



# The Molecular Paradigm of Human Complexity

Kallidaikurichi Venkatachalam\*

Professor of Biochemistry, College of Medical Sciences, Nova Southeastern University, USA

Ever since the crystal structure of DNA was discovered, humanity has been fascinated by the double helical structure of DNA. Scientists have estimated that if we were to extend the total of  $6 \times 10^9$  bp of DNA from the content of all 23 pairs of human chromosome, it would be a ladder that would go up formiles [1,2]. We were very pleased to believe that the greater length of the human genome could explain the complexity of human as an organism relative to simple bacteria that has a genome size of only  $4.7 \times 10^6$  bp, a 1000 fold difference. However, in the 1990s when the Human Genome Project gained momentum, we began to unravel that the total number of genes may not be as large a number as first believed. Now we know that the total number of protein coding genes in humans is amere  $\sim 22,700$  [3-5]. This number of  $\sim 23,000$  genes is only slightly higher than the simple worm *C. elegans*. This is startling, yet it is true. With this information we can conclude that it is neither the simple total length of the double helix nor the total number of protein coding genes responsible for the complexity of an organism. It is a conundrum as to how one could achieve a complex individual such as *Homo sapiens* given the fact there are only  $\sim 23,000$  (190,000 exons) protein coding genes. In other words, how do we achieve complexity with just 23,000 genes? We know that bacteria with just  $\sim 5000$  genes can perform pathways such as glycolysis, gluconeogenesis, pentose phosphate pathway, synthesis and degradation of glycogen, fatty acids, all 20 amino acids, purine, pyrimidine, and can perform more pathways to maintain life. Humans can *de novo* synthesize only 10 out of 20 amino acids and can perform many of the other anabolism/catabolism. This means with just  $\sim 5000$  or even fewer genes we could metabolically achieve all that a single celled bacterium can achieve. As an omnivour we depend on beneficial microbes, plants for dietary supplements such as vitamins and essential nutrients to satisfy the energy needs of bare existence, growth, and sustenance. With the remaining 17,000 or so genes we can do all that a simple mouse does (mice have a total of  $\sim 22,000$  protein coding genes). With the remaining  $\sim 1000$  genes that are different between mice and humans can we form a man? Well, let us compare with primates. With only 0.1% difference between human and primates we get 1 or fewer genes to come up with a difference of humans from apes. Speech in humans is controlled by a gene product of FOXD [6] (forkhead box D4), that is a transcription factor which is absent or negligibly present in apes. When we compare the liver functions of apes and humans there are very little differences. However, when we look at the differences of brain protein expressions of human and apes it is perhaps 1000 fold or more. At this point we are very tempted to say that the main differences come about not based on the total number of genes present; it could be due to how we use these genes. In other words, acquisition of differential splicing, expressions/suppressions of sets of genes according to topology, temporality and development can generate quite more complexity. Developmental/morphogenetic gene expression/suppression differences can obviously form unique organisms. When we compare the crucifer plant family, *Arabidopsis thaliana* for example, has slightly more genes than humans [7]. *A. thaliana* needs to perform more pathways including photosynthesis, a physical process that involves the conversion of light energy into chemical energy, and it has to synthesize secondary metabolites for defense purposes from scratch, which takes a lot more genes to code for proteins/enzymes. So the plant is well justified to have more genes because it, being an autotroph, has more tasks to perform in order to survive on this planet. Going back to *E. coli*, the bacteria, it manages to

live as an organism with a mere 4000 genes from a total genome size of  $4.7 \times 10^6$  bp of DNA, relative to the so called non-living thing-a virus that has a genome size range of 3200-800,000 bp with  $\sim 10$ -100 genes or even fewer. So the virus accomplishes its life cycle along with the help of other prokaryotes like bacteria or with the help of eukaryotes with a genome size of 3200 bp and a gene content of  $\sim 10$ . The non-living to free-living took 100 fold differences in the total genome/gene content to have a living autonomy. When we earlier compared *E. coli* with humans we said there are 5 fold higher numbers of genes and 1000 fold differences in the total genome. If we compare the total protein coding genes of *Drosophila melanogaster* (fruit fly), *Takifugu rubripes* (fugu fish), *C. elegans* (worm), *Arabidopsis thaliana* (simple plant), and *Homo sapiens* (human) the total ranged between  $>10,000$ - $<40,000$  [8,9]. Given these a priori numbers the protein coding genes can evolve a variety of eukaryotic organisms that can feed, metabolize, catabolize and perform some specialized functions. Hence it must be in large part the total genome size/content, not the total protein coding genes that controls the higher order complexity in humans. The protein coding gene amounts to less than 2% ( $3 \times 10^7$ ) of the total genome. Now, what is this extra part of DNA doing? Is it junk? Surely nature didn't evolve DNA just for wasteful reasons? Although during trial and error we probably kept some of these regions (the so-called junk DNA), sort of like an evolutionary relic. What functions can we attribute to this extra DNA? We do know that some of the introns ("junk DNA") can actually function as regulatory elements. These regulatory elements can add greater complexity. Perhaps nature is not all that silly.

