ANTICANCER RESEARCH *19:* 4853-4868 (1999)

Review

The Role of Molecular Discreteness in Normal and Cancerous Growth

JAMES MICHAELSON

Department of Pathology, Harvard Medical School, Department of Pathology, Massachusetts General Hospital, Massachusetts General Hospital Cancer Center, Bldg. 149, 13th Street, Charlestown (Boston) MA 02129, U.S.A.

Abstract. *The physicochemical events that underlie biological processes are inevitably either/or events. Either a growth factor molecule binds to a cell, or it doesn't. Either a site on a cyclin molecule is phosphorylated, or it isn't. Either a regulatory molecule binds to a DNA sequence, or it doesn't. These molecular either/or events lead to cellular either/or events. Either a cell divides, or it doesn't. Either a cell dies, or it doesn't. Either a cell turns on a particular gene, or it doesn't. Either a tumor cell stays where it is, or it forms a distant metastasis. By considering biological processes as the macroscopic aggregate results of these many individual microscopic either/or events, we can gain considerable insight into both normal and cancerous growth. In fact, as will be outlined here, such discrete modeling may allow us to see how the normal cellular populations of the body can grow to predictable sizes, at predictable times, and to predictable shapes. Such modeling can also allow us to gain insight into how normal cellular populations may become cancerous cellular populations. Indeed, such an approach allows us do a sufficiently good job of imitating the growth and spread of tumors as to be able to make estimates the most effective ways to both detect and treat cancer.*

INTRODUCTION

The discrete nature of atoms and molecules gives all chemical change a random quality

The potential for randomness in cellular behavior is latent in the simple but fundamental fact that all matter is made of ***discrete*** entities: electrons, atoms, molecules. There may be 0, 1, 3, or a trillion molecules of insulin bound to a cell, but never 1.7 molecules. Consider a chemical reaction between a single molecule of insulin and ***six*** insulin receptor molecules; at any single point in time, only ***one*** of the ***six*** receptor molecules can bind the single molecule of insulin. Which ***one*** of the ***six*** receptor molecules will bind the insulin molecule is just as random as which ***one*** of the ***six*** faces of a gambler's die will land upright. When such small numbers of molecules are at work on cells, or in cells, the intrinsic randomness of molecular discreteness expresses itself biologically.

Correspondence to: James Michaelson;
E-mail: michaelj@helix.mgh.harvard.edu

Key Words: Cancerous growth, carcinogenesis.

Cells are often exposed to small numbers of molecules

We can see a dramatic example of molecular discreteness by calculating how many growth factor molecules are at work in signaling a cell. Such a calculation is fairly straightforward, if we have information on growth factor concentration, affinity, number or receptors, *etc.* (Table I). The surprising result is that, in many instances, there are very few such ligand molecules at work in signaling individual cells (Table I). In fact, sometimes there are not even as many bound growth factor molecules as there are cells. For example, for the growth factor IGFII, at the ordinary physiological concentration of 10^{-11} mol/L, there is just ***one*** molecule of IGFII bound for every ***six*** cells with receptors for this ligand (Table I).

The discrete allocation of growth factor molecules among cells gives cell division an either/or quality

The discrete allocation of growth molecules imparts onto cells the ***either/or*** quality normally seen in individual molecular events. If it takes but one molecule of IGFII to make a cell divide, then the discrete allocation of IGFII among cells (Table I) would give every cell roughly a **1-in-6** chance (**~.166**) of dividing. If it takes ***two***, or ***three***, or ***four*** molecules of IGFII to make a cell divide, then the discrete allocation of IGFII among cells will give rise to this same random quality of cell division, with cells having probabilities of $\sim(1/6)^2$ [**~.0277**], or $\sim(1/6)^3$ [**~.0046**], or $\sim(1/6)^4$ [**~.00077**]. Molecular discreteness imparts a random quality to cell division (1).

Table I.

Ligand	Physiological concentration of the ligand in plasma ([L]) (moles/liter)	Dissociation constant (K_D) (moles/liter)	Number of receptors per cell (R_O)	Number of bound ligands per cell ($R_O[L]/K_D$)
Epinephrine			750	1 ligand per 17 cells - 1 ligand per 5 cells
Norepinephrine	$300\text{-}2800 \times 10^{-12}$	$.9 \times 10^{-6}$	750	1 ligand per 4 cells - 2 ligands per cell
Erythropoietin	$125\text{-}1.25 \times 10^{-12}$	$.09 \times 10^{-9}$	300	1 ligand per 2 cells - 4 ligands per cell
Insulin-like growth factor-	$1\text{-}18 \times 10^{-12}$	$.23 \times 10^{-9}$	250	1 ligand per cell - 18 ligands per cell
Insulin-like growth factor-II	$7\text{-}50 \times 10^{-12}$	1×10^{-8}	250	1 ligand per 6 cells - 1 ligand per cell

References: Epinephrine, Norepinephrine (44.45), Erythropoietin (46), Insulin-like growth factor-I, Insulin-like growth factor-II (47,48)

Method for calculating the average number of bound growth factor molecules per cell

It is possible to determine approximately how many ligand molecules are bound among a population of cells, starting with the Law of Mass Action (equilibrium form):

(1) $$[L][R]/[LR]=K_D$$

in which [L]=the physiological concentration of ligand, [R]=the concentration of growth factor receptors, [LR]=the concentration of growth factor molecules bound to receptors, and K_D=the dissociation constant of the receptor for its ligand. We may rearrange this to reveal the ratio of bound receptors to free receptors:

(2) $$[LR]/[R]=[L]/K_D$$

and by multiplying this ratio by the number of receptors per cell, R_O, we may determine the number of bound ligand molecules per cell, L_{BOUND}/N_R, where N_R=the number of cells with receptor for the ligand that make up the population and L_{BOUND}=the total number of ligands bound to all of the cells of the population. Thus:

(3) $$L_{BOUND}/N_R=R_O[L]/K_D$$

The molecular discreteness gives populations of cells a way to control their growth by creating and controlling the growth fraction (G)

While molecular discreteness imparts an ***either/or*** quality on the mitosis of individual cells, its effect on populations of cells is to cause mitosis to occur in just a fraction of the population (the ***growth fraction*** (G)). For example, for IGF-II at the physiological concentration shown in Table I, if the binding of a single molecule of IGFII is enough to make a cell divide, then the discrete allocation of that ***one*** molecule of IGFII binding among every ***six*** cells will result in just ***one*** in every ***six*** cells dividing ($G \sim 1/6 \sim 0.1666$). If it takes ***two*** molecules of IGFII to make a cell divide, then this same discrete allocation of IGFII molecules will result in a growth fraction of about **1/6x1/6** of the cells in the population ($G \sim (1/6)^2 \sim 0.0277$). In this way, molecular discreteness gives populations of cells a way to control their growth (2,3).

Cells can never escape the effects of molecular discreteness

How many molecules of growth factor must actually be bound to a cell to change the cell's behavior? A single photon is sufficient to set off a Gprotein cascade in a retinal photoreceptor (4). For some ligands, such as IGFII, at least two binding events, sometimes with receptor dimerization, are required, because receptor cross phosphorylation is necessary to set off a mitotic signaling cascade (5). Let us call this discrete threshold number "**h**". We shall call the number of inhibitory growth factor molecules required to prevent a cell from dividing "$\mathbf{h_I}$", and the number of stimulatory growth factor molecules required to induce a cell to divide "$\mathbf{h_S}$".

The discrete allocation of growth factor molecules among cells will be most dramatic when there are few ligand molecules, such as in the case of IGF-II. However, cells can never entirely escape the effects of this discrete process, no matter what the level of ligand binding might be (3). For example, should a population of cells have bound, either at one time, or over a period of time, an average of **1000** ligand molecules to its surface, then the discrete nature of these ligands would ensure that some cells will bind **1001** molecules, some **999**, some **997**, *etc.* If $\mathbf{h_S=1000}$, then about one half of the cells in the population will divide. The higher the value of $\mathbf{h_S}$, the lower will be the growth fraction (G) while the lower the value of $\mathbf{h_S}$, the higher will be the growth fraction (G).

Of course the concentration of growth factor may be so high, or value of **h** may be so low, as to make the ***growth fraction*** (**G**) close to **1**, and thus no longer a useful way to set the rate of growth. Might cellular signaling occur in such a range? The data in Table I suggest otherwise, but, of course, here is an area that is ripe for experimental analysis. Indeed, this provides one way to test the models proposed here.

GROWTH

As we have seen, molecular discreteness, working at the microscopic scale, can cause mitosis to occur in some cells but not others. In this way, molecular discreteness can create the ***growth fraction*** (**G**), and thus control growth. As we shall see next, with this ability to create and control the growth fraction, molecular discreteness can give rise to the macroscopic features of growth: growth to predictable ***sizes,*** at predictable ***times,*** and to predictable ***shapes.*** We can see this by considering several simple examples.

Cellular populations can use molecular discreteness to grow to predictable sizes by S-shaped growth curves

Let us consider a population of cells in a constant volume, such as a bird embryo in an egg, or an organ in an animal of a fairly constant size. We can call the number of cells in this population **N**, and the final size of the population $\mathbf{N_{MAX}}$. In this example, let us assume that proliferation will be controlled by an inhibitory growth factor, that is, a factor which, when it binds to a cell, will prevent cell division. In fact, inhibitory growth factors of this type, such as TGF-β are quite common. Let us also assume, in this example, that all of the cells of the population produce the factor, and that all are sensitive to the inhibitory action of the factor. Let us call the dissociation constant of the inhibitory growth factor for its receptor "$\mathbf{K_D}$", the number of receptors per cell "$\mathbf{R_O}$", the number of inhibitory growth factor molecules produced by each cell "**a**", the minimum number of inhibitory growth factor molecules required to prevent cell division "$\mathbf{h_I}$", and the volume of the organism "**V**".

