FACTS & FINDINGS
The Journal for Legal Assistants

August 2000 Vol XXVII Issue 2

The Human Genome Project: What a Legal Assistant Needs to Know

by Daniel Drell, Ph.D.
Life Sciences Division, Office of Biological and Environmental Research, U.S. Department of Energy

As you read this, the first phase of the Human Genome Project (HGP) will have just about been finished, its first major milestone achieved. A nearly complete, "draft" version of the entire human genome, all of the chromosomes containing the 3 billion base pairs of DNA, will have been determined and deposited in publicly accessible databases for anyone to view and use. Scientists will know almost all of the human DNA sequence, containing the essential information for all the "parts" that make up all of the approximately 1 trillion cells in an adult human body and carry out all the functions inside each of those cells.

Biologists also will know all the basic genetic building blocks, perhaps some 120,000 of them (referred to as "genes", each of which determines a protein, e.g. a biological "part") and all will be available for study and research. Computational biologists will have identified perhaps 30% of those genetic "parts" and at least tentatively "annotated" them, that is, assigned a reasonably certain biological function and purpose. Technologies for DNA sequencing will have progressed to the state that a highly organized and well equipped sequencing center can determine on the order of 30,000,000 base pairs of DNA sequence in a day, equivalent (if printed) to a 1600 page phone book. With this complete "parts list" in hand, and using modern DNA manipulation and analysis technologies, a scientist in her or his lab can find the location (and thus any gene) in the vast human genome and amplify, study, and characterize it.

The promise from this ambitious effort is huge and the outcomes will dramatically affect the economy of the United States. Along with the economy, DNA technologies and information are already and will continue to impact the judicial system, creating many challenges for judges, lawyers, juries, and legal assistants. These challenges will involve simply understanding what the new technologies and information mean, as well as anticipating some of the ways these technologies and information may find their ways into legal proceedings.

The Human Genome Project: What it is, Where it Came From, Where it is Going.

Let's start with some basic definitions. Every human cell (except for red blood cells and the platelets that are critical to normal blood clotting and wound healing) contains a discrete nucleus that has within the roughly 6 feet of a special chemical called deoxyribonucleic acid, or DNA. This molecule exists as a very long polymer, twisted in the familiar double helix discovered by James Watson and Francis Crick in 1953. DNA consists of a backbone of repeating sugar and phosphate units, off of each one of which sticks a simple chemical structure called a nucleotide (more commonly, a "base.")

There are exactly four kinds of bases that are found in DNA and, significantly, all DNA molecules in all life forms on Earth use the same 4 bases (abbreviated A for adenine, C for cytosine, G for guanine and T for thymine) and the same coding scheme to store genetic information. The key fact of DNA is that it stores information in the exact sequence of these bases, which can number in the millions along a single

DNA molecule. An analogy might be helpful here: imagine a cassette tape with some prerecorded music on it. The cassette tape is a passive storage device until it is inserted into a cassette player and played, at which time the information on it becomes expressed.

DNA is similar in that complex cellular machinery exists to express, in carefully controlled ways, the information that is on a DNA molecule. This is crucial because each and every nucleus-containing cell in an individual's body has the same full complement of DNA; in a liver cell, only a carefully regulated subset of that DNA is used while, in a nerve cell, a different subset is utilized. The aim of the Human Genome Project (HGP) is nothing less than to determine the entire sequence of a reference 3 billion DNA bases (the amount that is in a single cell's nucleus), called the genome. [One clarification needs to be made here: the genome of a typical human nucleated cell contains 6 billion base pairs of DNA. This is because at conception, 3 billion bases of DNA in an unfertilized egg are joined with another 3 billion bases of DNA from the single successful male sperm. On average, scientists think that the two sets of DNA differ by about one DNA base in every one thousand, differences that can be explored after a reference sequence has been determined.)

Since the research of Herman Muller in the 1920's, it had been known that radiation can lead to heritable effects, including cancers. Once the structure of DNA was established by Watson and Crick, it was immediately obvious not only that its exact sequence was the way it stored information but also that changes to the sequence (such as might result from exposure to radiation or DNA-damaging chemicals) could lead to altered biology including deleterious health effects. To understand what radiation did to the DNA sequence, and how it could cause heritable changes, scientists would have to sequence the DNA in its entirety. (Returning to the analogy of the prerecorded cassette tape, imagine passing a refrigerator magnet near the cassette tape. The magnet could easily alter the information that is recorded on the tape, without damaging the actual tape itself. To determine what the magnet may have done to the tape's content, you would need to play it. More fundamentally, in order to learn if the effects of the magnet mattered, you would need to know what had been on the tape before the magnet came near it. To complete the analogy, not all of the information on the tape might be of equivalent "value"; if the magnet damaged a soft passage in a piano sonata by Beethoven, the damage would be severe but if it created static in the middle of the transition from one movement to another, or in the middle of a song by an incompetent garage band, it would be a minor annoyance or even unnoticed.)

