Information Degradation and Evolutionary Games during Carcinogenesis

- Robert A. Gatenby, MD (Depts. Of Radiology and Applied Mathematics, University of Arizona
- B. Roy Frieden, PhD (Dept. of Optical Science, University of Arizona)
- Thomas L. Vincent, PhD (Dept of Aerospace and Mechanical Science, University of Arizona)

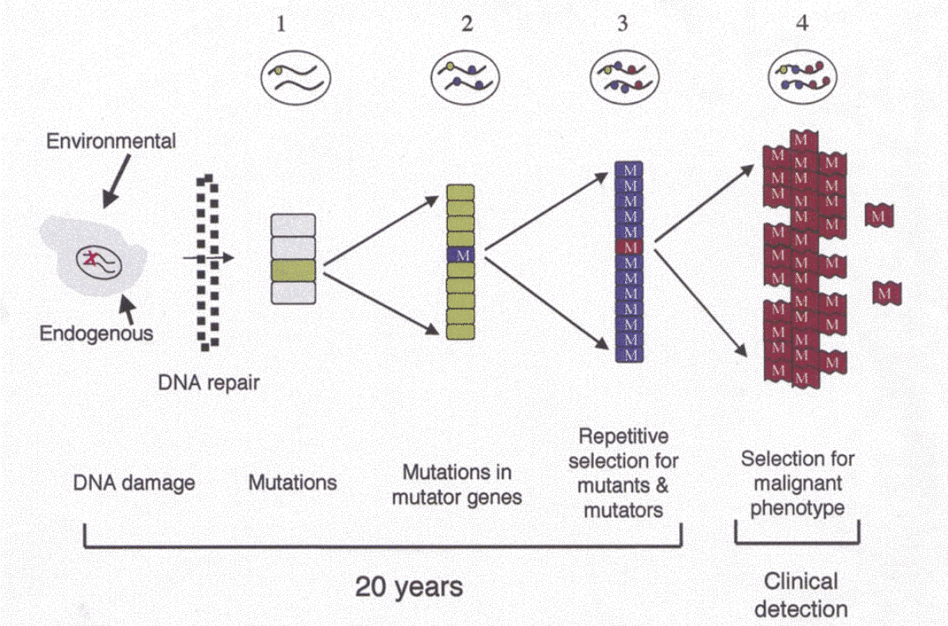
Literature

- Gatenby and Frieden "Application of Information Theory and Extreme Physical Information to Carcinogenesis" Cancer Research 62:3675-3684,2002
- Gatenby and Vincent "An Evolutionary Model of Carcinogenesis" Cancer Research 63:6212-6220,2003
- Gatenby and Vincent "Application of Quantitative Methods from Population Biology and Evolutionary Game Theory to Tumor Therapeutic Strategies" Molecular Cancer Therapeutics. 2(9):919-927, 2003
- Gatenby and Frieden "Information Degradation and Flow in Carcinogenesis" Mutation Research in press.

