The Laws of Complexity and Self-organization: A Framework for Understanding Neoplasia

Nat Pernick, M.D.
PathologyOutlines.com, Inc.
Bingham Farms, Michigan (USA)

Posted online: 6 December 2011, revised 7 December 2011
Email: NatPernick@hotmail.com

Abstract

Background: Current biologic research is based on a reductionist approach. Complex systems are broken down into combinations of simpler systems or parts, which can then be studied more readily. Although this approach is rational, it has failed to bring about the understanding necessary to substantially reduce cancer-related deaths. Complexity theory suggests that emergent properties, based on interactions between the parts, are important in understanding fundamental features of living systems. Applying complexity theory to neoplasia may yield a greater understanding of physiologic systems that have gone awry.

Methods and Findings: The laws of complexity and self-organization are reviewed and summarized, and applied to neoplasia:

1. In living systems, the whole is greater than the sum of the parts.
2. There is an inherent inability to predict the future, particularly for living systems.
3. Life emerges when the molecular diversity of a closed system exceeds a threshold of complexity.
4. Much of the order in organisms is due to generic properties that emerge from a network of gene products.
5. Numerous biologic pressures push cells towards disorder.
6. Organisms resist these biologic pressures towards disorder through multiple layers of redundant controls.
7. Neoplasia occurs as these multiple controls are breached. The resulting neoplastic process reflects the nature of the breaches, the individual's germline configuration and the network state of the cell of origin.

Conclusions: In the framework of the laws of complexity and self-organization, cells maintain order by redundant control features that resist the inherent biologic pressure in the cell towards disorder. Neoplasia can be understood as the accumulation of changes that undermine these control features, which may lead to dysregulated growth and differentiation. Studying the neoplastic process within this context may generate new approaches to treatment.

Introduction

On 23 December 1971, President Nixon signed the National Cancer Act of 1971, generally viewed as beginning the "war on cancer" in the United States (note 1). Since then, the U.S. National Cancer Institute has spent over $100 billion on this effort (New York Times, 2009 Jun 27). Death rates due to cancer have decreased for both sexes (J Natl Cancer Inst 2011;103:714), and the 5 year relative survival rate for all cancers has increased from 50% in 1975-1977 to 68% in 1999-2006 (Cancer Facts and Figures 2011, p. 2). However, in 2011, there will still be an estimated 571,950 cancer deaths just in the United States, many in otherwise healthy people (Cancer Facts and Figures 2011, p. 1).

Our current research efforts are based on a reductionist approach to understanding biology. Under this approach, analysis of complex systems can be reduced to studying combinations of simpler systems, which can themselves be reduced to simpler parts (**Reductionism: Interdisciplinary Encyclopedia of Religion and Science**). A cell can be thought of as a complex machine whose actions can be understood by analyzing each of the components. We study disease by finding and analyzing the defective components. For example, 80-90% of follicular lymphoma cases have the t(14;18)(q32;q21) chromosomal translocation, which brings the BCL2 proto-oncogene under the transcriptional influence of the immunoglobulin heavy chain gene. This leads to overexpression of a functionally normal bcl2 protein, which inhibits these cells from their usual condition of undergoing apoptosis. This allows additional genetic mutations to accumulate, which leads to neoplasia of follicular center cells in some patients (**Haematologica 2008;93:1773**).

The reductionist approach is logical and predictable, and has led to improvements in survival for some malignancies. However, many neoplastic processes cannot be described in so logical a manner, and this approach has failed to improve survival for the major causes of cancer related death. This may be because the reductionist approach ignores fundamental properties of living systems that are not necessarily logical and predictable. Understanding these properties may yield a greater understanding of physiologic systems that have gone awry during neoplasia. To this end, this paper summarizes the laws of complexity and self-organization, and applies them to neoplasia (**note 2**).

Table 1 - The Laws of Complexity and Self-Organization Relevant to Neoplasia

1. In living systems, the whole is greater than the sum of the parts.
2. There is an inherent inability to predict the future, particularly for living systems.
3. Life emerges when the molecular diversity of a closed system exceeds a threshold of complexity.
4. Much of the order in organisms is due to generic properties that emerge from a network of gene products.
5. Numerous biologic pressures push cells towards disorder.
6. Organisms resist these biologic pressures towards disorder through multiple layers of redundant controls.
7. Neoplasia occurs as these multiple controls are breached. The resulting neoplastic process reflects the nature of the breaches, the individual's germline configuration and the network state of the cell of origin.

