords: ALDH2, aldehydes, DNA damage, XRCC1, base excision repair, genetic instability, liver and lung carcinomas, mithramycin A

1 Introduction

Aldehydes are abundant organic molecules that can be assimilated from food or produced within cells as by-products of cellular metabolism. Among many cellular aldehydes, acetaldehyde (AcAl) is the most studied, particularly for its importance in cancer metabolism [1-3]. Endogenous AcAl is mainly produced by the cellular processing of alcohol and also as an intermediate in sugar metabolism [4]. Aldehydes are highly reactive molecules, generating a range of DNA modification products including DNA strand crosslinks and DNA-protein crosslinks [4]. Due to their toxicity, AcAls are processed in human cells by detoxification enzymes such as aldehyde dehydrogenase 2 (ALDH2), which oxidises AcAl to acetate [5]. Until recently, very little was known about how cells respond to DNA damage induced by AcAls. New studies suggest that AcAl-induced DNA damage is likely to be processed by multiple DNA repair pathways due to its complexity [6-9]. Two DNA repair pathways involved in repair of DNA crosslinks and DNA double strand breaks have been implicated in repair of AcAl-induced DNA damage thus far [6]. However, in addition to DNA strand crosslinks and DNA-protein crosslinks, AcAls also generate multiple DNA base modifications and neither of these pathways can deal with base modifications. N²-ethylidenedeoxyguanosine is the main adduct, and its reduced form N²-ethyldeoxyguanosine [10] has been directly linked to ALDH2, since its levels increase in ALDH2 knockout mice treated with ethanol [11].

Base excision repair (BER) is the primary mechanism responsible for repairing most endogenous DNA lesions, including numerous types of damage to nitrogenous bases that can be induced by AcAls [12]. BER is a well understood DNA repair pathway [13, 14]. It is initiated by a DNA glycosylase that recognizes the damaged DNA base and cleaves the N-glycosylic bond that binds the DNA base to the sugar-phosphate backbone. A new site without a base residue (also called an abasic site, an AP site or an apurinic/apyrimidinic site) is further processed by an AP endonuclease (APE1 in human cells) that cleaves the phosphodiester bond next to the AP site, thereby generating a DNA strand break containing a hydroxyl residue at the 3'-terminal and deoxyribose phosphate at the 5'-end. Further, DNA polymerase β , using its AP-lyase activity, removes deoxyribose phosphate from the 5'-end and simultaneously adds one nucleotide to the 3'-end of the nick. DNA repair is completed by sealing the ends by the complex of two proteins, XRCC1 and DNA ligase III α [12]. XRCC1 is a scaffold protein that is essential for the formation and stabilization of the ternary DNA polymerase β -XRCC1-DNA Ligase III α complex that is necessary to complete BER [15]. Hence, cells deficient in XRCC1 are characterized by reduced DNA repair and increased genomic instability (reviewed in [16]).

We previously showed that loss of BER triggers metabolic changes, notably by increasing the activity of the one-carbon cycle [17], which endogenously produces and metabolises formaldehyde, an even more reactive molecule than AcAl [2]. If BER-deficient cells increase the activity of this cycle, it is reasonable to suggest that this would in turn increase the concentration of formaldehyde in cells as both a waste product and a reaction intermediate. We thus hypothesised that BER-deficient cells may have increased levels of aldehydes. As a result, these cells would be in need of increasing an aldehyde detoxification system, such as ALDH2, to counteract AcAl and formaldehyde toxicity. Here, we tested this hypothesis along with its potential consequences on cancer aggressiveness and prognosis.

2 Materials and methods

2.1 Cell cultures and drug

The non-transformed bronchial epithelium cell line LIMM-NBE1 has previously been described in [18]. The normal human fibroblast cell lines TIG-1 and WI38 were obtained from the Coriell Institute Cell Repository. The non-small cell lung carcinoma-derived cell line H1299 was obtained from the American Type Culture Collection (ATCC), while the liver carcinoma-derived cell line JHH4 was kindly provided by Prof Ricky Sharma (University College London). All cell lines were cultured in DMEM (Life Technologies) supplemented with either 15% FBS (for the LIMM-NBE1 cell line) or 10% FBS (for the cancer cell lines) at 37°C in a humidified atmosphere with 5% CO₂. Cells were routinely checked for mycoplasma. Mithramycin A was purchased from Enzo Life Sciences.

2.2 siRNA transfection

siRNA transfections were carried out using the Lipofectamine RNAiMAX reagent (Life Technologies) according to the manufacturer's protocol. Unless otherwise indicated, cells were transfected with 30 nM siRNA and analysed 72 hours after transfection. siRNA oligonucleotides were obtained from Eurogentec. Sequences were as follows: XRCC1 5'-AGGGAAGAGGAAGUUGGAU-3', TDP1 5'-GACCAUAUCUAGUAGUGAU-3' and ALDH2 5'-CCUCAAAUGUCUCCGGUAU-3'. Control transfections were carried out using a non-targeting siRNA (Eurogentec, SR-CL000-005), referred to as NC.

2.3 Western blotting

Whole cell extracts for Western blotting were prepared as described previously [19]. The antibodies used were directed against XRCC1 (MS-1393-P0, Neomarkers), ALDH2 (102-10056, Cambridge Biosciences), β -actin (ab6276, Abcam) and α -Tubulin (T6199, Sigma Aldrich). Detection and quantification was carried out using an Odyssey image analysis system (Li-Cor Biosciences).

2.4 DNA-protein crosslink isolation and quantification

LIMM-NBE1 cells were grown to 100% confluency in 10-cm dishes in order to synchronise them in G1 phase, as XRCC1 siRNA stops the cell cycle in this cell line. DNA-protein crosslinks (DPCs) were detected using a modified rapid approach to the DNA adduct recovery assay [20]. In brief, 1.5 to 2 x 10⁶ cells were lysed in 1 ml M buffer (MB) containing 6 M Guanidine Thiocyanate, 10 mM Tris-HCl (pH 6.8), 20 mM EDTA, 4% Triton X100, 1% Sarkosyl and 1% Dithiothreitol. DNA was precipitated by adding 1 ml 100% ethanol and washed three times in wash buffer (20 mM Tris-HCl pH 6.8, 150 mM NaCl and 50% ethanol) after which DNA was solubilized in 1 ml 8 mM NaOH. A small aliquot of the recovered DNA was digested with 50 μ g/ml proteinase K (Invitrogen) for 3 hours at 50°C and quantified using PicoGreen dye (Invitrogen) according to the manufacturer's instructions. The DNA concentration was further confirmed by slot-blot analysis followed by immunodetection with an antibody directed against dsDNA. To quantify DPCs, DNA was digested with benzonase (Invitrogen) for 30 minutes at 37°C, after which proteins were precipitated by the standard Trichloroacetic Acid protocol [21], dissolved in Laemmli buffer, resolved by SDS-PAGE gel and visualized by silver staining.

2.5 Immunofluorescence assay

Immunofluorescence was carried out following standard procedures. Briefly, cells were fixed with paraformaldehyde (4% in PBS for 15 minutes). Permeabilisation was carried out using Triton X-100 (0.2% in PBS for 10 minutes at 4°C) after which cells were saturated with 5% bovine serum albumin (BSA) in PBS for 1 hour. Subsequent incubation with an anti-ALDH2 antibody (102-10056, Cambridge Biosciences) was carried out in 5% BSA-PBS supplemented with 0.01% Tween 20. Alexa Fluor 594-conjugated secondary antibodies (Life Technologies) were used for indirect detection of the antigens. Hoechst 33342 (Life Technologies) was used to visualise nuclei. Images were taken by confocal microscopy using a Zeiss LSM 710 microscope.

2.6 Quantitative real-time PCR (qRT-PCR)

Total RNA was extracted using a RNeasy kit (Qiagen) and cDNA was prepared using a SuperScript RT-PCR system (Life Technologies) as per manufacturer's instructions. qRT-PCR was performed using a Fast SYBR® Green Master Mix (Applied Biosystems) according to the manufacturer's protocol. Reactions were carried out using a 7500 Fast Real-Time PCR System (Applied Biosystems). The comparative CT method was applied for quantification of gene expression; GAPDH and B2M were used as endogenous controls, unless otherwise stated. The following Primers were used:

ALDH2 (For: CCTCTCCAGTGGACGGATT; Rev: CGAGGTCTTCTGCAACCAG)

GAPDH (For: AGCCACATCGCTCAGACAC; Rev: GCCCAATACGACCAAATCC)

B2M (For: ATGTCTCGCTCCGTGGCCTTA; Rev: ATCTTGGGCTGTGACAAAGTC)

TBP (For: CGGTTTGCTGCGGTAATCAT; Rev: TTTCTTGCTGCCAGTCTGGAC)

XRCC1 (For: CTGGGACCGGGTCAAAT; Rev: CAAGCCAAAGGGGGAGTC)

APE1 (For: CGGACAAGGAAGGGTACAGT; Rev: CAAATTCAGCCACAATCACC)

TDP1 (For: CGCTTGTTTCTTCAGCTCAG; Rev: ACAAGCAGGATTGGCTTCTT)

2.7 Clonogenic assay

500 cells were seeded per well in 6-well plates, with three technical replicates of each variable. In Figures 8D and E, siRNA transfections were carried out 24 hours before seeding. In Figure 9A, cells were allowed to attach for at least 18 hours before Sp1 inhibitor addition, at the indicated concentrations. Sp1 inhibitor was added for 48 total hours, replenishing the inhibitor at 24 hours. Surviving cells were allowed to proliferate for 10–14 days. Cells were fixed for 30 minutes in 0.4% methylene blue dye in methanol. Only colonies of more than 50 cells were counted using the automated system GelCount™ by Oxford Optronix. Each graph represents the mean of at least 3 independent experiments.

2.8 Bioinformatics analyses

Bioinformatics analyses were performed using publicly available data from the Oncomine® (www.oncomine.org), cBioPortal (<http://www.cbioportal.org>) [22, 23] and TCGA (RNAseq data) (cancergenome.nih.gov) databases. Studies using the Oncomine® database can be found in Table S1. All data extracted from cBioPortal (Coming originally from TCGA) can be found in Table S2 (Lung cancer [NSCLC]), Table S3 (Liver cancer) and Table S4 (Esophageal cancer).

2.9 Statistical analyses

Statistical analyses were performed by the two-tailed Student's t-test using either Microsoft Excel or SPSS (IBM).

