Quantification of the selective advantage of driver mutations is dependent on the underlying model and stage of tumor evolution

Ivana Bozic[1,2,*]

[1] Department of Applied Mathematics, University of Washington, Seattle, WA

[2] Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA

Running Title: Selective advantage of driver mutations

*Corresponding author

Mailing address:

Department of Applied Mathematics

University of Washington

Lewis Hall 201, Box 353925

Seattle, WA 98195

Email: ibozic@uw.edu

Authors' Disclosures: I. Bozic reports a grant from National Science Foundation during the conduct of the study; grant from Johnson & Johnson outside the submitted work; in addition, Dr Bozic has a patent for Use of clonal evolution analysis for ibrutinib resistance in chronic lymphocytic leukemia patients pending.

Abstract

Measuring the selective fitness advantages provided by driver mutations has the potential to facilitate a precise quantitative understanding of cancer evolution. However, accurately measuring the selective advantage of driver mutations has remained a challenge in the field. Early studies reported small selective advantages of drivers, on the order of 1%, whereas newer studies report much larger selective advantages, as high as 1200%. In this article, we argue that the calculated selective advantages of cancer drivers are dependent of the underlying mathematical model and stage of cancer evolution and that comparisons of numerical values of selective advantage without regard for the underlying model and stage can lead to spurious conclusions.

Introduction

Detecting and quantifying selection in cancer is one of the key goals of cancer genomics, as positively selected driver gene mutations present viable targets for development of novel cancer therapies[1]. Measuring the selective growth advantage provided by driver mutations could lead to a precise quantitative understanding of cancer evolution, allow for prognostication of cancer progression in patients and facilitate precision oncology.

Early works[2] reported small average selective advantage of driver mutations in pancreatic cancer and glioblastoma, on the order of 1%, with similar estimates obtained more recently for driver mutations in seemingly normal oesophagus and skin[3]. On the other hand, some newer works

report much larger selective advantages of driver mutations, on the order of 20% to 80% in multiple cancer types[4] and recently as high as 1200% in pancreatic cancer and even higher in thyroid and liver hepatocellular carcinoma[5]. Are the early estimates wrong? Are the new estimates too high? We demonstrate here that selective advantage of cancer drivers depends on the underlying model and stage of tumor evolution, and that comparison of numerical values of selective advantage across studies should be taken with care and regard to these issues. We will focus here mostly on the example of the *KRAS* oncogene in colorectal cancer (CRC), and show that its selective advantage can vary by two orders of magnitude, depending on the stage of colorectal tumor evolution in which the driver mutation is acquired. We will also demonstrate that different definitions of selective advantage across various studies can lead to an equally large variation in the reported selective advantage of driver mutations.

Selective Advantage Provided to Stem Cells

Initial stages of tumor evolution in many tissues are thought to occur in a stem cell compartment, such as within stem cells at the base of an intestinal crypt (Fig. 1a, top). The evolutionary dynamics of driver mutations in intestinal crypts were studied by Vermeulen and colleagues[6], who used experimental visualization of stem cell clones together with mathematical modeling to quantify the competitive advantage of common CRC driver mutations. They presented a model of stem cell dynamics in a crypt which can be thought of as a ring of N stem cells replacing their neighbors in a random fashion[6]. In this model, a wild-type stem cell has 50% chance of replacing a neighboring wild-type stem cell[6]. On the other hand, a stem cell with a mutation in *Kras* has been shown to have a 78% chance of replacing an adjacent wild-type stem cell (and 22% chance

of being replaced by a wild-type stem cell) [6]. In this setting, selective advantage of *Kras* can be defined as the relative increase in the chance of being chosen to replace a neighboring wild-type stem cell (rather than vice versa) conferred by the driver. Thus, selective advantage that *Kras* provides to a colorectal stem cell is equal to $s = 0.78/0.22 - 1 = 255\%$.

Selective Advantage Provided to Crypts

Oncogenic mutations lead not only to expansion of mutant stem cells within a single crypt, but also to increased levels of crypt fission[7] (Fig. 1a, bottom). Wild type crypt fission (division) rate in healthy human colonic tissue is measured to be 0.007/year[7], balanced by the rate of crypt fusion[8]. Division rate of *KRAS*-mutant crypts is increased to 0.07 per year[7], ten times higher than for wild-type crypts. Various models define selective advantage of cancer drivers as the relative decrease in death rate[2], relative increase in division rate[5] of relative increase in the net growth rate[4] (division - death). Defining selective advantage of *KRAS* as the relative increase in net growth rate of *KRAS*-mutant crypts over wild-type crypts would lead to an infinite selective advantage, as the net growth rate of wild type crypts is practically 0 in adult tissue. Defining selective advantage as the relative increase in crypt division (fission) rate would lead to selective advantage of *KRAS* of 900%.