1. In living systems, the whole is greater than the sum of the parts.

The reductionist approach is based on the premise that the whole is ***equal*** to the sum of the parts. Using this approach, a large system can be analyzed by breaking it down and studying the parts and the connections between the parts, which is assumed to provide an understanding of the entire system (**EMBO Rep 2004;5:1016**). For living organisms, we begin by examining large structures. To gain further understanding, we look deeper into organs, cells, biochemistry, chemistry and physics (**Reinventing the Sacred**, p. 10). As Francis Crick explained, "the ultimate aim of the modern movement in biology is to explain all biology in terms of physics and chemistry" (**note 3**).

This approach makes sense for studying complex mechanical structures created by man, such as automobiles, which have thousands of parts and utilize numerous engineering principles. By using the reductionist approach, even a non-engineer can begin to understand how an automobile works. However, an automobile lacks features of even the simplest living organism, such as the ability to respond to the environment, to reproduce or to evolve.

In living systems, as discussed below, it is the interactions between molecules that create life. Individually, the molecules can be considered as "dead." Collectively, these molecules develop emergent properties, the missing features that make the whole greater than its parts (**Complexity 2002;7:18**).

For example, the properties of a protein are not equivalent to the sum of the properties of each amino acid. Proteins can catalyze a chemical reaction or recognize an antigen not only because their amino acids are arranged in a specific order, but because their tertiary structure and function are additionally determined by external factors (**EMBO Rep 2008;9:10**). Yet the tertiary structure and function cannot be deduced from analysis of individual amino acids.

Emergent properties also arise in cell reproduction. Various molecules engage in linked processes that culminate in the reproduction of a cell as an entire entity. As Kauffman notes, "it is a closure of work tasks that propagates its own organization of processes" (**Reinventing the Sacred**, p. 94). Although no laws of physics are broken, it is not reducible to physics. How it works is something science "hardly knows how to talk about" (**Reinventing the Sacred**, p. 47). Thus, this first law of complexity and self-organization is really a recognition that these emergent properties exist and are not explained by the traditional reductionist approach.

2. There is an inherent inability to predict the future, particularly for living systems.

The deterministic nature of Newton's three laws of motion suggested that given initial and boundary conditions, one could exactly determine a system's evolution and predict its future (**note 4**). However, there are several reasons why we cannot predict the future, particularly for living systems.

First, many properties of living systems are emergent, and although they do not violate the laws of physics, they cannot be predicted by them. As indicated above, given the amino acid sequence of a protein, we cannot predict its tertiary structure. One explanation for this inability is that "this is a formidable problem given the very large number of degrees of freedom and hence very large conformational space of a polypeptide chain" (**J Mol Biol 2009;393:249**). However, a more basic explanation of our inability to predict is that the overall folding mechanism of proteins is an emergent property, based on properties between the parts (amino acids) not described by any known law of nature. Similarly, we cannot predict emergent properties at higher levels. Even with a complete map of its genome, we cannot determine basic features of an organism, such as what it will look like, or the impact of a genetic change. All we can do is to try to explain, after the fact, why things are the way they are.

Second, quantum physics dictates that there is inherent uncertainty in all matter at the subatomic level. Even the location of an electron can only be determined as a probability (**Relational Quantum Mechanics, Stanford Encyclopedia of Philosophy**). This is reflected in the Heisenberg Uncertainty Principle, which states that for an electron or any other particle, one cannot simultaneously measure, with an arbitrary degree of accuracy and certainty, both its present position and future momentum (**The Uncertainty Principle, Stanford Encyclopedia of Philosophy**). This is not a criticism of our ability to measure accurately, but a statement about the laws of physics. Thus, living and nonliving systems have features that cannot be predicted with precision because they are composed of subatomic particles whose features can only be determined as probabilities.

