Title: DNA damage repair and survival outcomes in advanced gastric cancer patients treated with first-line chemotherapy.

Authors: Livia Ronchetti [1,^], Elisa Melucci [1,^], Francesca De Nicola [2,^], Frauke Goeman [3,^], Beatrice Casini [1], Francesca Sperati [4] Matteo Pallocca [2], Irene Terrenato [4], Laura Pizzuti [5], Patrizia Vici [5], Domenico Sergi [5], Luigi Di Lauro [5], Carla Azzurra Amoreo [1], Enzo Gallo [1], Maria Grazia Diodoro [1], Edoardo Pescarmona [1], Ilio Vitale [6,7], Maddalena Barba [5,7,], Simonetta Buglioni [1], Marcella Mottolese [1], Maurizio Fanciulli [2], Ruggero De Maria [8,α,*], and Marcello Maugeri-Saccà [5,7,α,*]

^ These authors contributed equally to this work.

[α] These authors share senior authorship

[1] Department of Pathology, "Regina Elena" National Cancer Institute, Via Elio Chianesi 53, 00144, Rome, Italy

[2] SAFU Laboratory, Department of Research, Advanced Diagnostic, and Technological Innovation, "Regina Elena" National Cancer Institute, Via Elio Chianesi 53, 00144, Rome, Italy

[3] Oncogenomic and Epigenetic Unit, "Regina Elena" National Cancer Institute, Via Elio Chianesi 53, 00144, Rome, Italy

[4] Biostatistics-Scientific Direction, "Regina Elena" National Cancer Institute, Via Elio Chianesi 53, 00144, Rome, Italy

[5] Division of Medical Oncology 2, "Regina Elena" National Cancer Institute, Via Elio Chianesi 53, 00144, Rome, Italy

[6] Department of Biology, University of Rome "Tor Vergata", Rome, Italy.

[7] Scientific Direction, "Regina Elena" National Cancer Institute, Via Elio Chianesi 53, 00144, Rome, Italy

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/ijc.30668

[8] Institute of General Pathology, Catholic University of the Sacred Heart, Largo Agostino Gemelli, 10, 00168, Rome, Italy

*Corresponding author: Dr. Marcello Maugeri-Saccà, Division of Medical Oncology 2 and Scientific Direction, "Regina Elena" National Cancer Institute, Via Elio Chianesi 53, 00144 Rome, Italy. Tel. +39-06-52662724, E-mail: marcello.maugerisacca@ifo.gov.it. Prof. Ruggero De Maria, Institute of General Pathology, Catholic University of the Sacred Heart, Largo Agostino Gemelli, 10, 00168, Rome, Italy. Tel. +39-06-30154914, E-mail: demariaruggero@gmail.com

Running Title: DNA damage repair in gastric cancer

Keywords: DNA damage repair, γ-H2AX, pATM, TP53, ARID1A.

Abbreviations ATM: Ataxia-Telangiectasia Mutated; ATR: Ataxia Telangiectasia and Rad3-related protein; Chk1: Checkpoint kinase 1; Chk2: Checkpoint Kinase 2; DDR: DNA damage and repair; DNA DSBs: DNA double-strand breaks; DNA SSBs: DNA single- strand breaks; OS: overall survival; γ-H2AX: phosphorylated H2A Histone Family Member X; PFS: progression-free survival; pRPA32: phosphorylated replication protein A2; Wee1: Wee1-like protein kinase.

Word count: 2.769

Total number of figures and tables: 6

Novelty and Impact

DNA damage response (DDR) activation is common in gastric cancer (GC) and possibly fuelled by mutations affecting DNA damage repair ability (e.g. *TP53* and *ARID1A*). The subset of GC patients carrying activation of the DDR had adverse survival outcomes compared with their negative counterparts. *TP53* mutations did not modify the relationship between DDR biomarkers and inferior clinical outcomes, whereas *ARID1A* mutations did. Overall, this study identifies molecular factors associated with the efficacy of chemotherapy.

Abstract

The DNA damage response (DDR) network is exploited by cancer cells to withstand chemotherapy. Gastric cancer (GC) carries deregulation of the DDR and harbors genetic defects that fuel its activation. The ATM-Chk2 and ATR-Chk1-Wee1 axes are deputed to initiate DNA repair. Over-activation of these pathways in cancer cells may represent an adaptive response for compensating genetic defects deregulating G_1-S transition (e.g. *TP53*) and ATM/ATR-initiated DNA repair (e.g. *ARID1A*). We hypothesized that DDR-linked biomarkers may predict clinical outcomes in GC patients treated with chemotherapy. Immunohistochemical assessment of DDR kinases (pATM, pChk2, pChk1 and pWee1) and DNA damage markers (γ-H2AX and pRPA32) was performed in biological samples from 110 advanced GC patients treated with first-line chemotherapy, either in phase II trials or in routine clinical practice. In 90 patients this characterization was integrated with targeted ultra-deep sequencing for evaluating the mutational status of *TP53* and *ARID1A*. We recorded a positive association between the investigated biomarkers. The combination of two biomarkers (γ-H2AX^{high}/pATM^{high}) was an adverse factor for both progression-free survival (multivariate Cox: HR 2.23, 95%CI: 1.47-3.40) and overall survival (multivariate Cox: HR: 2.07, 95%CI: 1.20-3.58). The relationship between the γ-H2AX^{high}/pATM^{high} model and progression-free survival was consistent across the different *TP53* backgrounds and was maintained in the *ARID1A* wild-type setting. Conversely, this association was no longer observed in an *ARID1A*-mutated subgroup. The γ-H2AX^{high}/pATM^{high} model negatively impacted survival outcomes in GC patients treated with chemotherapy. The mutational status of *ARID1A*, but apparently not *TP53* mutations, affects its predictive significance.

