

# THE HUMAN GENOME PROJECT IN CLINICAL PERSPECTIVE: THE EMPEROR'S NEW CLOTHES?

*Almost all human disease (barring trauma) is influenced by our genetic make-up — either directly as an inherited metabolic disorder or by altering our susceptibility to pathogens.*



## RIK DE DECKER

*MSc, MB ChB, FCPaed (SA), Cert Med Genetics (SA), DCH (Lond)*

### Senior Specialist and Paediatric Cardiologist

*Division of Human Genetics and School of Child and Adolescent Health University of Cape Town and Red Cross Children's Hospital Cape Town*

*Rik De Decker was admitted to the Fellowship in Paediatrics (SA) in 1998, subsequently subspecialising in both paediatric cardiology and medical genetics. He is currently the senior consultant clinician in the Division of Human Genetics at the University of Cape Town, and a part-time paediatric cardiologist at Red Cross Children's Hospital. His research interest is the genetic aetiology of congenital heart disease.*

## THE SPECTRUM OF HUMAN DISEASE

Many human diseases are clearly Mendelian inherited disorders with readily discernable autosomal, X-linked, or mitochondrial inheritance patterns. For many of these disorders (e.g. achondroplasia or haemophilia) the genetic aetiology is well known, and diagnostic or predictive genetic tests are available. More frequently, however, diseases are caused by defects in many genes or in a gene controlling a complex pathway combining both genetic and environmental factors (e.g. spina bifida or cleft lip) — known as multifactorial disorders (Fig. 1). Even susceptibility to HIV infection is known to be genetically determined by the presence or absence of the T-cell CCR5 receptor to the HIV. Furthermore, an individual patient's response to a drug may be altered by genetic factors affecting the metabolism of that drug, rendering the therapy more or less effective or even toxic to that patient.

It is therefore clear that the more we know about our genetic material (our human genome), as well as the genomes of our pathogens, the better we should understand our susceptibility to disease and the more targeted our treatments may become. The Human Genome Project (HGP) aimed to produce the complete sequence of all the bases in all of human DNA, but in addition, to sequence the DNA of important pathogenic and research organisms. This knowledge should be valuable in defining our inherited susceptibility to disease and therefore be of fundamental importance to the understanding of the pathogenesis and treatment of almost all human disease (barring trauma).

What relevance does this knowledge have to today's clinician, and what help may he or she reasonably expect from this gushing yet bewildering fountain of knowledge? This review attempts to address these questions and aims to place the HGP in clinical perspective for today and the future.

## The HGP: what did it promise?

A comprehensive review of the aims, history and achievements of the HGP is given by Prof Raj Ramesar in this issue of *CME* (p.8). Its technological achievements are astounding; unfortunately these achievements were too often and too readily translated into predictions of *clinical* advancements: John Bell, founder of Oxford's Wellcome Trust Centre for Human Genetics, wrote in 1998 in the *BMJ*:<sup>1</sup> '*...within the next decade, genetic testing will be used **widely** for predictive testing in **healthy** people and for diagnosis and management of patients...*' [*my emphasis*].

In 1999 Francis Collins,<sup>2</sup> Director of the US National Human Genome Research Institute, wrote with passion of: '*...a new understanding of genetic contributions*

to human disease and the development of rational strategies for minimizing or **preventing disease phenotypes altogether**' [my emphasis].

The popular media followed suit. A *New York Times* editorial<sup>3</sup> in December 1999 proclaimed: 'Health care will shift from a focus on detection and treatment to a process of prediction and prevention... You can imagine having an infant tested at birth... and a result that says you are susceptible to disease A, B and C.'

It is easy to understand the eager enthusiasm behind these claims, but in the world of the clinician, where evidence-based medicine is demanded, perspective needs to be maintained. Therefore these enthusiastic claims soon had sceptics questioning their validity and scope, such as the comment below from Holtzman and Marteau<sup>4</sup> in 2000: 'Statements like these clothe medicine in a genetic mantle... [but] the genetic mantle may prove to be like the 'emperor's new clothes.'