Third, many aspects of living systems, whether cells or the weather, are best described by nonlinear equations, which have mathematical features known as chaotic properties (**note 5**). Although we can accurately model the weather by a series of equations, these equations are exquisitely sensitive to initial conditions (**Journal of Atmospheric Sciences 1963;20:130**). Lorenz, in

his pioneering work on chaos theory, found that his computer model of the weather experienced exponential divergence when he reran it substituting the figure 0.506 for 0.506127 (**Technology Review (MIT), March/April 2011**). This is often called "the butterfly effect," because a butterfly flapping its wings in one city can affect factors that eventually change the weather across the world (**Predictability: Does the Flap of a Butterfly's Wings in Brazil set off a Tornado in Texas, 1972 (AAAS)**). This inability to predict the future of well-understood systems is disturbing to many. It is not due to our lack of intelligence, and will not be altered by more powerful computers. It is an inherent property of the non-linear world in which we live.

The fourth limitation is that living systems may be based on incompressible algorithms. This means that there is no shorter means to predict what a living system will do than to simply observe what happens (**At Home in the Universe**, p. 23). We would like a model that provides a shorter, compressed description, such as an equation. However, if each part of the living system contributes to the whole, then a model of only part of the system may be markedly inaccurate.

Finally, we cannot predict the future of living systems because the function of some molecules is dependent on evolutionary pressures, which themselves cannot be predicted (**Cold Spring Harb Symp Quant Biol 1964;29:137**). For example, selection may favor individuals with the human sickle cell mutation, but only in geographic areas where falciparum malaria is endemic, because this mutation protects erythrocytes from infection (**Nature Education-Natural Selection**). However, we cannot predict the impact of this particular mutation on survival in the local environment without knowing the evolutionary pressures of all other human molecules, and how they reinforce or counteract each other. In addition, we must know how the physical environment will change over time, how the genomes of competing organisms will change over time, and how all these factors will interact. Each of these factors cannot be predicted because of the first four limitations above, and because we have no natural laws to predict specific evolutionary changes.

3. Life emerges when the molecular diversity of a closed system exceeds a threshold of complexity.

According to Kauffman, life is the emergent collective property of a modestly complex mix of molecules which catalyze each other's formation (**The Origins of Order**, Chapter 7). Individually each of the molecules is inert. However, with a large enough collection of molecules of sufficient complexity, confined to a small space so they are more likely to interact with each other, a self-sustaining network or web of reactions may be formed that can reproduce and evolve. Unlike some models of the origin of life, the molecules do not reproduce themselves. Rather, the network has the property that the last step in the formation of each molecule is catalyzed by some other molecule in the network.

Under the right circumstances, there is a high probability that a subset of a large set of complex molecules will form a network of molecules that catalyze the formation of each other. A complex protein has numerous clefts or grooves on its surfaces due to the tertiary structure of its polypeptide chain, which may bind the transition states of reactions, a requirement for a catalyst (**Proc Natl Acad Sci USA 1994;91:4103**). The likelihood that a specific molecule will catalyze a specific reaction is very low, but with a large number of molecules and possible reactions, the probability that *some* molecule will catalyze *some* reaction is high (**At Home in the Universe**, Chapter 3). This is analogous to the "birthday problem." In a group of 50 people, the probability that any of them will have the same birthday as the first is only 49/365 (13%), but the probability that any two people in the group will have a common birthday is 97% (**Wikipedia**).

As reactions are catalyzed, additional molecules are created, which further increases the diversity of molecules. The number of possible catalyzed reactions will increase faster than the number of molecules. Eventually, a threshold is crossed in which a network of molecules exists that constitutes a collectively autocatalytic set, in which each molecule's formation is catalyzed by other molecules. This reaction network will also couple exothermic and endothermic reactions

(**The Origins of Order**, Chapter 8). Many alternative pathways are possible to produce this autocatalytic network, but only one of them is required.

This model of the origin of life suggests an answer to the question of why free living cells have an apparent minimal complexity. *Mycoplasma genitalium*, the smallest known genome that constitutes a cell, contains approximately 470 genes, a large number for the simplest organism (**note 6**). Based on the above model, we can hypothesize that a lesser number of genes would lack the complexity to create a self-staining network. Other theories of the origin of life, such as the replicating RNA theory, provide no explanation for this minimal complexity.