In this example, when there are very few cells, there will be very few inhibitory growth factor molecules. The discrete allocation of these few growth factor molecules will mean that most of the cells will ***not*** have bound by $\mathbf{h_I}$ molecules of the inhibitory growth factor, thus allowing most to divide ($\mathbf{G \sim 1}$, when **N** is small). As the population grows in size, because the volume (**V**) in which it is growing is constant, there will be a gradual increase in the concentration of the inhibitory growth factor, and thus there will be an increase in the number of cells that have bound $\mathbf{h_I}$ inhibitory growth factor molecules, and an increase in the fraction of cells that will be prevented from dividing ($\mathbf{G<1}$ when $\mathbf{N<N_{MAX}}$). Eventually, the population will grow to a size where every cell will have bound $\mathbf{h_I}$ molecules of the inhibitory growth factor, at which point growth will finally stop. At this point, the $\mathbf{N_{MAX}}$ of the population will have been reached ($\mathbf{G=0}$ when $\mathbf{N=N_{MAX}}$) (Figure 1B). Such a gradual decline in growth rate to a predictable final size ($\mathbf{N_{MAX}}$), looks like an "S" when displayed on a graph of cell number-versus-time, as we can see by making a simple computer simulation of our example (Figure 1B). Such S-shaped growth is often found for the developing populations of cells that form the tissues, organs, and anatomical structures of the body (Figure 1A) (6,7), and thus Figure 1B illustrates how molecular discreteness can give rise to this familiar aspect of growth. It is fairly straightforward to calculate the equation of S-shaped growth and the value of $\mathbf{N_{MAX}}$, which results from such a discrete distribution of growth factor molecules. (For derivation the reader is referred to my webpapge; http://webm9120.ntx.net/BreastCancerMath.html):

(1) $$\textbf{rate of growth} = \mathbf{r\,N(1-[N/N_{MAX}])}$$
where $\mathbf{N_{MAX} = h_I[VK_D/R_Oa]}$

So, here we have arrived at the remarkable result that ***molecular discreteness can give a population of cells a way to grow by an S-shaped growth curve to a predictable size (N_{MAX}).***

Cellular populations can use molecular discreteness to grow at predictable times

The appearance of multiple organs and tissues, each at their familiar times, is a fundamental feature of embryonic development. Such timing may also be seen to be a natural consequence of the discrete allocation of growth factor molecules among cells. If, for example, there are two populations, $\mathbf{N_1}$ and $\mathbf{N_2}$ each produce self-acting inhibitory growth factors, and, in addition, the cells of the $\mathbf{N_1}$ population produce a stimulatory growth factor, without which the cells of the $\mathbf{N_2}$ population will not divide, then the discrete allocation of the stimulatory molecules will naturally lead to a delay in the appearance of population $\mathbf{N_2}$. Again, the essential mechanism at work here is the discrete allocation of growth factor molecules among cells. Such timing of the growth of a population may be shown mathematically, and the result is illustrated in Figure 1C.

Cellular populations can use molecular discreteness to grow to predictable shapes

Perhaps the most dramatic expression of embryogenesis is the creation of anatomical form. We have found that such growth to predictable ***shape*** can also be the consequence of the discrete allocation of growth factor molecules among cells. We were able to detect this potential of molecular discreteness by constructing a computer simulation model that followed the distribution of cells and growth factor molecules spatially. The programs for these simulations were developed by a student in my lab, ChaoWei Hwang, and run on a Silicon Graphics workstation. These simulations simply treated each cell and growth factor molecule as a discrete entity. In each

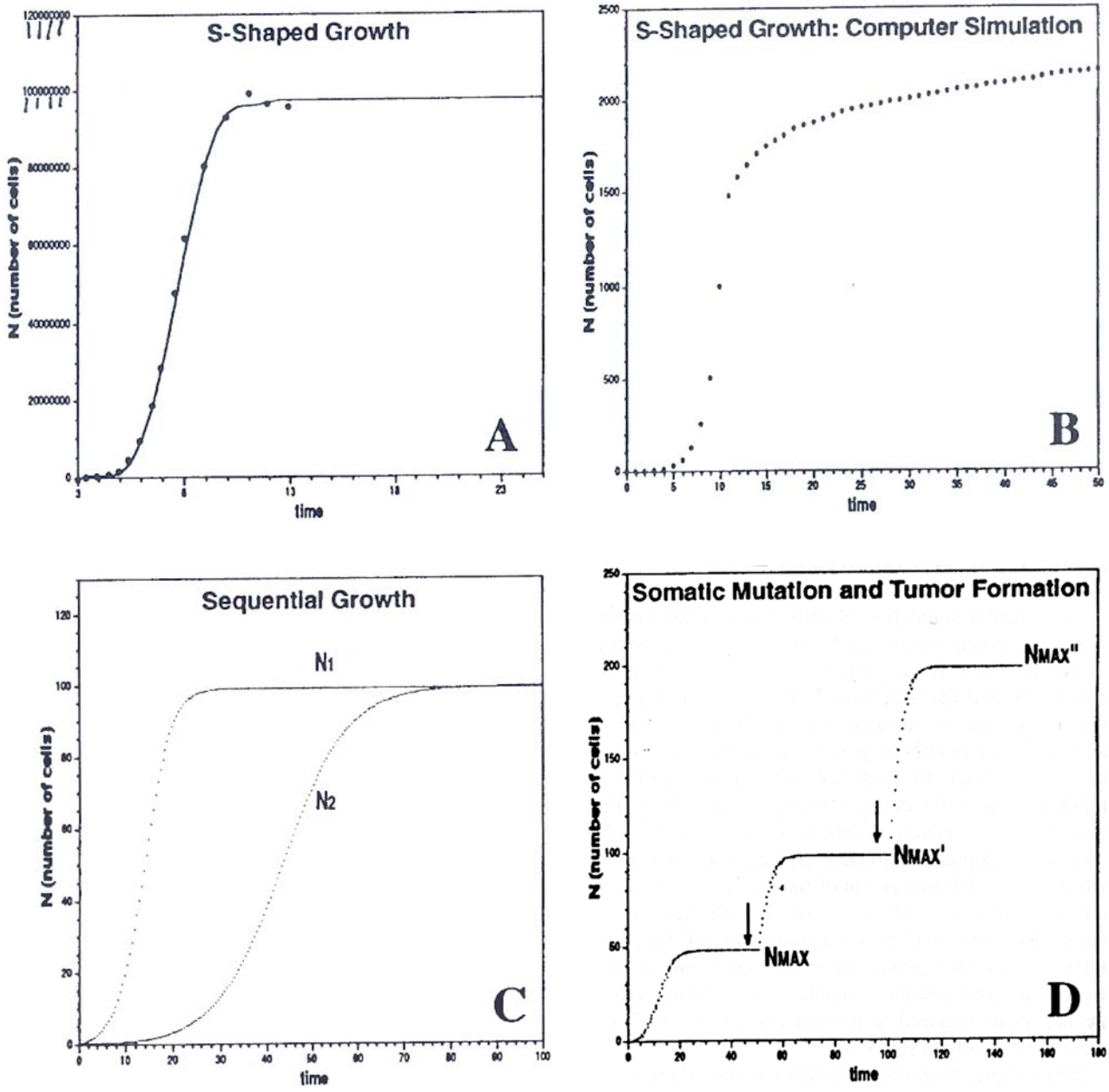


Figure 1A. *The empirical S-shaped curve of cellular proliferation Shown is data on the number of cells in the chick neural retina (4), but the growth of many tissues, organs anatomical structures, and whole organisms fit curves of this general form(4-8,25,26). S-shaped growth typically begins with exponential growth* ($dN/dt=rN$) *but gradually makes the transition to zero growth* ($dN/dt=0$), *when the cellular population has reached its maximum number cells* N_{MAX}. *In this example,* $N_{MAX} \sim 2\times 10^8$ *cells.*

1B. *S-shaped growth caused by an inhibitory growth factor by computer simulation Growth in this simulation is of a population of cells in a constant volume, where cell division is influenced by the action of an inhibitory growth factor.* ($h_I=10$, $a=10^6$ *molecules,* $R_O=10^7$ *receptors per cell,* $V=1cc$, $K_D=10^{-6}$ *M/liter* $\sim 10^{18}$ *molecules/liter).*

1C. *Sequential growth caused by the discrete allocation of stimulatory growth factor molecules As shown in the text, the discrete allocation of stimulatory growth factor molecules results in the growth of two populations, such that the size of the population* N_1 *will grow according to the logistic* $dN_1/dt = r\ (N_1 - (1/N_{MAX1})N_1^2)$ *with* $r=1$ *division per unit time,* $N_{MAX1}=a_iR_{Oi}/VK_{Di}$ *and* $V=1cc$, $K_D=10^{-6}$ *M/liter* $a_i = 10^6$ *molecules,* $R_O=10^7$ *receptors per cell), while the growth of the population* N_2, $dN_2/dt=r\ (N_2-(1/N_{MAX2})N_2^2)\ N(1-(1-1N)^{JN1N2})$ *where* $J= a_sR_{Os}/VK_{Ds}$ *and* $K_{Ds}=10^{-6}$ *M/liter* $a_s = .5 \times 10^6$ *molecules,* $R_{Os}=10^7$ *receptors per cell).*

1D. *The change in the growth curve caused by multiple somatic mutations The general progress to higher and higher values of* N_{MAX} *resembles the actual multi-step process of tumor formation. Times of somatic mutation indicated by arrows.*

iteration, the location of each molecule was assumed to be under the random influence of diffusion; there were no external spatial cues whatsoever. As in the examples cited in the previous paragraphs, whether or not a cell would divide depended upon whether it had bound a sufficient number of growth factor molecules. Surprisingly, we found that very complex structures would emerge quite spontaneously from this simple discrete allocation of growth factor molecules. For example, in the bottom panel in Figure 2 is shown the results of one such simulation, which examined the growth of a population of cells (all colored red in this figure) producing a self-acting inhibitory growth factor. Initially, the population grew as a rough ball, but after a period of time, "arms" began to appear, and the population took on the appearance of a starfish. Above this image in Figure 2, are displayed the results of another simulation, except involving two types of cells, shown in red and green, and in which each cell type was modeled as if it made an inhibitory growth factor that acted only on cells of the other type. Again, growth initially began as a rough ball comprised of a mix of cells of both cell types, but gradually differential cellular proliferation gave rise to a structure in which the two cell types came to comprise opposite sides of a "dumbbell" shaped structure. Repetition of both of these simulations reproducibly lead to shapes of the same general type. Thus, the growth of populations of cells to particular ***shapes***, like the growth of populations to particular ***sizes*** and at particular ***times,*** can be the direct consequence of the use of molecular discreteness.

