Institute of Science in Society

Science
Society
Sustainability

Relevant Links:
- The Human Genome--A Big White Elephant
- Where Genes Fail--Dietary Interventions for Alzheimer’s and Parkinson’s?
- Chief Scientist Bob May Lambastes Human Genetics Panel
- The Human Genome Map, the Death of Genetic Determinism and Beyond
- The Unnecessary Evil of "Therapeutic" Human Cloning

ISIS-TWN Report

Human Genome - The Biggest Sellout in Human History

Mae-Wan Ho
Institute of Science in Society and Dept. of Biological Sciences
Open University, Walton Hall, Milton Keynes MK7 6AA, UK

18 October 2000

Spanish Translation of this Report

- Abstract
- The human genome sellout
- The ‘bioinformatics’ gold rush
- The reality test
- Will it deliver?
- References
Abstract

Governments in the industrialized countries have handed over the human genome to private ownership together with the most triumphant hyperboles to boot, notwithstanding that it was mapped and sequenced at great public expense. A multi-billion bio-informatics 'goldrush' is on, as private companies scramble to mine the public database for genes to patent and to assemble their own proprietary databases which are sold at exorbitant fees to subscribers. Beneath the hype, bio-informatics is a desperate attempt to turn the exponentially increasing amount of information into knowledge. The human genome programme has dominated the scientific scene for the past ten years, raising hopes and fears in equal measure. Is it likely to deliver? No, especially if it continues to be misguided by a discredited genetic determinist paradigm that serves to divert attention and resources away from the real causes of ill-health and to stigmatize the victims. We are already witnessing the resurgence of genetic discrimination and eugenics that have blighted the history of much of the last century.

Bio-informatics suffers from the reductionist fallacy that knowledge will automatically arise once information is exhaustively listed. Molecular biology is suffocating from information overload. What we need is a quantum leap to a new paradigm for understanding the organism as a coherent whole. Otherwise, human genome research will remain a scientific and financial black hole that swallows up all public and private resources without any return either to investors or to improving the health of nations.

Key words: gene-patents, bio-informatics, proteonomics, gene-tests, gene-chip, susceptibility to disease, eugenics, genetic discrimination, genetic determinism, reductionist fallacy

The human genome sellout

"To-day, we are learning the language that allowed God to create life." That was how Clinton greeted the announcement of the human genome map on June 26 [1]. The Human Genome Project, (HGP) an international public consortium of research laboratories led by the United States, and Celera, a private American company, made the announcement jointly, ending months of competition to complete the first sequence of the human genome. Craig Venter, Director of Celera, referred to this "historical day in the 100,000 years of human history" when, for the first time, "the human species can read the letters of its own text." Not to be outdone, Francis Collins, head of the public project, called it "the revelation of the book of life".

Craig Venter claimed his discoveries could definitively cure cancer, thus securing Celera’s place in the private investment market, while Francis Collins stressed that "the real work is starting", thereby justifying the next round of major public finance.

French Research Minister, Roger-Gérard Schwartzenberg, hailed the event as " the victory of those who wanted knowledge to remain free" [2]. In reality, it is the biggest sellout in human history dressed up with the most far-flung hyperboles.

The human genome has been sequenced separately and independently with major public
finance, from the United States and the European Community. The US Government alone had earmarked $3 billion for the initiative. But that has not prevented the human genome from being owned and exploited by private companies. Earlier in March, Clinton and Blair released an ambiguous statement calling for open access to the human genome data. It sent biotech stocks on a downward slide, with some dropping 20% at the end of the day. In the weeks following, officials in the Clinton administration clarified that they still favor patents on "new gene-based health care products."

Celera’s genetic maps would eventually be available on the Internet, and the company will claim royalties from any commercial pharmaceutical application of its discoveries. In contrast, the gene sequences and gene maps produced by the public consortium have been deposited regularly within 24 hours of completion in GenBank, a public database set up in the early 1980s when DNA sequencing began, access to which is totally free. Celera kept its own human genome data secret while benefiting from free access to the public database throughout the period that the company was busy sequencing, thereby significantly reducing the time and effort needed to complete the task.

Celera is not the only company stealing from HGP’s Genbank[^3]. Others such as Incyte has mined the public data to help build its catalogue of genes and patents. At least 500 gene patents have already been awarded, while another 7000 have been applied for. Human Genome Sciences has won more than 100 gene patents and filed for roughly another 7000. There are some 20 000 patents on gene sequences pending at the US patent office[^4].