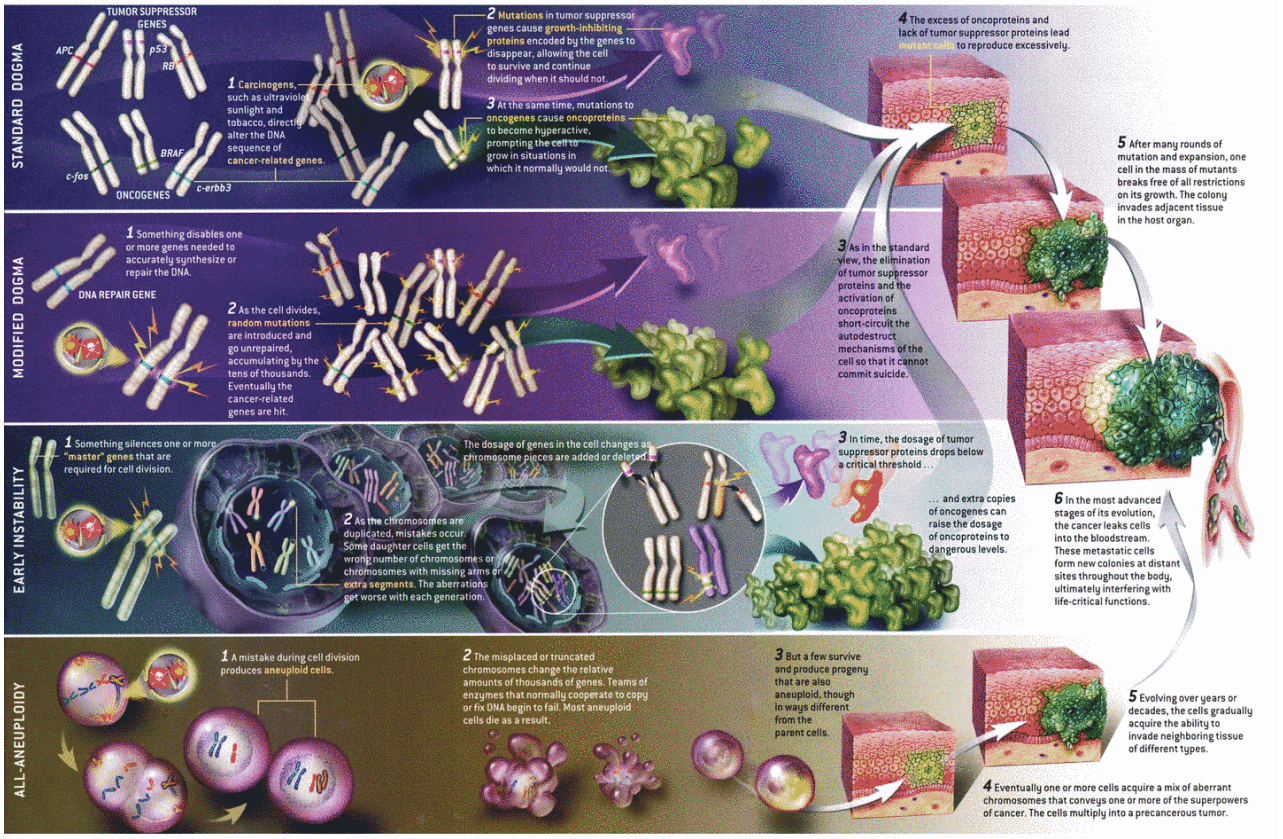
Background

- The development of sporadic cancer generally requires years or even decades with multi-step progression from normal tissue through increasingly disordered pre-malignant lesions such as colon polyps to invasive cancer.
- Carcinogenesis is often describe as “somatic evolution” driven by competition among different populations arising through random mutations with clonal selection determined by the properties of the tissue environment formally analogous to classical Darwinian dynamics.
- There is clear evidence of accumulating mutations during carcinogenesis – most transformed cells possess hundreds, thousands, or even hundreds of thousands of mutations. But, there is no prototypical cancer genotype – the genome of every sporadic cancer populations appears to be unique.
- The general conceptual model is that some genomic mutations confer “selective growth” advantage with clonal expansion. Over time these advantageous mutations accumulate until unconstrained growth results.
- Loeb and others hypothesize increased mutation rate due to chromosomal or microsattelite instability is necessary to drive carcinogenesis

Accumulating mutations is a central component of virtually all carcinogenesis theoretical models



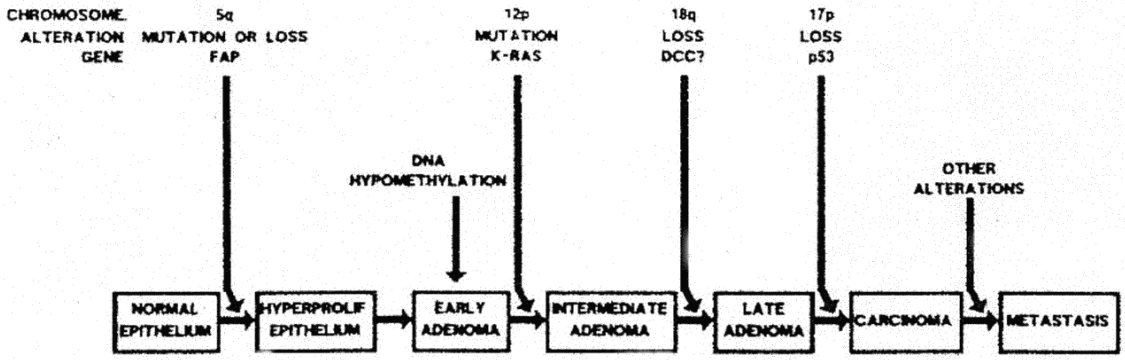
There are several conceptual models of mutation dynamics



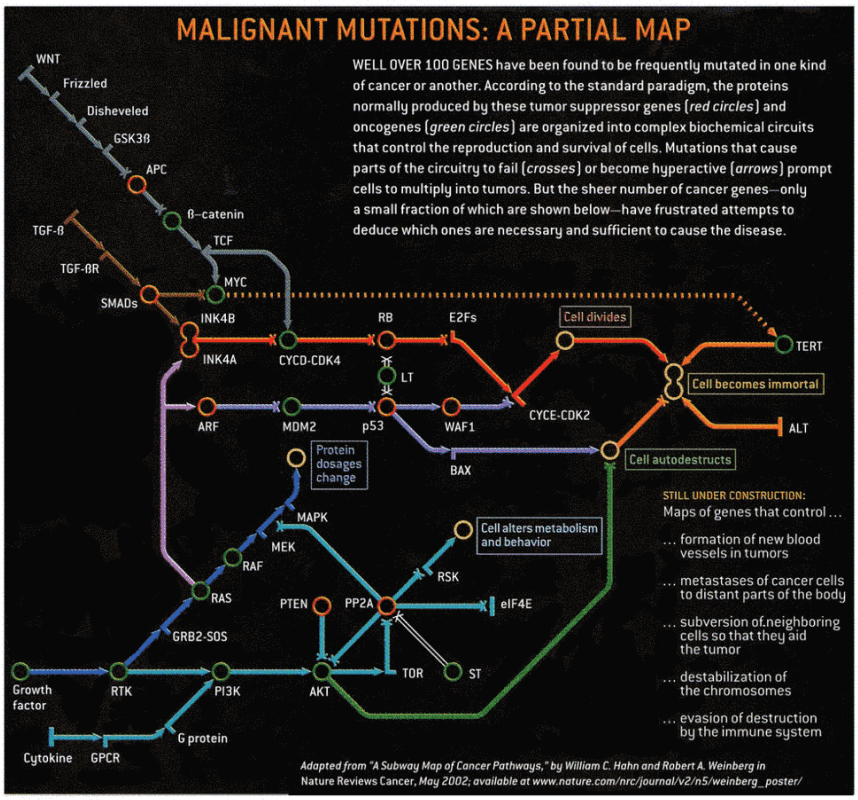
Intracellular factors are the primary focus of most models of carcinogenesis

- Cell-centric models of tumor development and growth typically include a sequence of genomic mutations that are synchronous with progressive drift of cellular populations from normal through premalignant lesions to invasive cancer. Loeb and others hypothesize increased mutation rate due to defects in chromosomal stability or DNA repair pathways is necessary for carcinogenesis. These theoretical models are supported by observations of high mutational frequency in specific oncogenes and tumor suppressor genes in different clinical cancers and a large number of random mutations in many cancer cells. Increased mutation rates have been reported in early colon and esophageal cancers.

