The evolution of lifespan and age-dependent cancer risk

Andrii I. Rozhok¹ and James DeGregori¹,²,³,⁴

¹Department of Biochemistry and Molecular Genetics, ²Integrated Department of Immunology, ³Department of Pediatrics, ⁴Department of Medicine, Section of Hematology, University of Colorado School of Medicine, Aurora, CO 80045

Corresponding Authors:
Andrii I. Rozhok: Andrii.Rozhok@ucdenver.edu
James DeGregori: James.DeGregori@ucdenver.edu

Abstract

The Armitage-Doll multi-stage model of carcinogenesis tremendously refocused cancer science by postulating that carcinogenesis is driven by a sequence of genetic changes in cells. Age-dependent cancer incidence thus has been explained in terms of the time necessary for oncogenic mutations to occur. While the multi-step nature of cancer evolution is well-supported by evidence, the mutation-centric theory is unable to explain a number of phenomena, such as the disproportion between cancer frequency and animal body size or the scaling of cancer incidence to animal lifespan. In this paper, we present a theoretical review of the current paradigm and discuss some fundamental evolutionary theory postulates that explain why cancer incidence is a function of lifespan and physiological, not chronological, aging.
The Armitage-Doll multi-stage paradigm of carcinogenesis in the light of current evidence

Over half a century ago, Carl Nordling postulated that carcinogenesis is driven by mutations that occur in cells [1]. Peter Armitage and Richard Doll generalized these ideas and proposed the multi-stage theory of carcinogenesis in a seminal paper [2], stating that carcinogenesis typically requires 6-7 mutations and/or other cell alteration to malignantly transform a cell, based on the evidence that the age-dependent exponential increase in cancer incidence follows mathematically the 6th power of age. They noted, though, that such a model holds only if the probability of oncogenic events stays the same throughout the human lifespan, which implies linear mutation accumulation with age. Later, Peter Nowell and David Hungerford provided further corroboration of the genetic origins of cancer [3]. Nowell also introduced the concept of clonal evolution of cancers, suggesting that it occurs through rounds of selection in tissues for progressively more advantageous pre-malignant cells and progressively more malignant phenotypes [4]. As pre-neoplastic cells transformed by mutations expand clonally, the likelihood that further oncogenic events will occur in already oncogenically initiated cells increases proportionally to the number of such dividing cells, and such a pattern thus is believed to ultimately provide for the exponential increase in the incidence of cancers with age.

These seminal theoretical concepts and findings tremendously refocused cancer science, leading to extensive discovery efforts for specific oncogenic events, a multitude of which has now been well characterized. Our current understanding of cancer development largely remains within this theoretical framework, with the assumption often made that oncogenic events typically confer certain defined selective (proliferative) advantages to cells over their normal wild-type counterparts (e.g. [5, 6]). However, as new evidence accumulates, it is becoming clear that carcinogenesis is a more complicated process regulated by many forces beyond the occurrence of oncogenic mutations.

The first theoretical problem with prevailing models arises with the recognition that somatic mutations do not accumulate as Armitage and Doll assumed. Data from various murine and human tissues convincingly demonstrate that a substantial portion of somatic mutations (up to 50%) accumulate early in life before full body maturation [7, 8]. The rate of accumulation significantly slows down thereafter (Fig. 1). Evidence from at least some tissues suggests that this slow-down occurs when stem cells transition from body building to body maintenance and thus dramatically reduce their division rates [9]. Within the current theoretical framework, the observations that cancer incidence increases exponentially with age and that mutation accumulation occurs largely during development stand in opposition to each other, which implies that this framework is incomplete. To address this, we suggest that selection for oncogenic clones is highly dependent on tissue context, which changes with age, and necessarily implies that clonal evolution is differentially regulated throughout life. As the probability of sequential mutation accumulation in one cell depends both on cell division rates and the number of dividing cells, early
in life when cells cycle rapidly, the lower observed risk of cancer should require that those cells that acquire mutations do not expand into multi-cell clones (which would increase the target size for the next oncogenic mutation) [10]. Late in life, the slow cell division program needs to be compensated by more profound clonal expansions of pre-malignant cells to provide for the elevated risk of cancers that require multiple mutations. Such logic implies that the strength of positive selection in tissues for oncogenically initiated cells should significantly differ between early and late life; otherwise the ultimate cancer incidence curve would resemble more closely the shape of mutation accumulation with age.