4. Much of the order in organisms is due to generic properties that emerge from a network of gene products.

Each cell coordinates the activities of approximately 20,000 genes and their products (**Nature 2004;431:931**). Not only does the cell coordinate complex activities such as mitosis, but these activities occur without any oversight. The molecules interact spontaneously with each other, and lead to the creation of new cells and new living organisms with the same properties. If we are to get a deeper understanding of diseases whose secrets have defied decades of determined research, we may need to understand the general principles underlying these processes.

The traditional view is that the sole source of order in organisms is natural selection as described by Darwin (**On the Origin of Species, 1859**); the patterns of a pine cone or the artistry of mitosis are due to "chance caught on the wing," a description of natural selection by Jacques Monod (**Le Hasard et la nécessité (Chance and Necessity, 1970**). An alternative view, advanced by Kauffman and others, is that order is an expected emergent property of molecular networks, based on structural properties of networks that are not dependent on details on the particular molecules (**The Origins of Order: Self-Organization and Selection in Evolution**, 1993). Genes, RNA and proteins form a complex parallel processing network in which each molecule is connected to some other molecule, and switches each other off and on. During ontogeny, gene products turn each other on and off to generate precursor cells. Once cells become differentiated, some of this switching continues, but the cells are relatively stable.

Consider a hypothetical gene product A, which could be considered active or not, based on its state of phosphorylation. Based on inputs from kinases or other gene products, protein A may change its properties in the next instant of time. Any change to gene product A may affect other gene products it is connected to in the network, and for which it acts as an input. If we consider all the gene products in a cell, we can consider all of their properties at a particular moment in time as the state of the cell. In the next instant, each gene product may change its properties based on relevant inputs, leading to a new state. The path that a network traverses over time, based on changes in the state of each gene product at each moment in time, is called its state space or state cycle.

Theoretically, a cell with 20,000 types of gene products, and numerous copies of each, could have an almost infinite length to its state space, and never return to its original state. For example, a network of N gene products, with only one copy and two possible properties for each gene product, would theoretically have a state cycle of length 2^N. For N = 20,000 gene products, this is approximately 10^{6000}. However, a state cycle this large does not happen, due to three network properties that induce order.

The first network property that induces order is the surprising finding that if each gene product is regulated by at most 2 inputs, which is relatively common, the median length of the state cycle is ***only*** the square root of the number of gene products, or 141 if N is 20,000 (**J Theor Biol 1969;22:437**, **The Origins of Order**, p. 479). This network property creates inherent stability, even in networks with large numbers of gene products, as the cell network is localized to a very small

percentage of its possible state space. In contrast, networks with 5 or more inputs per gene product have an exponential rate of growth and chaotic behavior (**The Origins of Order**, p. 194).

The second network property that maintains order is that most genes are regulated by "canalyzing" Boolean functions (**J Theor Biol 2011 Aug 25 [Epub ahead of print]**), which means that one input can completely determine the property of the gene. For example, the Boolean OR function is canalyzing. In the OR function, if one input is active, then the result is active, regardless of the value of the other input. A series of canalyzing functions, such as a network of genes regulated by the OR function, is considered a "forcing structure," because changing one input to active propagates the "active" status throughout the network. If there is a loop, then the network is frozen in this active state. Thus, genetic networks with canalyzing functions are stable, which further limits the length of the state cycle (**Proc Natl Acad Sci USA 2004;101:17102**).

The third network property that maintains order is the P parameter, which may provide order even when molecules are regulated by a large number of inputs, as long as the molecule has only two states, such as active and inactive. Assume that a Boolean function (such as AND or OR) exists to determine the state of a particular molecule. If there are K inputs for the molecule, there are 2^K possible combinations of input values. The P parameter measures the proportion of these combinations that create an active or inactive state, whichever is larger. Thus, it has a value of 0.5 to 1.0 (19). If the molecule exists 50% of the time in each state, the P parameter is 0.5. If the molecule exists 100% in one of the states, the P parameter is 1.0. As P moves from 0.5 to 1.0, the network is tuned to becoming more orderly. The P parameter helps explain the order induced by gene products with at most 2 inputs. In these networks, a large number of Boolean functions have P values of .75, which promotes stability. In contrast, as the number of inputs increases, functions with such high values of P become rare (**At Home in the Universe**, p. 84).