3 Results

3.1 ALDH2 expression is increased following BER deficiency

We previously showed that BER-deficient cells adjust their metabolism to prevent oxidative stress by producing anti-oxidant molecules such as GSH [17]. We hypothesised that BER-deficient cells operate in a similar way to prevent aldehyde-induced DNA damage by enhancing detoxification systems. To test this, we downregulated BER in the non-transformed cell line LIMM-NBE1 (hereafter NBE1) by reducing the expression of the key BER protein XRCC1, using siRNA. We observed a significant increase in ALDH2 expression both at the mRNA (Fig. 1A) and protein (Fig. 1B) level 72 hours after siRNA treatment. The localisation of ALDH2, as shown by immunofluorescence, was unchanged and the increase corresponds to a cytoplasmic localisation (Fig. 1C). The increase in ALDH2 expression following XRCC1 depletion was also observed in two additional normal human fibroblast cell lines TIG1 (Fig. 1D) and WI38 (Fig. 1E). APE1, POL β and XRCC1 are the essential components of BER. However, APE1 has diverse cellular functions and its depletion is highly toxic, creating many types of lesions, ranging from apurinic/apyrimidinic AP-sites to double-strand breaks [24, 25]. POL β knockout cells are BER deficient, but to a lesser extent than XRCC1 knockout cells [26]. We therefore considered XRCC1 knockdown as the optimal model of BER deficiency [17]. Nevertheless, it is also possible to simulate BER deficiency by creating unresolved single-strand breaks (SSBs) through the depletion of BER end processor protein tyrosyl-DNA phosphodiesterase 1 (TDP1), which catalyses the removal of Type I topoisomerase blocked at the 3-prime phosphate end of the DNA single-strand break [17]. Doing so, we again observed an increase in ALDH2 expression (Fig. 1F). This would suggest that it is the accumulation of SSBs that triggers ALDH2 up-regulation. Next, we investigated whether this balancing mechanism was linked to the detoxification of aldehydes and their products, DNA-protein crosslinks (DPCs [27]). Despite an increased amount of ALDH2, we observed a slight but significant increase in DPCs in XRCC1-deficient cells compared to control cells treated with unspecific siRNA (Fig. 1G, compare lanes 1 and 2). This might suggest that even with more ALDH2, DPCs are still occurring more frequently in BER-deficient cells. In positive control experiments, inhibition of aldehyde detoxification by the pan-ALDH inhibitor Disulfiram lead to substantial DPC accumulation in NBE1 cells (Fig. 1G, lane 3), confirming that ALDH2 detoxification activity can prevent DPC formation. From these results, we conclude that BER-deficient cells accumulate SSBs that might directly or indirectly, by increasing, for example, the amount of reactive aldehydes [26], provoke an increase in DPCs. In response, cells increase ALDH2 expression, potentially in order to prevent further accumulation of DPCs.

3.2 ALDH2 expression is low in most cancers whereas XRCC1 expression is high

DPCs are a major threat to DNA replication and genomic stability. Accumulation of DPCs has been observed during aging, neurodegeneration and cancer development [28]. In order to understand why cancer cells are unable to prevent the induction of DPCs and other DNA damage, caused by the accumulation of reactive aldehydes, we carried out bioinformatics analysis of ALDH2 mRNA levels in different cancer types using the OncoPrint[®] database (Fig. 2). Strikingly, we noticed that, when ALDH2 mRNA expression is significantly deregulated, it is

for a major part by downregulation (Fig. 2A and 2B). We found that the levels of ALDH2 mRNA are particularly low in melanoma, sarcoma, leukaemia and bladder cancer (Fig. 2A). It is also important to emphasize that ALDH2 mRNA is expressed at its highest levels in tissues such as the liver, brain and lungs [29]. All of these tissues, when cancerous, showed lower ALDH2 mRNA levels (Fig. 2A). Interestingly, we found that the levels of XRCC1 mRNA expression seem to generally reverse ALDH2 mRNA expression levels by being higher in most cancerous tissues (Fig. 2A and 2B). This is more obvious in sarcoma and bladder cancer, but more interestingly in liver, brain and lung cancers (Fig. 2A). Globally at the median value, we found that ALDH2 is downregulated by more than two-fold and XRCC1 is upregulated by more than two-fold across all cancer types (Fig. 2B). Altogether, this analysis suggests that the increase in DPCs in cancer might be caused by a lack of ALDH2 expression. This deficiency, however, coincides with a higher expression of XRCC1, which is required for the BER pathway.

3.3 Low ALDH2 levels predict a poor 5-year overall survival rate for lung and liver cancers but not for oesophagus cancer

To understand further the relevance of the interplay between XRCC1 and ALDH2 expression in cancer, we investigated whether the profile we observed in most cancers (i.e., low ALDH2 and high XRCC1 mRNA levels, Fig. 2) affects patient prognosis. To do so, we extracted provisional mRNA data from “The Cancer Genome Atlas” (TCGA) through the database cBioPortal [22, 23]. We selected data for lung (522 samples, Table S2) and liver (422 samples, Table S3) cancers in which the differences between ALDH2 and XRCC1 mRNA levels are pronounced (Fig. 3A and 4A) and compared these data to oesophageal cancer (186 samples, Table S4) in which XRCC1 and ALDH2 mRNA levels are very close (Fig. 5A). Looking at the overall survival rate at 5 years, we found that the XRCC1 mRNA level itself has no prognostic value in any of these cancers (Fig. 3B, 4B and 5B). On the contrary, we found that low ALDH2 mRNA levels, defined by separating value at the median, strongly predict a worse prognosis in lung and liver cancer (Fig. 3C and 4C). Interestingly, we found that the ALDH2 mRNA levels have no prognosis value in oesophageal cancer (Fig. 5C). Therefore, although XRCC1 mRNA levels alone have no prognostic value, low ALDH2 levels show a worse prognosis at 5 years, only in cancers showing a marked difference between ALDH2 and XRCC1 expression.

3.4 XRCC1 levels or XRCC1/ALDH2 ratios further stratify the effect of low ALDH2 expression on 5-year overall survival

In liver cancer patients, we observed a significant impact of ALDH2 mRNA expression on the overall survival rate, whereas the XRCC1 mRNA level itself had no predictive value (Fig. 4). Nevertheless, we stratified patients into two groups, low and high XRCC1 mRNA expression, and investigated whether this stratification improves the significance of ALDH2 mRNA expression on the prognosis. Strikingly, we found that the ALDH2 mRNA level has no prognostic value in the low XRCC1 mRNA expression group (Fig. 6A), whereas, in the high XRCC1 mRNA expression group, low ALDH2 mRNA expression was found to predict a very poor prognosis (Fig. 6B). Interestingly, we found that the XRCC1 mRNA levels are higher in the low ALDH2 mRNA samples (Fig. 6C) and that the ALDH2 mRNA levels are lower in the high XRCC1 mRNA group (Fig. 6D), confirming the above observed interplay between XRCC1 and ALDH2 expression (Fig. 1). In order to support the connection with the *in vitro* data of Fig. 1, we also used the APE1 and TDP1 mRNA levels to stratify ALDH2 expression and prognosis in liver cancer (Fig. S1). Interestingly, we found that APE1, whose depletion does not induce SSBs in their pure

form [25], does not stratify the ALDH2 prognosis value (Fig. S1A and S1B). However, TDP1, whose depletion does mimic XRCC1 loss (Figure 1F), was found to stratify the ALDH2 prognosis value in a similar fashion to XRCC1, although less powerfully (Figs. S1C and S1D). The POL β levels did not show any prognostic value in liver carcinomas (data not shown). Next, we investigated whether the XRCC1 mRNA/ALDH2 mRNA expression ratio could also be used as a prediction factor. We found that the XRCC1 mRNA/ALDH2 mRNA ratio is significantly better at predicting the 5-year overall survival rate for both lung (Fig. 7A, compare to 3C) and liver (Fig. 7B, compare to 4C) cancers than ALDH2 levels alone. In oesophageal cancer, where the XRCC1 and ALDH2 mRNA levels show no difference (Fig. 5A), the XRCC1 mRNA/ALDH2 mRNA ratio shows no improvement in predicting the overall survival rate (Fig. 7C, compare to 5C). We thus concluded that XRCC1 mRNA levels along with ALDH2 mRNA levels may be used for stratification in order to predict the 5-year overall survival rate in liver and lung cancer patients and to refine their treatment strategy.

3.5 Cancer-derived cell lines with low ALDH2 mRNA levels are sensitive to Sp1 inhibitor Mithramycin A

We showed that XRCC1 and ALDH2 mRNA levels are related and that a high XRCC1 mRNA/low ALDH2 mRNA ratio seems to be accompanied by tumour aggressiveness in patients. We then wondered whether destabilising this balance could have a deleterious effect on cancer cell survival. To do so, we selected two cancer-derived cell lines, i.e., H1299, a lung cancer cell line with a low ALDH2 expression and JHH4, a liver cancer cell line with a high ALDH2 expression (Fig. 8A). We then proceeded to knock down XRCC1 mRNA expression, ALDH2 expression, or both, and investigated the cellular survival using a clonogenic assay. We found that H1299 cells (ALDH2 low), do not tolerate decreases in both XRCC1 and ALDH2 proteins (Figs. 8B and 8D, last column), whereas JHH4 cells (ALDH2 high), cope better after a double knockdown of XRCC1 and ALDH2 (Figs. 8C and 8E, last column) most probably because they do not completely lose ALDH2 expression after knockdown (Fig. 8C). Interestingly, we found that ALDH2 knockdown alone promotes higher survival rates for both the H1299 (Fig. 8D, column 3) and JHH4 (Fig. 8E, column 3) cell lines, supporting our observation that lower ALDH2 levels are advantageous for cancer cells (Fig. 2). We also found that single knockdown of XRCC1 does not reduce the survival of H1299 cells (Fig. 8D, second column) even if their level of ALDH2 is already low. We believe that this may be due to there being some BER activity remaining as a result of incomplete siRNA knockdown of XRCC1 gene expression [25]. We thus hypothesise that H1299 cells may be sensitised by a global shutting down of the BER pathway. We have recently uncovered that inhibition of the Sp1 transcription factor decreases the expression of the BER protein APE1 [25], as well as other members of this pathway [30]. We, therefore, set out to investigate the survival of H1299 and JHH4 cells after treatment with the Sp1 inhibitor Mithramycin A. We found that H1299 cells are extremely sensitive to very low doses of Mithramycin A, whereas JHH4 cells are resistant (Fig. 9A). We also confirmed a decreased expression of XRCC1 and APE1 in JHH4 (Fig. 9B) and H1299 (Fig. 9C) cells after Sp1 inhibition by qRT-PCR. Because Sp1 regulates the whole BER pathway, Sp1 inhibition seems more potent in killing H1299 cells than depletion of XRCC1 alone. These data represent a proof of principle that, although low expression of ALDH2 gives cancer cells some growth advantage (Figs. 8D and 8E), they may still be dependent on an efficient BER pathway (Fig. 9A). In conclusion, we found that several types of cancer show a profile with a lower ALDH2 and a higher XRCC1 expression that is linked to a poor prognosis. Cancer cells with this profile could potentially be targeted by using BER enzyme inhibitors or by shutting down the pathway through BER gene expression inhibition.

