An Oncogene-Induced DNA Damage Model for Cancer Development

Thanos D. Halazonetis,¹⁺ Vassilis G. Gorgoulis,² Jiri Bartek³

Of all types of DNA damage, DNA double-strand breaks (DSBs) pose the greatest challenge to cells. One might have, therefore, anticipated that a sizable number of DNA DSBs would be incompatible with cell proliferation. Yet recent experimental findings suggest that, in both precancerous lesions and cancers, activated oncogenes induce stalling and collapse of DNA replication forks, which in turn leads to formation of DNA DSBs. This continuous formation of DNA DSBs may contribute to the genomic instability that characterizes the vast majority of human cancers. In addition, in precancerous lesions, these DNA DSBs activate p53, which, by inducing apoptosis or senescence, raises a barrier to tumor progression. Breach of this barrier by various mechanisms, most notably by p53 mutations, that impair the DNA damage response pathway allows cancers to develop. Thus, oncogene-induced DNA damage may explain two key features of cancer: genomic instability and the high frequency of p53 mutations.

In medicine, neoplasia (a Greek word meaning “new growth”) refers to new tissue composed of cells with the heritable capacity to grow beyond their normal confines. Neoplasias can be benign or malignant. The former remain localized, whereas the latter, also referred to as cancer, invade the host tissues. Some benign neoplasias are precancerous lesions and develop over time into overt cancers. Important goals of cancer research in the molecular biology era have been to understand the genetic basis of cancer and to explain its progressive nature. Here, we review a model that may contribute to better understanding of two features shared by most cancers: genomic instability and p53 mutations.

DNA DSBs in Cells Expressing Activated Oncogenes and in Human Precancerous Lesions and Cancers

Until recently, identifying DNA double-strand breaks (DSBs) with high sensitivity was not trivial. However, the identification of proteins that become recruited to large chromatin domains flanking the DNA DBSs has made it possible to visualize these breaks indirectly by immunofluorescence (1). (Additional references are listed in the Supporting Online Material, grouped by the section in this review to which they pertain.) The intracellular localization of one such protein, p53 binding protein 1 (53BP1), was examined in cancer cell lines grown in tissue culture and revealed 53BP1 nuclear foci (about 10 to 20 per cell), whose presence is indicative of DNA DSBs. No such foci were present in proliferating normal cells. Because cancer cells are typically not defective in DNA DSB repair, these results suggested a continuous cycle of DNA DSB formation and repair (1, 2). Further, in the panel of cancer cell lines examined, those cell lines with the highest number of 53BP1 foci per cell had p53 mutations, possibly reflecting a selection for p53 inactivation, because p53 induces apoptosis or senescence in response to DNA DSBs (1, 2).

The analysis of cancer cells in tissue culture is informative but may not accurately represent the state of cancer cells in human patients. Are DNA DSBs continuously being generated in human cancers? If so, at what stages of cancer development? And does DNA DSB formation precede the loss of p53 function, as would be expected if the DNA DSBs select for p53 inactivation? To address these questions, several research groups analyzed precancerous and cancerous lesions from human patients. In the precancerous lesions, before p53 mutations were acquired, 53BP1 localized at foci and histone H2AX, ataxia telangiectasia (ATM), Chk2, and p53 were phosphorylated, suggesting the presence of DNA DSBs (3, 4). In cancers, evidence for the presence of DNA DSBs, such as phosphorylated histone H2AX and 53BP1 foci, was again present, but the DNA damage checkpoint pathway was compromised, most often by p53 mutations and less often by loss of expression of various checkpoint proteins, such as ATM, 53BP1, Chk2, or p53 (2–4).

The presence of DNA damage was a feature that could distinguish precancerous lesions and cancers from normal tissues, irrespective of their proliferation rate (2–4). What could be causing the induction of DNA DSBs in the precancerous lesions and cancers? Eroded telomeres or mutations targeting genes required for genome integrity could be responsible, but neither of these characterizes all precancerous lesions (5–8).