The US and European Governments, in line with the private companies, are downplaying the free access to the public human genome database on grounds that raw genome sequence is useless. During the quarterly meeting of Research Ministers of the G8 in Bordeaux at the end of June, to which Mexico, Brazil, China and India were invited, all agreed that DNA sequences -- the fundamental data -- must not be patented, in recognition that they are discoveries and not inventions[^5]. This seems like a definite improvement over the previous situation in the United States where over 4 million patents on human genome sequences have already been granted[^6], the majority of which are on short fragments of DNA with no known function.

The US Patent and Trademark Office (USPTO) had tightened up the criteria for gene patents by issuing two new directives under section 101 "utility", and section 112 "written description requirements" last December[^7]. Under the new utility guidelines, the USPTO is looking for "specific utility" and "substantial utility". So, DNA fragments or express sequence tags (EST) will require a written description of their specific utility in order to be patented (though millions have been awarded patents already). Similarly, according to the current EU Directive on biotechnological inventions, genes and gene-sequences can still be patented if an "industrial application" is specified.

However, an "industrial application" may amount to no more than speculation based on similarity to gene sequences in the existing database. A notorious case involves the CCR5 gene patent awarded to Human Genome Sciences in the US this February. The company isolated the gene using automated computers to sequence it and software to determine that it belonged to a class of cell membrane receptors that pick up chemical signals in the body[^8]. A few months later, scientists at the Aaron Diamond AIDS Research Center in New York...
discovered that the AIDs virus requires the receptor to enter cells. A drug that can block the receptor would thus be a new weapon against AIDS.

Another industrial application for which many patents have been awarded is "association with condition X", where X is anything from cancer to criminality. There are already 740 patented gene tests on the market, among them BRCA1 and BRCA2, genes linked to breast cancer in women. Years after the tests were launched, scientists still do not know to what degree those genes contribute to a woman’s cancer risk \[3\]. But it is precisely this ignorance that is fueling the human genome goldrush in ‘bioinformatics’.

**The ‘bioinformatics’ gold rush**

The public GenBank holds sequence data on more than seven billion units of DNA, while Celera Genomics claims to have 50 terabytes of data in store, equivalent to 80 000 compact discs. The raw sequence data consist of monotonous strings of four letters -- A, T, C and G -that make up the 3 billion or so bases in the human genome. It is impossible to access the data or to make any sense of the sequences without special software. Some software are developed and made freely available in the public domain, but the databases of private companies are provided to paid-up subscribers only. Incyte launched an e-commerce genomics program in March that allows researchers to order sequence data or physical copies of more than 100 000 genes on-line. Subscribers to the company’s genomics database include drug giants such as Pfizer, Bayer and Eli Lilly. Celera’s gene notes, similarly, will cost commercial subscribers an estimated $5 to $15 million, and academics, $2000 to $15000 a year.

This first wave of the human genome goldrush, ‘bioinformatics’, is a fusion of information technology with biology \[9\] that promises to turn the raw genomic base-sequence data into knowledge for making even more lucrative new drugs. Bioinformatics is already a $300 million industry expected to grow to $2 billion within 5 years.

One of the most basic operations in bio-informatics is searching for similarity or homology between a new sequence and one in the database, which allows researchers to predict the type of protein encoded and its function, thus enabling the sequence to be patented. However, sequence homology is no guarantee of homology in function, as we have seen.

With the understanding of protein structure, it is possible to conduct searches for specific inhibitors and activators before carrying out actual biochemical experiments in the laboratory. Only 1% of proteins so far has had their structures determined (by X-ray crystallography).

Some bioinformatics companies cater to large users, aiming their products and services at genomics, biotechnology and pharmaceutical companies by creating custom software and offering consulting services. Lion Bioscience, in Heidelberg Germany, has a $100-million contract with Bayer to build and manage a bioinformatics capability across all of Bayer’s divisions. Other firms target small or academic users. Web businesses such as Oakland, California-based Double Twist, and e-Bioinformatics in Pleasanton, California, offer
one-stop internet shopping. These on-line companies allow users to access various types of databases and use software to manipulate the data. Large pharmaceutical companies have established entire departments to integrate and service computer software and facilitate database access across departments.

Close on the heels of bio-informatics, and possibly part of bio-informatics, is ‘proteomics’. Its focus is on when and where genes are active and on the properties of the proteins the genes encode. It attempts to make sense of the complex relationships between gene and protein and between different proteins [10], and has so far also attracted hundreds of millions in venture capital.