4 Discussion

Metabolic changes are a cornerstone of cancer development [31, 32]. Recent scientific advances in the understanding of cancer metabolism have shown the importance of defining the pathways that may limit cancer progression with the hope of designing reliable clinical strategies with benefits for the patient [32]. In this context, the pathways to be targeted must represent unique avenues taken by cancer cells to boost their proliferation and survival. By using these particular metabolic reactions, cancer cells leave traces or clues in the form of specific waste products that can be used to trace new metabolic pathways used by them [33]. BER is one of the first DNA repair pathways deregulated during cancer progression [25]. Here, we used XRCC1 levels as a readout for BER efficiency. This is because XRCC1 depletion is a model known for the purity of the type of lesion it causes, compared to APE1 depletion, and for the robust BER deficiency it induces, compared to POL β depletion [17, 25, 26]. We have previously shown that the loss of XRCC1 triggers metabolic changes resembling those found in many cancers [17, 34]. Here, we hypothesised that alterations in aldehyde metabolism could be part of these changes. Recently, aldehydes such as AcAl, methylglyoxal and formaldehyde have generated significant interest for their role as waste products of cancer metabolism [33, 35-37]. Interestingly, aldehyde production, when linked to beneficial metabolic pathways for cancer, acts as a double-edged sword associated with high genotoxicity [33, 35]. Cancer cells rely on detoxification enzymes such as ALDH2 that can neutralise aldehydes and keep a semblance of metabolic equilibrium [5, 29, 38]. Nevertheless, cancer cells frequently exhibit an excess of aldehydes and, therefore, a high occurrence of aldehyde-related DPCs and genetic instability [9, 28, 37]. Surprisingly, we observed in this study by using bioinformatics analysis that most cancers, especially lung and liver carcinomas, gamble with lowering their expression of ALDH2 along with overexpressing XRCC1. This observation correlates with the growing body of evidence indicating a high occurrence of DPCs and their link to genetic instability in cancer [27, 28]. Interestingly, DPC accumulation in the liver is particularly oncogenic as it has been found that patients lacking the fundamental DNA-protein crosslink repair metalloprotease SPRTN/DVC1 develop very early onset aggressive hepatocellular carcinoma [39]. Indeed, we show that inhibition of ALDH2 increases the occurrence of DPCs and that low expression of ALDH2 is associated with a poor prognosis for liver cancer patients. Surprisingly, despite the existence of different ALDH isoforms, we found that only ALDH2 expression has an impact on liver cancer (data not shown). This would suggest a fundamental function of ALDH2 in aldehyde detoxification, superior to any other ALDHs.

Several non-exclusive hypotheses can be made for a putative role of BER in the control of aldehyde-related damage. Firstly, our data support the idea that BER is involved directly in the repair of aldehyde-related damage. Aldehydes attack DNA bases and these alterations can be repaired by BER [40]. In addition, it has been proposed that BER may be involved in the repair of DNA inter-strand crosslinks that can be also caused by aldehydes [41, 42]. Finally, although there are very little data linking BER to the repair of DPCs, it is interesting to note that XRCC1 is involved in the removal and repair of lesions caused by trapped TOP1, and so is TDP1 [43]. If the TOP1 removal mechanism shares any similarity to DPC repair, this could potentially place BER downstream of SPRTN/DVC1 after the removal of the crosslinked protein. Altogether, this could mean that the BER pathway may be involved in the repair of aldehyde-related DPCs and this, in turn, could explain why low

ALDH2 expressing cancer cells have increased BER levels. Secondly, our data could be interpreted by using an alternative scenario, seeing loss of BER as a cause of aldehyde-related toxicity. We have previously shown that the loss of BER triggers metabolic changes, notably by increasing the activity of the one-carbon cycle [17]. More recently, it has been shown that endogenously produced formaldehyde plays a major part in the one-carbon cycle [37]. If BER-deficient cells increase the activity of this cycle, it is reasonable to suggest that this would, in turn, increase the concentration of formaldehyde in the cell as both a waste product and a reaction intermediate. Interestingly, we previously found that BER-deficient cells exhibit increased levels of the anti-oxidant GSH [20]. It is known that GSH can be used as an alternative to detoxify aldehydes [2]. Therefore, we propose that BER-deficient cells increase ALDH2 expression in order to prevent aldehyde-induced DNA damage. This would mean that low ALDH2 expressing cancer cells are at high risk of aldehyde overload if BER levels are not high enough. Thirdly, we could envisage that BER-deficiency, by creating numerous SSBs, some of which may contain aldehydes at the end of the strand break [43], can easily cause DPCs [26]. This hypothesis is reinforced by the fact that the two models used in this study, XRCC1 and TDP1 depletion, have in common a very substantial accumulation of SSBs. Finally, whatever the role of BER, knowing that any or all these hypotheses could be correct, we propose as proof of principle that cells may not survive by losing both BER and ALDH2 activities. By inhibiting Sp1, one of the main regulators of BER protein expression [25], we observed a higher sensitivity for BER loss in cells expressing low levels of ALDH2. Although Mithramycin A is quite toxic *in vivo* [44] and cannot be used to treat patients directly, the development of less toxic Mithramycin A derivatives or the use of other BER inhibitors [45-47] may open up a new avenues to target specific metabolic defects employed by some cancer cells.

In conclusion, we have discovered a previously unknown interconnection between the BER pathway and aldehyde detoxification. We demonstrated the importance of this relationship in the prognosis of liver and lung cancer and suggest that these observations may pave a new way towards targeting cancer metabolism.

Acknowledgements

This work was supported by grants from the Medical Research Council [H3RWGJ00.H302.1], Cancer Research UK [C5255/A15935] and the Federal Agency for Scientific Organizations (FASO Russia grant N° 0324-2016-0008) to G.L.D. XC was supported by the Jiangsu province grant [1402199C] and the Start-up fund from the Affiliated Hospital of Nantong University of Jiangsu Province, China. B.V and K.R were supported by the Medical Research Council [MC_EX_MR/K022830/1] to K.R.

Disclosure of conflict of interest:

None declared.

References

1. H.K. Seitz and F. Stickel, Molecular mechanisms of alcohol-mediated carcinogenesis. *Nat Rev Cancer*. **7**, 599-612 (2007) <https://doi.org/10.1038/nrc2191>
2. E.M. Tacconi, X. Lai, C. Folio, M. Porru, G. Zonderland, S. Badie, J. Michl, I. Sechi, M. Rogier, V. Matia Garcia, A.S. Batra, O.M. Rueda, P. Bouwman, J. Jonkers, A. Ryan, B. Reina-San-Martin, J. Hui, N. Tang, A. Bruna, A. Biroccio and M. Tarsounas, BRCA1 and BRCA2 tumor suppressors protect against endogenous acetaldehyde toxicity. *EMBO Mol Med*. **9**, 1398-1414 (2017) <https://doi.org/10.15252/emmm.201607446>
3. P.J. O'Brien, A.G. Siraki and N. Shangari, Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. *Crit Rev Toxicol*. **35**, 609-662 (2005)
4. G.P. Voulgaridou, I. Anastopoulos, R. Franco, M.I. Panayiotidis and A. Pappa, DNA damage induced by endogenous aldehydes: current state of knowledge. *Mutat Res*. **711**, 13-27 (2011) <https://doi.org/10.1016/j.mrfmmm.2011.03.006>
5. C.H. Chen, J.C. Ferreira, E.R. Gross and D. Mochly-Rosen, Targeting aldehyde dehydrogenase 2: new therapeutic opportunities. *Physiol Rev*. **94**, 1-34 (2014) <https://doi.org/10.1152/physrev.00017.2013>
6. K. Parmar and A.D. D'Andrea, Aldehydes Pose a Threat to BRCA2 Mutation Carriers. *Cell*. **169**, 979-981 (2017) <https://doi.org/10.1016/j.cell.2017.05.021>
7. J. Abraham, S. Balbo, D. Crabb and P.J. Brooks, Alcohol metabolism in human cells causes DNA damage and activates the Fanconi anemia-breast cancer susceptibility (FA-BRCA) DNA damage response network. *Alcohol Clin Exp Res*. **35**, 2113-2120 (2011) <https://doi.org/10.1111/j.1530-0277.2011.01563.x>
8. J.I. Garaycochea, G.P. Crossan, F. Langevin, M. Daly, M.J. Arends and K.J. Patel, Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature*. **489**, 571-575 (2012) <https://doi.org/10.1038/nature11368>
9. S.L.W. Tan, S. Chadha, Y. Liu, E. Gabasova, D. Perera, K. Ahmed, S. Constantinou, X. Renaudin, M. Lee, R. Aebersold and A.R. Venkitaraman, A Class of Environmental and Endogenous Toxins Induces BRCA2 Haploinsufficiency and Genome Instability. *Cell*. **169**, 1105-1118 e1115 (2017) <https://doi.org/10.1016/j.cell.2017.05.010>
10. H.S. Yu, T. Oyama, T. Isse, K. Kitagawa, T.T. Pham, M. Tanaka and T. Kawamoto, Formation of acetaldehyde-derived DNA adducts due to alcohol exposure. *Chem Biol Interact*. **188**, 367-375 (2010) <https://doi.org/10.1016/j.cbi.2010.08.005>
11. H. Nagayoshi, A. Matsumoto, R. Nishi, T. Kawamoto, M. Ichiba and T. Matsuda, Increased formation of gastric N(2)-ethylidene-2'-deoxyguanosine DNA adducts in aldehyde dehydrogenase-2 knockout mice treated with ethanol. *Mutat Res*. **673**, 74-77 (2009) <https://doi.org/10.1016/j.mrgentox.2008.11.009>
12. G.L. Dianov and U. Hubscher, Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res*. **41**, 3483-3490 (2013) <https://doi.org/10.1093/nar/gkt076>
13. G. Dianov and T. Lindahl, Reconstitution of the DNA base excision-repair pathway. *Curr Biol*. **4**, 1069-1076 (1994)
14. A. Klungland and T. Lindahl, Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J*. **16**, 3341-3348 (1997)
15. Dianova, II, K.M. Sleeth, S.L. Allinson, J.L. Parsons, C. Breslin, K.W. Caldecott and G.L. Dianov, XRCC1-DNA polymerase beta interaction is required for efficient base excision repair. *Nucleic Acids Res*. **32**, 2550-2555 (2004) <https://doi.org/10.1093/nar/gkh567>
16. K.W. Caldecott, XRCC1 and DNA strand break repair. *DNA Repair (Amst)*. **2**, 955-969 (2003)
17. E. Markkanen, R. Fischer, M. Ledentcova, B.M. Kessler and G.L. Dianov, Cells deficient in base-excision repair reveal cancer hallmarks originating from adjustments to genetic instability. *Nucleic Acids Res*. **43**, 3667-3679 (2015) <https://doi.org/10.1093/nar/gkv222>
18. L.F. Stead, S. Berri, H.M. Wood, P. Egan, C. Conway, C. Daly, K. Papagiannopoulos and P. Rabbitts, The transcriptional consequences of somatic amplifications, deletions, and rearrangements in a human lung squamous cell carcinoma. *Neoplasia*. **14**, 1075-1086 (2012)
19. G. Orlando, S.V. Khoronenkova, Dianova, II, J.L. Parsons and G.L. Dianov, ARF induction in response to DNA strand breaks is regulated by PARP1. *Nucleic Acids Res*. **42**, 2320-2329 (2014) <https://doi.org/10.1093/nar/gkt1185>
20. K. Kiianitsa and N. Maizels, A rapid and sensitive assay for DNA-protein covalent complexes in living cells. *Nucleic Acids Res*. **41**, e104 (2013) <https://doi.org/10.1093/nar/gkt171>
21. A.J. Link and J. LaBaer, Trichloroacetic acid (TCA) precipitation of proteins. *Cold Spring Harb Protoc*. **2011**, 993-994 (2011) <https://doi.org/10.1101/pdb.prot5651>