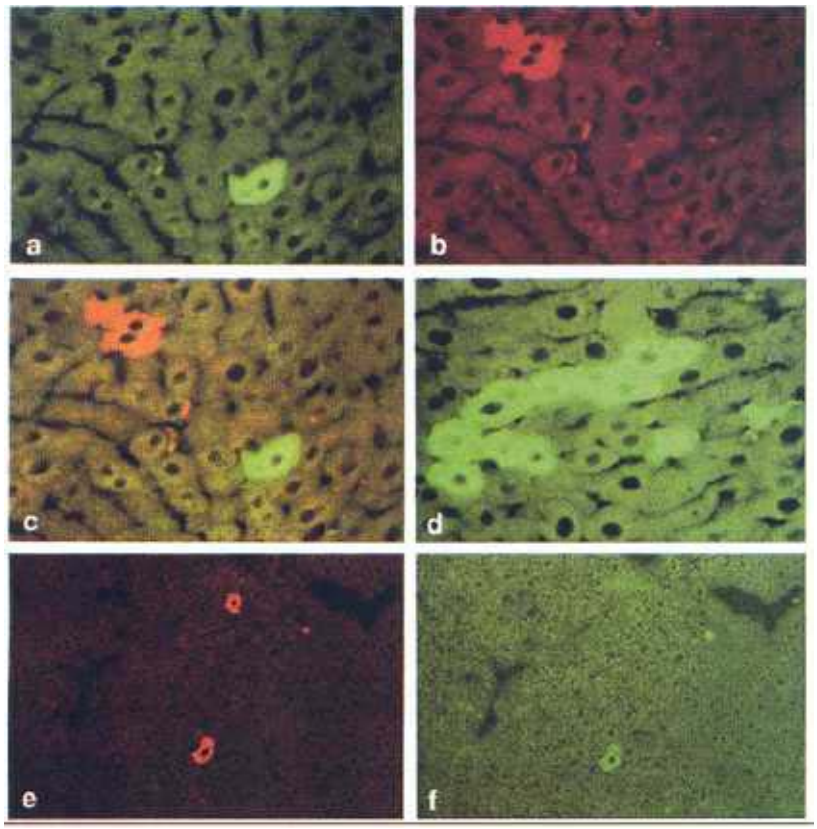
GENE EXPRESSION

Of course, growth factors are not the only proteins present in cells in small numbers of molecules. The gigantic macromolecules that we call genes and chromosomes are also present in cells in very few numbers. Every human cell, being diploid, has just ***two*** chromosome-6's, ***two*** globin genes, ***two*** nucleotide molecules at position #43 of a globin gene. These individual genes/molecules are subject to the same ***either/or*** quality that we have already seen to occur at the cell surface. ***Either*** a particular regulatory molecule binds to a particular DNA sequence, ***or*** it doesn't. ***Either*** a particular nucleotide in a gene is methylated ***or*** it isn't. Might we not expect, then, that this microscopic either/or chemistry of individual genes would be translated into an ***either/or*** quality seen in the expression of genes (8)? In fact, in a great variety of instances, where gene expression has been examined ***one cell at a time***, exactly this ***either/or*** quality of gene activation has been seen (8,9-24).

The first, and perhaps best known, instance of such random gene activation was identified by Till, McCulloch, and Siminovitch (9) in the creation of hematopoietic stem cells. They found that the choice of such cells to differentiate, or remain multipotential, is essentially a random process. Furthermore, by employing the statistical methodology developed by Till, the occurrence of random gene expression was subsequently found in a number of additional instances, including melanogenesis (as found by Bennett (10)), globin gene expression (as shown by Levenson and Housman (11), Gusella et al. (12) and Orkin, Harosi and Leder (13)), myogenesis (as revealed by NadalGinard (14)), and terminal differentiation (as has been described by Smith and Whitney (15)).

My colleagues and I have found a very dramatic example of this ***either/or*** quality in gene expression in the liver's synthesis of plasma proteins (8). If we homogenize a piece of liver, and run a northern blot, each plasma protein mRNA will be seen in amounts roughly reflecting their levels of synthesis. However, this overall level of plasma protein synthesis reflects a more subtle expression of cellular heterogeneity, as we have found by examining plasma protein gene expression by immunofluorescence (8). By this method, we have found that each plasma protein is present in, and presumably synthesized by, a small, separate, subpopulation of the liver's parenchymal cells (Figure 2). Indeed, my colleagues and I have found albumin to be in a bit less than 1% of the liver's hepatocytes, with each of the other plasma proteins present in separate populations of hepatocytes, in numbers reflecting their relative rates of synthesis (8). Furthermore we also found that the reason why each protein is expressed in a small number of cells is that plasma protein gene activation occurs with just the same sort of ***either/or*** quality that we have earlier seen for ligand binding at the cell surface. This could be seen in the livers of mice heterozygous for a structural polymorphism of albumin ($Alb1^a/Alb1^c$). We produced an antiserum that reacts with just one of the allelic forms (antiAlb1c) (8). With this reagent, we found that in the livers of $Alb1^a/Alb1^c$ heterozygous mice, some albumin-producing cells express the $Alb1^a$ form of albumin, while other albumin-producing cells express the $Alb1^c$ form of albumin, but there are few, if any, cells that express both forms (Figure 2e,f) (8). Such a situation could only occur if the albumin genes were turning on randomly.

The expression of just one of the two allelic forms of a gene, such as occurs in the liver, is striking, but, in fact there have been a whole series of precedents for this manifestation of the ***either/or*** quality in gene expression. A very similar case has been identified in the expression of olfactory receptor genes in the olfactory epithelium, which, like plasma protein genes, are each expressed in a small number of cells of this organ (16). By examining the reverse-transcribed, PCR expanded, material from single cells, Chess and colleagues have found that neural cells express mRNA from only one of the two allelic forms of an olfactory receptor gene (16). Similar examples of expression of just one of the two allelic forms of a gene, also occur for autosomal coat color genes. Such independent expression of allelic forms of coat color genes is apparent as a patchy, or variegated, appearance in the pelt of heterozygous animals, and has been seen for a half a dozen autosomes in the mouse, as well as for a variety of



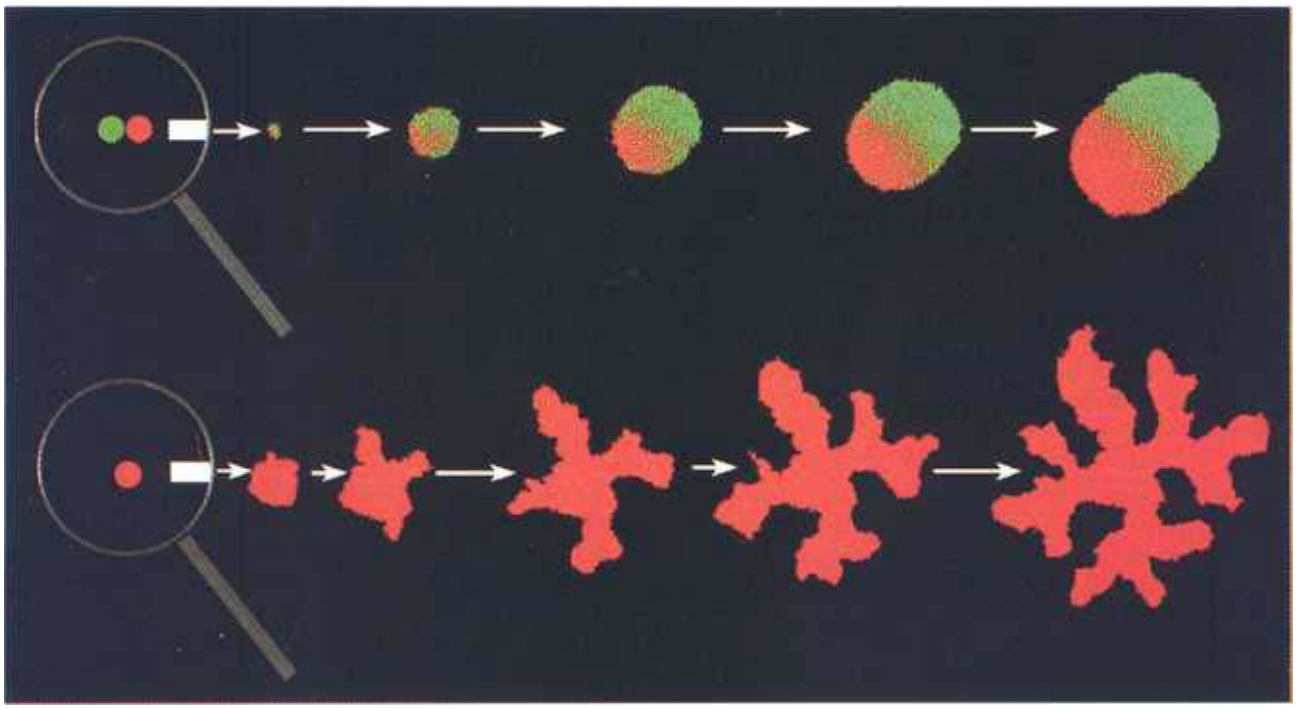


Figure 2a-g. *Immunofluorescence of mouse liver. a)Albumin-containing hepatocyte (green) identified by rabbit anti mouse albumin in mouse liver. b)Two complement component C3containing hepatocytes (red) in same field as figure 1a, identified by goat antimouse complement component C3 antiserum. c)Double exposure of same fields as Figures 1a and 1b, showing simultaneously albumin(green) and complement component C3(red) containing cells. d)Cluster of albumin-containing hepatocytes. e)Two albumin-containing cells in a liver from a* $(Alb1^c/.Alb1^a)_{F1}$ *mouse, stained red by a rhodamine-tagged heteroantibody to albumin. f)Reactivity of* $AntiAlb1^c$ *[*$(B6xA)_{F1}$ $antiB6.Alb1^c$ *serum] on the same two cells as in Figure 2e, only one of which is reactive. Of 735 such albumin-containing cells examined in 5 heterozygous mice, just 49% were Albc+.*

