Cancer - Mutational Resurrection of Prokaryote Endofossils

Wolfgang Sterrer

Affiliated with: Bermuda Natural History Museum, Flatts, Bermuda
Correspondence to: westerrer@gov.bm

Cancer Hypotheses 2016 **1**, 1, 1-15

Received: 24 June 2016, Accepted: 20 July 2016, Published: 21 August 2016

"*The greatest masterpiece in literature is only a dictionary out of order"*
Jean Cocteau

Abstract
I argue that cancer shares its roots with incipient *eukaryogenesis*. The eukaryote genome most likely originated within Archaea as a multiple of prokaryote genomes, acquired by massive horizontal gene transfer. Transferred genomes, many of them parasites, were attenuated by having their cell cycle 'drivers' cleaved and disassembled, then recombined into a new three-dimensional genome architecture and function, serving the emerging host as immune protection, and as the source ('toolkit') of most genetic innovation in eukaryote evolution.

Carcinogenesis in turn is initiated when, from among the many somatic mutations that accumulate with mitoses and cell line age, a first mutation restores the 'driver' of a constituent prokaryote endosymbiont. Cancer henceforth progresses akin to asexual evolution, generating more mutations, including additional drivers, accelerating cell proliferation and decay, and eventually overwhelming its host as a chaotically driven endoparasite.

Keywords: eukaryogenesis; prokaryotes; horizontal gene transfer (HGT); superinfection exclusion; carcinogenesis; mutations; prokaryote endofossil theory (PET)

1. "What if common mutations are not the key to cancer?"

Heng (1), in his book "Debating Cancer" (to which I shall refer throughout this essay as the most recent compendium), deplores the "unexpectedly poor progress in the war on cancer", despite the promises held by the widely accepted 'Somatic Mutation Theory (SMT)'. Instead, "after tens of thousands of samples have been sequenced, what has been found is not common cancer gene mutation patterns, but rather an overwhelming level of genetic heterogeneity that occurs as both inter- and intra-tumor heterogeneity". Heng then poses "the question that no one dared to ask: what if common mutations are not the key to cancer?", citing doubt (2, 3) whether the leading cancer paradigm for the last 50 years is addressing the right questions.

Indeed, several new alternatives to the 'Somatic Mutation Theory' have been proposed, to wit the 'Tissue Organization Field Theory' (TOFT), summarized as "development gone awry" (4). Duesberg et al. (5) ask whether carcinogenesis is a form of speciation, considering that "cancers

have individual clonal karyotypes (including aneuploidy), are immortal and evolve from normal cells treated by carcinogens only after exceedingly long latencies of many months to decades". Davies and Lineweaver (6) home in on "the genes of cellular cooperation that evolved with multicellularity about a billion years ago", suggesting they are "the same genes that malfunction to cause cancer. We hypothesize that cancer is an atavistic condition that occurs when genetic or epigenetic malfunction unlocks an ancient "toolkit" of pre-existing adaptations, re-establishing the dominance of an earlier layer of genes that controlled loose-knit colonies of only partially differentiated cells, similar to tumors".

The proposal I am outlining here harks back to the evolutionary transition from prokaryotes (mostly Archaea, Bacteria and viruses) to eukaryotes (organisms with nucleated cells, mainly protists, fungi, plants and Metazoa), and roots cancer in the prelude to eukaryogenesis. It is a Somatic Cancer Theory with a twist, restating Heng's question, but with the emphasis on 'common'. Perhaps we need to look for mutations that are *uncommon*, inasmuch as they restore rather than destroy latent genome functions, thereby empowering the endoparasite as which cancer asserts itself.

Cancer is usually associated with multicellularity. Surprisingly, genomic phylostratigraphy has revealed not one but "two strong peaks of the emergence of cancer-related protein domains, one at the time of the origin of the first cell and the other around the time of the evolution of the multicellular metazoan organisms" (7). Ciccarelli (8) points out that these two peaks are enriched in two distinct groups of genes, the first with 'caretakers' controlling genome stability, and the second with 'gatekeepers', which affect cell differentiation, growth and death. "The second mechanism is intimately connected to multicellularity and the interactions between cells within a complex organism. This observation puts cancer in the context of macroevolutionary transitions…" (8), and "confirms the ancient origin of gene functions involved in cancer" (7).

Regarding terminology in this essay, I shall refer to cancer as a genome syndrome endemic to eukaryotes, but prominently expressed in multicellular Metazoa. Except where otherwise stated, I refer to human cancer, and use 'driver' as shorthand for any gene/mutation that is essential for initiating carcinogenesis and confers clonal growth advantage (9). I propose the term 'prosex' for horizontal gene transfer (HGT) with recombination in prokaryotes, and 'eusex' for HGT (crossing over) in eukaryote meiosis and syngamy.

2. The 'Prokaryote Endofossil Theory' (PET) of cancer in a nutshell

It now appears that eukaryogenesis was prefaced by massive horizontal gene transfer (HGT) of prokaryote genomes to a candidate archaean (10). Transferred genomes, many of them parasitic, were attenuated (switched to the lysogenic cycle) by having critical (driver) genes disabled (cleaved) while providing immune protection to the host by "superinfection exclusion". Cleaved genomes were then dispersed and recombined across the host genome, thereby preventing reversal to virulence. Superinfection exclusion may thus have initiated the evolution of the complex eukaryote genome topology of split-and-separated genes, including the three-

dimensional interaction of coding genes with non-coding regulators in transcription, as outlined in 'the genome theory' (11).

