Review

The human genome project and the future of medical practice

Bennett C. Nwanguma

Centre for Biomedical Education, Department of Biochemistry, University of Nigeria, Nsukka, Nigeria. Email: bc_nwanguma@yahoo.co.uk.

Accepted 25 November 2003

Contrary to the scepticism that characterised the planning stages of the human genome project, the technology and sequence data resulting from the project are set to revolutionise medical practice for good. The expected benefits include: enhanced discovery of disease genes, which will lead to improved knowledge on the genetic basis of diseases; availability of DNA-based diagnostic methods, which will find widespread application in preimplantation diagnosis, carrier screening, presymptomatic testing and population screening; the availability of more effective and more tolerant drugs, which will result in more effective therapies characterised by higher potency and reduced incidence of adverse reactions. However, there are still a number of technological, ethical, legal and social obstacles that must be addressed before these medical advances are incorporated into routine clinical practice. To justify the huge investments in the human genome project, the new advancements should be affordable by all, and must not result in a further widening of the gap between the quality of healthcare available in the resource-rich and the resource-poor countries.

Key words. Human genome, genomics, medicine, pharmacogenomics, therapy, diagnosis.

INTRODUCTION

The 20th century was characterised by remarkable improvements in the understanding of the aetiology of diseases, which led to significant advancements in the fields of disease prevention, diagnosis and therapy (Table 1). Although these developments brought about a significant increase in the average life expectancy of human populations all over the world, the much-desired health for all by the year 2000 seemed a fleeting illusion as the twilight of the 20th century approached.

This was mostly because a number of diseases arising from different types of inherited and acquired genetic defects continued to plague a significant proportion of the human population. The best known of these are the thousands of monogenic conditions, e.g., cystic fibrosis (Donaldson and Boucher, 2003), sickle cell anaemia (Weatherall and Provan, 2000) and Huntington’s disease (MacDonald et al., 2003) caused by mutations that affect a single gene. These diseases have been catalogued in a book, ‘Mendelian Inheritance in Man’ (McKusick, 1998) and its online version, the Online Mendelian Inheritance in Man (OMIM), which can be accessed from the NCBI homepage: http://www.ncbi.nlm.nih.gov/omim. The others are the polygenic or multifactorial diseases, e.g., cardiovascular diseases, cancers, hypertension, diabetes, Alzheimer’s disease, Parkinson’s disease, arthritis and obesity, caused by mutations affecting more than one gene and in which the environment also plays a role (Carbonin et al., 2003; Dekker et al., 2003; Florez et al., 2003), the chromosomal abnormalities e.g., Down’s syndrome, caused by deletions, duplications or translocations of sections of the chromosome (Roizen and Patterson, 2003; Andrieux et al., 2002) and the
Table 1. Some of the major medical advances of the 20th century.

<table>
<thead>
<tr>
<th>Therapy</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery of antibiotics</td>
<td>X-ray diffraction</td>
</tr>
<tr>
<td>Development of vaccines</td>
<td>Ultrasound</td>
</tr>
<tr>
<td>Discovery of insulin</td>
<td>Electrocardiogram (ECG)</td>
</tr>
<tr>
<td>Discovery of blood groups</td>
<td>Magnetic Resonance Imaging (MRI)</td>
</tr>
<tr>
<td>Invention of the dialysis technique</td>
<td>Electroencephalogram (EEG)</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>MT Scan</td>
</tr>
<tr>
<td>Organ transplantation</td>
<td>Polymerase Chain Reaction (PCR)</td>
</tr>
<tr>
<td>Use of anaesthesia</td>
<td></td>
</tr>
<tr>
<td>Improved surgery techniques e.g., laser,</td>
<td></td>
</tr>
<tr>
<td>pinhole surgeries</td>
<td></td>
</tr>
</tbody>
</table>

mitochondrial diseases caused by mutations in the non-chromosomal DNA contained in the mitochondrion (Graff et al., 2002; Schapira, 2002).