As the steward of nuclear weapons production work, the Atomic Energy Commission, (predecessor to the present Department of Energy, DOE) was charged with conducting research on the health effects of radiation. Of particular concern was the fear that survivors of radiation exposure following a nuclear attack (or soldiers exposed as part of post nuclear combat operations) might pass on cancer-causing mutations to their children and so technologies for mutation detection were a high research priority. However, this was a very difficult challenge until the molecular biology revolution at the end of the 1970's. The HGP was begun in 1986 by scientists in the U.S. DOE as a way to explore newly developing DNA analytical technologies that might better assess mutations from radiation. Its network of National Laboratories, its experience with large projects (e.g. particle accelerators), and the availability of both interdisciplinary teams of scientists and powerful computational facilities, all made DOE a logical, if unexpected, agency to begin a massive effort to sequence the human genome. By 1988, the National Institutes of Health had joined the project and a joint effort was announced in 1990.

The HGP is considerably more than just determining a reference DNA sequence for a human. 3 billion As, Cs, Gs, and Ts don't reveal very much without the ability to discern the bits that code for proteins (the "parts" of a cell, or the "songs," referring once again to the cassette tape analogy) and then to determine what those parts actually do. It is an intriguing and presently mysterious fact that most of the DNA in the human genome does not code for "parts"; the functions of this DNA, which may exceed 95% of the total, is unknown at this time.

Figuring out the roles, in health and in disease, of all these parts, as well as the additional ("noncoding") DNA will be a huge challenge, and one that will last well into the 21st Century. One clear portent for

biology is that it must now become an information science and this will be a difficult transition for a community that has hitherto been very lab oriented and determinedly reductionist in its approach to learning about what constitutes cells and tissues. Biology must now, under the stimulus of the HGP, learn how to become a "reconstructionist" science where putting the parts together to make more complex subsystems, pathways, and operating systems is the goal.

Ethical, Legal, and Social Implications ("ELSI")

Each of us is a unique individual. Each of us has a genome, consisting of 3 billion DNA bases inherited from our mother, plus an additional (very similar) 3 billion inherited from our father. As noted earlier, each of us differs in DNA sequence by about 1 base in every thousand, for a total of between 3 and 6 million differences. These differences underlie our uniqueness and individuality. However, it is important to make the point right now that while these differences account for much of what makes us individuals, there are other contributing factors as well, many of them non-genetic (e.g. environmental). Numerous studies of identical twins have shown that identical twins, genetically identical from conception, are not identical for various traits and diseases. They are often observed to be more similar than either non-identical siblings or random individuals, but still very often non-concordant. The importance of this observation is that while we can expect to learn much from studying a person's genome, there is a limit to what it can tell us even when we "know" it in its entirety (what better genetic test can there be than to have an identical twin to observe?) So it is necessary to be extremely cautious and alert to "genetic determinism", the trap of assuming that more is due to genetic inheritance than in fact is.

Before discussing potential legal issues arising from the HGP, it is important to set an economic context since it is likely that numerous legal conflicts will arise from economic disputes. The US biotechnology sector has been estimated to contribute upwards of $14 billion annually and indisputably leads the rest of the world by a large margin. Although the production of a new drug is expensive (estimates vary widely, from $300 million to as high as $800 million), the HGP will generate thousands of new candidates. Many of these will take years to develop (with the attendant concerns about intellectual property protection, filing dates, etc.) We will see new genome- based drugs, precision characterization of individual variations (the genetic differences between one person and another) and their links (of lack thereof) with conditions, traits, and diseases. We will see new insights into the functioning of human cells under a wide range of circumstances, developmental states, and external challenges.

Genetic analyses of individual patients will lead to far more precise categorization of patients, more accurate diagnoses, and more precise understanding of underlying causes for disease. This will result in far more precisely targeted (and safer) therapies, tailored to an individual's condition and avoiding the nonspecific approach to treatment accompanied by all the risks of unexpected side effects. This will permit physicians to prescribe the right drug in the right dosages to the right people and can be expected to revolutionize medicine. Eventually, we will perfect technologies to replace a malfunctioning gene with a properly functioning version and gene therapy will emerge as a powerful tool in the armamentarium of the physician. While it won't happen overnight, and some investigations will take many years, we will eventually understand at the genetic level the origins of the most challenging of conditions and diseases, the complex ones, that are influenced by more than one gene as well as by the influences of the environment. The future for biological science and medicine looks very bright. However, difficult ethical, legal and social challenges will arise as well and the courts and allied personnel (e.g. lawyers and legal assistants) will be heavily involved.