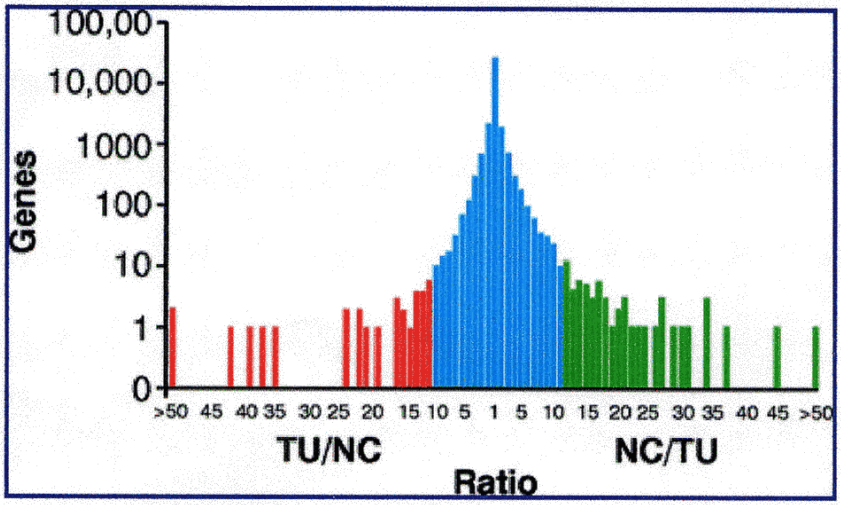
Fearon-Vogelstein conceptual model of colorectal carcinogenesis



But, it isn't that easy



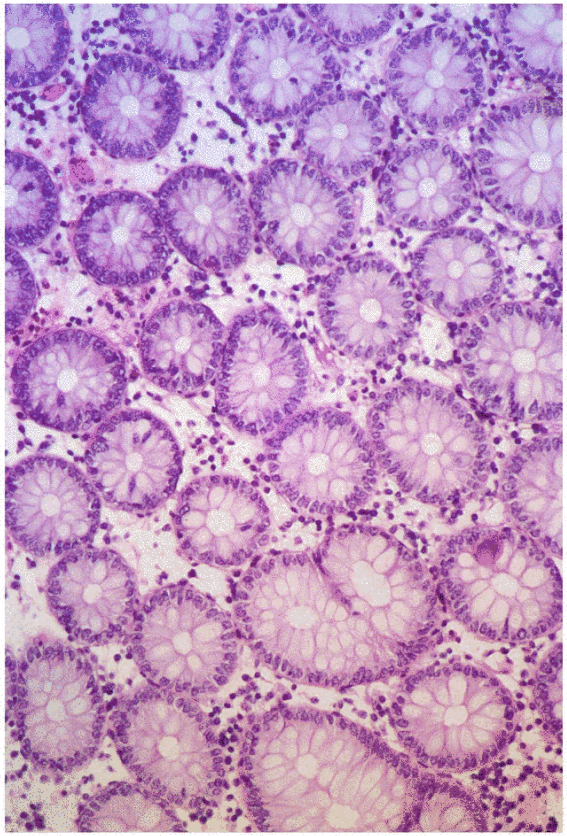
But the story is more that just mutation. Significantly different levels of expression are found in more than 500 unmuatated genes when normal colon cells are compared with colon cancer cells



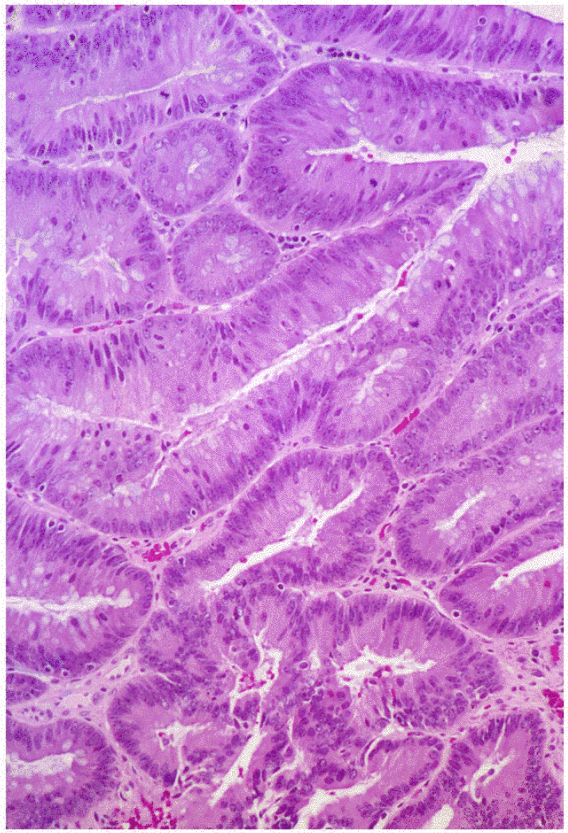
The mutator phenotype seems common, but is it necessary and sufficient? What is the role of the environment?

- Loeb and others hypothesize an increased mutation rate (the mutator phenotype) is both necessary and sufficient for carcinogenesis. This is supported by the large number of mutations typically found in cancer cells and by reports that increased mutation rate is an early event in development of colon and esophageal cancers.
- Tomlinson and others cite empirical evidence and mathematical models to argue a normal mutation rate is sufficient in conditions of strong clonal selection.
- Bissell and others have published numerous articles demonstrating alterations in the environment alter the phenotypic and growth of genomically identical tumor populations.
- Environmental factors including hypoxia and acidosis can increase the mutation rate.