INTRODUCTION

DNA damage and repair (DDR) pathways protect eukaryotic cells from genotoxic injuries ensuring the transmission of an unaltered genetic code to the progeny.[1] This complex genome-protecting network is aberrantly exploited by cancer cells to tolerate the high levels of DNA damage they accumulate. Indeed, the genome of cancer cells is threatened by a variety of endogenous substances and deregulated processes, including: i) the elevated production of reactive oxygen species related to the increased metabolic demands (oxidative DNA damage),[2] ii) the replicative stress imposed by activating mutations in oncogenes that control cell proliferation,[3] and iii) the defective nature of the G_1/S checkpoint (e.g. *TP53* mutations) that renders cancer cells "addicted" to intact cell cycle control systems for determining cell fate upon DNA damage.[4]

The crosstalking Ataxia-Telangiectasia Mutated (ATM)-Checkpoint Kinase 2 (Chk2) and ataxia telangiectasia and Rad3-related protein (ATR)-Checkpoint kinase 1 (Chk1)-Wee1-like protein kinase (Wee1) axes operate in the intra-S and G_2/M checkpoints.[5,6] When DNA single- and double-strand breaks (SSBs and DSBs) occur, ATR-Chk1-Wee1 and ATM-Chk2 pathways halt the cell cycle and coordinate DNA repair or self-elimination of irreversibly damaged cells, depending upon the entity of the damage and repair capabilities.[5] In cancer cells DNA damage often arises in a background of defective G_1/S transition; within this molecular frame activation of ATM-Chk2 and ATR-Chk1-Wee1 pathways becomes crucial to avoid entry into a defective and lethal mitosis. We hypothesized that this adaptive response may represent a sort of "molecular side effect" through which cancer cells are predisposed to efficiently handle exogenous (therapeutic) sources of DNA damage, such as chemotherapy.[7,8]

Molecular characterization of gastric cancer (GC) revealed ATM/ATR activation across all the molecular subtypes.[9] Moreover, a number of genes associated with their signaling are recurrently altered in GC, including genes linked to defective cell cycle control (e.g. *TP53*) and altered initiation of the DDR cascade (e.g. *ARID1A*).[9] We reasoned that aberrant activation of the DDR in

GC, supposedly driven or enforced by the underlying genetic portrait of the disease, may be exploited in the search for biomarkers predicting chemotherapy sensitivity/resistance. To test this hypothesis, biological samples from 110 GC patients treated with first-line chemotherapy in prospective phase II trials or in routine clinical practice,[10-13] were retrospectively evaluated by immunohistochemistry (IHC) for assessing the expression of DDR kinases (pATM, pChk2, pChk1 and pWee1) and DNA damage markers, namely the DNA DSB marker phosphorylated H2A Histone Family Member X (γ-H2AX) and the single-stranded DNA/SSB marker phosphorylated replication protein A2 (pRPA2, best known as pRPA32). Mutation analysis of *TP53* and *ARID1A* was conducted in 90 patients to investigate whether specific genetic events with elevated mutational rates impacted the predictive significance of most promising markers.

PATIENTS AND METHODS

Patients

Written informed consents were obtained by all the participants. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the "Regina Elena" National Cancer Institute of Rome. This study adheres to the REMARK guidelines.[14] For this analysis, 110 patients with histologically confirmed, inoperable locally advanced or metastatic cancer of the stomach or gastroesophageal junction who received first-line chemotherapy (December 2000-January 2015) were included. Median follow-up was 11 months (IQR 5-20 months). Patients were considered eligible if complete data on clinical features and treatment outcomes were available and the amount of biological materials from biopsies or surgical samples was sufficient for testing the entire panel of antibodies. Following trastuzumab approval for the treatment of GC patients,[15] HER2 overexpression/amplification was carried out in 20 patients. Chemotherapy regimens and schedules are detailed in Suppl. Table 1. Tumor responses were evaluated by Response Evaluation Criteria in Solid Tumors (RECIST) criteria v.1.1. Progression-free survival (PFS) was calculated as the time between the first cycle of chemotherapy until

radiological evidence of disease progression or death due to any cause. Overall survival (OS) was computed as the time from the first cycle of chemotherapy to death due to any cause.

Immunohistochemistry

The immunohistochemical assessment of γ-H2AX, pATM, pChk2, pRPA32, pChk1 and pWee1was performed in formalin-fixed paraffin-embedded (FFPE) tissues with the following antibodies: anti-phospho-H2AX (Ser139) (clone JBW301) mouse monoclonal antibody (MAb) (Upstate) at the dilution of 1:500 (pH8), anti-phospho-ATM (Ser1981) (clone 7C10D8) mouse MAb (Rockland) at the dilution of 1:200 (pH6), anti-phospho-Chk2 (Thr68) (clone C13C1) rabbit MAb (Cell Signaling) at the dilution of 1:200 (pH6), anti-phospho-Chk1 (Ser345) (clone 133D3) rabbit MAb (Cell Signaling) at the dilution of 1:150 (pH6), anti-phospho-Wee1 (Ser642) (clone D47G5) rabbit MAb (Cell Signaling) at the dilution of 1:100 (pH8), anti-phospho-RPA32 (Ser4/Ser8) rabbit polyclonal antibody (Bethyl) at the dilution of 1:100 (pH6). Immunoreactions were revealed by a streptavidin-biotin enhanced immunoperoxidase technique (Super Sensitive MultiLink, Leica, Milan, Italy) in an automated autostainer (Bond III, Leica).

To define positive and negative cases a classification comparable to that of our previous studies was used.[7,8] DNA damage markers (γ-H2AX and pRPA32) were classified as high and low/negative using the median score of all tumors as the cut-off points. Median percentages of nuclear-expressing cells for γ-H2AX and pRPA32 were 30% and 40%, respectively. For DDR kinases, samples were considered positive when ≥20% of the neoplastic cells showed a distinct nuclear immunoreactivity of any intensity. Immunoreactivity was evaluated by two investigators blinded to treatment outcomes (LR and EM) and discordant cases were reviewed by a third observer (MM and SB). Immunohistochemical staining of two representative cases is presented in Suppl. Figure 1.

