DNA repair, genome stability and cancer: a historical perspective

Article (Accepted Version)

DNA repair, genome stability and cancer – a historical perspective.

Penny A. Jeggo, Laurence H. Pearl and Antony M. Carr.

Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Brighton, BN1 9RQ, UK

Preface

The multi-step process of cancer progresses over many years. The prevention of mutations by DNA repair pathways led to an early appreciation of a role for repair in cancer avoidance. However, the broader role for the DNA damage responses (DDR) emerged more slowly. We reflect on how our understanding of the steps leading to cancer unravelled, focusing on the role of the DDR. We consider how our current knowledge can be exploited for cancer therapy.

Introduction.

Carcinogenesis has long been recognised as a state of uncontrolled growth of our own cells. The earliest models proposed the notion of a mutational event, even before Watson and Crick's seminal discovery of the structure of DNA. By the 1980s, the initiation step of carcinogenesis was understood to necessitate the generation of mutations, with the concept of environmental mutagens and failed DNA repair being central to many models. In contrast, an understanding of the basis underlying tumour progression or outgrowth unfolded relatively slowly and an appreciation of the critical role played by the DNA damage responses (DDR) took even longer to solidify. Indeed, even as the twenty first century began, DNA repair remained a relatively insignificant component of the broad field of cancer research. It is now appreciated that tumour progression necessitates the down-regulation of damage surveillance mechanisms and an increase in genetic/epigenetic instability to achieve uncontrolled proliferation and the adaptability associated with aggressive tumours. In this review, we describe the early concepts of mutation and cancer that predate knowledge of the structure of DNA, and how the links between mutagenesis and carcinogenesis were established. We go on to discuss the early studies of viral oncogenes and the insights they provided into carcinogenesis, leading to the much more recent but critical understanding that oncogene induced replicative stress promotes genomic instability. We provide a perspective for how the notions of tumour initiation and progression emerged, describing the key concept that tumour progression is inexorably linked to disruption of the DNA damage response (Figure 1). Finally, we consider the ironic conundrum that, while targeting the DDR can provide treatment strategies, it is the mis-regulation of the DDR that is often the route by which tumour cells evade therapy.

The emerging notion that mutagenesis underlies carcinogenesis

In the early 20th century, long before the structure of DNA was defined, Boveri proposed that a cancer cell was a changed normal cell and advanced the theory that tiny microscopic bodies called chromosomes were abnormally distributed in cancer[1]. As the notion of hereditary units or "determinants" and later "genes" emerged alongside their relationship to chromosomes, the idea that permanent changes to these "genes" (defined as mutations) could underlie heritable biological phenotypes became conceptualised. From there, it did not require a great leap

to appreciate that such mutations might underlie the origin of the biological variation observed in cancer (reviewed in[2]). In the 1930s it was recognised that a normal human cell has 46 chromosomes but in cancers they are abnormally distributed and frequently in excess of 46. Paradoxically, non-tumorous cells and plants with an asymmetric or unbalanced chromosome distribution grew less vigorously than normal cells, whereas cancer cells were characterised by an enhanced growth capacity. Work with Drosophila had revealed that chromosome aberrations correlated with the formation of genetic variants and, in 1927, Muller demonstrated that exposure of Drosophila to X-rays induced "transmutation" of a gene, causing both visible chromosome aberrations and heritable phenotypic variations[3, 4].

Intriguingly, in 1775, Percival Pott made the seminal observation that there was a high incidence of scrotal cancer in boys who assisted chimney sweeps and linked this finding to exposure to soot[5, 6, 7]. This represented the first evidence for a work-related and environmental cause of cancer. By 1955, shortly following the discovery of the DNA structure, it was well appreciated that exposure to chemical mutagens could enhance mutation rates, with a correlation between mutagenesis and carcinogenesis being hypothesised though certainly not consolidated[2]. Remarkably, the suggestion that there could be cancer susceptibility genes was also proposed[2].

The link between mutagens, carcinogens and DNA.

With the structure of DNA defined in 1953[8, 9] and the understanding that genetic mutation represented a change in the chemical structure of the DNA molecule, the first clear connections between the processes of mutation and carcinogenesis were made by Phil Lawley, working at the Chester Beatty Research Institute (now the Institute of Cancer Research). Working with Peter Brooks, he showed that many classic alkylating agents worked by reacting directly with DNA to form stable chemical adducts that could disrupt the template function of the DNA molecule[10, 11]. This led directly to their seminal observation that the carcinogenicity of polycyclic aromatic hydrocarbons – the likely causative agents of scrotal cancer in those chimney sweeps and also the classic carcinogenic components of tobacco smoke - was directly correlated with their ability to form DNA adducts, providing an unambiguous link between the initiation of cancer and chemical changes to DNA[12]. The significance of these early findings and functional link between mutagenesis and carcinogenesis is demonstrated by the development and later adoption of tests classifying carcinogens on the basis of this relationship[13].

Lawley and Brooks were also amongst the first to document biological repair processes for chemical and environmental damage to DNA[10], a theme very actively adopted by many laboratories[14-19] (see also[20] for a review). Over the next thirty years a plethora of repair pathways for chemical lesions of DNA, primarily identified in bacteria, was progressively revealed. Subsequently, human homologues for many of these repair enzymes and pathways were identified. For the most part, these systems - or possible defects in them - were not associated with the initiation or the progression of cancer in any significant way. There was, however, an emerging recognition of the role that DNA repair mechanisms might play in mediating resistance to alkylating agents used for cancer chemotherapy[21].

Insight from DNA repair disorders.

A significant exception to this picture was the seminal observation by Jim Cleaver in 1969 who was studying the rare (1:250,000) autosomal recessive cancer predisposition syndrome, xeroderma pigmentosum (XP). XP patients have a 1000 fold increased chance of skin cancer, but display almost normal levels of tumour presentation in other tissues[22]. Cleaver found that cells from XP patients were defective in the ability to repair DNA damage caused by ultraviolet (UV) light[23]. The DNA repair defects in most (though not all) XP cells were subsequently shown to result from mutations in components of the human nucleotide excision repair (NER) system[22, 23]. Critically, this process is responsible for the removal of helix-distorting UV-photodimers from DNA, explaining the highly specific skin cancer predisposition of XP patients.

A second clear example of a defective DNA repair pathway being responsible for cancer initiation emerged in the early 1990s: patients with Lynch Syndrome (aka hereditary non-polyposis colorectal cancer) - a familial pattern of colorectal cancer characterised by microsatellite repeat instability – were found to carry mutations in the human homologues of the bacterial mismatch repair (MMR) proteins MutS and MutL[24-27]. Both of these inherited diseases reinforced a model for cancer initiation in which random unrepaired point mutations eventually result in an alteration to the coding sequence of a key oncogene or tumour suppressor, initiating the first step in cell proliferation and enabling a subsequent cascade of mutagenic events in these precancerous cells.