Selective Advantage as Second Driver

Furthermore, *KRAS* often occurs as a second driver mutation on the way to CRC, following *APC* inactivation. It has been recently estimated[10] that *KRAS* typically increases the division (fission)

rate of colorectal crypts that have previously inactivated *APC* by 35%. Lahouel et al. recently reported an even smaller estimate for relative increase in division rate provided by *KRAS* as the second driver in colorectal cancer[5], of up to 15%. The examples above indicate that when *KRAS* is activated in previously normal colorectal stem cells, it provides selective advantage to the stem cells and colorectal crypts that is on the order of 100%-1000%. However, when it appears as the second driver in CRC, the typical selective advantage it provides is on the order of 10%.

We focused on the example of the *KRAS* mutation acquired during various stages of colorectal tumor evolution. Even though *KRAS* is typically thought to be the second driver mutation acquired on the way to CRC, it has been shown to also provide selective growth advantage at the earliest stages of colorectal tumor evolution[6,7], as discussed in the previous sections. This may seem in contrast with experimental findings that individual *KRAS*-mutated lesions are unlikely to progress to CRC[9]. Indeed, while the relative increase in the fission rate of *KRAS*-mutated crypts is large, their absolute fission rate remains fairly small (0.07 per year[7], corresponding to a doubling time of ~10 years). Thus, an individual *KRAS*- mutated lesion is expected to remain microscopic, and unlikely to obtain subsequent driver mutations necessary for malignant transformation within a human lifetime. However, recent work[10] helps quantify the expected number of *KRAS*- mutated lesions in the entire human colon and demonstrates that, due to the high estimated number of such lesions, *KRAS* may be the first driver mutation in CRC evolution in up to a third of CRC patients whose cancers harbor the mutation.

Selective Advantage Dependent on Model Formulation

When discussing selective advantage of cancer drivers, it is important to note that many mathematical models of cancer evolution disregard tissue hierarchy[2,4] and include populations of cells that can divide and die with some rates[2,4] (Fig. 1b). Such models are more suitable for later stages of cancer progression, when cancer cells are poorly differentiated or dedifferentiated, or for early stages of tumor evolution in tissues that are maintained by a single type of progenitor cells (such as skin[11]) or well-mixed stem cells (e.g. hematopoietic[12]).

Healthy tissues are at homeostasis, and have balanced overall division and death rates ($b = d$), leading to a net growth rate of 0. For simplicity we can assume that $b = d = 1$ (per some unit time). An oncogenic mutation can decrease cell death rate by $s = 1\%$ ($d_1 = d * (1 - s) = 1 - s$), leading to a clonal expansion with net growth rate equal to $b - d_1 = s$. This small decrease in the death rate[2] of 1% thus leads to a relative increase in the net growth rate[4] over homeostatic tissue that is infinite (as healthy tissue net growth rate is 0). A second mutation with the same decrease in death rate of $s = 1\%$, will lead to a clonal expansion with net growth rate $2s$, which represents 100% relative increase in net growth rate over a single mutation. A third mutation with the same decrease in the death rate of 1% would correspond to the relative increase in the growth rate of 50% and so on. The seeming discrepancy between small selective advantage (on the order of 1%) reported previously for driver mutations in pancreatic cancer and glioblastoma[2] versus the much larger advantages (20-80%) reported subsequently for late drivers in multiple cancer types[4] can be understood to stem from the difference in the models they were inferred with. The former

model defined selective advantage as the relative decrease in the death rate of cells due to driver mutations, while the latter defined selective advantage as the relative increase in the net growth rate. Both studies[2,4] assumed a well-mixed population of cells without tissue hierarchy.

The above examples demonstrate that the same driver mutation can have vastly different numerical values of selective advantage at different stages of tumor evolution and when using different underlying models, and that selective advantage of 1% defined in one way (e.g. as the relative decrease in the death rate) can correspond to virtually any value of selective advantage if it is defined as the relative increase in the net growth rate. Similarly, some model settings do not allow selective advantages greater than 100%. This occurs, for example, when selective advantage is defined as the relative decrease in the death rate of cells provided by a driver mutation[2]; or when selective advantage is defined as an increase in proliferative bias of stem cells towards symmetric renewal[3]. In contrast, in other settings selective advantage is theoretically unbounded, such as when drivers increase division rate[5] or net growth rate[4].