According to Mark J. Levin, CEO of Millennium Pharmaceuticals in Cambridge, Mass., large pharmaceutical companies need to identify between 3 and 5 new drug candidates a year in order to grow 10 to 20 percent -- the minimum increase shareholders will tolerate. Right now, they are only delivering a half to one a half a year. Millennium has a relationship with Bayer to deliver 225 pretested "druggable" targets within a few years. Celera is in negotiations with GeneBio, a commercial adjunct of Swiss Institute for Bioinformatics in Geneva to launch a company dedicated to deducing the entire human proteome. As the number of human genes could be as high as 100 000, it is estimated that the number of proteins could well be in the region of 1 million. Up to the mid 1970s, scientists had assumed, wrongly, that one gene codes for one protein. Instead, the relationship between genes and proteins are complicated by many layers of processing and editing starting before the genes are even transcribed [11].

Proteomics has spawned a number of technical innovations, among which is the Gene Chip, developed by Affy-metrix in Santa Clara, California. It consists of glass microarrays coated with cDNAs (complementary DNA) to identify which mRNA species are made (and hence which genes are expressed). One microarray allows researchers to identify more than 60 000 different human mRNAs. The US National Cancer Institute has been examining the mRNAs produced by various types of cancer cells in a Human Tumor Gene Index project involving government and academic laboratories as well as a group of drug companies including Bristol-Myers Squibb, Genetech, Glaxo Wellcome and Merck. So far, more than 50 000 genes have been identified that are active in one or more cancers.

The reality test

The sequencing of the human genome is undeniably a technical feat comparable perhaps to landing on the moon. And it is difficult not to be caught up in a frenzy of speculation on what can be achieved as genomics joins forces with the latest in information and nanotechnology.

According to John Bell at Oxford, within the next decade, ‘predictive’ gene testing will be widely used both in healthy people and for diagnosis and management of patients. Francis Collins, Director of the National Human Genome Research Institute in the US, has stated that the benefits of human genome mapping would include "a new understanding of genetic contributions to human disease" and "the development of rational strategies for minimizing
or preventing disease phenotypes altogether." [12].

Will predictive gene tests kill the insurance industry? That was one worrying aspect considered [13]. Apparently, during an industry conference held in Boston, senior executives from several of the world’s leading genomics concerns agreed that genomics, with its promise of being able to show who will be predisposed to what disease, would eventually give rise to universal healthcare in the United States. "This could happen especially if the defects in our genomes make us all uninsurable," said panelist Craig Venter.

"The good news about genomics is that we could soon be able to catch deadly diseases in their earliest stages, when many are still treatable and even curable. And genomics also holds the promise of being able to deliver a bold new generation of drugs that will work more effectively with our individual genetic quirks. The bad news is that everyone will learn they are a walking time bomb, in one way, shape or form.". But how reliable are gene tests in predicting what will happen to the individual?

Two medical geneticist writing in the New England Journal of Medicine [12], warned that the ‘genetic mantle’ currently put onto all diseases "may prove to be like the emperor’s new clothes."

As has been pointed out by many scientists, most diseases are complex, and correlations between genes and disease are therefore weak. Associations between a disease and a ‘genetic marker’ (of unknown function) can occur by chance and some have proved to be spurious. Although many disease-related genes have been mapped to regions of specific chromosomes, no clear markers for asthma, hypertension, schizophrenia, bipolar disorder, and other disorders have been found despite intensive efforts.

Searches for susceptibility genes in breast cancer, colon cancer, rare early-onset forms of type II diabetes, and Alzheimer’s disease have been more successful, but in each case these account for less than 3 percent of all cases. That is because the risk of disease depends not only on other genes but also on environmental factors. The problem of identifying susceptibility genes is compounded when different combinations of genes are implicated in a disease, for it means that finding enough patients to serve as research subjects in a study will be extremely difficult.

Holtzman and Marteau conclude, "In our rush to fit medicine with the genetic mantle, we are losing sight of other possibilities for improving the public health. Differences in social structure, lifestyle, and environment account for much larger proportions of disease. . . . Those who make medical and science policies in the next decade would do well to see beyond the hype."

Let us take stock of some of what is on offer. The human genome sequence, we are told, will enable geneticists to

- cure cancer
- understand more about diseases and thereby to design better drugs
- design customized cures based on our individual genetic makeup
- prescribe an individual’s lifestyle based on genetic makeup.
More contentious are the claims to

- diagnose all the bad genes that cause diseases
- identify all the good genes responsible for desirable qualities such as longevity, intelligence, being slim and beautiful, good at sports, and so on
- replace bad genes in ‘gene therapy’, including germline gene therapy
- create ‘genetic enhancement’ by introducing ‘good’ genes
- create ‘designer babies’ and superior human beings.