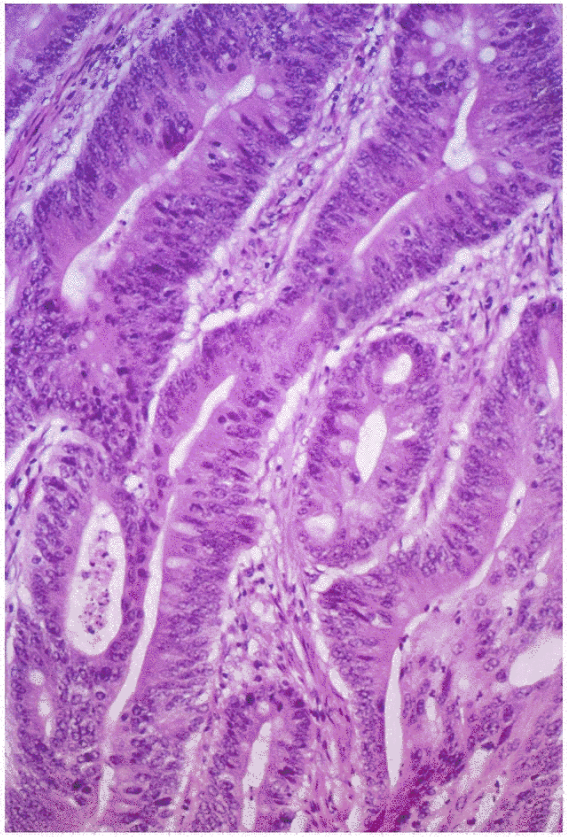
Normal colonic mucosa



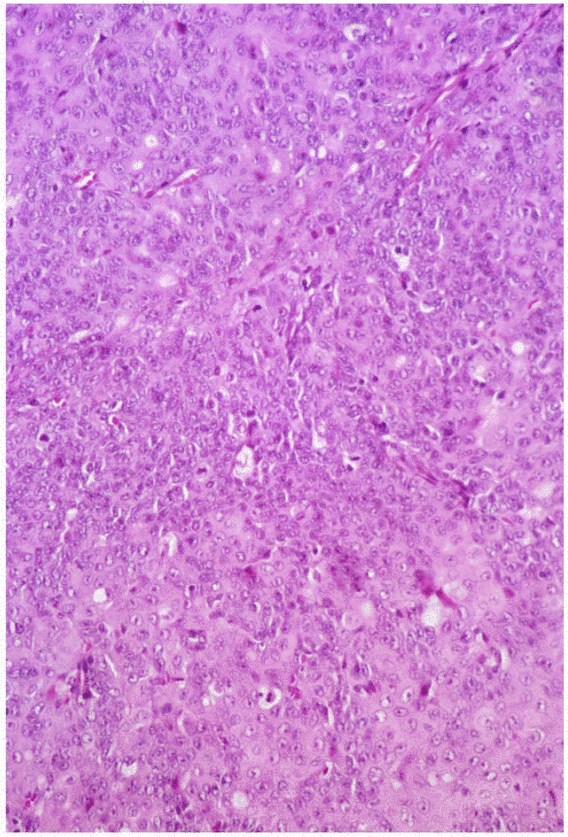
Colon polyp with low grade dysplasia



Colon polyp with high grade dysplasia



Colon cancer



In summary:

Transition from normal tissue to invasive cancer is a gradual, multistep process in which increasingly disordered cellular populations emerge over time. This is often described as "somatic evolution" as random mutations confer selective growth advantages but the underlying dynamics governing the Darwinian interactions of altered cellular genotypes with changing microenvironments remain unclear.

The basic question is really “How does multicellular tissue become cancer?”

- What are the necessary and sufficient conditions for evolution of invasive cancer?
- What are the dynamics at the genomic level of information storage and transmission that allow accumulating mutations to yield a malignant phenotype?
- What are the environmental selection parameters that control the evolution of specific properties of the transformed phenotype and how does these environment factors change over time as the tumor progresses?

Problems and Controversies in the Current Conceptual Model

What does "selective growth advantage" mean? Need to define the dynamics of environmental selection forces with the phenotypic expression of genetic mutations to understand the process.

Is the mutator phenotype necessary? Can invasive cancer evolve with the normal background mutation rate? There is evidence of non-random distribution of the mutations among different gene segment with complete degradation of some and stability of others (e.g. membrane transport proteins).

The role of the environment is poorly defined. Bissell et. al. have shown variations is the microenvironment can profoundly alter cellular phenotype and growth dynamics in the absence of alterations in the genome ("it takes a tissue to make a cancer").

The role of the mutagenic phenotype can be evaluated using information theory since genomic information generates and maintains the transmembrane entropy gradient.

Shannon entropy in each codon where r is the probability of each of the 64 possible configurations:

$$H^i = -\sum_{j=1}^{N} r_j^i \log_b r_j^i$$

Total information in a gene with m codons:

$$I_g = mH^i$$

Total information in the genome with G genes

$$I_c = \sum_{g=1}^{G} I_g$$

Cellular fitness u_c is sum of contribution from all of the genes where u_g is the fitness

contribution from each gene and is a function of the information content (i.e. reduction in the information content of the gene reduces its contribution to fitness) and the total number of gene products within the celll k which may be controlled by other genes acting as repressors or promoters

$$u_c = \sum_{g=1}^{G} k(I_c) u_g(I_g)$$

Cellular proliferation r_c is determined by the cellular fitness compared to the mean

fitness of its competitors within the community

$$r_c = f(u_c - \bar{u}) = f[\sum_{g=1}^{G} k(I_c) u_g(I_g) - \ 1/M \sum_{m=1}^{M} u_m]$$

Applying the Eigen-Schuster limit, information degradation in a specific gene follows its contribution to fitness:

$$\frac{du_c}{dt} = a(u_c - \bar{u}) \frac{d(I_{c\,max})}{dt}$$

Results

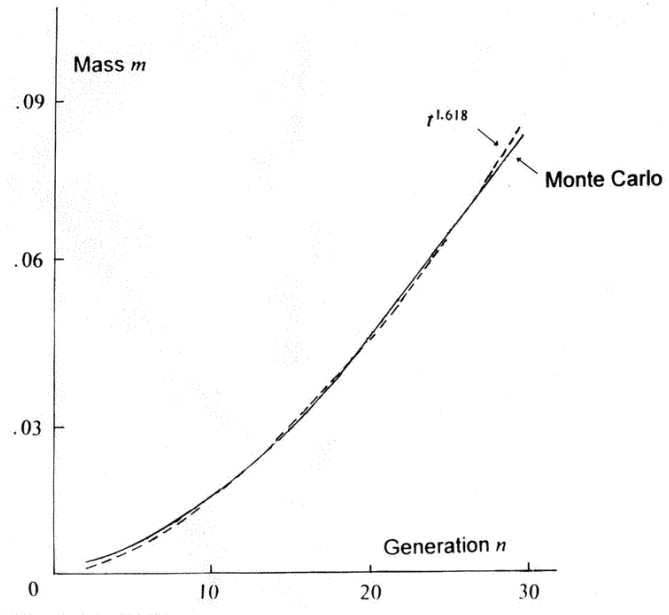
- Information loss because of accumulating mutations is constrained by the competitive stresses of the Darwinian environment in carcinogenesis protecting the genome from an “information crisis”
- Because of these dynamics, gene segments that decrease the fitness (proliferation) of the cells are subject to maximum degradation. This is manifested as loss of function of tumor suppressor genes and differentiation genes. The latter manifests as progressive de-differentiation
- Gene segments necessary for proliferation (such as oncogenes, membrane transporters, PFK) are protected by prompt clonally loss following a muation. The *observed* mutation rate in these genes will be minimal consisting primarily of gain-of-function mutations.
- The net effect is tumor cells will asymptotically approach a state of minimum information (minimal complexity) resulting in progressive loss of differentiated function but unbounded proliferation. This implies mechanism of tumor invasion must be simple