It is important to note that the confusion in reporting and comparing the magnitude of selective advantage across studies often arises when increases in fitness (or a similar quantity) are expressed in relative rather than absolute terms. Furthermore, knowledge of selective advantage expressed as a single number is insufficient for the full quantitative understanding of the population dynamics, which requires knowledge of parameters governing the growth of both populations, wild-type and mutant. The easiest way to avoid confusion is to report absolute growth rates (or fitness) of both wild-type and mutant populations.

Other Sources of Variation

Tissue

In addition to the stage of tumor evolution and the specific model employed, selective advantage of a driver mutation will also depend on the tissue in question[5]. Some drivers are frequently mutated in cancers of a single or a small number of tissues, while being infrequently mutated in most other cancer types, which may suggest that such mutations provide significant selective advantage only in specific tissues. An example of such a driver is the tumor suppressor *APC*, which is mutated in more than 80% of colorectal cancers, but is infrequently mutated in most other cancer types. *APC* has been shown to provide a very large selective advantage to previously normal colorectal crypts: the fission rate of crypts that have inactivated *APC* is roughly three times larger than the fission rate of crypts with activated *KRAS*[7,10], and it is likely that *APC* and *KRAS* are the two drivers that provide the largest advantage to previously normal colorectal crypts. The large selective advantage provided to colorectal crypts by *APC* inactivation seems to be behind both its frequency in CRC and the observation that it tends to be the first driver mutation in the majority of CRCs[10].

On the other hand, some driver genes seem to be mutated in many cancer types. *KRAS* and *TP53* are examples of such drivers. For example, Lahouel et al. have recently estimated that *KRAS* increases the division rate of pancreatic stem cells by 1200%, where it is typically the first driver

mutation on the way to pancreatic cancer. This selective advantage is of a similar order of magnitude as when *KRAS* is the first driver in CRC.

TP53 tumor suppressor gene is the most commonly mutated gene in human cancer. Does that imply that it confers a large selective advantage? The answer again depends on the stage of tumor evolution. While *TP53* does not seem to provide selective advantage as the first driver in CRC in normal conditions[6], it is crucial for transformation from benign to a malignant tumor, thus providing important advantage as a later driver. *TP53* is also typically a late driver in chronic lymphocytic leukemia (CLL), where it has been shown to provide a large growth advantage as a late, subclonal driver[13].

Metastasis, the dissemination and subsequent growth of tumor cells in other tissues, is also expected to play a crucial role in shaping selection, as disseminated cells must survive and proliferate in new environments. Recent work demonstrates that different metastatic sites in colorectal cancer (lymph node versus distant) are subject to different levels of selection, with stronger selection associated with formation of distant metastases[14]. Hu and colleagues report relaxed selective pressures in metastases relative to early cancer development in colorectal and breast cancer, but not in lung cancer[15].

Selective fitness advantage of a driver mutation also depends on possible competition or interaction between the populations of cells with and without the driver. While quantitative data on competition between different clonal populations is scarce in human tumor settings, there is

recent evidence that intestinal crypts carrying oncogenic mutations outcompete neighboring normal stem cells by secreting factors that promote their dedifferentiation[16].

Microenvironment

Tumor suppressor *TP53* is also a prime example of a driver gene that provides different selective advantage under different microenvironmental conditions. Vermeulen et al. showed that while mutated *TP53* does not provide advantage to previously wild-type colorectal stem cells under normal conditions, it does provide significant selective advantage when the microenvironment shows signs of inflammation, such as in colitis[6]. These findings are similar to those of Klein et al. who only find advantage for *TP53* mutation in irradiated skin, and not under normal conditions[11]. Virus infections such as HPV and mutational processes such as APOBEC mutagenesis have also been shown to affect selective advantage of driver mutations[17].

Aging is another process that can lead to significant changes in the tissue and tumor microenvironment, and which can have an effect on selective advantage of driver mutations. For example, there is experimental evidence that the same oncogenic mutations that are not selected in young organisms can become selected in older animals[18]. Immune cells represent an important part of the tumor microenvironment, potentially leading to negative selection of certain tumor clones (immunoediting) and exerting selective pressure on tumors to evolve immune-escape mechanisms[19].