In reality, the only concrete offering from mapping the human genome are the hundreds of patented gene tests. The high costs of the tests have prevented them from being used in cases where it might benefit patients in providing diagnosis [14]. At the same time, those healthy subjects who have tested positive are likely to suffer from genetic discrimination and risk losing employment and health insurance. The value of diagnosis for conditions for which there is no cure is highly questionable. The claim to identify putative ‘bad’ and ‘good’ genes is also fueling the return of eugenics, which has blighted the history of much of the 20th century. This is exacerbated by the dominant genetic determinist mindset that makes even the most pernicious applications of gene technology seem compelling.

A prominent band of scientists and ‘bioethicists’ are actively advocating human genetic engineering, not just in ‘gene therapy’ for genetic disease, but in positively enhancing and improving the genetic makeup of children of parents who can pay for the privilege, and have no qualms regarding human reproductive cloning either [15]. In many ways, this is the most subtle form of hype for business to prosper. It is no accident, therefore, that the Novartis Foundation has invited arch-eugenicist Arthur Jensen, to speak at a scientific meeting on intelligence [16]. Jensen is best known for his insistence that black people are genetically inferior in intelligence to white people, and hence all efforts at enhancing the education of disadvantaged black children are bound to fail.

It is clear that the promises as well as the threats remain largely in the realm of future potential if not outright fantasy. We were promised no less than “the blueprint for making a human being” by no less than Nobel laureate James Watson when the Human Genome Project was first touted, along with miracle cures for cancer and other diseases, and even immortality. Now, ten years and dozens of sequenced genomes later, it is all too obvious that geneticists haven’t got a clue of how to make even the smallest bacterium, or the simplest worm, let alone a human being. Nor has anyone been cured of a single disease on the basis of genes or genetic information.

Despite the proliferation of genetic tests, many of them are uninformative because the association between the genes and the diseases is tenuous in the first place. And even the most informative tests -- those associated with so-called single gene conditions -- cannot predict the age of onset or the severity of the disease, as pointed out by Wendy R. Uhlmann, president of the National Society of Genetic Counselors [3]. Indeed, an air of realism, if not disillusionment, pervades the scientific community in the public sector.

"For a long time, there was a big misconception that when the DNA sequence was done, we’d have total enlightenment about who we are, why we get sick and why we get old. . . . Well, total enlightenment is decades away." This remark is attributed to geneticist Richard K
Wilson of Washington University, one partner in the public consortium [17]. He should have said that the misconception has been perpetrated by the proponents of the HGP themselves. Still, he is promising "total enlightenment" in a matter of decades. But will the human genome project really deliver?

Rather than address the contentious claims of the human genome project, I want to concentrate on those offerings that are largely seen to be beneficial and uncontroversial; for if it cannot deliver on those, it can certainly not deliver on the rest.

Will it deliver?

The growth in ‘bioinformatics’ and ‘proteomics’ is an admission of the vast realms of ignorance that separate the 100 000 genes in the human genome from the living human being. It is also an acknowledgement that the genetic determinist paradigm, which has done so much to promote the human genome project, has failed miserably. There is no simple, linear causal chain connecting a gene to a trait, good or bad. Behind the hype is a desperate attempt to turn the exponentially increasing amount of information into knowledge that can pay off the heavy investments already sunk into the project.

Private ownership of the human genome is obviously not ever going to benefit those who cannot afford to pay. Proponents of human genetic engineering, indeed, see the creation of a ‘genetic underclass’ to be inevitable, as those who can afford to pay for genetic enhancement will become ‘gene rich’ relative to those who cannot afford to pay [15]. But can knowledge of the human genome really deliver the goods?

The fallacy of genetic determinism is widely recognized [18]. Genuine genetic diseases that can be attributed to single genes constitute less than 2% of all diseases. And more and more geneticists are coming around to the view that even those are subject to so many other genetic and environmental influences that there is simply no such thing as a single-gene condition. For the rest, the association between the condition and the specific genes or genetic markers reduces to tenuous ‘predispositions’ or ‘susceptibility’ (see above).

