Rowan University
Rowan Digital Works

Theses and Dissertations

4-28-1999

Changing destinies: an overview of the Human Genome Project

Jennifer Lynne Watson
Rowan University

Follow this and additional works at: https://rdw.rowan.edu/etd

Part of the Higher Education and Teaching Commons

Recommended Citation

Watson, Jennifer Lynne, "Changing destinies: an overview of the Human Genome Project" (1999). *Theses and Dissertations*. 1898.
https://rdw.rowan.edu/etd/1898

This Thesis is brought to you for free and open access by Rowan Digital Works. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Rowan Digital Works. For more information, please contact graduateresearch@rowan.edu.

المنارة للاستشارات

www.manaraa.com

CHANGING DESTINIES:
AN OVERVIEW OF THE HUMAN GENOME PROJECT

by
Jennifer Lynne Watson

A Thesis

Submitted in partial fulfillment of the requirements of the
Master of Arts Degree in the Graduate Division
of Rowan University
Spring 1999

Approved by
Dr. Richard J. Meagher

Date Approved April 28, 1999

ABSTRACT

JENNIFER LYNNE WATSON
CHANGING DESTINIES:
AN OVERVIEW OF THE HUMAN GENOME PROJECT
1999
DR. RICHARD J. MEAGHER
MASTER OF ARTS IN HIGHER EDUCATION

The Human Genome Project was initiated in 1988. Its purpose being to sequence the entire human genome, including the noncoding regions by 2005. Through the research conduced, due to the project, several genes associated with specific diseases have been identified. This paper is a review of the Human Genome Project and genes that have been associated with the onset of specific disease. Cystic Fibrosis, a disease affecting the respiratory system, has been isolated to a gene on chromosome 7. Alzheimer's Disease, the disease which is most responsible for dementia, has several genes associated with its onset. These genes lie on chromosomes 21, 19, 14 and 1. Two genes have been associated with Breast Cancer, Breast Cancer 1 (BRCA1), found on chromosome 17, and Breast Cancer 2 (BRCA 2), found on chromosome 13. A gene found on chromosome 4 is responsible for the onset of Huntington Disease, a disease that causes neural cell death. Although the gene responsible for Schizophrenia is not yet clear, there has been some speculation that chromosome 22 may be involved with its onset; however, there is still additional research that needs to be conducted. Through the information that has been discovered by genomic

research, new treatment strategies are being developed to treat these diseases and hopefully someday defeat them.

MINI-ABSTRACT

JENNIFER LYNNE WATSON
CHANGING DESTINIES:
AN OVERVIEW OF THE HUMAN GENOME PROJECT
1999
DR. RICHARD J. MEAGHER
MASTER OF ARTS IN HIGHER EDUCATION

The Human Genome Project was developed to sequence the entire human genome and identify diseases associated with specific genes. This paper is a review of the Human Genome Project and the genes that have been associated with five specific diseases: Cystic Fibrosis, Alzheimer's Disease, Breast Cancer, Huntington Disease and Schizophrenia.

TABLE OF CONTENTS

INTRODUCTION	1
LITERATURE REVIEW	2-11
METHODS AND MATERIALS	12
CHAPTER 1: CYSTIC FIBROSIS	13-19
CHAPTER 2: ALZHEIMER'S DISEASE	20-30
CHAPTER 3: BREAST CANCER	31-38
CHAPTER 4: HUNTINGTON DISEASE	39-46
CHAPTER 5: SCHIZOPHRENIA	47-50
CONCLUSION	51-52
LITERATURE CITED	53-60

INTRODUCTION

Over the centuries, man has been striving to unravel the mysteries of the human body. Through the advancement of science and technology what was once misunderstood is now considered common knowledge. From the original observation of cells to the discovery of the molecular structure of deoxyribose nucleic acid (DNA), scientists have challenged their own abilities to discover the truth about life. As more information about DNA became known, its influence on man's destiny was apparent. The challenge of trying to unravel the genetic code was set before scientists. If the code that makes us who we are could be deciphered, then problems within the code that lead to disease could be identified and possibly treated or altered. This ambitious undertaking is what is known today as the Human Genome Project.

LITERATURE REVIEW

The Human Genome Project is a research project that is being conducted on an international level. The purpose of this research is to construct both genetic and physical maps of the estimated 50,000 to 100,000 genes, which make up the human genome, with an original deadline of 2005 (Lyon and Gorner, 1995). In addition to the mapping of human genes, scientists are striving to sequence the 3 billion nucleotide pairs of the human genome, including all of the non-coding regions (Marshall, 1996). This complex map of the human genome will provide information that is needed to understand human biology and disease on a genetic level (Collins, et.al., 1995). Once specific genes are identified and linked to specific diseases, they can be examined through biological experimentation and new medical treatments can be initiated (Marshall, 1996). This research will help in the understanding of over 4000 genetic diseases that plague mankind, as well as many diseases in which genetic predisposition plays a role. (National Human Genome Research Institute, 1998) .

In the summer of 1985, the prospect of researching the human genome was first discussed seriously. Nobelist, Renato Dulbecco, former president of the Salk Institute, described two ways to combat the killers of the twentieth century, "Either to try to discover

the genes involved in malignancy by a piecemeal approach, or to sequence the whole genome of a selected species". Dulbecco encouraged the United States to initiate the search for the total sequence of the human genome and to do it with the same vigor that was used to 'conquer space'. Coincidentally, during the time when Dulbecco was pressing for the initiation of this research the Department of Energy was searching for new projects to fund. The Department of Energy did a great deal of research on this topic and ultimately Congress approved a fifteen-year project. The project had a funding potential of $200 million per year, $3 billion total ($1 per nucleotide sequence). Thus, the Human Genome Project was initiated and proclaimed the 'Holy Grail' of biology (Lyon and Gorner, 1995).

The initial United States involvement stemmed from the National Institute of Health. The director of the National Institute of Health, James Wyngaarden, created the National Center for Human Genome Research in October 1988. James Watson was asked, and agreed, to head the center. Watson, also the director of the Cold Spring Harbor Laboratory, embraced the challenge and stated that, "We used to think our future was in the stars, now we know it is in our genes". (Lyon and Gorner, 1995).

In accordance with Watson and legislation by Congress, the Human Genome Project consisted of four tiers which included: mapping the twenty three human chromosomes, sequencing these chromosomes in their entirety including the non-coding sequences, placing this information on computer databases and allowing scientists throughout the world to have access to the information, and finally, initiating an ethical component to the project so that the knowledge obtained is not used in any aspect of discrimination (Lyon and Gorner, 1995).

In the Fall of 1991, as the Human Genome Project continued progressing, James Watson found himself on unsettled ground. Bernadine Healy, Wyngaarden's successor, made it clear, at least in Watson's perception, that she wanted him to step down from the post due to conflicts of interest concerning the project. In April 1992 Watson quit, faxing his resignation to Healy (Lyon and Gorner, 1995).

The National Institute of Health did not replace Watson for almost a year, and many "insiders" doubted that Healy and her staff would be able to find a person adequate enough to fill Watson's shoes. In the fall of 1992 the announcement was made that Francis Collins, a universally respected scientist, was offered the position. In April 1993 Collins assumed command of the project (Lyon and Gorner, 1995).

Observations about the progress of the Human Genome Project in its first five years were made by Collins and Galas (1993). The first observation they made focused on technology. The continuing development of successful technological approaches that can be used in genomic and genetic research has been a contributing factor to the programs initial success, and will be a critical component in the future of the project. The second observation they made involved the level at which genomic research was being focused. Initially, a great deal of time was spent on Large-scale mapping which focused on the construction of chromosomal maps. Genetic maps and physical maps have been constructed using the information gathered from the Large-scale mapping efforts. However, other information gathered with the use of Small-scale mapping, as well as Large-scale mapping, will be useful for further study of the human genome. The third and final observation was that the identification of genes with the use of maps and sequences, which was an original goal of the

project, has developed into a definite activity of the project. The progress that has been made in the area of gene identification, along with new technology that has been developed for gene identification, will help with the production of more detailed maps (Collins and Galas, 1993).

Collins and Galas (1993) also modified the original five year plan for the Human Genome Project stating new goals that they hoped would be achieved in the next five years of the project. These goals were divided into separate categories, the first of which deals with "The Genetic Map". The five year goal is to complete a human genetic map that is fully connected and contains 'markers' that are on average spaced 2 to 5 centimorgans apart. A centimorgan (cM) is roughly one million base pairs. A genetic map is made by determining how often two 'markers' are inherited together. A 'marker' can be a physical trait, a specific medical condition or syndrome, or a DNA sequence that is detectable (The National Human Genome Research Institute, 1998). Rowen (1997) states that with the use of various techniques 'markers' have been placed across the genome every 200 kb on average. The use of genetic maps along with "florescence in-situ hybridization" and "radiation hybrid mapping" are techniques that have been used to locate 'markers' (Rowen, 1997).

The second goal stated by Collins and Galas (1993) deals with the creation of "The Physical Map". They hoped to complete an "STS-based physical map of the human genome with a resolution of 100 kb". STS refers to a "sequence-tagged site" which The National Human Genome Research Institute (1998) describes as a short DNA sequence that, through research, has been categorized as unique. In the past, it has been difficult for researchers to compare the results of physical maps generated by their colleagues because there were several

different mapping methods that were being used to generate physical maps. Initiating the STS technique will give researchers one common "language" to report their findings in, no matter what type of method they use. The idea behind this system is that each mapped element is represented by an STS and the order and spacing of the STS's are also shown. Rowen (1997) reports that by using this STS system researchers have been able to assemble a physical map for most of the human genome.

The third goal described by Collins and Galas (1993) involves DNA sequencing. New technology and improvements on current methods of sequencing are goals of the Human Genome Project in order to lower the cost of sequencing to $.50 per nucleotide. Other goals include developing technology in order to conduct "high throughput" sequencing, as well as sequencing a total amount of 50 Megabases (Mb) per year. In reality, the productivity of the researchers involved in sequencing the human genome has been very low. The director of the Human Sequencing Program, Mark Adams, at The Institute of Genomic Research in Rockville, Maryland, does not seem to think they will be able to achieve their stated goal of sequencing 100 Mb per year by 1999 or 2000. The Whitehead Institute of Biomedical Research in Cambridge, Massachusetts, expected to have completed 23 million bases by May 1998, and have actually completed less that 9 million. Similarly, The Genome Center at Baylor College of Medicine in Houston, Texas, had an original goal of having 15 million base pairs to submit to Genbank, a public database that contains sequencing data, but only submitted 8 million bases instead (Pennisi, 1998). Koonin (1998) states that less than 5% of the human genome has been sequenced. The traditional method used for sequencing DNA, called Sanger sequencing, involves using DNA precursors that are radioactively labeled with

32P. The DNA fragments are separated according to their length using polyacrylamide gel electrophoresis. The results of the DNA sequencing are determined by autoradiography (Russell,1992). Koonin (1998) states specific disadvantages of using the present technology for DNA sequencing. These drawbacks include the time and labor that is involved in preparing and running polyacrylamide gels, as well as the large amount of samples that are needed to obtain good results. According to Pennisi (1989), The National Human Genome Research Institute has funded six pilot centers giving them free reign to choose any strategy for sequencing that they prefer. There were only a few rules that these centers had to follow. The most important one being that the final data they presented had to be 99.99% accurate, that is to say, there could only be one error for every 10,000 bases they sequenced. The ultimate challenge presented to the centers is to be able to keep their standards high without having a significant increase in the cost of productivity. Ultimately, improvements on existing techniques, and development of new techniques based on increased technology and principles, will help in creating a complete sequence of the human genome (Pennisi, 1998).