In fact, evidence from the hematopoietic system corroborates the idea that positive selection in tissue for mutant cellular phenotypes is stronger late in life, as selection for particular oncogenic events has been shown to be substantially stronger in old murine hematopoietic microenvironments than in young [11, 12]. Increased oncogenic clonal expansions and clonal hematopoiesis have also been demonstrated in old mice [13, 14]. Moreover, after the age of approximately 40 for humans, exponentially increasing clonality is observed in hematopoiesis, bringing about large expanded clones that are not observed early in life [15-18], and these clonal expansions have been postulated to result from enhanced selection in the aged bone marrow microenvironment [19-21]. Interestingly, increased hematopoietic clonality late in life, whether or not associated with known oncogenic mutations, is associated with increased leukemia development as well as all-cause mortality [16], consistent with these clonal expansions being symptoms of an aged soma which contributes to alterations in oncogenic selection, as well as more general physiological decline.

A second major problem with the prevailing model(s) is the assumption that any mutation can confer a defined effect on a cell’s fitness (see Glossary), a prediction that contradicts evolutionary theory. A cell’s fitness represents the relative ability of stem and progenitor cells of a certain genotype (or epi-genotype) to contract, persist or expand in a cell population. Functional (non-silent) genetic alterations usually alter phenotype in a defined way. However, changes in fitness (selective value) arise only when this phenotype interacts with environment and can vastly differ in an environment-dependent manner. For example, a mutation conferring drug resistance to a cell produces a defined phenotype, but this phenotype will only have a selective advantage over other cells under drug exposure, and the degree of this advantage will also vary depending on the drug regimen. Likewise, a young and fit human tissue represents a dramatically different microenvironment than an aging one typically characterized by increased chronic inflammation, tissue integrity loss and other aging markers. Many experimental studies now demonstrate that selection for oncogenic mutants heavily depends on tissue microenvironmental context, which is influenced by environmental exposure (including carcinogens like cigarette smoke), age, and associated inflammation and altered tissue landscapes [11, 12, 22-25]. Therefore, understanding that selection for mutant cells, including oncogenic mutants, is not a stationary phenomenon but depends on other factors including aging-related factors that operate above the cellular level is
crucial to bring the current multi-stage theory of carcinogenesis in accord with evolutionary theory and modern evidence.

The evolution of lifespan as key to understanding aging-related cancer risk

In order to understand how and why physiological aging, not just chronological timing, can be directly implicated in cancer and other disease risk, it is key to consider how lifespans evolve. A substantial emphasis in medical and cancer-related literature is placed on time-dependent accumulation of cellular damage as a major driver of aging, thus positing aging as rate-limited by a certain mutation/damage “clock” [26, 27]. However, as mentioned previously, roughly half of all somatic mutations occur early in life when most cell divisions happen, but physiological aging does not follow this pattern and is significantly delayed, undermining the cell-intrinsic model of aging. The somatic damage accumulation model is also discordant with current views within evolutionary biology [28, 29]. The evolutionary theory of aging and lifespan states that aging and lifespan are determined by the rate at which selection is relaxed during the life of organisms, and is supported by substantial observational and experimental evidence [28, 29]. As proposed by Peter Medawar [30], alleles that increase organismal fitness during the pre- and reproductive period are strongly selected for in populations, while this selection is relaxed later in life because of lower chances of survival and reproduction (Figure 2A). In this way, the evolution of lifespan is dependent on the ecological niche of a species, which determines the likelihood of dying over time for any reason. For example, there are indications that reduced predation alone can substantially slow down the evolved rates of physiological aging in mammals [31], engendering a slower life strategy (delayed reproduction and longer somatic maintenance; see Figure 2A and 2B for a theoretical example). Therefore, evolution at the level of the organism (germline) basically creates two principal periods in mammalian lifespans: 1) the pre- and reproductive period where high body fitness is maintained and 2) post-reproductive period during which body fitness and tissue microenvironment declines progressively faster. Finally, the post-reproductive period does not refer to the ages where an animal is necessarily incapable of reproducing, but simply when reproduction becomes more and more unlikely (typically due to prior death from a variety of causes).