Carcinogenesis reverses the attenuation phase of this process, initiating a chaotic 'lytic cycle'. Among the numerous mutations that a somatic cell accumulates over its replication cycle, some may fortuitously restore ancient prokaryote genome topologies, fuse cleaved genes, re-empower disabled driver genes, or silence tumor suppressor genes. As oncogenes they subsequently accelerate the host cell cycle, generating more mutations that eventually drive cancer progression. Carcinogenesis in this scenario is the mutational resurrection of prokaryote endoparasite genomes, long ago disassembled and dispersed over what was to become the eukaryote genome. The heterogeneity and progression of cancer reflects the chaotic interplay between multiple, diverse (and possibly competing) ancient parasite drivers commandeering an unfamiliar host soma, and increasingly destabilizing it.

I am painfully aware of the risk of presenting two hypotheses in tandem - both unproven yet crucially reliant on each other; both pivoting on the 'just-so' story of the eukaryote genome having accreted from multiple dismembered prokaryote genomes - especially when linking cancer with eukaryogenesis, another "fundamental, forbidding evolutionary puzzle…"(10), and the origin of eusex, which is joined to eukaryogenesis like a Siamese twin, and "remains one of the great puzzles of evolution theory" (12).

Fortunately there is no need here to delve into eukaryogenesis (13, 14); its well-documented antecedents of "massive gene transfer" (10, 15), and the precocious complexity of the nascent eukaryote cell (16), with its peculiar gene (17) and genome architecture as the most conspicuous end product (18), may suffice for matching it to the retrograde nature of cancer. However, some background on the current view on cancer, and a list of essential differences between prokaryotes and eukaryotes including their sexual practices is needed before attempting to connect the dots, with emphasis on 'superinfection exclusion' as the key process in the path towards eukaryogenesis.

3. Background: Retrograde mutations

3.1. Cancer begins with a driver mutation, ends in uncontrolled growth

Cancer has been documented in most mammals and other vertebrates, indeed most Metazoa, and arguably in plants and single-cell ciliate protists (19). With rare exceptions (e.g., canine transmissible venereal tumor, 20) cancer is not contagious; nor is full-blown cancer inherited, but some 200 hereditary 'cancer susceptibility syndromes' have been described in humans, which may account for 10% of all cancer (20).

3.1.1. *Cancer* is always associated with mutations (alterations to the genome), which inexorably accumulate in somatic cells with an individual's age. In humans, 'the risk of suffering any cancer before the age of 40 is ~2%, but by age 80 this risk increases to 50%' (21). Of hundreds of thousands of mutations that may be found in a single cancer cell, only a few are 'drivers' that speed up clonal growth, the vast majority being 'passengers' hitching a ride.

3.1.2. *Carcinogenic mutations* can be caused by exogenous (e.g., UV light, certain chemicals) or endogenous agencies, particularly errors during genome replication. Mutations can also arise from horizontal gene transfer (HGT), especially in epithelia exposed to symbionts/pathogens (as in the intestinal tract), and from insertions by mobile transposable elements (22), viruses (e.g., human papilloma virus) or bacteria (e.g., *H. pylori*)(23).

3.1.3. *Mutations* may alter any level of genome organization, ranging from minor replication errors (base substitutions, indels - insertions/deletions of bases) to changes in gene copy or chromosome number (aneuploidy) and major genomic rearrangements (translocations). Many cancer genes are activated "through genomic rearrangement, which may join the sequences of two different genes to create a fusion gene, or it may position the cancer gene adjacent to regulatory elements from elsewhere in the genome, frequently across chromosomes" (9). 'The vast majority of somatic mutations in cancer are within non-coding regions, which comprise >98% of the genome' (24). Among epigenetic effects (i.e., in transcription or translation), DNA hypomethylation [i.e., over-activation of genes] and hypermethylation [i.e., excessive silencing of tumor suppressors], "its almost constant companion, just usually (but not always) in different sequences", are a ubiquitous feature in carcinogenesis (25). Other epigenetics related to carcinogenesis are only beginning to surface, such as changes in codon bias (i.e., the non-random use of synonymous codons in amino acid encoding) (26), and the reduction of DNA repair in progression to gastrointestinal cancer (27).

3.1.4. *Cancer progression* is akin to asexual evolution (28). It starts from a single somatic stem cell when a driver mutation speeds up the cell cycle. More mitogenic mutations are likely to result, including additional 'drivers', some recruited from among former 'passengers' by association. The mutation rate may vary, and cancer evolution may switch between punctuated and stepwise. Tumor growth is a highly dynamic process where emergent outlier subpopulations can greatly influence the pattern of progression and the direction of evolution. Genome level changes (such as translocations) have a greater impact on cancer evolution than individual gene mutations in most cancer types, as karyotype alteration often results in altered system inheritance which defines the network structure and even can change the meaning of individual genes (representing 'parts inheritance') by changing the gene context' (28). Different cancer types differ in their overall mutation rates, their predominant mutation types, and the distribution of mutations along their genomes. Substantial variation also exists in the epigenomic landscape of different tissues, specifically in patterns of chromatin accessibility, histone modifications, gene expression and DNA replication timing" (29). It has become increasingly evident that epigenetic abnormalities are not only widespread in many cancer types, but that some of them, either independently or in conjunction with genetic alterations, play a crucial role in establishing some of the hallmark of neoplasms (30).