‘Predispositions’ to cancer for example, conceals the fact that important environmental factors are left out of consideration. These include the hundreds of acknowledged industrial carcinogens polluting our environment. It is well-known that the incidence of cancer increases with industrialization and with the use of pesticides. Women in non-industrialized Asian countries have a much lower incidence of breast cancer than the women living in the industrialized west. However, when Asian women emigrate to Europe and the United States, their incidence of cancer jumps to that of the white European women within a single generation. Similarly, when DDT and other pesticides were phased out in Israel, breast cancer mortality in pre-menopausal women dropped by 30%. The overwhelming causes of ill-health are environmental and social. That is the conclusion of a growing body of research findings. Environmental influences swamp even large genetic differences.

The genetic determinist approach of the human genome programme is pernicious because it diverts attention and resources away from addressing the real causes of ill-health, while at
the same time stigmatizing the victims and fueling eugenic tendencies in society. The health of nations will be infinitely better served by devoting resources to preventing environmental pollution and to phasing out agrochemicals, rather than by identifying all the genes that ‘predispose’ people to ill-health. The UK Royal Society produced a report in July, calling for national and international coordination to deal with the dangers posed to humans and wildlife by endocrine-disrupting chemicals, substances thought to mimic or block natural hormones in amounts too minute to trigger a conventional toxic response [19].

But it is the inherent complexity of the human organism and the lack of a concept of the organism as a coherent whole that will continue to frustrate all attempts at understanding health and disease within the dominant, reductionist framework.

Despite the almost weekly hype on cancer cures, there is none, or none that has resulted from information on genes and gene sequences. As mentioned earlier, some 50,000 genes have been identified that are active in one or more cancers using the Gene Chip, which is half of the maximum number of gene predicted in the human genome!

In principle, knowing the genes that are over-expressed or inactive in individual cancers can allow specific genes to be targeted. But this is no different from interventions that have previously been available to single-gene defects such as sickle cell anaemia or cystic fibrosis, none of which has been cured as a result; which is why gene therapy has been attempted, equally to no avail so far. One obstacle to effective cure is that it is impossible to avoid unintended ‘side-effects’ in a system where proteins interact with one another and with the genes. But the main problem is the failure to recognize that just as health is a property of the organism as whole, so too is disease.

To try to understand disease in terms of genes and protein interactions is worse than trying to understand how a machine works in terms of its nuts and bolts, simply because the parts of the organism, unlike those of a machine, are inseparably tangled up with one another. Mechanistic understanding in terms of interacting parts is extremely unlikely to lead to the design of better drugs. For that, we require knowledge of the design of the human organism. And no amount of information on genes and protein interactions will ever add up to the complex, entangled whole that is the organism.

The promise of customized medicine and prescribed lifestyle based on an individual’s genetic makeup is a pipe-dream. The effect of each gene depends not only on external environmental factors, but on the genetic back-ground of all other genes in the genome. Individuals differ on average by one base per thousand in their DNA. This amounts to three million bases over the entire genome. As each gene is at least a thousand bases in length, it means that every gene will most probably be different. Assuming that only two variants exist in each gene, the number of different genotypes is already 3 (100,000). In fact, hundreds of variants are typically found for each gene. Consequently, every individual is genetically unique, except for identical twins at the beginning of development, before different genetic mutations can accumulate in each of the pair. That is why it is generally impossible to give accurate prognosis of even single gene diseases unless the genetic background is homogenous, as in an inbred laboratory strain of mice. And even then, the mice have to be raised in a uniform environment.
The population in Iceland is thought to approach genetic homogeneity, which is why the company deCode Genetics has acquired the genetic database of Iceland’s 270,000 inhabitants, linked, anonymously to medical records. The hope is to enable all the genes linked to a variety of diseases to be identified. Unfortunately, the results will be valid for the Icelandic population only, and will not be transferable to other populations. Thus, mutations in the gene giving rise to cystic fibrosis among Northern Europeans is associated with quite another condition among the Yemans; while conditions diagnosed as *bona fide* cystic fibrosis in the latter population is associated with mutations in another gene altogether [18]. Some geneticists, indeed, are beginning to think that better data for linkage to diseases might be found in genetically heterogeneous populations, such as those in Manhattan and London, rather than in homogeneous populations, such as those in Iceland and Finlands [20].

In classical genetic analysis, the net effects of a gene are determined over all environments as well as over all genetic backgroundss [21], in recognition that both environmental and genetic interactions have to be taken into account. So, the most reliable data are those obtained in large populations which are as heterogeneous as possible genetically as well as environmentally. But the predictive power of such genetic data is always limited to *population averages*. It is impossible, in principle, to predict anything based on any individual genome. Those who claim otherwise are ignorant of the most basic principles of population genetics.