The evolution of lifespan at the germline level has important implications for the stem and progenitor cells responsible for tissue renewal, which are subject to malignant transformation by oncogenic mutations (Figure 2B). During pre- and reproductive periods, selection for body fitness works to fine-tune the behavior of stem cells and their tissue microenvironment for optimal tissue renewal and function, so that stem cells divide and differentiate in a relatively stable environment to which they are well-adapted. In this environment, there still exists a chance that an oncogenic mutation can improve a cell’s fitness so that the mutated cell out-proliferates its normal counterparts in the stem cell pool. However, the probability and magnitude of such an effect will
dramatically increase in the post-reproductive tissue microenvironment, as stem cells have not adapted to this environment through germline evolution (Figure 3). This argument is consistent with evolutionary theory which holds that adaptation increases the strength of purifying selection and also reduces the ability of mutations to increase fitness [32].

Lifespan that significantly extend past the reproductive period are a consequence of modernity. Unlike modern humans, non-captive animals typically do not survive into the post-reproductive period, and we can only observe the effects of the “post natural survival” portion of their lifespan in captivity. For example, wild mice rarely survive beyond 6-7 months, but can live for more than 2 years in laboratories, and get the vast majority of cancers during the period that extends beyond their typical survival in nature [33]. As described above, different potential lifespans have evolved in different animals, dependent on environmental conditions, in order to maximize reproductive success. What then accounts for the vast differences in potential lifespans, and thus avoidance of diseases of old age, in different mammals? Is greater longevity achieved by the evolution of lower mutation rates or by enhancing other means of somatic maintenance? A clue is provided by studies of two DNA polymerase δ mutations in mice, L604K and L604G, both of which increase mutation rates more than 5-fold in the heterozygous state [34]. However, only the +/L604K genotype is associated with an accelerated incidence of cancer (occurring several months earlier) that mirrors a several month shortening of lifespan (Figure 4). In contrast, the +/L604G genotype is not associated with either an increase in cancer risk/rate or changes in lifespan. Although the reason underlying this difference is unknown, heterozygosity for L604K shows a (non-significant) trend to higher mutation rates relative to +/L640G. Regardless, increases in mutation rates clearly do not suffice to change the age-dependent cancer incidence curve (as seen in +/L640G mice), but may require other physiological changes (reflected in the shortened lifespan of +/L640K mice).

Without considering the effects of lifespan evolution on age-dependent cancer and other disease risk, the current multi-stage model of cancer creates seemingly unresolved paradoxes. The well-known Peto’s paradox is the inability of the current model to explain why animals with larger bodies and longer lifespans (both effectively increasing the target size of dividing stem cell pools for mutations) do not have proportionally higher incidence of cancers [35]. Multiple explanations have been proposed focusing mostly on the evolution of intracellular mechanisms that reduce the risk of cell transformation [36], but the paradox remains unresolved. Explanations have often focused on changes in mutation rates for different species, and yet the evolution of complex multicellularity has not been accompanied by reductions in mutation rate (in fact, mammals have higher mutation rates than many single-cell eukaryotes and bacteria [37]), and as described above, substantial increases in mutation rate can be engineered in mice without increasing cancer risk or decreasing lifespan. If carcinogenesis is viewed simply as a process depending on time and number of cell divisions, this paradox might indeed have no solutions. However, considering that lifespan and body size primarily evolve based on the species ecological niche, and that cancer and other disease risk is lifespan-dependent, resolves this paradox. Similar sized animals within very
closely related groups, such as rodents, often have very different lifespans. Evolutionary theory holds that this occurs because of different species-specific probabilities of late-life survival and reproduction. This evolution of divergent lifespans is unlikely to be caused by a very rapid evolution of mutation rates, extra tumor suppressor genes or other cancer-related factors that are often proposed [36]. Very long lifespans, such as those of whales, Greenland sharks, Galapagos tortoises and elephants, appear to typically evolve in animals that are under reduced risk of predation or death by other external factors as adults. There is a common life strategy that is acted upon by natural selection when these external risk factors change - somatic maintenance; the “tuning” of this investment will either increase or decrease the period of somatic maintenance depending on environmental selective forces in order to maximize reproductive success in the new environment (Figure 2). This somatic maintenance strategy has implications for the timing of diseases of old age like cancer, should an animal somehow manage to survive beyond periods typical in nature (as humans do today).