A prediction arising from these studies of patients with hereditary predisposition to cancer was that mutations in DNA repair genes might frequently arise in cancer cells. As will be discussed below, this has certainly proved to be the case. However, the early studies were

carried out when there was not a comprehensive understanding of DNA repair pathways and DNA damage responses in humans and when sequencing technology was significantly less sophisticated, and thus the attempts to address this prediction were not very revealing. In these early studies, polymorphisms and tumour-associated mutations in DNA base excision repair enzymes such as 8-oxoguanine DNA glycosylase (Ogg1) and APE1 and in components of the downstream generic XRCC1-based part of base-excision repair were identified in some tumour cells[28, 29]. However the penetrance of such polymorphisms is weak and the clinical relevance of these to the overall cancer burden was unclear[30]. Subsequently several complex conditions in which cancer predisposition is a feature, such as Bloom's and Werner's syndromes, and Fanconi anaemia, have been shown to arise from genetic defects in DNA repair systems, as have subsets of familial breast, ovarian, prostate and pancreatic cancers[31-36].

Contribution from studies of radiation exposure.

That X-ray exposure confers an increased risk of malignant disease, including leukaemia and skin cancers, became accepted within a few decades after the discovery of X-rays in 1895. However, radiation studies were disappointing when it came to gaining mechanistic insight into the aetiology of cancer. Nonetheless, reports by the International Committee on Radiological Protection (ICRP) and Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) provide an invaluable source for rationalising the emerging concepts (e.g.[37]). In particular, the analysis of the atomic bomb survivors provided a wealth of epidemiological data, such as revealing that there can be a long induction period prior to the onset of cancer, and in UNSCEAR 1958 it was concluded that radiation-induced mutations are biologically relevant for carcinogenesis[37]. However, the relationship between chromosome aberrations/rearrangements (which X-rays avidly induce) and point mutational changes (which X-rays inefficiently induce) remained puzzling. A further important concept emerged from these early radiation studies: the carcinogenic effect of radiation does not correlate with its cell-lethal effects. It was observed that, while cell killing consistently increased as the linear energy transfer (LET – a measure of radiation quality) increased because of the enhanced complexity of the DNA damage, the frequency of cell transformation, a correlate for carcinogenicity, did not follow the same pattern: it increased initially but peaked and decreased at very high LET.

These early studies evaluating the response to radiation damage raised an additional dilemma: cells from ataxia telangiectasia (A-T) patients, a cancer prone and profoundly radiosensitive human disorder, were not hyper-mutated by X-rays and, despite their marked radiosensitivity, A-T cells did not display an obvious defect in repairing X-ray induced DNA damage. The cause of this apparent paradox only became clear after the discovery that the gene defective in A–T patients, ATM, encodes a protein kinase that triggers a signalling cascade that regulates cell cycle arrest and cell death responses rather than a DNA repair enzyme. (Nonetheless, ATM signalling can indirectly influence DNA repair processes). This important distinction between signalling responses and direct DNA repair has proved to be critical in the context of cancer avoidance. Indeed, the wider response to DNA damage (known as the DNA Damage Response; DDR) is now usefully sub-divided into damage response signalling and direct DNA repair. Significantly, the DDR signalling response frequently has a greater impact on genomic stability in response to DNA damage in contrast to DNA repair pathways, which more dramatically influence survival.

Battles between competing models and research fields.

The concept that cancer involved at least one genetic mutation was well accepted by the 1970s. However, the notion that oncogenesis is a multistage process was proposed by Berenblum and Shubik as early as 1948, based on studies showing that tumour cells induced by carcinogen treatment could remain in a latent stage until outgrowth was promoted by subsequent treatment[38]. From 1970 through to the turn of the century, a range of studies including epidemiological analysis of atomic bomb survivors, studies in mice and cell culture models of transformation all provided strong evidence that cancer was a multistage process[39]. It was understood that cancer incidence increased dramatically with age and that exposure to ionising radiation brought forward the age of onset of most cancers, but still involved a marked lag period. Two contrasting models arose to explain these observations: cancer involved a mutagenic initiation step, following by a long period of outgrowth, or cancer was a multistage process, involving multiple mutational hits[40, 41]. Slowly, the concept of a multistage process became the predominant model, supported in part by the observation that the transformation of cultured cells occurred more rapidly, and at higher frequency, if cells were transfected with two versus a single oncogene[42] (see for example the review written in 1993[43]).

The marked number of mutational and chromosome changes present in cancer cells, evident from the early studies, and the fact that the median number of rearranged chromosome arms correlated with cancer prognosis[44] played significant roles in shaping models and thoughts. The multistep nature of carcinogenesis coupled with the evident chromosome changes led to several models which, at their core, involved clonal evolution, i.e. progressive selection of rare mutated cells[45]. Two extreme models were proposed. In the first, carcinogenesis required the activation of multiple oncogenes and/or the inactivation of tumour suppressor genes and each rearrangement contained an amplification, or loss, of a specific gene. This was supported by the identification of p53 as a tumour suppressor, the loss of which enabled the evolution of rare mutated cells[46]. Such hypotheses also suggested that elevated genome instability would not necessarily be required (each acquired mutation could increase growth potential). The alternative extreme model suggested that the vast majority of rearrangements had no phenotypic consequence, but rather represented "passenger mutations". To explain this, the notion of a mutator phenotype was proposed, which though controversial, remains actively discussed as a working model today[47-49]. Current advances in single cell sequencing procedures are demonstrating the enormous sequence changes between cells within a single tumour, and have shown that the level of plasticity within a tumour correlates with aggression[50] . However, these findings do not entirely allow the distinction to be drawn as to whether multiple mutations and a mutator phenotype are causal of malignancy or rather a consequence of malignancy. Critical to this is deducing the stage at which instability arises.

In parallel to the emerging concepts that carcinogenesis necessitated DNA sequencing changes, thinking also focused on the fact that cancer is predominantly a disorder of deregulated growth likely involving changed patterns of differentiation or de-differentiation. By this time it was generally accepted that differentiation during development was epigenetic[51]. This led to experiments where tumour cell nuclei with a normal modal chromosome number were transplanted into anucleated eggs to generate adult animals[52-54]. Significantly, such injections allowed for the development of normal animals, potentially demonstrating developmental totipotency that suggested a non-mutational basis for transformation to malignancy. In the context of current knowledge, such experiments were likely flawed, or at least exceptional. However, developments in the DNA methylation field provoked research into methylation changes associated with cancer, leading to proposals that methylation changes drive expression changes and thus cancer development[55]. Indeed, we now know that epigenetic changes are commonly found in

cancer cells and provide a route, like mutagenesis, to changed gene expression and thus function.