Targeted DNA Deep Sequencing

Genomic DNA was extracted from 5µm FFPE tissue sections using the AllPrep DNA/RNA FFPE kit (Qiagen, Valencia, CA, USA). To perform the targeted DNA resequencing of *ARID1A* and *TP53*, a custom panel employing DesignStudio from Illumina was designed. The library was prepared using the TruSeq Custom Amplicon Low Input Kit following the manufacturer's instructions. Sequencing was carried out on a NextSeq 500 (Illumina, Inc., San Diego, CA, USA) in paired-end mode, sequencing from each side 150 bp. Primary analysis encompassing FASTQ file generation, alignment and variant calling was performed on the Illumina BaseSpace Cloud environment, using the Truseq Amplicon analysis pipeline version 2.0. TSV files were generated from VCFs with the Illumina Variant Studio software. From TSV files, we filtered out low-coverage (<200x), dbSNP annotated variants (MAF >5%) and mutations that did not lead to an amino acid change of the protein. Afterwards, *TP53* and *ARID1A* mutations were further filtered on the basis of the following criteria: i) allele frequencies ≥10%, ii) having a COSMIC ID, iii) annotated as oncogenic in cBioPortal Version 1.2.5, last accessed on September 21st. The experimental workflow is summarized in Suppl. Figure 2.

Statistical Analyses

The Pearson's Chi-squared test of independence (2-tailed) or the Fisher Exact test, depending upon the size of the groups compared, were employed for investigating the relationship between categorical variable. The correlation between DDR biomarkers, when analyzed as continuous variables (percentage of nuclear-expressing cancer cells), was assessed with the Pearson's correlation coefficient. Survival curves were estimated with the Kaplan-Meier product-limit method and compared by log-rank test. Variables potentially affecting PFS and OS were tested in univariate Cox proportional hazard models. Multivariate Cox models were built with variables testing significant at the univariate analysis and by stepwise regression through backward elimination. The related estimates were reported as hazard ratio (HR) and 95% confident interval (CI). Cross-

validation was used for assessing the robustness of the molecular background of interest ($H2AX^{high}$/$pATM^{high}$); univariate Cox analysis was carried out on the group of patients treated in phase II trials (training set), and then on the group of patients treated in routine clinical practice (validation set). Level of significance was defined as $p<0.05$. Statistical analyses were carried out using SPSS version 21.0 (SPSS Inc., Chicago, Illinois, USA).

RESULTS

Baseline characteristics of the patients and expression pattern of DDR biomarkers

Baseline characteristics of the 110 patients included in the present analysis are summarized in Table 1. Median age at diagnosis was 60.8 years. Seventy-two (65.5%) patients had metastatic disease, whereas 38 (34.5%) patients received chemotherapy for inoperable locally advanced disease. Regarding chemotherapy regimens, 68 (61.8%) patients received three-drug regimens, and taxane-containing chemotherapy was administered to 61 (55.5%) patients. Fifty-six patients were treated in prospective phase II trials. Distributions of DDR biomarkers did not significantly differ by treatment received in the first-line setting, i.e., number of therapeutic agents and taxane administration (data available upon request). We recorded a significant positive association between DDR biomarkers (Figure 1, panel A); a significant correlation between the various markers was also observed when they were analyzed as continuous variables considering the percentage of nuclear-expressing tumor cells (Figure 1, panel B).

DDR biomarkers and PFS, sensitivity analysis and outliers analysis

Patients whose tumors had elevated expression levels of γ-H2AX experienced significant shorter PFS compared with their negative counterparts (log rank $p=0.001$) (Figure 2, panel A), and a non significant association was observed for pATM (log rank $p=0.081$) (Figure 2, panel B). A model of double positivity (γ-$H2AX^{high}$/$pATM^{high}$) was also significantly associated with shorter PFS (log rank $p<0.001$) (Figure 2, panel C). Similar results emerged from a sensitivity analysis carried out on

patients with metastatic disease, thus excluding patients with locally advanced tumors (Suppl. Figure 3). In the entire cohort, we did not appreciate any association between pChk1, pWee1, pRPA32 and pChk2 and PFS (data available upon request). Nevertheless, a significant association was observed between γ-H2AX, pATM and pChk1 expression and negative and positive outliers, namely patients in the lowest (PFS<3 months) and highest (PFS>10 months) quartile (N=54), with a trend toward statistical significance for pRPA32 (p=0.057, Suppl. Table 2). Thus, an extensive activation of upstream branches of the DDR machinery may delineate a fraction of patients with intrinsically chemoresistant GC.

Uni- and multivariate Cox regression models for PFS and OS, cross-validation

As our results suggested a prominent role of DSB repair avenues in feeding chemoresistance, the γ-H2AX^{high}/pATM^{high} phenotype was verified in uni- and multivariate Cox models for PFS and OS. Uni- and multivariate Cox regression analyses revealed that patients with double-positive tumors were at increased risk of progression (multivariate Cox: HR 2.23, 95%CI: 1.47-3.40, p<0.001) (Table 2). The consistency of the H2AX^{high}/pATM^{high} model was further confirmed upon cross-validation (Suppl. Figure 4). Uni- and multivariate Cox models for OS (Table 3) yielded comparable results (univariate Cox: HR 1.71, 95%CI: 1.03-2.86, p=0.039; multivariate Cox adjusted for variables significant at univariate analyses: HR 1.87, 95%CI: 1.00-3.50, p=0.050; multivariate Cox with stepwise backward elimination: HR: 2.07, 95%CI: 1.20-3.58, p=0.009).