Instead, the presence of activated oncogenes distinguishes precancerous lesions from normal tissues, and in various model systems activated oncogenes are capable of inducing both DNA DSBs and a DNA damage response (DDR). For example, activated ras induces DNA DSBs in NIH3T3 fibroblasts within a single cell cycle, and other oncogenes, including myc, cyclin E, mos, cdc25A, and E2F1, have similar effects in various cell types and in animal models (4, 9–11). In a human skin xenograft hyperplasia model, sustained delivery of growth factors also led to induction of DNA DSBs and a DDR, in the absence of telomere erosion (3). In all these models, the majority of cells overexpressing the oncogene exhibited a DDR, suggesting that a second stochastic event was not required for induction of DNA damage.

A Tumorigenesis Barrier in Human Precancerous Lesions

In various cell and animal models, activation of the DNA damage checkpoint induces p53-dependent cell cycle arrest, apoptosis, or senescence (12). The question then arises whether apoptosis and senescence are present in human precancerous lesions and, if so, whether they subside during cancer progression, when defects in the p53 pathway are acquired. Indeed, the precancerous tissues show evidence of apoptosis or senescence, and both of these are suppressed during cancer progression (3, 4, 10, 11, 13).

This general pattern is observed in multiple tumor types, but with interesting tissue-specific differences (3, 10, 13). In the lung, the normal bronchial epithelium has a very low proliferation index and negligible apoptosis; the precancerous lesions exhibit high levels of apoptosis and a high proliferation index; whereas lung carcinomas have a high proliferation index but low levels of apoptosis. In the skin, normal melanocytes exhibit a low proliferation index and negligible senescence; dysplastic nevi, the precancerous lesions, exhibit senescence and a low proliferation index (because senescent cells do not proliferate); and in melanoma there is loss of the senescent phenotype and a high proliferation index. In colon, the normal tissue has a very high proliferation index; progression to adenoma, the precancerous lesion, leads to senescence and a significant decrease in the proliferation index; whereas further progression to carcinoma is associated with escape from senescence and a high proliferation index similar to that of the normal tissue. Plotting the apoptosis or senescence indices of these three tissues (3, 10, 13) shows a sharp peak in precancerous lesions, representing a barrier that is raised specifically in that stage of cancer development and which is eroded upon further tumor progression (Fig. 1).

The concept of a tumorigenesis barrier acting in precancerous lesions contrasts with earlier models, which considered that precancerous lesions were less aggressive than cancers simply
because they had fewer oncogenic mutations (14). Instead, the oncogene-induced DNA damage model suggests that the less aggressive nature of precancerous lesions is in part due to the tumorigenesis barrier imposed by the DNA damage checkpoint. This concept is best illustrated by the example of colon adenomas, which have a lower proliferation index than that of normal colon (Fig. 1).

**The DNA Damage Checkpoint as an Important Mediator of the Tumorigenesis Barrier**

Much evidence points to p53 as a key protein imposing the tumorigenesis barrier in precancerous lesions: Apoptosis and senescence are known p53-dependent responses, escape from the tumorigenesis barrier correlates with p53 mutations in human cancers, and the transition from precancerous lesions to cancer is accelerated in p53-deficient mice (3, 4, 12, 13). In fact, in some p53-null mouse models, tumor development bypasses the senescence stage (13). The question that remains to be answered is whether, in precancerous lesions, p53 is activated by the DNA damage checkpoint pathway, as the oncogene-induced DNA damage model predicts, or by one of the other pathways that activate p53, with the most serious contender being the pathway that includes the alternative reading frame (ARF) tumor suppressor protein (15).

Perhaps the strongest evidence in favor of p53 being activated by the DNA damage checkpoint pathway in human cancers comes from analysis of clinical material. Analysis of colon and bladder precancerous lesions by immunohistochemistry showed that the senescence markers coincide with the DDR markers (10). The spectrum of genetic mutations present in cancer patients also suggests that p53 is principally activated by the DNA damage checkpoint. Activation of p53 by DNA DSBs is mediated principally by the kinases ATM and Chk2 (12). In an analysis of 518 protein kinase genes in 210 human cancers, the ATM gene ranked third in terms of mutation frequency (16). Further, heterozygous chk2 germline mutations predispose one to cancer and can occasionally be associated with a Li-Fraumeni syndrome that is indistinguishable from the one caused by p53 germline mutations (17). Lastly, expression of 53BP1 and MDC1, two DDR proteins that act upstream of Chk2 and p53, is suppressed in subsets of melanomas, breast and lung carcinomas (3). Mutations targeting arf are also frequent in human cancer, but they are typically deletions also targeting the tumor suppressor genes INK4a and INK4b (15). Point mutations targeting arf without affecting INK4a are found with about one-twentieth the frequency of point mutations targeting exclusively INK4a.