The role of the minimum information state in cancer biology can be determined using Extreme Physical Information (EPI)

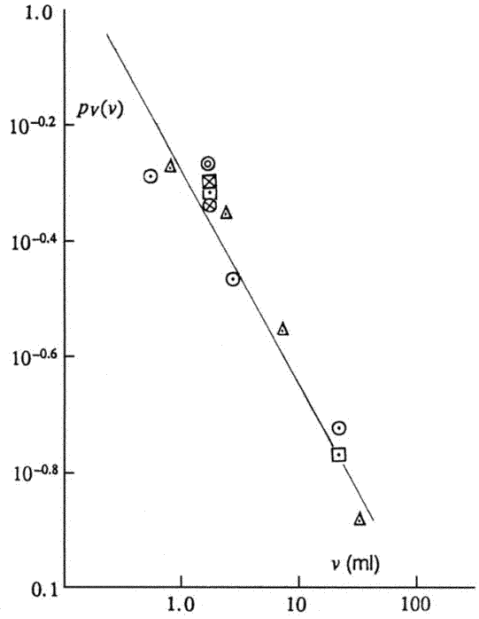
- Define $p_c(x,t)$ as the probability that any observed cell in some volume of tissue (x) is a tumor cell. This marginal probability also represents, by the law of large numbers, the relative number of cancer cells in any space x. The time dependence of $p_c(x,t)$ defines tumor growth.
- In EPI data in any measurement is the result of flow of Fisher Information from an information source to a sink $J \_ I$
- Here J is the extracellular information produced by the presence of the cancer cell about the age of the tumor. I is the quantity of information that reaches normal cells.
- Assuming only that cancer cells exist at a minimum information state and proliferate in a “free field” the prediction is:

$$p(t)=Ft^{-} \quad F=\text{constant} \quad \_ \approx 1.62$$

Tumor growth dynamics from EPI assuming state of minimal information and through Monte-Carlo simulations



The EPI predictions can be compared to the growth rate of small breast cancers obtained using sequential mammograms. Six studies were found in the literature – all showed power law growth with _ of 1.72, 1.69, 1.47, 1.75, 2.17, and 1.61 (mean 1.73 ± 0.23)



Other EPI predictions

- In-vitro growth of tumor populations will be exponential – this is observed in 3-dimensional co-culture experiments
- There is a potentially large discrepancy between the tumor "age" and size. That is, tumors will generally be older that one would assume by working backward from their size.
- Based on recent literature demonstrating the metastatic phenotype develops relatively early in cancer progression, the latter results may place an absolute limit on the efficacy of screening approaches that rely on size-thresholds for detection (ie mammography)

Basic Evolution Equations

Assume a volume of tissue contains distinct cellular populations designated by $x_i, i = 1, \ldots, n_s$. Each population is defined by a phenotype vector $\mathbf{u}_i$ composed of multiple scalar components that include cellular properties and interactions with the microenvironment (other cells, ECM, nutrients etc.). We define population and mean phenotype vectors:

$$\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_{n_s} \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_{n_s} \end{bmatrix}$$

"mean phenotype" assumes limited diversity within each population due to the small background mutation rate and environmental perturbations. This has been observed in clonal populations of both normal and transformed cells .

Cellular fitness is defined by clonal proliferative capacity and determined by fitness-generating or G-functions with a virtual variable, v. Setting the virtual variable equal to the phenotype of a population produces the fitness for that population, which is a function of $\mathbf{x}, \mathbf{u}$, and substrate concentration R. The relationship between fitness and the G-function is given by

$$G(v, \mathbf{u}, \mathbf{x}, R)_{v=\mathbf{u}_i} = H_i(\mathbf{u}, \mathbf{x}, R) \quad i = 1, \ldots, n_s.$$

The population dynamics maybe written either in terms of the fitness function or the fitness generating function

$$\dot{x}_i = x_i H_i(\mathbf{u}, \mathbf{x}, R) = x_i G(v, \mathbf{u}, \mathbf{x}, R)|_{v=\mathbf{u}_i}$$

While the G-function does not provide a conceptual advantage from simply writing down equations of motion, it is critical for understanding how systems evolve.

A single G-function model with scalar strategies - defining "somatic ecology":

$$G(v, \mathbf{u}, \mathbf{x}, R) = B_n \left(1 - \frac{\sum_{i=1}^{n_s} a(v, \mathbf{u}) x_i}{K(v)}\right)\left(\frac{E(v)R^2}{R_0^2 + R^2} - m\right)$$

Cell populations in-vivo are subject to two general growth constraints:
1."Organizational" controls encompassed in $K(\mathbf{v})$ including intracellular processes such as senescence and interactions with the extracellular environment including other cells and environmental factors that resutl from their phenotypes such as ECM, growth promoters and suppressors $[a(v, \mathbf{u})]$. 2. Substrate availability (second term) - cells must obtain substrate in excess of basal metabolic demand m to supply energy and macromolecules for proliferation. B_n is a constant converting excess substrate into new cells.