I do not believe that the HGP will prove to be the 'emperor's new clothes', a false promise, but its output must be placed in context of current and future clinical and therefore genetic clinical practice. To do this one must start with the advances made by clinical genetic research before the publication of the HGP in 2001.

### RECENT ADVANCES — BEFORE THE HGP

Classic clinical genetic teaching has generally focused on the simpler, yet rarer Mendelian disorders such as achondroplasia, Duchenne muscular dystrophy, cystic fibrosis, etc. The genetic mechanisms for these disorders are generally well understood and genetic testing and counselling are available. Many chromosomal disorders such as Down syndrome and Turner syndrome are also amenable to testing and prediction of recurrence risks, and are classic subjects in clinical genetic teaching. The remainder of

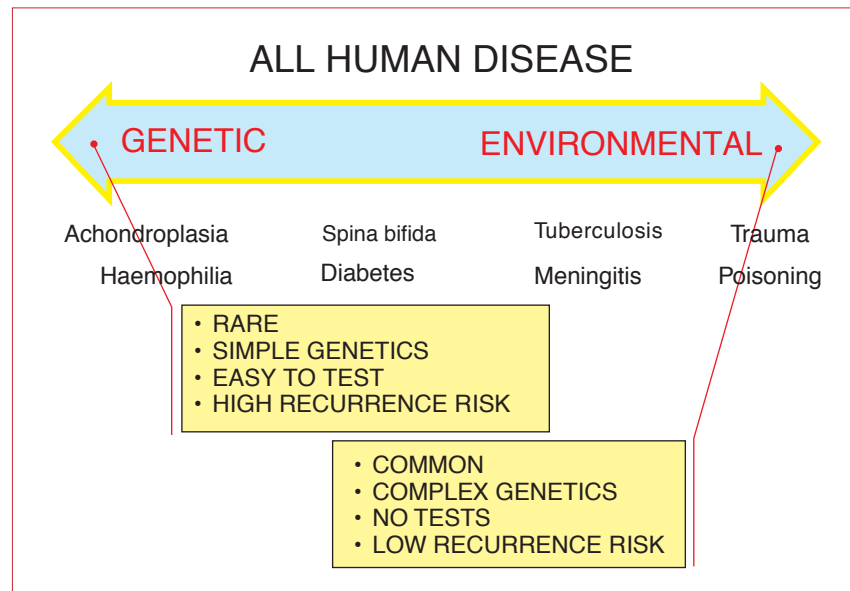


Fig. 1. All human disease lies on a spectrum from a clearly defined genetic aetiology to a complex interaction between genetic predisposition and environmental factors.

clinical genetics seemed to be concerned with the identification and management of rare obscure dysmorphic syndromes, seldom encountered by general practitioners.

However, more recently, many noteworthy advances have been made in the understanding of common genetic disorders, which do not readily fall within the ambit of classic chromosomal or autosomal genetic disorders. Some advances include:

- the fragile X syndrome: an atypically X-linked disorder due to a dynamic mutation of the FMR1 gene and the commonest cause of inherited mental retardation (see the article by Dr Karen Fieggen, (p.29 of this issue)
- the 22q deletion syndrome: a chromosomal microdeletion disorder which is the second most common cause of congenital heart disease after Down syndrome
- other microdeletion disorders, such as Williams syndrome
- uniparental disomy and imprinting disorders such as Prader-Willi and Angelman syndromes
- trinucleotide repeat disorders such as Huntington disease and spinocerebellar ataxia
- skewed X-inactivation as an explanation for the unusual occurrence of X-linked recessive disorders in girls.

All these advances, and others, particularly in the field of epigenetic (non-DNA) control of gene expression, have come in spite of, or preceded the publication of the human genome sequence in 2001. Yet few have entered the ambit of the general physician.

### The HGP: what has it not achieved?

Disappointingly, despite knowledge of the full sequence of all human DNA, the exact aetiology of some common genetic disorders has remained frustratingly elusive. For example, the full sequences of DNA of the two smallest chromosomes (chromosomes 22 and 21) were published in *Nature* in 1999<sup>5</sup> and 2000,<sup>6</sup> respectively. A microdeletion with the loss of some 30 genes in chromosome 22 gives rise to the 22q deletion syndrome; three copies of chromosome 21 cause Down syndrome.