3.1.5. The most recent – and comprehensive – whole-genome analysis of the somatic mutation landscape in 560 breast cancers (31), reconfirms the Somatic Mutation Theory in principle. The study "detected 3,479,652 somatic base substitutions, 371,993 small indels and 77,695 rearrangements, with substantial variation in the number of each between individual samples", finding "that 93 protein-coding cancer genes carried probable driver mutations". Some recurrent rearrangements "were found beyond the numbers expected, …[but] the significance of these rearranged regions remains unclear."

3.1.6. In summary, cancer emerges from a normal somatic (stem) cell acquiring driver mutations that may occur at any level of genome organization. The first drivers precipitate a self-accelerating cascade of mutations that collectively promote clonal growth along stochastic pathways while noisily destabilizing the host genome. Martincorena and Campbell (21) conclude that "cancer is virtually inevitable in complex, long-lived, multicellular organisms."

3.2. Prokaryotes and eukaryotes - worlds apart?

The average eukaryote cell is about 10,000 times larger than a prokaryote, which in turn has 10,000 times the volume of a virus. Apart from structural and functional changes which, in eukaryogenesis, were brought about by the symbiogenic acquisition of organelles (mitochondria and chloroplasts), increasing size and complexity, novel lifestyles (e.g., predation/photosynthesis), and eventually multicellularity, the most glaring differences between prokaryotes and eukaryotes lie arguably in the architecture, mode of gene activation, and mode of recombination (sex) of their respective genomes.

3.2.1. *Prokaryote genes cluster and line up in a circular chromosome*

Prokaryotes lack a nucleus; their usually single, circular chromosome may contain 100,000 to 5 million base-pairs coding for a few thousand genes, most of which are transcribed. Genes typically form linearly arranged operons in which a promoter is immediately followed by several genes that are transcribed together. Bacterial operons often consist of genes that are functionally related, being part of the same metabolic pathway.

Prokaryotes defend themselves against intruding parasites such as bacteriophages by restriction enzymes (the 'work horses of molecular biology', 32), of which many hundreds have been identified (https://www.neb.com), each protecting the host from invading genomes by cleaving them at a target site "with immaculate specificity in sequence recognition" (33). Prokaryotes are also protected by CRISPR-Cas systems, endonuclease machineries akin to adaptive immunity because they 'remember' previous viral infections.

3.2.2. *Eukaryote genes are cleaved, spread over many chromosomes, and three-dimensionally regulated*

The eukaryote genome is sequestered within a porous nuclear membrane. Fairly typical for eukaryotes, the human genome contains 3 billion DNA base-pairs, with some 20,000 genes distributed over 23 pairs of linear chromosomes. Only a small percentage of the total genome is transcribed, whereas the remainder (which used to be called 'junk DNA', 34) is now known to hold elements that regulate genes during transcription. Most remarkably, genome elements in eukaryotes, while not randomly distributed (35), are rather 'scattered' along and among chromosomes; and are split, with coding regions (exons) interrupted by non-coding introns that must be excised before transcription begins. In addition, operons (or clusters) are rare, and the coding gene is often widely separated from its enhancers, which may even be located on a different chromosome (36). Transcriptional initiation, termination and regulation are consequently mediated by 'DNA looping', which brings together promoters, enhancers, transcription factors and RNA processing factors inside special nuclear regions ('transcription factories') to accurately regulate gene expression. Correct and timely activation of a gene not only depends on its sequence integrity, but as critically on genome topology and the three-dimensional configuration of the chromosomes within the nucleus.

Proposed by Heng (1, 18), the genome-centric concept, or genome theory, views a genome as a hierarchic system in which karyotypes define system inheritance, and genes define parts inheritance. “This explanation is often easiest to comprehend with the aid of analogy. Imagine that each individual gene is a building material – red brick, lumber, tile, etc. These are necessary to build any kind of building yet, depending on how they are arranged, the final results will be drastically different - a house, a skyscraper, a laboratory. In this analogy, the genome is the blueprint that determines how the various materials (genes and their encoded pro ducts) will come together to ultimately form the complete structure (the cell). The three-dimensional architecture of the genome therefore governs the structure of the genetic and protein networks” (18). In this system view, random mutations – rather than being near-universally destructive – may have a better chance of tinkering with existing modularity - throwing switches, substituting materials, changing quantities or time sequences – and constructing novelty. The building materials may be of prokaryote heritage while their innovative modular use, especially at the time of multicellular diversification, was the hallmark of eukaryotes.

The resourcefulness of the eukaryote genome emerges when we read that “de novo origin of protein-coding genes from non-coding DNA is a consistent feature of eukaryotic genomes, having been discovered in organisms as diverse as yeast, plants, flies, mammals, primates and even in recent human evolution” (37) – including some that are implicated in human neuroblastomas (38).

