Introduction

When we think of a disease, most of us imagine a nasty virus or bacterium, something that is an outside force. But in fact, one of the most frightening diseases is caused within our own cells. Cancer is a terrifying disease because, essentially, it is a disease of the body attacking itself. It is mysterious, insidious, and can strike seemingly randomly at any time. The human genome controls the cellular make-up of our bodies, and the genome has the answer to what causes cancer.

According to The National Cancer Institute, “all cancer arises from the accumulation of genetic changes within a cell.” In the fight against cancer, the ability to study the genetic profile of a cell during cancer development is a key step to advances in cancer prevention, early detection, diagnosis, and finding new drugs to cure cancer. The goal is to identify precisely what is different between a normal cell and a cancer cell and to follow the genetic changes that make a normal cell turn into a cancerous one. Eventually, by knowing the molecular differences among tumors, we can determine which tumors will respond to therapy, and if so, which therapies they respond to, and whether a tumor will metastasize or not. The National Cancer Institute (NCI) has established the Cancer Genome Anatomy Project (CGAP), a program with a goal to “achieve the comprehensive molecular characterization of normal, precancerous, and malignant cells.” In order to accomplish this, CGAP is establishing an index of all genes that are expressed in tumors (In Silico Analysis of cancer through the Cancer Genome Anatomy Project, by Strausberg).

CGAP has made sure that all project resources, data, informatics tools, and technology would be immediately accessible to the entire research community. This has helped me greatly in finding resources and information for my project. My research project studies the source of cancer: the genes that are involved with the formation of cancer cells. Using CGAP, I am researching genes that play an important role in breast cancer. The question of my project is: Do chromosome aberrations give rise to differentially expressed genes in cancer?

For such a difficult question to be answered, a basic understanding of cancer is needed. Certain genes regulate cell growth and division, and if one of these genes is mutated, the cell will not be able to regulate its growth and division, leading to cancer. A gene mutation may be spontaneous, or caused by environmental influences such as, X-rays, viruses or chemical carcinogens.

Here is an example of how cancer may start from a carcinogen. Carcinogenic substances are breathed into lungs usually. They also can enter in food and by absorption through the skin. Carcinogens, however, do not in their original form cause cancer. They must undergo a molecular modification inside a human cell before acquiring cancer causing ability. This process is called activation. A set of enzymes that use cells to detoxify alien substances converts them to a form that can be excreted in urine. Harry Gelboin of the National Cancer Institute has found, however, that some people have enzymes that change carcinogenic molecules the wrong way. The carcinogens are altered so that they can easily enter the cell’s nucleus and bind to DNA. This is the first step to a cancer causing mutation. The cells still have a natural line of defense – a DNA repair mechanism. Special molecules in the nucleus are able to detect abnormalities such as alien molecules attached to the DNA. The enzymes cut out the damaged DNA and allow the DNA to be repaired with new nucleotides. The double helix of DNA allows one strand to serve as a template for the damaged strand. If DNA repair occurs before the cell undergoes its next division into daughter cells, there will be no cancer. If the cell does divides before it is repaired, the portion of the genetic message that is bound to by a carcinogen may be copied abnormally. The daughter cells inherit a gene with a mutation, which cannot be repaired because now the genetic message is hidden within the molecularly normal DNA. There are 30,000 genes in a cell, so it is very likely that a bound carcinogen will bind to a gene that has nothing to do with cancer. However, if the carcinogen has bonded to a gene that controls cell division, cell growth, or any important characteristic that protects tissue from tumors, then cancer is likely to occur. These genes are called proto-oncogenes, which mean that if they are mutated they have the potential to cause cancer. (Cancer: The New Synthesis, by Boyce Resberger)

A proto-oncogene is normal cellular gene that will become an oncogene when it gains a dominant function mutation. There are several ways in which a proto-oncogene can be mutated to become an oncogene: either by rearrangement of DNA within the genome, by amplification of a proto-oncogene, or by point mutation. Rearrangement of DNA is what occurs when malignant cells are frequently found to contain chromosomes that have broken and rejoined incorrectly, translocating fragments from one chromosome to another. This may cause important genes to be placed in the wrong location, so the cell has no regulation and can become cancerous. The second way a proto-oncogene can be mutated into an oncogene is amplification. Increases in the number of copies of an oncogene will cause too many signals to be made, too many proteins to be built, and so the cell may become cancerous. The third way is a point mutation. A point mutation changes the gene’s protein product to one that is more active or more resistant to degradation than the normal protein.

The tumor-suppressor gene is a normal cellular gene that will become an oncogene when there is a recessive loss-of-function mutation. Some tumor suppressor genes make protein products that inhibit cell division, and its protein products normally help prevent uncontrolled cell growth. If these genes are not expressed, then excessive growth will lead to cancer. Some tumor suppressor proteins normally repair damaged DNA, which prevents cancer causing mutations. To have a tumor suppressor gene that repairs damaged DNA become mutated is traumatic for the cell because now it has no protection against common mutations. Another tumor suppressor gene controls adhesion of cells to each other or to the extra cellular matrix. If one of these genes is mutated, the cancer cells will not have proper cell anchorage, and are likely to pile loosely on top of each other.

Cancer cells do not respond normally to the body’s control mechanisms. By dividing excessively they can kill the organism. Normal cells exhibit density dependent inhibition, which means they will stop dividing and the tissue will stop growing after a certain point. But cancer cells do not stop dividing. Culture cancer cells will divide indefinitely if they are given proper nutrients. The breast cancer cells from Henrietta Lacks are still dividing today in culture since 1951 (Biology, 221). The immortality of many cancer cells is caused by a gene that regulates telomerase on the ends of chromosomes. Normally, a telomere functions as a tandem array of a short DNA sequence, TTAGGG. Telomeres provide the solution of the inability of DNA polymerases to completely replicate the end of a double stranded DNA molecule. In cancer, the enzyme telomerase prevents erosion of the ends of the chromosomes, thus removing a natural limit on the number of times the cells can divide. Specific inhibitors of telomerase have been suggested as cancer therapeutic agents.

