Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A DNA damage response-like phenotype defines a novel subset of colon cancer

Stefania Mauro
Vimercate Hospital, ASST-Vimercate
Maddalena Bolognesi
Università di Milano-Bicocca https://orcid.org/0000-0002-8238-2369
Nicoletta Villa
San Gerardo Hospital, ASST-Monza
Giulia Capitoli
Università di Milano-Bicocca
Laura Furia
European Institute of Oncology IRCCS
Francesco Mascadri
Universita' degli Studi Milano-Bicocca
Nicola Zucchini
San Gerardo Hospital, ASST-Monza https://orcid.org/0000-0003-4997-2807
Mauro Totis
San Gerardo Hospital, ASST-Monza
Mario Faretta
European Institute of Oncology IRCCS
Stefania Galimberti
Università di Milano-Bicocca
Giorgio Bovo
Vimercate Hospital, ASST-Vimercate
Giorgio Cattoretti (giorgio.cattoretti@unimib.it)
Universita' degli Studi Milano-Bicocca https://orcid.org/0000-0003-3799-3221

Article

Keywords: DNA Damage, DNA Mismatch Repair, Colonic Neoplasms, T-Lymphocytes, Multiplex Immunostaining, chromosome 20q amplification, Interferon Type I.

Posted Date: July 29th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1868531/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Colon adenocarcinoma (COAD) has a limited range of diversified, personalized therapeutic opportunities, besides DNA hypermutating cases; thus both new targets or broadening existing strategies for personalized intervention are of interest. Routinely processed material from 246 untreated COADs with clinical follow-up was probed for evidence of DNA damage response (DDR) by multiplex immunofluorescence and immunohistochemical staining for DDR complex proteins (γH2AX, pCHK2, pNBS1) and for type I interferon response, T-lymphocyte infiltration (TILs), mutation mismatch repair defects (MMRd). FISH analysis for chromosome 20q copy number variations was obtained. 33.7% of COAD display a coordinated DDR on quiescent, non-senescent, non-apoptotic glands, irrespective of TP53, chromosome 20q abnormalities, type I IFN response. Clinicopathological parameters did not differentiate DDR + cases from the other cases. TILs were equally present in DDR and non-DDR cases. DDR + MMRd cases were preferentially retaining wild-type MLH1. The outcome after 5FU-based chemotherapy was not different in the two groups. DDR + COAD represents a subgroup not aligned with known diagnostic, prognostic or therapeutic categories, with potential new targeted treatment opportunities.

Introduction

Cancer is in essence a genetic disease [1] and DNA replication and maintenance is central to its origin and sustainment. During cancer development and life history, DNA damage may occur subsequent to a multitude of events, endogenous (oxidative damage, replication stress, mitotic dysfunction, etc.) or exogenous (gamma X rays, genotoxic drugs) [2–4].

A response to DNA damage (DDR), which leads to DNA repair, is essential for mammalian cells to maintain genome integrity, cell cycle proficiency and to guarantee tissue homeostasis and function [4]. Repair of the DNA damage is a complex, highly coordinated task which involves the sequential recruitment of sensors, mediators, repair complex organizers and new DNA generation at the break site [4].

Phosphorylation of histone H2AX (γH2AX) by ATM (Ataxia telangiectasia, mutated) at the DNA double strand break sites (DbSB) is the first and reproducibly detectable phenotype of the response, followed by activation of CHK2 (pCHK2), which in turn participate to the cascade, leading to the repair complex assembly, cell cycle pausing etc. A TP53-binding protein, 53BP1, is required for the repair complex formation and for the aggregation of all the actors at the break site, forming discrete foci, containing the molecules mentioned above and other DDR-associated proteins such as Rad51 and NBS1 [5].

Besides known DNA damaging drugs, the induction of replication arrest by hydroxyurea or UV irradiation [6], the growth factor-dependent stimulation of tissue hyperplasia [7] or the controlled induction of transcription from an HIV promoter [8] causes the focal accumulation of a DNA damage response complex identifiable by γH2AX, pCHK2, 53BP1 etc. The formation of DDR foci requires and intact ATM response [9].

In the presence of a conserved DDR, the cell cycle is paused and, depending on the extent of DNA damage and the level of TP53 induction [10], resolution of the stalling, senescence [9,11] or apoptosis may ensue. Cancer cells overcome a senescence-dependent permanent replication arrest by silencing this barrier via mutation of several members of the response and continuing proliferation [9,11].

COAD contains a variety of genetic lesions of the DNA damage repair (DDR) genes: 44% bear lesion of at least one Direct Repair (DR) gene, 21% and 22% of Homologous Recombination (HR) and Mismatch Repair (MMR) genes respectively, 17% of Damage Sensor (DS) genes [12]. These are potential targets for personalized therapy [3,13–15] However, at the present time, there is no established therapy and no recognized COAD subset which can be targeted via DDR abnormalities.

While the dynamic interplay of the DDR machinery has been demonstrated in great detail in artificial in vitro experiments, relatively scanty data are available for established tumors, mostly limited to single or double tissue stainings or by extractive methods; the single cell landscape of the many molecular actors is still obscure, thus we wanted to analyze the DDR in untreated colon cancer. What we found is at variance of the established narrative and we show in a third of untreated colon adenocarcinomas (COAD) foci of DDR in quiescent cells, marked by γH2AX, pCHK2 and pNBS1, unrelated to TP53, p21/CDKN1A or p16/CDKN2A expression.

Materials And Methods

Human specimens

Untreated colon cancer biopsies and associated clinicopathologic data for 246 patients were extracted from the laboratory information systems of both Hospitals by the Authors with clinical privileges and anonymized. Paraffin blocks and the area to be analyzed were selected by a Pathologist after a review of the Hematoxylin and Eosin (H&E) stain. All fresh surgical pathology specimens (COAD and leftovers such as pediatric tonsils) were fixed overnight at RT in buffered 4% formaldehyde (Bio-Optica Milano Spa, Milano, Italy), processed through a graded ethanol gradient, then in xylene and embedded in molten paraffin for sectioning.

A pilot study involved matched cryosections from isopentane-frozen tissue at -80° C and sections from routinely processed tissue (formalin-fixed, paraffin embedded; FFPE) from a small number of consecutive anonymous cases, with the scope of validating the staining conditions for both as reported previously [16].

The study has been approved by the Institutional Review Board Comitato Etico Brianza, N. 3204, “High-dimensional single cell classification of pathology (HDSSCP)”, October 2019. Patients consent was obtained or waived according to article 89 of the EU general data protection regulation 2016/679 (GDPR) and decree N. 515, 12/19/2018 of the Italian Privacy Authority.

Tissue microarrays

Tissue microarrays (TMA) were prepared as previously published on a Tissue Microarrayer Galileo model TMA CK4600 (Integrated System Engineering srl, Milan, Italy). Three cores of 1 mm per patient were used, from non necrotic, random tumor areas.

Antigen retrieval

Antigen retrieval (AR) was performed placing the dewaxed, rehydrated sections [17] in a 800 ml glass container filled with the retrieval solutions (EDTA pH 8; 1 mM EDTA in 10 mM Tris-buffer pH 8, Merck Life Science S.r.l.,Milano, Italy; cat. T9285)), irradiated in a household microwave oven at full speed for 8 min, followed by intermittent electromagnetic radiation to maintain constant boiling for 30 min, and cooling the sections to about 50° C before use.