"Gene Identification" is another goal of the Human Genome Project. Collins' and Galas' (1993) goals for gene identification include identifying genes using more proficient methods, as well as placing the genes that have been identified on physical maps. Since the inception of the Human Genome Project, the availability of genetic and physical maps of the human genome, and also of the genomes of certain model organisms, have increased. Using these genetic maps and physical maps that have been constructed as a result of the Human Genome Project has helped in the discovery of genes that underlie Mendelian disorders (Collins et. al., 1997).

"Technology Development" is an important component of the Human Genome project. The goal in the area of technology development is to expand the current technology for all aspects the Human Genome Project, such as DNA sequencing, mapping, and gene identification (Collins and Galas, 1993). Specific types of technology, such as robotics, have been experimented with. Eric Lander, head of the Whitehead Center at MIT, estimated a 30%-40% savings in labor costs with the use of robots. Later Lander realized that this estimate was "over optimistic" (Pennisi, 1998). The original plan that was established by Watson aimed to develop technologies and refine them by sequencing the genome of simpler organisms. This strategy did work in that when the Human Genome Project began, Watson estimated that to sequence one base pair the cost was approximately $10 with a potential to be reduced to $1 per base pair within 5 years (Marwick, 1989).

The next set of goals that Collins and Galas (1993) focused on were those associated with "Model Organisms". A great deal of progress has been made with respect to model organisms which are being researched in conjunction with the human genome. The original goals that were set in this area are expected to be reached and possibly even surpassed. The revised goals include completing an "STS map of the mouse at 300 kb resolution". In addition to the mouse genome, the goals include completing the sequence of the Escherichia coli and Saccharomyces cerevisiae genome.

The continued development of new databases and database tools for storage and retrieval of information gathered during the course of the Human Genome Project is an ongoing goal. Specifically, Collins and Galas (1993) have goals in the area of "informatics" to create databases that have easy access, as well as information exchange among databases.

In addition, they would like to see the consolidation of software, as well as offering tools that will allow for comparing genome information from one database to another (Collins and Galas, 1993). The three databases that are of essential importance to the Human Genome Project are GenBank, Genome Data Bank, and Protein Information Resources. The information they provide consists of DNA sequences (GenBank), chromosome mapping information (Genome Data Base), and protein sequence and structure (Protein Information Resources). The Department of Energy and The National Institute of Health established a Joint Informatics Task Force (JITF) to help in the development of "detailed informatics programs" (The National Human Genome Research Institute (1998).

There has also been modifications made in the area of the "Ethical, Legal and Social Implications" of the Human Genome Project. In October 1988, when James Watson officially announced his intentions for initiating the Human Genome Project, one of the first questions that he was asked was about the social implications of this research. In response to this issue there was a committee formed, the National Center for Human Genome Researcher's Committee on Ethical, Legal and Social Implications (ELSI) which received 5% of the budget designated for human genome research (Marshall, 1996). The initiation of this committee established the largest biomedical ethics program in this country and possibly the largest in the world (Zylke, 1992). This research would address issues such as genetic discrimination, genetic elitism, and issues of eugenics. With the increasing amount of information that has been gathered since the inception of the Human Genome Project, the goals for the ELSI component of the project have been modified. One of the goals is to develop policies involving services with regard to genetic testing. Another goal of the ELSI is to reduce

genetic discrimination allowing for greater acceptance to genetic variability. In addition, there is a goal to increase public, as well as professional education to the issues that are raised by researching the human genome (Collins and Galas, 1993).

The last three categories Collins and Galas (1993) have discussed are in the areas of "Training", "Technology Transfer" and "Outreach". The goal that they have established for training is to encourage scientists that are involved in various fields of study, related to the Human Genome Project, to become familiar with the techniques used in genome research. The goal of technology transfer is to encourage transfer of information within and between various genome centers. Finally, the goals concerning outreach include sharing information within 6 months of its discovery by way of a public database, and the development of "distribution centers" that would process and circulate information about the Human Genome Project (Collins and Galas, 1993).

Since the inception of the Human Genome Project, it has been a federally funded program with the most substantial portion of the budget provided by The National Institute of Health and The Department of Energy (Lyon and Gorner, 1995,). Recently, J. Craig Venter, former researcher at The National Institute of Health, (Lyon and Gorner, 1995) and Perkin-Elmer Corporation have teamed up to create a new company, The Institute of Genomic Research (TIGR). The goal of this private company is to sequence the human genome in three years with a cost of $200 million to $300 million (Travis, 1998). The technique that Venter is using is referred to as "shotgun sequencing". This technique involves "randomly breaking DNA into segments of various sizes and cloning these fragments into vectors". "This project aims to produce a highly accurate, ordered sequence that spans more

than 99.9% of the human genome" (Venter et. al., 1998). The shotgun approach to sequencing does not include any mapping as does the Human Genome Project. Instead, it enlists the help of a computer to sort through and "line up" the thousands of DNA sequences that it is presented with. This strategy will result in a sequence that is not complete, leaving 3,000 to 4,000 gaps in the finished product, each gap consisting of approximately 50 base pairs each. Venter et. al. (1998) proposes that the regions that will be left are areas that consist of "repetitive DNA" which is of little interest. Venter et.al. (1998) also supports the fact that the genome sequence his company constructs will be just as accurate as its competitors. (Travis, 1998). The establishment of this new company has sparked the leaders of the Human Genome Project to change the deadlines they have set for the completion of the project. They now plan to have a completed sequence of the human genome by 2003 with a "working draft" available by 2001. Francis Collins, director of the National Human Genome Research Institute, states that increasing the productivity of DNA sequencing "eclipses all other priorities" (Holden, 1998).

Although a complete sequence of the human genome has not been constructed thus far, there has still been a great deal of progress made in the area of gene identification and the involvement of these genes in the area of medical genetics. In this paper five different disorders will be examined: Cystic fibrosis, Alzheimer's disease, Breast cancer, Huntington disease and Schizophrenia. The information presented will include what discoveries have been made on a genetic level about these disorders, the biochemistry of the genes involved in these disorders, and also innovative therapies that are being conducted to help in the prevention and treatment of these disorders.

MATERIALS AND METHODS

Cystic Fibrosis, Alzheimer's Disease, Breast Cancer, Huntington Disease and Schizophrenia all have genes that have been associated with their onset. In this paper each disease will be examined. The genes that influence the onset of these diseases will also be discussed. Included in the discussion will be the strategies that were employed to isolate these specific genes. In some cases the function of specific genes have been identified. These functions will be incorporated in the analysis of the disease. Also, in some cases, various treatment strategies have been developed to combat the inevitable fate that these genes lead to.

CHAPTER 1: CYSTIC FIBROSIS

In August 1989, Francis Collins and a group of his scientific colleagues were rushed to Washington D.C. to attend several press conferences. The reason for these meetings was to announce that they had identified the gene and the protein defect which results in the devastating disease Cystic Fibrosis (CF). This announcement thrilled thousands, especially parents of CF victims. The discovery of this gene would make it possible to develop drug therapies that would treat the problem at the source, not just the symptoms (Lyon and Gorner, 1995).

Cystic Fibrosis, regarded as the most commonly occurring autosomal disease in the Caucasian population, has a frequency of 1 in 2000 live births and a carrier population of about 5%. Clinically, the major symptom of CF is the obstruction of airways by thick mucus that is very sticky. This results in respiratory tract infections (Collins, 1992). "Chronic pulmonary disease, pancreatic exocrine insufficiency, and an increased concentration of sweat electrolytes" are also common symptoms of the disease (Rommens et al., 1989). A common method for the management of cystic fibrosis is chest percussion which helps to clear infected mucus secretions from the respiratory tract. A great deal of importance is also placed on the

nutritional habits of the patient. The chances for survival have increased over the past 4 decades. The average survival age is roughly 29 years (Collins, 1992).

Cystic Fibrosis has long been a mystery to physiologists and biochemists; however, more recently they have been able to solve the riddle behind this disorder. In patients with CF the mechanism that controls the efflux of chloride ions across the apical membrane does not function properly. There is an elevation of adenosine 3',5'-monophosphate (cyclic AMP), and cyclic AMP-dependent kinase (PKA) is activated by cyclic AMP in CF cells, but PKA fails to activate a chloride conductance. This information was new to CF researchers, but was not sufficient to identify the protein product of the gene that is defective in CF patients (Collins, 1992). A more aggressive strategy was employed to identify the CF gene, positional cloning, which was triumphant in the identification of the CF gene (Collins, 1992). The introductory stages of this process included the mapping of the CF gene to chromosome 7. This was achieved by examining DNA from affected individuals, as well as a panel of polymorphic DNA markers through linkage analysis. Through this process the CF locus was unambiguously assigned to the long arm of chromosome 7, band q31. The identification of MET and D7S8, closely linked flanking markers, made it possible to pinpoint the CF gene. The gene cloning strategies employed included chromosome jumping from flanking markers, and pulse field gel electrophoresis to clone DNA fragments from defined physical regions. "Chromosome microdissection" and saturation cloning of a large number of DNA markers from the q31 region were also used. Saturation mapping was used to identify two DNA markers, D7S122 and D7S340. These two markers, along with MET and D7S8, helped to create genetic and physical maps of the CF locus. Their order on the CF locus are as follows:

MET-D7S340-D7S122-D7S8 with intervals of 500-10-98 Kilo-Base (kb) respectively. As indicated by the genetic data, D7S122 and D7S340 are in close proximity on the CF locus. Next, a large portion of the DNA surrounding these markers needed to be cloned to seek out candidate gene sequences. Through the use of chromosome 'walking' and 'jumping' techniques, ten genomic libraries were constructed, as well as a restriction map of the cloned human DNA (Rommens et al., 1989).

The entire cloned region of the CF locus extended over 500 kb. To ensure that the DNA segment isolated by the chromosome 'walking' and 'jumping' techniques was collinear with the genomic sequence, each segment was examined using the following techniques: "Hybridization analysis with human-rodent somatic hybrid cell lines to confirm localization on chromosome 7, pulse field gel electrophoresis, and comparison of the restriction map of the cloned DNA to that of the genomic DNA." The evaluations indicated that most phage and cosmid isolated DNA segments were correct, although a few inconsistencies resulted from cloning artifacts, as well as cross-hybridizing sequences that are found in other areas of the human genome (Rommens et al., 1989).

Through the use of long range restriction mapping it was indicated that the entire CF locus was contained on a 380 kb Sal I fragment. The restriction sites derived, from the pulse field gel electrophoresis analysis, were aligned with overlapping genomic codes to indicate that the CF locus was actually about 250 kb (Rommens et al., 1989) with a messenger RNA (mRNA) reduced to approximately 6.5 kb. This mRNA encode for a protein containing 1480 amino acids (Collins ,1992).

The final proof that was needed to confirm that the CF gene was identified was the mutation that distinguished normal individuals from affected individuals. This was done by comparing cDNA sequences acquired from CF and unaffected (N) individuals.. The most apparent difference between the CF and N sequences was a 3 Base-Pair (bp) deletion in Exon 10 of the candidate gene (Collins, 1992). This deletion results in the loss of the amino acid phenylalanine (∧F508) in the CF polypeptide that was predicted. Kerem et al. (1989) studied the proportion of CF patients carrying this deletion (∧F508). Genomic DNA samples were taken from patients and their parents. These samples were amplified by PCR and hybridized to 32P labeled oligonucleotides specific for the normal and mutant sequences. The results of their study indicated that about 68% of the CF chromosomes in the patient population had the ∧508 mutation. Conversely, none of the N chromosomes had the deletion. The results that they obtained suggested that this 3 bp deletion is specific to CF and is the major mutation which results in the disease (Kerem et al., 1989). Presently, there are over 170 mutations that have been described in the CF gene, many of these mutations have only been identified in a single individual (Collins, 1992).