Substrate dynamics

We assume population numbers for each phenotype x_1 are normally determined by $K(\mathbf{v})$. That is, normal cells under physiologic conditions are not subject to substrate limitations. Pathological exceptions include acute or chronic ischemia such as stroke, myocardial infarction or diabetic ulcers.
Substrate dynamics include Michaelis-Menten uptake:

$$\dot{R} = r - \sum_{i=1}^{n_s} \frac{E(v)R^2}{R_0^2 + R^2} x_i$$

where r is substrate delivery rate

$$r = r_e \left(m_n N_1 + m_t \sum_{i=2}^{n_s} N_i \right)$$

r_e represents local physiologic control that modulates flow through the vascular network and must be > 1 for cell proliferation (i.e. delivery must exceed basal demand). We assume maximum substrate delivery is limited by local vascularity:

$$\text{if } r > r_{\max} \text{ then } r = r_{\max}$$

The model becomes evolutionary by defining the critical growth parameters as distribution functions

$$E(v_1) = E_{\text{mean}} \exp\left(-\frac{(v_1 - u_{E_{\max}})^2}{2\sigma_E^2} \right)$$

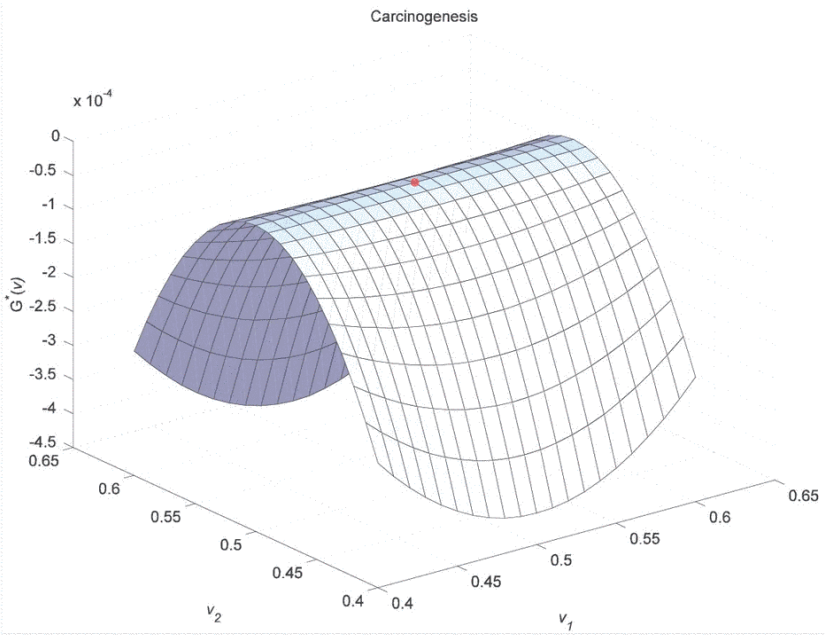
and the second component of v determines carrying capacity according to

$$K(v_2) = K_1^{mean} \exp\left(-\frac{(v_2 - K_{ij}^{\max})^2}{2\sigma_K^2} \right)$$

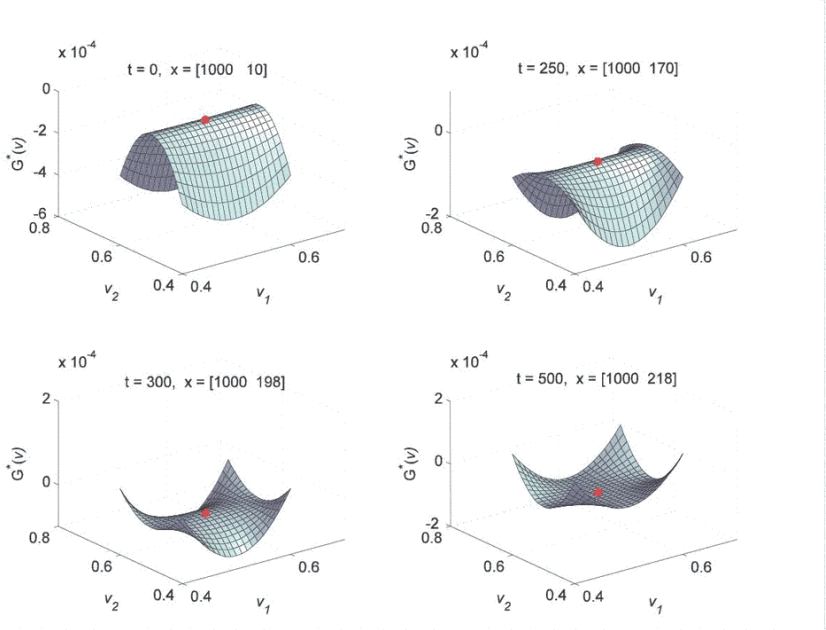
where

$K_1^{\max}$ = Maximum number of cells
K_1^{mean} = Mean tissue carrying capacity of tumor cells
$E_{n_{\text{mean}}}$ = Mean substrate uptake for normal cells
$E_{t_{\text{mean}}}$ = Mean substrate uptake for tumor cells
u_{x_t} = Value of u_{i2} for largest $x_1^{\max}$ $t = 2, \ldots, n_s$
$u_{E\max}$ = Value of u_{i1} for maximum E
σ_j = Variance in K distributions
σ_E = Variance in E distributions

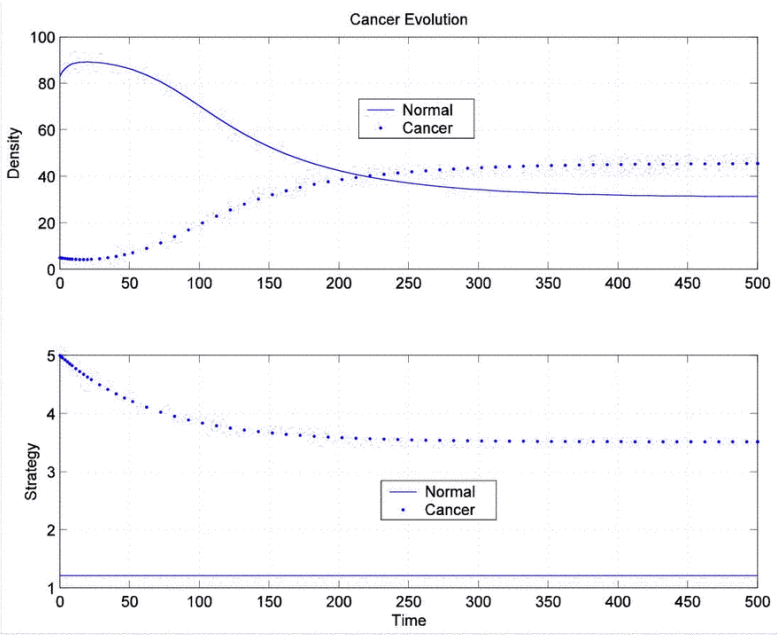
Fitness landscape with only normal cells. The ridge-like maximum allows non-competitive coexistence of multiple phenotypes – a state necessary for formation of functioning, multicellular tissue. Problem – this is not a proper maximum and is subject to invasion by fitter phenotypes.