22. J. Gao, B.A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S.O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander and N. Schultz, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. **6**, p11 (2013) <https://doi.org/10.1126/scisignal.2004088>
23. E. Cerami, J. Gao, U. Dogrusoz, B.E. Gross, S.O. Sumer, B.A. Aksoy, A. Jacobsen, C.J. Byrne, M.L. Heuer, E. Larsson, Y. Antipin, B. Reva, A.P. Goldberg, C. Sander and N. Schultz, The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. **2**, 401-404 (2012) <https://doi.org/10.1158/2159-8290.CD-12-0095>
24. G. Antoniali, L. Lirussi, M. Poletto and G. Tell, Emerging roles of the nucleolus in regulating the DNA damage response: the noncanonical DNA repair enzyme APE1/Ref-1 as a paradigmatic example. *Antioxid Redox Signal*. **20**, 621-639 (2014) <https://doi.org/10.1089/ars.2013.5491>
25. M. Poletto, A.J. Legrand, S.C. Fletcher and G.L. Dianov, p53 coordinates base excision repair to prevent genomic instability. *Nucleic Acids Res*. **44**, 3165-3175 (2016) <https://doi.org/10.1093/nar/gkw015>
26. J.K. Horton, M. Watson, D.F. Stefanick, D.T. Shaughnessy, J.A. Taylor and S.H. Wilson, XRCC1 and DNA polymerase beta in cellular protection against cytotoxic DNA single-strand breaks. *Cell Res*. **18**, 48-63 (2008) <https://doi.org/10.1038/cr.2008.7>
27. B. Vaz, M. Popovic, J.A. Newman, J. Fielden, H. Aitkenhead, S. Halder, A.N. Singh, I. Vendrell, R. Fischer, I. Torrecilla, N. Drobnitzky, R. Freire, D.J. Amor, P.J. Lockhart, B.M. Kessler, G.W. McKenna, O. Gileadi and K. Ramadan, Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. *Mol Cell*. **64**, 704-719 (2016) <https://doi.org/10.1016/j.molcel.2016.09.032>
28. B. Vaz, M. Popovic and K. Ramadan, DNA-Protein Crosslink Proteolysis Repair. *Trends Biochem Sci*. **42**, 483-495 (2017) <https://doi.org/10.1016/j.tibs.2017.03.005>
29. M.J. Stewart, K. Malek and D.W. Crabb, Distribution of messenger RNAs for aldehyde dehydrogenase 1, aldehyde dehydrogenase 2, and aldehyde dehydrogenase 5 in human tissues. *J Investig Med*. **44**, 42-46 (1996)
30. S.C. Fletcher, C.P. Grou, A.J. Legrand, X. Chen, K. Soderstrom, M. Poletto and G.L. Dianov, Sp1 phosphorylation by ATM downregulates BER and promotes cell elimination in response to persistent DNA damage. *Nucleic Acids Res*. **46**, 1834-1846 (2018) <https://doi.org/10.1093/nar/gkx1291>
31. D. Hanahan and R.A. Weinberg, Hallmarks of cancer: the next generation. *Cell*. **144**, 646-674 (2011) <https://doi.org/10.1016/j.cell.2011.02.013>
32. M.G. Vander Heiden and R.J. DeBerardinis, Understanding the Intersections between Metabolism and Cancer Biology. *Cell*. **168**, 657-669 (2017) <https://doi.org/10.1016/j.cell.2016.12.039>
33. L.B. Sullivan, D.Y. Gui and M.G.V. Heiden, Altered metabolite levels in cancer: implications for tumour biology and cancer therapy. *Nat Rev Cancer*. **16**, 680-693 (2016) <https://doi.org/10.1038/nrc.2016.85>
34. A.J. Legrand, M. Poletto, D. Pankova, E. Clementi, J. Moore, F. Castro-Giner, A.J. Ryan, E. O'Neill, E. Markkanen and G.L. Dianov, Persistent DNA strand breaks induce a CAF-like phenotype in normal fibroblasts. *Oncotarget*. **9**, 13666-13681 (2018) <https://doi.org/10.18632/oncotarget.24446>
35. G. Burgos-Barragan, N. Wit, J. Meiser, F.A. Dingler, M. Pietzke, L. Mulderrig, L.B. Pontel, I.V. Rosado, T.F. Brewer, R.L. Cordell, P.S. Monks, C.J. Chang, A. Vazquez and K.J. Patel, Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. *Nature*. **548**, 549-554 (2017) <https://doi.org/10.1038/nature23481>
36. M.J. Nokin, F. Durieux, P. Peixoto, B. Chiavarina, O. Peulen, A. Blomme, A. Turtoi, B. Costanza, N. Smargiasso, D. Baiwir, J.L. Scheijen, C.G. Schalkwijk, J. Leenders, P. De Tullio, E. Bianchi, M. Thiry, K. Uchida, D.A. Spiegel, J.R. Cochrane, C.A. Hutton, E. De Pauw, P. Delvenne, D. Belpomme, V. Castronovo and A. Bellahcene, Methylglyoxal, a glycolysis side-product, induces Hsp90 glycation and YAP-mediated tumor growth and metastasis. *Elife*. **5**, (2016) <https://doi.org/10.7554/eLife.19375>
37. E.M. Tacconi, X. Lai, C. Folio, M. Porru, G. Zonderland, S. Badie, J. Michl, I. Sechi, M. Rogier, V. Matia Garcia, A.S. Batra, O.M. Rueda, P. Bouwman, J. Jonkers, A. Ryan, B. Reina-San-Martin, J. Hui, N. Tang, A. Bruna, A. Biroccio and M. Tarsounas, BRCA1 and BRCA2 tumor suppressors protect against endogenous acetaldehyde toxicity. *EMBO Mol Med*. (2017) <https://doi.org/10.15252/emmm.201607446>
38. D.E. Maessen, C.D. Stehouwer and C.G. Schalkwijk, The role of methylglyoxal and the glyoxalase system in diabetes and other age-related diseases. *Clin Sci (Lond)*. **128**, 839-861 (2015) <https://doi.org/10.1042/CS20140683>
39. D. Lessel, B. Vaz, S. Halder, P.J. Lockhart, I. Marinovic-Terzic, J. Lopez-Mosqueda, M. Philipp, J.C. Sim, K.R. Smith, J. Oehler, E. Cabrera, R. Freire, K. Pope, A. Nahid, F. Norris, R.J. Leventer, M.B. Delatycki, G. Barbi, S. von Ameln, J. Hogel, M. Degoricija, R. Fertig, M.D. Burkhalter, K. Hofmann, H. Thiele, J. Altmuller, G. Nurnberg, P. Nurnberg, M. Bahlo, G.M. Martin, C.M. Aalfs, J. Oshima, J. Terzic, D.J. Amor, I. Dikic, K. Ramadan and C. Kubisch, Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nat Genet*. **46**, 1239-1244 (2014) <https://doi.org/10.1038/ng.3103>

40. R.A. Sharma and G.L. Dianov, Targeting base excision repair to improve cancer therapies. *Mol Aspects Med.* **28**, 345-374 (2007) <https://doi.org/10.1016/j.mam.2007.06.002>
41. A. Kothandapani, A. Sawant, V.S. Dangeti, R.W. Sobol and S.M. Patrick, Epistatic role of base excision repair and mismatch repair pathways in mediating cisplatin cytotoxicity. *Nucleic Acids Res.* **41**, 7332-7343 (2013) <https://doi.org/10.1093/nar/gkt479>
42. A. Kothandapani and S.M. Patrick, Evidence for base excision repair processing of DNA interstrand crosslinks. *Mutat Res.* **743-744**, 44-52 (2013) <https://doi.org/10.1016/j.mrfmmm.2012.11.007>
43. B.B. Das, S.Y. Huang, J. Murai, I. Rehman, J.C. Ame, S. Sengupta, S.K. Das, P. Majumdar, H. Zhang, D. Biard, H.K. Majumder, V. Schreiber and Y. Pommier, PARP1-TDP1 coupling for the repair of topoisomerase I-induced DNA damage. *Nucleic Acids Res.* **42**, 4435-4449 (2014) <https://doi.org/10.1093/nar/gku088>
44. B.J. Kennedy, Metabolic and toxic effects of mithramycin during tumor therapy. *Am J Med.* **49**, 494-503 (1970)
45. L.E. Nunez, S.E. Nybo, J. Gonzalez-Sabin, M. Perez, N. Menendez, A.F. Brana, K.A. Shaaban, M. He, F. Moris, J.A. Salas, J. Rohr and C. Mendez, A novel mithramycin analogue with high antitumor activity and less toxicity generated by combinatorial biosynthesis. *J Med Chem.* **55**, 5813-5825 (2012) <https://doi.org/10.1021/jm300234t>
46. A. Malek, L.E. Nunez, M. Magistri, L. Brambilla, S. Jovic, G.M. Carbone, F. Moris and C.V. Catapano, Modulation of the activity of Sp transcription factors by mithramycin analogues as a new strategy for treatment of metastatic prostate cancer. *PLoS One.* **7**, e35130 (2012) <https://doi.org/10.1371/journal.pone.0035130>
47. M. Poletto, A.J. Legrand and G.L. Dianov, DNA Base Excision Repair: The Achilles' Heel of Tumour Cells and their Microenvironment? *Curr Pharm Des.* **23**, 4758-4772 (2017) <https://doi.org/10.2174/1381612823666170710123602>

Figure Legends

Figure 1. XRCC1 depletion causes an increased expression of ALDH2 and more DNA-protein crosslinks

(A) qRT-PCR analysis of ALDH2 transcript levels in NBE1 cells after 72h of XRCC1 knock down (KD). (B) Representative Western blot analysis of ALDH2 and XRCC1 levels in NBE1 cells after 72h of XRCC1 KD. (C) Representative immunofluorescence analysis of ALDH2 in NBE1 cells after 72h of XRCC1 KD. (D) qRT-PCR analysis of ALDH2 and XRCC1 transcript levels in TIG1 cells after 72h of XRCC1 KD. (E) qRT-PCR analysis of ALDH2 and XRCC1 transcript levels in WI38 cells after 72h of XRCC1 KD. (F) qRT-PCR analysis of ALDH2 and TDP1 transcript levels in TIG1 cells after 72h of TDP1 KD. qRT-PCR reference genes are B2M and GAPDH for A, D and E, and TBP and GAPDH for F. (G) Left panel, representative silver staining analysis of protein crosslinked onto DNA in NBE1 cell extracts after 72h of XRCC1 KD or 24h of disulfiram treatment (10 μ M). Right panel, densitometric quantification of the data. Data are expressed as mean \pm SD from at least three independent experiments * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Figure 2. XRCC1 and ALDH2 mRNA levels mirror each other in most cancer types

(A) Boxplot summarising the fold changes in mRNA levels for ALDH2 (blue) and XRCC1 (red) depending on the cancer type. (B) Boxplot representing the overall variation in XRCC1 and ALDH2 levels in all cancer types. A list of the studies used and the associated mRNA fold changes can be found in Table S1.

Figure 3. Low ALDH2 expression predicts a poor prognosis in lung cancer

(A) Boxplot representing the overall variation in XRCC1 and ALDH2 expression levels in lung cancers (Oncomine). (B) Kaplan-Meier analysis of the 5-year overall survival of lung cancer patients (provisional TCGA) stratified for XRCC1 mRNA levels (cut at the median). (C) Kaplan-Meier analysis of the 5-year overall survival of lung cancer patients (provisional TCGA) stratified for ALDH2 mRNA levels (cut at the median).