Predictive significance of DDR biomarkers according to TP53 and ARID1A mutations

Basal DDR activation may stem from the underlying genetic portrait of the disease, as a number of frequently-mutated genes in GC were preclinically tied to the activation of the intra-S and G_2/M checkpoints. To address this issue, targeted deep resequencing was conducted to evaluate *TP53* and *ARID1A* mutational status in 90 patients with sufficient biological materials. As shown in Figure 3

(panel A), 53% and 47% of the samples harbored *TP53* and *ARID1A* mutations, respectively (detailed information in Suppl. Table 3). In order to investigate whether the genetic status of *TP53* and/or *ARID1A* impacted the predictive ability of the γ-H2AX^{high}/pATM^{high} model, univariate Cox analyses were carried out in the different genetic contexts. The relationship between the γ-H2AX^{high}/pATM^{high} model and PFS was independent on *TP53*, albeit to a borderline significant extent in the *TP53*-mutated setting (p=0.052) (Figure 3, panel B). Likewise, the H2AX^{high}/pATM^{high} model continued to be associated with an increased risk of disease progression in the *ARID1A* wild-type background. Conversely, this relationship was lost in the *ARID1A*-mutated setting (Figure 3, panel B).

DISCUSSION

We herein examined the expression pattern of central orchestrators of the DDR response in a relatively large series of advanced GC patients treated with first-line therapy. Approximately half of these patients received chemotherapy in prospective phase II trials.[10-13] As activation of the ATM-Chk2/ATR-Chk1-Wee1 cascade may be related to, or enforced by, genetic events altering cell cycle progression and DNA repair efficiency, the mutational status of top-ranking mutated genes in GC (*TP53* and *ARID1A*) was assessed. Our results suggest that: i) a subset of GC is characterized by a robust DDR activation, ii) activation of the system that signals the presence of DSBs may be detrimental for these patients conferring chemoresistant features, as denoted by uni- and multivariate Cox models for PFS and OS, and iii) the clinical relevance of this process may be independent on *TP53* status, whereas it seemed affected by protein-damaging *ARID1A* mutations.

Our study, which is hypothesis-generating in nature, raised a number of points that may streamline the identification of efficient DDR-related predictive factors. In the search for molecular determinants anticipating the benefits patients receive from chemotherapy, the focus was mostly placed on single effectors acting within a specific repair avenue deputed to correct a given type of genetic lesion.[16,17] Our results suggest that molecular mechanisms through which cancer cells sense

DNA damage and the way this information is transmitted/amplified should not be underestimated in future investigations.

The lesson learned from preclinical investigations striving to elucidate how cancer cells exploit the DDR to survive chemotherapy is that protein-protein interactions should be viewed in light of genetic contexts that, at the level of protein function, intersect these communications.[18,19] For instance, pharmacological inhibitors of upstream DDR kinases were originally advocated as synthetically lethal therapeutics for targeting p53-defective tumors.[18,19] We neither observed an impact of *TP53* mutations on the predictive ability of DDR markers, nor any association between *TP53* status and the investigated molecular endpoints was found (data available upon request). Albeit these data may question the idea of an intra-S/G_2-M checkpoint dependency of p53-defective tumors, two not mutually exclusive molecular mechanisms deserve mention. First, *TP53* mutations not invariably result in a defective protein, but rather some missense mutations (overall accounting for ~75% of all *TP53* mutations) encode for a protein with oncogenic potential (gain-of-function mutations).[20] Even though these mutated p53 forms exert a dominant-negative effect on wild-type p53, they also endow the new protein with tumor-promoting abilities.[20] Second, GC harbors amplification of cyclins (e.g. *CCNE1* and *CCND1*) and cyclin-dependent kinases (e.g. *CDK6*) that drive cells from G_1 into S phase, suggesting the existence of another level of deregulation at the G_1-S transition beyond loss of p53 function.[9] Tailored preclinical investigations are warranted for the correct interpretation of how the different spectrum of *TP53* lesions, together with other defects in the G_1-S transition machinery, lead to intra-S and G_2/M checkpoint dependency and fuel chemoresistance.

Next, the tumor suppressor ARID1A, a subunit of the SWI/SNF chromatin remodeling complex, was identified as frequently mutated in a broad spectrum of tumors.[21] It was recently demonstrated that ARID1A is recruited to DNA DSBs via ATM/ATR signaling where it facilitates the processing of DNA lesions.[22] ARID1A-deficient cells displayed impaired initiation and maintenance of the G_2-

M checkpoint and overall a suboptimal response to DSBs. Moreover, ARID1A suppression resulted in reduced non-homologous end joining activity and sensitized cell lines to cytotoxic agents. [23] As we did not observe any association between DDR markers and an increased risk of disease progression in the *ARID1A*-mutated setting, we speculate that defective ARID1A function may interfere with molecular mechanisms conferring chemoresistant traits. In order to provide further ground to this observation, we strive to establish a collection of patient-derived cancer (stem) cell lines and xenografts. There are several reasons behind this choice: i) even though our data suggest a connection between *ARID1A* status, the DDR and sensitivity/resistance to chemotherapy, we were unable to conduct mechanistic *in vitro* studies given the limited availability of commercially available (e.g. ATCC) GC cell lines carrying *ARID1A* mutations, ii) prospectively isolated patient-derived GC cells and xenografts are expected to better recapitulate the molecular portrait of the parental tumors compared with commercial cell lines,[24] which expose investigators to a number of potential biases (e.g. cross-contamination, molecular artifacts stemming from multiple passages in culture) and whose clinical usefulness is the focus of intense debate in the scientific community.[25] Moreover, prospective isolation and characterization of patient-derived GC cells will enable us to combine mechanistic and correlative studies, iii) the majority of commercially available GC cell lines were obtained from patients of Asian ethnicity, and it is known that non-negligible differences exist between GC arising in Asians and Caucasians, both at the molecular and clinical level. [26-28]