At the outset of the HGP, several people recognized that ethical, legal and social issues (quickly abbreviated "ELSI") would be challenges. Much of the initial credit goes to Dr. James Watson, co-discoverer of the structure of the DNA molecule, and first director of the Genome Office at the National Institutes of Health. At the establishment of his office, Watson noted the concerns about ELSI issues and said that some funding, initially 3%, should be set aside to study ELSI issues that would arise from the HGP. DOE soon followed setting aside funds from its genome research budget as well and participating in joint coordination activities between the two agencies' programs.

One of the first tasks was to define a research agenda for the ELSI programs. It quickly became apparent that the range and diversity of issues was considerable and that many of them were not particularly new. Table 1 shows a list, from 1990, of the issues that were identified then.

Table 1: Research Agenda for the ELSI Program of the HGP.

A) Science Policy Issues

- Organization and dynamics of genome research
- Intellectual property policy
- Commercialization practices
- Human Subjects research issues

B) Clinical Practice Issues

- Technology assessment criteria
- When and who to test
- Information disclosure practices
- Counseling practices

C) Health Policy Issues

- Public health role of genetic services
- Allocation of resources to genetic services

D) Privacy Issues

- DNA databanking policies
- DNA identification standards
- Legal definitions of "genetic privacy"
- Intra-familial communications issues

E) Civil Rights Issues

- Employment testing policies
- Health insurance underwriting practices
- Life and disability insurance practices
- Social discrimination/stigmatization
- Genetics and race

F) Educational Policy Issues

- Clinical education models
- Public education models
- Health professional training

Many medical uses for HGP-derived knowledge can be identified, but so too, can numerous non-medical uses and these are a concern. These non-medical uses can include identification of individuals, whether in the forensic context or in the context of paternity disputes, numerous insurance issues (as more information about an individual's future health status is known or knowable, it may be harder to obtain insurance and an insurance company, answerable to its investors, may elect not to underwrite "high risk" individuals), employment issues (should an employer be able to learn of an employee's potential genetic susceptibilities so the employer can assign her or him to work where exposure to a hazard is less likely?), personal injury litigation (can one party to a dispute request genetic information about the potential longevity of the other party with the intent of seeking reduction of potential damages due to a genetically-

shortened life span?), commercial transactions (should a bank or loan agency be able to request genetic information from a borrower to ensure she or he will survive long enough to pay back the loan?), domestic relations (adoption or child custody cases), educational settings (can a school require genetic information about students in order to sequester or even refuse any student thought to have a predisposition to disruptive behavior?) and issues in the criminal justice setting (can a defendant argue that "my genes predisposed me to this behavior" and thus seek mitigation of punishment post-conviction with the equivalent of a "not guilty by reason of genetic inheritance" assertion?) Presently, most of these questions represent untested legal scenarios and also presuppose much highly questionable science, but it is hard to imagine that they won't be brought to courts around the country over the next decade. Courts and their allied personnel will need to face these issues when they arise.

Faced with this plethora of ELSI issues, each one an important and difficult challenge, and a finite budget limit, DOE elected to downselect from this list those topic areas considered most critical and within the traditional interests of the agency. DOE has had a long and consistent record of supporting science education so promoting education to various groups about both the promise and concerns raised by genomics was a natural area for the DOE ELSI program to enter. The DOE ELSI program has supported numerous curriculum units for high school biology students, programs for Institutional Review Boards (IRBs, responsible for overseeing human subjects research), documentaries (both public TV and radio) for general audiences, as well as a major program of workshops for members of the judiciary. To date, nearly 1400 judges have been introduced to some of the basics of genetics so that when (not if) cases appear that include some genetic issue, or a dispute in which genetic evidence is introduced, courts will be better prepared to deal with it.