## Mitochondrial Genome: Evolutionary Relic?

During the course of evolution it is thought that the bacterial genome got incorporated into eukaryotic cells and morphed into mitochondrial genome. The circular mitochondrial genome of humans is composed of  $\sim 16,569$  bp comprised of 37 genes. These 37 genes in total encode 13 proteins, 22 tRNAs, and 2 rRNAs. The transcriptional and translational machinery of mitochondrial genome is very close to that of bacteria. In fact the side effect that we experience from the antibiotic targeted therapy toward prokaryotic organisms (for example, bacterial infections) is due to cross reactivity with our mitochondrial metabolism. By acquiring perhaps the entire organism such as bacteria and modifying it to an extent into a new organelle we have evolved and have added complicated ATP generating machinery to it. Thus we evidence a genetic re-use to generate additional complexity. In conclusion, through genetic re-use of various processes that had been mastered by variety of organisms including the ocean-dwelling

\*Corresponding author: Venkatachalam K, Professor, Department of Biochemistry, College of Medical Sciences, Health Professions Division, Nova Southeastern University, Ft. Lauderdale, FL-33328, USA, Tel: (954)262-1335; E-mail: [venk@nova.edu](mailto:venk@nova.edu)

Received January 12, 2014; Accepted January 16, 2014; Published January 18, 2014

Citation: Venkatachalam K (2014) The Molecular Paradigm of Human Complexity. Clon Transgen 3: e109. doi:10.4172/2168-9849.1000e109

Copyright: © 2014 Venkatachalam K. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

microbes and other creatures, the human organism has perfected its need as well as complexity during speciation.

### Non Coding RNA (ncRNA): A Role on Cell Control, Cell Growth and Differentiation, and Human Development

The control of gene expression during the course of development, spatial differences etc., is now known to be controlled by various types of non-coding RNAs such as MicroRNA (MiRNA), small nuclear RNA (snRNA), and medium and large non coding RNA (lnRNA).

### Epigenetic Regulation (Chromatin Methylation)

Chromatin, especially the histone proteins are modified by a) phosphorylation at the serine residues b) ubiquitination and N-acetylation of lysine residues and c) methylation of lysine and arginine residues. There are several Histone Methyl Transferases (HMT's) that use SAM as the methyl group donor to methylate specific lysines and arginines of histones which forms the characteristic epigenetic markers or signatures. The heterochromatin to euchromatin switch also involves changes in methylation patterns. The chromatin can also undergo remodeling by the removal of specific methyl groups (demethylation) catalyzed by demethylases. The methylation patterns on the histones of chromatins are recognized by specific transcription factors that can activate transcription [i.e., heterochromatin (transcriptionally inactive) to euchromatin (transcriptionally active)] switch. Thus the three processes a) methylation (writing) b) demethylation (erasing) c) methyl pattern recognition (reading) is tightly controlled. Any deregulation in these three processes of methylation, stimulated by environmental agents, dietary contents, or infections, can lead to dysregulation in the normal chromatin biology, uncontrolled cell division and eventually dysregulations such as cancer. The normal aging process also involves gradual changes in the chromatin morphology. Now we have one more issue to deal with: How do we achieve the differences among *Homo sapiens*? Aside from Single Nucleotide Polymorphisms (SNP's), can epigenetic factors such as DNA, RNA and histone methylation control the variety? Perhaps yes. Methylation pattern differences of histones among individuals are quite different—that could perhaps explain individual behaviors. The knowledge contained on the individual genome sequences and epigenetic methylomes can help formulate personalized medicine,

especially for cancers and age-related degenerative diseases and behavior-related issues. The role of methylation in epigenetic control mechanisms on various aspects such as behavioral fixation/plasticity could be very profound for human biology.

Unlike specific mutations of the cell division genes, there are changes on the chromatins that occur above and beyond the gene (Epi) and hence are called epigenomic changes [10]. Aside from the 23,000 protein-coding genes of humans, normal epigenetic changes on the methylation could play a crucial and major role in cell differentiation, organ development, organismal growth and evolution. Thus, it is not just the inherited genes that shape a child's health development, but epigenetics also play a major role. The epigenetic processes are factors that are of concern to molecular cloning and transgenic studies of an organism as well.

### Acknowledgement

I couldn't possibly include all references since there are thousands. Instead I would like to acknowledge the contributions made by all scientists in the field of human genomics, comparative genomics, epigenetics and the related areas of my article. All of the relevant background articles are available at the NCBI website, particularly the PubMed.

### References

1. Watson JD (1965) Molecular biology of the gene. New York.
2. Watson JD (2012) The annotated and illustrated double helix. Simon & Schuster, New York.
3. Venter JC (2007) A Life Decoded. Penguin Group, New York.
4. Venter JC (2003) A part of the human genome sequence. Science 299: 1183-1184.
5. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. (2001) The sequence of the human genome. Science 291: 1304-1351.
6. Tuteja G, Kaestner KH (2007) SnapShot: forkhead transcription factors I. Cell 130: 1160.
7. Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815.
8. Brenner S, Elgar G, Sandford R, Macrae A, Venkatesh B, et al. (1993) Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. Nature 366: 265-268.
9. Brenner S (2013) Brenner's Encyclopedia of Genetics. Elsevier.
10. Tollefsbol OT (2012) Epigenetics of Human Disease. Elsevier.