Despite knowledge of the sequence of the sequence of every missing or extra base pair in these two common disorders, the pathogenesis of neither is well understood. In both, certain genotype-phenotype correlations have been defined, but full genetic explanation

remains a mystery. This concern reaches far beyond mere academic aetiological curiosity, and touches clinician and patient directly. Antenatal testing for both conditions is generally available, but the imprecise knowledge of how these genetic aberrations give rise to malformation precludes accurate antenatal counselling with respect to phenotypic expression. This is of particular concern in the 22q deletion syndrome, which has an unpredictable and wide phenotypic spectrum ranging from near normality to multiple severe anomalies. Similar comments can be made of other common or 'well-known' genetic disorders, which despite full genomic definition, have unpredictable expression.

**THE AETIOLOGY PUZZLE: WHY?**

How is it possible to have complete knowledge of the fundamental building blocks, yet remain ignorant of their consequences? There are two predominant reasons for this failure, and both lie in the interpretation of the effects of changes (mutations) in the sequence:

- defining a significant, disease-causing mutation in a known sequence is not simple, but more importantly,
- predicting the effect of that mutation on the clinical phenotype is even more difficult.

**The output from the HGP**

The genomic output from the HGP consists of no more than a very long sequence of 4 letters (A, C, T and G) in a very precise order. This complete sequence of more than 3 billion letters (the human genome) is stored in immense databases for further analyses (see the article by Dr George Rebello, p.30 of this issue). The challenge lies in determining what the sequence means when transcribed and translated into functional proteins. By far the majority (99.9%) of the sequence is not transcribed; it is in the remaining fraction of only 0.1% that most normal human variation lies, as well as the foundations of most genetic disease. An even larger challenge lies in determining the effects of alterations



Fig. 2. The output from the Human Genome Project. The significance of a single base change is not readily apparent!

of the transcribed sequence of our approximately 40 000 genes on health or disease. Furthermore, a single alteration does not necessarily cause disease: approximately 1 per 1 000 base-pairs is altered without being a disease-causing mutation (a so-called single nucleotide polymorphism or SNiP). The difficulty lies in showing that a specific sequence alteration sufficiently alters the gene's protein product to have a significant effect or be pathogenic (Fig. 2).

**The gene mutation model**

Classic teaching is that an altered (or mutated) gene produces an altered protein which in some small or large way may have an effect (sometimes deleterious) on the organism bearing that gene (Fig. 3). In addition, the mutation may be in the organism's germline (ova or sperm in humans) and may therefore be transmitted to offspring — the typical scenario for an inherited trait or disease.

Once a mutation has been shown to exist in a transcribed gene, it is still necessary to place that altered gene in context. The gene's protein product may have multiple disease-causing downstream effects; it may be a regulator or trigger for other genes, it may be a DNA damage repair protein, it may play varying roles during differ-

ent developmental pathways, alter the susceptibility of other genes to mutation, trigger apoptotic or cell growth genes, and so on.

Furthermore, the effect of the mutation on the protein structure will need to be defined: the effect of this alteration may be subtle or profound, which depends critically on the site of the mutation in that gene. A mutation may truncate a protein to a functionless remnant, while a neighbouring mutation may merely induce a subtle change in its folding characteristics, inducing variable functioning under different conditions. This is the burgeoning science of proteomics. Merely demonstrating an alteration of the sequence produced by the HGP is insufficient to explain human disease; many more complex processes need to be understood to begin to grasp the genetic underpinnings of a multifactorial disorder.

**HOW DOES THE HGP HELP THE PHYSICIAN?**

Where does this complexity leave the medical practitioner? And how can the HGP be helpful to the clinician? In everyday reality, the HGP is of little use to the clinician. It is little more than a tool, albeit a very powerful

