HUMANKIND'S MOST EXCITING JOURNEY: THE HUMAN GENOME PROJECT

The shaping of life forms through the principles of mutation and natural selection have resulted in the biological diversity, both past and present, on this planet.

RAJ RAMESAR
MSc, PhD
Professor and Head
Division of Human Genetics
Faculty of Health Sciences
University of Cape Town

Raj Ramesar serves as Director of the MRC Human Genetics Research Unit and CANSA's Colorectal Cancer Research Consortium. He is principal investigator on the retinal degenerative disorders research project. Apart from being involved directly with several established research projects aimed at elucidating the genetic basis of diseases in South Africa, he is currently channelling his energy into setting up research into understanding the genetic basis of the more complex yet common chronic disorders (e.g. hypertension) in our populations.

The plasticity of life forms is shaped by biological possibilities against a vast range of environmental pressures. This plasticity is a feature of the mutability of the fundamental building blocks of life, namely DNA. From the beginning of life, to the evolution of humanity, is a testimony to this creative process, eloquently described in Richard Dawkins' book, *The Blind Watchmaker*.[1]

The origins of humanity, its dispersal over the surface of the planet and the emergence of civilisation are the result of humankind's innate exploratory nature. Having emerged from the great plains of Africa and migrated to various parts of the planet, humankind has obviously been exposed to an unimaginable number of selective pressures. If one could not appropriately mobilise in the face of imminent physical danger, one would succumb; similarly if one were ill and could not keep up with the migrations, one would have been selected against. The move from hunter-gatherer to relatively stable civilisations is reflective of mankind's grasp of the very powerful concepts of livestock domestication and plant propagation; subsequent agricultural breeding for desired traits may have been our first true genetic experiments.

Inquisitiveness, needed for survival in various, often testing, environments, selected for the most imaginative individuals who could think of new ways to mobilise, subsist and exploit their terrain. This process might have resulted in the establishment of this trait as the hallmark of humanity.

Over the aeons, the quest for new terrain and knowledge has led to an increasing number of disciplines that continue to grow as a reflection of our intrinsic inquisitiveness. These disciplines spawn sub-disciplines, which in turn spawn sub-sub-disciplines, and so on.

SCIENCE IS A JOURNEY

As much as there has been a need to explore our physical environment, on earth, there has been an urge to understand what lies within our oceans and skies, and beyond. Ferdinand Magellan, no doubt preceded by countless of thousands of unsung explorers all over the world, was the first European to cross the Pacific and navigate the globe. In a similar exploratory manner, Galileo's passion for exploration was turned heavenward; and one of the more recent technologically advanced space telescopes (Galileo) bears his name. The Hubble telescope and several other more recent astronomical developments are very rapidly allowing us a view across the universe and to moments back in time, in an unprecedented way.

At a microscopic level the advances have been at least as remarkable as the macroscopic developments. The invention of the microscope by Antonie von Leeuwenhoek was a necessary prerequisite to the vast array of high-definition

The origins of humanity, its dispersal over the surface of the planet and the emergence of civilisation are the result of humankind's innate exploratory nature.

The invention of the microscope by Antonie von Leeuwenhoek was a necessary prerequisite to the vast array of high-definition microscopic instruments currently available, including the electron microscope.

microscopic instruments currently available, including the electron microscope. Developments in the field of imaging, generally, have allowed us to view aspects of the human body unimaginable only a few years ago. Magnetic resonance imaging (MRI), pulse emission tomography (PET) and computed axial tomography (CAT) scans allow us to visualise the real-time structure and function of the different aspects of our physical bodies. The plethora of medical interventions currently available, ranging from the array of drugs to the various surgical technologies that are possible, is nothing short of mind-boggling.

The advances in medicine have moved from poultices in the days of yore, to current designer drugs targeted at specific components of biochemical pathways, known to cause disease. In surgery, the practice has evolved from roughly severing an extremity (when the practice of the surgeon and barber coincided) to sophisticated laser-based excisions even at a cellular level. All of these exciting clinical journeys converge with the developments in molecular medicine. This quest for better understanding pathologies and coming up with rational therapies has led to the Human Genome Project (HGP), man's most exciting journey.

GENOMICS: A NEW QUEST

Genetics refers to the study of patterns of inheritance of specific traits, and the mechanisms of heredity and biological variation. A more recently coined term, genomics, on the other hand, refers to the study of the sequence, structure, and function of the genome, an organism's 'biological blueprint' of DNA, chromosomes and genes. Although genetics may well have been used for aeons, formal literature recognises Charles Darwin for setting the scene for genetics. His description of 'Evolution' in 1859 and ideas of random variation and selection prepared the soil for genetics. The independent contribution of the Augustinian monk, Gregor Mendel, through his description of 'the basic laws of heredity and independent assortment' in 1862 sowed the seeds for the concept of genetic mutations. This work lay largely unrecognised until 1900, when it was rediscovered.