A third level of genetic characterization that should be pursued refers to oncogene-induced replication stress. This condition is characterized by increased numbers of stalled or collapsed replication forks that elicits an ATM-ATR-dependent response.[3] *MYC* overexpression/amplification is among the best characterized condition underlying oncogene-induced replication stress, and inhibition of DDR kinases was proposed as a synthetic lethality-based approach for targeting MYC-driven tumors.[29,30] Approximately 15% of GCs carry *MYC* amplification/mutation and a "target of MYC activation" signature was found to be active in all the four GC subtypes.[9]

A mention deserves the positive association in the expression of the investigated markers (Figure 1), which suggests a robust activation of the DDR response in a subset of GC patients. This finding potentially holds therapeutic perspective in light of the number of cell cycle checkpoint inhibitors that have entered clinical development over the past years.[18] Even though results from pioneering clinical trials dampened the expectations around these compounds,[18] these studies were weakened by lack of biomarker analysis. Our study may represent a first step toward delineating the target population for clinical trials with cell cycle checkpoint inhibitors, enforcing the concept that extensive DDR pathway analysis should be carried out both at the protein and genetic level to streamline the development of these compounds.

Finally, our interest toward the identification of DDR-linked predictive biomarkers prompted us to initiate a further level of molecular characterization, with special emphasis being placed on the Fanconi Anemia (FA) pathway. The FA pathway is involved in DNA interstrand crosslink (ICL) repair,[31] and its over-activation was found to be a shared trait across the four molecular subtypes identified by the TCGA.[9] To this end, we activated a multicenter, retrospective study in an expanded case series with the aim of investigating the predictive significance of various FA pathway markers. More specifically, we planned the study of FA pathway components involved in lesion recognition (e.g. FANCM), components of the FA core complex (e.g. FANCA to FANCT), and the ERCC4-ERCC1 complex, which is central in unhooking ICL lesions.[31-33]

In conclusion, our data pointed to the activation of upstream regulators of the DDR machinery as potential biomarkers for predicting the efficacy of chemotherapy in GC patients. Prospective validation of these encouraging findings, together with a deeper characterization of genetic events tied to the processing of DNA lesions, is advised in order to gain a better understanding on how this information can be transferred into the clinical setting.

Authors' contributions: MM-S and RDM conceived and designed the study. LR, EM, BC, CAA, EG, EP, MGD, SB and MM were involved in molecular pathology analysis. FDN, FG and MF carried out targeted DNA deep sequencing. FS, MP, IT and MB performed bioinformatic and statistical analyses. LP, PV, DS and LDL acquired the data related to clinical features, treatment administered and therapeutic outcomes. IV was involved in study design and provided critical revision of the manuscript for important intellectual content. MM-S wrote the final version of the manuscript. All authors read and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Acknowledgements: We thank Tania Merlino and Ana Maria Edlisca for editorial assistance. IV is supported by the Italian Association for Cancer Research (AIRC, MFAG), Ministero Italiano della Salute, and the Programma per i Giovani Ricercatori "Rita Levi Montalcini" 2011. Financial support: This study was supported by an intramural research grant to the "Gastrointestinal Tumors Translational Research Group".

Conflict of interest disclosure statement: The authors declare no potential conflicts of interest.

References

[1]. Jeggo PA, Pearl LH, Carr AM. DNA repair, genome stability and cancer: a historical perspective. Nat Rev Cancer. 2016;16:35-42.

[2]. Scott TL, Rangaswamy S, Wicker CA, Izumi T. Repair of oxidative DNA damage and cancer: recent progress in DNA base excision repair. Antioxid Redox Signal. 2014;20:708-26.

[3]. Gaillard H, García-Muse T, Aguilera A. Replication stress and cancer. Nat Rev Cancer. 2015;15:276-89.

[4]. Dixon H, Norbury CJ. Therapeutic exploitation of checkpoint defects in cancer cells lacking p53 function. Cell Cycle. 2002;1:362-8.

[5]. Bartek J, Lukas J. Chk1 and Chk2 kinases in checkpoint control and cancer. Cancer Cell. 2003;3:421-9.

[6]. Vriend LE, De Witt Hamer PC, Van Noorden CJ, Würdinger T. WEE1 inhibition and genomic instability in cancer. Biochim Biophys Acta. 2013;1836:227-35.

[7]. Vici P, Buglioni S, Sergi D, et al. DNA Damage and Repair Biomarkers in Cervical Cancer Patients Treated with Neoadjuvant Chemotherapy: An Exploratory Analysis. PLoS One. 2016;11:e0149872.

[8]. Vici P, Di Benedetto A, Ercolani C, et al. Predictive significance of DNA damage and repair biomarkers in triple-negative breast cancer patients treated with neoadjuvant chemotherapy: An exploratory analysis. Oncotarget. 2015;6:42773-80.

[9]. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513:202-9.

[10]. Di Lauro L, Vici P, Belli F, et al. Docetaxel, oxaliplatin, and capecitabine combination chemotherapy for metastatic gastric cancer. Gastric Cancer. 2014;17:718-24.

[11]. Di Lauro L, Giacinti L, Arena MG, et al. Phase II study of epirubicin, oxaliplatin and docetaxel combination in metastatic gastric or gastroesophageal junction adenocarcinoma. J Exp Clin Cancer Res. 2009;28:34.

[12]. Di Lauro L, Nunziata C, Arena MG, et al. Irinotecan, docetaxel and oxaliplatin combination in metastatic gastric or gastroesophageal junction adenocarcinoma. Br J Cancer. 2007;97:593-7.

[13]. Di Lauro L, Belli F, Arena MG, et al. Epirubicin, cisplatin and docetaxel combination therapy for metastatic gastric cancer. Ann Oncol. 2005;16:1498-502.

[14]. McShane L, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor MARKer prognostic studies (REMARK), Eur J Cancer, 2005, 41, 1690-1696.