3.2.3. *Prosex is unilateral HGT on demand or by ‘parasitic rape’*

The core purpose of sex is recombination of non-identical genomes; a new genotype. Sex in prokaryotes is the acquisition and recombination of foreign DNA by unilateral HGT modes, such as transformation (uptake of naked DNA from the environment), conjugation (transfer of a plasmid from donor to recipient), and transduction (infection by a bacteriophage, frequently including some adjoining bacterial DNA from a previous host).

Operating in the “no man’s land between heredity and infection, between physiology and pathology” (39), HGT is now considered the main mechanism for the astonishing diversity and adaptability of bacterial and archaean communities. While HGT events may not be much more common at a per capita per gene rate than mutation, it is estimated that 81% of prokaryotic genes have been involved in HGT at some point in their history (40). Prokaryote HGT can be highly promiscuous, extending across species, bacterial divisions and even across the three domains of life. Transferred DNA is mostly short, one to several genes (an operon), but can also involve gene clusters or significant parts of genomes. Acquired DNA is routinely cleaved into smaller sections, usually at highly specific sites, before recombination occurs.

3.2.4. *Eusex is bilateral HGT between consenting conspecifics*

“Prokaryotic recombination leads to [horizontal] pangenomes, and eukaryotic recombination leads to vertical inheritance” (41). Eusexual recombination is typically transacted twice per cell cycle, in meiosis (between grandparental genomes) and in fertilization (between parental genomes), when partner DNA strands pair up and cross over at least once per chromosome before recombining, each generating novel genotypes (42). Unlike prosex, eusex is always followed by genome and cell replication. In meiosis and fertilization, it may be described as ‘safe sex between consenting, conspecific partners’ in which HGT has been internalized, preserving species identity (43), while generating offspring with unique new genotypes. The eukaryote individual thus becomes both a uniquely new message (genotype) and its incarnation

(phenotype) being tested by selection (44). This double role makes recombination (45) the compromise of choice between the freedom to change and the necessity to safeguard genome integrity and copying fidelity - both paramount in Darwinian evolution.

3.2.5. *Genome 'go-betweens' and HGT*

There are also tiny mobile elements such as *viruses and plasmids*, interpreted either as evolutionary precursors (escapees) or contemporaries of prokaryotes (46). Not being able to reproduce by themselves, they invade hosts and act as go-betweens in HGT, shuttling between often unrelated hosts, frequently taking adjoining host DNA with them. This may result in substantial amounts of bacterial DNA being of prophage origin, suggesting an ancient, ongoing two-way coevolution of viral (parasite) and bacterial (host) genomes (47).

Transposable elements (TEs, 31) act like internal versions of phages; some are indeed considered endogenous retroviruses, i.e., remnants of viruses that have lost the ability to re-infect cells (48). TEs can mutate the genomes of their hosts either by 'jumping' to new locations or by facilitating chromosomal rearrangements through homologous recombination. Almost half of the human genome is composed of TEs but only a small fraction is active; the rest is considered 'extinct' or 'epigenetically silenced' (49). TEs have been implicated in tumorigenesis (50). Ewing (51) "found somatic L1 insertions not only in all cancer types and metastases but also in colonic adenomas, well-known cancer precursors. Our results demonstrate that extensive somatic insertional mutagenesis occurs very early during the development of GI tumors, probably before dysplastic growth".

3.2.6. *'Superinfection exclusion' - a 'Jekyll and Hyde' syndrome*

The co-evolutionary arms race between bacteria and their viruses - bacteriophages - may well be the oldest and longest battle in life's history, continuously pitting bacterial defenses against the emergence of new phage infectivity. Yet out of antagonistic interactions between prokaryote parasites competing for the same host a process emerged as ubiquitous as it is therapeutically used in veterinary and human medicine (52). Known as 'superinfection exclusion' or competitive exclusion, it may be related to phenomena described as lysogenic conversion (53), attenuation (54), bacterial interference (55), barrier effect (56), and surface exclusion in plasmids (57). The net effect is that a resident, 'attenuated' parasite protects the host from further infection as its genome recombines with the host's, and is passed on to the next generation (58-60).

Lysogeny in viruses provides the most inclusive model for this process. 'Temperate' phages infecting bacteria can alternate between two lifestyles:

i) a lysogenic cycle (in which the phage genome (as prophage) recombines, and is passed on to daughter cells together with the host genome (which renders the host effectively immune to subsequent infection by additional phages), and

ii) a lytic cycle, switched on by an environmental insult such as UV light, in which phages proliferate, killing the host and swarming out in search of new hosts. Indeed, many bacteria harbor a multitude of prophages, with each phage-encoded virulence or fitness factor making an incremental contribution to the fitness of the lysogen. "It is important to note that lysogenic conversion is only one of at least five different ways by which temperate phages affect bacterial fitness: (i) as anchor points for genome rearrangements, (ii) via gene disruption, (iii) by protection from lytic infection, (iv) by lysis of competing strains through prophage induction, and (v) via the introduction of new fitness factors (lysogenic conversion, transduction) (79)."

