Context-dependent selection as the keystone in the somatic evolution of cancer

Vibishan B.[1,2] **and Milind G. Watve**[1,2,*]

[1] Department of Biology, Indian Institute of Science Education and Research (IISER), Pune

This manuscript was compiled on November 14, 2018

Please provide an abstract of no more than 250 words in a single paragraph. Abstracts should explain to the general reader the major contributions of the article. References in the abstract must be cited in full within the abstract itself and cited in the text.

Somatic evolution | Mutation accumulation | Epidemiology | Cancer etiology |

Over the past 60 years or so, ideas in the field of cancer epidemiology have evolved significantly, with the first models starting from the Armitage-Doll multi-stage models (1), which were essentially statistical fits to available data of age-specific incidence, and predicted a power law relationship of cancer risk with age. The connection between the multiple stages and sequential genetic mutation events was established definitively for retinoblastoma with the two-hit hypothesis (2). This finding has since directed a lot of attention to mutational processes and genetic instability within cells as fundamental forces in cancer, and their signatures in population-level datasets. Tomasetti et al. have made the argument for cancer risk being largely determined by random mutations (3, 4). Others have studied the impact of selectively-neutral or deleterious passenger mutations on the expansion and progression of advantageous mutant clones (5), mutation accumulation rates across tissue types (6), or dependencies between mutations (7). On the other hand, an old debate in the theory of evolution is how the simple process of random mutations and natural selection can lead to compled structures such as the eye, that need coordinated action of several genes. This is often perceived as a monkey-on-a-typewriter paradox (8, 9); how likely is it that a monkey sitting at a typewriter and hitting keys at random would end up typing a meaningful sentence? The problem of cancer is qualitatively similar to this, but quantitatively even more difficult. No single mutation is known to make a cell cancerous. All cancers are necessarily a combination of different types of genomic changes including point mutations, aneuploidy, and other chromosomal aberrations. The cancer phenotype has a large number of distinguishing characeters, encapsulated by the notion of the "hallmarks" of cancer (10–12), and the wide range of these characterisitics that the hallmarks include make it astonishing that so many alterations in cell properties come together in cancers purely out of chance, especially since most cancers must evolve independently in each individual organism.

Within the level of the organism, clonal expansion is another process that is implicated in carcinogenesis. Every component mutation on the way to a cancerous phenotype causes the mutant clone to expand, and as the mutant population increases, the probability of a second component mutation increases proportionately (13, 14). Implicit in this theory is the assumption that every component mutation has a selective advantage over the normal cell. Since most changes involved in carcinogenesis relate to evading growth regulatory mechanisms, it is considered logical that any mutation that allows for such evasion will have a natural selective advantage. However, the final word is far from conclusive on the view that cells with component mutations are always at an advantage, and evidence has been accumulating over the past few years that the fitness advantage of a mutant is largely dependent on the tissue micro-environment (15, 16). Studies in mice (17) and humans (18) have demonstrated the effect of contingent factors, such as behavioural profiles and lifestyle parameters, on cancer progression. Such findings provide clear indications that the selective forces which determine mutant clone fitness can vary considerably across individual organisms, leading to *context-dependent clonal expansion* of potentially oncogenic mutants. This aspect of cancer progression, that moves beyond mutagenesis, has also seen some modeling effort in the past few years, towards attributing cancer risk to environmental factors (19), or other forms of selection (20–22). These efforts reveal the incorporation of additional complexity to causal factors underlying cancer, although a thoroughly unified picture of cancer etiology is still emerging. So far however, what we identify here as context-dependent clonal expansion has not been incorporated explicitly in models of cancer incidence.

Across biological levels then, there at least three processes that seem to play a major role in cancer progression and etiology: (1) random mutagenesis, or the "bad luck" hypothesis, (2) expansion of mutant clones within organisms, and (3) context-dependent selection acting across organisms. Since well-curated data is available for human cancer incidence patterns, we develop models of these three processes, and compare their predictions with the epidemiological picture of cancer in the human population. This also includes an examination of how these different models explain the well-known Peto's paradox (23, 24), and the observed relationships with stem cell number (3, 4) and mutation rates (25).

Significance Statement

Authors must submit a 120-word maximum statement about the significance of their research paper written at a level understandable to an undergraduate educated scientist outside their field of speciality. The primary goal of the Significance Statement is to explain the relevance of the work in broad context to a broad readership. The Significance Statement appears in the paper itself and is required for all research papers.

The authors declare that there are no conflicts of interest

[2] V.B. and M.G.W. contributed equally to this work.

[*] To whom correspondence should be addressed. E-mail: milind@iiserpune.ac.in

The "bad luck" model

This hypothesis assumes that the required set of driver mutations accumulate in a cell by chance alone. This may happen over a period of time, or in a single large-scale event, like chromothripsis (26).