Figure 4. Low ALDH2 expression predicts a poor prognosis in liver cancer

(A) Boxplot representing the overall variation in XRCC1 and ALDH2 levels in liver cancers (Oncomine). (B) Kaplan-Meier analysis of the 5-year overall survival of liver cancer patients (provisional TCGA) stratified for XRCC1 mRNA levels (cut at the median). (C) Kaplan-Meier analysis of the 5-year overall survival of liver cancer patients (provisional TCGA) stratified for ALDH2 mRNA levels (cut at the median).

Figure 5. ALDH2 and XRCC1 expression have no predictive value in oesophagus cancer

(A) Boxplot representing the overall variation in XRCC1 and ALDH2 levels in oesophageal cancers (Oncomine). (B) Kaplan-Meier analysis of the 5-year overall survival of oesophageal cancer patients (provisional TCGA) stratified for XRCC1 mRNA levels (cut at the median). (C) Kaplan-Meier analysis of the 5-year overall survival of oesophagus cancer patients (provisional TCGA) stratified for ALDH2 mRNA levels (cut at the median).

Figure 6. XRCC1 levels increase the predictive value of ALDH2 expression in liver cancer

(A) Kaplan-Meier analysis of the 5-year overall survival of liver cancer patients (provisional TCGA) pre-stratified for low XRCC1 mRNA levels and stratified for ALDH2 expression (cut at the median). (B) Kaplan-Meier analysis

of the 5-year overall survival of liver cancer patients (provisional TCGA) pre-stratified for high XRCC1 mRNA levels and stratified for ALDH2 expression (cut at the median). (C) Boxplot representing the overall variation in XRCC1 mRNA levels in liver cancers (provisional TCGA) depending on ALDH2 expression group (cut at the median). (D) Boxplot representing the overall variation in ALDH2 mRNA levels in liver cancers (provisional TCGA) depending on XRCC1 expression group (cut at the median).

Figure 7. XRCC1/ALDH2 mRNA level ratio has an increased predictive value compared to that of ALDH2 expression alone in lung and liver cancer

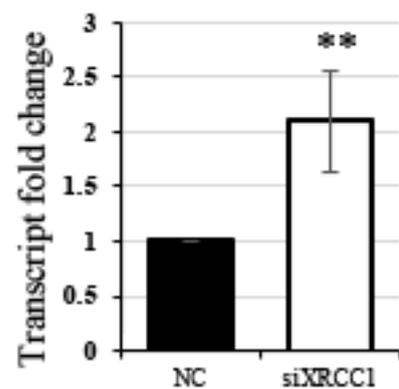
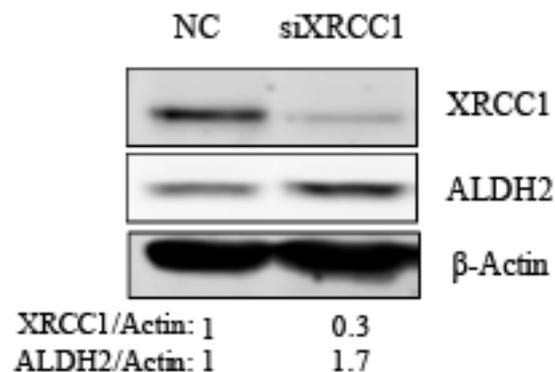
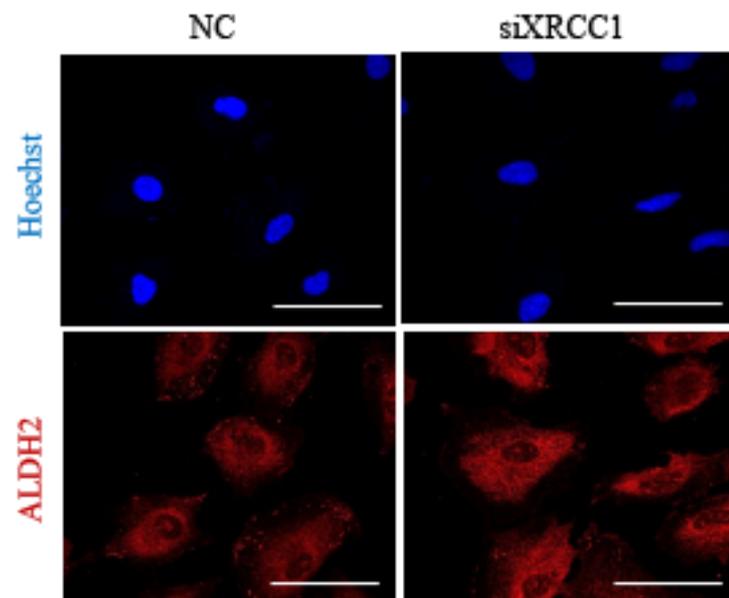
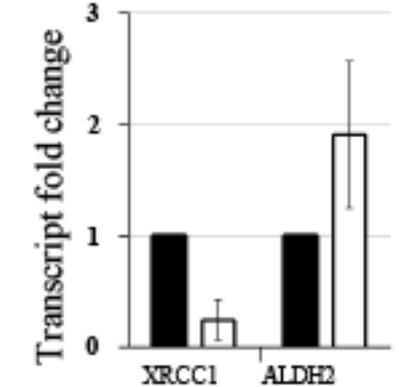
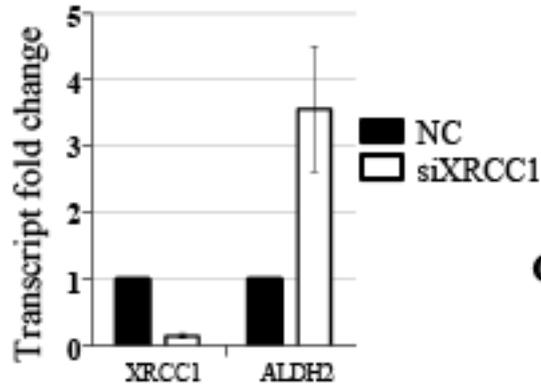
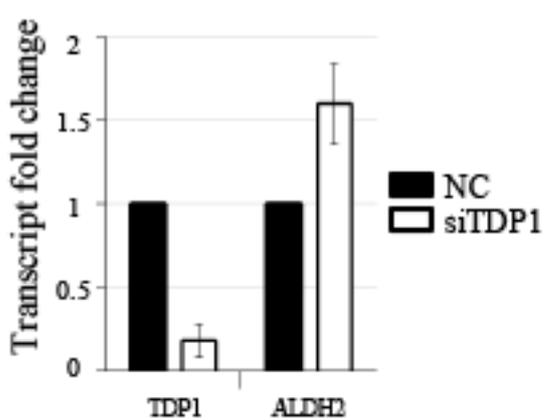
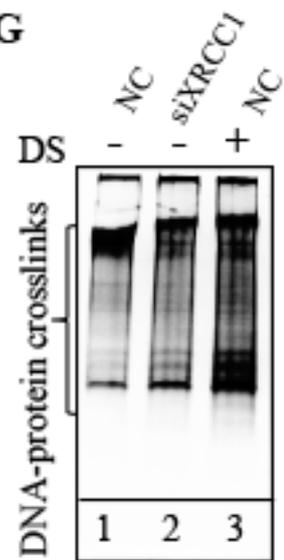
(A) Kaplan-Meier analysis of the 5-year overall survival of lung cancer patients (provisional TCGA) stratified for a XRCC1/ALDH2 mRNA levels ratio (cut at the median). (B) Kaplan-Meier analysis of the 5-year overall survival of liver cancer patients (provisional TCGA) stratified for a XRCC1/ALDH2 mRNA levels ratio (cut at the median). (C) Kaplan-Meier analysis of the 5-year overall survival of oesophageal cancer patients (provisional TCGA) stratified for a XRCC1/ALDH2 mRNA levels ratio (cut at the median).

Figure 8. Co-depletion of XRCC1 and ALDH2 sensitises low ALDH2-expressing cancer cells

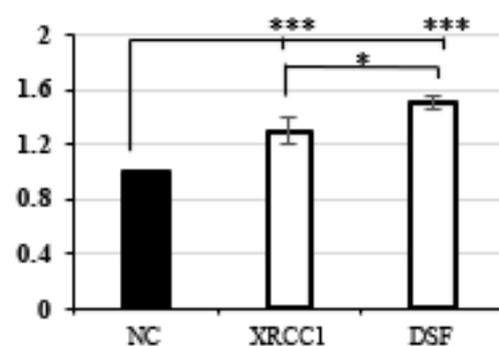
(A) Representative Western blot analysis of ALDH2 and XRCC1 levels in H1299 and JHH4 cells. Densitometric quantifications of the data are presented below. (B) Representative Western blot analysis of ALDH2 and XRCC1 levels in H1299 cells after 72h of XRCC1 and/or ALDH2 knock down (KD). (C) Representative Western blot analysis of ALDH2 and XRCC1 levels in JHH4 cells after 72h of XRCC1 and/or ALDH2 KD. (D) Clonogenic analysis performed in H1299 cells after 72h of XRCC1 and/or ALDH2 KD. (E) Clonogenic analysis performed in JHH4 cells after 72h of XRCC1 and/or ALDH2 KD. Data are expressed as mean \pm SD from at least three independent experiments * $p < 0.05$; N.S.: non-significant.

Figure 9. Sp1 inhibitor Mithramycin A kills low ALDH2 expressing cancer cells

(A) Clonogenic analysis performed in H1299 and JHH4 cells after two subsequent 24 h treatments (total 48 h) with Mithramycin A. (B) qRT-PCR analysis of XRCC1 and APE1 mRNA levels in JHH4 cells after treatment with 50 nM or 500 nM Mithramycin A for 24 h. (C) qRT-PCR analysis of XRCC1 and APE1 mRNA levels in H1299 cells after treatment with 50 nM or 500 nM Mithramycin A for 24 h. qRT-PCR reference genes are B2M and GAPDH. Data are expressed as mean \pm SD from at least three independent experiments. *** $p < 0.001$.