Since the Supreme Court case of *Daubert v. Merrill Dow Pharmaceuticals* (509 U.S. 579 (1993)) made judges in Federal jurisdictions responsible for managing the introduction of scientific evidence, it has been even more imperative that judges have some understanding of, and comfort with, the nature of scientific evidence. No attempt is ever made, in these workshops, to instruct judges what they should think (that would almost surely be impossible anyway); rather, the aim is to reduce the apprehension and anxiety that novel and unfamiliar science can produce. DOE also selected genetic privacy as an area of importance, largely due to the fears that had been voiced that this was one of the greatest concerns about the project. Within this broad area, the DOE ELSI program has focused on privacy in workplace settings (at many DOE worksites, hazardous materials were used in nuclear weapons manufacturing, and some of these materials could affect past and present workers in ways influenced by genetic susceptibilities) and issues surrounding computerized databases of genetic information which can be a tempting (and relatively easy) target for a moderately sophisticated hacker to penetrate.

In 1991, patents on partial gene fragments were filed with the Patent and Trademark Office, initiating a controversy that still persists and the DOE HGP elected to add this topic to its ELSI agenda. As economic returns from investments in biotechnology have become more attractive, the financial stakes from control of intellectual property (through patents or copyrights), resources, technologies and (recently) databases of human genetic variations and test results, have gotten larger and more precious; this can be expected to generate cases that courts will be called upon to decide, as will the inevitable regulatory disputes as the applicable rules and responsible agencies are identified. Lastly, issues associated with complex traits have been identified as an area of importance for ELSI studies since these encompass the conditions (diabetes, heart diseases, cancers, mental illnesses, etc.) that affect the majority of Americans but to which genetic background makes an uncertain contribution. These ELSI issues will be particularly difficult not only for scientific reasons, but also for economic reasons since they affect so many people.

While the HGP raises many difficult issues that courts will need to prepare for, we should not lose sight of the many ways in which this exciting science will make a better world for all of us and our descendents. It is perhaps too easy to offer scenarios of a world in which no privacy exists, each of our genomes is an open book that anyone, individual, company, or government can read. One of the great features of the genome is its mind-numbing complexity and how little, at least today, we understand about how it functions. The simplest known microbe has a genome that has only 580,000 bases (about 32 pages of that

phone book at the top of this article), determining less than 500 "parts," yet we are a long way from understanding how this microbe lives and functions. From 500 parts to 120,000 parts is a vast gulf that will take many years to begin to cross. For a scientist, it is the next frontier, analagous to the voyages of Star Trek's Enterprise.

So what should you, a Legal Assistant, do to prepare yourself for the advent of the genomic age of the 21st Century? I will make several suggestions. First, learn all you can about what the Human Genome Project is and the numerous societal implications that will accompany it (a superb web site that you can start with is: *http://www.ornl.gov/hgmis*). Keep in mind that the completion of the sequencing of the human genome is not the end, but only the beginning, that the HGP is really an infrastructure upon which the next scientific advances can build. The completion of the HGP, probably in about 2003, will furnish science with a resource that all can use. Henceforth, it will be unnecessary for a researcher to devote time and effort to a "construction" task that once done, shouldn't need to be done again. The talents and energy of the scientific community should strive to understand what the human genome is telling us, not what it is.

Second, keep in mind that the genome stores information, much as the music on that prerecorded cassette tape. Information, whether printed, electronic, taped or genomic can be misunderstood, misused, bought, sold, stolen, forged, damaged in copying, and (it is to be hoped someday) repaired. Information is key to normal commerce, but when it is someone's private property, rules need to be defined to control and regulate what happens with it and how it is used.

Third, biology is hard and a complex biological system is subject to many impacts, from the sequence of environments it experiences, that we cannot yet measure or even detect. This means that, as in the case of the identical twins experiencing different diseases, there are many answers we cannot look to genomics to provide. This is a conundrum for courts that must decide disputes on the basis of the evidence presented there and then. Fourth, understand the rudiments of statistics. Much genetic risk information (e.g. the added risk a woman has for breast cancer if she is positive for one of the higher-risk variants at her Breast Cancer-1 gene) is statistical and applies to populations not to any one individual. Yet both in the physician's office and the courtroom, it is the fate of the individual that is the focus of concern. It is critically important to know just how much weight to put on a genetic test result. Finally, most of the challenging ELSI issues that arise from the HGP are beyond any one community to solve and that means that all of us must address them. No one community in our society is better placed than any other to determine how genetic information should be used and what is the wisest course forward, consistent with our fundamental principles of governance and fundamental liberties, rightly involves everyone.

Daniel Drell is a biologist with the Human Genome Program, Life Sciences Division, Office of Biological and Environmental Research, U.S. Department of Energy. His duties include oversight of research issues involving ethical, legal, and social implications, particularly questions of privacy and confidentiality of genetic information, intellectual property protection related to genetic information, education, and the societal implications of complex multigenic traits.

The online presentation of this publication is a special feature of the Human Genome Project Information Web site.