During the latter part of the 20th century the different models tended to be considered as exclusive, generating unfortunate conflicts that also influenced the battle that emerged regarding a viral aetiology for cancer. The identification of numerous oncogenes from work on viruses, plus revelations that many viruses encode proteins related to human growth factors whose expression could promote deregulated growth, led to a widespread belief that the majority of cancers were of viral origin[56, 57]. The viral community, in part because of its huge contributions to oncogene discovery (see below), gained significant influence. In hindsight, the strength of the arguments made by the viral community discouraged full consideration of a genomic instability model for cancer development. With our current knowledge, a model proposing a genomic instability origin for cancers would predict that viral infections could be carcinogenic, given the ability of many viruses to subvert components of the DDR (e.g.[58]). Any remaining controversy ultimately depends on the magnitude of the viral aetiological contribution, which cannot, for example, easily explain diet and smoking related tumorogenesis.

Oncogenes, their significance and oncogene induced stress.

Oncogenes were first identified by studying retroviruses: the prototype oncogene, src, is a chicken Rous sarcoma virus gene that was hijacked from the chicken genome[59]. It soon became clear that a defined, but significant, number of oncogenes were involved in cancer initiation and that oncogene expression caused neoplastic transformation[60]. In the early 1990s it was reported that genome instability occurred rapidly after Ras oncogene expression and subsequent reports demonstrated that this was not an isolated phenomenon[61-63]. By the late 1990s, it had been shown that p14ARF, the product of a second open reading frame (ORF) within the INK4 locus that binds to Mdm2 and upregulates p53, responded to Ras and Myc expression by activating the p53 - p21 pathway to drive senescence/apoptosis[64, 65]. Since it was known that DNA damage treatment activated p53 - p21 to drive senescence/apoptosis, this led to the suggestion that oncogene expression directly caused DNA damage[66].

It was initially proposed that deregulated metabolism due to Myc-dependent transcription increased reactive oxygen species (ROS) and thus DNA damage, an idea consistent with models postulating that a mutator phenotype underpinned cancer

development[66]. The link between oncogene expression and DNA damage and/or its repair generated significant interest: for example, Myc expression in non-cancerous cells was shown to reduce DNA repair efficiency and induce p53 and its ATM-ATR dependent (but p14ARF-independent) phosphorylation[67, 68]. Concurrently, p53 and p21 were shown to prevent cell proliferation when Cyclin E/Cdk2 was over-expressed and that this operated through an ATM/ATR-dependent and p14ARF independent mechanism[69]. Cyclin E expression had been previously shown to cause chromosome instability[70] and later it was demonstrated to interfere with replication[71]. The scene was thus set: oncogene-induced proliferation of otherwise normal cells perturbed DNA replication mechanisms, producing measurable DNA damage and genome rearrangements and activating p53 - p21 via ATM/ATR-dependent mechanisms. In 2005, two key papers clearly demonstrated both the activation of DDR signalling, including proteins required for cell cycle checkpoint arrest, and increased DNA damage markers in precancerous tissue and proposed that this reflected oncogene-induced damage arising from replication stress, synthesising the cell culture data and demonstrating a direct relevance to clinically derived cancer tissue[72, 73].

Tumour progression requires DNA damage response down-regulation.

As discussed above, the notions that endogenous and exogenous DNA damage cause mutations leading to carcinogenesis, and that cancer avoidance necessitates active DNA repair mechanisms emerged as key early concepts in the aetiology of cancer. What emerged more slowly, however, was an appreciation that DDR mechanisms in general, as distinct from specific repair pathways *per se*, were essential for cancer avoidance. Initially based on studies of apoptosis, an important concept for our understanding cancer onset emerged – it was not necessarily the DNA damage itself that killed (or prevented the growth of) a cell, but rather the signalling pathways activated by such damage. Cell cycle checkpoint pathways, initially defined as systems that monitor genome integrity, are now understood to be pivotal in precluding the continued proliferation of damaged cells[74]. p53 plays a key role in this and the frequent loss of p53 function in tumours contributed to the emerging notion that DDR pathways must be down regulated to allow uncontrolled proliferation[75].

In 1997 Serrano proposed that expression of the Ras oncogene led to p53/p21 activation that drives senescence or apoptosis. Thus, to achieve proliferation, the p53/p21 pathways must become down-regulated in Ras expressing cells[64]. Similar findings were

observed following expression of the Myc oncogene, although in this case, proliferation necessitated down regulation of the ARF-Mdm2-p53 pathways[76]. Slightly distinct but related examples also followed, such as the demonstration that, although tumours in BRCA2+/- mice undergo loss of heterozygosity, the proliferation of homozygous BRCA2-deficient tumour cells demands additional mutations in spindle checkpoint genes[77]. The full breadth of the relevant pathways that require down-regulation and their significance in contributing to tumour progression subsequently slowly unravelled, as did the realisation that down-regulation of these pathways could create a "mutator phenotype" causing genomic instability, as postulated many years earlier.

Lessons from history

Current models of cancer would argue that an initiating event(s), often caused by a mutation, leads to oncogene activation and ensuing replication stress. However, this must be coupled with subsequent down-regulation of DDR mechanisms - possibly by genetic alterations as a consequence of replication stress - to allow continued proliferation and continued genome instability - a prerequisite for a cancer cell to rapidly adapt to its ever-changing microenvironment. Whilst this historical reflection has considered the time-line at which seminal concepts emerged, this does not reflect the order of events in the aetiology of cancer (Figure 1). Initiating event(s) causing oncogene activation most likely precede a state of replication stress, but it remains unclear if the DDR down-regulation is always directly caused by errors arising from replication stress or if this can precede it. The findings of Bartek and Halazonetis[72, 73] that the up-regulation of the DDR occurs in precancerous lesions and that p53 mutations occur subsequently in late stage tumours, strongly support an order of events in which the onset of replication stress represents an early event, promoting the subsequent mutations that allow outgrowth.

However, a more recent study involving ultradeep sequencing of cancer genes in sun-exposed normal skin biopsies revealed a substantial accumulation of cancer driver mutations (with the characteristic signature of UV-induced mutations) that had undergone positive selection in the absence of evidence for cancer formation[78]. This suggests a different aetiology in which there is a strong initial selection to up-regulate growth-enhancing genes, and cells with such changes then await further steps leading to a genetically unstable state. The enhanced cancer predisposition caused by mutations in genes affecting both the early

(e.g. mutations enhancing an initiation event such as in xeroderma pigmentosum) and perceived late steps of cancer, (damage surveillance mechanisms such as in Li Fraumeni Syndrome) would be consistent with there not being a defined order of events leading to carcinogenesis.