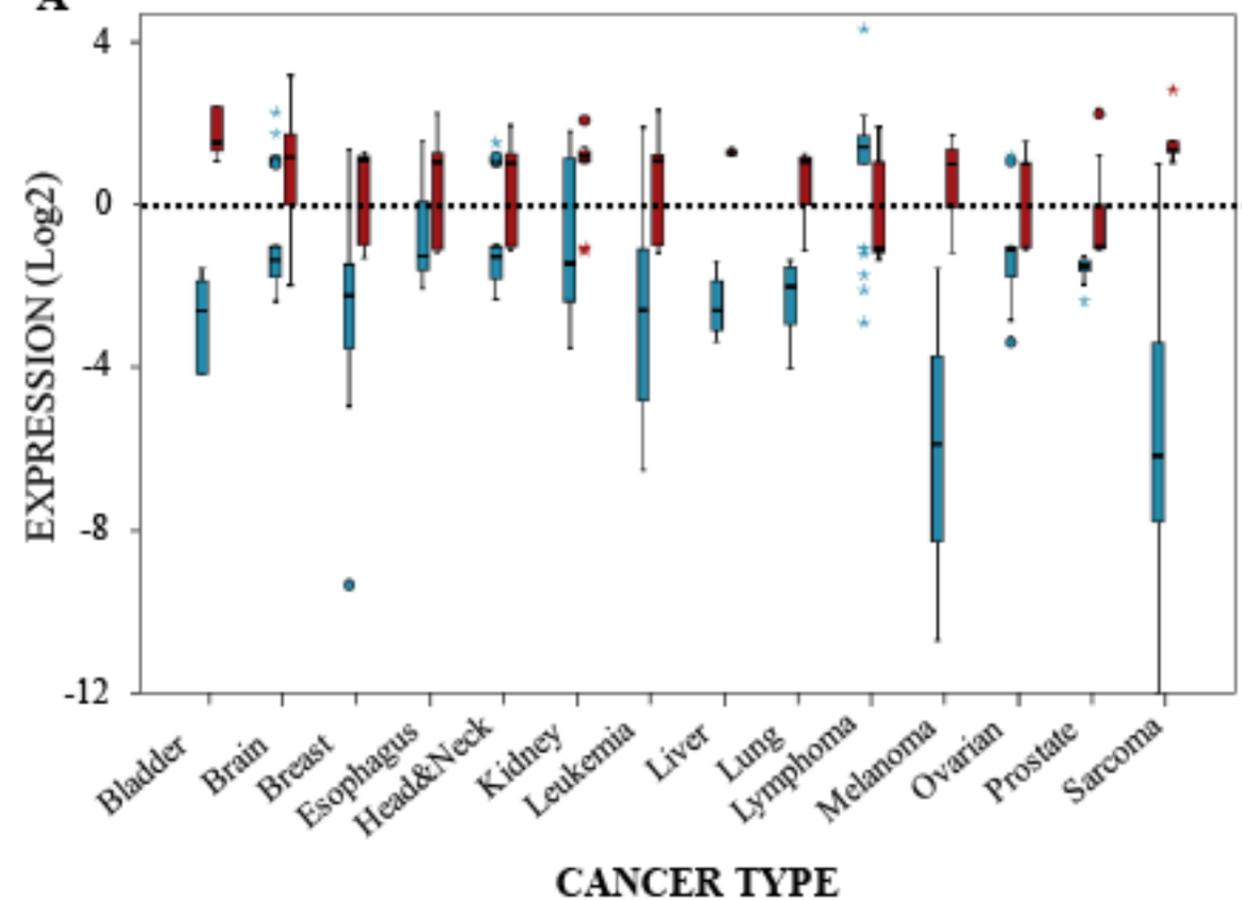
A NBE1 - ALDH2**B** NBE1**C****D** TIG1**E** Wi38**F** TIG1**G**