From this perspective it is not lifespan that depends on cancer-free survival, but rather that the age-dependence of cancer incidence should be viewed as a function of lifespan, which itself is dependent on the curve of body fitness decline (Figure 2B). This perspective helps explain why animals rarely have cancer in the wild, since many animals simply do not survive to the age when cancer becomes a risk factor (Figure 2C). Senescence appears to be common in some groups of wild animals [38, 39], however cancer demography data are lacking from wild populations and more research is needed to elucidate if cancer rates are higher in aged wild animals. It is also unclear how senescence in wild populations affects the overall odds of reproductive success, which is the key regulator of the age-dependent strength of germline selection. Recent evidence also indicates that different tissues in the body have different rates of aging, which might potentially be a contributor to the differential tissue-specific cancer risk [40]. Finally, while still speculative, understanding cancer as being determined by evolved life history may help explain the earlier incidence of a number of cancers, such as brain, bowel, chronic myeloid leukemia, chronic lymphocytic leukemia and acute myeloid leukemia, in males relative to females (cancers which are neither sex-specific or substantially due to smoking were queried; Cancer Research UK; http://www.cancerresearchuk.org). Males (for humans and many other species) age earlier, and this earlier decline in tissue microenvironments could potentially elicit earlier somatic oncogenic adaptation relative to females.

**Understanding carcinogenesis as rate-limited by physiological aging-dependent processes**

Carcinogenesis is currently understood as a series of oncogenic mutations that increase the relative fitness of the cells in which they occur and lead to a series of clonal expansions of these oncogenically-initiated cells [5]. Stem and progenitor cells have been shown to compete for niche space [41], so fitness can be considered as the ability of dividing cells to remain in the competitive stem and progenitor pool and expand their progeny instead of differentiating, dying or senescing.
If a mutation increases the probability for a cellular genotype to become more frequent in a competitive population, then this mutation increases cell fitness. In this light, carcinogenesis is recognized as a Darwinian process of mutation and selection for the fittest among dividing somatic cells. Within the current multi-stage paradigm, oncogenic mutations are deemed to provide immediate and measurable fitness advantage to the affected cells and thus drive carcinogenesis [5, 6]. However, multiple lines of evidence indicate that this process is not so simple, and oncogenic mutations experience very different fates depending on the tissue environmental context in which they occur, and the evolutionary forces, including drift and selection, to which they are subjected [20]. For example, in small populations (of cells or animals), drift can be the dominant evolutionary force. Indeed this is the case in the gut epithelium, where epithelial stem cells are split into small clusters of 15-20 cells and their clonal dynamics in the gut crypts have been shown to be essentially neutral [42, 43]. Even potent oncogenic mutations, such as those in p53, do not confer significant selective advantage to gut stem cells, unless they occur in an inflamed environment [23]. On the other hand, in large competitive stem cell pools, such as hematopoietic stem cells (HSC), selection has been shown to dominate clonal dynamics [12, 44, 45]. Studies indicate that selection acts differentially on oncogenic mutations in young and healthy versus old and damaged tissue contexts [11, 12, 46], with the former being dominated by stabilizing selection that prevents most genetic changes from conferring fitness advantages to cells. As noted above, because of the co-evolution at the germline level of stem cells with their tissue microenvironment that regulates cell fates, stem cells in young tissues are closer to their peak fitness than cells in aged or damaged tissues (Figure 3). This adaptation to specific tissue environments leads to the increased role of stabilizing selection which reduces the probability that a random genetic mutation can be advantageous. In adult stem cell pools of a stable size, the average frequency of differentiation per stem cell division is close to 0.5, and stabilizing selection acts to increase the rate of differentiation of cells bearing phenotype-altering mutations, as shown for multiple oncogenic mutations engineered in HSC in mice [47]. In this case regardless of the effect of oncogenic mutations on cell cycle, differentiation rates above 0.5 per cell division will always lead to the elimination of the mutant clone from the pool. For example, deletion of the PTEN gene substantially reduces the self-renewal of mouse HSC [48, 49]; while deletion of PTEN simultaneously in all HSC leads to leukemia, the more natural context of gene disruption in a single HSC would be expected to result in clonal exhaustion by differentiation. Evidence from the human hematopoietic system also shows that during early life (until the age of 40) clonality above a few percent of the pool is practically non-existent, regardless of the fact that roughly half of all mutations that could accumulate in a full lifetime have already occurred by the age of ~18 years [15-18]. This suggests that positive selection for particular mutant phenotypes in general only has a major presence in late life during post-reproductive ages.