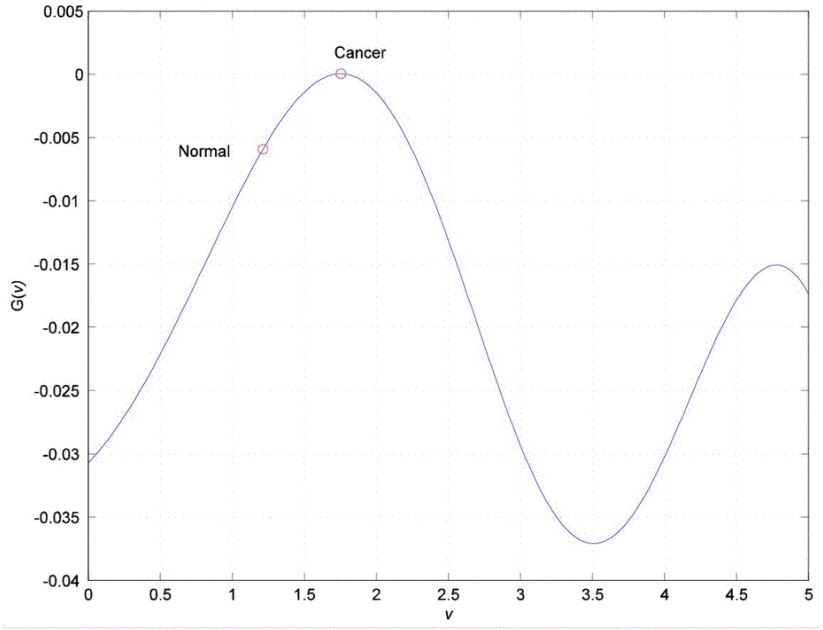
The introduction of a mutant population with a higher fitness warps the fitness landscape so that the extant they come to occupy a minimum adjacent to unoccupied maxima.



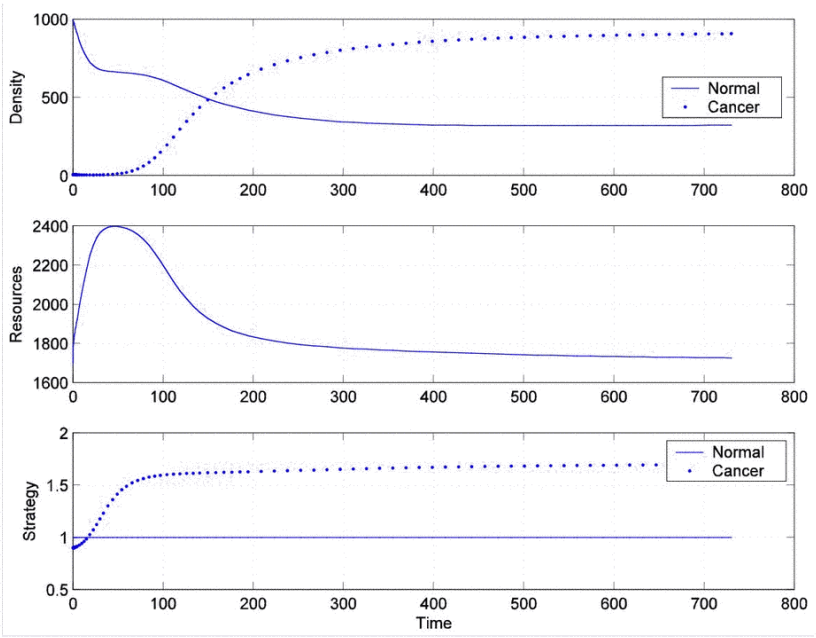
Since normal cells are constrained by tissue factors, only mutations that increase the value of *K* increase fitness allowing growth in early carcinogenesis. So early mutations must involve tumor oncogenes and suppressor genes. However, even when *K* evolves to a maximum only self-limited growth, such as a polyp, is observed due to development of substrate limitation.



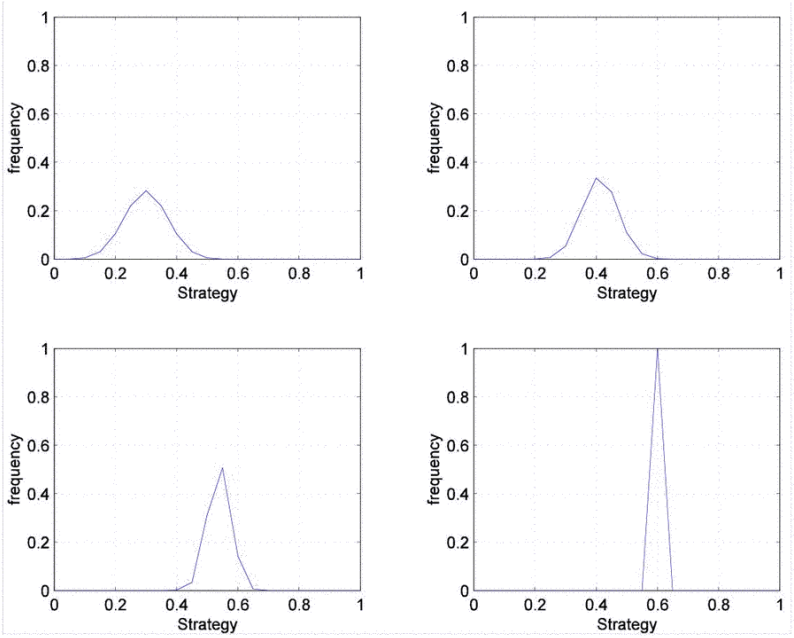
Tumor populations reach a maximum only by altering the dynamics of substrate uptake by acquiring the angiogenic and glycolytic phenotypes during a second phase of carcinogenesis dominated by subtstrate competition.