Our historical reflection considers the many models or factors that have been proposed to contribute to carcinogenesis – viruses, epigenetic changes, DNA damaging agents, replication stressors and oxidative stress. Our current knowledge suggests that indeed all these factors are valid contributors, and they all merge into a model that leads ultimately to the generation of a genetically unstable state (Figure 2). Strikingly, this pinpoints the enormous significance of the DDR processes: their importance was evident in the earliest studies, but has emerged to be far more substantial than originally predicted. Whilst early studies demonstrated that cells can recover from exposure to external DNA damaging agents revealing that they harbour DNA repair mechanisms[20], perceptive scientists also saw that the DNA structure revealed by Watson and Crick could accumulate endogenously arising DNA damage, predicting that DNA repair pathways might be essential even during normal growth and metabolism[79]. However, these early studies did not predict that the DDR mechanisms encompass not only DNA repair processes that seek to avoid the initiator mutations, but also DNA damage surveillance mechanisms that preclude the proliferation of genetically unstable cells. Further, an efficient replication machinery that restricts replication errors, and processes that allow recovery from the inevitable difficulties encountered during replication in a manner that maintains genomic stability are also required (Box 1). We now know that cancer cells not only down-regulate these pathways but often subvert them to fast track evolution and gain adaptability, the ultimate driver of carcinogenesis and metastasis[80].

The future

Our historical reflections highlight the significance of the role played by the DDR processes in cancer aetiology. However, the plethora of DNA integrity surveillance, repair and signalling pathways, alongside their profound interconnectedness, has only been appreciated more recently. Similarly, the advent of tumour genome profiling by deep sequencing has only recently demonstrated that DDR genes are frequency mutated in all cancer types, with many individual pathways or genes found to be mutated in more than 50% of a specific cancer type (e.g. greater than 50% of ovarian cancers harbour mutations in HR genes)[80, 81] (Figure 3).

The steps during carcinogenesis can be summarised as shown in Figure 2. While an understanding of these steps is of significant academic interest, it also provides a key tool for informed, targeted cancer therapy[80]. The enhanced sensitivity of many cancer cells to DNA damaging agents, including radiation exposure, has been evident for many years, and indeed exploited for therapeutic benefit. The rationale for such sensitivity was poorly understood and unsatisfactorily often attributed to the more rapid growth of tumour cells. Consequently, the approach relied on serendipity, coupled with random trial and error. Our current knowledge of how the DDR pathways are changed in cancers is providing routes for more specific and rationally targeted therapy. A significant concept in this regard is that of synthetic lethality, where the goal is to identify a drug that will cause lethality to a cancer cell with inherent DDR defects whilst not harming a normal cell[80], The foremost and very elegant example of exploiting a synthetically lethal genetic relationship is the treatment of breast cancers arising in BRCA1/2 deficient patients[82]. The key insight came from the realisation that BRCA1 and 2 function during homologous recombination (HR), a key process that promotes replication fork stability, and that drug targeted inhibition of a specific enzyme (PARP1) confers a reliance on HR, and hence drug sensitivity[82]. Although such an approach might be anticipated to uniquely benefit BRCA1/2 defective patients, current studies revealing that HR can be downregulated in around 50% of ovarian cancers, dramatically expands the scope for such therapy[81]. A plethora of related studies are in progress, which include strategies to promote synthetic lethality based on such changes as the subversion of apoptosis, altered pathways of DNA double strand break repair, and loss of checkpoint arrest in cancer cells[80, 82].

Finally, to fully exploit the genome instability that is now recognised as an inherent property of most, if not all cancers, it is critical that we enhance both our understanding of the DDR pathways and exploit imaginative ideas to progress cancer treatment. Ironically, however, our understanding of the stages of carcinogenesis has also provided an explanation for why our targeted therapies will frequently fail.

References.

1. Boveri, T. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *J Cell Sci* **121 Suppl 1**, 1-84 (2008).
2. Burdette, W.J. The significance of mutation in relation to the origin of tumors: a review. *Cancer Res* **15**, 201-26 (1955).
3. Muller, H.J. Artificial Transmutation of the Gene. *Science* **66**, 84-7 (1927).
4. Muller, H.J. The Production of Mutations by X-Rays. *Proc Natl Acad Sci U S A* **14**, 714-26 (1928).
5. Hall, E.J. From chimney sweeps to oncogenes: the quest for the causes of cancer. *Radiology* **179**, 297-306 (1991).
6. Pott, P. [The first description of an occupational cancer in 1777 (scrotal cancer, cancer of chimney sweeps)]. *Bull Soc Liban Hist Med*, 98-101 (1993).
7. Waldron, H.A. A brief history of scrotal cancer. *Br J Ind Med* **40**, 390-401 (1983).
8. Franklin, R.E. & Gosling, R.G. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* **172**, 156-7 (1953).
9. Watson, J.D. & Crick, F.H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-8 (1953).
10. Brookes, P. & Lawley, P.D. The reaction of mustard gas with nucleic acids in vitro and in vivo. *Biochem J* **77**, 478-84 (1960).
11. Brookes, P. & Lawley, P.D. The reaction of mono- and di-functional alkylating agents with nucleic acids. *Biochem J* **80**, 496-503 (1961).
12. Brookes, P. & Lawley, P.D. Evidence for the Binding of Polynuclear Aromatic Hydrocarbons to the Nucleic Acids of Mouse Skin: Relation between Carcinogenic Power of Hydrocarbons and Their Binding to Deoxyribonucleic Acid. *Nature* **202**, 781-4 (1964).
13. Ames, B.N. A combined bacterial and liver test system for detection and classification of carcinogens as mutagens. *Genetics* **78**, 91-5 (1974).
14. Setlow, R.B. The photochemistry, photobiology, and repair of polynucleotides. *Prog Nucleic Acid Res Mol Biol* **8**, 257-95 (1968).
15. Pettijohn, D. & Hanawalt, P. Evidence for Repair-Replication of Ultraviolet Damaged DNA in Bacteria. *J Mol Biol* **9**, 395-410 (1964).
16. Friedberg, E.C. & Goldthwait, D.A. Endonuclease II of E. coli. *Cold Spring Harb Symp Quant Biol* **33**, 271-5 (1968).
17. Laval, J. & Laval, F. Enzymology of DNA repair. *IARC Sci Publ*, 55-73 (1980).