Normalised DPC Levels

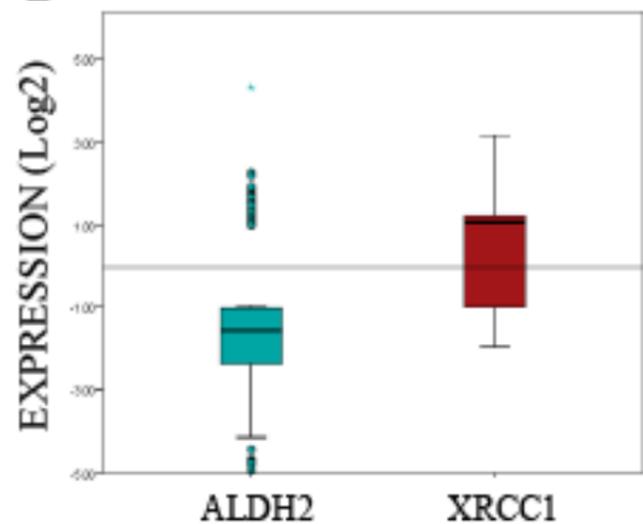


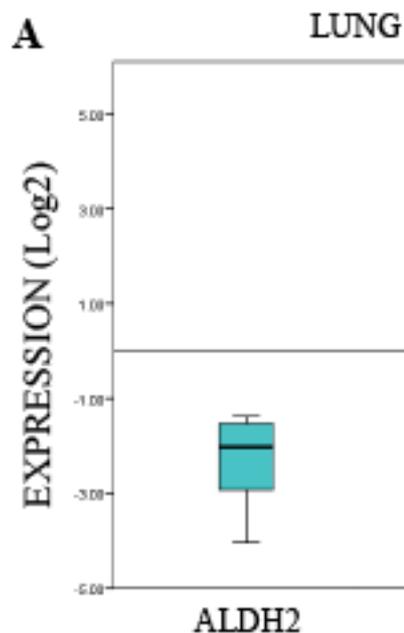
GENE: ■ ALDH2 ■ XRCC1

A

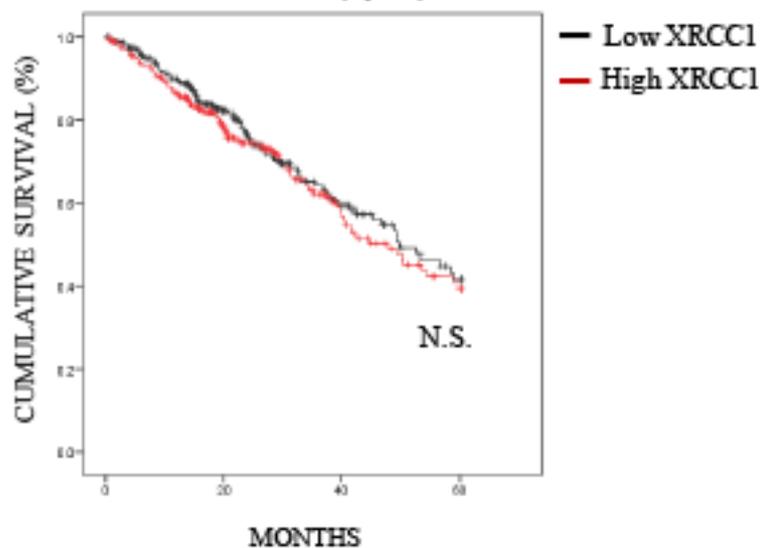


B

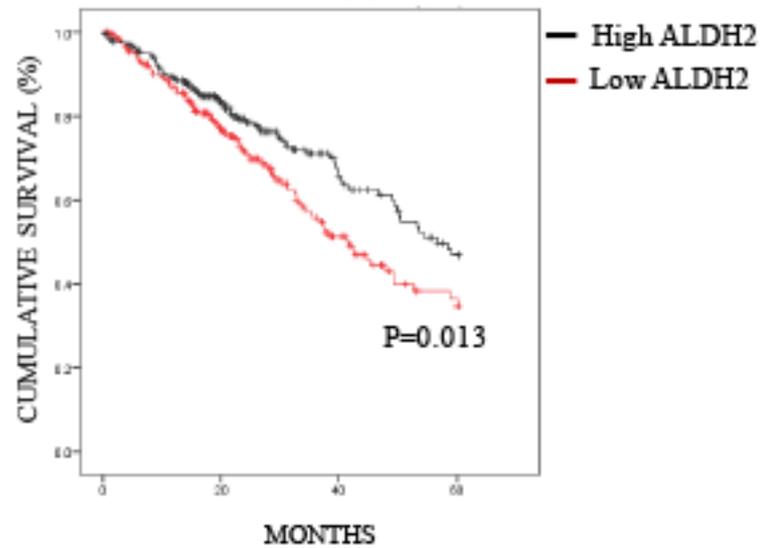


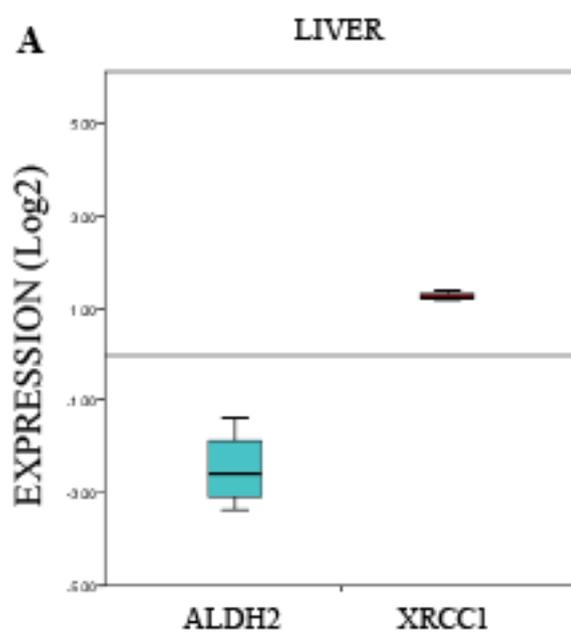


B OVERALL SURVIVAL (5 years) - LUNG

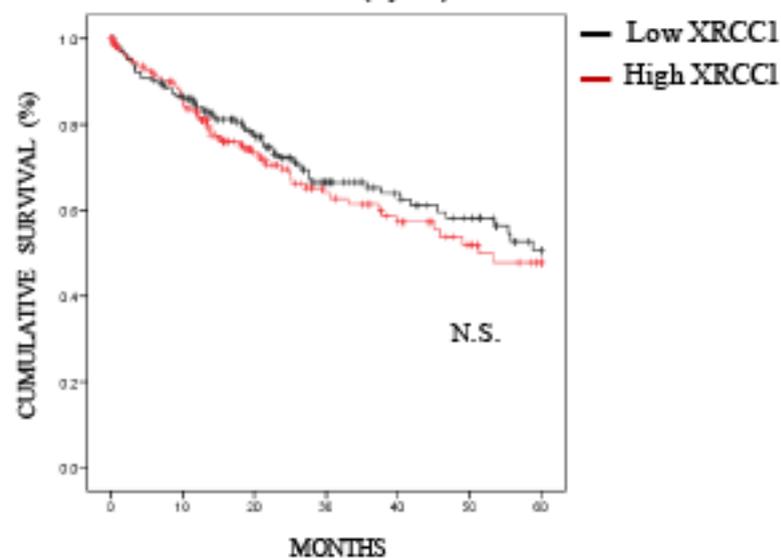


C OVERALL SURVIVAL (5 years) - LUNG

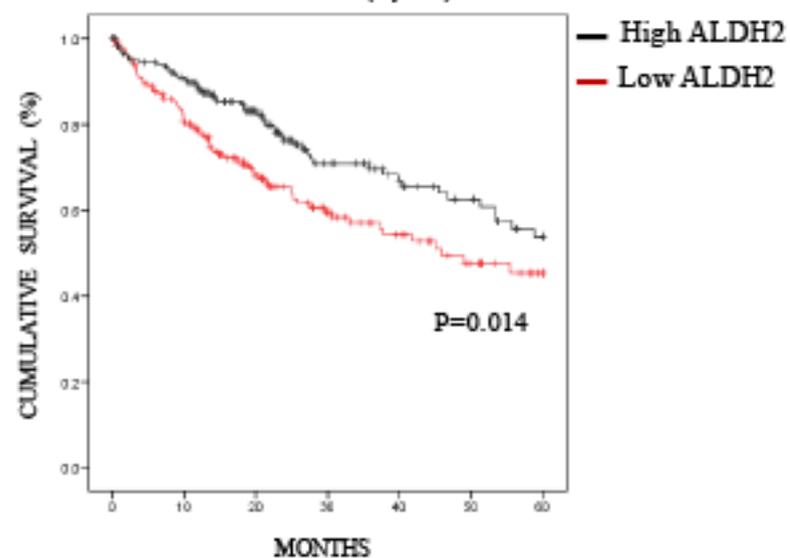


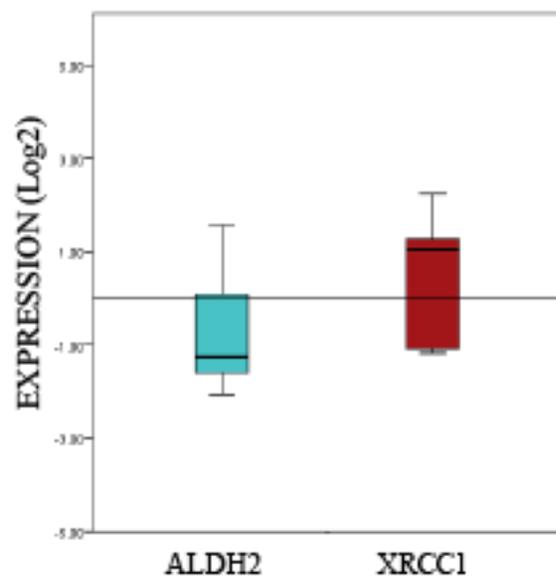
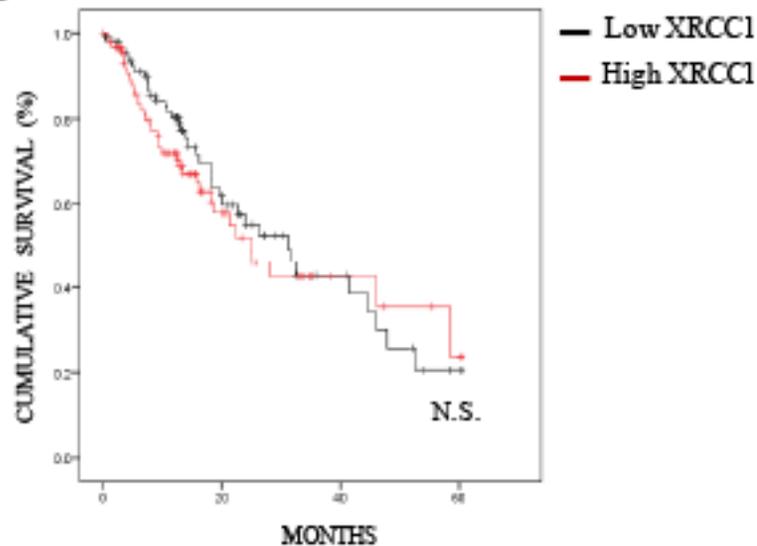
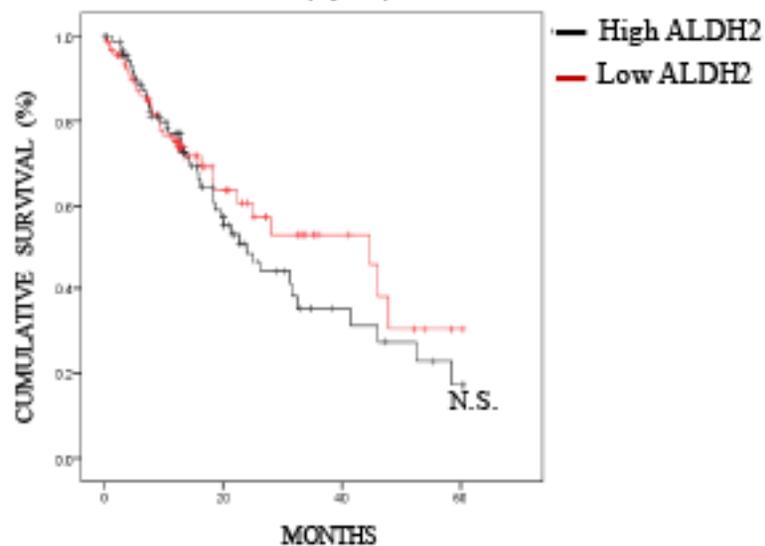


B OVERALL SURVIVAL (5 years) - LIVER

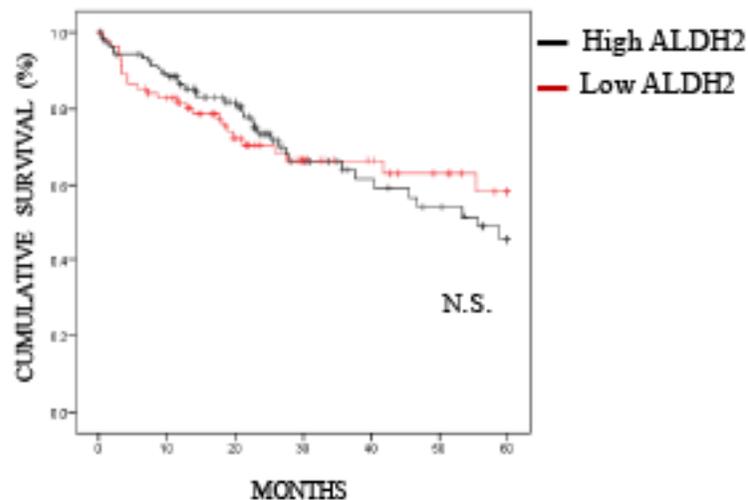


C OVERALL SURVIVAL (5 years) - LIVER

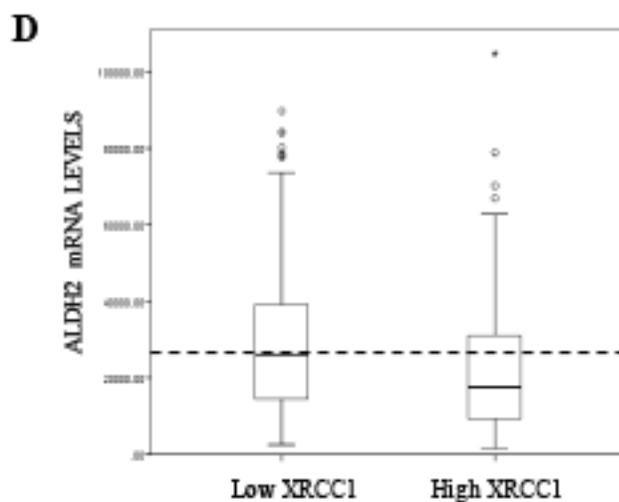
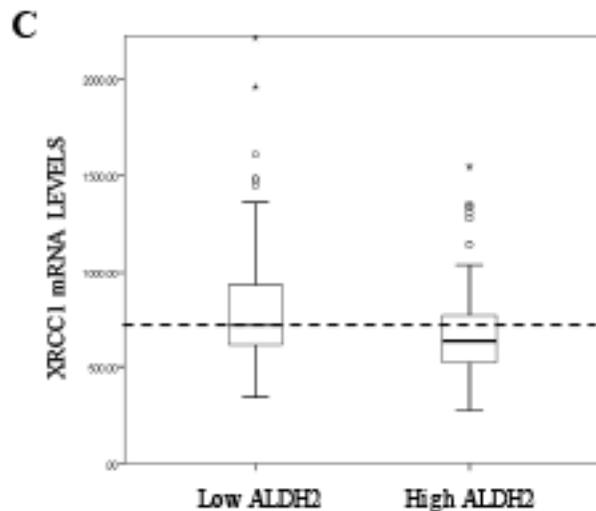
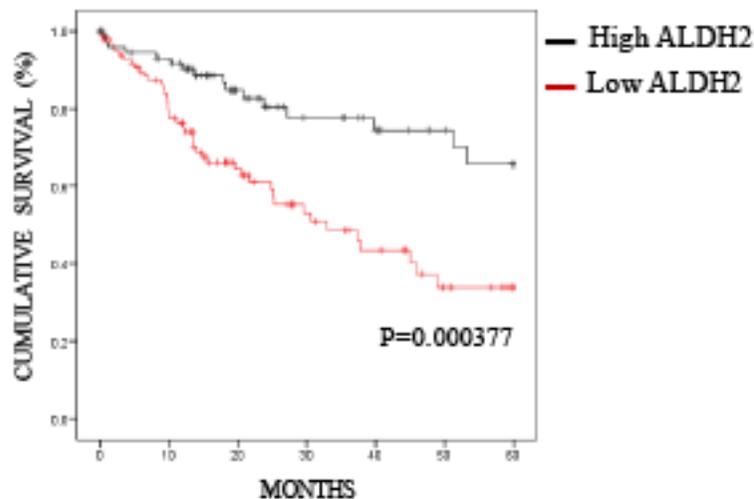