Immunohistochemistry and immunofluorescence

For immunohistochemistry (IHC), optimally diluted, validated primary antibodies (Supplemental Table S1) were applied overnight, washed in 50mM Tris-HCl buffer (pH 7.5) containing 0.01% Tween-20 (Merck) and 100 mM sucrose (TBS-Ts) [18], counterstained with a horseradish peroxidase–conjugated polymer (Vector Laboratories, Burlingame, CA, USA), washed, developed in DAB (Dako, Glostrup, DK), lightly counterstained and mounted.

Multiple immunofluorescent (IF) labeling and cyclic staining and stripping was previously described in detail [17].

Briefly, the sections were incubated overnight with optimally diluted primary antibodies in combination of up to four, washed and counterstained with specific distinct fluorochrome-tagged secondary antibodies (Supplemental Table S1) [17]. The slides, counterstained with DAPI and mounted, were scanned on an S60 Hamamatsu scanner (Nikon, Italia) at 20x magnification. The filter setup for seven color acquisition (DAPI, BV480, FITC, TRITC, Cy5, PerCp, autofluorescence/AF) was as published [19]. After a successful image acquisition, the sections were stripped according to the MILAN method [18] and stained with another round.

Preparation of immunofluorescent images for single cell analysis.

After the stainings were acquired, digital slide images (.ndpi) were imported as .tiff and registered with the AMICO software [20], based on ImageJ (RRID:SCR_003070). All DAPI images from different staining rounds were registered together and their coordinates of rotation and translation were used to align the individual marker images of each round. Once all images were aligned, autofluorescence was subtracted from FITC, TRITC and PerCp channels.

Nuclear segmentation was performed on DAPI using CellPose [21] and subsequently by eroding 2 pixel from each nuclear contour identified (ImageJ plugin (LabelsToRois) [22] (Supplementary Fig.S1).

Median value of each marker inside every single nucleus segmented (with related spatial coordinates) was extracted, recorded in a .csv file and scaled. Then .csv files were uploaded to the R Studio software (version 1.4) for a more detailed analysis.

Only Keratin + nuclei were considered for the analysis (See Supplementary Methods)

IHC scoring criteria

Sections were scored by two independent observers (SM and GC), which were unaware of the clinicopathologic data of the patients scored. Exclusive nuclear staining was considered for γH2AX, pNBS1 and pCHK2. Positive foci were assigned whenever there was a focal increase in nuclear staining above the background value of the other tumor cell nuclei and the stroma. The three-tiered intensity staining score was 0, + (1), ++ (2) for each marker. A case would be defined as DDR + if the sum of the scores of γH2AX, pNBS1 and pCHK2 was ≥ 4 or ≥ 3 if pNBS1 was at least ++. The number of positive cells was not recorded.

MSI was scored in the epithelial component as suggested by the antibody provider (Agilent) in the presence of nuclear-positive inflammatory and stromal cells. TP53 staining intensity was scored as suggested by Cole et al. [23] (Supplementary Fig.S2C). MX1 staining was scored in the epithelial compartment semiquantitatively (neg, focal/weak, strong) (Supplementary Fig.S2B).

Tumor infiltrating lymphocytes (TILs) quantification

The TMAs were immunostained for CD3 (rabbit polyclonal, Cat. GA503, Agilent, Santa Clara, CA) with an OMNIS autostainer (Agilent). Whole slide images, acquired with the S60 scanner at 20x (see above) were color deconvoluted [24] with a Fiji plugin [25], the DAB staining thresholded with the MaxEntropy algorithm, the watershed function was applied and particles ranging from 70 to 1,000 square pixels in size, 0.00–1.00 circular were counted. Results are expressed as the number of CD3 per mm^2 (Supplementary Fig.S2A).

In-situ hybridization for Chr20q

Two COAD microarrays were hybridized for chromosome 20 Fluorescent in-situ Hybridization (FISH) probes, long arm and centromeric (SPEC BCL2L1/CEN 20 Dual Color Probe, ZytoLight, ZytoVision GmbH, Bremerhaven, D) as recommended by the manufacturer. Three areas were selected from each case, containing 2 DDR + and 1 DDR- or three random DDR- areas, unbeknown to the geneticist who scored the samples.

Hybridization signals for D20Z1 (green) and BCL2L1(orange) probes were counted in preselected areas and in non overlapping nuclei. A variable number of nuclei (from a minimum of 20 to a maximum of 80) was counted for each sample area, depending on hybridization efficiency and presence of tumor cells. The cohybridization of BCL2L1 probe with the chromosome 20 centromeric probe allowed to distinguish polysomy from amplification (Supplementary Figure S3). Normal disomic nuclei showed two green and two red signals. Polysomic nuclei had a number of red and green signals per nucleus ≥ 3 and a ratio

BCL2L1/ D20Z1 of about 1. Amplification was defined when BCL2L1 signals were more represented than the centromeric ones and the ratio BCL2L1/D20Z1 was ≥ 2.

Disease-free survival analysis.

Event Free Survival was defined as the time from the date of diagnosis to the first cancer-related adverse event (progressive disease or relapse at any site), or last follow-up (censored observation), whichever occurred first. Follow-up was updated as of 2021. Survival curves were estimated with the Kaplan-Meier method, whereas the log-rank test was applied to compare the survival of different groups (i.e. DDR vs notDDR). The analysis was carried out using the open-source R software v.3.6.0 (R Foundation for Statistical Computing, Vienna, Austria).

Results

A pilot study on colon adenocarcinomas (COADs) employing frozen tissue sections, was run with a limited multiplex combinationations directed at DDR molecules, with antibodies compatible with frozen material [16] (Supplemental Table S1). It showed two distinct staining patterns: on one side tumors with rare, mostly isolated γH2AX + nuclei and absence of pNBS1 and pCHK2; on the other side, neoplastic glands containing γH2AX, pNBS1 and pCHK2, non proliferating (Supplemental figure S4).

In order to detail the complete phenotype and clinical correlation of the DDR in COAD, we collected 246 additional untreated cases as FFPE material and performed IHC and multiplex high-dimensional analysis via the MILAN method.

Focal areas comprising single neoplastic glands displaying a coordinated staining of γH2AX, pNBS1 and pCHK2 on serial sections were identified in 83/246 (33.7%) COADs (Table 1); these cases were defined DDR+. DDR + cells were cleaved Caspase 3 (clCasp3) negative and cleaved PARP (clPARP) negative (Fig. 1 and Supplementary Fig. S5). One single case in the cohort displayed scattered clCasp3 + malignant cells, which were densely γH2AX+, pCHK2+, pNBS1-, RAD51-, indicative of apoptosis (Fig. 1 and Supplementary Fig. S5); this case was not considered DDR+.

Variegation of the γH2AX staining was observed: about half of the DDR + cases, 37/84 (15% of the total) showed reduced γH2AX staining ($DDR^{\gamma H2AXLo}$), compared to the robust punctuated nuclear staining seen in the other half ($DDR^{\gamma H2AXHi}$) (Fig. 1 and Supplementary Fig. S5).

Staining for RAD51 was negative in all cases.

Immunostaining for mutated B-RAF V600E and molecular analysis for mutated KRAS was performed in 36 cases, not enough for an analysis of significance (not shown).