18. Lindahl, T. An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues. *Proc Natl Acad Sci U S A* **71**, 3649-53 (1974).
19. Ljungquist, S. & Lindahl, T. A mammalian endonuclease specific for apurinic sites in double-stranded deoxyribonucleic acid. I. Purification and general properties. *J Biol Chem* **249**, 1530-5 (1974).
20. Friedberg, E.C. A brief history of the DNA repair field. *Cell Res* **18**, 3-7 (2008).
21. Laval, F. Repair of methylated bases in mammalian cells during adaptive response to alkylating agents. *Biochimie* **67**, 361-4 (1985).
22. Lehmann, A.R., McGibbon, D. & Stefanini, M. Xeroderma pigmentosum. *Orphanet J Rare Dis* **6**, 70 (2011).
23. Cleaver, J.E. Xeroderma pigmentosum: a human disease in which an initial stage of DNA repair is defective. *Proc Natl Acad Sci U S A* **63**, 428-35 (1969).
24. Leach, F.S. et al. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* **75**, 1215-25 (1993).
25. Fishel, R. et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**, 1027-38 (1993).
26. Bronner, C.E. et al. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* **368**, 258-61 (1994).
27. Papadopoulos, N. et al. Mutation of a mutL homolog in hereditary colon cancer. *Science* **263**, 1625-9 (1994).
28. Cui, Z. et al. Association between polymorphisms in XRCC1 gene and clinical outcomes of patients with lung cancer: a meta-analysis. *BMC Cancer* **12**, 71 (2012).
29. Yamtich, J., Nemec, A.A., Keh, A. & Sweasy, J.B. A germline polymorphism of DNA polymerase beta induces genomic instability and cellular transformation. *PLoS Genet* **8**, e1003052 (2012).
30. Karahalil, B., Bohr, V.A. & Wilson, D.M., 3rd. Impact of DNA polymorphisms in key DNA base excision repair proteins on cancer risk. *Hum Exp Toxicol* **31**, 981-1005 (2012).
31. Polani, P.E. DNA repair defects and chromosome instability disorders. *Ciba Found Symp*, 81-131 (1979).
32. Moldovan, G.L. & D'Andrea, A.D. How the fanconi anemia pathway guards the genome. *Annu Rev Genet* **43**, 223-49 (2009).

33. Gray, M.D. et al. The Werner syndrome protein is a DNA helicase. *Nat Genet* **17**, 100-3 (1997).
34. Hartman, A.R. et al. Prevalence of BRCA mutations in an unselected population of triple-negative breast cancer. *Cancer* **118**, 2787-95 (2012).
35. Cavanagh, H. & Rogers, K.M. The role of BRCA1 and BRCA2 mutations in prostate, pancreatic and stomach cancers. *Hered Cancer Clin Pract* **13**, 16 (2015).
36. Banerjee, T. & Brosh, R.M., Jr. RECQL: a new breast cancer susceptibility gene. *Cell Cycle*, 0 (2015).
37. Lindell, B. & Sowby, D. The 1958 UNSCEAR report. *J Radiol Prot* **28**, 277-82 (2008).
38. Berenblum, I. & Shubik, P. The role of croton oil applications, associated with a single painting of a carcinogen, in tumour induction of the mouse's skin. *Br J Cancer* **1**, 379-82 (1947).
39. Farber, E. The multistep nature of cancer development. *Cancer Res* **44**, 4217-23 (1984).
40. Farber, E. Carcinogenesis--cellular evolution as a unifying thread: Presidential address. *Cancer Res* **33**, 2537-50 (1973).
41. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, 197-200 (1975).
42. Weinberg, R.A. The action of oncogenes in the cytoplasm and nucleus. *Science* **230**, 770-6 (1985).
43. Vogelstein, B. & Kinzler, K.W. The multistep nature of cancer. *Trends Genet* **9**, 138-41 (1993).
44. Fearon, E.R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759-67 (1990).
45. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **194**, 23-8 (1976).
46. Chen, P.L., Chen, Y.M., Bookstein, R. & Lee, W.H. Genetic mechanisms of tumor suppression by the human p53 gene. *Science* **250**, 1576-80 (1990).
47. Loeb, L.A., Springgate, C.F. & Battula, N. Errors in DNA replication as a basis of malignant changes. *Cancer Res* **34**, 2311-21 (1974).
48. Loeb, L.A. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res* **51**, 3075-9 (1991).

49. Fox, E.J., Prindle, M.J. & Loeb, L.A. Do mutator mutations fuel tumorigenesis? *Cancer Metastasis Rev* **32**, 353-61 (2013).
50. Fox, E.J. & Loeb, L.A. Cancer: One cell at a time. *Nature* **512**, 143-4 (2014).
51. Gurdon, J.B. & Uehlinger, V. "Fertile" intestine nuclei. *Nature* **210**, 1240-1 (1966).
52. McKinnell, R.G., Deggins, B.A. & Labat, D.D. Transplantation of pluripotential nuclei from triploid frog tumors. *Science* **165**, 394-6 (1969).
53. Mintz, B. & Illmensee, K. Normal genetically mosaic mice produced from malignant teratocarcinoma cells. *Proc Natl Acad Sci U S A* **72**, 3585-9 (1975).
54. Illmensee, K. & Mintz, B. Totipotency and normal differentiation of single teratocarcinoma cells cloned by injection into blastocysts. *Proc Natl Acad Sci U S A* **73**, 549-53 (1976).
55. Bird, A.P. The relationship of DNA methylation to cancer. *Cancer Surv* **28**, 87-101 (1996).
56. Deuel, T.F., Huang, J.S., Huang, S.S., Stroobant, P. & Waterfield, M.D. Expression of a platelet-derived growth factor-like protein in simian sarcoma virus transformed cells. *Science* **221**, 1348-50 (1983).
57. Waterfield, M.D. et al. Platelet-derived growth factor is structurally related to the putative transforming protein p28sis of simian sarcoma virus. *Nature* **304**, 35-9 (1983).
58. Lees-Miller, S.P. et al. Attenuation of DNA-dependent protein kinase activity and its catalytic subunit by the herpes simplex virus type 1 transactivator ICP0. *J Virol* **70**, 7471-7 (1996).
59. Bishop, J.M. Enemies within: the genesis of retrovirus oncogenes. *Cell* **23**, 5-6 (1981).
60. Cooper, G.M. Cellular transforming genes. *Science* **217**, 801-6 (1982).
61. Denko, N.C., Giaccia, A.J., Stringer, J.R. & Stambrook, P.J. The human Ha-ras oncogene induces genomic instability in murine fibroblasts within one cell cycle. *Proc Natl Acad Sci U S A* **91**, 5124-8 (1994).
62. Mai, S., Fluri, M., Siwarski, D. & Huppi, K. Genomic instability in MycER-activated Rat1A-MycER cells. *Chromosome Res* **4**, 365-71 (1996).
63. Felsher, D.W. & Bishop, J.M. Transient excess of MYC activity can elicit genomic instability and tumorigenesis. *Proc Natl Acad Sci U S A* **96**, 3940-4 (1999).

64. Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D. & Lowe, S.W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88**, 593-602 (1997).
65. Zindy, F. et al. Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. *Genes Dev* **12**, 2424-33 (1998).
66. Vafa, O. et al. c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. *Mol Cell* **9**, 1031-44 (2002).
67. Karlsson, A. et al. Defective double-strand DNA break repair and chromosomal translocations by MYC overexpression. *Proc Natl Acad Sci U S A* **100**, 9974-9 (2003).
68. Lindstrom, M.S. & Wiman, K.G. Myc and E2F1 induce p53 through p14ARF-independent mechanisms in human fibroblasts. *Oncogene* **22**, 4993-5005 (2003).
69. Minella, A.C. et al. p53 and p21 form an inducible barrier that protects cells against cyclin E-cdk2 deregulation. *Curr Biol* **12**, 1817-27 (2002).
70. Spruck, C.H., Won, K.A. & Reed, S.I. Deregulated cyclin E induces chromosome instability. *Nature* **401**, 297-300 (1999).
71. Ekholm-Reed, S. et al. Deregulation of cyclin E in human cells interferes with prereplication complex assembly. *J Cell Biol* **165**, 789-800 (2004).
72. Bartkova, J. et al. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* **434**, 864-70 (2005).
73. Gorgoulis, V.G. et al. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* **434**, 907-13 (2005).
74. Wahl, G.M. & Carr, A.M. The evolution of diverse biological responses to DNA damage: insights from yeast and p53. *Nat Cell Biol* **3**, E277-86 (2001).
75. Baker, S.J. et al. Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science* **244**, 217-21 (1989).
76. Eischen, C.M., Weber, J.D., Roussel, M.F., Sherr, C.J. & Cleveland, J.L. Disruption of the ARF-Mdm2-p53 tumor suppressor pathway in Myc-induced lymphomagenesis. *Genes Dev* **13**, 2658-69 (1999).
77. Lee, H. et al. Mitotic checkpoint inactivation fosters transformation in cells lacking the breast cancer susceptibility gene, Brca2. *Mol Cell* **4**, 1-10 (1999).
78. Martincorena, I. et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880-6 (2015).

79. Lindahl, T. The Croonian Lecture, 1996: endogenous damage to DNA. *Philos Trans R Soc Lond B Biol Sci* **351**, 1529-38 (1996).
80. Pearl, L.H., Schierz, A.C., Ward, S.E., Al-Lazikani, B. & Pearl, F.M. Therapeutic opportunities within the DNA damage response. *Nat Rev Cancer* **15**, 166-80 (2015).
81. Cancer Genome Atlas Research, N. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-15 (2011).
82. Lord, C.J., Tutt, A.N. & Ashworth, A. Synthetic Lethality and Cancer Therapy: Lessons Learned from the Development of PARP Inhibitors. *Annu Rev Med* **66**, 455-70 (2015).
83. Hill, R.F. A radiation-sensitive mutant of Escherichia coli. *Biochim Biophys Acta* **30**, 636-7 (1958).
84. Hartwell, L.H. & Weinert, T.A. Checkpoints: controls that ensure the order of cell cycle events. *Science* **246**, 629-34 (1989).
85. Kinzler, K.W. & Vogelstein, B. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* **386**, 761, 763 (1997).
86. Hanahan, D. & Weinberg, R.A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
87. Murphree, A.L. & Benedict, W.F. Retinoblastoma: clues to human oncogenesis. *Science* **223**, 1028-33 (1984).
88. Kerr, J.F., Wyllie, A.H. & Currie, A.R. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* **26**, 239-57 (1972).
89. Taylor, A.M. et al. Ataxia telangiectasia: a human mutation with abnormal radiation sensitivity. *Nature* **258**, 427-9 (1975).
90. Miki, Y. et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66-71 (1994).
91. Moynahan, M.E., Chiu, J.W., Koller, B.H. & Jasin, M. Brca1 controls homology-directed DNA repair. *Mol Cell* **4**, 511-8 (1999).
92. Futreal, P.A. et al. BRCA1 mutations in primary breast and ovarian carcinomas. *Science* **266**, 120-2 (1994).
93. Sharan, S.K. et al. Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature* **386**, 804-10 (1997).

Figure Legends.

Figure 1: Time line showing when key concepts and findings relating to cancer development emerged. (currently shown as a and b but will be merged).

Figure 2: How the DDR pathways impact upon steps leading to cancer.

A depiction of how changes in the DDR pathways promote critical steps in the aetiology of carcinogenesis.

Figure 3: Many DDR proteins are mutated in cancer.

Data from the Cancer Genome Atlas has revealed the extent to which mutations in DDR proteins are observed in cancers. These changes differ for different tumour sites. A) shows a radial plot, where the radius length indicates the proportion of patients with protein coding mutations predicted to be impacting in each cancer type. All mutations included are present in at least two distinct patient samples. The concentric circles indicates the percentage of patients affected. B) shows copy-number variation in the different DDR pathways (red, loss of gene-copies; blue, amplification of genes C) expression level variation in DDR pathways. Red, decreased expression; Blue, increased expression. Pathways are: AM, alternative mechanism for telomere maintenance; BER, base excision repair; CPF, checkpoint factor; CR, chromatin remodelling; CS, chromosome segregation; DR, direct repair; FA, Fanconi anaemia pathway; HR, homologous recombination; MMR, mismatch repair; NER, nucleotide excision repair; NHEJ, non-homologous end joining; OD, other double-strand break repair; TLS, translesion synthesis; TM, telomere maintenance; UR, ubiquitylation response.

Box 1: DDR processes of relevance for Cancer

Data for the time line figure (currently shown as two figures a, b – they will be combined)

Green are key concepts: yellow are key findings

1. Environmental exposure can cause cancer.
1775
[6]

2.. Concept of hereditary material
1900-1950.
[1-4].

3. Helical structure of DNA
1953
[8, 9]

4. Link between mutagenesis and carcinogenesis. Cancer proposed to be caused by DNA mutations.
1900-1961 specific link to 1961
[2, 10, 11].

5. Alklyating agents react with and damage DNA
1960
[10]

6. DNA can be damaged by endogenous and exogenous agents.
1900-1961
[4, 10, 79].

7. DNA damage can be repaired.
Circa 1958
[83](for a review see [20])

8 XP identified as the first DNA repair disorder.
1969,
[23]

9 DNA repair defective disorders are cancer prone; including Bloom's Syndrome and Fanconi anaemia. Hence failure to repair DNA contributes to cancer
1969-2015
[23, 31]

10. Viruses proposed as a major cause of cancer.
1975-1985
[56, 57].

11. Cell cycle checkpoints proposed.

1989
[84]

12. Concept of caretaker and gate keeper genes.
1997
[85]

13. Acquired capability of cancers.
2000
[86]

14. Concept of oncogenes.
1981
[59].

15. Concept of tumour suppressors
1984
[87]

16 Significance of replication stress and replication fork stability appreciated.
2005
[72, 73].

17 Oncogene expression causes replication instability.
1999
[70]

18. Wylie and co-workers define the process of apoptosis.
1972.
[88]

19. AT is a radiosensitive disorder with cancer predisposition.
1975
[89]

20. Microsatellite instability identified in LS tumours and shown to be due to :MMR deficiency. MSH2 identified as 1st LS locus.
1993
[24, 25]

21. Mutator phenotype for cancer cells proposed.
1974
[47]

22, BRCA1/2, which are mutated in hereditary breast cancer, function in HR.