Concluding remarks
The fate of somatic mutations (including oncogenic ones) should be governed by the same basic evolutionary principles that govern natural populations, whereby adaptation leads to higher fitness and increased stabilizing and purifying selection, while positive selection is driven by environmental change and lower population fitness [20, 47]. From this perspective, carcinogenesis should be viewed as a function of physiological aging, whereby aging and altered tissue microenvironments lead to selection for previously accumulated random mutations, some of which confer a fitness advantage under these altered conditions. This framework explains many previously unexplained phenomena, such as the rarity of cancers among wild animals (which do not typically live beyond their reproductive ages), the scaling of cancer incidence curve to lifespan, and the fact that longer-lived and larger animals do not have proportionally more cancers. It is also easier to explain from this perspective why cancers that require different numbers of oncogenic mutations and originate from tissues with different numbers of stem cells and cell division profiles all follow a similar age-dependent timing of incidence increase, including cancers that appear to require just one mutation, such as chronic myeloid leukemia. In this regard, while solid evidence supports the general paradigm that carcinogenesis is a Darwinian multi-stage process of rounds of mutation and positive selection, the current paradigm oversimplifies the evolutionary forces that drive these selection steps and determine their timing. It thus lacks key theoretical foundations of modern evolutionary theory and requires revision (see Outstanding Questions Box). Understanding somatic evolution in the context of the evolution of animal life histories and from a tissue-level ecological perspective provides immense explanatory power to better understand a very complex disease: cancer.

Acknowledgments:

We would like to thank Hannah Scarborough, Roberta Pelanda, Eric Pietras and Richard Davis of the University of Colorado for critical review of the manuscript. These studies were supported by National Cancer Institute grants (R01CA180175 and R21CA179501) to J.D.

References:

45. Rozhok, A.I. et al. (2014) Stochastic modeling indicates that aging and somatic evolution in the hematopoietic system are driven by non-cell-autonomous processes. Aging 6 (12), e1-16.

**Figure 1. Age-dependent physiological events.** Roughly half of all mutations and epigenetic changes occur before full body maturation [7]. The accumulation of genetic and epigenetic changes significantly slows down later, followed by a slowdown in stem cell division rates when stem cells transition from body maturation to body maintenance (example based on HSC) [9]. Both elevated aging rates and cancer incidence are significantly delayed and do not follow the mutation accumulation curve. However, cancer incidence increase mirrors the curve of physiological decline (aging).

**Figure 2. The evolution of lifespan and cancer incidence.** (A) The lifespan of animals in the wild (dark green curve) is usually defined by external hazards, such as predators, food availability, and infectious diseases (dark green arrows). Their maximum lifespan, limited by their rates of physiological aging (blue line), can only be observed in captivity or in modern humans.
Germline selection acts to maintain high body fitness (solid blue arrows) for the duration of likely survival and reproduction in the wild. Physiological aging in humans and captive animals accelerates after the time of their likely survival in the wild promoted by progressively relaxing germline selection for body fitness. (B) If external hazards are reduced, animals survive into more advanced ages in the wild and trigger germline selection to extend (hollow blue arrows) their maximum lifespan, pushing the curve of physiological aging rightwards on the age axis and leading to the evolution of a longer potential lifespan. (C) Cancer is rare in wild animals, as well as in humans and captive animals during the period of their likely survival in the wild (natural lifespan), with carcinogenesis being suppressed by higher tissue and stem cell fitness relative to ages past natural lifespan. A dramatic increase in cancer incidence coincides with higher rates of physiological decline.