Single cell phenotype by IF multiplexing

By multiplex immunofluorescence [17], the complete phenotype of individual cells and neoplastic glands was investigated. Coexpression at the single cell level of γH2AX, pCHK2, pNBS1 was noticed in all cases with a DDR + phenotype (Fig. 2 and Supplementary Fig. S6, S7), but not in the DDRnot ones. However, a significant variegation of each marker relative to the others was noticed, with pCHK2 being the most extensively present, pNBS1 the least. DDR + foci showed absence or significantly reduced proliferation (Ki-67) compared to the adjacent epithelium or neighboring glands (Fig. 2 and Supplementary Fig. S6, S7). No p16/INK4A/CDKN2A or p21/CDK21A immunofluorescent staining was noticed in correspondence to the DDR + cells, however p21/CDK21A was focally visually appreciated in well differentiated glands (Fig. 3), in the absence of the DDR complex, in quiescent cells. In DDR + foci, TP53 was wild type or mutated (overexpressed or absent) (Fig. 2 and Supplementary Fig. S6, S7). When wild type, a decrease in intensity was noticed in the DDR + foci (Fig. 2 and Supplementary Fig. S4). No association with nuclear βCatenin and DDR + foci was appreciated (not shown). Apoptosis by cleaved PARP staining was not present in the DDR + foci (not shown).

To verify if the observed phenotype was common to all cases as a continuum or discrete (either DDR + or DDRnot), we analyzed whole tissue sections from six cases, which were judged as negative on the three 1mm TMA cores: no clusters or isolated γH2AX nuclei were observed.

Next, we profiled single tumor cells from each case to confirm in a global fashion the results gleaned from the multiplex IF images. Pairwise comparison of fluorescence intensity for each marker showed γH2AX being negatively correlated with p21/CDK21A, p27/CDKN1B, TP53 or Ki-67 (Supplementary Figure S8), while is correlated with pCHK2 and pNBS1 (in DDR + cases). p21/CDK21A and p27/CDKN1B co-regulate, as expected, and p27 is expressed at low levels on Ki-67 + tumor cells (Supplementary Figure S8).

Correlations with biological variables

The DDR can be activated as a consequence of mitotic dysfunction [3] or during progress through the cell cycle [26], although in this latter instance, no other molecule besides γH2AX has been identified.

In order to evaluate if an ongoing chromosomal instability would justify the DDR phenotype observed, leading to copy number variation limited to the DDR positive areas, we did FISH for the most frequent CNV abnormality in COAD, chromosome 20q amplification [27].

FISH was performed on 26 COAD samples (Table 2 and Supplementary Figure S3), of which 16 (62%) were disomic for chr20q and 8 (31%) showed polysomy (Table 2). Amplification was present in just one case (UPN45) in a fraction of the nuclei (Table 2) and monosomy was demonstrated in two cases (Table 2).

Abnormalities of chromosome 20q were equally distributed among DDR + and DDRnot cases.

No spatial heterogeneity of the karyotype was noted, limited to the probes used, including cases with subclonal alterations.

DDR may result in an immune activated phenotype via multiple mechanisms leading to a type I interferon-driven response [28, 29]. To assess such response within the epithelial malignant cells we stained 30 cases for MX1/MxA, a cytoplasmic protein induced by IFNα or IFNβ [30]. Three out of 14 DDR + and 8 out of 16 DDRnot cases expressed substantial cytoplasmic MX1 in tumor epithelium (chi square = 0.11), thus MX1 expression was unrelated to the DDR status of the tumor (Supplementary fig. S2B). An attempt to detect immunostimulatory cytoplasmic single-strand DNA [29] by antibody staining (Supplemental Table 1) on frozen and FFPE material was unsuccessful (not shown).

Infiltration by CD3 + T lymphocytes was equally represented in DDR+ (mean 370.4 ± 214.6 per sqmm) and DDRnot cases (mean 345.1 ± 233.1 per sqmm), after removing MMR deficient cases. The difference is statistically not significant (Wilcoxon rank test p = 0.61) (Supplementary fig. S2A).

Clinical correlations

None of the histopathological or clinicopathologic features was differentially distributed in a statistically significant fashion among DDR + and DDRnot cases (Table 1), except the frequency of MMR deficiency due to MSH2/6, which was overrepresented in the DDR + cases and significantly in DDR$^{\gamma H2AXLo}$ cases (Fig. 4). The differential distribution of all the MMRd cases among the three groups (DDR$^{\gamma H2AXHi}$, DDR$^{\gamma H2AXLo}$, and DDRnot) was highly significant (chi square p < 0.001)

The majority of patients were treated with post-surgery adjuvant chemotherapy, which included the DNA-damaging agent 5-Fluorouracyl (5FU) (154 patients) or combinations containing 5FU (leucovorin, 5FU and oxaliplatin; FOLFOX) (25 patients). Cancer recurrence occurred in 49 of these patients. The event-free survival probability of treated DDR + and DDRnot patients was statistically not different (p-value = 0.11) (Fig.Supplementary fig. S2F). Survival of untreated patients was not analyzed.

Discussion

A third of untreated colorectal cancers (COAD) at first diagnosis show focal evidence of activation of the DNA damage machinery, as shown by colocalization of multiple independent members of the repair complex. This phenotype represents an all-or-nothing feature and not a continuum among all cases. There are no distinct histopathologic, immunologic or clinicopathologic characteristics associated with this phenotype. The outcome after the post-surgery adjuvant chemotherapy treatment of these patients with DNA-damaging drugs (5FU etc.) was not different from the rest of the patients.

This phenotype was previously reported.

Oka et al. [31] reported a gradual increase in phosphorylation of ATM, H2AX and CKH2 from normal tissue, to adenoma and to COAD in 55 patients. Oxidative, replicative and mechanical stress was the suggested cause for DDR activation in non-apoptotic cells during cancer progression. Takabayashi et al. [32] found variable co-localization of γH2AX and 53BP1 in foci in 33.9% of 56 COAD, but did not elaborated about

the cause for this phenotype. Neither manuscript provided additional phenotypic or clinicopathologic data.

DDR is a restraining checkpoint on the road to cancer, as shown in four seminal papers dating 2005 and 2006 [7, 9, 11, 33]. Since an intact ATM response is required to assembly the DDR foci [9], we assume that in the DDR + group of COAD this arm of the response must be preserved, at least in an hemizygous state, as shown by the co-assembly of several of its members (γH2AX, pATM, pCHK2, 53BP11, pNBS1) ([31, 32] and this report). Indeed, ATM inactivating mutations have been found in only 7% of non-hypermutated, TP53 wild-type COADs [34].

Immunostaining on serial sections suggested a coexpression of proliferation and DDR markers [7, 33] and a dissociation in cancer between response scores and senescence [9].

Our high-plex IF data show a different scenario, made of a majority of COADs which are non-informative because of lack of DDR phenotype, apoptosis or senescence, and a substantial minority which displays a DDR phenotype which is at variance to the expectation.

These latter COAD cases contain DDR + glands which are non-proliferating, negative for senescence markers (p16, p21) and in which TP53 expression is irrelevant or diminished, compared to the surrounding tissue.

To this point, a more nuanced picture has emerged over the years, adding up to the findings published in 2005 and 2006.

Individual lesions of members of the DDR (ATM, ATR, CHK1, CHK2), in the presence of replication stress response, cause an incomplete escape from quiescence or apoptosis, allowing the cells to resume growth [35]. While γH2AX and pCHK2 remain active, p21/CDKN1A is post-transcriptionally downregulated via MDM2 in a MEK/ERK dependent, TP53 independent fashion [35].