Bottom of Panel: Morphogenesis in a computer? Shown are images of two simulations of cell growth under the influence of inhibitory growth factor molecules diffusing in space. TOP: several selected iterations of the growth from two cells (one red, the other green) in which each type of cells make inhibitory growth factor molecules that act on the other type. BOTTOM: The growth from a single red cell (red), in which cell growth is inhibited by the action of a self produced, self acting inhibitory growth factor.

autosomal genes in guinea pigs, cattle, rabbits, dogs and pigs (8). Globin genes have also been found to express this independent expression of allelic genes, as has been observed in the erythroid precursors from children heterozygous for a fetal globin gene (17). An absence of coordination of allelic genes has been identified with respect to the state of DNA methylation, as has been shown by Silva and White (18). Finally, the independent expression of allelic genes has been found for expression of the interleukin-2 locus by Hollander *et al* (19)

Of course, the ***either/or*** quality of gene expression that can be seen when cells are examined ***one cell at a time*** may be hard to see when we examine the large cellular populations that make up the tissues, organs, and anatomical structures of the body. Many millions of cells are required for most biological and molecular-biological assays, such as SDS gels, northern, or western blots, and thus it may be hard to see the fine scale cellular heterogeneity present among individual cells.

Another reason why the signs of the ***either/or*** quality of gene expression may escape our notice is that differential cellular proliferation, occurring after the creation of cellular heterogeneity, may obliterate the signs of the earlier events of stochastic gene expression. The classic example of this may be seen in the immune system, where a great variety of lymphocytes is generated by random assembly and activation of immunoglobulin and T-cell receptor genes. Unless we examine lymphocytes one cell at time, we may not see this stochastic quality, because it is erased by the clonal selection that organizes the immune response. My colleagues and I have seen an analogous, if perhaps simpler, example of multicellular organization by cellular selection among the plasma protein producing cells of the liver (Figure 2). Thus, we have found that during inflammation, the liver is able to produce more fibrinogen because differential cellular proliferation of fibrinogen-producing hepatocytes causes the liver to contains an increased number of these cells (8). Likewise, we have also found that young mice and rats are able to produce high levels of α-fetoprotein early in life because the α-fetoprotein producing cells divide more rapidly, and thus become more plentiful, than the other plasma protein producing cells in the liver (8). This raises the question of whether cellular organization by cellular selection is really an exclusive specialty of the immune response. During embryonic development especially, differential cellular proliferation is a widespread occurrence that creates the tissues, organs and anatomical structures of the body (8,20-24). Perhaps early metazoans have turned to this simple, if inelegant, process to create many aspects of multicellular organization. If this is so, then perhaps one key to multicellular organization lies in the ability to link the randomizing action of molecular discreteness on the genome to the organizing action of molecular discreteness on cellular proliferation.

CARCINOGENESIS

What is cancer

As we noted above, many of the normal cellular populations of the body grow to their final sizes by S-shaped growth curves. Surprisingly, when tumors have been subject to the same sort of analysis, they have also usually been found to display S-shaped growth curves, but with N_{MAX}'s that are so large as to be lethal (8,25,26). For example, the BICR/A8 murine adenocarcinoma grows by an S-shaped curve to a projected $N_{MAX}$of 700 grams (25); it's not the shape of this tumor's growth curve, but the size of it's N_{MAX} that is lethal to a 50 gram mouse! Human breast cancers have also been found to grow by S-shaped growth curves growing towards projected N_{MAX}'s of $\sim 3 \times 10^{12}$ cells, but they cause death at about 10^{12} cells (26). It's almost as if tumors make a half-hearted effort at growth control that is normal in quality, but lethal in quantity.

By considering molecular discreteness at work in mitotic signaling, we can examine, in a quantitative fashion, how somatic mutations lead to cancerous growth

The math that we have developed above allow us to gain insight into what causes the S-shaped growth curves of cancers, with their disastrously large N_{MAX}'s. Let us return to Equation#1, where we saw how the simple discrete allocation of inhibitory growth factor molecules among cells can result in S-shaped growth to an N_{MAX} whose value is:

(1a) $$N_{MAX} = (VK_D/a)\, h_I\, (1/R_O)$$

Recall that V is the volume of the organism, K_D is the dissociation constant of the ligand for its receptor, **a** is the number of ligand molecules that are produced by each cell, h_I is the minimum number of bound ligand molecules necessary to prevent a cell from dividing, and R_O is the number of receptors per cell for the ligand.

Equations #1 and #1a give us a way to see, in a quantitative fashion, how somatic mutations can lead to changes in the N_{MAX}'s of cellular populations. Let us consider the case where there are ***two*** autosomal ***pairs*** of inhibitory growth factor receptor genes, for a total of ***four*** homologous receptor genes. In terms of Equation #1a, we would say that $R_O=4q$, where **q** is the number of cell surface molecules produced per gene. (We shall assume for the simplicity of this example that all four genes are identical in type.) Equation #1a allows us to see the effect on N_{MAX} if any of these genes became mutated:

(1b) $$N_{MAX} = (VK_D/a)h_I / [4q]$$

In fact, mutations for the genes for the receptor of the inhibitory growth factor TGF-β have been found to be

associated with cancer development (27), so our example is not unrealistic. Let us consider what would happen if there were a mutation resulting in the deletion of just one of the four hypothetical inhibitory growth factor receptor genes:

(1c) $N_{MAX}' = (VK_D/a)h_I / [3q] = 4/3\ N_{MAX}$

Note that such a change will result in the rather minor **33%** increase in the cellular population's N_{MAX}. If there were a second deletion:

(1d) $N_{MAX}'' = (VK_D/a)h_I / [2q] = 4/2\ N_{MAX}$

this will result in an increase in the value of N_{MAX} by a more impressive **100%** increase in size. Yet another mutation will result in:

(1e) $N_{MAX}' = (VK_D/a)h_I / [1q] = 4/1\ N_{MAX}$

That is, a population growing to an N_{MAX} that is *four times* bigger than the initial size.

Note how Equations #1b-1e show how each sequential mutation raises the value of N_{MAX} by a notch (Figure 1D). ***This is remarkably reminiscent of the stepwise progression from preneoplastic disease to outright cancer, where a clonal cellular population evolves from its normal N_{MAX}, to N_{MAX}'s that lead to hyperplasia, and finally to N_{MAX}'s that are so large as to be incompatible with life (43).***

With a discrete model of the chains of mitotic signaling that go on within cells, it is possible to estimate, in a quantitative fashion, the effects of oncogene mutations on growth

Of course, Equation #1 describes the N_{MAX} for a very simple example of growth. However, the very same discrete approach can be used to model systems of mitotic signaling that are far more complex and realistic (28). The general method used to assemble such more real-life cases is analogous to the method used to construct Equation #1, although more tedious, and readers who are curious to see how these equations are strung together are referred to my webpage (http://webm 9120.ntx.net/BreastCancer Math. html) for details. Here, I will illustrate just one example, the case of the **Ras** pathway. Following what we know of the biology of this pathway, (29,30), we shall consider the case where ligand binding event leads to receptor dimerization and phosphorylation, then **Shc** phosphorylation, then **Grb2** binding, then **Sos** binding, then **Ras** binding, then **Ras-GTP** association, then **Raf** binding, then **MEK** phosphorylation, then **MAP** kinase phosphorylation of transcriptional proteins (perhaps **fos, c-jun, c-myc, or myb**), which induce the expression of **cyclin D**:

(2) $$N_{MAX} = (VK_D/R_Oa)\ (1/Cn)\ (V_C^2 / k_b\ E_b)\ Rf_{RAS}$$

where $Rf_{RAS} = y\ k'' (\beta_1 / V_C)(\gamma / V_C)$

γ = the number of cyclin D genes

and where $\beta_1 = [M\text{-}P]V_C = \{(\beta_2/V_C)(V_{MAX2})(S_M/V_C)(1.49t_{1/2M\text{-}P})\}/K_{M2} + (S_M/V_C)$

where $\beta_2 = [MAP\text{-}P]V_C = \{(\beta_3/V_C)(V_{MAX3})(S_{MAP}/V_C)(1.49t_{1/2MAP\text{-}P})\}/K_{M3} + (SMAP/V_C)$

where $\beta_3 = [MEK\text{-}P]V_C = \{(\beta_4/V_C)(V_{MAX4})(S_{MEK}/V_C)(1.49_{t1/2\ MEK\text{-}P})\}/K_{M4} + (S_{MEK}/V_C)$

where $\beta_4 = [Ras\text{-}GTP\text{-}Raf]V_C = k''' (\beta_5/V_C)(1.49_{t1/2Ras\text{-}GTP\text{-}Raf})$

where $\beta_5 = [Ras\text{-}GTP]V_C = \{(\beta_6/V_C)(V_{MAX5})(S_{Ras\text{-}GDP}/V_C)(1.49t_{1/2Ras\text{-}TP\text{-}P})\}/K_{M5} + (S_{Ras\text{-}GDP}/V_C)$

where $\beta_6 = [Grb\text{-}P\text{-}SOS]V_C = k'''' (\beta_7/V_C)(1.49_{t1/2Grb\text{-}P\text{-}SOS})$

where $\beta_7 = [Grb\text{-}P]V_C = \{(\beta_8/V_C)(V_{MAX6})(S_{Grb}/V_C)(1.49_{t1/2Grb\text{-}P})\}/K_{M6} + (SGrb/V_C)$

where $\beta_8 = [Shc\text{-}P]V_C = \{(\beta_9/V_C)(V_{MAX7})(S_{GShc}/V_C)(1.49_{t1/2Shc\text{-}P})\}/K_{M7} + (S_{Shc}/V_C)$

where β_9 = concentration of dimerized receptor molecules activated by ligand binding = $k''''' (q/2)(E/V_C)2(1.49t_{1/2receptor\ dimer})$

Although this expression is far more complicated, and far more realistic, than Equation #1, with it we can use exactly the same approach to examine how somatic mutations can lead, step by step, from normal N_{MAX}'s, to N_{MAX}'s that are hyperplastic but benign, to N_{MAX}'s that are lethal. Furthermore, expressions like Equation #2 allow us to predict the behavior of actual, rather than hypothetical, biological situations. ***Thus, the mathematics of molecular discreteness gives us a practical arithmetic with which we can examine the molecular biology of cancer in quantitative terms.***

SCREENING FOR CANCER

Once groups of cells have become cancerous, that is have converted to having N_{MAX}'s that are lethal, they can still be eradicated if detected early enough for the tumor to be removed entirely. Perhaps the greatest success in this approach has been achieved for breast cancer, where early detection has been shown to result in remarkable reductions in breast cancer death. Indeed, controlled randomized trials have shown that mammogramic examination, carried out every one to three years, will improve the chances of surviving breast cancer by 30% (31-37). This is a considerable achievement, but it also raises the question of whether there might have been a way to save the lives of the 70% of women whose tumors were ***not*** discovered before their cancers spread. As we shall see next, the mathematics of cellular discreteness provides just the method that we need to address such questions (34).

Data are available for breast cancer which makes it possible to estimate that probability, 1/P, that a cell will leave the primary tumor and form a distant metastasis

We can tackle the problem of breast cancer screening by asking whether the spread of a cancer cell from a primary tumor to a distant site might have just the same sort of ***either/or*** event as ligand binding or cell division. Let us call **1/P** the chance, per day, that a breast cancer cell will leave the

primary tumor and form a distant lethal metastasis. Thus, **P** may be thought of as the number of "cells" for the number of "days" it takes until there will be, ***on average, 1*** metastasis per tumor. Thus **P** carries the units which we shall call "celldays". Let us also call **C** the number of "celldays" that a tumor of **N** cells has accumulated. The number of cells (**N**) in a tumor diameter (**d**) can be roughly estimated by the expression $N \sim [4/3\pi(d/2)^3]10^8$ **cells/cc.** To calculate number of "celldays", **C**, that a tumor of **N** cells has accumulated requires a consideration of the rate of tumor growth, which can be estimated from the results of a number of studies (33,39). For example, Peer et al. (39) have found that tumors in patients less than 50 years of age have a median doubling time of 80 days, while the tumors of patients 50-70 years of age had a median doubling time of 157 days, and tumors in patients more than 70 years of age had a median doubling time of 188 days. Such growth information can be incorporated into a simple computer simulation, and the relationship between tumor size, **N**, and the number of "cell-days", **C**, can be estimated. (Again, interested readers may find further detail on my webpage, http://webm9120.ntx.net/BreastCancerMath.html)

Now consider a group of patients whose tumors have grown to a size such that they have accumulated exactly **P** "cell-days" (**C=P**). It would be expected in such a group of patients that, ***on average***, there would be **1** metastasis per patient, since **C x 1/P = P x 1/P = 1.** Of course, this is just an ***average***. While some of these patients would have **1** metastasis, others would have **2** metastases, or **4** metastases, and some patients would have **0** metastases, and will be free of metastatic disease. Using the Poisson distribution, it is possible to calculate that when the ***average number of metastases per patient=1***, the fraction of patients with **0** metastases, that is, ***fraction of metastatic-disease free patients=1/e***, or about **37%**, of patients. On the other hand, for a different group of patients, whose tumors were sufficiently larger to have accumulated, for example, **C=3P** "cell-days", that is, for which the **average number of metastases per patient=3**, then it follows that the ***fraction of metastatic-disease free patients=$1/e^3$***. Thus, in general, if ***C=the number of "cell-days" a tumor has accumulated and P=the number of "cell-days" required to cause 1 metastasis per patient***:

(3) ***the fraction of metastatic-disease free patients*** $= 1/e^{C/P}$
or
(4) ***ln (the fraction of metastatic-disease free patients)*** $= -C/P$

It follows from Equation #4 that if **1/P** is constant, then a graph of ***the LOG of the fraction of metastatic-disease free patients versus the number of "cell-days" that a tumor has accumulated (C)*** will form a straight, downwardly pointing line, which forms an intersection with the origin. Two groups, Tabar *et al* and Tubiana *et al* (31,32,37,38), have collected data on the relationship between the size of the primary breast cancer at the time of surgical removal (from which the tumor's value of **C** can be calculated) and the subsequent incidence of distant metastatic disease. This information can be seen in Figure 3A. These data conforms fairly well to the log-linear predication of Equation#4, and thus provides empirical support for the idea that **1/P** is, in fact, fairly constant (Figure 3A).

Of course, from a biological standpoint, this relative constancy of **1/P** is what would be expected for mechanisms of metastasis formation that are dependent upon processes such as mutation or simple mechanical events such as detachment from the primary tumor, dissémination, survival, and re-engraftment. However, what is uniquely powerful about the information provided by Tabar *et al* and Tubiana *et al* (31,32,37,38) is that it has provided us with a way to measure the ***value*** of this probability, **1/P**, which Equation #4 indicates will correspond to the point on the X-axis where ***the fraction of metastatic-disease free patients=1/e***, or about **37%.** By such an approach, it appears that the value of $1/P \sim 10^{12}$ metastases/"cellday".

We can use our estimate of the probability, 1/P, that a cell will leave the primary tumor and form a distant metastasis, to calculate the time-course of breast cancer metastasis

With this estimate of **1/P,** it is possible to make an initial rough estimate of the appearance of metastases over time (FIGURE 3). After all, the ***average number of metastases formed each day*** is simply the probability that a single cell will form a metastasis **(1/P)** ***times*** the number of cells in the tumor **(N).** Using such an approach, it is fairly straightforward to estimate the probability of metastatic disease from the time of the first breast cancer cell, as well as from the time when the breast cancer can first be detected. The remarkable result of these simulations, and the likely explanation as to why screening can reduce the death rate, is that metastasis occurs quite late in the time course of breast cancer growth, and generally after the minimal sizes usually detectable by mammography (about 1 to 10mm) (34) have been reached.

By a discrete computer simulation of breast cancer growth it is possible to determine the effect of various screening intervals on the reduction in breast cancer death

It is necessary to add several additional features to the simulation illustrated in Figure 3 before it is possible to make practical estimates of the influence of breast cancer screening on the reduction in metastatic disease among women in the population as a whole. ***First,*** breast cancers that appear between screens are detected if they become palpable. Since the median size of the breast cancers seen in Tubiana's studies (37,38) in the pre-mammographic era was about 10^{10} cells, the simulation assumed that tumors which had escaped detection by screening would be identified and removed when they reached this size. ***Secondly,*** because tumors arise with a variety of growth rates and probabilities of metastasis **(1/P),**

the simulation must incorporate a consideration of this feature of breast cancer biology. Tumor doubling times and distributions were taken from Peer *et al* (39) and Spratt *et al* (33), while the justification for the distribution of values for $1/P(\sim 10^{-12})$ may be found on my webpage (http://webm 9120.ntx.net/BreastCancerMath.html). The incidence of metastasis for each of the various types of tumors was calculated, and the weighted contributions of each of the groups were summed to derive an estimate of the incidence of metastasis among women as a whole. ***Thirdly,*** in a screening program, tumors detected by screening come to the attention of the physician at the time of the examination, and not at the time when they reach minimum detectable size. Thus, the simulation was adapted to examine the case where a breast cancer might reach minimum detectable size on any day between mammographic examinations.

The discrete modeling of breast cancer reveals that most breast cancer death should be avoidable if screening is carried out often enough

The results of these more realistic simulations are shown in Figure 3C and Table II. For a screening method that could detect tumors larger than 3mm in diameter ($\sim 10^7$cells), used at intervals such as those now commonly in practice (1 to 3 years), the simulations yields reductions in the incidence of metastatic disease (14%-51%) generally in agreement with the actual experience of randomized mammographic screening (31-37).

The results of the simulation shown in Figure 3 and Table II are of most interest where they show the consequences of the use of screening at intervals not presently employed. In fact, these simulations suggest that women screened every 9 months would have a 66% reduction in metastatic disease in comparison to unscreened women, while women screened every 6 months would have a 78% reduction in the incidence of metastatic disease, and women screened every 3 months would have a 96% reduction in the incidence of metastatic disease.