The ability to choose between lifestyles may be advantageous to the parasite, but not to the host. In the long run probably only those hosts survived who could lock the switch in the lysogenic position, effectively attenuating the parasite while retaining its immune protection. Endosymbiotic gene transfer (41) served a similar goal: regardless of how the 'wild' precursors of mitochondria (61) and plastids (62) were acquired, lasting symbiosis was only achieved by transferring critical symbiont genes to the host, resulting in the loss of symbiont autonomy (63). The genetics of switching have only recently been explored, suggesting, for example, a role for CRISPR regions in modifying the effects of lysogeny on *P. aeruginosa* (64). Goldberg (65) shows "that the *Staphylococcus epidermidis* CRISPR-Cas system can prevent lytic infection but tolerate lysogenization by temperate phages."

I propose that 'superinfection exclusion' by disarming driver genes of incoming genomes followed by recombination was the prevailing HGT mechanism in the preface to eukaryogenesis, resulting in massive gene gain with immune protection to the emergent host. Attenuation may have been such a vital selective advantage that it set the incipient eukaryote on a monophyletic course towards its hallmark cleaved-and-scattered, three-dimensionally functioning genome architecture.

4. Genomes within the genome

4.1. *From prokaryotes to eukaryotes…*

Molecular phylogenetic analyses of gene clusters from marine planktonic archaea showed that nearly 30% of genes had been acquired from bacteria in ancient and recent HGT events (66). In yeast and other eukaryotic genomes bacterial genes outnumber genes of archaeal origin by a factor of about 3:1 (41). Koonin (10) summarizes the current view that eukaryotes originated within archaea (specifically the TACK superphylum), citing "homologues of signature eukaryotic proteins [that] have been detected in diverse archaea", and evidence that "the origin of the major archaeal phyla involved massive acquisition of bacterial genes". Hard in Koonin's tracks, Spang (15) describe the new phylum Lokiarchaeota, "which forms a monophyletic group with eukaryotes... and whose genomes encode an expanded repertoire of eukaryotic signature proteins… demonstrating that many components that underpin eukaryote-specific features were already present in [the archaean] ancestor. This provided the host with a rich genomic 'starter-kit' to support the increase in the cellular and genomic complexity that is characteristic of eukaryotes". Probing the characteristics of the LECA (Last Eukaryotic Common Ancestor), Wickstead and Gull (16) consider that "one of the most surprising results…has been how much of the biological complexity in extant cells can be traced back to this ancestral cell…. This "complexity early" model of eukaryotic evolution is mirrored in the cytoskeleton, where, somewhere in the evolutionary space between prokaryotes and the LECA, single proto-tubulin and proto-actin molecules diversified into multiple specialized forms".

An increasing number of complex homologies, such as genes and metabolic pathways, are now being identified - the 'building materials' in Heng's analogy (18). Taylor (67) mused that "the eukaryote cell is a multiple of the prokaryotic cell", and Treangen and Rocha (68) noted that prokaryotes "invented most of life's biochemical diversity", collectively sharing their major metabolic pathways with eukaryotes (69). Wilkins (70), reviewing the evolution of meiosis from mitosis, remarked that "it is striking how much of the molecular machinery that [homolog synapsis in meiosis] brings into play is conserved between prokaryotes and eukaryotes and

between mitosis and meiosis". Lane (71) concluded that "virtually every 'eukaryotic' trait can also be found in prokaryotes".

Overall, this suggests that the nascent eukaryote somehow inherited the bulk of the prokaryotes' toolkit for tackling life's challenges and reorganized it before going on to construct and operate complex organisms – with cancer in pursuit. In an analogy from Latin etymology it is as if prokaryotes had evolved the roots of verbs (such as *–duce* = to lead) and prefixes and what not, but the eukaryote re-combined them for new meanings, such as *ad-, de-, pro-* or *re-duce*. It would not come as a surprise, then, if a typographical error (mutation) spelling the sequence "re-pro-duce-me-now!" initiated a neoplastic cell cycle (just as the command "re-duce me now!" might in-duce apoptosis…).

4.2. *…and back to cancer*

The 'Prokaryote Endofossil Theory' matches and connects heterogeneity - the key feature of cancer (1) - with heterogeneity and multiple inheritance in the processes leading to eukaryogenesis. The numbers of candidate ancestral prokaryote genomes (including phages), and the astonishing diversity and specificity (33) of restriction enzymes and CRISP-Cas systems - all in the hundreds or thousands – meet in the cleaved-and-dispersed nature of the eukaryote genome. Quite possibly the original prokaryote 'owner' (phage?) of a resurrected driver gene may be long extinct, justifying the use of 'endofossil'.

If disrupting/dispersing former parasite driver genes was indeed the universal mechanism for achieving lasting attenuation in nascent eukaryogenesis, then we should expect mutations that change genome topology and/or epigenetic regulation, such as rearrangements, to have a higher probability for restoring oncogenes, as seems to be the case.

Conversely, it has been proposed that recombination sex arose to minimize the formation of potential 'meiotic drive' allegiances (72), which have been found in all genetically well-studied species (73). Since recombination is also common in mitosis, where it functions as a DNA repair mechanism (74), the possibility of 'oncogene allegiance' cannot be excluded. Interestingly, a 'mitotic drive' has been shown in myotonic dystrophy (75), as "a rapidly proliferating mutant…which shifted toward further expansion by 'step-wise' mutations, becoming the major allele population [in culture]" – all of which fairly describes cancer progression. Melamed (76) only just uncovered a fascinating "genetic similarity between cancers and co-morbid Mendelian diseases" sharing candidate driver genes. This opens up the possibility that cancer is not the only syndrome governed by the mechanisms of the PET.