[15]. Bang YJ, Van Cutsem E, Feyereislova A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. Lancet. 2010;376:687-97.

[16]. Friboulet L, Olaussen KA, Pignon JP, et al. ERCC1 isoform expression and DNA repair in non-small-cell lung cancer. N Engl J Med. 2013;368:1101-10.

[17]. Olaussen KA, Dunant A, Fouret P, et al. DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. N Engl J Med. 2006;355:983-91.

[18]. Maugeri-Saccà M, Bartucci M, De Maria R. Checkpoint kinase 1 inhibitors for potentiating systemic anticancer therapy. Cancer Treat Rev. 2013;39:525-33.

[19]. Weber AM, Ryan AJ. ATM and ATR as therapeutic targets in cancer. Pharmacol Ther. 2015;149:124-38.

[20]. Oren M, Rotter V. Mutant p53 gain-of-function in cancer. Cold Spring Harb Perspect Biol. 2010;2:a001107.

[21]. Wu RC, Wang TL, Shih IeM. The emerging roles of ARID1A in tumor suppression. Cancer Biol Ther. 2014;15:655-64.

[22]. Shen J, Peng Y, Wei L, et al. ARID1A Deficiency Impairs the DNA Damage Checkpoint and Sensitizes Cells to PARP Inhibitors. Cancer Discov. 2015;5:752-67.

[23]. Watanabe R, Ui A, Kanno S, et al. SWI/SNF factors required for cellular resistance to DNA damage include ARID1A and ARID1B and show interdependent protein stability. Cancer Res. 2014;74:2465-75.

[24]. Tentler JJ, Tan AC, Weekes CD, et al. Patient-derived tumour xenografts as models for oncology drug development. Nat Rev Clin Oncol. 2012;9:338-50.

[25]. Gillet JP, Varma S, Gottesman MM. The clinical relevance of cancer cell lines. J Natl Cancer Inst. 2013;105:452-8.

[26]. Strong VE, Song KY, Park CH, et al. Comparison of gastric cancer survival following R0 resection in the United States and Korea using an internationally validated nomogram. Ann Surg. 2010;251:640-6.

[27]. Hundahl SA, Phillips JL, Menck HR. The National Cancer Data Base Report on poor survival of U.S. gastric carcinoma patients treated with gastrectomy: Fifth Edition American Joint Committee on Cancer staging, proximal disease, and the "different disease" hypothesis. Cancer. 2000;88:921-32.

[28]. Lin SJ, Gagnon-Bartsch JA, Tan IB, et al. Signatures of tumour immunity distinguish Asian and non-Asian gastric adenocarcinomas. Gut. 2015;64:1721-31.

[29]. Murga M, Campaner S, Lopez-Contreras AJ, et al. Exploiting oncogene-induced replicative stress for the selective killing of Myc-driven tumors. Nat Struct Mol Biol 2011;18:1331-5.

[30]. Brooks K, Oakes V, Edwards B, et al. A potent Chk1 inhibitor is selectively cytotoxic in melanomas with high levels of replicative stress. Oncogene. 2013;32:788-96.

[31]. Ceccaldi R, Sarangi P, D'Andrea AD. The Fanconi anaemia pathway: new players and new functions. Nat Rev Mol Cell Biol. 2016;17:337-49.

[32]. Hodskinson MR, Silhan J, Crossan GP, et al. Mouse SLX4 is a tumor suppressor that stimulates the activity of the nuclease XPF-ERCC1 in DNA crosslink repair. Mol Cell. 2014;54:472-84.

[33]. Klein Douwel D, Boonen RA, Long DT, et al. XPF-ERCC1 acts in Unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. Mol Cell. 2014;54:460-71.

Figure legends

Figure 1: Relationship between the expression levels of six DDR biomarkers (γ-H2AX, pATM, pChk2, pRPA32, pChk1 and pWee1). Associations are shown in the oncoprint in panel A, correlations in panel B (*N*=110).

Figure 2: Kaplan-Meier survival curves of progression-free survival. γ-H2AX^{high} vs γ-H2AX$^{low/neg}$ (A); pATM^{high} vs pATM$^{low/neg}$ (B); γ-H2AX^{high}/pATM^{high} vs single-positive tumors/double-negative tumors (C) (*N*=110).

Figure 3: Panel A: Oncoprint showing the distribution and type of *TP53* and *ARID1A* mutations in 90 gastric cancer patients. Panel B: Forest plot for subgroup analysis displaying the relationship between the γ-H2AX^{high}/pATM^{high} model and progression-free survival in the different *TP53* and *ARID1A* genetic backgrounds (wild-type and mutated).

Supplementary Figure 1: Representative examples of immunohistochemical expression of six DNA damage and repair biomarkers in gastric cancer patients. Upper panel: a triple-positive tumor with nuclear γ-H2AX(a), pATM (b), and pChk2 (c) immunoreactivity. Lower panel: a triple-positive tumor with nuclear pChk1 (d), pRPA32 (e) and pWee1(f) immunoreactivity.

Supplementary Figure 2: Experimental workflow for filtering *TP53* and *ARID1A* mutations in 90 gastric cancer patients.

Supplementary Figure 3: Kaplan-Meier survival curves of progression-free survival in patients with metastatic disease. γ-H2AX^{high} vs γ-H2AX$^{low/neg}$ (A); γ-H2AX^{high}/pATM^{high} vs single-positive tumors/double-negative tumors (B) (*N*= 72).

Supplementary Figure 4: Univariate Cox models for progression-free survival (cross-validation) in the group of patients treated in phase II trials (training set) and in the group who received chemotherapy in real-world clinical practice (validation set).