Finally, therapy typically radically alters the fitness landscape of cancer. This is particularly evident in the case of targeted therapies, which often succeed in eradicating many if not most tumor subpopulations, but ultimately fail due to survival and proliferation of treatment-resistant cells, which often carry well-defined resistance mutations[20,21].

Conclusion

Quantifying parameters of cancer evolution in individual patients, including selective growth advantages of individual and combinatorial driver lesions, holds great promise for prognosticating individual tumor trajectories, which could facilitate personalized treatment selection and precision oncology. However, care must be taken that, when dealing with estimates of selective growth advantages of driver mutations, the underlying model is clearly stated, as well as the stage of tumor evolution that it applies to, which includes any genetic background upon which the driver mutation may appear. Otherwise, comparison of values of selective advantage across studies without careful consideration of these issues can lead to unnecessary confusion and unjustified conclusions.

Acknowledgements

The author thanks Bert Vogelstein for careful reading and comments on the manuscript. This work is supported by the National Science Foundation grant DMS-2045166.

References

1. Rabadan R., Mohamedi Y, Rubin U, Chu T, Alghalith AN, Elliott O, et al. Identification of relevant genetic alterations in cancer using topological data analysis. *Nat Commun* 2020;**11**:3808.
2. Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, Chen S, et al. Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci USA* 2010;**107**:18545-18550.
3. Williams MJ, Zapata L, Werner B, Barnes CP, Sottoriva A, Graham TA. Measuring the distribution of fitness effects in somatic evolution by combining clonal dynamics with dN/dS ratios. *eLife* 2020;**9**:e48714.
4. Williams MJ, Werner B, Heide T, Curtis C, Barnes CP, Sottoriva A, et al. Quantification of subclonal selection in cancer from bulk sequencing data. *Nat Genet* 2018;**50**:895-903.
5. Lahouel K, Younes L, Danilova L, Giardiello FM, Hruban RH, Groopman J, et al. Revisiting the tumorigenesis timeline with a data-driven generative model. *Proc Natl Acad Sci USA* 2020;**117**:857-864.
6. Vermeulen L, Morrissey E, van der Heijden M, Nicholson AM, Sottoriva A, Buczacki S, et al. Defining stem cell dynamics in models of intestinal tumor initiation. *Science* 2013;**342**:995-998.
7. Nicholson AM, Olpe C, Hoyle A, Thorsen A, Rus T, Colombe M, et al. Fixation and spread of somatic mutations in adult human colonic epithelium. *Cell Stem Cell* 2018;**22**: 909-918.e8.
8. Baker A, Gabbutt C, Williams MJ, Cereser B, Jawad N, Rodriguez-Justo M, et al. Crypt fusion as a homeostatic mechanism in the human colon. *Gut* 2019;**68**:1986-1993.

9. Kinzler KW, Vogelstein B. Lessons from hereditary colon cancer. *Cell* 1996;**87**:159-170.
10. Paterson C, Clevers H, Bozic I. Mathematical model of colorectal cancer initiation. *Proc Natl Acad Sci USA* 2020;**117**:20681-20688.
11. Klein AM, Brash DE, Jones PH, Simons BD. Stochastic fate of p53-mutant epidermal progenitor cells is tilted toward proliferation by UV B during neoplasia. *Proc Natl Acad Sci USA* 2010;**107**:270-275.
12. Watson CJ, Papula AL, Poon GYP, Wong WH, Young AL, Druley TE, et al. The evolutionary dynamics and fitness landscape of clonal hematopoiesis. *Science* 2020;**367**:1449-1454.
13. Gruber M, Bozic, I, Leshchiner I, Livitz D, Stevenson K, Rassenti L, et al. Growth dynamics in naturally progressing chronic lymphocytic leukemia. *Nature* 2019;**57**:474-479.
14. Reiter JG, Hung W, Lee I, Nagpal S, Giunta P, Degner S, et al. Lymph node metastases develop through a wider evolutionary bottleneck than distant metastases. *Nat Genet* 2020; **52**:692-700.
15. Hu Z, Li Z, Ma Z, Curtis C. Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nat Genet* 2020;**52**:701-708.
16. Chia S, DeGregori J. Cancer stem cells in the gut have a bad influence on neighbouring cells. *Nature* 2021;**594**:340-341
17. Cannataro VL, Gaffney SG, Sasaki T, Issaeva N, Grewal NKS, Grandis JR, et al. APOBEC-induced mutations and their cancer effect size in head and neck squamous cell carcinoma. *Oncogene* 2019;**38**:3475-3487.