Although there had never been much doubt that the information needed to find a cure for these diseases laid hidden in the human genome, the scientific community lacked the conviction to embark on the ‘adventure’ of unravelling the estimated 3 billion base pairs that make up the human genome until the last decade of the 20th century. The need to identify these ‘errors’ in genes and to unravel the mechanisms through which they contribute to the different diseases was the overriding justification of the human genome project (Human Genome Project Information, 2003).

With the release of the first draft of the human genome sequence in June 2000 (HGPI, 2000; Lander et al., 2001; Venter et al., 2001), the scepticism that characterised the planning and initial phase of the project gave way to optimism as the emphasis shifted to ways to optimise the benefits of the genome information arising from the project (Chakravarti, 2001; Collins and McKusick, 2001; Futreal et al., 2001; Jimenez-Sanchez et al., 2001). The subsequent publication of the final genome sequence on April 14 2003 has attracted worldwide interest on how medical practice will change in the post genomic era (Burke, 2002; Collins et al., 2003; Guttmacher and Collins, 2002; Guttmacher and Collins, 2003; Khoury et al., 2003; Weinshliboum, 2003). This paper is a contribution to the increasing pile of reviews on how medical practice will change as a result of the completion of the human genome project.

A BRIEF HISTORY OF THE HUMAN GENOME PROJECT

The human genome project (HGP) was initiated in 1990 – after about 5 years of planning - against a background of scepticism (Roberts, 2001, Service, 2001), as a multinational project (involving 20 research institutions located mostly in China, France, Germany, Great Britain, Japan and the United States) to determine the entire sequence and develop the genetic and physical maps of the entire human genome.

The Human Genome Project was designed as a 15-year project (1999 – 2005) with a number of major goals in mind (HGPI, 1990; Collins and Gallas, 1993):

1. To unravel the entire human DNA sequence and to make the data accessible to the public and scientific community;
2. To develop efficient technology to sequence human DNA;
3. To identify the variations in the genetic code of humans that predispose to different diseases, particularly the single nucleotide polymorphisms (SNPs) that are responsible for interindividual differences;
4. To understand how specific genes and/or groups of genes interact exclusively or with some environmental factors to determine predisposition of individuals to disease or good health;
5. To also decipher the genetic code of some model organisms like yeast (C. cerevisiae), roundworm (C. elegans) and the fruit fly (D. melanogaster), which is expected to facilitate the understanding of gene functions in a more complex genome like that of humans;
6. To identify the ethical, legal and social implications of genome research and other obstacles that would be overcome before the results of the HGP can be successfully integrated into medical practice;
7. To develop Bioinformatics tools and manpower to facilitate the gathering, analysis and subsequent storage of the increasing amounts of sequence data that will arise from the project (National Human Genome Research Institute, 2003).

The first draft of the sequence was released in June 2000, while the final draft was released in April 2003 - two years ahead of the projected 2005 date. Other projected targets in terms of technology and sequencing
of the genomes of the model organisms were also met and even exceeded (HGPI, 2003a). The surprising revelations of the genome project include 1), the discovery that there are only about 30,000 – 35,000 genes in the human genome, and 2). The realisation that any two unrelated people are thought to share about 99.9% of their genome, and that the remaining 0.1% is responsible for all the interindividual variations observed in people. Cataloguing these variations known as single nucleotide polymorphisms (SNPs) and matching them with their respective phenotypes has been identified as the next milestone for the human genome project.

Since the completion of the project, scientists have been relatively united in the opinion that human genome project will live up to its promise as ‘the single most important project undertaken by biomedical scientists. There is no doubt that the impact of the genome project on biology, biomedical research, biotechnology and clinical practice will be unprecedented and durable (Collins et al., 2000). Aspects of medical practice being revolutionised by the human genome project include disease discovery, disease diagnosis and prediction, discovery of new drugs and therapies.