This phenotype, which we describe for the first time in vivo, may allow the progression to cancer and may be retained in a third of full-blown COAD, while is being changed to a DDRnot phenotype in the others. Worth noting that cancers stem cells from continuously growing cell lines seems to show this behavior [35].

What is the cause of the phenotype? DNA replication stress [2] seems the unifying motif for the many aspects which may lead to γH2AX phosphorylation (the broadest and least specific hallmark of a constellation of mechanisms) and the loading of DDR molecules into the nucleus. DNA replication stress may lead to multiple abnormalities which include replication forks stalling, DNA breaks, chromosome missegregation, immune activation and cell cycle arrest. This latter aspect which we observed in our COAD cases may derive by multiple actors such as availability of metabolic constituent of the DNA replication [29, 36] or out-titration of essential cofactors [2]. Furthermore, asynchrony of the causes and

effects of the replication stress have been observed and may cause the microheterogeneity of the phenotype we have described.

We did not look for genetic lesions underlying the DDR phenotype; worth noting that no single genetic abnormality alone may identify a third of COADs. It is however conceivable that multiple separate lesions may contribute to cause DNA replication stress [2, 12].

Consistent with published data, 12% of our COAD cohort had a MSI-H phenotype and in the majority of the cases (87%) MLH1 was absent. The MSI-H phenotype was equally represented in the DDR + and DDRnot groups, however none of the MSI-H DDR$^{\gamma H2AXHi}$ was MLH1 protein negative. The presence of alternative MSI gene abnormalities was over-represented in the DDR$^{\gamma H2AXLo}$ phenotype.

MSH6, but not MLH1 neither PMS2, is part of the DDR complex loaded on the replication stress sites [8]. Besides this, there is little evidence for a role of the MLH1 in the DDR complex, thus the suggestion of the requirement of an intact MLH1 to express the full DDR phenotype remains unexplained.

We reasoned that the DDR + tumor areas might be the site of punctuated karyotipic evolution [37] in COAD with ongoing copy number variation. To prove that, we used a FISH probe for the most represented chromosome alteration in COAD, chr20qAmp [27] and we targeted the analysis to the DDR + phenotype-containing areas (and controls).

For the first time, ch20qAmp has been assessed by multi-region FISH analysis, which provides spatial dimension to the chromosome status, as well as a distinction between subchromosomal amplification and polysomy. In only one case of the tested, 20qAmp was confirmed in a subclone, the remaining being polysomic.

Different high-throughput technologies (ArrayCGH, SNP Microarray, NGS) have been used to study chromosomal copy number aberrations in colorectal cancer and chromosome 20q gain/amplification is the recurrent anomaly in microsatellite stable subtype. FISH, a low- throughput molecular cytogenetic test developed in the 1980s, remains the only test capable of discriminating between amplification and polysomy.

Unexpectedly, we found uniformity in chr20q asset across all areas of all tumors, without any evidence of variegation or clonal evolution, at least with the probes we used. These data are confirmed by Bollen et al. [37], where the chr20qAmp seeding core for a progeny of organoids did not modify the chr20q status over multiple rounds of divisions and punctuated karyotypic evolution of other segments. It is possible that clonal aneuploidies, such as chr20q, represent the fixed outcome of a selection process [38] which is not subject to further evolution. Private subclonal copy number variations or rearrangements in a fully established tumor may require single cell in-situ genomics [39] in order to investigate the occurrence of punctuated evolution at DDR + sites.

Worth considering that both in the senescence model [40] and in the experiments showing a mitosis-associated γH2AX increase [26], no other DDR proteins are colocalized.

Evidence of a type I interferon response or accumulation of intratumoral lymphocytes remained elusive, as well as linking a DDR phenotype with response to DNA damaging therapy. Regarding this latter aspect, personalized therapy and basket innovative trials have produced remarkable therapeutic advances in subsets of cancers otherwise chemotherapy-resistant [41], thus the evidence of a substantial number of potentially treatable COAD cases [3, 14, 15] should prompt additional translational work.

Our work shows a robust assay for this subpopulation of cancers, which can be determined on routinely processed material.

This study has limitations. An IHC assay on routinely processed material or acquisition of bidimensional IF images at relatively low numerical objective aperture (NA 0.75) do not reach the resolution required to resolve the subcellular aggregates of DDR members, because of the combination of optical constraints, antigen availability and image quality (e.g. zeta stacks, confocal images). The analytical balance of this work is tilted toward a robust assay, which can be applied in routine diagnostics, at expense of greater analytical sensitivity and a detailed, more complex set of data.

Another limitation is the inability to reproduce data obtained with the same or equivalent reagents, however this falls in the problematic of science reproducibility, particularly evident with antibody based assays [42].

Lastly, assessing the DDR within a complete multi-omics picture would require not only single cell genomics and epigenomics but single cell proteomics, to address post-translational changes of labile regulatory proteins such as the ones involved in DNA damage response.

The identification of a novel COAD subgroup, based on a post-transcriptional DDR + phenotype, detectable by routine diagnostics methods, is of biological interest and may be susceptible of therapeutic targeting if the molecular mechanism is further explored in detail.

Declarations

Competing interests: the authors declare no conflicts of interest, financial or otherwise.

Acknowledgements: We wish to thank all the lab technicians for the excellent skills and support.

Ethics Approval: The study has been approved by the Institutional Review Board Comitato Etico Brianza, N. 3204, “High-dimensional single cell classification of pathology (HDSSCP)”, October 2019. Patients consent was obtained or waived according to article 89 of the EU general data protection regulation 2016/679 (GDPR) and decree N. 515, 12/19/2018 of the Italian Privacy Authority.

Author's contribution: SM: study design, case selection, data curation; MMB: image analysis, data curation, writing; NV: genetic analysis, data curation; GiuC: statistical analysis, data curation; LF: data generation; FM: data generation; NZ: case selection, histopathology; MT: case selection, clinical follow-up; MF: image analysis, data curation; SG: statistical analysis, data curation; GB: study design, case selection, histopathology, data curation; GioC: study design, data curation, writing.

Grant numbers and source of support: This work was supported by the Departmental University of Milano-Bicocca funds and by Regione Lombardia POR FESR 2014–2020, Call HUB Ricerca ed Innovazione: ImmunHUB to Giorgio Cattoretti; by the Italian Ministry of Health with Ricerca Corrente and 5x1000 funds to Laura Furia and Mario Faretta.

Supplementary data and data availability:

Supplementary data are available at Bicocca Open Archive Research Data (BOARD).

Mauro, Stefania; Bolognesi, Maddalena M; Villa, Nicoletta; Capitoli, Giulia; Laura Furia; Francesco Mascadri; Zucchini, Nicola; Totis, Mauro; Faretta, Mario; Galimberti, Stefania; Bovo, Giorgio; Cattoretti, Giorgio (2022), "A DNA damage response-like phenotype defines a novel subset of colon cancer. Supplementary data.", Mendeley Data, V1, doi: 10.17632/r9927r57bw.1

Additional raw data are available upon reasonable request for academic use.