In his personal recollection of some 30 years of cancer research, Robert Weinberg (77) muses how "by the mid-1970s, with rare exception, tumor virologists had come up empty-handed in their search for human retroviruses… The work went on nevertheless, because of a simple inescapable fact: that small numbers of viral genes could somehow overrule the vastly larger genomes of cells, forcing the latter into a neoplastic state. This, on its own, represented a truly revolutionary concept!"

The perspective of cancer, finally, adds poignancy to Lane's (78) vision that "the secret of complex life lies in the chimeric nature of the eukaryotic cell—a hopeful monster, born in an improbable merger 2 billion years ago, an event still frozen in our innermost constitution and dominating our lives today".

Acknowledgments
I am indebted to my home institution and the Bermuda Zoological Society for continuing support, and to Henry Heng for his encouragement when it was most needed.

References

1. Heng HH. Debating cancer: the paradox in cancer research: World Scientific Publishing Co. Pte. Ltd. ; 2016. 440 p.
2. Soto A, C. S. The somatic mutation theory of cancer: growing problems with the paradigm? BioEssays. 2004;26:1097-107. doi: DOI: 10.1002/bies.20087.
3. Frank SA, Nowak MA. Problems of somatic mutation and cancer. BioEssays 2004;26:291–9. doi: DOI 10.1002/bies.20000.
4. Baker SG. A Cancer Theory Kerfuffle Can Lead to New Lines of Research. JNCI J Natl Cancer Inst 2014;107(2). doi: doi: 10.1093/jnci/dju405.
5.Duesberg P, Daniele Mandrioli, McCormack A, Nicholson JM. Is carcinogenesis a form of speciation? Cell Cycle 2011;10(13):2100-14 doi: DOI: 10.4161/cc.10.13.16352.
6. Davies PCW, Lineweaver CH. Cancer tumors as Metazoa 1.0: tapping genes of ancient ancestors. Phys Biol. 2011;2011 Feb(8):1-7. doi: 10.1088/1478-3975/8/1/015001.
7. Domazet-Loso TT, D. Phylostratigraphictracking of cancer genes suggests a link to the emergence of multicellularity in metazoa. BMC Biology 2010. 2010;8. doi: 10.1186/1741-7007-8-66.
8. Ciccarelli FD. The (r)evolution of cancer genetics. BMC Biology 2010. 2010;8. doi: 10.1186/1741-7007-8-74.
9. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458(9 April 2009):719-24. doi: 10.1038/nature07943.
10. Koonin EV. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? Philosophical Transactions of the Royal Society B, Biological Sciences. 2015;370(1678). doi: 10.1098/rstb.2014.0333.
11. Heng HHQ, Guo Liu, Joshua B. Stevens, Bremer SW, Karen J. Ye, Batoul Y. Abdallah, et al. Decoding the genome beyond sequencing: The new phase of genomic research. Genomics. 2011; 98 242–52.
12. Smith JM. The evolution of sex: Cambridge University Press; 1976.
13. Williams TA, Embley TM. Changing ideas about eukaryotic origins. Philos Trans R Soc Lond B Biol Sci 2015 Sep 26; . 2015;370(1678). doi: 10.1098/rstb.2014.0318.
14. Booth A, Doolittle WF. Eukaryogenesis, how special really? PNAS. 2015;112(33):10278-85. doi: 10.1073/pnas.1421376112.
15. Spang A, Saw JH, Jørgensen SL, Katarzyna Zaremba-Niedzwiedzka, Joran Martijn, Anders E. Lind, et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature. 2016;521(14 May 2016):173-9. doi: 10.1038/nature14447.
16. Wickstead B, Gull K. The evolution of the cytoskeleton. Journal of Cell Biology. 2011;194(4):513-25. doi: 10.1083/jcb.201102065.
17. Lynch M. The origins of eukaryotic gene structure. Molecular Biology and Evolution. 2006;23(2):450-68. doi: 10.1093/molbev/msj050.
18. Heng HH. The genome-centric concept: resynthesis of evolutionary theory. BioEssays. 2009;31:512–25. doi: 10.1002/bies.200800182.
19. Lauckner G. Diseases of Protozoa through Mollusca. In: Kinne O, editor. Diseases of marine animals. 1. Chichester: John Wiley and Sons, Ltd; 1980.