A role of the DNA damage checkpoint in mediating a tumorigenesis barrier is also supported by analysis of cultured primary cells overexpressing oncogenes. Such cells exhibit apoptosis or senescence or both, but these responses are suppressed when ATM is inhibited (10, 11). In mouse tumor models, apoptosis induced by oncogenes, such as myc, also depends on the DDR pathway, and inhibiting DDR proteins, such as Chk2, ATM, or Tip60, facilitates tumor progression (10, 11). Tumors also develop in mice deficient for atm or chk2 and in mice in which two residues in p53 that are phosphorylated by ATM and Chk2 are substituted with alanines (18).

Nevertheless, there are also observations that seem to contradict the notion that the DNA damage checkpoint is an important barrier to tumorigenesis. First, why is p53 mutated much more frequently in human cancer than any other DNA damage checkpoint gene? One explanation is that single amino acid substitutions targeting the p53 DNA binding domain can easily inactivate the protein (because of its very low melting temperature) and at the same time generate mutants with dominant-negative activity. Further, unlike inactivation of some upstream DDR proteins, such as ATM, p53 inactivation does not compromise the G2/M-phase DNA damage checkpoint or DNA DSB repair, upon which cancer cells are likely to rely to optimize their survival (12).

A second question is why knockout of the atm and p53 genes in mice leads to a shorter latency in spontaneous tumor formation (median survival age of about 55 days) than the individual knockouts (medians of about 120 and 140 days for the atm and p53 knockouts, respectively) (18). Further, why do tumors that develop in atm-null, p53-heterozygote mice tend to lose the remaining wild-type p53 allele? These observations appear to contradict the notion that p53 is activated by ATM in precancerous lesions. However, analysis of the median survival ages of mice with all combinations of atm and p53 genotypes shows that the effect of knocking out one or both alleles of p53 is much smaller in an atm-null background than in heterozygote or wild-type atm backgrounds, which is consistent with ATM activating the tumor suppressor function of p53 (18). The residual tumor suppressor function of p53 in atm-null mice can be explained in part by activation of p53 by ATR, an ATM-related kinase that, like ATM, also responds to DNA damage (12). Further, the residual tumor suppressor function of ATM in p53-null mice can be explained by ATM having multiple substrates in addition to p53, some of which, for example, NBS1 and BRCA1, are also tumor suppressors (12).

A third question relates to the role of arf as an alternate pathway activating p53 in cancer. Indeed, many studies examining the role of arf in cultured cells and mouse models support a role in oncogene-induced apoptosis and senescence (15), although there are exceptions. Interestingly, in the same mouse model, a lymphoma induced by a myc transgene, lymphomagenesis can be accelerated by knocking out either arf or atm. This suggests that both the DDR and the arf pathways have tumor suppressor functions but leaves open the question whether the tumor suppressor function of arf is p53-dependent. Studies to answer this question in mice have led to conflicting conclusions (15).

Some observations suggest that part of the tumor suppressor effect of arf is p53-independent. First, in mice, arf is a much weaker tumor suppressor than atm or p53: The median survival ages of mice null for arf, atm, and p53 are about 350, 120, and 140 days, respectively (15, 18, 19). This is not consistent with arf being the major activator of the tumor suppressor function of p53. Second, analysis of tumor formation in mice with various combinations of arf and p53 genotypes suggests that arf and p53 are acting additively and independently of each other. Specifically, the tumor suppressor function of p53 is not compromised in an arf-null background (19), unlike what was observed in an atm-null background (18). Third, arf may be exerting its tumor suppressor function at a different stage of cancer development than does p53. In a chemical carcinogen-
esis mouse skin cancer model, inactivation of arf promotes formation of large papillomas, whereas inactivation of p53 promotes progression of papillomas to carcinomas (15).