THE GENETIC BASIS OF DISEASES

Although the genetic basis of diseases has been known for millennia, genetic diseases were perceived as rare diseases for the most of the 20th century, and the field of genetics was studied as a specialty different from medicine. Recent advances in medical research have, however, continued to reveal that genes interact with the environment in yet unknown ways to affect the severity and predisposition of individuals to many common health conditions, many of which were not thought to have a genetic component (Thomson and Esposito, 1999; Munnich and Feingold, 2003).

The number of diseases known to have a genetic component has increased significantly in recent years, and includes cardiovascular diseases (Aouizerat et al., 2002), diabetes (Florez et al 2003), stroke (Rosand and Altshuler, 2003), cancers (Antoniou and Easton, 2003), Alzheimer’s disease (Ezquerra et al., 2003) Parkinson’s disease (Dekker et al., 2003) arthritis (Svendsen et al., 2003; Olofsson and Holmdahl, 2003), hypertension (Morris et al., 2003) and obesity (Boutin and Froguel, 2003). Many of these are ranked amongst the leading causes of death, especially in the developed countries (Gray, 2000; Guttmacher, 2000). Similarly, genetic variations have been linked with differing susceptibilities to many other diseases, including the infectious diseases, malaria (Omi et al., 2002; 2003) and HIV/AIDS (Michael, 1999), which are the major causes of death in developing countries of Africa and Asia, and the newly identified new Variant Creutzfeldt-jakob disease and related prion diseases (Weissman, 1996; Nwanguma and Orisakwe, 2003). Thus, genetic illnesses are no longer perceived as ‘rare’ disorders, which arise directly as a result of gene mutations and chromosomal aberrations, but include the many common diseases in which genetics plays a role in determining an individual’s susceptibility or the severity of disease suffered by an individual. It is believed in some circles that all diseases have a genetic component, and that some of the so-called monogenic diseases may actually be polygenic because of the contribution of other genes in determining the severity of the disease in individual cases.

DISEASE DISCOVERY AND DIAGNOSES

The fields of disease discovery and diagnostics are two of the many areas in which the human genome project will bring about a major revolution in the 21st century, because the identification and eventual location of all human genes will reveal the mechanism and pattern of inheritance of all monogenic diseases and the many other common diseases in which genetics plays a role in increasing an individual’s resistance or vulnerability. Significant achievements have already been recorded in the area of disease discovery since the commencement of the human genome project. For example, the number of identified human disease genes has increased from less than 100 in 1990, when the human genome project started, to over 1,400 in 2003 (Drell and Adamson, 2003). These include genes associated with a wide variety of diseases including the neurodegenerative diseases - Alzheimer’s disease, multiple sclerosis and Parkinson’s disease (Mathisen, 2003), asthma (Van Eerdewegh, 2002) dyslexia (Francks et al., 2002) cystic fibrosis, diabetes (Walder et al., 2003), Duchenne’s muscular dystrophy, myotonic dystrophy, neurofibromatosis, retinoblastoma, a variety of cancers (Egland et al., 2002; Suzuki et al., 2002), etc. This is because the availability of the human genome data and associated technologies - including position cloning (Nickel et al., 1997; Collins, 1992), sequence based gene discovery and retroviral tagging (Suzuki et al., 2002) - have made the process of gene discovery easier and a lot faster. With these technologies, especially positional cloning, it is now possible to discover the genetic basis of a disease and identify the disease gene without knowing the exact function of the gene (Collins, 1999; Olofsson and Holmdahl, 2003).