20. Nagy R, Sweet K, Eng C. Highly penetrant hereditary cancer syndromes. Oncogene. 2004;23:6445-70. doi: 10.1038/sj.onc.1207714.
21. Martincorena I, Campbell PJ. Somatic mutation in cancer and normal cells. Science. 2015;349(6255):1483-9.
22. Chénais B. Transposable elements and human cancer: A causal relationship? Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 2013;1835(1): 28–35.
23. Parsonnet J. Bacterial infection as a cause of cancer. Environmental Health Perspectives. 1995;103(Supplement 8):263-8.
24. Piraino SW, Furney SJ. Beyond the exome: the role of non-coding somatic mutations in cancer. Annals of Oncology. 2016;27(2):240-8. doi: 10.1093/annonc/mdv561.
25. Ehrlich M. DNA hypomethylation in cancer cells. Epigenomics. 2009;1(2):239–59. doi: 10.2217/epi.09.33.
26. Félez-Sánchez M, Bedhomme S, González-Bravo MI, Kamp C, Bravo IG. Cancer, Warts, or Asymptomatic Infections: Clinical Presentation Matches Codon Usage Preferences in Human Papillomaviruses. Genome Biol Evol 2015;7(8):2117-35. doi: 10.1093/gbe/evv129.
27. Bernstein C, Bernstein H. Epigenetic reduction of DNA repair in progression to gastrointestinal cancer. World J Gastrointest Oncol 2015;7(5):30-46. doi: http://dx.doi.org/10.4251/ wjgo.v7.i5.30.
28. Horne SD, Wexler M, Joshua B. Stevens, Heng HHQ. Insights on processes of evolutionary tumor growth. Atlas of Genetics and Cytogenetics in Oncology and Haematology. 2015:12.
29. Polak P, Rosa Karlićć, Amnon Koren, Robert Thurman, Richard Sandstrom, Lawrenc MS, et al. Cell-of-origin chromatin organization shapes the mutational landscape of cancer. Nature. 2015;518:360-4.
30. Podlaha O, Markus Riester, De S, Michor F. Evolution of the cancer genome. Trends Genet 2012;28(4):155-63. doi: 10.1016/j.tig.2012.01.003.
31. Nik-Zainal S, Helen Davies, Johan Staaf, Manasa Ramakrishna, Dominik Glodzik, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature. 2016:8 pp. doi: 10.1038/nature17676.
32. Roberts R. How restriction enzymes became the workhorses of molecular biology. Proceedings of the National Academy of Sciences. 2005;102(17):5905–8. doi: 10.1073/pnas.0500923102.
33. Vasu K, Nagaraja V. Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. Microbiology and Molecular Biology Reviews. 2013;77(1):53-72. doi: 10.1128/MMBR.00044-12.
34. Doolittle W. Is junk DNA bunk? A critique of ENCODE. PNAS. 2013;110(14):5294–300.
35. Hurst LD, Pál C, Lercher MJ. The evolutionary dynamics of gene order. Nature Reviews Genetics. 2004;5:299-310.
36. Spilianakis C, Lalioti MD, Town T, Lee GR, RA F. Interchromosomal associations between alternatively expressed loci. Nature. 2015;435:637-45.
37. McLysaght A, Guerzoni D. New genes from non-coding sequence: the role of de novo protein-coding genes in eukaryotic evolutionary innovation. Philos Trans R Soc Lond B Biol Sci 2015;370(1678). doi: 10.1098/rstb.2014.0332.
38. Suenaga Y, S. M. Raf iqul Islam, Alagu J, Yoshiki Kaneko, Mamoru Kato, Yukichi Tanaka, et al. NCYM, a Cis-Antisense Gene of MYCN, Encodes a De Novo Evolved Protein That Inhibits GSK3b Resulting in the Stabilization of MYCN in Human Neuroblastomas. PLOS Genetics. 2014;10(1). doi: 10.1371/journal.pgen.1003996.
39. Jacob F, Wollman E. Viruses and Genes. Readings from Scientific American. New York, NY: W.H. Freeman & Co; 1961.

40. Dagan T, Artzy-Randrup Y, W. M. Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. Proc Natl Acad Sci USA. 2008;105:10039–44.
41.Ku C, Shijulal Nelson-Sathia, Mayo Roettger, Sriram Garg, Einat Hazkani-Covo, Martina WF. Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. PNAS August 18. 2015;112(33):10139–46. doi: 10.1073/pnas.1421385112.
42. Sterrer W. On the origin of sex as vaccination. J theor Biol. 2002;216:387-96. doi: 10.1006/yjtbi.3008.
43. Heng HH. Elimination of altered karyotypes by sexual reproduction preserves species identity. Genome. 2007;50:517-24.
44. Sterrer W. Prometheus and Proteus: the creative, unpredictable individual in evolution. Evolution and Cognition. 1992;1:101-29.
45. Heyer W, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. Annu Rev Genet 2010;44:113–39. doi: 10.1146/annurev-genet-051710-150955.
46. Nasir A, Kim KM, Caetano-Anollés G. Viral evolution - Primordial cellular origins and late adaptation to parasitism. Mobile Genetic Elements. 2012;2(5):247–52.
47. Canchaya C, Fournous G, Brüssow H. The impact of prophages on bacterial chromosomes. Molecular Microbiology. 2004;53(1):9-18. doi: 10.1111/j.1365-2958.2004.04113.x.
48. O'Donnell K, Burns K. Mobilizing diversity: transposable element insertions in genetic variation and disease. Mobile DNA. 2010;1(21):1-10.
49. Fedoroff N. Transposable Elements, Epigenetics, and Genome Evolution. Science 2012;338(9 November):758-67. doi: 10.1126/science.338.6108.758.
50. Lee E, Iskow R, Network TCGAR. Landscape of somatic retrotransposition in human cancers. Science. 2012;337(6097):967-71. doi: 10.1126/science.1222077.
51. Ewing A, Gacita A, Wood LD, Florence Ma, Xing D, Min-Sik Kim, et al. Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. Genome Research. 2015:10.
52. Pirnay J-P, Bob G. Blasdel, Bretaudea L, Angus Buckling, Nina Chanishvili, Jason R. Clark, et al. Quality and Safety Requirements for Sustainable Phage Therapy Products. Pharm Res. 2015;32:2173–9. doi: 10.1007/s11095-014-1617-7.
53. Brüssow H, Carlos Canchaya, Hardt W-D. Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. Microbiol Mol Biol Rev 2004;68(3):560–602. doi: 10.1128/MMBR.68.3.560-602.2004.
54. Minor PD. Live attenuated vaccines: Historical successes and current challenges. Virology. 2015;479–480(May 2015):379–92. doi: 10.1016/j.virol.2015.03.032.
55. Falagas ME, Petros I. Rafailidis, Makris GC. Bacterial interference for the prevention and treatment of infections. International Journal of Antimicrobial Agents. 2008;31(6):518–22. doi: 10.1016/j.ijantimicag.2008.01.024.
56. Guarner F, Malagelada J-R. Gut flora in health and disease. The Lancet. 2003;361(February 8, 2003):512-19.
57. Ham LM, Cram D, Skurray R. Transcriptional analysis of the F plasmid surface exclusion region: Mapping of traS, traT, and traD transcripts. Plasmid. 1989;21(1):1–8. doi: 10.1016/0147-619X(89)90081-4.
58. Ewald P. Transmission modes and the evolution of the parasitism-mutualism continuum. Ann N Y Acad Sci. 1987;503:295--306.
59. Bull J, Rice W. Distinguishing mechanisms for the evolution of cooperation. J theor BioI. 1991;149:63-74.
60. Frank SA. Models of parasite virulence Q Rev Biol. 1996;71:37-78.
61. Gray M, Burger G, BF L. Mitochondrial evolution. Science 1999;283:1476-81.