Kauffman terms the results of these three network properties "order for free." As a result of these properties, cellular networks actually have only a limited number of states. Mutations and other perturbations can change the functional connections between some of the molecules in the network, but this usually does not change the stability of the network very much due to these order inducing properties.

These generic network properties have profound implications. Cellular networks, once they arise, have an inherent order independent of the particular molecules or reactions involved. This order is part of the interactions between the parts described above that makes the reductionist approach inadequate for understanding biologic systems.

5. Numerous biologic pressures push cells towards disorder.

How do these laws of complexity and self-organization relate to neoplasia? Tension exists in living systems between order and disorder, an example of the tradeoffs inherent in living systems to achieve compromise between conflicting interests. A certain amount of order is required within cells, tissues and organs for living systems to function. However, numerous pressures push cells towards disorder.

First, all systems, living and nonliving, tend towards increasing disorder or entropy, based on the second law of thermodynamics (**Wikipedia**). Without external energy, entropy tends to increase, which means cells have to spend their limited resources on countering this pressure towards disorder. Cells can use entropy to drive useful tasks. For example, entropy promotes DNA self-assembly by favoring the packing of many small molecules, which outweighs the cost of forming loops with thousands of base pairs (**Biophys J 2006;90:3712**). However, using entropy in this way still requires significant cellular resources.

Second, the process of creating an autocatalytic network, discussed above, promotes disorder. As a closed network of molecules begins reacting with itself, it produces an increasing number of

new molecules, which catalyze further reactions. This exponential expansion of the network creates continuing disorder, and itself has no natural stopping point until limiting factors are reached.

Third, Kauffman suggests that the earth's environment is supracritical, which puts pressure on living systems to breach their orderly constraints (**At Home in the Universe**, Chapter 6). In an uncontrolled nuclear reaction, a uranium 235 atom absorbs one neutron, leading to nuclear fission, which releases more neutrons, which then act on other uranium atoms. Similarly, it is believed that life on Earth is expanding at an exponential rate. Life may have begun with a small collection of amino acids (**Science 1953;117:528**, **Astrobiology 2003;3:291**), but today, over 10 million different organic molecular structures have been cataloged (**Wikipedia**). The presence of these molecules throughout our biosphere puts constant pressure on living systems to interact with them in novel ways and abandon their ordered ways.

Fourth, natural selection disfavors rigid order in living systems. Organisms do have the ability to maintain genetic information unchanged over long time periods. For example, the 16S and 23S ribosomal RNA genes are the most conserved DNA sequence segments across phylogeny, with minimal changes over billions of years of evolution (**Orig Life Evol Biosph 2008;38:517**). They apparently are under strong evolutionary pressure and positive selection, as changes would be deleterious to the organism.

However, although conservation of DNA is possible, cells must also be flexible to respond to a complex environment, to facilitate development, and to exhibit minor modifications when mutated. In addition, cells must have the ability to undergo change sufficient to form new species. From the Cambrian explosion onwards, the fossil record shows a number of large extinction events which wiped out a significant fraction of Earth's species, possibly caused by loss of habitat and environmental changes (**Extinctions and Biodiversity in the Fossil Record 2002**). Rigid order would doom species amidst these changes.

Finally, the ability of living systems to maintain viability after mutational changes demonstrates an inherit flexibility not present in a completely ordered regime. Many systems are incapable of sustaining small changes without substantial loss of function. For example, some scientists speculate that life began with an "RNA world" dominated by RNA polymerase (**Nature 1986;319:618**). However, a simple network with this type of polymerase is subject to "error catastrophe" (**J Virol 2006;80:20**) because a small change in the polymerase will most likely lead to erroneous replication if no error correction mechanisms are in place. Similarly, computer programs are also exquisitely sensitive to changes. Almost any change to valid computer code will lead to a nonfunctioning program. Thus, the ability of complex systems to survive small changes and actually improve in function should not be taken for granted.