References

1. Vogelstein B, Kinzler KW. The multistep nature of cancer. Trends Genet 1993; 9:138–41.
2. Saxena S, Zou L. Hallmarks of DNA replication stress. Molecular cell 2022; 82:2298–314.
3. Chabanon RM, Rouanne M, Lord CJ, Soria J-C, Pasero P, Postel-Vinay S. Targeting the DNA damage response in immuno-oncology: developments and opportunities. Nature Reviews Cancer 2021:1–17.
4. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. Nature 2009; 461:1071–8.
5. Abraham RT. Checkpoint signalling: focusing on 53BP1. Nature Cell Biology 2002; 4:E277-9.
6. Ward I, Chen J. Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. The Journal of biological chemistry 2001; 276:47759–62.
7. Gorgoulis V, Vassiliou L, Karakaidos P, Zacharatos P, Kotsinas A, Liloglou T, *et al.* Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. Nature 2005; 434:907–13.
8. Salifou K, Burnard C, Basavarajaiah P, Grasso G, Helsmoortel M, Mac V, *et al.* Chromatin-associated MRN complex protects highly transcribing genes from genomic instability. Sci Adv 2021; 7.
9. Bartkova J, Rezaei N, Liontos M, Karakaidos P, Kletsas D, Issaeva N, *et al.* Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. Nature 2006;

444:633–7.

10. Chen X, Ko LJ, Jayaraman L, Prives C. p53 levels, functional domains, and DNA damage determine the extent of the apoptotic response of tumor cells. Genes Dev 1996; 10:2438–51.
11. Di Micco R, Fumagalli M, Cicalese A, Piccinin S, Gasparini P, Luise C, *et al.* Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. Nature 2006; 444:638–42.
12. Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, *et al.* Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. Cell reports 2018; 23:239 – 54.e6.
13. Pearl LH, Schierz AC, Ward SE, Al-Lazikani B, Pearl FMG. Therapeutic opportunities within the DNA damage response. Nature Reviews Cancer 2015; 15:166–80.
14. Lee MS, Kopetz S. Are Homologous Recombination Deficiency Mutations Relevant in Colorectal Cancer? J Natl Cancer Inst 2022; 114:176–8.
15. Reilly NM, Novara L, Di Nicolantonio F, Bardelli A. Exploiting DNA repair defects in colorectal cancer. Molecular oncology 2019; 13:681–700.
16. Bolognesi MM, Mascadri F, Furia L, Faretta M, Bosisio FM, Cattoretti G. Antibodies validated for routinely processed tissues stain frozen sections unpredictably. Biotechniques 2021; 70:137–48.
17. Bolognesi MM, Manzoni M, Scalia CR, Zannella S, Bosisio FM, Faretta M, *et al.* Multiplex Staining by Sequential Immunostaining and Antibody Removal on Routine Tissue Sections. J Histochem Cytochem 2017; 65:431–44.
18. Cattoretti G, Bosisio F, Marcelis L, Bolognesi MM. Multiple Iterative Labeling by Antibody Neodeposition (MILAN). https://protocolexchange.researchsquare.com/article/nprot-7017/v5. 2019.
19. Mascadri F, Ciccimarra R, Bolognesi MM, Stellari F, Ravanetti F, Cattoretti G. Background-free Detection of Mouse Antibodies on Mouse Tissue by Anti-isotype Secondary Antibodies. J Histochem Cytochem 2021; 69:535–41.
20. Furia L, Pelicci PG, Faretta M. A computational platform for robotized fluorescence microscopy (I): high-content image-based cell-cycle analysis. Cytometry Part A: the journal of the International Society for Analytical Cytology 2013; 83:333–43.
21. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nature methods 2021; 18:100–6.
22. Waisman A, Norris AM, Elias Costa M, Kopinke D. Automatic and unbiased segmentation and quantification of myofibers in skeletal muscle. Sci Rep 2021; 11:11793.
23. Cole AJ, Dwight T, Gill AJ, Dickson K-A, Zhu Y, Clarkson A, *et al.* Assessing mutant p53 in primary high-grade serous ovarian cancer using immunohistochemistry and massively parallel sequencing. Scientific reports 2016; 6:26191.
24. Ruifrok AC, Katz RL, Johnston DA. Comparison of quantification of histochemical staining by hue-saturation-intensity (HSI) transformation and color-deconvolution. Applied immunohistochemistry &

molecular morphology: AIMM / official publication of the Society for Applied Immunohistochemistry 2003; 11:85–91.

25. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, *et al.* Fiji: an open-source platform for biological-image analysis. Nature methods 2012; 9:676–82.
26. McManus KJ, Hendzel MJ. ATM-dependent DNA damage-independent mitotic phosphorylation of H2AX in normally growing mammalian cells. Molecular biology of the cell 2005; 16:5013–25.
27. Zhang B, Yao K, Zhou E, Zhang L, Cheng C. Chr20q Amplification Defines a Distinct Molecular Subtype of Microsatellite Stable Colorectal Cancer. Cancer Research 2021; 81:1977–87.
28. Hong C, Schubert M, Tijhuis AE, Requesens M, Roorda M, van den Brink A, *et al.* cGAS-STING drives the IL-6-dependent survival of chromosomally instable cancers. Nature 2022.
29. McGrail DJ, Pilié PG, Dai H, Lam TNA, Liang Y, Voorwerk L, *et al.* Replication stress response defects are associated with response to immune checkpoint blockade in nonhypermutated cancers. Science translational medicine 2021; 13:eabe6201.
30. Simon A, Fah J, Haller O, Staeheli P. Interferon-regulated Mx genes are not responsive to interleukin-1, tumor necrosis factor, and other cytokines. J Virol 1991; 65:968–71.
31. Oka K, Tanaka T, Enoki T, Yoshimura K, Ohshima M, Kubo M, *et al.* DNA damage signaling is activated during cancer progression in human colorectal carcinoma. Cancer Biol Ther 2010; 9:246–52.
32. Takabayashi H, Wakai T, Ajioka Y, Korita PV, Yamaguchi N. Alteration of the DNA damage response in colorectal tumor progression. Human Pathology 2013; 44:1038–46.
33. Bartkova J, Horejsi Z, Koed K, Kramer A, Tort F, Zieger K, *et al.* DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 2005; 434:864–70.
34. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012; 487:330–7.
35. McGrail DJ, Lin CC-J, Dai H, Mo W, Li Y, Stephan C, *et al.* Defective Replication Stress Response Is Inherently Linked to the Cancer Stem Cell Phenotype. Cell reports 2018; 23:2095–106.
36. Gemble S, Wardenaar R, Keuper K, Srivastava N, Nano M, Macé A-S, *et al.* Genetic instability from a single S phase after whole-genome duplication. Nature 2022; 604:146–51.
37. Bollen Y, Stelloo E, van Leenen P, van den Bos M, Ponsioen B, Lu B, *et al.* Reconstructing single-cell karyotype alterations in colorectal cancer identifies punctuated and gradual diversification patterns. Nature Genetics 2021; 53:1187–95.
38. Trakala M, Aggarwal M, Sniffen C, Zasadil L, Carroll A, Ma D, *et al.* Clonal selection of stable aneuploidies in progenitor cells drives high-prevalence tumorigenesis. Genes Dev 2021; 35:1079–92.
39. Zhao T, Chiang ZD, Morriss JW, LaFave LM, Murray EM, Del Priore I, *et al.* Spatial genomics enables multi-modal study of clonal heterogeneity in tissues. Nature 2022; 601:85–91.
40. Pospelova TV, Demidenko ZN, Bukreeva EI, Pospelov VA, Gudkov AV, Blagosklonny MV. Pseudo-DNA damage response in senescent cells. Cell cycle (Georgetown, Tex) 2009; 8:4112–8.