Thus the results of the simulations shown in Figure 3 and Table II tell us that we should be able to eliminate most breast cancer death, if mammographic screening was utilized often enough! This is a very surprising and encouraging result, which needs now to be tested, but it flows directly out of the mathematics cellular discreteness. Furthermore, although here we have examined breast cancer, the method is sufficiently general that it should be applicable to the screening of other tumors as well.

CANCER CHEMOTHERAPY

We can use the same mathematics of molecular discreteness to calculate the most efficient ways to use chemotherapeutic drugs to treat cancer.

Table II. *Computer-generated estimates of the reduction in the incidence of metastatic disease (in comparison to unscreened women) as a function of the frequency of the screening examination.*

Frequency of examination	Reduction in the incidence of metastatic disease by a method that can detect tumors larger than 10^8 cells ($\sim$12 mm)	Reduction in the incidence of metastatic disease by a method that can detect tumors larger than 10^7 cells ($\sim$3 mm)	Reduction in the incidence of metastatic disease by a method that can detect tumors larger than 10^6 cells ($\sim$1.2 mm)
every third year	7%	14%	16%
every second year	14%	22%	27%
every year	33%	51%	56%
every 6 months	62%	78%	84%
every 4 months	77%	90%	93%
every 3 months	86%	96%	98%
every 2 months	87%	97%	98%

By modeling the discrete interactions of drug molecules and cells, we can understand the effect of anticancer agents on cells over short time periods

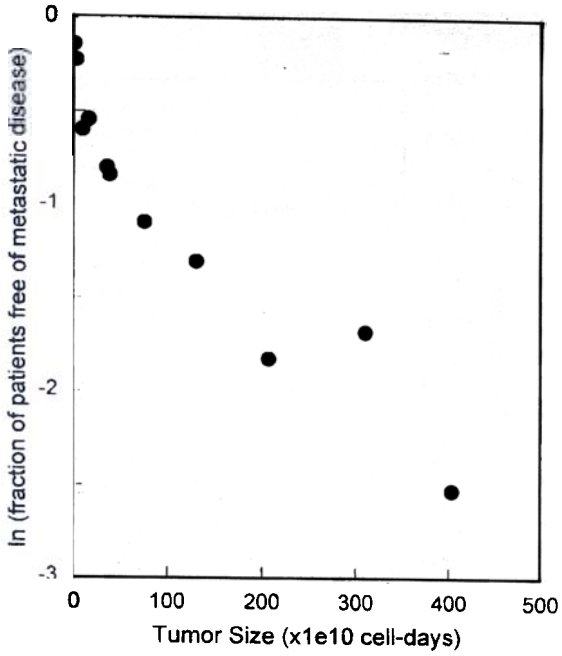
Anticancer agent molecules and cells interact ***one-at-a-time,*** much as growth factor molecules and cells interact ***one-at-a-time.*** Let us examine this interaction. If the chemotherapeutic drug is an alkylating agent (40,41), it will interact with DNA:

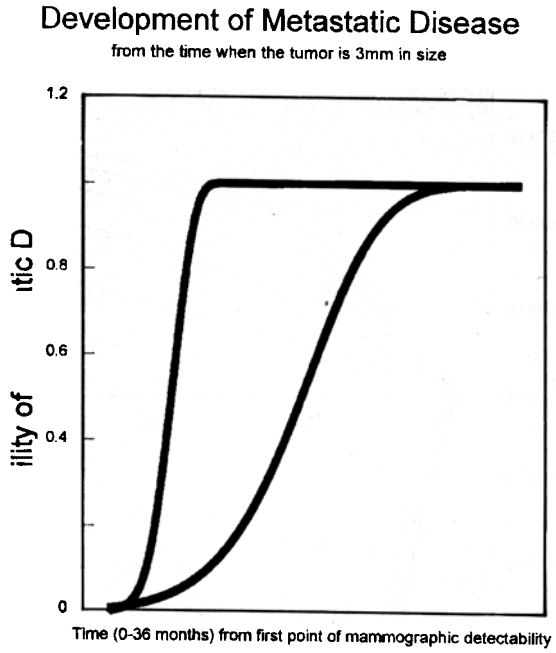
[alkylating agent] + [potentially lethal sites on DNA] ➜ [alkylating agent -DNA]

Chemical reactions of this type may be modeled using the Law of Mass Action, which in its equilibrium form is:

$$(5)\quad k = \frac{[\text{alkylating agent}]\,[\text{potentially lethal sites on DNA}]}{[\text{alkylating agent -DNA}]}$$

Where **k** is a constant. Since most alkylating agents diffuse into cells in a fairly rapid fashion, the intracellular and extracellular concentrations of the alkylating agent can be expected to be approximately the same. Let us define **Q** as ***the number of potential lethal sites on DNA molecules per cell,*** and A as the concentration of the alkylating agent. Since the **DNA**





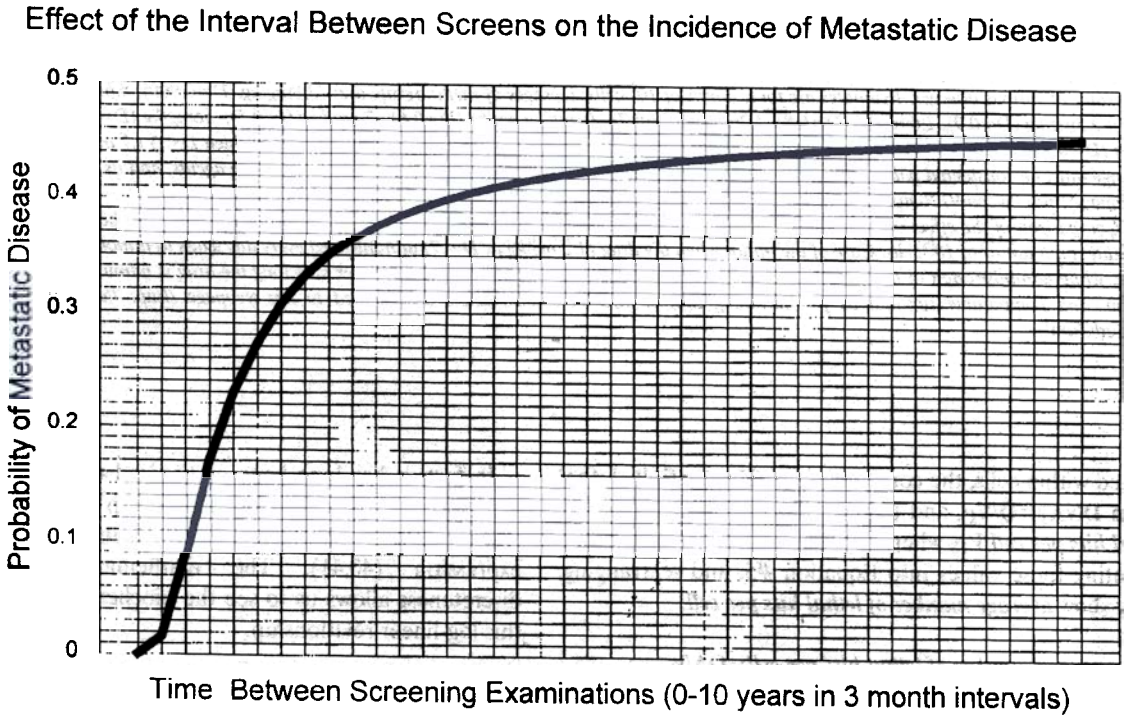


Figure 3. *Top left. Relationship between tumor size (in "cell-days") and the log of the incidence of metastatic disease. Data taken from Tabar et al. and Tubiana et al. (31,32,37,38). For details of the method for calculating the number of celldays for tumors of each size, and the possible variation of values of 1/P compatible with this data, readers are referred to my webpage (http://webm9120.ntx.net/BreastCancerMath.html). Top right. The probability of metastatic disease from the time when the primary tumor is first detectable (10^7 cells). Shown are simulations for tumors with doubling times of 20, 40 and 80 days. Most breast cancers probably fall into one of these categories (see 33,39). $1/P \sim 10^{-12}$. Bottom. Computer simulation estimate of the effect of the interval between screens on the incidence of metastatic disease. For details of the simulation, and its code, readers are referred to my webpage (http://webm9120.ntx.net/BreastCancerMath.html).*