Kauffman believes that organisms maintain a position between order and disorder that he terms the "edge of chaos," an evolution derived compromise between order and surprise that may be optimal to coordinate complex activities and to evolve further (**At Home in the Universe**, p. 86; **Proc Natl Acad Sci USA 2005;102:13439**). This "edge of chaos" is a type of phase transition, similar to that of water and ice at the freezing point.

6. Organisms resist these biologic pressures towards disorder through multiple layers of redundant controls.

It appears that organisms have multiple layers of redundant controls that resist the above pressures towards disorder.

First, the structure of biologic networks inherently resists these pressures towards disorder. Based on interactions between the components, a large "frozen" component forms, whose state does not change over time even as the states of other molecules change (**The Origins of Order**, p.

206; Physica D: Nonlinear Phenomena 1990;42:135). In addition, transient changes to the state of a gene product tend not to propagate to the remainder of the network. This not only maintains network stability against minor changes, but also provides the basis for cellular homeostasis. Although cells can theoretically exhibit an almost infinite number of states, there are actually only approximately 300 cell types, which maintain a consistent phenotype due to the relatively stable expression of a large percentage of their gene products (The Origins of Order, p. 467). These homeostatic states are called attractors, which limit the behavior of the network to a small portion of its state space. Thus, a cardiac myocyte cannot transform to a glial cell, and only rarely do mature cells change their state of differentiation at all.

Second, cells have external membranes with transporter proteins that act as "border controls" (Science 2008;322:709). This limits the entry of novel molecules into the cell that might create new reactions or alter existing ones. In addition, molecules are compartmentalized to limit their possible reactions. For example, hematopoietic stem cells physiologically produce the c-Abl tyrosine kinase, which is sequestered in the cytoplasm by binding to 14-3-3 proteins (Nat Cell Biol 2005;7:278). In response to DNA damage, c-Abl is targeted to the nucleus, but only under tightly controlled conditions (Cancer Control 2009;16:100).

Third, cells have robust processes to limit errors. This prevents altered protein structure and function, which might disturb existing network interactions or create new ones. Transcription is a complex process, with an estimated error rate of 1 per 100,000 nucleotides (Nature Education 1(1)-DNA Replication and Causes of Mutation). DNA repair processes, including proofreading and mismatch repair, correct most of these errors. Cells also limit known causes of DNA damage, such as reactive oxygen species, which would otherwise increase the error rate of transcription and translation (J Biol Chem 1969;244:6049, J Biol Chem 2001;276:38084).

Fourth, cells have several mechanisms to respond to injury or DNA damage, which might eventually alter proteins. Through an underdetermined intracellular audit, some injured cells undergo apoptosis through activation of caspases and bcl2 family members (Mol Neurobiol 2010;42:4). In addition, transcription factors and signaling molecules can recognize potentially lethal stimuli and initiate cycle arrest, autophagy or protein synthesis shutoff (Carcinogenesis 2011;32:955). The importance of the apoptotic pathway is suggested by the high frequency of a defective p53 response in most cancers, either by p53 gene mutations or deletions, or by other alterations in the p53 pathway (Science 1991;253:49).

Fifth, key cellular processes have numerous controls that tightly regulate their activity. For example, various mitotic checkpoints ensure genomic integrity by delaying cell cycle progression in the presence of DNA or spindle damage (Mol Cell Biol 2010;30:22, DNA Repair (Amst) 2009;8:1047). These multiple levels of control prevent inappropriate initiation of mitosis or other key processes, which reduces the propagation of transformed cells.

Finally, humans have an extensive and varied immune surveillance system of B and T cells, NK cells and macrophages. These immune cells have both innate and adaptive properties, and maintain order by destroying cells with disordered properties (Cell 2010;140:883). Their importance is suggested by the impact of immunosuppression, which is usually associated with a markedly elevated risk of malignancy (Nat Rev Nephrol 2010;6:511).

7. Neoplasia occurs as these multiple controls are breached. The resulting neoplastic process reflects the nature of the breaches, the individual's germline configuration and the network state of the cell of origin.

