Life AND DNA

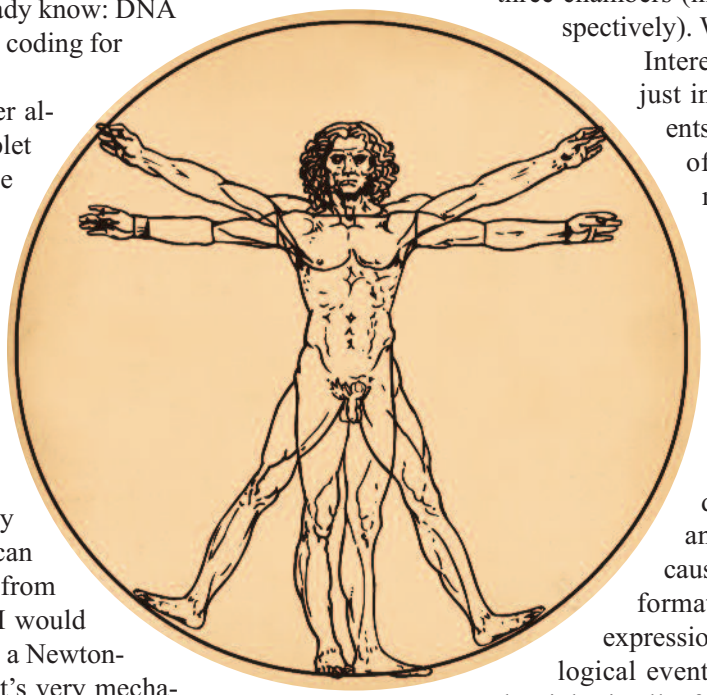
By Michael Harwood

This article was catalysed by "Who's your daddy" by Frieda Wong in the November issue of *MC²*. Her essay caught my interest when she discussed the "deeper genome" and triple and quadruple stranded DNA. I'm going to write about some of the related ideas that I've come across in my layman excursions into biochemistry, and I apologize up front for the technical jargon. This article would be a lot longer if I took time to explain all of the background.

First of all, the results of the ENCODE project, which has found that at least 80 per cent of our DNA is biochemically active, are in some ways not surprising at all, as I'll discuss later. On the one hand, yes, it has ruffled some feathers because it flat out contradicts the idea that we have a lot of junk DNA left over from our evolutionary past. There are some scientists who are very unhappy with the results of the ENCODE project and are trying to find flaws in it, which, one must admit, is part of the way science progresses.

However, even before this data was released in 2010, the phrase "junk DNA" was already falling out of favour: Google's Ngram viewer and my own searches show it declining after 2003. Biochemists have been discovering more and more functions for non-coding DNA (e.g. Kozak sequences, promoter regions, enhancer regions, genetic switches, etc). The ENCODE project just builds on what we are already know: DNA is a lot more than just a bunch of genes coding for protein and RNA enzymes.

Our understanding of the four-letter alphabet of DNA (C,A,T,G) and the triplet code for amino acids is an incredible discovery. It works just like a tiny machine reading instructions from a tape and making the protein that corresponds to the instructions fed into it. It is really astonishing that a three-dimensional structure can be made from the 46 pieces of one-dimensional DNA in our chromosomes. We have a complete understanding of how DNA makes proteins, but this is only just scratching the surface. As far as I can tell, no one has any clue how we get from proteins to a complete 3D organism. I would liken this stage in our understanding to a Newtonian stage (by analogy with physics). It's very mechanistic and quite easy to understand. Physics has undergone two further major revolutions: Einsteinian or relativistic, and quantum mechanical. Our understanding of the information content of our DNA, our genome, needs equivalent seminal revolutions.



One of the interesting consequences of the results from ENCODE is that, assuming that almost all of our DNA is necessary, how do we cope with random mutations? Formerly it was thought that so little of our DNA is vital that most mutations happened in the non-coding 98 per cent (junk DNA) and so had no harmful effect on the organism. (By the way, no one ever actually verified the "junk DNA" hypothesis by removing all of the DNA that was thought to be junk and seeing if a viable cell could be created.) We know that the nucleus puts forth tremendous resources to try to eliminate every single mutation (the error rate in DNA replication is less than one in a billion). However, given that almost all mutations are harmful (a fairly safe assumption), our DNA must be degrading much faster than we would have imagined. Maybe we are actually getting stupider and more prone to genetic diseases than our ancestors 40,000 years ago. With their larger brains, Neanderthals probably were a lot smarter than we are.

There is a lot that we don't know about DNA. For example, what all of the genetic code means and how it got there in the first place. We only have a complete understanding of the two per cent that comprises our genes. Consider this: Every human being has a four-chambered heart that looks exactly the same: two ventricles, two atria, valves, the aorta, etc. A set of instructions for making a heart specifically according to this pattern must be somewhere in our DNA, but no one knows where. We know it is not the case that all hearts just automatically form like this (akin to how a bubble is automatically spherical due to surface tension), since there are hearts with two chambers and three chambers (in fish and amphibians, respectively). Where is this data stored?