41. Domchek SM, Aghajanian C, Shapira-Frommer R, Schmutzler RK, Audeh MW, Friedlander M, *et al.* Efficacy and safety of olaparib monotherapy in germline BRCA1/2 mutation carriers with advanced ovarian cancer and three or more lines of prior therapy. Gynecol Oncol 2016; 140:199–203.
42. Bradbury A, Plückthun A. Reproducibility: Standardize antibodies used in research. Nature 2015; 518:27–9.

Tables

TABLE 1 Clinicopathological data.

	All cases	DDRnot	any DDR+	gH2AX^{Hi}	gH2AX^{Lo}
	246 (100%)	162 (65.9%)	83 (33.7%)	48 (19.5%)	37 (15%)
Age (mean)	72.6	72.5	72.6	72.7	72.6
Sex (M)	130 (52.8%)	85 (52.5%)	44 (53%)	27 (56.3%)	18 (48.6%)
≤50 yrs	12 (4.9%)	8 (4.9%)	4 (4.8%)	2 (4.2%)	2 (5.4%)
Right colon	73 (29.7%)	50 (30.9%)	22 (26.5%)	12 (25%)	10 (27%)
T (≥4)	45 (18.4%)	34 (21%)	10 (12%)	7 (14.6%)	3 (8.1%)
N (=0)	129 (52.4%)	83 (51.2%)	45 (54.2%)	25 (52.1%)	21 (56.8%)
High grade (G3)	17 (6.9%)	11 (6.8%)	5 (6%)	3 (6.3%)	2 (5.4%)
Mucinous histology	45 (18.3%)	28 (17.3%)	17 (20.5%)	7 (14.6%)	11 (29.7%)
MMRd	30 (12.2%)	21 (13%)	8 (9.6%)	2 (4.2%)	6 (16.2%)
MLH1 def.*	26 (86.7%)	20 (95.2%)	5 (62.5%)	0 (0%)	5 (83.3%)
TP53 wt	147 (59.8%)	101 (62.3%)	46 (55.4%)	24 (50%)	22 (59.5%)

*chisquare p = 0.022

Table 2 Chromosome 20q status

UPN	Chr20q	Ratio 20q/cen	Mean 20q copies	MMRstatus	DDR+
UPN36	Normal CN	*1.1*	*1.7*	proficient	DDRnot
UPN37	Normal CN	*1.1*	*1.8*	proficient	DDR+
UPN38	Normal CN	*1.0*	*1.8*	proficient	DDRnot
UPN39	Normal CN	*1.1*	*1.9*	proficient	DDR+
UPN40	Normal CN	*1.0*	*1.8*	proficient	DDR+
UPN41	Normal CN (Monosomy 20%)	*1.0*	*1.8*	proficient	DDRnot
UPN42	Monosomy (34%)	*1.2*	*1.6*	proficient	DDR+
UPN43	Polysomy [3-5 copies] (75%)	*1.0*	*3.3*	proficient	DDRnot
UPN44	Normal CN	*1.0*	*1.7*	deficient	DDRnot
UPN45	Polysomy [3-7 copies] (35%); amp (51%)	*1.8*	*4.3*	proficient	DDRnot
UPN46	Polysomy [4-7 copies] (100%)	*1.0*	*5.4*	proficient	DDR+
UPN48	Normal CN	*1.0*	*2.2*	proficient	DDR+
UPN49	Polysomy [3-6 copies] (57%)	*1.1*	*3.0*	proficient	DDRnot
UPN50	Polysomy [3-6 copies] (79%)	*1.1*	*3.4*	proficient	DDR+
UPN52	Normal CN	*1.1*	*1.9*	proficient	DDR+
UPN53	Normal CN	*1.1*	*1.7*	deficient	DDRnot
UPN54	Polysomy [3-6 copies] (77%)	*1.1*	*3.8*	proficient	DDR+
UPN55	Polysomy [3-6 copies] (74%)	*1.1*	*3.7*	proficient	DDR+
UPN56	Normal CN	*1.0*	*2.0*	proficient	DDRnot
UPN57	Normal CN	*1.0*	*2.0*	proficient	DDR+
UPN58	Normal CN	*1.0*	*2.4*	proficient	DDRnot
UPN59	Normal CN	*1.1*	*2.2*	proficient	DDR+
UPN60	Polysomy [3-4 copies] (39%)	*1.1*	*2.5*	proficient	DDRnot
UPN61	Polysomy [3-5 copies] (37%)	*1.1*	*2.5*	proficient	DDRnot
UPN62	Normal CN	*1.0*	*1.9*	proficient	DDRnot
UPN63	Normal CN	*1.0*	*2.3*	proficient	DDRnot

The chromosome 20q status by FISH is shown. UPN = unique patient number; Normal CN = normal copy number; cen = centromeric probe; MMR = mismatch mutation repair; DDR = DNA damage Repair.