The period of substrate competition promotes the metabolic changes in cancer allowing tumor cells to function in environments of extreme subtrate limitation. Invasive cancer emerges following this final stage of evolution. Consistent with data from Folkman and from FDG-PET showing the glycolytic phenotype is universal in clincal invasive cancer and is evident simultaneous with the transition from polyp to cancer in CRC



Although phenotypic diversity is favored during carcinogenesis, once a population has reached a fitness peak, diversity decreases



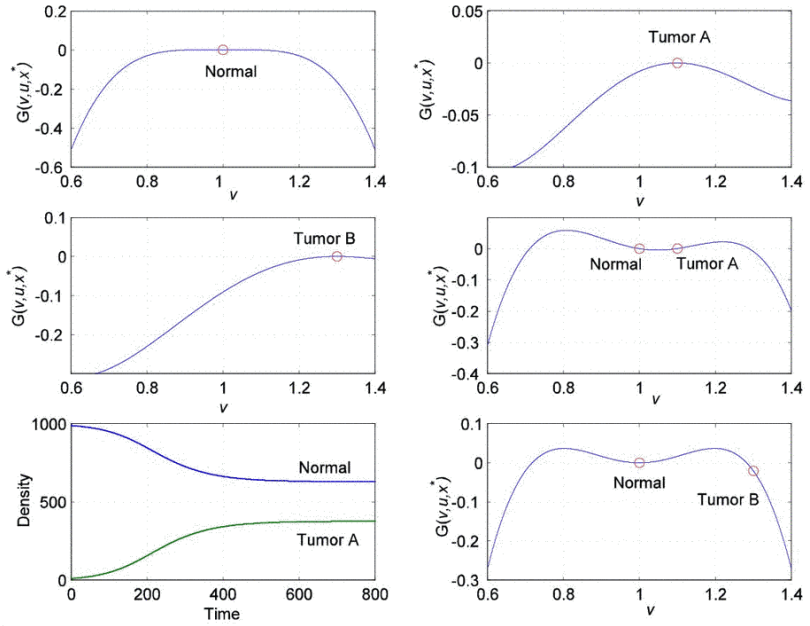
So, invasive cancer is an inevitable consequence of the fitness landscape necessary for multicellular life. The rate of evolution is critical – if it is sufficiently fast to approach the fitness maximum within the lifetime of the host, cancer develops. Otherwise, no problem.

Using appropriate assumptions, the rate with which a population evolves to a fitness maximum is:

$$\dot{u}_i = \sigma_i \frac{\partial G}{\partial v}\bigg|_{v=u_i}$$

This rate equation demonstrates the evolutionary dynamics determining clinical emergence of invasive cancer are governed by both intra and extracellular factors: phenotypic variance σ_i dependent on the mutation rate and 2. the local slope of the fitness curve, $\partial G/\partial u_i$ determined by clonal selection forces in the microenvironment.

The critical role of cellular mutations in evolution of the malignant phenotype is well accepted. The equally critical role of the environment is less well known but consistent with multiple studies by Bissell et al and observed clinically in Neuroblastoma IVs and probably in metastases.



Conclusions:

- Application of information theory to carcinogenesis demonstrates the important role of Darwinian competition among mutant clones in constraining the effects of accumulating genomic mutations.
- The constraints also produce variations in the observed mutation rate of different gene segments and add a caveat to interpretation of experiments designed to measure the mutation rate and understand the role of specific mutations in tumor biology. Specifically: 1. using the observed mutation rate in any gene to infer the global genomic mutation is possible only with precise knowledge of the contribution of that gene to cellular fitness. 2. tumor cells in-vitro are subject to environmental selection pressures different from those in-vivo. Mutations observed in cell lines may be irrelevant to the same tumor cells when they are in-situ
- During carcinogenesis, cellular information asymptotically approaches a minimum. This state of minimum complexity can be used to accurately predict tumor growth dynamics and suggests fundamental limitations in tumor screening strategies.

Conclusions (cont)

- Evolutionary game theory can be used to define the interactions of the phenotypic properties generated by accumulating mutations and environmental selection properties.
- Growth constraints in normal tissues favor initial mutations that alter cellular reception or processing of growth promoter and inhibitory signals such as mutations in oncogenes and tumor suppressor genes.
- As these mutations accumulate, the populations, although unconstrained by local growth factors, exhibit only self-limited growth due to substrate limitations. This results in transition to a previously unknown phase of carcinogenesis dominated by competition for substrate and provides the evolutionary dynamics for development of the angiogenic and glycolytic phenotypes.
- The cellular properties of invasive cancer populations represent the summation of both stages of evolution.

Conclusions (cont)

- In general, the evolutionary models demonstrate malignant phenotypes will inevitably emerge from the fitness landscape necessary for maintenance of multiple phenotypes in a cooperative, non-competitive environment. That is, tumors are the price of the environment necessary to maintain functioning multicellular organisms. This is evident in the number of benign lesions such as colon polyps or skin nevi that increase monotonically with age. These mutant populations, although benign, have the potential to form cancers if they can evolve to a local fitness maximum.
- The development of a clinical cancer from a premalignant lesion is dependent on the speed of evolution. If a tumor population approaches a fitness maximum within the lifetime of the host, he/she develops cancer. Otherwise, the tumor is insignificant.
- The evolutionary rate is determined by the mutation rate and the clonal selectivity of the environment. This combines the cell-centric approach and the environmental approach into a single conceptual model of carcinogenesis.
- Alterations in the environment may substantially alter tumor growth even in the presence of a stable genome. Consider strategies that treat normal cells and alter the adaptive landscape rather than treating just tumors.