**Figure 3. The effect of fitness landscapes on selection processes.** (A) A cellular phenotype (red dot) is driven by the evolution at the germline level to adapt to the tissue microenvironmental fitness landscape (represented here by the simplest Shelford curve of tolerance model [50]), moving the phenotype towards the fitness peak. The strength of selection triggered by a mutation is proportional to the fitness difference the mutation creates. Germline selection reduces the number of mutations capable of increasing fitness (green area), limits the strength of positive selection that such a mutation can confer (compare the fitness differential between wild type versus A-mutant and wild type versus the negatively selected B-mutant), and increases the number of negatively selected mutations (orange area). (B) In an altered environment, the intensity of many environmental factors is changed (represented here with a rightward-shifted Shelford curve). In an altered environment to which a cell is not adapted through germline evolution, the same cellular phenotype results in lower cellular fitness. Also, the number of mutations capable of elevating fitness is greater and the strength of positive selection they can trigger is increased (compare fitness differential between wild type versus A-mutant and wild type versus B-mutant with those in panel A). The directionality of selection triggered by a specific mutation can also change, such as represented here by mutation B which is negatively selected in the normal microenvironment, but strongly positively selected in the altered one.

**Figure 4. Rates of aging and cancer incidence in DNA polymerase mutant mice.** Mutations L604K and L604G in the heterozygous state both increase mutations rates more than 5 times in mice. Mutation L604G does not have any effect on lifespan and cancer incidence. Mutation L604K, however, causes a shortening of the lifespan and a symmetrical acceleration of the cancer incidence curve relative to +/- (wild type) and +/-L604G [34].
Glossary

**Adaptive** – Increases fitness (e.g. a mutation that increases fitness is adaptive).

**Allele** – a variant of a gene. Different alleles for the same gene will be largely homologous, but can modestly differ in sequence and can encode for proteins with different activities.

**Drift** – random assortment of genetic alleles. The influence of drift becomes greater as population size decreases.

**Fitness** – a measure of reproductive success; classically, the ability of an organism to pass its genes on to future generations of that organism. For somatic tissues, a measure of the ability of a cell to pass its genotype or epigenotype on to future cell generations.

**Germline** – the genetic lineage that is maintained in the next generation following reproduction.

**Lifespan evolution** – changes in somatic maintenance programs that are selected for in different environments to maximize reproductive success (part of Life History Theory).

**Microenvironment** – the tissue environment for somatic cells.

**Phenotype** – the characteristics of an organism or a cell that are determined by its genotype.

**Selection** – the process of differential survival of individuals in a population that impacts on allele frequency based on their fitness; depends on genetic and environmental contexts.

  - **Positive selection** – increases the frequency of alleles that increase fitness
  - **Purifying (or negative) selection** – decreases the frequency of alleles that reduce fitness
  - **Stabilizing selection** – maintenance of a well-adapted phenotype, as mutations will mostly reduce fitness

**Somatic** – cells and tissues of an organism (the soma), excluding gametes (e.g. sperm and eggs)

**Somatic maintenance** – the maintenance of cells, tissues and organs through replacement and repair.
While the increasing cancer incidence with age is currently understood from the perspective of time-dependent mutation accumulation, this model fails to account for key observations, such as early life mutation accumulation and the scaling of cancer incidence to the lifespans of different animals.

Animal evolution has selected for somatic maintenance strategies that maximize reproductive success, and these strategies determine the evolution of maximal lifespans.

External conditions, including food availability, disease and predators, influence somatic maintenance strategies, as higher odds of earlier death will promote strategies that reduce long-term somatic maintenance and thus lead to earlier senescence when removed from these hazards.

Somatic maintenance strategies have a major impact on cancer incidence, as well-maintained tissues in young animals promote stabilizing selection (disfavoring oncogenic mutations), while tissue decline in post-reproductive periods promotes positive selection for adaptive mutations (including oncogenic ones).
Outstanding Questions Box

- Would expanding field data to different groups of mammals and other vertebrates significantly enhance our understanding for how lifespan and age-related disease risk, and the risk of cancer in particular, are linked?
- How do the processes of aging determine the strength and directionality of selection driven by oncogenic mutations?
- What are the key changes in tissue microenvironments with old age (or in smokers, etc.) that affect the fate of somatic mutations? Can strategies be developed that restore more youthful microenvironments, thus dampening selection for oncogenic events?
A. Adaptive landscape evolved through germline evolution

B. Adaptive landscape altered by tissue aging or insults