Now that the CF gene had been identified, it was necessary to find the protein that it codes for. Lyon and Gorner (1995) state that the CF gene codes for a protein which consists of 1479 amino acids, and the N gene codes for a protein consisting of 1480 amino acids. The absence of the 3 bp sequence results in a loss of the amino acid phenylalanine. Researchers named this protein CFTR, the cystic fibrosis transmembrane conductance regulator (Lyon and Gorner, 1995). With regard to genetic data and predicted properties, it is feasible to resolve that CFTR is responsible for CF (Riordan et al., 1989).

The amino acid sequence of CFTR is very similar to a family of proteins which are involved in active transport of substances across cell membranes. These proteins can be found in a wide range of species, from bacteria to mammals. The subfamily these proteins belong to is called the traffic adenosine triphosphatase family (ATPase). They have a common characteristic in that they have one or two hydrophobic transmembrane domains. They also have one or two nucleotide binding folds (NBFs) which binds, as well as cleaves, ATP. This provides the energy resource for transport. CFTR was placed in this subfamily because it has two transmembrane domains and two NBFs. Another characteristic of CFTR is its highly charged central domain. This domain contains serine residues which are thought to be targets of PKA-mediated phosphorylation (Collins, 1992).

The mystery of CF was beginning to unravel; however, to most researchers, the most elusive characteristic of CF was how one genetic flaw could produce such diverse symptoms. Interestingly, some researchers discovered that the cells that line the respiratory tract in CF patients cannot transport chloride ions across the cell to the outside. This affects CF patients by trapping water inside the cell making it impossible to dilute the mucus in their respiratory tract; therefore, the mucus remains thick (Lyon and Gorner, 1995).

Through further studies and experimentation it was concluded that CFTR is a chloride channel instead of the initially speculated active transport mechanism (Collins, 1992). The structure of the protein is similar to that of a family of proteins which act as channels (Lyon and Gorner, 1995). Riordan et al. (1989) suggests another possibility, the CFTR may not be an ion channel, but may regulate the activities of an ion channel. In either situation, the fact that the ATP-binding domain is present on CFTR implies that the hydrolysis of ATP is

necessary for transport functions. The Λ508 may result in the improper binding of ATP resulting in an ineffective transport of ions. Bradbury et al. (1992) confirmed that the CFTR protein does not function properly in patients with CF, and that the N CFTR gene works as a chloride ion channel. They state that "CFTR can function as a cyclicAMP-activated chloride channel within epithelial cells, where it regulates plasma membrane chloride conductance."

The research that has been conducted thus far to unravel the mystery of CF has initiated the development of new methods of treatment, such as aerosol therapy, as well as new ways to identify and treat the disease through neonatal screening. In the area of aerosol therapy, recombinant human deoxyribonuclease I (rhDNase) has been used to treat CF. The rhDNase is administered by aerosol and acts in-vivo by attaching to high-molecular-weight DNA to increase the elimination of mucus, and may be a possibility for the short-term management of lung function for CF patients (Hubband et al., 1992). In addition, adenovirus vectors have been developed to fight CF. Adenoviruses have many unique qualities that make them desirable, such as being able to infect cells in any phase, unlike retroviruses. This characteristic is essential to the lungs because the number of cells going through division at any one time is relatively low. In this procedure N CFTR genes are administered by aerosol into the CF patients lungs in the form of an adenovirus. The N CFTR combines with the CF CFTR allowing the cells chloride ion channels to function normally improving airway function (Lyon and Gorner, 1995). Collins (1992) points out that this treatment may be transitory and treatments may need to be repeated often.

There has also been success in the area of Neonatal Screening. This screening has been used to enhance the early diagnosis of CF neonatally, as well as offer early treatment options to the parents. Through CF screening, mothers have been able to change their nutritional habits to improve the health of the baby. The CF screening also allows the family to prepare for the arrival of a child with CF and to help them plan for the best medical treatments possible. The infants who were screened and treated neonatally had less pulmonary inflammation and less deterioration of lung function. These infants suffered no long-term adverse affects due to the screening program. This should be encouraging for families who are deciding if they want to go through the CF screening process (Dankert-Roelse and Meerman, 1997).

There are many advances that have been made to help fight CF. The research that has been done and the new treatments that are being employed are helping families cope, and also giving them new hope to fight this terrible disease.

CHAPTER 2: ALZHEIMER'S DISEASE

Alzheimer's disease, the disease which is most responsible for dementia, was initially described in 1906 by Alois Alzheimer, a German physician. Through the use of autopsy techniques employed at the time Alzheimer discovered that the cerebral cortex of a 55-year-old woman, who had suffered from dementia, had nerve cell abnormalities. These abnormalities consisted of tangled fibers and clusters of degenerated nerve endings, or plaques (Check, 1989). These two abnormalities described by Alzheimer are seen in most deceased Alzheimer's disease victims and are today described as hallmarks of the disease (Lyon and Gorner, 1995).

Plaques, which are developed at the end of nerve cells, result in the blunting of the tips of the 'branch like spines', which are known as axons and dendrites. The nerve cells communicate through these axons and dendrites by chemicals referred to as neurotransmitters. It is known today that plaques consist of a tough fibrous protein, *B*-amyloid, surrounding either dead or dying nerve tissue. Tangles, the other abnormality described by Alzheimer, are filaments within the nerve cells that are twisted or frayed, similar to a wire that has been 'burnt out' (Lyon and Gorner, 1995).

Plaques and tangles are found prominently in the cells of the cortex and hippocampus, which are the two major areas of the brain associated with speech, memory, and reasoning. However, it is not known whether plaques and tangles are what causes Alzheimer's disease, or if they are merely an effect (Lyon and Gorner, 1995).

In addition to plaques and tangles, Alzheimer's disease patients produce less acetylcholine, a neurotransmitter found in the brain. This neurotransmitter is made in the basal nucleus, a brain structure, and is disbursed to the cortex and hippocampus. The degree to which Alzheimer's disease victims lack acetylcholine is directly related to the severity of memory loss they experience. In addition to acetylcholine, other neurotransmitters, somatostatin, noradrenaline, and serotonin, are not produced in normal amounts in Alzheimer's disease patients. Presently, the lack of specific neurotransmitters is a consequence of Alzheimer's disease, rather than a cause (Lyon and Gorner, 1995). So then, what is the cause of Alzheimer's disease?

There have been many theories as to the cause of Alzheimer's disease, including a viral cause which suggests that a viral infection in the brain causes selective brain cells to die. Another hypothesis suggests that a toxic accumulation of aluminum in the brain may cause Alzheimer's disease. This suggestion was made because Alzheimer's disease patients had a higher than normal concentration of aluminum salts in their brain tissue (Lyon and Gorner, 1995). An additional theory which hypothesizes genetic causes for Alzheimer's disease is presently the most active area of research that is being pursued (Check, 1989).

In some instances Alzheimer's disease affects more than one person in a family. It has been suggested that this form of Alzheimer's disease, familial Alzheimer's disease, is caused

by a genetic defect (St George-Hyslop et al., 1987). In a report published by St George-Hyslop et al. (1987) the chromosomal location of a defective gene contributing to Alzheimer's disease was announced. The focus of their study was on chromosome 21. The rational for this being that a large number of Down's Syndrome (trisomy 21) patients who reached older age developed "Alzheimer-like neuropathologic changes." Their hypothesis was that a gene on chromosome 21 could influence the development of Alzheimer's disease.

In their study, St. George-Hyslop et al. (1987), used genetic linkage of polymorphic DNA markers to find the chromosomal location of familial Alzheimer's disease. Through their study it was shown that the familial Alzheimer's disease gene does lie on chromosome 21 in the region of 21q11.2-21q21. Several DNA markers were used to try to pinpoint the location of familial Alzheimer's disease on chromosome 21. The data was statistically analyzed using the logarithm of likelihood of linkage (lod score) to familial Alzheimer's disease. Two markers, D21S1/D21S11, gave positive lod scores for the location of a familial Alzheimer's disease allele on chromosome 21 suggesting that the gene is located in the area of these two markers. St. George-Hyslop et al's (1987) study also confirmed that the familial Alzheimer's disease gene is not situated on the Down's Syndrome region (21q22) of chromosome 21. Rather, it is just outside the region falling between 21q11.2-21q21. In addition, St. George-Hyslop et al. (1987) also suggested that the familial Alzheimer's disease gene found on chromosome 21 is associated primarily with early-onset (30 to 60 years) familial Alzheimer's disease. Although St. George-Hyslop et al. (1987) reportedly located the cite of familial Alzheimer's disease on chromosome 21, these findings were not initially confirmed by other researchers. It was suggested then that familial Alzheimer's disease could

be genetically heterogeneous, and chromosome 21 may be responsible for only a subset of familial Alzheimer's disease cases (Schellenberg et al. 1991).

Schellenberg et al. (1991) conducted a study to support St. George-Hyslop et al's. (1987) findings. This study was designed to test the genetic linkage of familial Alzheimer's disease in 48 families. The families included in the study had 2 or more generations with either early-onset or late-onset (60 years or older) Alzheimer's disease. Also involved in the study were 7 Volga German families whose ancestors were immigrants from the Hesse region of Germany to Russia in the 1760's. These families then moved to the United States in late 1800's. This subset of families is thought to be genetically homogeneous for familial Alzheimer's disease. In their study Schellenberg et al. (1991) used several DNA markers including D21S16, D21S13, D21S110, and D21S1/S11. These markers were used to test for genetic linkage to the 21q11-q21 region of chromosome 21. The results that they obtained from their study indicated that as a sum the families did not show genetic linkage to 21q11.2-q21; however, some subsets of kindred did show genetic linkage to 21q11-q21. Schellenberg et al. (1991) analyzed the results of the late-onset kindred, their research indicated negative lod scores for all of the DNA markers used in the study. Early-onset non-Volga Germans were then analyzed as a subgroup, the lod scores for D21S13 and D21S16 markers were positive in some families. This data supports St. George-Hyslop et al's (1987) hypothesis that early-onset familial Alzheimer's disease is linked to 21q11-q21. The Volga German families showed congruous negative results for linkage to the 21q11-q21 region, supporting the hypothesis that the familial Alzheimer's disease locus for these kindred is not located on chromosome 21q11-q21 (Schellenberg, et al., 1991). Their research also supports

the hypothesis that 21q11-q21 may be responsible for only a subset of familial Alzheimer's disease cases (Schellenberg et al., 1991).

Interestingly, the kindred who show genetic linkage to chromosome 21 also have a point mutation in the *B*-Amyloid Precursor Protein (APP) gene. This gene is very close to the area on chromosome 21 where linkage to the familial Alzheimer's disease gene has been demonstrated (Goate et al., 1991). In their study Goate et al. (1991) analyzed the APP gene, beginning with the sequencing of exon 17, which indicated a cytosine to thiamine transition at base pair 1249. This transition causes a change in the amino acid 717 from valine to isoleucine. Chartier-Harlin (1991) reported a thymine to guanine transition at base pair 2150 which results in an amino acid change from valine to glycine at codon 717. In addition to the valine to isoleucine and valine to glycine shifts, there is also a valine to phenylalanine shift at codon 717. These missense mutations have been linked to some cases of early- onset Alzheimer's disease (Tanzi et al. 1992). Since these mutations all occur at the same codon, Chartier-Harlin et al. (1991) surmise that this could be a coincidence, although it is unlikely. The substitution of valine with isoleucine, glycine, and phenylalanine increase the amount of *B*-amyloid protein that is deposited due to a change in the APP processing, but the mechanism for deposition is not yet known (Citron et al., 1992).

It seems evident that the APP gene located on chromosome 21 has a vital role in familial Alzheimer's disease, but accounts for less than 1% of all Alzheimer's disease cases. Therefore, additional research needed to be conducted to exploit other chromosomes involved with familial Alzheimer's disease (Lyon & Gorner, 1995).