Contradicting the results presented above (19), two recent studies concluded that p53 did not have a tumor suppressor function in arf-null mice (20). In one study, the function of a tamoxifen-regulatable p53 protein was turned on or off, and the incidence of radiation-induced lymphoma was monitored. In the other study, 3-methylcholanthrene–induced tumor formation was monitored in mice with two or three wild-type p53 alleles. In both studies, p53 status affected tumor kinetics in a wild-type arf genetic background but not in an arf-null background. However, when tumor kinetics were considered in all genotypes, the results were consistent with p53 and arf acting independently and additively of each other. The status of p53 had no effect in the arf-null background because in these models the tumors develop fast in the absence of arf, before any effect of p53 can become apparent. Interestingly, in the tamoxifen-regulatable p53 model, administration of tamoxifen for a short time window a few weeks after irradiation, when precancerous cells are presumably present, affords tumor protection. Because the tamoxifen-regulatable p53 protein requires both tamoxifen and a DNA damage signal to induce apoptosis (20), this result is consistent with the precancerous cells in these mice having DNA DSBs.

**DNA Replication Stress Underlies DNA DSB Formation and Genomic Instability in Human Precancerous Lesions**

A key question that follows from the results reviewed so far relates to the mechanism by which activated oncogenes continuously induce DNA DSBs in human precancerous lesions and cancers. Most oncogenes deregulate entry into the cell cycle and do so by directly or indirectly enhancing the activities of the cyclin-dependent kinases (CDKs) that function in the G1 and S phases of the cell cycle (7, 21). In yeast, deregulation of CDK activity compromises DNA replication and leads to formation of DNA DSBs and genomic instability (22). By analogy, oncogenes could induce a state of DNA replication stress in human precancerous lesions leading to the formation of DNA DSBs (3, 4, 10, 11).

The term DNA replication stress was originally used to describe the state of DNA replication arrest that is induced by deoxynucleotide depletion and is characterized by activation of the DNA replication checkpoint pathway and inability to complete DNA replication in the absence of this checkpoint. Through a better understanding of the biology of DNA replication forks, it is now possible to define DNA replication stress as any systemic state in which there is fork collapse of DNA replication forks as an early event in DNA replication stress. This can be induced by an increase in the number of stalled DNA replication forks (some of which will inevitably collapse) mediated by agents that inhibit DNA replication or by a decrease in the stability of stalled forks (because there is some fork stalling in normal S phases) mediated by deregulation of DNA replication checkpoint proteins. Irrespective of how DNA replication stress is induced, the collapse of DNA replication forks occurs preferentially at specific chromosomal loci called common fragile sites (23). These loci are prone to formation of microdeletions and gross chromosomal rearrangements, because the pathways that are induced to complete DNA replication after fork collapse often involve recombinogenic processes and formation of DNA DSBs.

The presence of DNA replication stress in human precancerous lesions can be monitored by exploiting the fact that the recombinogenic processes that are activated after fork collapse have the potential to lead to loss of heterozygosity (LOH). Indeed, in both human precancerous lesions and in a human skin xenograft hyperplasia model there is LOH that occurs predominantly at the common fragile sites, invoking the presence of DNA replication stress (3, 4). Expression of oncogenes in nontransformed cells also leads to DNA replication stress. Such cells have prematurely terminated DNA replication forks, DNA DSBs that form specifically in S phase, stretches of single-stranded DNA, and LOH targeting preferentially the common fragile sites (10, 11).

A key prediction of the oncogene-induced DNA damage model is that genomic instability in human precancerous lesions and in cancer is induced by the oncogenes themselves and is a feature of cancer from its earliest stages. How does this prediction fit with our current understanding of the mechanisms leading to genomic instability in human cancer? In advanced cancers, multiple mechanisms, including hypoxia, inflammation, and cell-matrix detachment, may contribute to genomic instability, but in precancerous lesions fewer mechanisms are likely to be implicated.