The laws of complexity and self organization give us a framework to better understand neoplasia. Cells are not just the sum of 20,000 gene products interacting with each other in a predictable way (law 1), but networks of interactions that emerge between these molecules, which cannot be predicted (laws 2-3). These networks possess a great deal of stability, independent of the

specific nature of the molecules (law 4). On the other hand, these cells are also under tremendous pressure to breach the control mechanisms that maintain order by resisting the inherent biologic pressures in the cell towards disorder (law 5). In fact, a certain amount of disorder is promoted by natural selection. Thus, it is not surprising that cells will escape some of these controls. Only the presence of multiple redundant controls at various levels (law 6) maintains sufficient order in cells and organisms to allow them to function, consistent with the "multiple-hit" theories that more than one mutation is necessary for neoplasia (**Br J Cancer 1953;7:68**, **Proc Natl Acad Sci USA 1971;68:820**).

To illustrate the key idea of the inherent biologic forces promoting disorder, consider a plan to put a large number of teenagers in a small area, and get them to create some kind of ordered community. A large number of adults would be required to get the teenagers to obey the rules. Yet teenagers inherently don't like being told what to do, and even with multiple layers of enforcement, there will be frequent attempts by the teenagers to work together to break the rules, with occasional success, and with varied responses from the adults. This type of tension exists in our cells on a regular basis. Of course, molecules cannot think, but they have an inherent tendency to want to react with other molecules in novel ways. Another difference is that the control mechanisms are also part of the molecules being controlled, and not a separate class. When order starts to break down, it can create chaos with the teenagers, and neoplasia within the cell.

Nature of breaches

It may be possible to classify neoplasia based not just on morphologic changes, but on the type of controls that have been breached. For example, low grade gastric MALT lymphoma is strongly associated with chronic *Helicobacter pylori* infection, and *H. pylori* eradication provides an excellent long-term outcome (**Gut 2011 Sep 2 [Epub ahead of print]**). It is possible that over long periods of time, these inflammatory changes induce changes to the cellular network that is otherwise "frozen," which leads to neoplasia.

Some genetic changes alter the distribution of molecules within the cell, allowing them to interact with other molecules previously restricted. The t(9;22) translocation leads to a bcr-abl fusion protein in hematopoietic stem cells, which leads to constitutive tyrosine kinase activation and translocation of the fusion protein from the nucleus to the cytoplasm. As a result, the fusion protein now has numerous additional targets for catalysis based on its new location, which may disrupt existing networks (**Cancer Control 2009;16:100**).

Similarly, alterations in networks associated with error control (**Mech Ageing Dev 2008;129:391**), apoptosis (**Cancer Epidemiol Biomarkers Prev 2009;18:1680**), other key pathways (**Science 2008;321:1801**) and immune surveillance (**Int J Biol Sci 2011;7:651**) can be viewed as breaches in the physiologic mechanisms that resist pressures towards disorder. Tumors with similar types of breaches may have common clinical or morphologic features, or respond to similar types of treatment, an area of future investigation.

Germline configuration

The nature of the neoplasia is also affected by the germline configuration of the organism. A large percentage of genes have more than one allele, and many individuals are heterozygous for a large fraction of these alleles (**RC Lewontin, The Genetic Basis of Evolutionary Change, 1974**). Individual variation in these alleles means individuals have inherent differences in the genetic networks in their cells. These differences may involve genes known to be important control mechanisms. To date, at least 54 familiar cancer syndromes have been identified, with mutations in p53 or other known oncogenes (**J Natl Cancer Inst Monogr 2008;(38):1**).

Some allelic variations may affect important control mechanisms only under specific circumstances. For example, the CCR5 gene encodes a chemokine receptor whose role in

normal immune function is unclear, but which is the principal receptor for macrophage-tropic viruses, including HIV1. Approximately 80% of Caucasians have this allele, rendering them susceptible to HIV1 infection, which if untreated, impairs immune surveillance and leads to increased rates of various malignancies. However, 20% of Caucasians are heterozygous for the CCR5 delta 32 mutation, which produces a shorter nonfunctional gene that does not migrate to the cell surface. As a result, these patients are relatively resistant to HIV1 infection (**Cell 1996;86:367**, **PLoS One 2011;6:e22215**). Thus, the CCR5 delta 32 mutation has an impact on neoplasia, but only in the presence of an HIV epidemic.