62. Martin W, Bettina Stoebe, Vadim Goremykin, Sabine Hansmann, Hasegawa M, Kowallik KV. Gene transfer to the nucleus and the evolution of chloroplasts. Nature. 1998;393(14 May):162-5. doi: 10.1038/30234.
63. Ochman H, Moran N. Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. Science. 2001;292:1096-8.
64. Zegans ME, Jeffrey C Wagner, O'Toole GA. Interaction between Bacteriophage DMS3 and Host CRISPR Region Inhibits Group Behaviors of Pseudomonas aeruginosa. Journal of bacteriology. 2008;191(1):210-9. doi: 10.1128/JB.00797-08.
65. Goldberg GW, Wenyan Jiang, David Bikard, Marraffini LA. Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. Nature. 2014;514(7524):633–7. doi: 10.1038/nature13637.
66. Deschamps P, Zivanovic Y, Moreira D, Rodriguez-Valera F, López-Garcia P. Pasngenome evidence for extensive interdomain horizontal transfer affecting lineage core and shell genes in uncultured planktonic Thaumarchaeota and Euryarchaeota. Genome Biol Evol. 2014;6(7):1549-63. doi: 10.1093/gbe/evu127.
67. Taylor F. Implications and extensions of the serial endosymbiosis theory of the origin of eukaryotes. Taxon. 1974;23:229-58.
68. Treangen TJ, Rocha EPC. Horizontal Transfer, Not Duplication, Drives the Expansion of Protein Families in Prokaryotes. PLoS Genetics. 2011;7(1). doi: 10.1371/journal.pgen.1001284.t001.
69. White D. The Physiology and Biochemistry of Prokaryotes. New York: Oxford University Press, Inc.; 2000. 565 p.
70. Wilkins AS, Holliday R. The evolution of meiosis from mitosis. Genetics. 2009;181(1):312.
71. Lane N, Martin W. The energetics of genome complexityNature. Nature. 2010;467(7318). doi: 10.1038/nature09486.
72. Hurst L, Atlan A, Bengtsson B. Genetic conflicts. Q Rev Biol. 1996;71:317-64.
73. Bell G. Selection. The Mechanism of Evolution. New York: Chapman & Hall; 1997.
74. Symington LS, Rothstein R, Lisby M. Mechanisms and Regulation of Mitotic Recombination in *Saccharomyces cerevisiae*. Genetics. 2014;198(3):795-835.
75. Khajavi M, Tari AM, Patel NB, Tsuji K, Siwak DR, Meistrich ML, et al. "Mitotic drive" of expanded CTG repeats in myotonic dystrophy type 1 (DM1). Human Molecular Genetics. 2001;10(8):855-63.
76. Melamed RD, Kevin J. Emmett, Chioma Madubata, Andrey Rzhetsky, Rabadan R. Genetic similarity between cancers and comorbid Mendelian diseases identifies candidate driver genes. HHS Public Access. 2015. doi: 10.1038/ncomms8033.
77. Weinberg RA. Coming Full Circle—From Endless Complexity to Simplicity and Back Again. Cell. 2014;157:267-71. doi: 10.1016/j.cell.2014.03.004.
78. Lane N. Power, Sex, Suicide: Mitochondria and the Meaning of Life. Oxford: Oxford University Press; 2005.
79. Brüssow H, Canchaya C, Hardt WD. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol Mol Biol Rev. 2004 Sep; 68(3): 560–602. doi: 10.1128/MMBR.68.3.560-602.2004