Progress in Biophysics and Molecular Biology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

# The Importance of Henry H. Heng's Genome Architecture Theory

*Keywords:*
Evolutionary mechanism of cancer
Genome mediated macroevolution
Genome
Architecture theory
Genomic information
Karyotype coding
Phase transition

A B S T R A C T

A recent symposium on cancer and evolution has bought many innovative thinkers together to challenge the status quo of current cancer research. Professor Henry Heng's presentation considers cancer as a new system emerging via macro-evolution, where genome chaos-mediated information creation and maintenance plays an important role. This concept departs from the neo-Darwinian influenced somatic mutation theory of cancer. To appreciate his theory, it is helpful to briefly review several of his heterodox findings in the fields of oncology and evolutionary biology. This letter summarizes and highlights these findings and calls for a medical and scientific reckoning as well as integration within and between these fields.



## 1. Introduction

In the Review Article (Heng and Heng, 2021a), the authors share several findings of critical importance for cancer evolution which will surprise most oncologists and most evolutionary biologists:

1. The primacy of the whole genome architecture over individual genes and collections of genes in explaining the phenomenon of cancer and predicting its initiation and progression.
2. Cancer evolution as “punctuated equilibrium”, not a step-wise, gradual accumulation of mutations (as Darwin might have hypothesized).
3. The discovery of a new evolutionary mechanism (genome chaos) which is responsible for the punctuated phase of cancer evolution.

While these findings should by themselves be revolutionary for the field of oncology, my aim with this letter is to shine a light on the breadth, depth and profound importance of Dr. Henry Heng's research over the last 40 years, as summarized in *Genome Chaos* (Heng 2019). A thorough reading of this book will illuminate the Review Article as the tip of an iceberg, which is that evolutionary biology in general, and oncology in particular, is founded on a scientific framework which is incomplete and increasingly disconnected from empirical reality.

The tip of the iceberg regards cancer's conundrum: we are no closer to curing cancer than ever (Leaf 2004; Heng, 2016; Raza 2019, Roser and Ritchie 2019) despite the hundreds of billions of dollars spent each year on research (NCI 2021) and treatment (Statistica 2021). Hidden below the waterline is the old framework, known as neo-Darwinism and the Modern Synthesis. For a good explanation of why and how these frameworks came to be dominant, see Denis Noble's Illusions of the Modern Synthesis (Noble 2021) as well as James Shapiro's analysis (Shapiro 2021).

## 2. Genome Architecture Theory

Henry Heng proposes a new framework — Genome Architecture Theory, sometimes referred to as genome theory or genome chaos — which addresses what's missing in the old framework, reconciles anomalous data, and unifies evolutionary biology with the phenomenon of cancer.

In a nutshell, the old framework focuses on the role of individual genes, and misses the critical point that life depends on the *interactions* and *interdependencies* of genes - and these are encoded in the genome architecture. The implications of the shift from gene to genome are sweeping and profound for biology, evolution and medicine.

Here I will summarize the highlights of the Genome Architecture Theory, and its logical consequences, as it relates to the old framework and brings it up to date with the newly emerging empirical data.

1) Genome architecture defines the species.

There is no agreed upon definition of species which unifies all of biology. This is known as the “species problem”. Without being able to agree on something as fundamental as the definition, the origins of new species and the mechanisms of evolution are impossible to understand. To get a sense of the complexity of the species problem and the diversity of the proposed definitions, see the Wikipedia page on Species (Wikipedia 2021).

Heng's definition cuts through the complexity and puts a stake in the ground which anchors the rest of his framework: *a species is a population which shares the same genome architecture.* Heng asks us to think of genes as bricks, and the genome as the architectural form the bricks are organized into, e.g. a house, a garden path, an archway, and so on. Each new rearrangement of genes represents a different species. In an arbitrary population of multicelled

https://doi.org/10.1016/j.pbiomolbio.2021.08.009
0079-6107/

organisms, all members of the same species will share the same architecture. But the architecture will differ between species. To bring this point home, consider that the protein-coding regions of a mouse are 85% identical with those of humans. Yet their karyotypes (i.e. genome architectures) are quite different.

While the old framework focuses on the inheritance of the parts (individual genes), Genome Architecture Theory focuses on the inheritance of the whole system (the genome).

2) Cancer represents a new species of cells in the organismal body.