Researchers focused their attention on chromosome 19, linking the disorder to a specific gene, apolipoprotein E (APOE), located on chromosome 19q13.2. The APOE locus has 4 alleles: $APOE_2$, $APOE_3$, and $APOE_4$. People who have the $APOE_4$ allele seem to have an increased risk of Alzheimer's disease (Lyon & Gorner, 1995). As reported by Corder et al. (1993), 80% of familial and 64% of sporadic late onset Alzheimer's disease victims have at least one $APOE_4$. These statistics implicate $APOE_4$ as the cause of more than one half of all Alzheimer's disease cases. Corder et al. (1993) suggest that an increase in the gene dose of $APOE_4$ increases one's risk for Alzheimer's disease and also results in an earlier onset. The presence of each additional $APOE_4$ results in an increase risk of Alzheimer's disease by a "factor of 2.84". In contrast, Lyon and Gorner (1995) state that the $APOE_4$ gene is not responsible for Alzheimer's disease directly; rather, it is the lack of $APOE_3$ and $APOE_2$ that increase a persons vulnerability to Alzheimer's disease. This suggests that $APOE_2$ and $APOE_3$ have some protective factors. It is not clear what involvement APQE as in the nosogenesis of Alzheimer's disease; however, the protein, apoE isoform E, that is encoded by $APOE_4$ was found to be immunoreactive in plaques as well as tangles of Alzheimer's disease patients. These findings, as well as statistical analysis, suggest that $APOE_4$ is directly involved with Alzheimer's disease (Corder et al., 1993).

In addition to chromosome 21 and 19, chromosome 14 seems to have an association with familial Alzheimer's disease. In a study conducted by Schellenberg et al. (1992) genetic linkage of familial Alzheimer's disease on chromosome 14 was tested. In their research, Schellenberg et al. (1992) studied families consisting of 3 or more affected individuals spanning two or more generations. Their kindred also had at least one family member who

had died and had confirmed Alzheimer's disease based on autopsy results. These families were divided into three groups, Volga Germans, non Volga Germans, and Volga Germans and non-Volga Germans combined. DNA markers D14S1, D14S43, and D14S53 were tested and lod scores were calculated. The D14S43 and D14S53 markers gave positive lod scores for non-Volga German families. Conversely, the Volga German families showed negative lod scores for all of the DNA markers used in the study. Through this study it was suggested that an early-onset familial Alzheimer's disease locus was located on chromosome 14q24.3 and could account for a large percentage of early-onset familial Alzheimer's disease (Sherrington et al., 1992). Once the initial mapping of chromosome 14q24.3 was done the locus that was detected near D14S43 and D14S53 markers was given the name AD3 locus. The AD3 locus has been associated with an aggressive form of early-onset familial Alzheimer's disease. Through linkage mapping, Sherrington et al. (1992) were able to define the regions containing the AD3 gene and isolate approximately 19 transcripts that are encoded in this region. One transcript, S182, is correspondent to a novel gene who's product is thought to consist of 'multiple transmembrane domains' and is reminiscent of an integral membrane protein. The S182 transcript, which is 3.0kb, was sequenced revealing 5 different mutations which result in the following amino acid changes: Methionine-leucine at codon 146; histidine-arginine at codon 163; alanine-glutamic at codon 246; leucine-valine at codon 286; and finally cysteine-tyrosine at codon 410. These mutations occur independently, but all seem to result in early-onset familial Alzheimer's disease . Although there is strong evidence that the mutations in S182 gene results in early-onset familial Alzheimer's disease Sherrington et al. (1995) proposed two questions. First, what function does the gene have; secondly, "how do

these mutations confer a dominantly inherited familial Alzheimer's disease phenotype"?

Sherrington et al. (1995) suggest that the S182 protein is an integral membrane protein, such as a receptor, a channel protein, or possibly a structural membrane protein. Its function could possibly include the 'docking of other membrane bound proteins' or fusion budding of membrane bound vesicles during protein transport. If the function of S182 is as speculated, then the mutations found in S182 could result in abnormal transport or processing of APP, as well as aberrant interactions with cytoskeletal proteins. Although the functions and processes of S182 are not fully understood the fact that no mutations of S182 are present in normal controls indicates that it is highly probable that these missense mutations are responsible for AD3 (Sherrington et al., 1995).

Through all of the research that has been conducted on familial Alzheimer's disease, one group, the Volga German kindred, has shown consistent negative results as to genetic linkage on chromosome 21, 14, and 19. Volga Germans, along with some families in which familial Alzheimer's disease is inherited as an autosomal dominant trait, do not have a known loci. Volga German are German immigrants in Russia that did not intermarry with the surrounding population and remained culturally distinct. The average age of onset for Alzheimer's disease in Volga Germans is 50.2 to 64.8 years of age (Levy-Lahad et al., 1995). In a study conducted by Levy-Lahad (1995) chromosome 1 is explored for a possible connection to familial Alzheimer's disease. DNA from 139 individuals of Volga German kindred were prepared for study, 36 subjects had been clinically diagnosed with Alzheimer's disease. A total of 21 DNA markers were used in the study, D1S279 had the most significant lod score. Through further study it was discovered that the 112 bp allele was seen in the

affected subjects compared to the controls. This data localized the familial Alzheimer's disease locus to 1q31-q42. Levy-Lahad et al's (1995) next strategy was to try to find the mutations in the gene candidate for this locus. The first gene candidate which was tested encodes cathepsin E. Cathepsin E is a protease that is found in the brains of Alzheimer's disease patients in increased amounts. Interestingly, the second gene candidate that was tested is a homolog to S182, which was a candidate gene in AD3. An expressed sequence tagged of 475 bp was sequenced and compared to S182. This sequence, referred to as T03796, is 80.5% identical to S182. T03796 is located between D1S439 and D1S479 markers on chromosome 1, this is consistent with their previous finding of a 112 bp allele observed in affected kindred. The gene corresponding to the T03796 sequence, referred to as the E5-1, gene was cloned and analyzed (Rogaev et al., 1995). In Levy-Lahad et al's (1995) work the E5-1 gene is referred to as F (Full length)-T03796. The results indicate that E5-1 has a 448 amino acid open reading frame that is homologous to S182. The E5-1 gene was then analyzed for mutations, revealing two. The first mutation resulted from a adenine-guanine substitution at nucleotide 1080 resulting in a methionine-valine change in codon 239. Secondly, a change in nucleotide 787 from adenine to thymine was seen at codon 141, this results in an amino acid change from asparagine to isoleucine. This second mutation is referred to as N141I; interestingly, all of the subjects that had the N141I mutations also had the 112 bp allele. These mutations were not found in normal controls (Levy-Lahad et al., 1995).

The predicted protein from E5-1 has seven transmembrane domains and is called STM2 ("the second seven transmembrane gene associated with Alzheimer's disease"). The

similarities between E5-1 and S182 are amazing, suggesting that the mutations in these genes could possibly be negatively affecting the attachment or anchoring of their proteins into membranes. Although the normal cellular function of E5-1 and S182 are not yet known, it is possible that they have similar roles in cellular function and their mutations may hinder normal cellular function (Levy-Lahad, 1995).

Rogaev et al. (1995) suggest that since the E5-1 and S182 genes are so similar, sharing a 63% similar amino acid sequence, they may be members of a family of genes called presenilins. According to Strooper et al. (1998) presenilins promote a "proteolytic activity that cleaves the integral membrane domain of APP." When there is a mutation there is an increased production of "amyloidogenic peptide amyloid" which results in the manifestation of Alzheimer's disease.

The research that has been conducted on chromosomes 21, 19, 14 and 1 offers a promising future for the detection and treatment of Alzheimer's disease. There are currently few treatments for Alzheimer's disease. One possible treatment involves the use of nerve growth factor (NGF), which is a protein involved in the growth of nerve cells (Lyon & Gorner, 1995). In studies conducted in rats and monkeys NGF is injected into the brain. The results indicate that nerve cell destruction in the "acetylcholine pathway of the basal nucleus and hippocampus had been stopped." NGF stopped the shrinkage of nerve calls and also promoted new connections between nerve cells (Lyon & Gorner, 1995).

Although the results seem positive, researchers are being cautious in the treatment of Alzheimer's disease with NGF. The hypothesis is that NGF may make the disease worse, specifically, it is thought that NGF may promote the growth of nerve cells, strangling the

existing cells. The formation of new nerve cells could possibly lead to more *B*-Amyloid deposition in the brain (Lyon & Gorner, 1995).

Although a great deal of progress has been made in understanding Alzheimer's disease, its origin and processes, there is still much that remains a mystery. Hopefully there will be a day when Alzheimer's disease is referred to as a disease of the past, not the present and future, and the suffering will end for the victims of this disease as well as their families.

CHAPTER 3: BREAST CANCER

Breast cancer is a common disease that affects 180,000 women per year and results in approximately 50,000 deaths each year in the United States alone (Lyon and Gorner, 1995). The disease, which primarily affects women and is mostly sporadic, having no genetic linkage, is estimated to develop in one out of eight women who live to the age of 95 (Miki et al., 1994). Early detection is essential for proper medical management of the disease. If breast cancer is diagnosed in more advanced stages, treatments may cause disfigurement or not be successful at all. There is evidence that suggests that genetic factors do contribute to a small percentage of the disease. Miki et al. (1994) estimate that approximately 5% of all breast cancer cases are a result of genetic predisposition. These familial type cases of breast cancer are thought to result from two different genes, BRCA1 (Breast Cancer 1, located at 17q12-q21) and BRCA2 (Breast Cancer 2, located at 13q12-q13), which result in approximately 85% of all familial breast cancer cases. BRCA1 has also been implicated in some sporadic cases of breast cancer, as well as the early-onset of ovarian cancer (Korenberg et al., 1995). Genetic predisposition results in a higher risk of developing the disease. The percentage of risk for those inheritably susceptible to breast cancer are as follows: 37% by

age 40, 66% by age 55, and an 82% risk over ones entire lifetime. In contrast, the risk of developing breast cancer in non-genetically predisposed persons are as follows: .4% by age 40, 2.8% by age 50, and an 8.1% risk over an entire lifetime (Hall et al., 1990).

BRCA1 and BRCA2 both contribute to the early-onset (onset before 50 years of age) of breast cancer. It is estimated that BRCA1 contributes to 45% of familial breast cancer, leaving the other 55% as a result of BRCA2 and also some sporadic factors (Miki et al., 1994).

BRCA1 was first mapped to chromosome 17 by Hall et al. (1990). In their study a total of 329 participants from 23 families were analyzed. In these families there were 146 reported cases of breast cancer. The families shared characteristics that are seen in familial breast cancer cases including early-onset of the disease, and an increased number of men with the disease. Hall et al. (1990) used various DNA markers in their study, including D17S74. D17S74 had the highest lod score of all of the DNA markers and is located on chromosome 17q21. Additional research isolated BRCA1 to a 1 to 2 megabase region of chromosome 17. Through further research BRCA1 was compressed further to a 600 kb region between D17S1321 and D17S1325 markers (Miki et al., 1994). Within the 600 kb region 65 candidate expression sequences were recognized, these expressed sequences were analyzed through DNA sequencing. Through this process BRCA1 was narrowed down to a small area in the 600 kb region previously mentioned and is associated with the DNA marker D17S855 (Miki et al., 1994). In the next stage of analysis cDNA of the full length of BRCA1 was constructed, and upon evaluation one long open reading frame encoding a protein consisting of 1863 amino acids and 22 exons was discovered (Miki et al., 1994). Also included in BRCA1 are two zinc finger domains and an acidic COOH-terminal domain (Chen et al., 1995).

