The Universality of Cancer

Carson C Chow, Yanjun Li, Vipul Periwal

***Short Abstract* —** Cancer is a constellation of diseases differing in underlying mutations and on cellular environments. The stochastic process of carcinogenesis has been studied for sixty years, but there is no accepted model. We show that the hazard rates of all cancers are characterized by a simple dynamic stochastic process on a half-line, with a universal linear restoring force balancing a universal simple Brownian motion starting from a universal initial distribution. Only a critical radius defining the transition from normal to tumorigenic genomes distinguishes between different cancer types when time is measured in cell-cycle units. Reparametrizing to chronological time units introduces two additional parameters: the onset of cellular senescence with age and the time interval over which senescence takes place. Thus, there may exist a finite separation between normal cells and tumorigenic cells in all tissue types that may be a viable target for both early detection and preventive therapy.

***Keywords* — cancer, DNA replication, DNA damage, DNA repair, senescence**

I. Purpose

CANCER is part of life for multi-cellular organisms when individual fitness in propagation overcomes the checks and balances required for collective fitness. It is a multifaceted disease where the phenotypic similarities of tumor progression are a veneer over a multitude of possible underlying genetic alterations[1]. 75-80 % of all cancers are sporadic. As an organism ages, the accumulation of mutations increases the likelihood of alteration in an oncogene or in a tumor suppressor gene, which in turn can lead to an accumulation of mutations and other alterations allowing for unchecked proliferation. The process of carcinogenesis has been modeled for over 60 years[2]. Tomasetti and Vogelstein[3] showed that the lifetime risk of cancers of many different types is correlated with the total number of divisions of the normal self-renewing cells maintaining each tissue's homeostasis. Thus most cancer is due to random mutations arising during DNA replication in normal, noncancerous stem cells. This motivates the existence of a simple universal quantitative physical model for this stochastic process of tumorigenesis. We posited that DNA replication could be described by a continuous diffusive process on a mean mutational distance, or error, coordinate from an initial genome. Following the Ornstein-Uhlenback process[4], we hypothesized that DNA replication error correction could be modeled by the simplest possible restoring force, just Hooke's law with a universal spring constant in this coordinate.

II. Methods

Harding, Pompei and Wilson[2] have analyzed the Surveillance, Epidemiology and End Results (SEER, specifically SEER 9) cancer registries to compile age-specific incidence rates, with particular care accorded to the data on the very elderly (ages > 80 years). [2] suggested that tissue and cellular senescence are the likely biology mechanisms for the observed drop off in cancer incidence in the very elderly. These incidence rate curves provide a test for our model. We defined cancer incidence as the likelihood that the diffusing DNA has moved beyond a cancer-type independent critical threshold radius, R, in this coordinate, with the cancer incidence rate defined as the derivative of this likelihood. As the number of replicating stem cells differs from tissue to tissue[3], we explicitly set the maximum value of the incidence rate to the maximum value of the incidence rate of the data for each cancer type[2].

III. Results

We could fit the age-specific incidence rates for all cancers with our models. The best model selected by the Bayes Information Criterion has a width of the initial distribution about 0.16 for all tissues in units where the equilibrium Ornstein-Uhlenbeck distribution has width 1.

IV. Conclusions

The space of mutational histories has a natural diffusion away from the initial starting distribution. The rate of moving beyond a relatively sharp tissue-specific threshold is the incidence rate for all cancers. An interval between normal cells and tumors that could serve as a target for early detection, and a relatively sharp demarcation between tumors and normal cells, are concrete predictions of our model of tumorigenesis. Combining cancer mutation data and epidemiology to find an appropriately weighted cell-type specific mutational burden that could serve as the tissue-specific error coordinate would be of great value.

Acknowledgements: This work was funded by Intramural Research Program of the NIH, NIDDK.

Laboratory of Biological Modeling, National Institute of Diabetes Digestive and Kidney Disorders, NIH. E-mail: carsonc@niddk.nih.gov, yanjun.li@nih.gov, vipulp@mail.nih.gov

References

1. M. O'Dell and M. Stubblefield, Cancer Rehabilitation: Principles and Practice, Demos Medical Publishing (NY), 2009.
2. C. Harding, F. Pompei and R. Wilson, Peak and Decline in Cancer Incidence, Mortality and Prevalence at Old Age, Cancer (2012) 118:1371-1386, DOI: 10.1002/cncr.26376.
3. C. Tomasetti, B. Vogelstein, Variation in cancer risk among tissues can be explained by the number of stem cell divisions, Science (2015) 347:78-81, DOI: 10.1126/science.1260825.
4. G.E. Uhlenbeck and L.S. Ornstein, On the theory of Brownian motion, Phys. Rev. 36 (1930) 823-841