In studying the relative importance of chromosomal aberrations versus genetic mutations, Heng puts forth the idea that cancer cells are a different species than the host cells, illustrated by drastically changed karyotypes. Furthermore, the cancerous process leads to the emergence of novel cellular systems. For example, many cancer cell lines can live in the lab for decades. Moreover, contagious cancers do exist in the wild, including canine transmissible venereal tumors and Tasmanian devil facial tumors. This transmission of cancer between individuals suggests that cancer is its own entity separate from the host.

3) The whole genome (not the gene) is the primary unit of heritable information upon which evolution by natural selection works.

The evolutionary process has been formally modeled as consisting of three core elements: (A) a population of organisms which (B) have heritable variability (C) such that the variability leads to differential survival in the population. Nothing in this model necessitates or predicts DNA, genes, or genomes. Yet there is an implication of something heritable, which may either change or not change in the process. We know that it can't be the physical body because that always changes in the replication or birthing process (i.e. the child is not the same entity as the parent). The logical conclusion is that there is some kind of *information* which gets replicated (with variability) which encodes for the development of the physical body (or cell in the case of single celled organisms).

We now know that DNA is the way life on Earth encodes the heritable information. And we know that DNA is not a simple structure but a very complex one, involving separate chromosomes, 3D folding structure with temporal dynamics, protein coding segments (i.e. genes), regulatory and repair segments, epigenetics, and even non-locality (as we need to include not just nuclear DNA but also mitochondrial). All of this and more make up the genome architecture.

The old framework focused on the individual genes and how they seem to compete with one another for heritability. And while this is true under certain conditions, it misses the larger picture. The organism (be it a single cell or multicellular) cannot survive or replicate without the entire genome being passed to the next generation. The whole genome is **required** for heredity. Thus it is not merely an empirical question but a logical certainty that Darwinian evolution by natural selection is driven by heritable variation at the entire genome level.

This is the key point that is overlooked by neo-Darwinism. Genome Architecture Theory acknowledges the importance of individual genes in the evolutionary process, but the genome is the primary way life encodes and transmits genetic information from one generation to the next, in a species-specific manner.

Furthermore, from an information theory perspective, the genomic architecture represents a new form of biological information which cannot be gleaned by reductionist methods of studying genes, gene products, genetic pathways and so on. A complex systems view is required.

4) The origin of new species is NOT the gradual result of environmental pressures over many generations.

Darwin's theory about the origin of the various species (and how one species evolves into another) was predicated on gradualism and linear accumulation of small changes (the mechanism of natural selection). He ignored the countervailing fossil record which shows a "punctuated equilibrium": long periods of gradual change within species ("micro-evolution"), punctuated by short bursts of massive changes resulting in (many) new species coming into existence at once ("macro-evolution").

Until Heng's Genome Architecture Theory, nobody has been able to resolve the paradox with a plausible biological explanation (genome chaos, as explained below). Connecting the above, Heng observes that genetic variation leads to micro-evolution and genomic variation leads to macro-evolution. In other words, new species are formed in sudden bursts, not gradually as Darwin hypothesized.

A newly generated species — if viable and adaptive — then becomes preserved and can dominate the population in as little as a single generation (!) through the function of the sexual filter, as will be explained below.

5) Genome chaos is the heretofore unrecognized evolutionary mechanism which drives speciation and cancer metastasis (both being forms of macro-evolution).

Within a given species (including cells within a single organism) the genome is relatively homogeneous and stable. Darwinian micro-evolution can occur, selecting for or against pre-existing or latent traits allowed for by the genome. But this does not lead to the creation of new species or metastatic disease. What does is a sudden spike of massive genomic heterogeneity known as *genome chaos*. Each new architecture created during genome chaos represents a new species. And while the vast majority are not viable, the survivors proliferate and adapt via micro-evolution, potentially forming a stable new species population.

6) Contrary to the standard belief in biology, sexual reproduction makes it harder for new species to evolve, not easier.

The neo-Darwinian explanation of sex is that it provides genetic variation as grist for the evolutionary mill. This is a compelling idea, after all, ecologists have long understood that diversity is the key to a resilient population. However, the mechanics of genetic recombination makes it vanishingly unlikely that genetic mutations get passed down. In other words, sexual reproduction acts as a very strong filter against genomic diversity, due to the requirement of chromosomal pairing during meiosis. Thus Heng argues, sex preserves the species by perserving the genome architecture within the population. As elucidated above, genome chaos (not sex) is the missing link in the speciation puzzle, providing massive amounts of genomic diversity all at once.

Until Heng, the role of sex in preserving adaptive genotypes was lost on everyone - including Gould and Eldredge who first proposed the punctuated equilibrium model. And without a clear causal mechanism, Gould eventually gave up on the significance of punctuated equilibrium later in his career.