Consider an organism with a population of n stem cells, each with a mutation rate per cellular generation per genome, p. The probability that at least one cell acquires one mutation at a given point of time can be given as $1-(1-p)^n$. If k such mutations are requried for cancer onset, the probability of cancer according to the bad luck model can be given as below, based on an algebraic formulation (27):

$$p_{can} = 1-(1-p^k)^n \qquad [1]$$

For the sake of simplicity, we have ignored the cost of lethal and/or passenger mutations, and assume that the mutations occur together at any given point of time; although the model predictions change qualitatively, we see that the latter assumption does not affect the general inferences we draw.

On the whole, the algebraic form of 1 predicts a threshold relationship of p_{can} with both n and p;

Author Affiliations. Include department, institution, and complete address, with the ZIP/postal code, for each author. Use lower case letters to match authors with institutions, as shown in the example. Authors with an ORCID ID may supply this information at submission.

Submitting Manuscripts. All authors must submit their articles at PNAScentral. If you are using Overleaf to write your article, you can use the "Submit to PNAS" option in the top bar of the editor window.

Format. Many authors find it useful to organize their manuscripts with the following order of sections; Title, Author Affiliation, Keywords, Abstract, Significance Statement, Results, Discussion, Materials and methods, Acknowledgments, and References. Other orders and headings are permitted.

Manuscript Length. PNAS generally uses a two-column format averaging 67 characters, including spaces, per line. The maximum length of a Direct Submission research article is six pages and a Direct Submission Plus research article is ten pages including all text, spaces, and the number of characters displaced by figures, tables, and equations. When submitting tables, figures, and/or equations in addition to text, keep the text for your manuscript under 39,000 characters (including spaces) for Direct Submissions and 72,000 characters (including spaces) for Direct Submission Plus.

References. References should be cited in numerical order as they appear in text; this will be done automatically via bibtex, e.g.. All references should be included in the main manuscript file.

Data Archival. PNAS must be able to archive the data essential to a published article. Where such archiving is not possible, deposition of data in public databases, such as GenBank, ArrayExpress, Protein Data Bank, Unidata, and others outlined in the Information for Authors, is acceptable.

Table 1. Comparison of the fitted potential energy surfaces and ab initio benchmark electronic energy calculations

Species	CBS	CV	G3
1. Acetaldehyde	0.0	0.0	0.0
2. Vinyl alcohol	9.1	9.6	13.5
3. Hydroxyethylidene	50.8	51.2	54.0

nomenclature for the TSs refers to the numbered species in the table.

Language-Editing Services. Prior to submission, authors who believe their manuscripts would benefit from professional editing are encouraged to use a language-editing service (see list at www.pnas.org/site/authors/language-editing.xhtml). PNAS does not take responsibility for or endorse these services, and their use has no bearing on acceptance of a manuscript for publication.

Digital Figures. Only TIFF, EPS, and high-resolution PDF for Mac or PC are allowed for figures that will appear in the main text, and images must be final size. Authors may submit U3D or PRC files for 3D images; these must be accompanied by 2D representations in TIFF, EPS, or high-resolution PDF format. Color images must be in RGB (red, green, blue) mode. Include the font files for any text.

Figures and Tables should be labelled and referenced in the standard way using the `\label{}` and `\ref{}` commands.

Figure **??** shows an example of how to insert a column-wide figure. To insert a figure wider than one column, please use the `\begin{figure*}...\end{figure*}` environment. Figures wider than one column should be sized to 11.4 cm or 17.8 cm wide. Use `\begin{SCfigure*}...\end{SCfigure*}` for a wide figure with side captions.

Tables. In addition to including your tables within this manuscript file, PNAS requires that each table be uploaded to the submission separately as a âTableâĪ file. Please ensure that each table .tex file contains a preamble, the `\begin{document}` command, and the `\end{document}` command. This is necessary so that the submission system can convert each file to PDF.

Single column equations. Authors may use 1- or 2-column equations in their article, according to their preference.

To allow an equation to span both columns, use the `\begin{figure*}...\end{figure*}` environment mentioned above for figures.

Note that the use of the `widetext` environment for equations is not recommended, and should not be used.

Supporting Information (SI). Authors should submit SI as a single separate PDF file, combining all text, figures, tables, movie legends, and SI references. PNAS will publish SI uncomposed, as the authors have provided it. Additional details can be found here: policy on SI. For SI formatting instructions click here. The PNAS Overleaf SI template can be found here. Refer to the SI Appendix in the manuscript at an appropriate point in the text. Number supporting figures and tables starting with S1, S2, etc.

Authors who place detailed materials and methods in an SI Appendix must provide sufficient detail in the main text methods to enable a reader to follow the logic of the procedures

and results and also must reference the SI methods. If a paper is fundamentally a study of a new method or technique, then the methods must be described completely in the main text.

SI Datasets. Supply Excel (.xls), RTF, or PDF files. This file type will be published in raw format and will not be edited or composed.