Interestingly enough, we don't just inherit DNA from our parents, but also inherit a number of other things. We inherit mitochondria with their own rings of DNA, but mostly from our mothers, since the sperm cells' mitochondria don't make it inside the egg cell. We inherit the histones upon which the DNA is wound into condensed chromosomes – and these are important because they carry epigenetic information which modifies the expression of genes due to physiological events in the parent's life. We also inherit all of the other organelles that make up human cells. As far as I know, it hasn't been proven that these other biochemical structures (e.g. centrioles) don't also contribute to the specific inherited characteristics of the

offspring.

However, I don't see any of these other inherited structures and molecules as having the requisite information-storage ability to hold the blueprints for the organs that make up our body.

Frieda Wong mentioned new research into triple and quadruple stranded DNA. Personally, I doubt that the information is stored in these more complex 3D constructs made of DNA. It would seem too easy for the DNA to be misaligned with other strands, and also we don't know of any mechanism that could read a triple strand of DNA, whereas RNA polymerase does just fine with the ordinary double strands. I suspect that the information that I see as essential to creating a human being is stored in the non-coding DNA, in the vast mass of DNA that doesn't code for proteins and whose function we don't understand.

How could the data to design organs be stored in our DNA? A friend of mine suggested that it could be holographically stored across multiple chromosomes. We know that proteins that affect the heart's development and function are not all on one chromosome. I don't know enough about holographic information storage, but we know that one fragment of a hologram still contains the complete image. It's very unlikely that this is possible with data stored in a one-dimensional structure, and holograms are definitely not one-dimensional. Our DNA is a lot more analogous to computer code. I suspect that there is data compression built into it as well as redundant features and error checking. We know that blood vessels and tree crowns have a fractal nature. Fractals excel at making branching tree-like structures with a very small set of instructions. This is one type of very efficient data compression that must exist in our genome. Our bodies clearly don't have specific instructions of where every single blood vessel and capillary must be placed – it would take up way too much data space to specify all 100,000 kilometres of our blood vessels. However, somehow the fractal nature is guided and switched off and on at the appropriate places. For example, we all have an aorta and the left subclavian artery is always the third branch off it. No room for diversity there.

The location of the large amount of information needed to form our organs and organ systems is currently unknown. Our best guess is that it is indeed in our DNA and that is why I've always thought that all our DNA is functional, and thus the ENCODE results are not at all surprising to me. We need to move from the Newtonian stage of understanding DNA to the Einsteinian stage. We won't know how many stages there are until we proceed to the next one and clearly look at what exactly it is that we still don't understand. Yes, the Newtonian stage of Watson and Crick is an earth-shattering discovery, but it's only the beginning.

There are two other interesting considerations about data storage in DNA to consider. First of all, it's well known that most amino acids have more than two possible codons that code for them (e.g. glycine is made by any of GGT, GGC, GGA, GGG). In September, Dr. Yi Liu of UT Southwestern demonstrated that the choice of codon changes the speed at which the protein is made, and this will change the folding of the resultant protein and thus its function, while the actual protein is still made of exactly the same chain of amino acids! Astounding.

We had never imagined this possibility — although it clearly happens in baking: Take the same ingredients and you'll get different cakes if you bake them fast or slow. There seems to be an extra layer of information based on which of the redundant codons is actually in the gene for a particular protein.

Secondly, renowned biologist Denis Noble has shown that the majority of essential proteins have more than one way to be produced in the body. If one gene gets destroyed, there is built-in redundancy to bring other dormant genes into function and keep the body alive. In a 2012 talk, he gave an example of genes involved in voltage regulation of our heart's pacemaker. (See YouTube "British Biologist Denis Noble Debunks Neo-Darwinism"). This sort of extra layer of backup and redundancy is mind-boggling to ponder. Noble proposes a systems biology approach that overlaps with some of what I've been talking about here. He does go a bit further into speculative areas than I'm comfortable with, but I think he's entirely correct in that the reductionist approach will not get us much further.

We need to admit to our very limited understanding of biochemistry and DNA. The situation is dire enough that there really isn't a solid biochemical underpinning to evolution. Since we can't actually explain in a complete step-by-step, molecule-by-molecule procedure, how an existing organ is formed, how can we hope to explain how something like a two-chambered fish heart evolved into a three-chambered frog heart? We have enough problems explaining this logically even without involving biochemistry.

The research now being done on DNA will probably require quite significant changes in the theory of evolution and how it relates to biochemistry. I'm looking forward to the day when we can read and understand DNA in the same way that people can read blueprints for a Boeing 747 and can assemble it — rivet by rivet, strut by strut.

(Part 2 will appear in the next issue)

Michael Harwood is a Mensan who lives in London, Ont.