A range of subsequent advances that contributed to the development of genetics as a formal science included the discovery of chromosomes by Walther Fleming in 1882, and the recognition by Thomas Hunt Morgan that some traits may be sex-linked and reside on chromosomes. The discovery of the double helix structure of DNA by James Watson and Francis Crick (1953) was the first major step towards the field of molecular genetics, which was a necessary prerequisite for genomics. In 1964, Charles Yanofsky showed the congruity of the sequence of nucleotides in DNA with amino acids in proteins. In 1969, James Shapiero of Harvard University, working with Jonathan Beckwith, isolated the first integral (bacterial) gene. In 1972, Paul Berg's success with recombinant DNA methods was a major influence on the development of the technologies that have come to be relied on in any genetic/genomics endeavour. Fred Sanger's subsequent description of a method for sequencing DNA was a major breakthrough that enabled later researchers to map and sequence the genomes of various species. These major steps led to the plethora of subsequent discoveries, including that of specific genes which were cloned and used for therapeutics, e.g. insulin. These technologies led to a new era of gene mapping and identification of specific genes causative of human disease (e.g. Huntington disease and haemophilia), and for various therapies that were gene based and included gene therapy.

GENES AND DISEASE

From a human disease perspective, most developmental and degenerative disabilities, including blindness, deafness, crippling and mental retardation, have proven recalcitrant to scientific investigations aimed at their treatment or cure. The recognition of genetics as a fundamental tool for identifying fragments of genetic material that are shared by affected individuals (either in a family or community) has made it possible to identify the gene/s responsible for the morbidity attributable to a particular disease. The HGP, which rivals the international space programme in cost and complexity, has been aimed at exactly this. This programme has made available the 'flat human genomic DNA sequence' at the fingertips of researchers, and has gradually populated this flat sequence with known and previously unknown genes. Other genes still wait to be discovered. Although fundamental to the identification of disease-causing genes, this genetic 'catalogue or reference manual' does not of itself point to disease-causing genes. Many common diseases, which may appear clinically identical in unrelated patients, are heterogeneous, or caused by any one of several genes or genetic defects, placed seemingly randomly on the 23 pairs of chromosomes that make up our genome. It is worth considering that this phenomenon of heterogeneity may be the reason why individuals with apparently the same

Inherited diseases in large families provide the means to identify shared fragments of the genome, and the eventual identification of the 'culprit gene' that lies at the heart of the biological process of the disease.

The HGP, in effect, is over. However, not all genomic secrets have been revealed.

disorder or phenotype respond very differently to the same drug.

Inherited diseases in large families provide the means to identify shared fragments of the genome, and the eventual identification of the 'culprit gene' that lies at the heart of the biological process of the disease. Once genes are located and identified, their role in the cellular and tissue milieu may be determined through cell biology studies, and work towards finding new highly targeted therapeutic strategies can begin in earnest.

With the recognition that inherited diseases reflect shared regions of the genome, geneticists explored the possibility of identifying genomic reagents to identify specific regions of the genome. With the emphasis on human disease, the developing technologies in recombinant DNA allowed a few intrepid travellers to consider localising the genes for such disorders. The relatively slow pace of disease gene identification led a group of researchers to devise a strategy of comprehensively mapping many disease genes, and then to identify all genes in humans to the level of DNA sequencing — this was the HGP.

THE HUMAN GENOME PROJECT

All of these advances led to the capacity to map and identify single human genes underpinning disease traits, where large enough numbers of individuals, related by blood line, were identified with the same disorder. The relatively slow pace of identifying human genes led to a watershed meeting (in the 1980s) between the Department of Energy (DOE) and the National Institutes of Health (NIH) in the USA and resulted in plans for a larger scale onslaught towards identifying the entire genomic complement of human genetic material. This was the advent of the HGP, which relied heavily on extrapolated developments, since the extant technologies were relatively unsuitable for the envisaged gigantic scale-up operations. The era of genomics was thus born. There was a realisation that a well-serviced informatics environment was an absolute necessity in order to provide the complex annotations and systematic storage and retrieval capacity for this scale of information. The technology being developed for the HGP afforded researchers the capacity to sequence the genomes of other organisms — for various reasons, primarily that of comparing biological systems and functions, and the amenability of these other organisms to genetic experimentation (**http://www.ceolas.org/VL/mo/**).