18. Laconi E, Marongiu F, DeGregori J. Cancer as a disease of old age: changing mutational and microenvironmental landscapes. *Br J Cancer* 2020;**122**:943–952.

19. Vitale I, Shema E, Loi S, Galluzzi L. Intratumoral heterogeneity in cancer progression and response to immunotherapy. *Nat Med* 2021;**27**:212-224.

20. Bozic I, Allen B, Nowak MA Dynamics of targeted cancer therapy. *Trends Mol Med* 2012;**18**:311-316.

21. Misale S, Arena S, Lamba S, Siravegna G, Lallo A, Hobor S, et al. Blockade of EGFR and MEK intercepts heterogeneous mechanisms of acquired resistance to anti-EGFR therapies in colorectal cancer. *Sci Transl Med* 2014;**6**:224ra26.

FIGURE AND FIGURE LEGEND

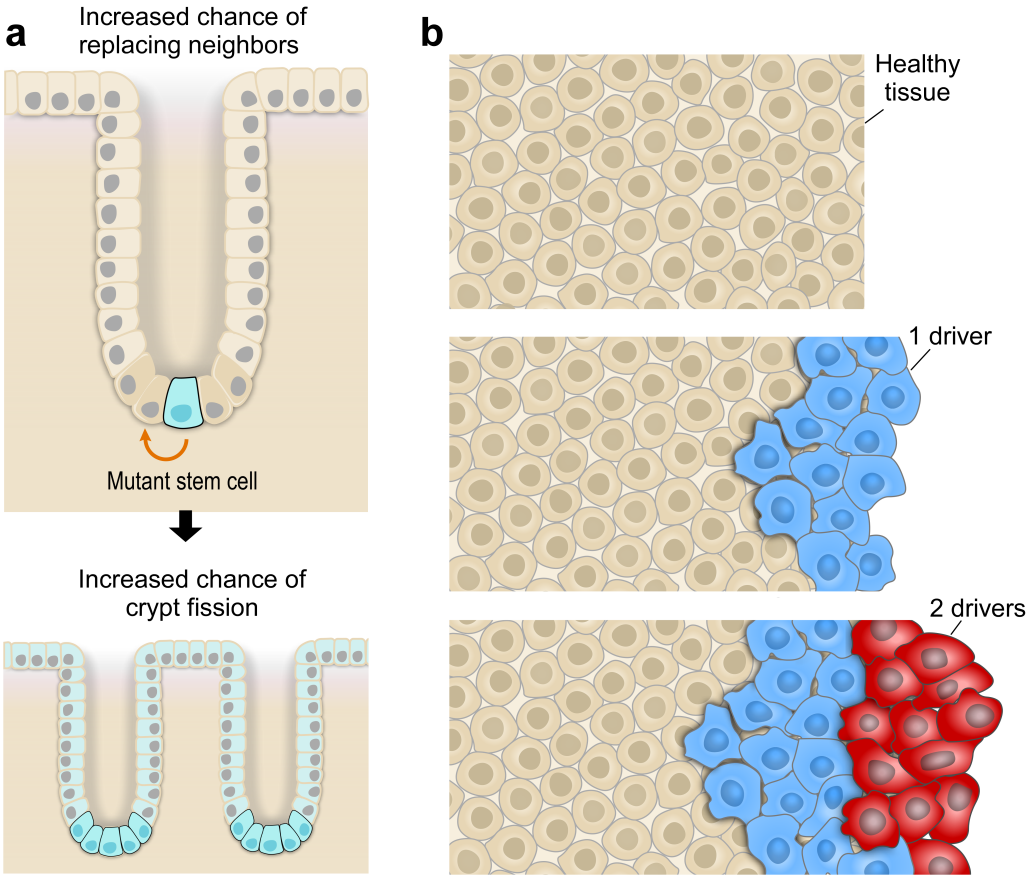

Figure 1. a, At the earliest stages of tumor evolution, a driver mutation can act by increasing the chance that a crypt stem cell will replace its neighboring wild type stem cell (top), and/or by increasing the chance of crypt fission (bottom). Cells with driver mutation are shown in blue. **b,** Time series of tumor evolution. Initially, healthy tissue is at homeostasis (balanced cell division and cell death, top). A driver mutation can act by lowering the death rate and/or increasing the division rate of cells, leading to a positive net growth rate and clonal expansion of cells carrying it (middle). Subsequent driver mutations can appear that lead to further clonal expansions (bottom).