Network state of cell of origin

The phenotype and state of differentiation of the cell of origin, a reflection of its network state, constitute a somewhat permanent variation to the germline configuration that may also influence the nature of the neoplastic process. Cellular differentiation is due to the capacity of genes to modify the activity of other genes. The network properties that provide "order for free," discussed above, also create variability in the effect of breaches of control mechanisms. In many cell types, a change in a molecule's function will have minimal impact on the existing network in that cell, because the molecules they would typically affect are "frozen" in a particular state by other members of the network. For example, t(14;18) is important only in lymphocytes that produce IgH; in other cells, the altered molecules cannot propagate within the network.

The control mechanisms may also change due to hormones, environmental influences, or unknown maturational factors. Thus, mutations may have a different impact in neonates, children and adults. Although cells may look similar morphologically at different points in development, there may be varying degrees of homeostasis based on different control mechanisms. Perhaps part of the reason Wilm's tumor or neuroblastoma occur primarily in children/young adults is that the cells in older adults have developed additional control mechanisms that make these cells more resistant to the changes required to develop these neoplastic conditions.

Summary

In the framework of the laws of complexity and self-organization, cells maintain order via redundant control features that resist the inherent biologic pressure in the cell towards disorder. Neoplasia can be understood as the accumulation of changes that allow cells to override these control features, leading to altered networks, which may lead to dysregulated growth and differentiation. The nature of the neoplasia may be affected by the nature of the breaches, the germline configuration and the network state of the cell of origin. A better understanding of how these factors interact with each other may generate new approaches to treatment.

Notes

Note: all links were retrieved on 30 November 2011 unless otherwise indicated.
Note: links in **green** are to free full text references.

(1) The National Cancer Act followed President Nixon's pledge in his State of the Union address on 22 January 1971:

I will also ask for an appropriation of an extra $100 million to launch an intensive campaign to find a cure for cancer, and I will ask later for whatever additional funds can effectively be used. The time has come in America when the same kind of concentrated effort that split the atom and took man to the moon should be turned toward conquering this dread disease. Let us make a total national commitment to achieve this goal.

See **National Cancer Institute pages #1**, **#2**, **video of President Nixon signing the bill into law**

(2) This paper is based primarily on these works of Dr. Stuart A. Kauffman. The Origins of Order contains detailed discussions supporting the underlying theories:

Kauffman, SA. **The Origins of Order: Self-Organization and Selection in Evolution**, Oxford University Press, 1993

Kauffman, S. **At Home in the Universe: The Search for the Laws of Self-Organization and Complexity**, Oxford University Press, 1996

Kauffman, SA. **Investigations**, Oxford University Press, 2002

Kauffman, SA. **Reinventing the Sacred: A New View of Science, Reason, and Religion**, Basic Books, 2010

(3) Crick FHC. Of Molecules and Men (1966), reprinted by Prometheus Books (2004), **http://www.amazon.com/exec/obidos/ASIN/1591021855/pathologyoutl-20**

(4) Newton, Isaac; The mathematical principles of natural philosophy (Principia), Volume 1, see page 19 of the **1729 translation** by Andrew Motte (type "19" in the box to the right of "contents")

See also Steinbach, Peter J: Classical and Quantum Mechanics - in a Nutshell, **http://cmm.cit.nih.gov/intro_simulation/node1.html**

(5) For examples of chaos in biologic systems, see **Phys Rev E Stat Nonlin Soft Matter Phys 2010;82:011924** (CNS neurons), **Phys Rev E Stat Nonlin Soft Matter Phys 2009;80:016213** (blood flow), **Med Biol Eng Comput 2008;46:433** (coronary circulation).

(6) *Mycoplasma genitalium* is reported to have from 475-82 protein encoding genes (see **Science 1995;270:397** and **Wikipedia**). Smaller genomes have been documented in bacterial endosymbiots, but they are not capable of independent life, and are considered a transition to becoming an organelle.

-- End --