Table 1: Baseline characteristics of gastric cancer patients included in this study (N=110)

Characteristics		N (%)
Age at diagnosis	Median (min-max) [IQ range]	60.8 (28-79) [51.3-67.4]
Gender	Male	58 (52.7)
	Female	52 (47.3)
ECOG PS	0	58 (52.7)
	1-2	52 (47.3)
Stage	Locally Advanced	38 (34.5)
	Metastatic	72 (65.5)
Previous Surgery	No	53 (48.2)
	Yes	57 (51.8)
Neoadjuvant/Adjuvant chemotherapy	No	82 (74.5)
	Yes	28 (25.5)
Lauren classification	Intestinal	50 (45.5)
	Diffuse	48 (43.6)
	Mixed	12 (10.9)
Grade	G2	24 (21.8)
	G3	82 (74.5)
	Unknown	4 (3.7)
Localization	Esophagogastric Junction (EOJ)	10 (9.1)
	Stomach	100 (90.9)
Agents (N)	2	42 (38.2)
	3	68 (61.8)
Taxanes	No	49 (44.5)
	Yes	61 (55.5)

Table 2 : Uni- and multivariate Cox regression models for PFS (N=110)

		Univariate Cox regression model		Multivariate Cox regression model[§]		Multivariate Cox regression model**	
		HR (95%CI)	**p-value**	**HR (95%CI)**	**p-value**	**HR (95%CI)**	**p-value**
γ-H2AX^{pos}/pATM^{pos}	**positive vs other**	2.23 (1.47-3.40)	<0.001	2.11 (1.35-3.28)	0.001	2.23 (1.47-3.40)	<0.001
ECOG-PS	**1-2 vs 0**	1.28 (0.85-1.92)	0.238				
Stage	**Met vs loc adv**	1.23 (0.80-1.89)	0.344				
Localization	**Stomach vs EOJ**	0.66 (0.33-1.33)	0.243				
N° metastatic sites	**2-3 vs 1**	1.54 (0.99-2.37)	0.052	1.22 (0.77-1.93)	0.390		
Peritoneal metastasis	**Yes vs No**	0.70 (0.47-1.06)	0.092				
Taxanes	**Yes vs No**	0.88 (0.58-1.33)	0.551				

[§] Adjusted for variables significant at the univariate analysis.

**Backward stepwise exclusion.

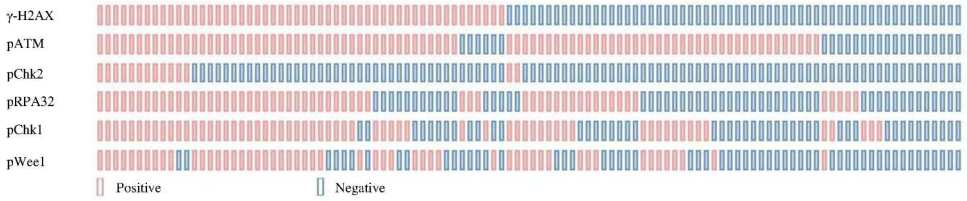
Table 3 : Uni- and multivariate Cox regression models for OS (N=110).

		Univariate Cox regression model		Multivariate Cox regression model[§]		Multivariate Cox regression model**	
		HR (95%CI)	**p-value**	**HR (95%CI)**	**p-value**	**HR (95%CI)**	**p-value**
γ-H2AX^{pos}/pATM^{pos}	**positive vs other**	1.71 (1.03-2.86)	0.039	1.87 (1.00-3.50)	0.050	2.07 (1.20-3.58)	0.009
ECOG-PS	**1-2 vs 0**	1.05 (0.62-1.75)	0.860				
Stage	**Met vs loc adv**	1.17 (0.68-2.01)	0.574				
Localization	**Stomach vs EOJ**	0.54 (0.23-1.28)	0.161				
N° metastatic sites	**2-3 vs 1**	1.67 (0.99-2.79)	0.054	1.24 (0.65-2.35)	0.515		
Peritoneal metastasis	**Yes vs No**	0.69 (0.41-1.15)	0.156				
Taxanes	**Yes vs No**	0.56 (0.33-0.93)	0.027	0.55 (0.31-0.99)	0.048	0.54 (0.30-0.97)	0.041
Response to first-line therapy	**Yes vs No**	0.42 (0.25-0.73)	0.002	0.48 (0.26-0.88)	0.018	0.50 (0.28-0.90)	0.022
2nd line chemotherapy	**Yes vs No**	0.27 (0.15-0.47)	<0.001	0.41 (0.23-0.73)	0.002	0.41 (0.23-0.72)	0.002

[§] Adjusted for the variables significant at the univariate analysis.

**Backward stepwise exclusion.

A



γ-H2AX and pATM: Chi2 p=0.013
γ-H2AX and pChk2: Fisher p=0.003
γ-H2AX and pRPA32: Chi2 p<0.001
γ-H2AX and pChk1: Chi2 p<0.001
γ-H2AX and pWee1: Chi2 p<0.001
pATM and pChk2 : Fisher p=0.037
pATM and pRPA32: Chi2 p=0.031
pATM and pChk1: Chi2 p=0.002
pATM and pWee1: Fisher p<0.001
pChk2 and pRPA32: Fisher p=0.009
pChk2 and pChk1: Fisher p<0.001
pChk2 and pWee1: Fisher p=0.003
pRPA32 and pChk1: Chi2 p<0.001
pRPA32 and pWee1: Chi2 p=0.002
pChk1 and pWee1: Chi2 p<0.001

B

Pearson's Correlation Coefficient	**γ-H2AX**	**pATM**	**pChk2**	**pRPA32**	**pChk1**
pATM	0.32**				
pChk2	0.50**	0.33**			
pRPA32	0.54**	0.32**	0.43**		
pChk1	0.48**	0.42**	0.38**	0.53**	
pWee1	0.44**	0.36**	0.39**	0.41**	0.51**

** p<0.01

Figure 1: Relationship between the expression levels of six DDR biomarkers (γ-H2AX, pATM, pChk2, pRPA32, pChk1 and pWee1). Associations are shown in the oncoprint in panel A, correlations in panel B (N=110).