A ESOPHAGUS**B** OVERALL SURVIVAL (5 years) - ESOPHAGUS**C** OVERALL SURVIVAL (5 years) - ESOPHAGUS

A OVERALL SURVIVAL (5 years) – LIVER
LOW XRCC1 LEVELS

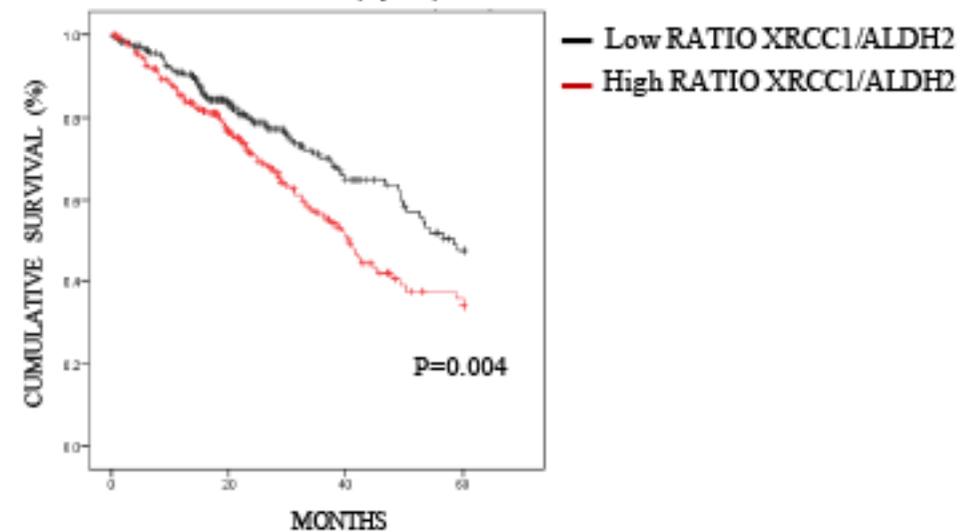


B OVERALL SURVIVAL (5 years) – LIVER
HIGH XRCC1 LEVELS

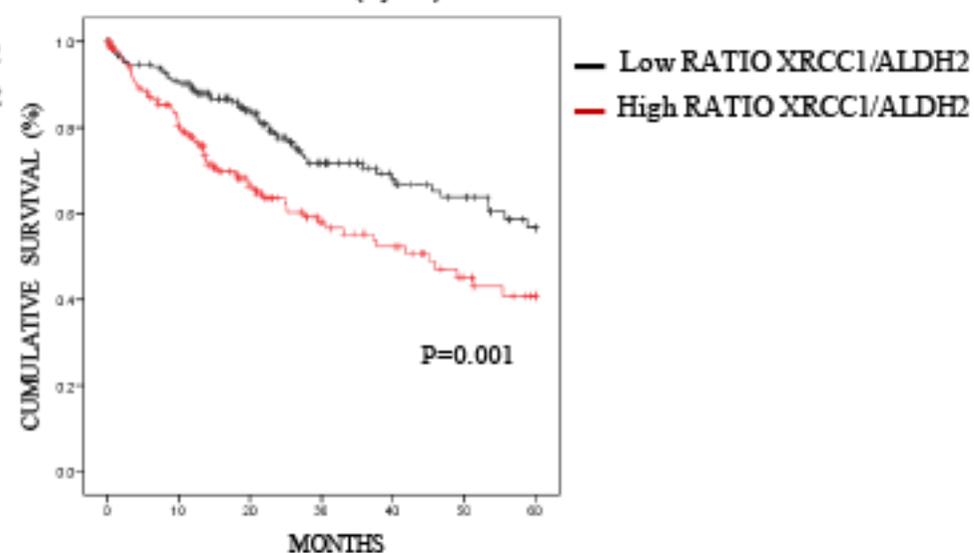


A

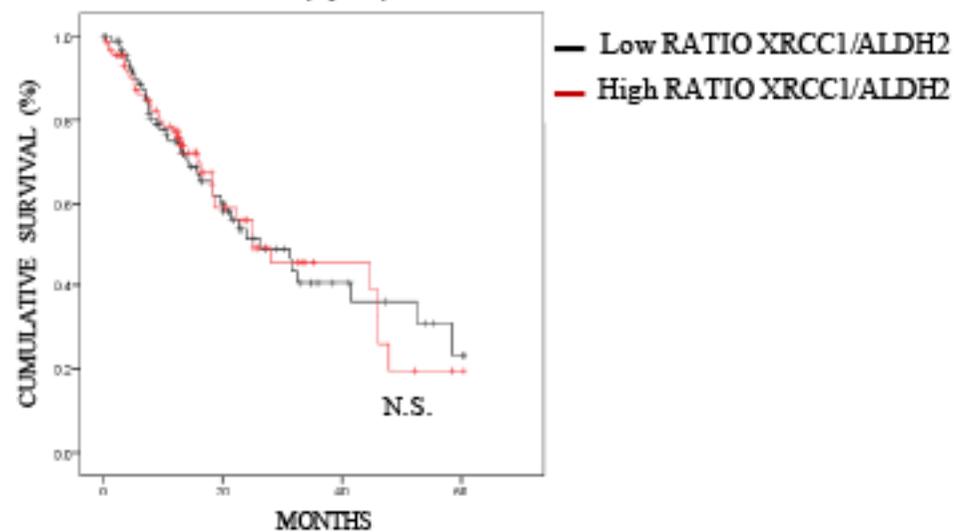
OVERALL SURVIVAL (5 years) - LUNG

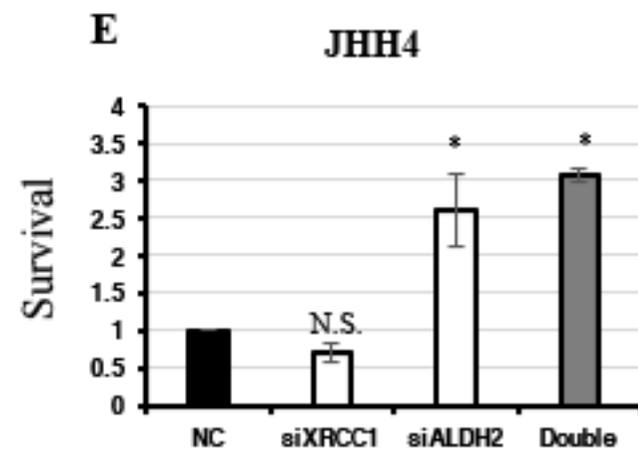
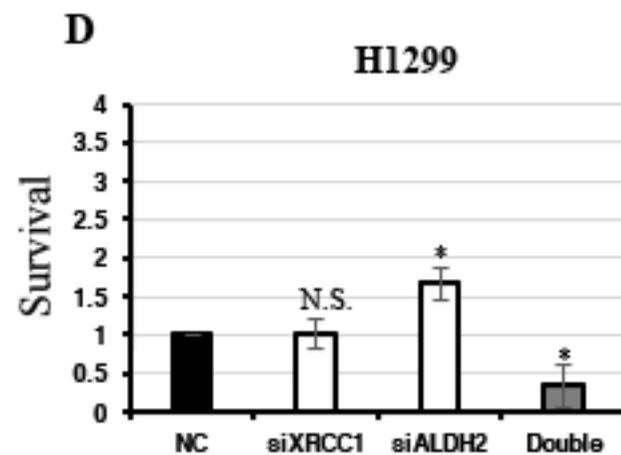
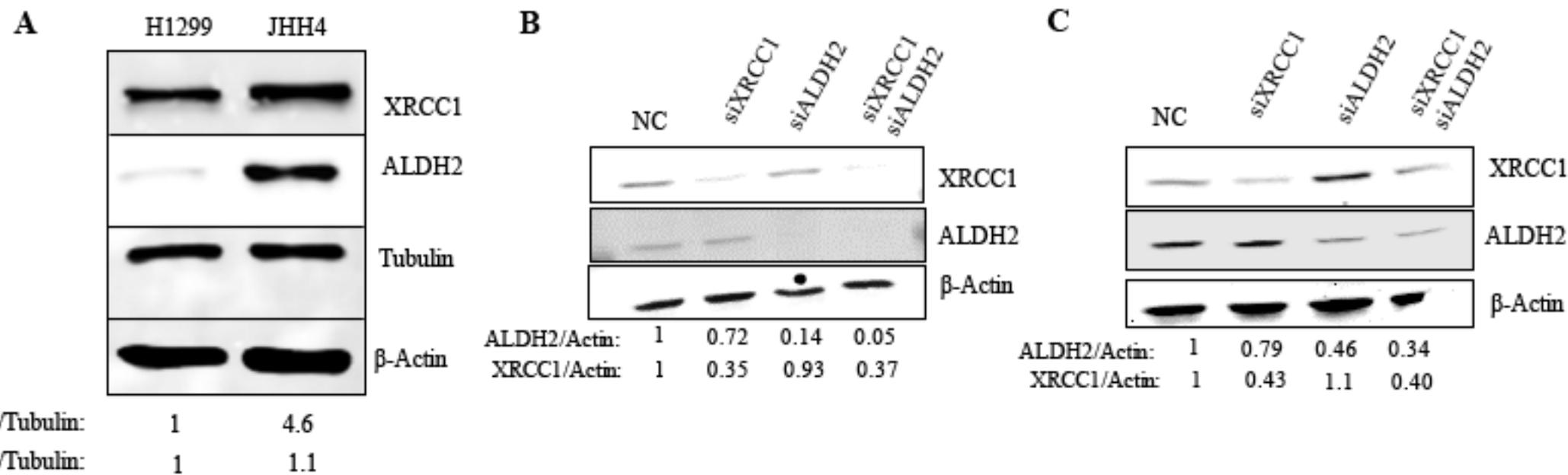
**B**

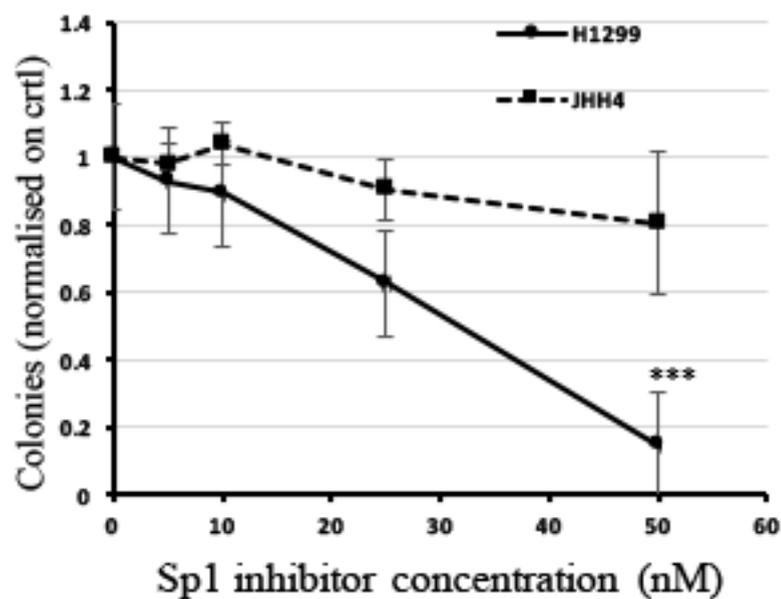
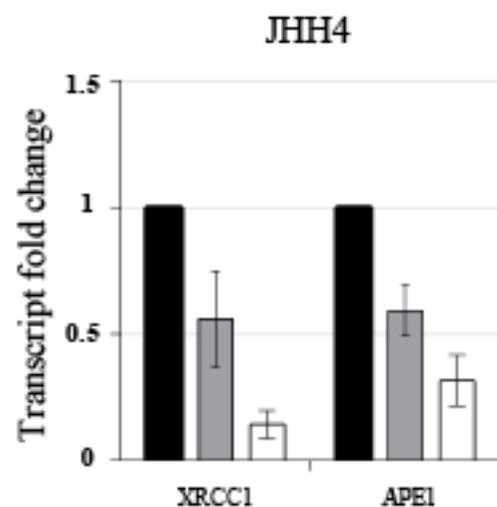
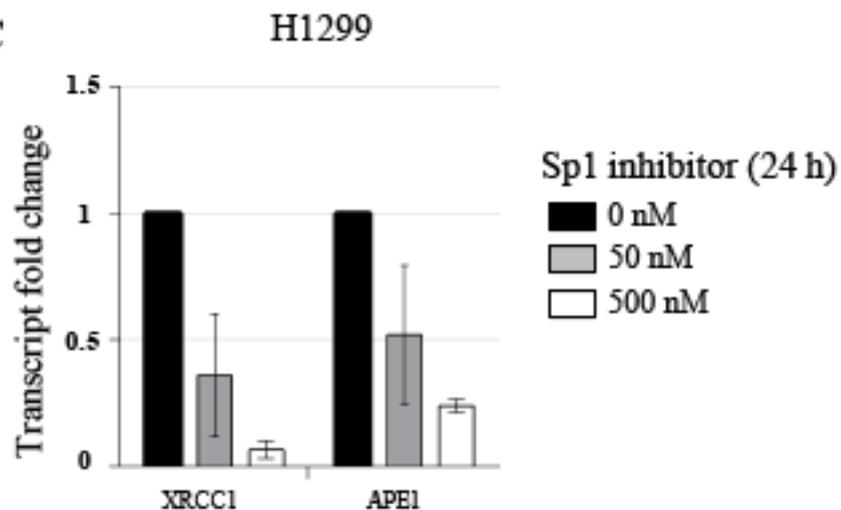
OVERALL SURVIVAL (5 years) - LIVER

**C**

OVERALL SURVIVAL (5 years) - ESOPHAGUS





A**B****C**

Sp1 inhibitor (24 h)

0 nM

50 nM

500 nM

Figure S1