7) Even the "healthy" cells in our bodies do not have the exact same genetic code.

I learned in high school that each of our cells has an identical copy of the unique combination of DNA we inherited from our parents. Turns out this is not true. In fact the genome that is passed to

Please cite this article as: R. Furst, The Importance of Henry H. Heng's Genome Architecture Theory, Progress in Biophysics and Molecular Biology, https://doi.org/10.1016/j.pbiomolbio.2021.08.009

the zygote begins to differentiate in subtle (epigenetic) and not so subtle (structural) ways throughout our lives. Healthy adults have a heterogeneous genome! Systemically unhealthy adults (e.g. cancer and autoimmune patients) often have massive heterogeneity. Whether in healthy or unhealthy organisms, this somatic heterogeneity is adaptive and creates resiliency. At some point if the body cannot adapt to bring back integrity/health through the micro-evolutionary process (for instance via the adaptive immune system), the organism goes into crisis, and genome chaos is initiated.

10) Genome chaos may be an adaptive evolutionary meta-strategy.

In the reductionist paradigm, Life is a one-way journey from disorder to order. Instability and chaos are the constant enemy. So the notion that genome chaos and chromosomal instability are a strategic response to crisis is unthinkable to most biologists. Yet this is what Heng's data suggests.

Complexity theorists should be very comfortable with this notion though, as it jives well with concepts like self-organized criticality and life being a process "at the edge of chaos". Computer scientists will recognize genome chaos as a globally optimal strategy for rapidly searching the space of possible new genome configurations.

Intuitively, each peak in the fitness landscape represents a different species. Micro-evolution is a linear, gradual search algorithm which climbs the nearest hill but tends to get stuck in local maxima. In periods where the environment is relatively stable, micro-evolution allows a population to adapt by tweaking individual genes which then compete for clonal dominance. But in highly dynamic/stressful environments where extinction of the entire population is at risk, the genome chaos strategy allows the population to generate radically different genomes; most of these will die off, but the chances are increased of finding a new genome (which represents a higher fitness peak) and the population may thus survive.

11) Inheritance is not precise, it's fuzzy.

With the realization that the organismal genome is not as homogeneous as we once thought, Heng coined the term "fuzzy inheritance" to describe what goes on in mitosis when the body's cells replicate. The heterogeneity of the somatic genome maps to a range of possible cellular phenotypes that may manifest. Thus what's acquired in mitotic replication is not an exact set of traits but rather a range of possible traits, where the range is defined by the heterogeneity of the genome.

The story is slightly different at the organismal population level since the germ line and somatic cell lines are separated at birth, and the sexual filter keeps the germ line pure from a genome architecture perspective. However, practically speaking epigenetic changes (we now know) can be acquired and then passed to subsequent generations, effectively creating some genomic heterogeneity even within the species. And while it's unknown how permanent epigenetic inheritance is, Heng's work shows that at some level, inheritance within a population of organisms is also fuzzy.

The fuzzy inheritance concept provides a causal mechanism for biological resilience, as modest amounts of heterogeneity creates the possibility of adaptation through micro-evolution.

12) Cooperation is just as important as competition in evolution.

The neo-Darwinian lexicon (though not Darwin's himself) is rife with notions of individualism and competition: "survival of the fittest," "nature, red in tooth and claw," "selfish genes," just to name a few. And while it's true that individualism and competition for resources and mates is part of the evolutionary dynamic, the importance of cooperation and collective behaviors is often overlooked. Genome Architecture Theory with its emphasis on the whole genome completes the picture. For implicit in the notion of a whole system is the integrated cooperation between the parts in service of the whole. We see this in pathways and gene interactions, cell tissue, and so on.

Indeed, multicellular life is the manifestation *par excellence* of the cooperative dynamic, as every cell in our body is in fact an individual life form which has evolved mechanisms that have it give up its own interests and reproductive potential in service of the organism. Importantly, the cooperative dynamic is encoded at the genome level.

The balance between individualism and collectivism is essential. Too much of the former results in cancer, too much of the latter results in fragility and lack of resilience.

13) Is it possible that genes code for cellular machinery, and genomes code for living organisms?

The causal link between gene products and biological structure/function at the sub-cellular level seems well-established. Yet, when it comes to explaining how multicellular organisms gain their unique forms and functions, we don't have a good explanation. Even if we restrict ourselves to the single aspect of vertebrate morphology, it has been noted that there's an absence of plausible biological explanation in the literature (Edelman et al., 2016).