1997-1999.
90-93

23. Ames test established to identify carcinogens via analysis of mutagens.
1974
13

24. p53 mutations identified in cancers. Role of p53 dependent surveillance pathways recognised as suppressing cancers.
1990.
46, 74

25. Oncogenes expression leads to p53/p21 activation, senescence or apoptosis.
1997
64.

26. Oncogene expression causes deregulated metabolism leading to ROS and DNA damage.
2002-2003
66, 67

27. DDR is an anti cancer barrier in early stage tumourogenesis; mutations in DDR genes in later stage tumours
72, 73

28. Multiple mutations in DDR genes identified in cancers.
2014
80, 81

Box 1. DDR processese of relevance for Cancer.

1) DNA repair.

There are multiple DNA repair pathways, with sub-pathways providing lesion specificity. Nucleotide Excision Repair (NER) removes bulky DNA lesions; DNA non-homologous end-joining and homologous recombination (HR) repair DNA double strand breaks (DSBs); Mismatch repair (MMR) corrects mismatched base pairs; Base excision repair (BER) repairs damaged bases and links to single strand break (SSB) repair. Mutations in these pathways in patients enhance cancer susceptibility.

2) Damage Response Signalling.

There are two DDR signalling pathways. ATM-dependent signalling is activated by DSBs; ATR-dependent signalling is activated by single stranded DNA. DDR signalling can activate apoptosis, checkpoint arrest and influence DNA repair. Mutations in ATM signalling components in patients confers cancer susceptibility. However, ATR-deficient mice do NOT get tumours.

3) Cell cycle checkpoints.

DNA integrity is constantly monitored with DNA damage triggering a checkpoint response that prevents cell cycle progression. Arrest can be permanent or transient. Checkpoints prevent entry from G1/S, G2/M and an intra-S phase checkpoint regulates fork progression or origin firing. Many tumours have inactivated checkpoint responses.

4) Apoptosis.

Apoptosis represents a programmed cell death pathway, which functions in some tissues during normal development but also prevents proliferation of damaged cells. Apoptosis can be p53 dependent or independent. *p53* is commonly mutated in cancer.

5) Fidelity of replication.

Multiple processes function to maintain the accuracy of replication and enhance recovery from replication fork stalliing or collapse. Homologous recombination (HR) plays a key role and HR genes are commonly mutated in cancers.

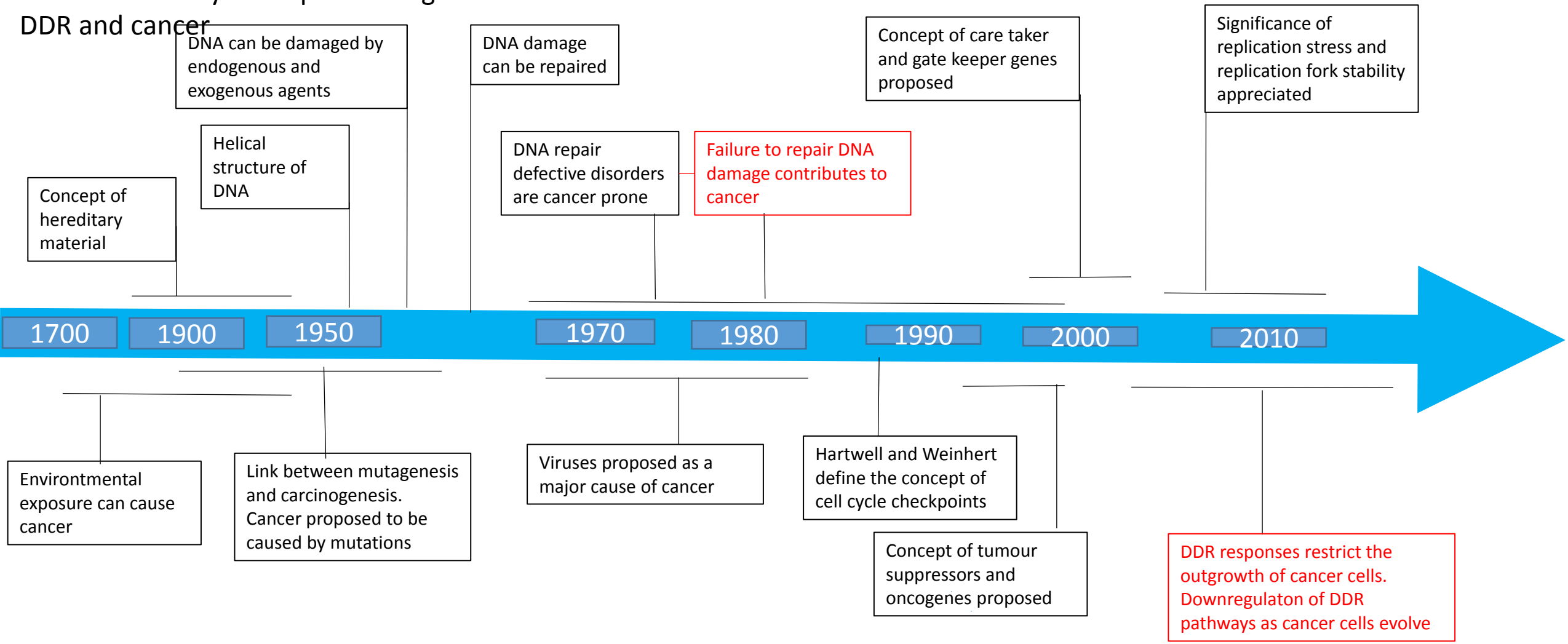
6) DNA re-replication.