Figures

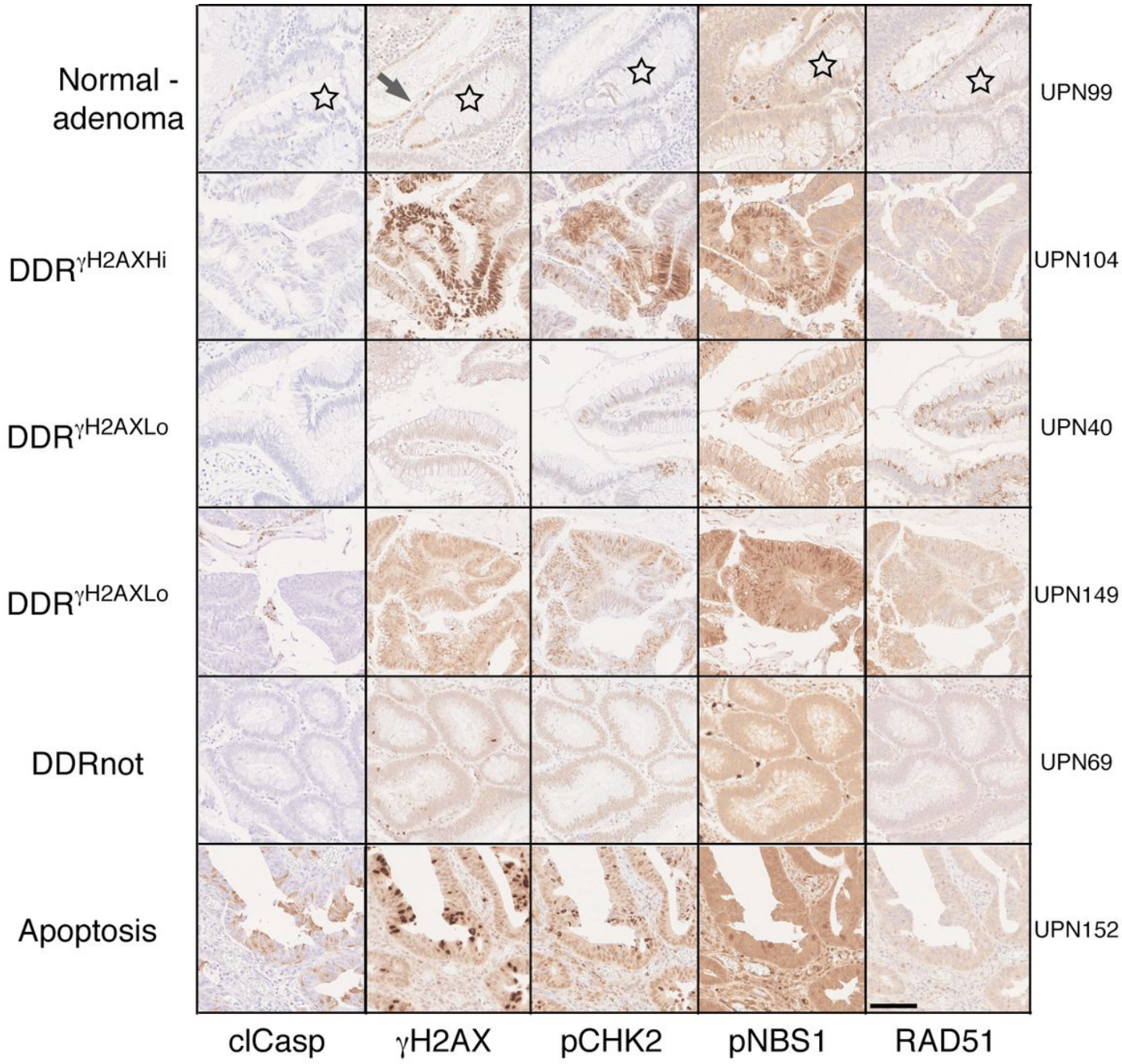

Figure 1

Expression of the DDR phenotype in six representative COAD cases.

Serial sections from six cases stained in IHC for five markers (bottom) are shown at high power, each representing one of the DDR phenotypes encountered. The star identifies a normal gland. γH2AX+ adenomatous gland is arrowed. Scale bar = 500μm. UPN = unique patient number.

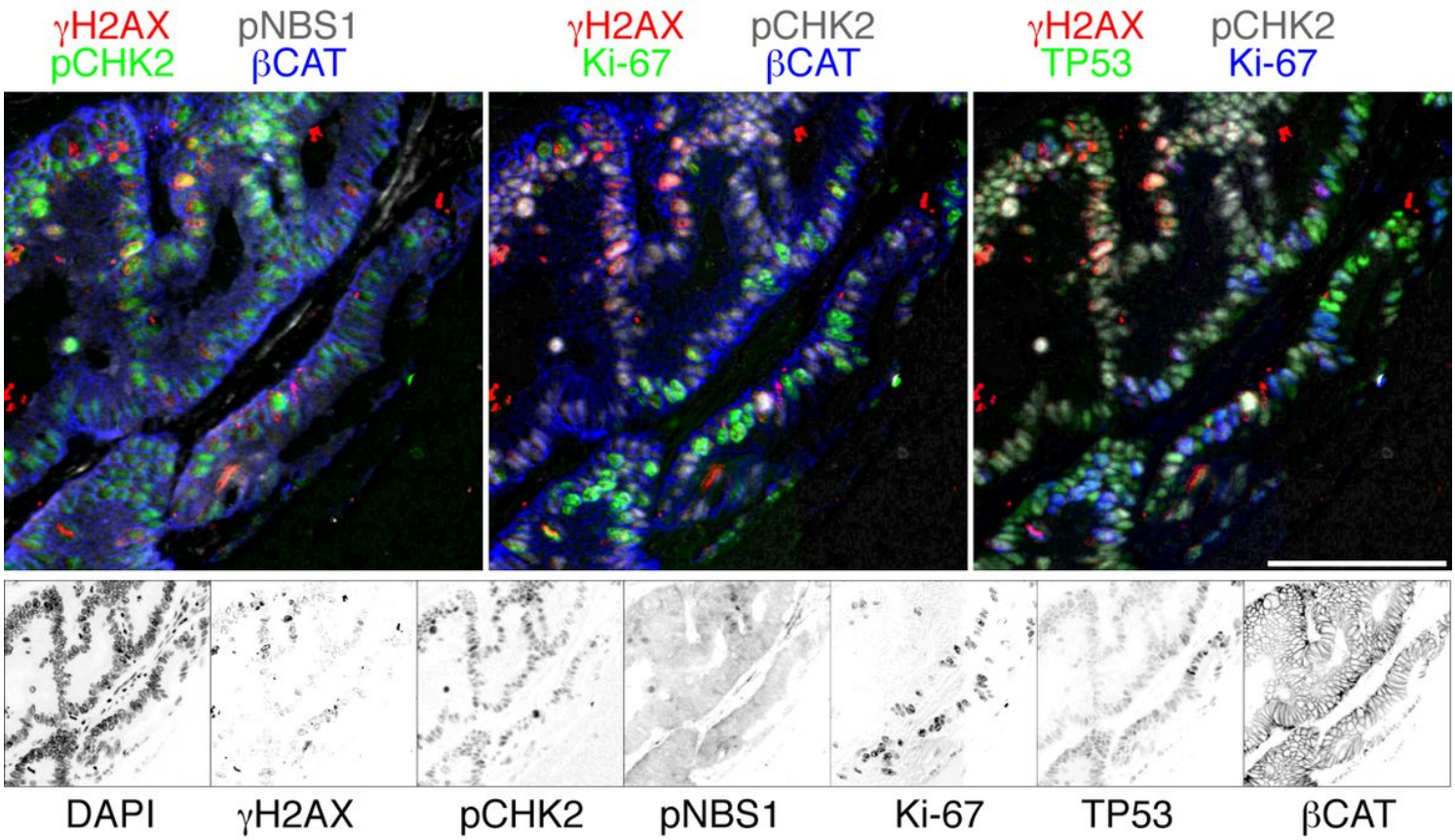

Figure 2

Multiplex phenotype of DDR$^{\gamma H2AXHi}$ COAD.

Six selected markers are shown as four color composites (top) and inverted grayscale single images (bottom). Color-coded or plain markers are shown next to each image. Note in the βCatenin+ tumor gland in the upper left part of the composite, the variegated coexpression of γH2AX, pCHK2 and pNBS1 in the absence of proliferation (KI-67) and diminished wild type TP53. Scale bar = 500μm. UPN 47.