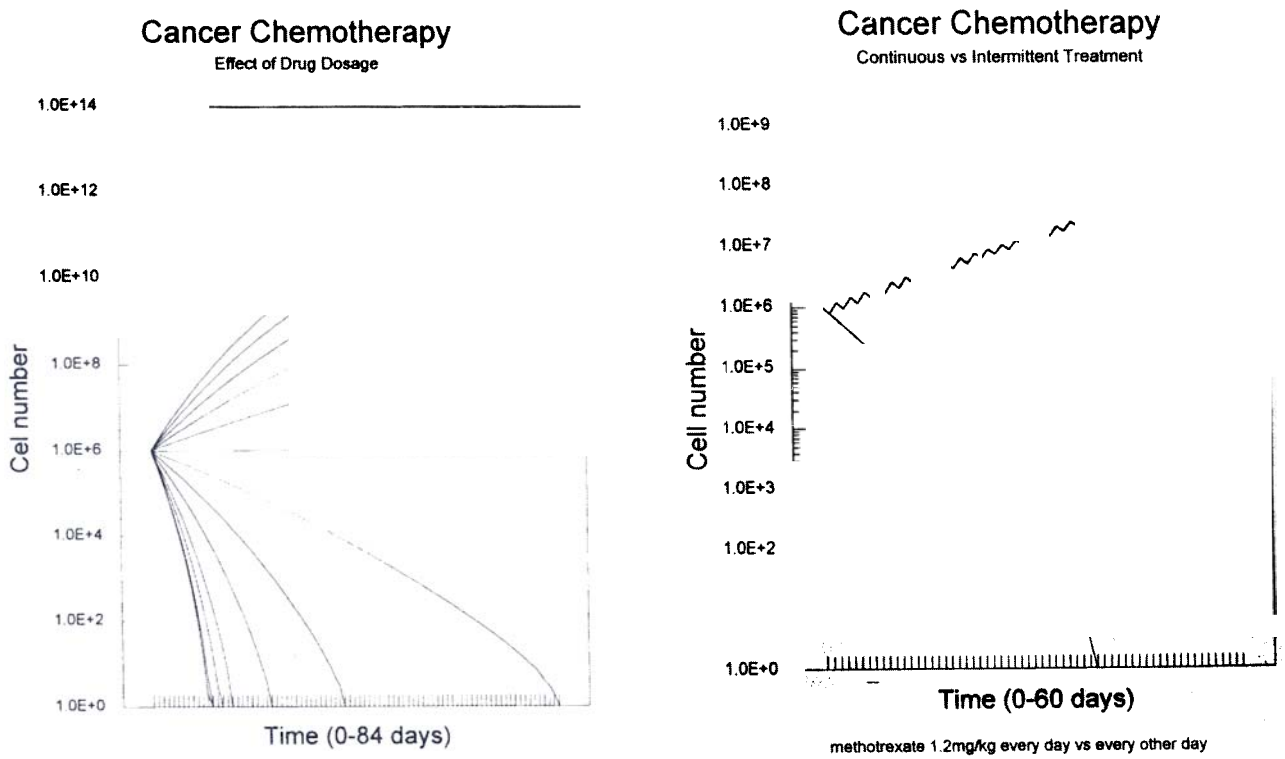


Figure 4. *Top. Calculations of the change in tumor cell number under the influence of a continuous dose of a reversible chemotherapeutic agent (methotrexate) at various dosages. Estimates are for a tumor of 10^6 cells. Dosages of methotrexate are: 13.8mg/kg.(D=0.999), 4.6mg/kg, (D=0.99), 3mg/kg (D=0.95), 2.3mg/kg (D=0.9), 1.6 mg/kg (D=0.8), 1.2mg/kg (D=0.7), .91mg/kg mg/kg (D=0.6), .7 mg/kg (D=0.5), .56mg/kg (D=0.4), .42 mg/kg (D=0.3), .28 mg/kg (D=0.2), .14 mg/kg (D=0.1). Note that the tumor will continue to grow when exposed to dosages for which D<.5 (.7 mg/kg), and thus treatment in this range must be considered palliative. For details of the simulation, and comparable simulations for alkylating agents, and its code, readers are referred to my webpage (http://webm9120.ntx.net/BreastCancerMath.html).*
Bottom. Comparison of the effect of continuous versus intermittent dosages of a reversible chemotherapeutic agent. Shown is the action of a drug such as methotrexate, on a tumor of 10^6 cells. Cure of the tumor when the drug (1.2 mg/kg or D=.7) administered every day leads to tumor eradication in 42 days of treatment, while administration of the same dose, but given every other day, is not curative, no matter how long the drug is administered. For details of the simulation, and comparable simulations for alkylating agents, and its code, readers are referred to my webpage (http: //webm9120.ntx.net/ Breast CancerMath.html).

is located within cells, the concentration of **[potentially lethal sites on DNA]=Q/V_C**, and **[ALKDNA]=*the average number of lethal hits per cell*/V_C**, where **V_C**= the volume of the cell. By substituting these values into Equation #5, and rearranging, we may show ***average number of lethal hits per cell:***

(6) ***average number of lethal hits per cell*=QA/k**

Let us call **D** the fraction of cells killed by a chemotherapeutic drug, and thus **1-D** is the fraction of cells which survive. Using the Poisson distribution, we know that when ***average number of lethal hits per =QA/k***, we can expect $1\text{-}D=1/e^{AQ/k}=e^{-AQ/k}$, or:

(7) ***ln* (1-D) = -n A** where **n = Q / k**

In fact, it has long been recognized that, when examined over short time periods, the relationship between cell number and drug dose often conforms to just such a log-linear expression (40,41). The mathematics of molecular discreteness allows us to see the biochemical explanation for this log-linear relationship.

A discrete, one-day-at-a-time, estimate of the effects of drugs on cells allows us to model the competing forces of tumor growth and tumor kill which determine the effects of cancer chemotherapy over long time periods

Equation #7 is sufficient if we want to know the effect of a drug over short periods of time, but when we use these drugs over the days, weeks, and months required to treat cancer, we must also take into account the countervailing influence of

the tumor's regrowth. We can calculate the effect of chemotherapy in such a long-term setting if we break the problem down into single day increments. That is, if a tumor, today, is comprised of N_{today} cells, then tomorrow the tumor will be comprised of $N_{tomorrow}$ cells, which is simply:

(8) $N_{tomorrow}$ =
(the number of today's cells)
- (the number of cells that would die anyway)
- (the number of cells killed by the drug)
+ (the number of cells created by the mitosis of the cells not killed by the drug)

or:

$$(9)\ N_{tomorrow} = N_{today} - [L\ N_{today}] - [D\ G\ (N_{today} - LN_{today})] + [(1\text{-}D)\ G\ (N_{today} - LN_{today})]$$

Where **L** is defined as the natural cell loss fraction, and **G** is the growth fraction, as was discussed at the outset of this paper. Of course, if Equation†#9 allows us to calculate the number of cells present tomorrow, we should be able to repeat the process to calculate the number of cells the ***day*** after ***tomorrow,*** and so on, until the tumor is eradicated. Thus, by following the shrinking of the tumor, one-day-at-a-time, we are able to calculate the number of days that it takes to achieve the cure of a tumor, t_{CURE}.

Let us look at an example of this ***discrete,*** one-day-at-a-time, method for drugs, such as methotrexate and 5FU, that only kill cells that are dividing at the time they are exposed to the drug. (The alkylating agents interact irreversibly with cells, and to calculate their effect on tumors requires a somewhat more complicated method). Let us also consider what would happen to a tumor for which there is little cells loss (*i.e.* $L \sim 0$). There are a number of methods for estimating the value of G, but for this example, we can use the Gompertz equation, $[G=1(\ln N/\ln N_{MAX})]$, where $N_{MAX}=10^{12}$, which has been found to describe the behavior of breast cancer (26; see also my webpage "http://webm9120.ntx.net/Breast Cancer Math.html" for further details). Finally, as we have shown above, **D** can be estimated with the expression $\ln(1D)=An$, where **A** is the concentration of the chemotherapeutic agent and **n** is a constant, for which empirical analysis for methotrexate has been shown to be ~**2.3mg/kg** (42). Let us, then, put these features into Equation #9.

$$(10)\ N_{tomorrow} = N_{today} - [D\ (1(\ln N_{today}/\ln 10^{12}))\ (N_{today})] + [(1\text{-}D)\ (1(\ln N_{today}/\ln 10^{12}))\ (N_{today})]$$

Should, for example, the tumor contain **1,000,000** cells (*i.e.* $N_{today}=10^6$), and be treated with a dose of drug that would kill **85%** of all dividing cells (**D=.85, A=1.9 mg/kg**), then using a calculator, we may easily determine that $N_{tomorrow}=650{,}000$ cells. The process can then be repeated to determine the number of cells the ***day-after tomorrow*** (**415,406** cells). In fact, it takes **22** such calculations before the number of cells is reduced to below **1** cell, so the time required to cure this tumor (t_{CURE}) is **22** days.

Of course, it is much easier to do these calculations by computer simulation, and results of these simulations are shown in Figure 4. Using such simulations, we have calculated the t_{CURES}'s for tumors with these properties of various sizes (N_0's), and at various doses of the chemotherapeutic drug (Figure 4 Bottom). For example, these simulations show that it is possible to achieve the cure of such a tumor of 10^6 cells by treatment with methotrexate if given at a dose of 3.0 mg/kg (**D=.95**) for **15** days, or **2.3** mg/kg (**D=.90**) for **18** days, or **1.9** mg/kg (**D=.85**) for **22** days. These predictions are highly testable in experimental tumors in animals, and should also be applicable to the actual treatment of cancer in people.

These computer simulations also allow us to examine the effect of ***drug scheduling*** on tumor cure (Figure 4). For example, while treatment of a tumor of 10^6 cells with 1.6 mg/kg of drug (**D=.8**) given every day will result in the cure of the tumor after 26 days of treatment (Figure 4) shows that treatment of the ***same*** tumor with the ***same*** drug dose given ***every other day*** will result in growth, rather than cure, of the tumor. ***These results make dramatically clear that administration of a chemotherapeutic drug every day is vastly more efficient than an intermittent schedule of the same drug at the same dose.***

SUMMARY

Let us recap. The things of which we are made - electrons, atoms, molecules - are ***discrete*** things. This molecular discreteness imparts an inevitable ***either/or*** quality onto the reactions of individual molecules, and then onto the behavior of individual cells. We have seen a dramatic example of this for the growth factor IGFII, for which there is normally just about ***one*** molecule of IGFII bound for every ***six*** cells, and many other mitotic signaling proteins also act in this realm of small numbers of molecules (TABLE I). The consequences of these microscopic discrete events for organisms as a whole can be far reaching. The discrete allocation of growth factor molecules among cells can provide multicellular organisms with a way to limit cell division to a fraction of cells, and thus to control growth. Indeed, as we have seen, populations of cells may use this consequence of molecular discreteness to achieve all of the essential features of organized growth: growth to predictable ***sizes,*** and at predictable ***times,*** and to predictable ***shapes.***

Of course, it is not only mitotic signaling molecules that are subject to the effects of molecular discreteness. The huge molecules that we call genes are also present in small numbers per cell, and are subject to the same ***either/or*** quality of molecular discreteness. Whether or not a specific site on a specific gene will bind a specific regulatory factor, or be methylated, or bind a polymerase, will be an ***either/or*** event. Apparently, these molecular ***either/or*** events affecting the chemistry of the gene are translated into cellular ***either/or*** events of gene expression, as has been seen in a great variety of cases when gene

expression has been examined one cell at a time (8,20-24).

