# Implications of fractal organization of DNA on disease risk genomic mapping and immune function analysis

Simons, M., Pellionisz, A.

In: ASEATTA Proc. of Australasian and Southeast Asian Tissue Typing Association

30th scientific meeting 22-24 November 2006, Chiangmai, Thailand (Peer-invited and peer-reviewed Keynote lecture)

*The late Dr. Malcolm J. Simons (1939-2012), as the award-winning and aptly titled documentary movie "Genius of Junk" documents, was one of the earliest advocates (not the first, but likely to be the most controversial) that mis-labeling "Junk DNA" was a fatal mistake. (He died of a "Junk DNA disease" just as billions of people have died and continue to be devastated by his and other cancers, etc.) The documentary movie, for copyright reasons, was never aired in the USA. However, full a transcript was made available by Australian ABC at http://www.abc.net.au/catalyst/stories/s898887.htm and a video-copy was entrusted by Dr. Simons to be requested for personal use from Dr. Pellionisz at holgentech@gmail.com . In the movie, Dr. Simons said that his meticulous study concluded that non-coding DNA was certainly "not random". He did use the word in the movie "chaotic" but in the colloquial sense of disorder and claimed that evolution would not have conserved such an overwhelming amount unless it did have some function.*

*We probably all agree, that the "everything but" is crucial, but it opens the Pandorra-box of mathematics of full DNA (hologenome) regulation. Malcolm, R.I.P., gravely ill but intensely curious (and curiously intense) has travelled from Melbourne to me in Silicon Valley, California three times for many weeks and to co-chair in Budapest, Hungary to learn and agree that "the genome was fractal". His co-chairing in Hungary and this plus our additional papers testify to this:*

*http://www.junkdna.com/postgenetics_science_paper_and_electronically_enhanced.pdf*
*http://www.junkdna.com/genomics_morphogenesis_and_biophysics.pdf*

FULL ABSTRACT PDF below

Australasian and Southeast Asian Tissue Typing Association
30th scientific meeting 22-24 November 2006, Chiangmai, Thailand

# Implications of fractal organization of DNA on disease risk genomic mapping and immune function analysis

**Malcolm J. Simons (Haplomic Technologies, Melbourne)**
**Andras J. Pellionisz (HelixoMetry, Sunnyvale, California)**

Genetic risk for over 100 diseases has been localized to the MHC. While many of these associations have been established for more than 30 years, the mechanisms remain virtually unknown. In the last century, it was expected that gene characterisation, projected to culminate with the completion of the Human Genome Project, would at last reveal all its causal secrets. With the realization that genetic risk can not be attributed to genes alone, the era of Genetics, with exclusive focus on genes, and dismissal of non-protein coding DNA as "Junk", has now ended. In this new century of Post-Genetics attention is being increasingly directed to non-coding intron and intergenic contributions to genetic phenomena; Genetics beyond Genes. Now a major challenge is to understand the 'language' of all DNA, including non-protein coding sequence. One possibility, first considered in the 1990's, is that DNA is fractal, with the corollary that sequence similarity search algorithms assuming Euclidean geometry may fail to reveal sequence patterns with non-Euclidean properties. Very recently, Rigoutsos and colleagues at IBM reported the presence of over 66 million motifs in the human genome with over 128,000 repetitive short-length sequences. This strongly suggests a multi-component composition of the genome involving non-coding as well as coding elements.

Beyond characterization of DNA sequence patterns, the task will also be to discern functional implications. A fractal approach, first proposed by one of us (AJP, 2002) causally connected fractal structure with fractal processes. Application to the prokaryote Mycoplasma genitalium, as the smallest genome of free living organisms, showed that whole genome accords with the Zipf/Mandelbrot parabolic fractal distribution (personal communication by AJP). This establishes a fractality in the Mycoplasma genome.

In M.genitalium, near-contiguous iterative repetitions of fractosets composing fractogems occur in the intergenic sequences. An immunological phenomenon that directly impinges on genomic organization involves these intergenic repeat elements, which are in coding gene-reading frame but lack an open reading frame. Antigenic phase variation occurs as a result of homologous recombination between FractoGem-containing intergenic sequence and coding genes, in turn subserving escape from immune surveillance and chronicity of infection.

We are now applying to the MHC these genomic fractal analysis tools which the two of us earlier espoused (2006). In the first instance we are examining the 300 kb psoriasis risk region. The goal is to determine whether any fractal composition of multi-SNP haplotypes reveals disease associated risks greater than those observed with recombinant haplotype analysis, thereby refining the mapping of disease genetic risk.

Pellionisz, A. (2002) FractoGene: Utility to use self-similar repetitions in the language-like genetic information as fractal sets. *US Patent Application (Aug. 1st, 2002)*

Rigoutsos, Isidore, Tien Huynh, Kevin Miranda, Aristotelis Tsirigos, Alice McHardy, and Daniel Platt (2006) Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes *PNAS April 25, 2006 vol. 103 no. 17 6605-6610*

Simons, M, Pellionisz, A;(2006) Genomics, morphogenesis and biophysics: Triangulation of Purkinje cell development. *The Cerebellum; 5(1): 27-35*