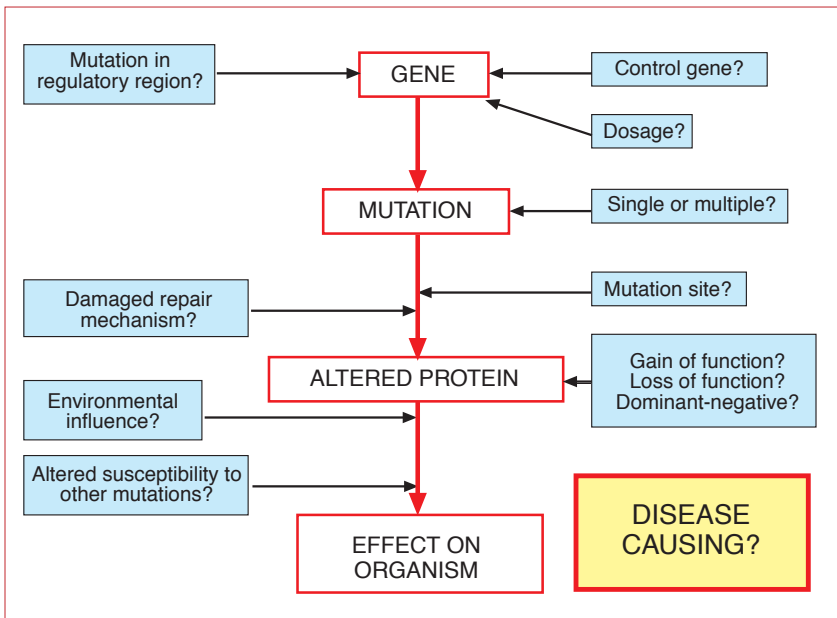


Fig. 3. The classic gene mutation model. A mutation must be placed in context and clearly shown to be disease causing.

one, in the hands of the research geneticist. It is of enormous value in assisting with the dissection of the genetic fundamentals of health and disease. The gap between genetic research and routine clinical care can now begin to be bridged by knowledge of the complex interplay between inherited and environmental factors which give rise to variations and disease phenotypes. Already many previously ill-understood diseases have been explained in terms of their basic genetic aetiology, allowing targeted therapies to be designed, following precise knowledge of their abnormal physiology.

### THE HGP: FUTURE BENEFITS

As these inroads into genetic disease are being made, the clinician will be able to reap benefits from improved understanding of disease processes. Some examples of possible future clinical benefits include:

- susceptibility to pathogens (e.g. HIV, TB)
- susceptibility to teratogens (e.g. fetal alcohol syndrome (FAS), warfarin embryopathy)

- prognosis in variable phenotypes (e.g. 22q deletion syndrome)
- definition of tumour subtypes, allowing more precise management
- classification of complex disorders (e.g. Ehlers-Danlos syndrome)
- targeting of drugs by individual susceptibility (pharmacogenomics)
- targeting and repair of specific gene defects (gene therapy)
- more accurate counselling of inheritance risks and offspring phenotypes.

The enthusiastic claims above may have been somewhat premature, but it is certain that genetics will become an integral component of all medical disciplines. The HGP will have an ever-increasing impact on the elucidation of common multifactorial (polygenic) disorders, in the rapid detection of disease-causing mutations and determination of individual drug targeting. Finally, without precise knowledge of the genetic aetiology of disease, the elusive frontier of routine direct gene therapy will never be breached.

References available on request.

### IN A NUTSHELL

Since all cellular development and function is controlled by our genetic make-up, it follows that our genome (all our DNA) has a fundamental influence on almost all human disease.

The Human Genome Project (HGP) aimed to produce the complete sequence of all the bases in all of human DNA, knowledge which should prove invaluable to the understanding of the pathogenesis and treatment of disease.

However, its output must be placed in context of current and future clinical practice.

Many recent genetic advances have come in spite of, or preceded, the publication of the human genome sequence in 2001.

The exact aetiology of some common genetic disorders has remained frustratingly elusive, despite knowledge of the full sequence of all human DNA.

Merely demonstrating an alteration of the sequence produced by the HGP is insufficient to explain human disease; complex genetic, protein and environmental interactions need to be elucidated to fully grasp the genetic underpinnings of disease.

By elucidating the basic genetic aetiology of disorders, targeted therapies may be designed, based on precise knowledge of abnormal physiology.

The HGP will have an ever-increasing impact on the understanding of common multifactorial disorders, in the rapid detection of disease-causing mutations and in determination of individual drug therapy.