The stochastic feature of gene expression that has been found when cells are examined one at a time (8,2024) may seem to us messy and surprising, but it's not without its uses to the organism. One of the tasks of embryonic development is to generate a great variety of cell types, and stochastic gene activation can achieve this end (8,21,22). The immune system has long been recognized to use just such a stochastic mechanism to generate the variety of lymphocytes that it requires. It then relies upon differential cellular proliferation to the create the cellular order that lies behind an immune response. Perhaps metazoans also use this inelegant mechanism of stochastic gene activation to produce the many varieties of cells that make up our tissues, organs and anatomical structures of the body. If so, then differential cellular proliferation, rather than control of gene expression, may be the driving force behind much of the order of the multicellular organization (8,2024).

We have been able to "eavesdrop" on the body's use of molecular discreteness by mathematical modeling and computer simulation. As we have seen, such discrete modeling has given us insight into the behavior of both of normal and cancerous cellular populations. With this approach, we have been able to create a computer simulation can try out, in a few seconds, the effectiveness of various regimens of cancer screening and treatment that would take decades to assess by clinical trial. In short, by modeling molecular discreteness, we have a very practical way to gain insight into, and power over, the populations of cells of which we are comprised.

Acknowledgements

Much of this paper was written in the wonderful library of the Marine Biological Laboratory, Woods Hole, Massachusetts.

References

1 Brooks RA: Random transitions in the cell cycle. Progress in Clinical Biology and Research *66:* pt a A593-601, 1981.

2 Bray D: Signaling complexes: biophysical constraints on intracellular communication. Annu Rev Biophys Biomol Struct *27:* 59-75, 1998.

3 Nurse P: Reductionism: The ends of understanding. Nature *387:* 657-658, 1997.

4 Baylor DA, Lamb TD, Yau KW: Response of retinal rods to single photons. J Phsiol *288:* 613-634, 1979.

5 Schlessinger J, Bar-Sagi D: Activation of Ras and other signaling pathways by receptor tyrosine kinases Cold Spring Harbor Symposium on Quantitative Biology *LIX:* 173-179, 1994.

6 Zeiger SL, Harlow SD: Mathematical models from laws of growth to tools for biological analysis: Fifty years of Growth. Growth *51:* 1-21, 1987.

7 Bertanalfy M: Principles and Theory of Growth. *In:* Fundamental Aspects of Normal and Malignant Growth (Wiktor Nowinski, ed) 1960.

8 Michaelson J: Cellular selection in the genesis of multicellular organization. Laboratory Investigation *69:* 136-151, 1993.

9 Till JE, McCulloch EA, Siminovitch L: A stochastic model of cell proliferation based upon the growth of spleen colonyforming cells. PNAS *51:* 29-36, 1964.

10 Bennett DC: Differentiation in mouse melanoma cells: initial reversibility and an onoff stochastic model. Cell *34:* 445-453, 1983.

11 Levenson R, Housman D: Commitment: How do cells make the decision to differentiate? Cell *25:* 5-6, 1981.

12 Gusella J, Geller R, Clarke B, Weeks V, Housman D: Commitment to erythroid differentiation by Friend erythroleukemia cells: a stochastic analysis. Cell *9:* 221-229, 1976.

13 Orkin SH, Harosi FI, Leder P: Differentiation in erythroleukemic cells and their somatic hybrids. Proc Natl Acad Sci *72:* 98-102, 1975.

14 Nadal-Ginard B: Myoblast differentiation in the absence of DNA synthesis. Cell *11:* 855-864, 1978.

15 Smith JR, Whitney RG: Intraclonal variation in proliferative potential of human diploid fibroblasts: stochastic mechanism for cellular aging. Science *208:* 82-84,,1980.

16 Chess A, Simon I, Cedar H, Axel R: Allelic inactivation regulates olfactory receptor gene expression. Cell *78:* 823-34, 1994 .

17 Kidoguchi K, Ogawa J, Karam D, Schneider G, Carpentieri U: Synthesis of Hemoglobin F-maltaI in culture by adult erythropoietic precursors. Blood *55:* 334-337, 1980.

18 Silva AJ, White R: Inheritance of allelic blueprints for methylation patterns Cell *54:* 145-52, 1988.

19 Hollander GA, Zuklys S, Morel C, Mizoguchi E, Mobisson K, Simpson S, Terhorst C, Wishart W, Golan DE, Bhan AK, Burakoff SJ: Monoallelic expression of the interleukin-2 locus Science *279:* 2118-2121, 1998.

20 Till JE: Cellular diversity in the blood forming system. American Scientist *69:* 522-527, 1981.

21 Michaelson J: Cell selection in development. Biol Rev *62:* 115-139, 1987.

22 Michaelson J: The significance of cell death. *In:* Apoptosis (Cope FO, Tomlei LD eds). Cold Springs Harbor Press, 1991.

23 Kupiec JJ: A Darwinian theory for the origin of cellular differentiation. Mol Gen Genet *255:* 201-208, 1997.

24 Atamas S: Self-organization in computer simulated systems. Biosystems *39:* 143-151, 1996.

25 Steele G G: Growth Kinetics of Tumors. Clarendon Press, Oxford, 1977.

26 Norton L: A Gompertzian model of human breast cancer growth. Cancer Research *48:* 7067-7071, 1988.

27 Markowitz S, Wang J, Myeroff L, Parsons R, Sun L, Lutterbaugh J, Fan RS, Zborowska E, Kinzler KW, Vogelstein B, Brattain M, Willson JKV: Inactivation of the type II TGF-β receptor in colon cancer cells with microsatellite instability. Science *268:* 1336-1338, 1995.

28 Nurse P: Reductionism: The ends of understanding. Nature *387:* 657-658, 1997.

29 Eagan SE, Weinberg RA: The pathway to signal achievement. Nature *365:* 781-783, 1993.

30 MacAra IG, Lounsbury KM, Richards SA, McKierna C, Bar-Sagi D: The Ras family of GTPases. FASEB Journal *10:* 625-631, 1996.

31 Tabar L, Fagerberg G, Chen H-S, Duffy SW, Smart CR, Gad A, Smith RA: Efficacy of breast cancer screening by age. Cancer *75:* 2507-2517, 1995.

32 Tabar L, Fagerberg G, Duffy SW, Dat NE, Gad A, Grontoft O: Update in the swedish two-county program of mammographic screening for breast cancer. Radiologic Clinics of North America *30:* 187-210, 1992.

33 Spratt JA, Von Fournier D, Spratt JS, Weber EE: Mammographic assessment of human breast cancer growth and duration. Cancer *71:* 2020-2026, 1993.

34 Michaelson J, Halpern E, Kopans D: A computer simulation method for estimating the optimal intervals for breast cancer screening. Radiology. In press, 1999.

35 Smart CR, Byrne C, Smith RA, Garfinkel MA, Letton H, Dodd GD, Beahrs OH: Twenty-year follow-up of the breast cancers diagnosed during the breast cancer detection demonstration project. CA-A Cancer Journal for Clinicians *47:* 135-149, 1997.

36 Kopans DB: Updated results of the trials of screening mammography.

Surgical Oncology Clinics of North America *6:* 233-263, 1997.
37 Tubiana M, Koscielny S: Natural history of human breast cancer: recent data and clinical implications. Breast Cancer Research and Treatment *18:* 125-140, 1991.
38 Tubiana M, Koscielny S: The natural history of human breast cancer: implications for a screening strategy Int J Radiation Oncology Biol Phys *19:* 1117-1120, 1990.
39 Peer PGM, Van DickJ AAM, Hendricks JHCL Holland R, Verbeek ALM: Agedependent growth of primary breast cancer. Cancer *71:* 3547-3551, 1993.
40 Frei E, Antman KH: Combination chemotherapy, dose and schedule. *In:* Cancer Medicine (Holand, J.F. Lea & Febiger edsw) Philadelphia p631-639, 1993.
41 Chabner B: Clinical Strategies for drug treatment. in: Chemotherapy: Principles and Practices. p115 (Chabner BA and Collins eds) JM Philadelphia Lippincott, 1990.
42 Teicher BA, Holden SA, Brann TW, Jones SM, Frei E: influence of scheduling on alkylating agent cytotoxicity *in vitro* and *in vivo*. Cancer Research *49:* 5994-5998, 1989.
43 Morris VB Cowan RC: A growth curve of cell numbers in the neural retina of embryonic chicks. Cell and Tissue Kinetics *17:* 199-208, 1984.
44 Atlas D, Ster ML, Levittzki A: stereospecific binding of propranolol and catecholamines to the β-adrenergic receptor PNAS *71:* 4246-4248, 1974.
45 Levittzki A, Atlas D, Ster ML: The binding characteristics of β-adrenergic receptors on turkey erythrocytes. PNAS *71:* 1773-1776, 1974.
46 Kranz SB: Erythropoietin. Blood *77:* 419-435, 1991.
47 Kooijman R, Willems M De, Haas CJ Rijkers, GT Schuurmans AL, Van Buul-Offers SC, Heijnen CJ, Zegers BJ: Expression of type I insulin-like growth factor receptors on human peripheral blood mononuclear cells. Endocrinology *131:* 2244-2250, 1992.
48 Bennet A, Wilson DM, Liu F, Nagashima RC, Rosenfeld RC, Hinz RL: Levels of insulin-like growth factors I and II in human cord blood. J Clin Endocrinol Met *57:* 609-612, 1983.

Received June 8, 1999
Accepted July 5, 1999