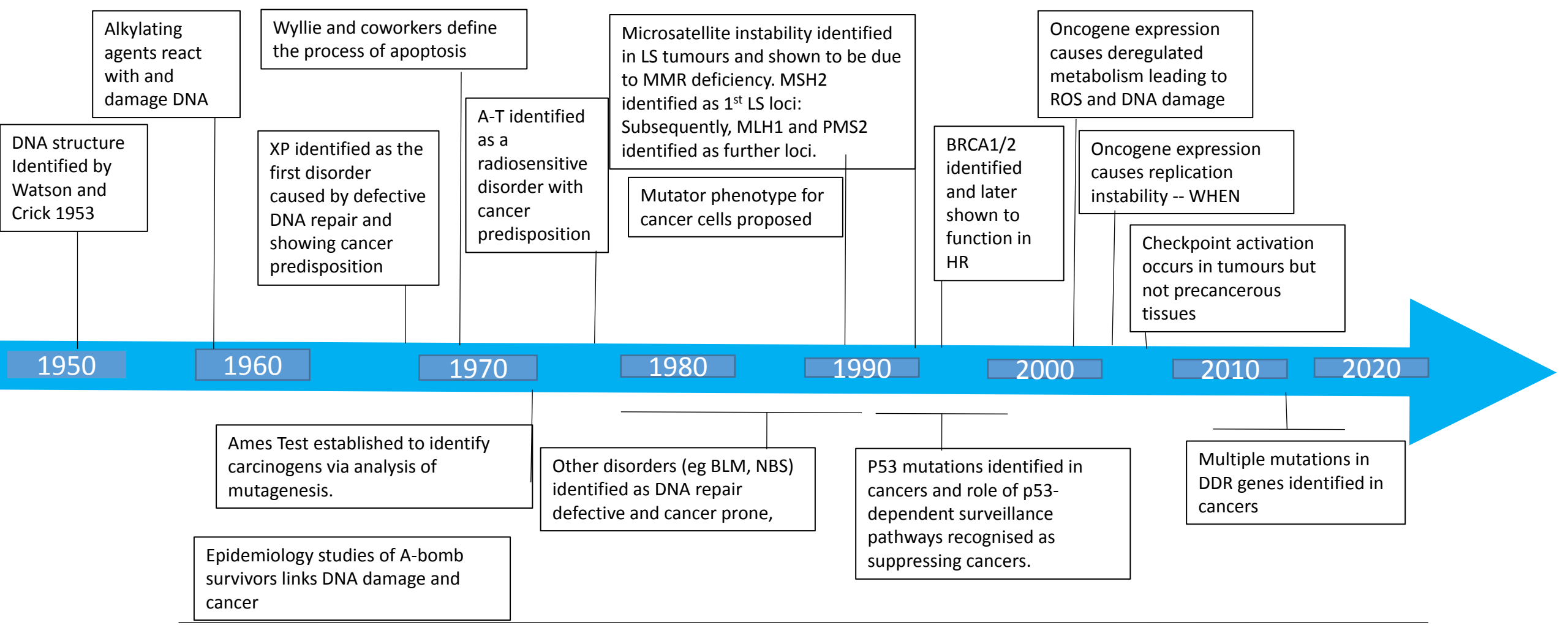
Re-replication can cause aneuploidy and subsequently genomic instability. Several mechanisms prevent DNA re-replication.

7) Telomere length.

Shortened telomeres lead to senescence; cancer cells need to maintain telomere length. Activation of telomerase or an alternative pathway (ALT) to maintain telomere length in cancers is common.

Time line 1a: Key concepts relating to DDR and cancer





Time line 1b: Key findings relevant to DDR pathways and cancer

How DDR pathways impact upon Steps leading to cancer

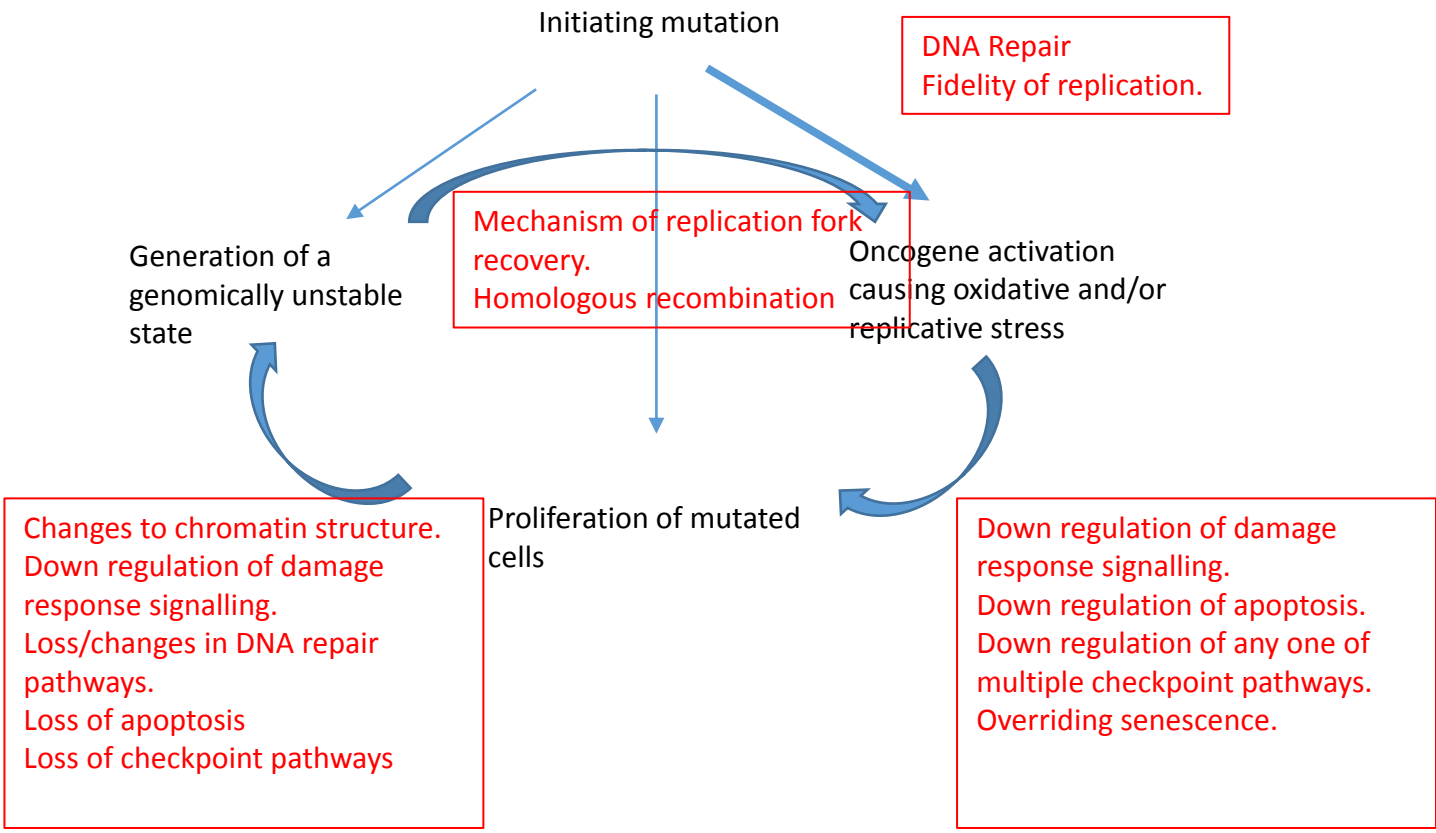


Figure 2

Figure 3.

Figure to be adapted from Figure 4 Nature Reviews in Cancer 15, 166-180 (2015)

Our reference 80.