In case you still think that the blueprint for making a human being is written in our genome, just take note that up to 95% of the human genome may be ‘junk’ DNA, so called because no one knows what its function is. The same is true of all genomes of higher organisms. The rough draft of the human gene map announced in June is only 85% complete for the coding (functional) regions only.

It is difficult to see any definite strategy within either bioinformatics or proteomics that can pay off, either in terms of basic understanding the human organism as a whole, or in terms of miracle cures and wonder drugs. There is nothing beyond the proliferation of more and more detailed information on genes and proteins that have been spilling out of the pages of scientific journals for the past decade. The one million proteins encoded by the 100,000 genes interact with one another, with the genes themselves, and small molecular weight ‘cofactors’ and ‘messengers’. Those interactions vary in different cells and tissues at different times, subject to feedback from the environment. Feedback from the environment can alter the genes themselves, and hence the entire cascades of interactions involved. All that is the reality of the fluid and adaptable genomes [11], which the moguls of genomics and bioinformatics have yet to come to grips with. The prospect of understanding the human being by a detailed description of its molecular parts is essentially nil. This reductionist fallacy has been exposed in different forms, starting with the physicist Walter Elsassers [22].

Elsasser pointed out that there is no unique correspondence between the states of the molecules within the system and the macroscopic condition of the organism, say, whether it is ill or well, simply because there are infinitely more molecular states than macroscopic states. Hence, a detailed description of all of the molecules, even if it were possible, will not enable one to determine the condition of the system as a whole.

If we were to define the state of the human organism in terms of its 100,000 genes simply as
to whether each gene is active or not in each of its 70 trillion cells, the total number of possible states for each cell is 2 (100 000). And that does not include the proteins, nor the interactions among genes, proteins and cofactors. We need a computer large enough to represent the states of all the molecules and their interactions in each cell, and fast enough to give a description of how they change in real time as the entire organism goes about its business of living. But even then, we would still be left with no understanding of what is being described. Current computation is unable to handle the dynamics of one single protein folding, even given all the information on the amino-acid sequence and the final shape of the folded protein. It takes the computer four hours to find a solution that is at best 70% accurate; whereas the protein itself folds to perfection within a fraction of a seconds [23].

What we need is a quantum leap to a new paradigm for understanding the organism as a coherent whole [24]. Without that, human genome research will remain a scientific and financial black hole that swallows up all public and private resources without any return either to investors or to improving the health of nations.

References

2. Interview by Corinne Bensimon. Biotech Activists July 20 posted by Genetics@gn.apc.org.
5. "The Minister of Research Rejects Patents onDNA Sequences. No One Can Own a Gene" Interview by Corinne Bensimon, Biotech Activists July 20 posted by Genetics@gn.apc.org.
7. "Stricter criteria for patents may lead to many rejections" ISIS News#4, March/April, 2000 <www.i-sis.org.uk>.


The human genome sellout. "To-day, we are learning the language that allowed God to create life." That was how Clinton greeted the announcement of the human genome map on June 26 (1). The Human Genome Project, (HGP) an international public consortium of research laboratories led by the United States, and Celera, a private American company, made the announcement jointly, ending months of competition to complete the first sequence of the human genome. The human genome has been sequenced separately and independently with major public finance, from the United States and the European Community. The US Government alone had earmarked $3 billion for the initiative. The human genome is a complete set of nucleic acid sequences for humans, encoded as DNA within the 23 chromosome pairs in cell nuclei and in a small DNA molecule found within individual mitochondria. These are usually treated separately as the nuclear genome, and the mitochondrial genome. Human genomes include both protein-coding DNA genes and noncoding DNA. Haploid human genomes, which are contained in germ cells (the egg and sperm gamete cells created in the meiosis phase of sexual reproduction The Human Genome Project revealed how the 3.2 billion base pairs contained in the human genome are ordered. Learn more about the Human Genome Project. Scientists have long known that these four letters provide the recipes for proteins, which carry out numerous bodily functions. But there are still questions to be answered, including how the 3.2 billion base pairs contained in the human genome are ordered. (The human genome is a person's entire bundle of DNA divided unevenly among 23 pairs of chromosomes.) To that end, the Human Genome Project (HGP) was launched in 1990. Some of the project's ambitious goals included: Advertisement. Advertisement. Sequencing the entire human genome. Identifying human genes. Charting variations across