SI Movies. Supply Audio Video Interleave (avi), Quicktime (mov), Windows Media (wmv), animated GIF (gif), or MPEG files and submit a brief legend for each movie in a Word or RTF file. All movies should be submitted at the desired reproduction size and length. Movies should be no more than 10 MB in size.

3D Figures. Supply a composable U3D or PRC file so that it may be edited and composed. Authors may submit a PDF file but please note it will be published in raw format and will not be edited or composed.

Materials and Methods

Please describe your materials and methods here. This can be more than one paragraph, and may contain subsections and equations as required. Authors should include a statement in the methods section describing how readers will be able to access the data in the paper.

Subsection for Method. Example text for subsection.

ACKNOWLEDGMENTS. Please include your acknowledgments here, set in a single paragraph. Please do not include any acknowledgments in the Supporting Information, or anywhere else in the manuscript.

1. Armitage P, Doll R (1954) The Age Distribution of Cancer and a Multi-Stage Theory of Carcinogenesis. *British journal of cancer* 8(1):1–12.
2. Knudson AG (1971) Mutation and Cancer: Statistical Study of Retinoblastoma. *Proceedings of the National Academy of Sciences* 68(4):820–823.
3. Tomasetti C, Vogelstein B (2015) Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* 347(6217):78–81.
4. Tomasetti C, Li L, Vogelstein B (2017) Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* 355(6331):1330–1334.
5. McFarland CD, Korolev KS, Kryukov GV, Sunyaev SR, Mirny LA (2013) Impact of deleterious passenger mutations on cancer progression. *Proceedings of the National Academy of Sciences* 110(8):2910–2915.
6. Blokzijl F, et al. (2016) Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* 538(7624):260–264.
7. Mina M, et al. (2017) Conditional Selection of Genomic Alterations Dictates Cancer Evolution and Oncogenic Dependencies. *Cancer Cell* 32(2):155–168.e6.
8. Dawkins R (1996) *The blind watchmaker : why the evidence of evolution reveals a universe without design.* (W. W. Norton & Company), pp. xvii, 358.
9. Garvin MR, Gharrett AJ (2014) Evolution: Are the monkeys' typewriters rigged? *Royal Society Open Science* 1(2):140172–140172.
10. Hanahan D, Weinberg R (2000) The hallmarks of cancer. *Cell* 100(1):57–70.
11. Schäfer M, Werner S (2008) Cancer as an overhealing wound: An old hypothesis revisited. *Nature Reviews Molecular Cell Biology* 9(8):628–638.
12. Hanahan D, Weinberg R (2011) Hallmarks of Cancer: The Next Generation. *Cell* 144(5):646–674.
13. Nowell P (1976) The clonal evolution of tumor cell populations. *Science* 194(4260):23–28.
14. Aktipis CA, Nesse RM (2013) Evolutionary foundations for cancer biology. *Evolutionary Applications* 6(1):144–159.
15. Hanahan D, Coussens LM (2012). *Cancer Cell* 21(3):309–322.
16. Pietras K, Östman A (2010) Hallmarks of cancer: Interactions with the tumor stroma. *Experimental Cell Research* 316(8):1324–1331.
17. Cao L, et al. (2010) Environmental and Genetic Activation of a Brain-Adipocyte BDNF/Leptin Axis Causes Cancer Remission and Inhibition. *Cell* 142(1):52–64.
18. Rundqvist H, et al. (2013) Effect of Acute Exercise on Prostate Cancer Cell Growth. *PLoS ONE* 8(7):e67579.
19. Hochberg ME, Noble RJ (2017) A framework for how environment contributes to cancer risk. *Ecology Letters* 20(2):117–134.
20. Nagy JD, Victor EM, Cropper JH (2007) Why don't all whales have cancer? A novel hypothesis resolving Peto's paradox. *Integrative and Comparative Biology* 47(2):317–328.
21. Caulin AF, Maley CC (2011) Peto's Paradox: Evolution's prescription for cancer prevention.
22. Tollis M, Boddy AM, Maley CC (2017) Peto's Paradox: How has evolution solved the problem of cancer prevention? *BMC Biology* 15(1):60.
23. Peto H, Roe FJ, Lee PN, Levy L, Clack J (1975) Cancer and ageing in mice and men. *British Journal of Cancer* 32(4):411–426.
24. Noble R, Kaltz O, Hochberg ME (2015) Peto's paradox and human cancers. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370(1673):20150104.
25. Hao D, Wang L, Di Lj (2016) Distinct mutation accumulation rates among tissues determine the variation in cancer risk. *Nature Publishing Group* (January):1–5.
26. Stephens PJ, et al. (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 144(1):27–40.
27. Calabrese P, Shibata D (2010) A simple algebraic cancer equation: calculating how cancers may arise with normal mutation rates. *BMC Cancer* 10(1):3.