While Heng does not explicitly address this issue, one of the pillars of Genome Architecture Theory is that genes encode the parts while genomes encode the whole. Thus, a possible corollary of Heng's theory which deserves further attention is that by and large, genes (and collections of genes) encode for cells, whereas the genome architecture encodes for the organism, including its morphology, tissue architecture, immune system, neural architecture, and so on. Recently, Heng has articulated this idea using the concepts of system information flow and management (Heng and Heng 2021b).

## 3. Conclusion

Roughly 10 Million people die of cancer each year. Without a clear understanding of what cancer is and what it's not, we have no hope of effectively treating or preventing the disease. While there are many scientists who are now openly challenging and pointing out the flaws in the Modern Synthesis, we need a new framework which works better. Heng's Genome Architecture Theory is the best candidate we have so far. It is empirically derived, and explains, predicts and integrates the causal mechanisms of both cancer etiology and Darwinian evolution.

At the theory level, Heng's two-phased evolutionary model reconciles the function of both gene and genome. On a practical level, we should rethink the use of cytotoxic treatments (like chemotherapy), as they are responsible for inducing rapid drug resistance and metastasis via the mechanism of genome chaos.

Evolution is an emergent phenomenon which occurs any time there are populations of entities which can replicate with heritable differences. Seen in this light, cultures evolve too, via the heritable differences of replicating memes, and meme complexes (i.e. "memetic genomes"). The memetic complex called Science is no exception. Heng's work neatly places us at a crisis point. Can we

Please cite this article as: R. Furst, The Importance of Henry H. Heng's Genome Architecture Theory, Progress in Biophysics and Molecular Biology, https://doi.org/10.1016/j.pbiomolbio.2021.08.009

embrace a radical reshuffling of the old framework in service of the higher truth, and better outcomes for cancer patients?

## Declaration of competing interest

None.

## References

Edelman, D., McMenamin, M., Sheesley, P., Pivar, S., 2016. Origin of the vertebrate body plan via mechanically biased conservation of regular geometrical patterns in the structure of the blastula. Prog Biophys Mol Biol 212–244. https://doi.org/10.1016/j.pbiomolbio.2016.06.007.

Heng, H.H., 2016. Debating Cancer: the Paradox in Cancer Research. World Scientific, Singapore. https://www.worldscientific.com/worldscibooks/10.1142/8879.

Heng, H.H., 2019. Genome Chaos: Rethinking Genetics, Evolution, and Molecular Medicine. Academic Press, San Diego, CA. https://doi.org/10.1016/C2016-0-05291-9.

Heng, J., Heng, H.H., 2021a. Two-phased evolution: genome chaos-mediated information creation and maintenance. Prog Biophys Mol Biol S0079–6107 (21). https://doi.org/10.1016/j.pbiomolbio.2021.04.003. 00033-X. Advance online publication.

Heng, J., Heng, H.H., 2021b. Karyotype coding: the creation and maintenance of system information for complexity and biodiversity. Bio Systems 208, 104476. https://doi.org/10.1016/j.biosystems.2021.104476.

Leaf, C., 2004. Why we're losing the war on cancer (and how to win it). Fortune magazine, march 22. https://fortune.com/2004/03/22/cancer-medicines-drugs-health/.

NCI, 2021. NCI budget and appropriations. Updated. https://www.cancer.gov/about-nci/budget. (Accessed 13 January 2021).

Noble, D., 2021. The Illusions of the modern Synthesis. Biosemiotics 14, 5–24. https://doi.org/10.1007/s12304-021-09405-3.

Raza, A., 2019. The First Cell: and the Human Costs of Pursuing Cancer to the Last. Basic Books, New York.

Roser, M., Ritchie, H., 2019. Cancer. OurWorldInData.orgPublished Date: July 2015; updated in April 2018 with latest cancer data for the year 2017 – and revised substantially in November, 2019. https://ourworldindata.org/cancer#is-the-world-making-progress-against-cancer.

Shapiro, J.A., 2021. What can evolutionary biology learn from cancer biology?, 00028-6 Prog Biophys Mol Biol S0079–6107 (21). https://doi.org/10.1016/j.pbiomolbio.2021.03.005. Epub ahead of print. PMID: 33930405.

Statistica, 2021. Global oncology spending from 2011 to 2022. Snapshot as of July 21, 2021. URL. https://www.statista.com/statistics/696208/oncology-costs-worldwide/.

Wikipedia, 2021. Species. Snapshot: august 9, 2021. URL. https://en.wikipedia.org/wiki/Species.

Rafe Furst
*Boulder, Colorado*
*E-mail address:* rafe@rafefurst.net.

22 July 2021
Available online xxx