254x190mm (300 x 300 DPI)

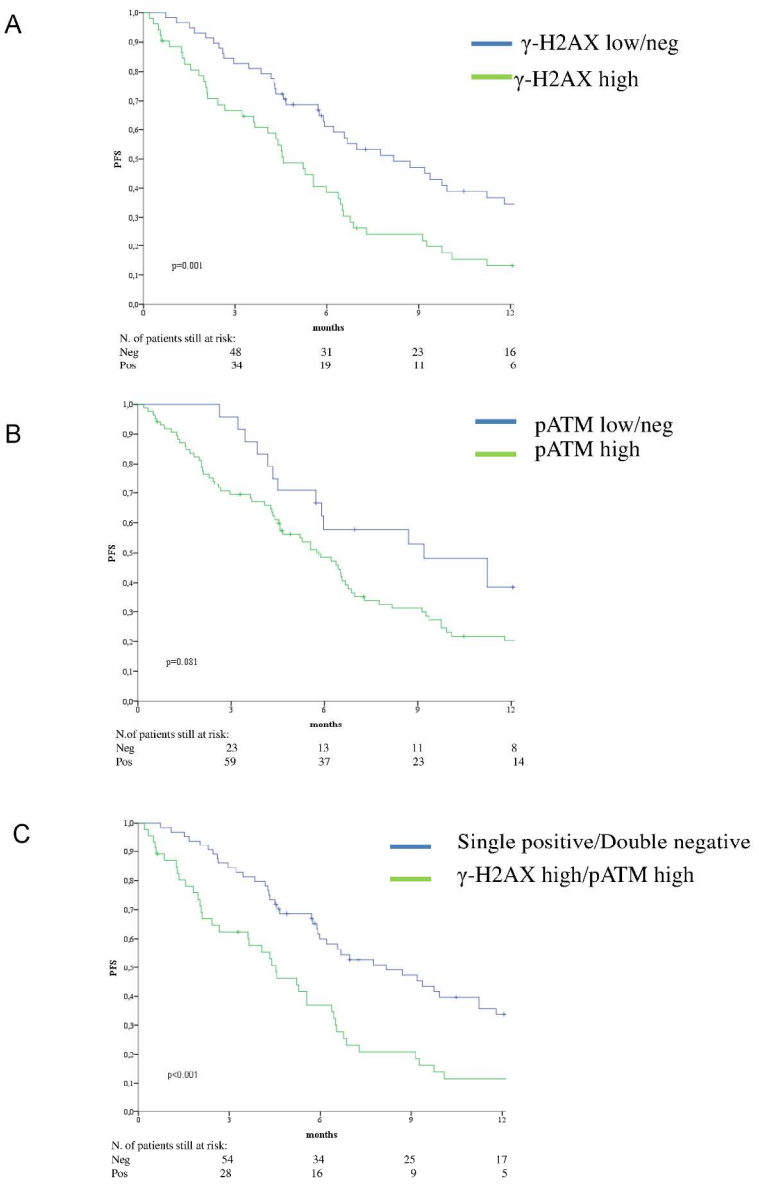


Figure 2: Kaplan-Meier survival curves of progression-free survival. γ-H2AXhigh vs γ-H2AXlow/neg (A); pATMhigh vs pATMlow/neg (B); γ-H2AXhigh/pATMhigh vs single-positive tumors/double-negative tumors (C) (N=110).

209x296mm (300 x 300 DPI)

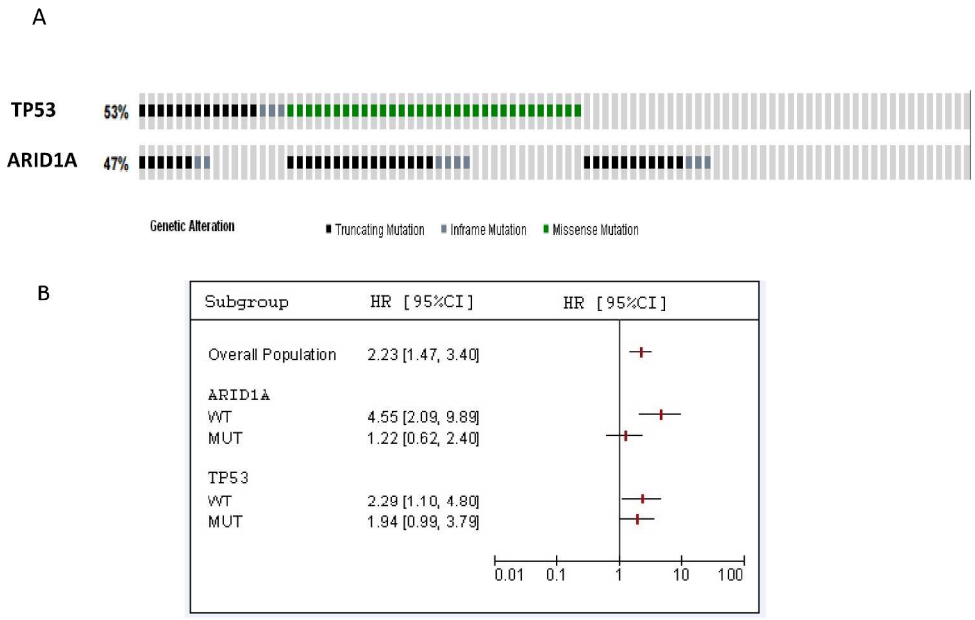


Figure 3: Panel A: Oncoprint showing the distribution and type of TP53 and ARID1A mutations in 90 gastric cancer patients. Panel B: Forest plot for subgroup analysis displaying the relationship between the γ-H2AXhigh/pATMhigh model and progression-free survival in the different TP53 and ARID1A genetic backgrounds (wild-type and mutated).

254x190mm (300 x 300 DPI)