Further evidence of the rapid rate of progress in the field of disease discovery is presented by the rate at which the database, Online Mendelian Inheritance in Man available at the website http://www.ncbi.nlm.nih.gov/omim is updated. With time, it is believed, that the mechanisms by which these and several other genes cause diseases independently, or in association with a number of critical environmental factors (known and unknown) will also be uncovered. Already, a new venture known as the
International ‘HapMap’ project – involving scientists at the John Hopkins University and an International Consortium of Scientists from six countries - Canada, China, Japan, Nigeria, United States of America and The United Kingdom – has been initiated since 2002 to facilitate the discovery of the genes that contribute to such complex diseases as cancer, diabetes, heart disease, etc. (NHGRI, 2002)

One direct clinical impact of the reported progress in the discovery of disease genes is the development of gene or DNA-based diagnostic tests for the respective genetic disorders. Already, DNA-based tests or (gene tests), which involve the direct examination of the DNA molecule (Jurkovic et al., 1995), are being used to test for certain genetic disorders (Table 2). Such DNA-based tests are more effective than biochemical methods, which are based on the measurement of gene products (like enzymes) and cytological methods based on the microscopic examination of labelled or stained chromosomes (Wolpert, 2000). In DNA – based tests, a patient’s DNA, which could be obtained from any tissue, is scanned for mutated sequences. For some types of gene tests, probes, whose sequences are complementary to the mutated sequence, are used to detect the presence or absence of a genetic disease. Another type of DNA testing involves comparing the sequence of DNA bases in a patient’s gene to a normal version of the gene. Some clinical applications of gene tests are listed in Table 3.

**Table 2.** A list of some of the diseases for which DNA-based gene tests are currently available.

<table>
<thead>
<tr>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>Ataxia</td>
</tr>
<tr>
<td>Cystic fibrosis</td>
</tr>
<tr>
<td>Gaucher disease</td>
</tr>
<tr>
<td>Duchenne muscular dystrophy</td>
</tr>
<tr>
<td>Fragile X syndrome</td>
</tr>
<tr>
<td>Huntington’s disease</td>
</tr>
<tr>
<td>Inherited breast cancer</td>
</tr>
<tr>
<td>Inherited ovarian cancer</td>
</tr>
<tr>
<td>Myotonic dystrophy</td>
</tr>
<tr>
<td>Phenylketonuria</td>
</tr>
<tr>
<td>Sickle cell anaemia</td>
</tr>
<tr>
<td>Tay-sachs disease</td>
</tr>
<tr>
<td>Thalassemias</td>
</tr>
</tbody>
</table>

In addition to use in routine diagnoses, the newly developed DNA tests will find wide applications in prenatal and postimplantation diagnoses, newborn screening, carrier testing, presymptomatic testing, or for estimating the underlying risk of developing adult diseases like Huntington’s disease and Alzheimer’s disease confirmatory diagnosis of a symptomatic individual and in forensic medicine (Williams and Schutte, 2000; Burke, 2002; Pierce et al., 2003). The list of diseases for which gene tests exist has increased significantly in recent times, and currently includes Alzheimer’s disease, Duchenne muscular dystrophy, Huntington’s disease, sickle cell anaemia, Adult polycystic kidney disease, phenylketonuria, cystic fibrosis, Gaucher’s disease, inherited breast and ovarian cancer, thalassamias and haemophilia (Burke, 2002, Gene Tests, 2003). An advanced form of gene testing, involves the use of the so - called, DNA chips (microarrays) to detect the presence in the DNA of an individual of a wide variety of genetic defects which cause or predispose to a number of diseases (Altman, 2001). In addition to mutation detection, these biochips, like they are called, also offer enormous opportunities for use in biomedical research in the analysis of protein expression, (especially in cancer research) the understanding of gene function and gene discovery (Clark et al., 2000; Miyaoika et al., 2001. In addition, based on the revealed information on the human genome and the genome of the model organisms, including those of mouse and rat, ideal animal models can be created for all human diseases. Such models, which will enhance the understanding of gene functions and disease mechanisms (through comparative genomics), currently exist for a number of diseases, including diabetes, sickle cell anaemia and cancer (Fortini and Bonini, 2000). Such models will also be useful in animal trials for new drugs and therapies meant for genetic disorders (Boelsterli, 2003).