The structure of BRCA1 was beginning to unravel, it was necessary to find specific mutations in BRCA1 that may alter the function of the gene. Initial research by Futreal et al. (1994) discussed three mutations found on the BRCA1 gene. The first mutation is glutamic acid to Stop at nucleotide 1541, this stop codon creates a truncated protein which is missing 323 amino acids. The second and third mutations are missense mutations, alanine to glutamic acid at nucleotide 1708 and methionine to arginine at nucleotide 1775. These mutations are all absent in normal population controls. Interestingly, Futreal et al. (1994) did not discover any sporadic mutations on BRCA1. They suggest that some mutations may have been missed in the screening process or possibly some sporadic mutations are located in noncoding sequences. Gayther et al. (1996) report that over 65 specific mutations have been detected scattered throughout the BRCA1 gene with clusters of mutations in various areas. In their study, Gayther et al. (1996) used a technique called multiplex heteroduplex analysis (MHA) to search for mutations that occur frequently with a screening process that requires only a single step. In the study 162 families with a history of breast and/or ovarian cancers from various geographic locations were analyzed. Mutations from 35 families were identified, the majority of which were frame shift mutations resulting from insertions or deletions. Although the region of BRCA1 analyzed by MHA was only one-quarter of the entire coding sequence, 50% of the mutations that have been reported to date are found in this region. This study also indicated that specific geographic and ethical origins influence BRCA1 mutations. For example, an insertion of cytosine at nucleotide 5382 was seen in 9 out of 19 families from Russia. A deletion of adenine and guanine at nucleotide 185 is a common mutation in the Ashkenazi Jewish population. MHA is a technique that has a promising future in the

detection of mutations of the BRCA1 gene; however, this procedure is not sensitive enough to be used for testing on a clinical basis (Gayther et al., 1996).

Miki et al. (1994) speculate that BRCA1 is a tumor suppressor gene, which is a protein that functions to negatively regulate tumor growth. Familial breast cancer is inherited as a dominant trait; therefore, it is only necessary to inherit a single copy of a mutant allele to be genetically predisposed to the disease. In BRCA1 the wild type allele is overcast by a dominant mutant allele, this disrupts the normal function of BRCA1 as a tumor suppression gene (Miki et al., 1994).

BRCA1 was found in all breast cell lines which were tested. In normal breast cells there were no mutations to BRCA1. Chen et al. (1995) located BRCA1 in normal breast cells primarily in nuclei. In the breast cancer cells BRCA1 was detected primarily in the cytoplasm and also appears to be completely mislocated in most end-stage breast cancer cells. The mislocation of BRCA1 suggests that when germ-line mutations alter BRCA1, these abnormalities help in the initiation and development of breast cancer (Chen et al., 1995).

As previously stated, BRCA1 was originally thought to be a tumor suppression gene (Miki et al., 1994). However, Marx (1997) indicates that new research suggests BRCA1 maybe involved in DNA repair, and any mutations in the gene may disrupt its ability to function properly.

In 1994 Wooster et al. (1994) performed a study trying to genetically link 15 high-risk breast cancer families to BRCA1. These families did not show any linkage to chromosome 17q21. Through their research a second "breast cancer susceptibility locus" was discovered. BRCA2 located on chromosome 13q12-q13 indicated a high risk to breast cancer, but not

ovarian cancer as seen in BRCA1. The BRCA2 gene was mapped using various DNA markers including D13S260 and D13S263. Using two-point and multipoint lod scores the BRCA2 gene was narrowed down to the area of DNA marker D13S171 encompassing 600 kb. The cDNA of BRCA2 was analyzed revealing an open reading frame which encodes a protein consisting of 3418 amino acids and 27 exons (Tavtigian et al., 1996)

In further research Wooster et al. (1995) detected several germ-line mutations in breast cancer families. These mutations altered the open reading frame of BRCA2, hindering its suspected function as a tumor suppression gene. In a study conducted by Weber et al. (1996) various BRCA2 germ-line mutations were identified including a 1 bp deletion in exon 11 at nucleotide 2882, the addition of 9 amino acids, and a "translocation termination at codon 894". Schubert et al. (1997) suggest that most mutations in BRCA2, which alters its normal function and cause disease, are a result of premature Stop codons.

As in BRCA1, geographic and ethnic origins influence mutations that are observed. For example, Thoracius et al. (1996) identified a 5 bp deletion in exon 9 of BRCA2 that has been seen in 16 families of Icelandic origin. Schubert et al. (1997) report that a deletion of a tyrosine at nucleotide 6174 is a common mutation among Ashkenazi Jewish families. Interestingly, unlike BRCA1, there is a relatively high occurance of male breast cancer associated with BRCA2. In a study conducted by Friedman et al. (1997) male breast cancer patients were analyzed to see how frequent BRCA1 and BRCA2 germ-line mutations occurred. The results they obtained indicated no mutations on the BRCA1 gene. There were 2 BRCA2 germ-line mutations that resulted in truncated proteins. Although only 1% of all breast cancer cases are seen in men, and an even smaller percentage is due to germ-line

mutations in BRCA2, it is vital to try to resolve the influence of BRCA2 on breast cancer for both males and females. Interestingly, there was also an "ovarian-cancer cluster" identified on the BRCA2 gene. In previous reports it has been suggested that BRCA2 does not influence the onset of ovarian cancer (Friedman et al., 1997).

In BRCA2, as with BRCA1, there are many mutations that have been seen in various families. The major question is, what is the role of BRCA1 and BRCA2 with relation to tumor development? Initially, they were put into the category of tumor suppressors, or gatekeepers (Kinzler and Vogelstein 1997). Gatekeepers as described by Kinzler and Vogelstein (1997) are genes which monitor tumor growth by inhibiting growth of the tumor or promoting death. Inactivation or mutations of these gatekeepers lead to specific cancers, in this case breast cancer. The other type of gene proposed by Kinzler and Vogelstein (1997) are caretakers. These genes influence the cell by insuring genetic stability of the cell, making repairs to DNA. Mutations in the caretakers result in the loss of genetic integrity which leads to the occurance of tumors. Kinzler and Vogelstein (1997), along with Marx (1997), propose that BRCA1 and BRCA2 are caretakers whose function is to repair DNA, keeping cell growth in check. When these caretakers are altered by mutations they can not properly function. It is proposed that BRCA1 and BRCA2 are caretakers because mutations of these genes are uncommon in sporadic cases of breast cancer. It is important to determine if BRCA1 and BRCA2 act as gatekeepers, or the more recently supported function, caretakers, for practical reasons as well as theoretical ones. If it is determined that BRCA1 and BRCA2 are definitely caretakers, there are additional avenues of treatment that can be pursued (Kinzler and Vogelstein, 1997). For example, it was determined that BRCA2 genes are

"sensitive to gamma-irradiation". Thus, patients with BRCA2 mutations could have favorable results with such treatments (Sharan et al, 1997).

As an interesting side note, BRCA1 has been connected to prostate cancer, a common form of cancer among men. Langston et al. (1996) assessed the role of BRCA1 germ-line mutations on the pathology of prostate cancer. BRCA1 was analyzed for germ-line mutations in a group of men with the disease. One specific mutation, a frame shift mutation in exon 2 resulting in a premature Stop codon in BRCA1, had previously been described and was also observed in their study. If this mutation proves to be common among men with a family history of prostate cancer, a familial connection to the disease may be identified (Langston et al., 1996). Although the frequency of germ-line mutations in BRCA1 which lead to prostate cancer is small, the ability to undergo genetic counseling and possibly isolate one's risk at an early age could aid in the treatment of the disease (Langston et al., 1996).

With the advent of BRCA1 and BRCA2 and their influence on familial breast cancer, genetic counseling for men and women with a family history of the disease has increased. In the past, women who had many members of their families diagnosed and eventually die from breast cancer, considered having full mastectomies although they were not diagnosed with the disease. Today there is increased technology and genetic counseling techniques to help people assess their risk and make decisions about their course of treatment (Lyon and Gorner, 1995). As more and more information about the normal function of BRCA1 and BRCA2 becomes apparent, scientists will be able to compare this information with what is known about the function of mutated BRCA1 and BRCA2. Scientists will also be able to observe what the most common germ-line mutations in BRCA1 and BRCA2 for various

geographic and ethnic groups are. Also, what effects specific mutations in these genes have on the initiation and development of breast cancer. In conjunction with continued analysis of the BRCA1 and BRCA2 genes, genetic counseling techniques will become more refined to evaluate risk assessment and aid in the early diagnosis of the disease. The ability to assess one's risk and diagnose the disease in its early stages will facilitate the treatment of some forms of breast cancer projecting a more successful outcome.

CHAPTER 4: HUNTINGTON DISEASE

In 1872 George Huntington, a Long Island physician, published a paper, "On Chorea". In his article Huntington first described the disease we today call Huntington Disease (Conneally, 1984). Huntington's description of the disease included characteristics such as a late-onset disorder that didn't afflict patients until mid-life. In his description he also used such words as "bowing", "twisting" and "grimacing" as he observed patients with the disease (Lyon and Gorner, 1995). Today, Huntington Disease is described as a progressive disorder afflicting 5-10 in 100,000 persons. The disease affects motor function, personality and cognition, eventually resulting in death (Brinkman et al., 1997). From the time of onset of the disease to death, the average duration of Huntington Disease is 15-20 years (Brinkman et al., 1997). In general, the first symptoms of Huntington Disease are seen between the third and fifth decade of life. The symptoms that are associated with Huntington Disease are a result of premature death of neural cells, primarily in the basal ganglia (Gusella et al., 1983).

Nancy Wexler, a clinical psychologist, and a potential Huntington Disease carrier, formed the Huntington Disease Collaborative Research Center in 1984, which is a group of researchers that would work collectively to try to find the genetic cause of Huntington

Disease. Wexler, whose mother and several other relatives died from Huntington Disease, was intent on promoting research efforts to find the genetic cause and potential treatment of Huntington Disease. Wexler's passion has lead her to organize medical expeditions to Venezuela every year since 1981. Specifically, three villages in Venezuela were the targets of the expeditions. San Luis, Barranquitas and Lagueneta have the highest frequency of Huntington Disease cases in the world. During these expeditions blood samples were taken from the villagers and a detailed lineage of the disease was formulated. The results that were obtained linked Huntington Disease back to the early 1800's to a woman appropriately named Maria Concepcion. The mystery of the origin of Huntington Disease was beginning to unravel, but the genetic basis of the disease was still not understood (Lyon and Gorner, 1995).

In 1983 Gusella et al. (1983) conducted a study to find a genetic marker associated with Huntington Disease. Through their research they isolated a 17,000 nucleotide patch of DNA that could be associated with the gene that results in Huntington Disease (Lyon and Gorner, 1995). Through further research twelve polymorphic DNA markers were used to screen for linkage to the Huntington Disease gene. One marker, G8, also referred to as D4S10, had the highest lod score and corresponded to chromosome 4 (Gusella et al., 1983).

Gusella et al. (1983) then used a technique called in-situ hybridization to isolate the location of the Huntington Disease gene on chromosome 4. The probe used was the G8 DNA marker which hybridized to the short arm of chromosome 4, narrowing down the location of the gene. However, this technique was not sensitive enough to indicate the precise location of the Huntington Disease gene.

Interestingly, in Gusella et al's. (1983) research blood samples of both American and Venezuelan kindred were used to isolate the location of the Huntington Disease gene. In both pedigrees the G8 marker had the highest lod score and hybridized to the short arm of chromosome 4. This data indicates that the gene responsible for Huntington Disease may be the same universally with no geographic or ethnic influences (Gusella et al., 1983).