One mechanism involves telomere erosion, which can lead to transient surges in genomic instability, when telomeres become critically short but before telomerase expression is induced (5). Telomere erosion almost certainly contributes to genomic instability in human cancer, but whether it contributes to genomic instability in human

---

**Fig. 2.** Oncogene-induced DNA damage model for cancer development and progression. Genomic instability and tumor suppression are direct outcomes of oncogene-induced DNA replication stress and are both present from the beginning of cancer development, before the transition from precancerous lesion to cancer.
precancerous lesions is less clear. Telomere erosion is common in human precancerous lesions (6), but human skin xenografts already show genomic instability at common fragile sites a few weeks after induction of hyperplasia, before any changes in telomere length become evident (3).

A second mechanism involves permanent increases in genomic instability induced by mutations in genes whose function is to preserve genomic integrity. These so-called caretaker genes include DNA repair and cell cycle checkpoint genes (7). Indeed, one form of genomic instability, referred to as microsatellite instability (MIN), is caused by mutations in DNA mismatch repair genes (8). However, most human cancers do not have MIN, but rather have chromosomal instability (CIN), and the caretaker genes, whose deregulation is responsible for CIN, have not been identified (24). p53 was considered initially a “guardian of the genome,” but p53-null mice do not have overt genomic instability (25). In experimental models, candidate caretaker genes have been identified, but none of them are frequently mutated in sporadic human cancers (25). Further, large-scale sequencing analysis of all known coding regions in breast and colorectal cancers has failed to identify other genes that could account for CIN in a large fraction of human cancers (26). Because DNA DSBs can lead to CIN, it is possible that oncogene-induced DNA damage, rather than inactivation of caretaker genes, may contribute to CIN in human cancer, a hypothesis that could also explain why, in cell fusion experiments, CIN, unlike MIN, is a dominant trait (8).

Genomic instability is considered critical for cancer development because it would be difficult otherwise for a normal cell to accumulate all the mutations necessary to become a cancer cell (27). It is still unclear whether genomic instability induced by DNA replication stress has any role in cancer progression, but, in oral precancerous lesions, LOH at the common fragile site FRA3B predicted progression to cancer much better than any other marker, including p53 mutations and LOH at the INK4a/arf locus (28).

Unanswered Questions and Future Prospects

An oncogene-induced DNA damage model for cancer development and progression can help explain many features of cancer, including its tendency to progress, the presence of genomic instability, and the tumor suppressor role of DNA DSB checkpoint proteins such as p53 (Fig. 2). Of course, some aspects of the model need to be better defined. What is the mechanism by which oncogenes induce DNA replication stress? Under what conditions do oncogenes induce DNA replication stress, and is the strength of the oncogenic signal a critical factor, as suggested (29)? What is the role of ARF in human cancer development? What fraction of genomic instability is due to DNA replication stress and what fraction to other factors, such as telomere erosion? Does this model apply to most cancers, or are there tumor types, perhaps those characterized by low levels of genomic instability, in which this model does not apply? Lastly, does this model suggest new ways to diagnose and treat cancer?

With regard to the last question, one can envision new diagnostic methods for precancerous lesions. LOH at the common fragile site FRA3B shows promise in identifying high-risk precancerous lesions (28). It may also be possible to establish immunohistochemistry assays for DNA damage responses in a clinical setting on the basis of the assays currently available for research (3, 4, 13). An interesting possibility would be to develop cancer therapies, capitalizing on the presence of DNA replication stress specifically in cancer cells. DNA replication needs to resume after replication fork collapse. We suspect that cancer cells would not fare well if resumption of DNA replication could be inhibited.

References and Notes


30. We apologize for not being able to include all relevant concepts and references. We thank all our collaborators for their contributions to the development of the oncogene-induced DNA damage model. The authors are supported by grants from the Swiss National Foundation and the U.S. National Cancer Institute (T.D.H.); the Greek General Secretariat for Research and Technology Program to Enhance Research (V.G.G.); and the Danish Cancer Society, the Danish National Research Foundation, the Czech Ministry of Education (MSM6198959216), and the European Commission “Active p53” and “Mutant p53” Integrated Projects (J.B.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5868/1352/DC1

References

10.1126/science.1140735