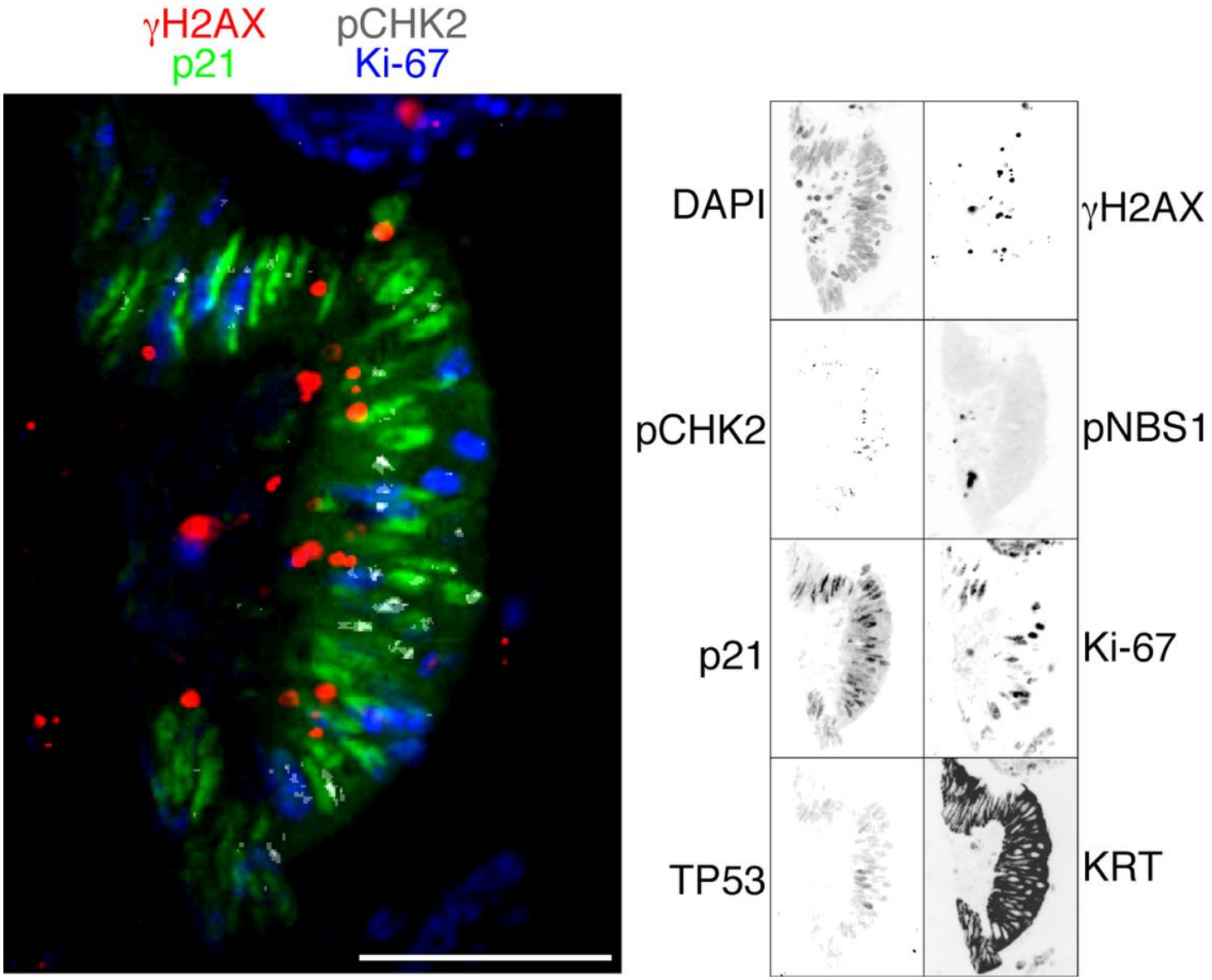

Figure 3

P21/CDKN1A expression in DDR^{not} COAD.

A single tumor gland expresses P21/CDKN1A in quiescent cells (Ki-67 neg) in the absence of γH2AX or pCHK2. Scale bar = 100µm. UPN 35.

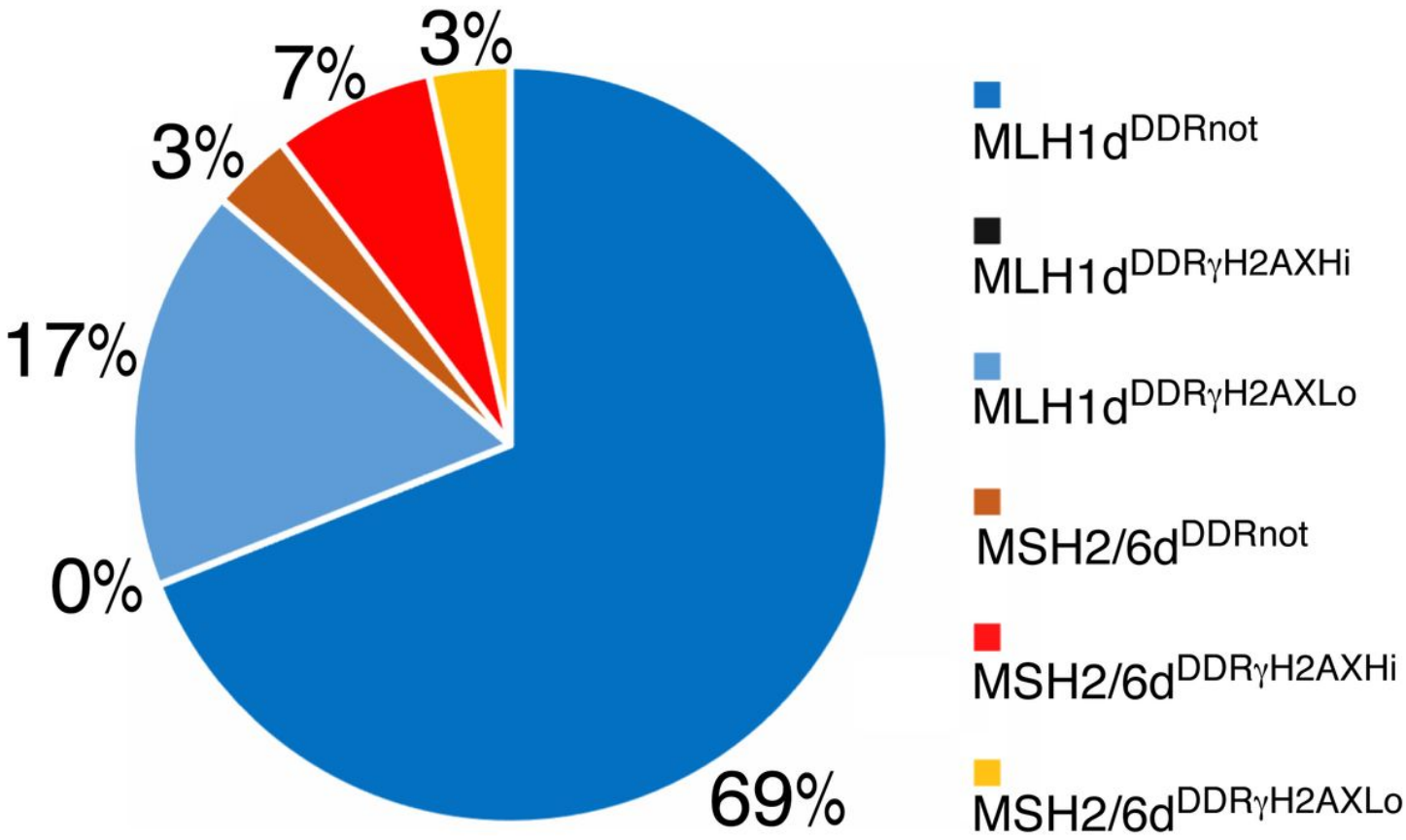

Figure 4

Distribution of MMR deficient cases among the DDR phenotypes.

The pie chart show the percentage distribution of MMR deficient cases, MLH1/PMS2 or MSH2/MSH6, among DDR^{not} , DDR+ ($DDR^{\gamma H2AXHi}$) or DDRalt ($DDR^{\gamma H2AXHLo}$) cases. Note the absence of cases with the MLH1d, DDR+ ($DDR^{\gamma H2AXHi}$) phenotype (0%).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SUPPLEMENTARYDATA.pdf