The HGP, in effect, is over. However, not all genomic secrets have been revealed. The human genome contains 3 164.7 million chemical nucleotide bases (A, C, T, and G), with the average gene consisting of 3 000 bases. However, gene sizes vary greatly, with the largest known human gene being *titin* at 10 million bases. The total number of genes is estimated at between 20 000 and 25 000, which is much lower than previous estimates of 80 000 - 140 000 that had been based on extrapolations from gene-rich areas of the genome. Almost all (99.9%) nucleotide bases are exactly the same in all people. The functions are unknown for over 50% of discovered genes. The basic information emerging from the genomic project has spawned a plethora of fields of endeavour; the most important one, which is aimed at showing the animation of the cell, through an accurate inventory of interacting proteins, is called proteomics. A range of other '- ics' reflect their derivation from the 'genomics' enterprise.

HGP BENEFITS: CLINICAL AND BEYOND

The envisaged benefits of access to genome-based medicine include improved diagnosis of disease, earlier detection of genetic predisposition to disease, rational drug design, gene therapy and control systems for drugs and pharmacogenomics or more effective 'customised drugs'.

Technological advances in microbial genomics will result in rapid detection and treatment of pathogens (disease-causing microbes). Manipulation of microbial genomes will also result in important new energy sources (biofuels), and the possibility of safe, efficient toxic waste cleanup. Genomic sciences will also afford the capacity for environmental monitoring to detect pollutants and protection from biological and chemical warfare.

The use of DNA technology in forensics has already seen ready application, and is commonly used to identify potential suspects by DNA matching with evidence left at crime scenes. Equally this technology may be used to exonerate persons wrongly accused of crimes. These technologies will also help to identify disaster victims accurately, while affording the capacity to establish paternity and other family relationships.

The use of DNA technology is useful to identify endangered and protected species as an aid to wildlife officials (could be used for prosecuting poachers), while authenticating consumables such as caviar and wine, and to determine pedigree for seed or livestock breeds.

Molecular genetics is being used in areas previously unthought of: bioarchaeology, evolution and anthropology. The emerging knowledge based on DNA technology has led to major advances in understanding human

migration patterns. Also, agriculture, livestock breeding, and bioprocessing are benefiting from the current technologies.

Of particular interest during the development of the HGP was recognition of the need for a strong ethical framework within which technology was developed and used, and how this would be translated to society. The HGP has formally recognised and funded a programme aimed at understanding the ethical, legal and social implications of using genetic and genomic information. Some of the pertinent issues include:

- **The need for privacy and confidentiality of genetic information.** Genetic information is regarded as a more definitive means of risk assessment than previously available. Furthermore, genetic information from one individual of necessity provides information on other members of the same genetic lineage. Because there is no consensus on the impact of information on genetic predisposition to disease, fairness in the use of genetic information by insurers, employers, courts, schools, adoption agencies, and the military, among others, needs to be ensured.

- **The process of doing genetic research** is necessarily linked to an overarching process, including adequate and informed consent to facilitate the use of genetic information in reproductive decision making.

- **Clinical issues** include the education of doctors and other healthcare providers, people identified with genetic conditions, and the general public about capabilities, limitations, and social risks, and implementation of standards and quality control measures.

Uncertainties associated with gene tests for susceptibilities and complex conditions (e.g. heart disease, diabetes, and Alzheimer's disease) should be dealt with ethically, and with due consideration before being used in any clinical environment. Fairness in access to advanced genomic technologies is pertinent in a setting such as South Africa, where the derived information is at risk of being used only or mostly in economically privileged settings. A range of conceptual and philosophical implications regarding human responsibility, free will versus genetic determinism and concepts of health and disease needs to be investigated and addressed before the application of genetic information. No doubt, health and environmental issues concerning genetically modified (GM) foods and microbes, and the commercialisation of products including property rights (patents, copyrights, and trade secrets) and accessibility of data and materials should be addressed as part of the greater genomic endeavour in any society.

Genetics, and its much more powerful offspring, genomics, have led humanity onto a very slick road that has little traction. Finding the correct vehicle to navigate this track will probably be the greatest challenge. Without checks and balances, the promise that is better health, food security and safe biotechnological approaches to humankind's needs will be endangered, as much as a conventional car is on ice.

Reference

1. Dawkins R (1986). *The Blind Watchmaker: Why the evidence of evolution Reveals a Universe Without Design*. Vermont: WW Norton.

IN A NUTSHELL

Arrival of humankind on earth has been marked by adventurous explorations in many different directions; many of these journeys have been against incredible odds.

Collectively, the advances have placed humankind on the pile of technological society, which continues to rise at a remarkable rate.

The macroscopic discoveries and inventions have been matched by remarkable and innovative means of microscopic resolution and imaging.

The Human Genome Project, which started off as a fairly limited journey to identify disease-causing genes, has turned into one of the most exciting adventures of humankind that aims to plumb the depths of who and what we are as biological beings.