In a study by Wang et al. (1986) a more precise location of the Huntington Disease locus was reported. In their research in-situ hybridization techniques were also employed. The results indicate that the Huntington Disease locus lyes within 4p16.1-4p16.3 of chromosome 4 (Wang et al., 1986). In further attempts to narrow down the location of the Huntington Disease gene, Gilliam et al. (1987) report that by using multipoint analysis techniques, D4S10 was more refinely mapped to 4p16.3 of chromosome 4.

Since the first DNA marker corresponding to the Huntington Disease gene was isolated by Gusella et al. (1983), many other researchers have tried to find additional markers associated with the gene. Gilliam et al. (1987) identified D4S43, another polymorphic DNA marker, which could potentially be closer to the Huntington Disease gene than D4S10. Using Multipoint analysis the two markers, D4S10 and D4S43, were tested. The results indicate that D4S43 is closer to the Huntington Disease locus than D4S10. This finding was important and implicated D4S43 as a possible candidate for use in predictive testing assays (Gilliam et al., 1987).

Research by Robbins et al. (1989) suggests that two more DNA markers that are even closer to the Huntington Disease gene were discovered. Using somatic hybrid cell mapping and genetic linkage analysis the two markers, D4S95 and D4S90, were detected on the distal

portion of chromosome 4 having a close linkage to the Huntington Disease gene. Robbins et al. (1989) state that the D4S95 marker is the most useful of the two markers and could be used for predictive testing for Huntington Disease (Robbins et al., 1989).

While other researchers were trying to find the location of the gene responsible for Huntington Disease, Myers et al. (1985) was trying to identify factors that influence the age at onset of the disease. In their study, Myers et al. (1985), researched parental factors than may influence the age at onset. In their report they state that the sex of the affected parent is a factor that influences the onset of Huntington Disease. They suggest that when the gene is inherited from the father, the age at onset is earlier than when the gene is inherited from the mother (Myers et al., 1985).

There was a great deal that was known about Huntington Disease, including factors influencing age at onset as well as the location of the gene which causes the disorder. It was now time to try to figure out what mutation(s) influence the expression of the disease.

In 1993 Gusella's lab was working feverishly to clone the 200,000 base pair Huntington Disease gene in order to find possible mutations. In February, Marcy MacDonald, Gusella's senior postdoc, struck gold. MacDonald finally cloned the gene and had sequenced a portion of it. While analyzing the sequence MacDonald noticed something unique. There was a sequence of forty eight repeating cytosine-adenine-guanine (CAG) nucleotide sequences. When compared to the sequence of a normal control, MacDonald found that the affected population had significantly more trinucleotide repeats compared to the unaffected controls. Gusella and his colleagues (those that make up the Huntington Disease Collaborative Research Center) performed numerous assays sequencing the DNA of

Huntington Disease victims. The results indicated that they all had the same phenomena, a gene that always began with a long series of CAG repeats (Lyon and Gorner, 1995). It was as 'simple' as that, a stutter in the gene. Further research indicated that the CAG repeats found on exon one are found in all individuals, affected and unaffected. The number of CAG repeats in unaffected individuals range from 6 to 35 with only 1% of unaffected individuals having between 30 and 35 repeats. In contrast, affected individuals have CAG repeats ranging from 36 to 121 (Brinkman et al., 1997). In order to identify the number of CAG repeats an individual has, a technique which uses a Polymerase Chain Reaction (PCR) protocol was developed. This technique is what is currently used for predictive testing and diagnosis of Huntington Disease. Interestingly, it has been shown that a significant inverse relationship has been seen between the number of CAG repeats and the age at onset of Huntington Disease (Brinkman et al., 1997). In research conducted by Brinkman et al. (1997), the goal was to estimate one's probability of developing Huntington Disease by a specific age based on the number of CAG repeats one has. In their study a large number of individuals with varying numbers of CAG repeats were analyzed. Through their research it was shown that the number of CAG repeats a person has is the major factor of the age at onset of Huntington Disease. It was once thought that Huntington Disease was 100% penetrant, now through their study Brinkman et al. (1997) show that in persons with less than 42 CAG repeats the risk of penetrance is not 100%. Therefore, if a person inherited between 36 and 42 CAG repeats, they may be spared the unforgiving fate of Huntington Disease. In their study they found that individuals with 50 CAG repeats had a median age at onset of 27, whereas individuals with 43 CAG repeats had a median age at onset of 44. Brinkman et al.

(1997) caution against the use of the number of CAG repeats as a predictive tool for age at onset because it varies from individual to individual.

The CAG repeats produce the amino acid glutamine, the protein which the gene codes for is called huntingtin. The increased number of CAG repeats results in the formation of more glutamine, thus an expansion of huntingtin. The discovery of CAG repeats and their influence on the expression of Huntington Disease has not enabled researchers to understand what causes the nerve cells of the brain to sicken and ultimately die. Scientists would first need to understand the normal function of the gene before they could determine how the mutant gene functions. In a report by Williams (1996) a possible connection was made. Huntingtin interacts with a protein called glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which is a key enzyme involved in glycolysis. The theory is that the mutated form of huntingtin interferes with the normal function of GAPDH and stunts energy production. It is suggested that since the brain depends predominately on energy produced from the breakdown of glucose, the disabling of a key component of glycolysis could harm and eventually destroy nerve cells in the brain by interfering with their source of energy (Williams, 1996).

Scientists now have the ability to provide accurate predictive testing procedures as well as determine, roughly, the age of onset of Huntington Disease. However, this is all they know. There is no treatment for Huntington Disease, the biochemical mechanisms that contribute to the disease are not known. It is an ironic dilemma, to be able to confirm that someone will someday begin to show the symptoms of Huntington Disease, but to also inform them that there is nothing that can be done to alter the fate that lies before them.

Initially, predictive testing was done with polymorphic DNA markers that were discovered before the gene responsible for Huntington Disease was isolated. There were certain risks associated with this type of predictive testing. For example, when only one polymorphic DNA marker, D4S43, was used for testing the results were not very accurate; whereas, when more polymorphic DNA markers were discovered the predictive test became more accurate. In one case, a man with a family history of Huntington Disease was tested when D4S43 was used for predictive testing. The man was told that he had a 17% risk of inheriting the gene responsible for Huntington Disease. However, when the predictive tests were improved due to the discovery of more markers, he was retested and told that his risk for inheriting the Huntington Disease gene was 96%. Although predictive testing is an effective tool for predicting risk, it may be advantageous to wait until more is known about the genetic basis of a disease before testing for inheritance. In another case, a woman was told that she had an 85% risk of inheriting the gene for Huntington Disease. When the gene for Huntington Disease was isolated she was retested. The test indicated that she had 35 CAG repeats and was told that it was very unlikely that she would develop Huntington Disease because her CAG repeats were within normal limits (Almqvist et al., 1997).

Another controversy with predictive testing for Huntington Disease is whether children below the age of 18 should be tested. Bloch and Hayden (1990) state that guidelines established for predictive testing programs allow only individuals over 18 years of age to be tested. In the case of Huntington Disease some parents request that their children have predictive testing done, primarily because of their own curiosity and anxiety. Bloch and Hayden (1990) argue that children cannot make decisions about predictive testing, and the

only way it should be done is if it can be demonstrated that it is for the child's benefit. However, since there is currently no treatment for the prevention of Huntington Disease, predictive testing for a child would not be for his or her benefit. In contrast, if a child knows that they will eventually develop Huntington Disease, it may influence their self esteem and their ability to interact socially and is not at all recommended (Bloch and Hayden, 1990).

Predictive testing is beneficial for individuals who want to know their fate for personal reasons. For example, it may influence a couples decision on whether or not to have children.

Although there is no treatment for Huntington Disease, the information that has been discovered in recent years is very encouraging. With the continuation of research for this disease, hopefully someday there will be a course of action that scientists and physicians can take to combat Huntington Disease.

CHAPTER 5: SCHIZOPHRENIA

Schizophrenia is a disease which affects approximately 1% of the population of the United States. Symptoms of the disease include hallucinations, paranoia, delusions, disrupted thought processes and frequent inability to control emotions. It is thought that several factors influence the onset of the disease, including inheritance and environmental factors (Pulver et al. 1994).

The strongest evidence that schizophrenia is indeed inherited comes from adoption and twin studies. In studies involving schizophrenics that were adopted into families from the general population, it has been shown that the adopted children develop schizophrenia even though they are not raised by schizophrenic families. In twin studies, monozygotic twins showed a higher concordance, having both been diagnosed with schizophrenia, than dizygotic twins. These twin and adoption studies support the hypothesis that schizophrenia is heritable (Carlson, 1991).

Although it is thought that genetic factors are involved in the pathogenesis of schizophrenia, it is not clear how it is inherited. There has been extensive research to try to find a locus for schizophrenia; however, there are many variables that make this difficult.

Primarily it is hard to develop family lineage for the disease, because in many cases schizophrenics do not reproduce. Schizophrenics are also less likely to volunteer to be part of a study about the disease. There is also a lack of understanding of the genetic mechanism that results in biochemical changes associated with schizophrenia. A great deal of research has been conducted to try to find potential chromosomal linkage to schizophrenia and develop a better understanding of its inheritance (Pulver et al., 1994).

In a study by Pulver et al (1994) analysis of 39 families, all diagnosed with schizophrenia based on criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R), was conducted. In the study 240 polymorphic DNA markers that are randomly distributed over the genome were used to identify potential linkage to schizophrenia. Two markers gave the highest lod scores, D22S268 and D22S307 spanned approximately 23 centi-Morgans (cM) from 22q12-q13.1. This data seemed very promising, but in order to obtain more concrete results the study needed to be replicated. Pulver et al. (1994) conducted a follow-up study to try to replicate linkage of schizophrenia to chromosome 22q12-q13.1. In this study 217 families from various geographic locations were analyzed. Again, all subjects in the study met DSM-III-R criteria for schizophrenia. Likelihood of linkage of two polymorphic DNA markers, D22S268 and D22S307, was based on results of linkage analysis of chromosome 22. The results of this study indicated that there was no linkage of these markers to chromosome 22q12-q13.1 for the 217 families. These results do not support the initial research of Pulver et al. (1994) that suggested possible linkage of schizophrenia to chromosome 22q12-q13.1 (Pulver et al., 1994).

In further research conducted by Coon et al (1994) linkage analysis using 329 polymorphic DNA markers was conducted. In this research 9 families were analyzed, the linkage analysis revealed high lod scores for markers associated with chromosome 4, D4S35; chromosome 14, D14S17; chromosome 15, D15S1; and chromosome 22, D22S84 and D22S55. According to Conn et al. (1994) the regions that have been identified in this study are worthy of additional research to try to isolate a specific schizophrenia locus. Coon et al. (1994) also suggest that there may be several genes that interact and are collectively responsible for the onset of schizophrenia.