**DRUG DEVELOPMENT AND THERAPY**

It is widely believed that drug therapy often fails to achieve the desired cure, may not be efficacious in a significant proportion of treated patients, and causes severe, and, sometimes, life threatening toxicities in some patients (Lee, 2003). Adverse drug effects, is believed to be responsible for a significant proportion of deaths recorded in hospitals (Lazarou et al 1998., Goldstein, 2003.). Although factors like differences in age, sex, weight and physiological state, concomitant therapy and drug interactions affect drug response in individuals, genetic variability amongst individuals has been recognised as a major determinant of variable drug effect and, therefore, therapy outcome (Pedley and Hirano, 2003). The newly emerging field of Pharmacogenomics (coined from pharmacology and genomic) studies how the entire genetic constitution of an individual determines his body’s response to drugs (Rioux, 2000). Thus, pharmacogenomics deals with genetic variations (polymorphisms) in drug receptors (List and Habener, 2003), transporters (Siddqui et al., 2003),
Table 3. Clinical applications of gene tests.

<table>
<thead>
<tr>
<th>Type of test</th>
<th>Clinical application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier screening</td>
<td>Identification of people carrying a single copy of a gene that requires two copies for disease manifestation. E.g., in sickle cell anaemia.</td>
</tr>
<tr>
<td>Preimplantation genetic diagnosis</td>
<td>Screening embryos for genetic defects before implantation during <em>in vitro</em> fertilisation</td>
</tr>
<tr>
<td>Prenatal diagnosis</td>
<td>Testing the foetus for the presence of genetic defects. E.g., screening for Down’s syndrome and the haemoglobinopathies.</td>
</tr>
<tr>
<td>Newborn screening</td>
<td>Testing of newborns for inborn genetic defects. E.g., phenylketonuria</td>
</tr>
<tr>
<td>Presymptomatic testing</td>
<td>Detection of a predisposing gene in an individual before the manifestation of disease symptoms, e.g., in the detection of adult onset diseases like Alzheimer’s disease and Huntington’s disease.</td>
</tr>
<tr>
<td>Confirmatory diagnosis</td>
<td>The confirmation of the presence of a disease in a symptomatic individual.</td>
</tr>
<tr>
<td>Identity testing (Forensics)</td>
<td>Matching organ donors with recipients during transplantation, confirmation of paternity and family relationships, Identification of accident victims, crime victims and suspects, etc.</td>
</tr>
</tbody>
</table>

other drug targets and the drug-metabolising enzymes, which are responsible for the interindividual variations in rates of drug disposition and response (Evans and Relling, 1999; Evans and McLeod, 2003). All 30 families of drug metabolising enzymes known are believed to show such variations capable of resulting in varying drug response. Following the completion of the human genome project, it would be possible to identify and distinguish individuals who can metabolise a drug rapidly form those who metabolise it slowly or who do not metabolise it at all, and to distinguish those who show an adverse reaction to a drug from those who do not (Lewis and Manning, 1999; Gray, 2000; Turner and Boerwinkle, 2001).

The anticipated benefits of pharmacogenomics will revolutionise drug therapy in a number of ways. In due course, pharmaceutical companies would be able to create more specific, more efficient and less toxic drugs. In addition to being able to prescribe the most suitable drug for a patient without first trying out a variety, doctors would be able to base drug dosages on the patient’s genetic ability to metabolise the drug and no longer on age and weight as is currently the case. Another direct outcome of the human genome project will be in the identification of novel drug targets (Chan et al., 2002; Williams, 2003) and the eventual development of more efficient vaccines, which will be made from DNA or RNA. Nucleic acid vaccines combine all the desirable attributes of existing vaccines without the associated problems, especially the risk of zoonotic infection, which is a major problem with current vaccines derived from animal tissues. In addition, DNA vaccines will be more stable and easier to preserve, especially in the developing countries where vaccine preservation is a source of worry in public health sectors. DNA vaccines could also be engineered to carry multiple strains of a pathogen at once.