Although the chromosomal location which causes, or triggers, the onset of schizophrenia is not known, researchers have an understanding of how schizophrenia functions neurologically. An initial hypothesis suggested that schizophrenia is a result of overactive dopaminergic synapses, suggesting that schizophrenics have an abnormal amount of dopamine, a neurotransmitter, in their brain. However, new research by Moghaddam and Adams (1998) reports that there is evidence that the onset and progression of schizophrenia is due to glutamatergic mechanisms. In schizophrenics there is a higher than normal level of glutamate, another neurotransmitter. If the higher than normal glutamate levels can be reduced and controlled, there may be positive behavioral changes. Currently, neuroleptics, which is medication given to schizophrenics, function by disabling the action of dopamine. This treatment does help to reduce such symptoms as paranoia and hallucinations; however, it does not give much relief to other symptoms of the disease including cognitive and emotional instability (Wickelgren, 1998). In Moghaddam and Adams' (1998) study, a new drug, LY354740, was used in rodent trials to see if it influenced the symptoms of

schizophrenia in rats. LY354740 functions by stimulating a "subgroup of metabotropic glutamate receptors" which act to regulate glutamate levels. In rodent trials it has been observed that LY354740 regulates the output of glutamate when levels of the neurotransmitter become to high; however, LY354740 does not influence glutamate receptors when levels of the neurotransmitter are normal. In the rodent experiments, phencyclidine (PCP), a psychoactive drug which "induces schizophrenic-like symptoms" and causes an increase in levels of glutamate transmission, was administered to control and experimental rats. In the control group the PCP caused glutamate levels to rise and rats developed symptoms of schizophrenia; frantic running, incessant head twisting and jerking. In the experimental group LY354740 was administered as well as PCP. In this group there was no elevation of glutamate and there were no signs of schizophrenic behavior. The experimental group was also able to successfully navigate mazes and showed the same cognitive ability as the normal group of rats. In rodent trials LY354740 appears to be very promising, but many drugs that are effective in animal trials fail in human trials. In order to see if LY354740 is effective in humans, it needs to undergo testing on schizophrenia patients (Moghaddam and Adams, 1998).

Although the genetic basis of schizophrenia in not yet understood, new approaches to treat the disease appear to have a promising future. If schizophrenia patients can get relief from the unforgiving symptoms of schizophrenia, it would be a milestone in the treatment of the disease.

CONCLUSION

Identifying genes that are responsible for the onset of particular diseases is an overwhelming accomplishment. However, as the technology for identifying genes improves, so does the potential for genetic discrimination. When James Watson announced the plans to sequence the Human Genome, he allocated funds to research the Ethical, Legal and Social Implications (ELSI) of the project. The ELSI work group was developed and is responsible for evaluating issues that are associated with the project, such as privacy, genetic discrimination, and health care. In an article by Hudson et al (1995) the issue of genetic discrimination involving health care was discussed. In some cases patients were denied health coverage based on their genetic susceptibility to a particular disease. It is frightening for people to consider loosing their health coverage based on their genetic predisposition. In some instances patients undergo genetic testing secretively, so that their insurance carriers will not find out. In some states laws have been passed which protect individuals from being denied health insurance based on genetic information; however, such laws need to exist throughout the nation (Hudson et al., 1995). Interestingly, it has been recommended that laws be implicated to prevent positive discrimination. This would discourage insurance

companies from giving low risk clients lower insurance rates. This type of positive discrimination could lead to a societal division of those with "good" DNA versus those with "bad" DNA (Doby, 1995). Genetic discrimination is not limited to the health care industry; it could venture into many areas of society such as the work place and schools. Nancy Wexler, the leader of the ELSI work group, states that "Genetic information itself is not going to hurt the public, what could hurt the public is existing social structures, policies and prejudices." (Lyon and Gorner, 1995). Ours is a society where discrimination and breach of privacy issues seem to be the norm, rather than the exception. That is why it is imperative that as new discoveries are made to advance the Human Genome Project, the ELSI work group stays in stride to keep the project in check, strengthening its moral backbone.

LITERATURE CITED

Almqvist, E., Adam, S., Bloch, M., Fuller, A., Welch, P., Eisenberg, D., Whelan, D., Macgregor, D., Meschino, W. And Hayden, M. (1997). Risk reversals in predictive testing for Huntington Disease. American Journal of Human Genetics, 61, 945-952.

Bradbury, N. A., Jilling, T., Berta, G., Sorscher, E. J., Bridges, R. J., & Kirk. K.L. (1992). Regulation of plasma membrane recycling by CFTR. Science, 256, 530-531.

Brinkman, R., Mezei, M., Theilmann, J., Almqvist, E. and Hayden, M. (1997). The likelihood of being affected with huntington disease by a particular age, for a specific cag size. American Journal of Human Genetics, 60, 1202-1210.

Bloch, M. and Hayden, M. (1990). Opinion: Predictive testing for Huntington Disease in childhood: Challenges and implications. American Journal of Human Genetics, 46, 1-4.

Carlson, N. (1991). Physiology of Behavior. Boston, Massachusetts: Allyn and Bacon.

Chartier-Harlin, M.S., Crawford, F., Houlden, H., Warren, A., Hughes, D., Fidani, L., Goate, A., Rosser, M., Rogues, P., Hardy, J. & Mullan, M. (1991). Early-onset Alzheimer's disease caused by mutations at codon 717 of the *B*-amyloid precursor protein gene. Nature, 353 844-845.

Check, W. A. (1989). The encyclopedia of health: Alzheimer's disease. New York, New York: Chelsea House Publishers.

Chen, Y., Chen, C. F., Riley, D. J., Albert, C., Chen, P. L., Von Hoff, D., Osborne, C. K., Lee, W. H. (1995). Aberrant subcellular localization of brca 1 in breast cancer. Science, 270, 789-791.

Citron, M., Oltersdorf, T., Haass, C., McConlogue, L.,Hung, A.Y., Seubert, P., Vigo-Pelfrey, C., Lieberburg, I. & Selkeo, D. (1992). Mutation of the *B*-amyloid precursor protein in familial Alzheimer's disease increases *B*-protein production. Nature, 360, 672-674.

Collins, F.S., & Galas, D, (1993). A new five-year plan for the U.S. human genome program. Science, 262, 43-46.

Collins, F. S. (1992) Cystic fibrosis: Molecular biology and therapeutic implications. Science, 256, 774-779.

Collins, F. S., Guyer, M.S., & Chakravarti, A. (1997). Variations on a theme: Cataloging human DNA sequence variation. Science, 278, 1580-1581.

Collins, J. E., Cole, C. G., Smink, L. J., Garrrett, C. L., Leversha, M. A., Soderlund, C. A., Masien, G. L., Evertt, L. A., Rice, K. M., Coffey, M. J., Gregory, S. G., Gwilliam, R., Dunham, A., Davies, A. F., Hassock, S., Todd, C. M., Lehrach, H., Hulsebos, J. M., Weissenbach, J. Morrow, B., Kucherlapati, R. S. Wadey, R., Scambler, P. J., Kim, U. J., Simon, M. I., Peyrard, M., Xie, Y., Carter, N. P., Durbin, R., Dumanski, J. P., Bentley, D. R., & Dunham, I. (1995). A high-density YAC contig map of human chromosome 22.Nature, 377, 367-379.

Conneally, P. (1984). Huntington disease: Genetics and epidemiology. American Journal of Human Genetics, 36, 506-526.

Coon, H., Jensen, S., Holik, J., Hoff, M., Worsley, M., Reimherr, F., Wender, P., Waldo, M., Freedman, R., Leppert, M. and Byerley. (1994). Genomic scan for genes predisposing to schizophrenia. American Journal of Medical Genetics, 54, 59-71.

Corder, E.H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., Roses, A. D., Haines, J. L. & Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein e type 4 allele and the risk of alzheimer's disease in late onset families. Science, 261, 921-923.

Dankert-Roelse, J. E. & Meerman, G. J. (1997) Screening for cystic fibrosis: Time to change our position. The New England Journal of Medicine, 337, 997-999.

Duby, J. (1995). Genetic Discrimination. Science, 270, 1282-1283.

Friedman, L. S., Gayther, S. A., Kurosaki, T., Gordon, D., Noble, B., Casey, G., Ponder, B. A. and Culver, H. A. (1997). Mutation analysis of brca1 and brca2 in a male breast cancer population. American Journal of Human Genetics, 60, 313-319.

Futreal, P, A., Liu, Q., Eidens, D., Cochran, C., Harshman, K., Tavtigan, S., Bennett, L. M., Strano, A. H., Swensen, J., Miki, Y., Eddington, K., McClure, M., Frye, C., Feldhaus, J., Ding, W., Gholami, Z., Soderkvist, P., Terry, L., Jhanwar, S., Berchuck, A., Iglehart, J. D., Marks, J., Ballinger, D., Barrett, J. C., Skolnick, M. H., Kamb, A. and Wiseman, R. (1994). BRCA1 mutations in primary breast and ovarian carcinomas. Science, 266, 120-122.

Gayther, S. A., Harrington, P., Russell, P., Kharlevich, G., Garkavtesa, R. F., Ponder, B., A., and the UKCCCR Familial Ovarian Cancer Study Group. (1996). Rapid detection of regionally clustered germ-line brca1 mutations by multiplex heteroduplex analysis. American Journal of Human Genetics, 58, 451-456.

Gilliam, T., Bucan, M., MacDonald, M., Simmer, M., Haines, J., Cheng, S., Pohl, T., Meyers, R., Whaley, W., Allitto, B., Farynlarz, A., Wasmuth, J., Frischauf, A., Conneally, M., Lehrach, H. And Gusella, J. (1987). A dna segment encoding two gene very tightly linked to Huntington disease. Science, 238, 950-952.

Goate, A., Chartier-Harlin, M.S., Mullan, M., Brown, J., Crawford, F., Fidani, L., Giuffra, L., Haynes, A., Irving, N., James, L., Mant, R., Newton, P., Rooke, K., Roques, P., Talbot, C., Pericak-Vance, M. A., Roses, A., Williamson, R., Rossor, m., Owen, M. & Hardy, J. (1991). Segregation of missense mutation in the amyloid precursor protein gene with familial alzheimer's disease. Nature, 349, 704-706.

Gusella, J., Wexler, N., Conneally, P., Naylor, S., Anderson, M., Tanzi, R., Watkins, P., Ottina, K., Wallace, M., Sakaguchi, A., Young, A., Shoulson, I., Bonilla, E. and Martin, J. A polymorphic dna marker genetically linked to Huntington disease. Nature, 306, 234-239.

Hall, J. M., Lee, M., K., Newman, B., Morrow, J. E., Anderson, L. A., Huey, B. And King, M. C. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. Science, 250, 1684-1689.

Holden, C. (1998). Sequence, sequence, sequence. Science, 280, 1515.

Hubband, R. C., McElvaney, N. G., Birrer, P., Shak, S., Robinson, W. R., J. C., Wu,

M., Chernick, M. S., Crystal, R. G. (1992) A preliminary study of aerosolized recombinant human deoxyribonuclease I in the treatment of cystic fibrosis. The New England Journal of Medicine, 326, 812-815.

Hudson, K. L., Rothenberg, K. H., Andrews, L. B., Kahn, M. E. and Collins, F. S. (1995). Genetic discrimination and health insurance: An urgent need for reform. Science, 270, 391-393.

Kerem, B. S., Rommens, J. M., Buchanan, J. A., Markiewicz, D., Cox, T. K., Chakravarti, A., Buchwald, M., & Tsui, L.C. (1989). Identification of the cystic fibrosis gene: Genetic analysis. Science, 245, 1073-1079.

Kinzler, K. W. and Vogelstein, B. (1997). Gatekeepers and caretakers. Nature, 386, 761-763.

Koonin, S. (1998). An independent perspective on the human genome project. Science, 279, 36-37.

Korenberg, J. L. and Rimoln, D. L. (1995). Medical genetics. Journal of the American Medical Association, 273, 1692-1693.

Langston, A. A., Stanford, J. L., Wicklund, K. G., Thompson, J. D., Blazej, R. G. and Ostrander, E. A. (1996). Germ-line brca1 mutations in selected men with prostate cancer. American Journal of Human Genetics, 58, 881-885.

Levy-Lahad E., Wijsman, E., Nemens, E., Anderson, L., Goddard, A.B., Weber, J., Bird, T. Schellenberg, G. D. (1995). A familial Alzheimer's disease locus on chromosome 1. Science, 269, 970-972.

Levy-Lahad, E., Wasco, W., Prokaj, P., Romano, D., Oshima, J., Pettingell, W. H., Yu, C., Jondro, P., Schmidt, D., Wang, K., Crowley, A. C., Fu, Y. H., Guenette, S. Y., Galas, D., Nemens, E., Wijsman, E., Bird, T. D., Schellenberg, G. D. & Tanzi, R. E. (1995). Candidate gene for the chromosome 1 Familial Alzheimer's disease locus. Science, 269, 973-976.

Lyon, J., & Gorner, P. (1995). Altered fates: Gene therapy and the retooling of human life. New York, New York: W.W. Norton & Company, Inc.

Marshall, E. (1996). The genome program's conscience. Science, 274, 488-490.

Marwick, C. (1989). Technology, cost, cooperation, ethics challenges face genome mapping plan. Journal of the American Medical Association, 2652, 3247.

Maghaddam, B. And Adams, B. (1998). Reversal of phencyclidine effects by a group II metabotropic glutamate receptor agonist in rats. Science, 281, 1349-1351.

Marx, J. (1997). Possible function found for breast cancer genes. Science, 276, 531-532.

Miki, Y., Swensen, J., Eldens, D. S., Futreal, A., Harshman, K., Tavtigan, S., Lie, Q., Cochran, C., Bennett, L. M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Strano, A. H., Katcher, H., Yakamo, K., Ghalami, S., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayananth, P., Ward, J., Tonon, P., Narod, S., Bristow, P. K., Noris, F. H., Helvering, L., Neuhausen, S., Albright, L. C., Goldgar, D., Wiseman, R, Kamb, A. and Skolnick, M. H. (1994). A strong candidate for breast and ovarian cancer susceptibility gene brca1. Science, 266, 66-71.

Myers, R., Cupples, L., Schoenfeld, M., D'Agostino, R., Terrin, N., Goldmaker, N. and Wolf, P. (1985). Maternal factors in onset of huntington disease. American Journal of Human Genetics, 37, 511-523.

Pennisi, E. (1998). DNA sequencers' trial by fire. Science, 280, 814-817.

Pulver, A., Karayiorgou, M., Wolyniec, P., Lasseter, V., Kasch, L., Nestadt, G., Antonarakis, S., Housman, D., Kazazian, H., Meyers, D., Ott, J., Lamacz, M., Liang, K., Hanfelt, J., Ullrich, G., DeMarchi, N., Ramu, E., McHugh, P., Alder, L., Thomas, M., Carpenter, W., Manschreck, T., Gordon, C., Kimberland, M., Babb, R., Puck, J. and Childs, B. (1994). Sequential strategy to identify a susceptibility gene for schizophrenia: Report of potential linkage on chromosome 22q12-q13.1: Part 1. American Journal of Medical Genetics, 54, 36-43.

Pulver, A., Karayiorgou, M., Lasseter, V., Wolyniec, P., Kasch, L., Antonarakis, S., Housman, D., Kazazian, H., Meyers, D., Nestadt, G., Ott, J., Liang, K., Lamacz, M., Thomas, M. and Childs, B. (1994). Follow-up of a report of a potential linkage for schizophrenia on chromosome 22q12-q13.1: Part 2. (1994). American Journal of Medical Genetics, 54, 44-50.

Riordan, J. R., Rommens, J. M., Merem, B. S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J. L., Drumm, M. L., Iannuzzi, M. C., Collins, F. S., & Tsui, L. C. (1989). Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. Science, 245, 1066-1072.

Robbins, C., Theilmann, J., Youngman, S., Haines, J., Altherr, M., Harper, P., Payne,

C., Junker, A., Wasmuth, J. and Hayden, M. (1989). Evidence from family studies that the gene causing huntington disease is telomeric to d4s95 and d4s90. American Journal of Human Genetics, 44, 422-425.

Rogaev, E. I., Sherrington, R., Rogaeva, E. A., Ikeda, M., Liang, Y., Chi, H., Lin, C., Holman, K., Mar, L., Sorbl, S., Nacmlas, B., Placentini, S., Amaducci, L., Chumakov, I., Cohen, D., Lannfelt, L., Fraser, P. E., Rommens, J. M. & St. George-Hyslop, P. H. (1995). Familial Alzheimer's disease kindreds with missense mutations in a gene on chromosome 1 related to alzheimer's disease type 3 gene. Nature, 376, 775-778.

Rommens, J. M., Iannuzzi, M. C., Kerem, B. S., Drumm, M. L., Melmer, G., Dean, M., Rozmahel, R., Cole, J. L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J. R., Tsui, L. C., & Collins, F. S. (1989). Identification of cystic fibrosis gene: Chromosome walking and jumping. Science, 245, 1059-1065.

Rowen. L. (1997). Sequencing the human genome. Science, 278, 605-607.

Russell, P. J. Genetics (3rd Ed.). (1992). New York, New York: HarperCollins Publishing Inc.

Schellenberg, G., Pericak-Vance, M. A., Wijsam, E. M., Moore, D. K., Gaskell, P. C., Yamaoka, L. A., Bebout, J. L., Anderson, L., Welsh, K. A., Clark, C. M., Matin, G. M., Roses, A. D. & Bird, T. D. (1991). Linkage analysis of familial Alzheimer's disease, using chromosome 21 markers. American Journal of Human Genetics, 48, 563-583.

Schellenberg, G. D., Bird, T. D., Wijsman, E. M., Orr, H. T., Anderson, L., Nemens, E., White, J. A., Bonnycastle, L., Weber, J., Alonso, E., Potter, H., Heston, L. L. & Martin, G. M. (1992). Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. Science, 258, 668-671.

Schubert, E. L., Lee, M. K., Mefford, H. C., Argonza, R. H., Morrow, J. E., Hull, J., Dann, J. L. And King, M. C. (1997). BRCA2 in american families with four or more cases of breast or ovarian cancer: Recurrent and novel mutations, variable expression, penetrance, and the possibility of families whose cancer is not attributable to brca1 or brca2. American Journal of Human Genetics, 60, 1031-1040.

Sharan, S. K., Morimatsu, M., Albrecht, U., Lim, D. S., Regel, E., Dinh, C., Sands, A., Eichele, G., Hasty, P. and Bradley, A. 91997). Embryonic lethality and radiation hypersensitivity mediated by rad51 in mice lacking brca2. Nature, 386, 804-810.

Sherrington, R., Rogaev, E. I., Liang, Y., Rogaeva, E. A., Levesque, G., Ikeda, M.,

Chi, H., Lin, C., Holman, K., Tsuda, T., mar, L., Foncin, J., Bruni, A. C., Montesi, M. P., Sorbi, S., Rainero, I., Pinessi, L., Nee, L., Chumakov, I., Pollen, D., Brookes, A., Sanseau, P., Polinsky, R. J., Wasco, W., Silva, H., Haines, J. L. Pericak-Vance, M. A., Tanzi, R. E., Roses, A. D., Fraser, P. E., Rommens, J. M. & St. George-Hyslop, P. H. (1995). Cloning of a gene bearing missense mutations in early-onset Familial Alzheimer's disease. Nature, 375, 754-760.

St. George-Hyslop, P. H., Tanzi, R. E., Polinsky, R. J., Haines, J. L., Nee, L., Watkins, P. C., Myers, R. H., Feldman, R. G., Pollen, D., Drachman, D., Growdon, J., Bruni, A., Foncin, J., Salmon, D., Frommelt, P., Amaducci, L., Sorbi, A., Piacentini, S., Stewart, G., Hobbs, W. J., Conneally, W. & Gusella, J. F. (1987). The genetic defect causing familial alzheimer's disease maps on chromosome 21. Science, 235, 885-889.

Strooper, B. D., Saftig, P., Craessaerts, K., Vanderstichele, H., Guhde, G., Annaert, W., VonFigura, K. & VanLeuven, F. (1998). Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. Nature, 391, 387-390.

Tanzi, R. E., Vaula, G., Romano, D. M., Mortilla, M., Huang, T. L., Tupler, R. G., Wasco, W., Hyman, B. T., Haines, J. L., Jenkins, B. J., Kalaitsidaki, M., Warren, A. C., McInnis, M. C., Antonarkis, S. E., Crapper-Mclachlan, D. R., Gusella, J. F. & St. George-Hyslop, P. H. (1992). Assessment of amyloid *B*-protein gene mutations in a large set of familial and sporadic alzheimer's disease cases. American Journal of Human Genetics, 51, 273-282.

Tavtigian, S. V., Simard, L., Rommens, J., Couch, F., Eidens, S., Neuhausen, S., Merajver, S., Thorlacius, S., Offit, K., Lyonnet, S., Belanger, C. and Bell, R. (1996). Nature Genetics, 12, 333-337.

Thoralacius, S., Olafsdottir, G., Tryggvadottir, L., Neuhausen, S., Jonasson, J. G., Tavtigian, S. V., Tulinius, H., Ogmundsdottir, H. M. and Eyfjord, J. E. (1996) A single brca2 mutation in male and female breast cancer families from iceland with varied cancer phenotypes. Nature Genetics, 13, 117-122.

Travis, J. (1998). Another human genome project. Science News, 153, 334-335.

Travis, J. The human genome race heats up. Science News, 154, 239.

Understanding Our Genetic Inheritance [Internet Information]. (1998). The National Human Genome Research Institute.

Venter, J.G., Adams, M. D., Sutton, G. G., Kerlavage, A. R., Smith, H. O., &

Hunkapiller, M. (1998). Science, 280, 1540-1542.

Wang, H., Greenberg, C., Hewitt, J., Kalousek. and Hayden, M. (1986). Subregional assignment of the linked maker g8 (d4s10) for huntington disease to chromosome 4p16.1-16.3. American Journal of Human Genetics, 39, 392-396.

Weber, B. H., Brohm, M., Stec, I., Backe, J. and Caffier, H. (1996). A somatic truncating mutation in brca2 in a sporadic breast tumor. American Journal of Human Genetics, 59, 962-964.

Wickelgren, I. (1998). A new route to treating schizophrenia? Science, 281, 1264-1265.

Williams, M. (1996). An intriguing new lead on Huntington disease. Science, 271, 1233- 1234.

Wooster, R., Neuhausen, S. L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., Fields, P., Marshall, G., Narod, S., lenoir, G. M., Lynch, H., Feunteun, J., Devilee, P., Cirnelisse, C. J., Menko, F. H., Daly, P. A., Ormiston, W., McManus, R., Pye, C., Lewis, C. M., Albright, L. A., Peto, J., Ponder, B. A., Skolnick, M. H., Easton, D. F., Goldgar, D. E. And Stratton, M. R. (1994). Localization of a breast cancer susceptibility gene, brca2, to chromosome 13q12-13. Science, 265, 2088-2090.

Wooster, R. Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., Barfoot, R., Hamoudi, R., Patel, S. Rice, C., Biggs, P., Hashim, Y., Smith, A., Connor, F., Arason, A., Gudmundsson, J., Ficenec, D., Keisell, D., Ford, D., Tonin, P., Biship, D. T., Spurr, N. K., Ponder, B. A., Eeles, R., Peto, J., Devilee, P., Cornelisse, C., Lynch, H., Narod, S., Lenoir, G., Egilsson. V., Barkadottir, R. B., Easton, D. F., Bentley, D. R., Futreal, P.A., Ashworth, A. and Stratton, M. R. (1995). Identification of the breast cancer susceptibility gene brca2. Nature, 378, 789-792.

Zylke, J. W. (1992). Examining life's (genomic) code means reexamining society's long-held code. Journal of the American Medical Association, 267, 1715-1716.

BIOGRAPHICAL DATA

Jennifer Lynne Watson
Born - November 29, 1971 - Hunterdon Medical Center - Raritan Township, New Jersey
Elementary School - Lester D. Wilson Elementary School - Frenchtown, New Jersey
High School - Clearview Regional High School - Mullica Hill, New Jersey
College - Lebanon Valley College - Annville, Pennsylvania