Following from these, the process of drug discovery and testing will become less demanding, as target populations for therapy and clinical testing will be chosen more specifically based on their genetic make-up. Ultimately, these changes will bring about an eventual decrease in the overall cost of health care. Pharmacogenomics thus holds the promise that in the future drugs can be “tailor-made” or “designed” to match the genetic constitution of an individual. Such personalised drugs will expectedly be of the desired efficacy and will overcome the problem of adverse reactions some of which are potentially fatal (Human Genome Project Information, 2000c; Sadee, 2002).

Gene therapy - the use of genes for the treatment of diseases - is another type of therapy that would receive a boost from the completion of the human genome Project. Gene therapy, which involves the use of normal genes to supplement or replace a defective gene, or to suppress the expression of an undesirable gene, e.g., tumour genes, is already a rapidly evolving field in therapeutic medicine and holds great hopes for the treating and curing of both genetic and acquired diseases (Noguchi, 2003; Prchal, 2003) In addition to the spectacular
success recorded in the treatment of severe combined immunodeficiency (Engel et al., 2003; Hacein-Bey-Abina et al 2003), significant progress has been made in the development of gene therapy protocols for a number of diseases including cystic fibrosis, HIV infection, type 1 diabetes, a number of haemoglobin disorders, a variety of cancers and cardiovascular diseases (Nettelbeck et al., 2000; Romano et al., 2000., Russell and Peng, 2003., Griesenbach et al., 2003; Moon et al., 2003., Persons, 2003; Zhang and Lui, 2003). It is thought that the efficiency of gene therapy will improve in the near future and the procedure may become routine in the treatment of single gene disorders in less than 2 decades from now (Dzau, 2003; Prchal, 2003).

**ETHICAL, LEGAL AND SOCIAL ISSUES**

Expectedly, the different potential applications of genomics already discussed raise a number of ethical, social and legal issues, which must be addressed before the full potential impact of the human genome project on medicine can be realised. In anticipation of these problems, between 3 and 5% of the annual human genome project budget of the U.S. Department of Energy (DOE) and the National Institutes of Health (NIH) was devoted to studying the ethical, legal, and social issues (ELSI) that may arise from the availability of genetic information as a result of the genome project (HGPI, 2003b). The greatest concern seems to be expressed about genetic testing because of the need to distinguish the ‘significance’ of such tests in the diagnosis of the highly penetrant genetic disorders, where the identification of a mutation serves as a very reliable predictor of disease, and the diagnosis of the multifactorial diseases, where the detection of a mutation can only predict an increased or reduced susceptibility to a particular disorder (Bove et al., 1997; Foster and Sharp, 2000). There are also strong concerns that the availability of information on the entire genetic make up of individuals, including their predisposition to potentially fatal illnesses, for which there are still no cures, based on which predictions of a potentially shortened life expectation can be made, would lead to discriminations in the hands of employees and insurance companies (Burgermeister, 2003; Traynor, 2003).

**CONCLUSION**

Despite the progress already recorded in the different areas of medicine and the potentials discussed, it will still take a while and a lot of research before the full medical benefits of the human genome project will be realised, particularly in the field of pharmacogenomics and gene therapy, where progress has been rather slow (Dunham, 2000; Romano et al., 2000). It is estimated that many of the new DNA-based diagnostic tests will become common, and gene therapy for monogenic genetic disorders will become routine clinical practices by 2020 (Drell and Adamson, 2003).

There are also genuine fears that the cost of clinical care in the post genomic era would rise above what the average person can afford. If this is not addressed, only the rich can afford the new therapies. At an estimated cost of $1000 USD for the sequencing of the complete genome of an individual and the likely claim of patent rights for some of the gene tests, this is likely to be the case. The implication of this is that the already worrying gap between standard of health care in the developed and developing countries will widen even further in the post genomic era.

**REFERENCES**