Rapidly dividing cells form a tumor, a mass of abnormal cells within normal tissue. The problem begins when a single cell undergoes transformation, and if it escapes body’s immune system, it will continue to divide rapidly to form a tumor. The tumor is benign if it remains in one site, and the lump can be removed by surgery. The tumor is malignant if it can impair the functions of organs. This is defined as cancer. Cancer cells may separate from the original tumor, enter blood and lymph vessels, and proliferate to form more tumors, which is a process called metastasis.

Right now there are limited treatments for cancer patients. High energy radiation targets the tumor with a laser, and chemotherapy uses drugs that kill rapidly dividing cells, including the cancer, but also stomach lining cells and hair follicle cells, leaving patients with stomach pains and loss of hair.

The need for new cancer treatments is being concentrated on the study of the genome. Cancer runs in some families because their genomes are similar to each other’s and may be more likely to develop certain cancer causing genes.

Breast cancer is second most common type of cancer in U.S. and studies show that breast cancer is more likely to affect women who have a family history of breast cancer. BRCA 1 and 2 are important genes involved in breast cancer. Mutations in either gene increase the risk for developing breast cancer, because both are tumor suppressor genes, and their wild-type alleles suppress breast cancer (Biology, Mitchell, Reece, and Campbell).

The study of these and other genes associated with inherited cancer may lead to new methods for early diagnosis and treatment of all cancers. One example, the Ras oncogene, has a point mutation that leads to the hyperactive version of the Ras protein, which leads to excessive cell division. The normally functioning Ras protein relays a growth signal from a growth factor receptor on the plasma membrane of a cell, directing the synthesis of other proteins that stimulate the cell cycle. Ras proto-oncogene is critical to the regulation of cell division, so if it is amplified, the excessive cell division will lead to a tumor. Ras is found in 30% of human cancers, showing that this gene indeed plays an important role.

The p53 protein is a tumor suppressor gene found in 50% of human cancers.

p53 gene is normally expressed when there is damage to the cell’s DNA. p53 activates a signal to halt the cell cycle to allow time for the cell to repair the damaged DNA. If the DNA is irreparable, p53 activates suicide genes to cause cell death by apoptosis. Therefore, if p53 is missing, damaged DNA will go uncorrected, and cancer may ensue.

It takes, however, more than one mutation to cause cancer. In fact, multiple mutations underlie the development of cancer. Since cancer results from an accumulation of mutations, the older an organism is, the more likely it is to get cancer. For a cell to be cancerous, there must be at least one oncogene, and the mutation or loss of several tumor-suppressor genes. Loss of tumor suppressor genes is recessive, so both alleles must block tumor suppression, but oncogenes behave as dominant alleles.

Cancers due mainly to dominant oncogenes are the most likely targets for drug therapy. Overexpression can be attacked by antibodies. Tumors caused by mutations in tumor-suppressor genes are harder to treat because these result from the loss of a normal protein. Among the approaches now being attempted is a reintroduction of the p53 gene into a tumor. With p53, the damaged DNA will be repaired and abnormal cells will die by apoptosis. (Molecular Cell Biology Ch.24)

A substantial number of oncogenes and tumor-suppressor genes have already been discovered. However, genome analysis techniques suggest that the number of such genes may be strikingly large, and there is much to be discovered. It is important to find regions in the genome of recurrent aberrations in chromosomes. An aberration is a location on a chromosome that has deviated from the normal. A number of such regions have been identified in human cancers but the functional consequences of most of these abnormalities are not yet known. Identification of the affected genes in these regions, knowledge of their functions, and association of these genes with tumor progression is essential to fully understand the growth and progression of tumors. (Genome Changes and Gene Expression in Human Solid Tumors, by Joe W. Gray and Colin Collins)

My research project is to identify chromosome aberration patterns and identify genes that occur significantly in breast cancer tissue. I am concentrating on specific locations on chromosomes that have already been noted by CGAP to be potentially important locations for breast cancer causing genes.

There are twenty-two different chromosomes, and each have a double copy. There are also two sex chromosomes for a total of forty-six chromosomes in every cell. To specify where a chromosomal aberration is, the chromosome’s number is stated. The chromosome can further be divided because it has two arms, the longer arm is called q, and the shorter arm is called p. Each cytogenetic location is given a number to specify the exact location the aberration has occurred on the chromosome. When dyeing a chromosome, there will be many stripes, altering dark and light colors. These divisions on the chromosomes are divided into bands. For example, 11p15 means that the cytogenetic location is on chromosome number 11, it is on the shorter arm of the chromosome, and is located at band number 1, and sub-band number 5. There are many cytogenetic locations on chromosomes that are thought to be involved in the development of cancer. There may be genes in these regions that are differentially expressed through mutations.

The first step of my project is to choose 3 to 4 important recurrent chromosome aberrations in breast cancer from the Mitelman Database of Chromosome Aberrations. This is done by selecting aberrations by the criteria of how frequently they occur and by being in a region that contains many genes. Second, I will find the list of genes that are mapped to the cytogenetic band of each chromosome aberration, by using the Gene Finder on CGAP. A second program called Virtual Northern has data on the relative expression levels of each gene in normal versus cancer tissue. The expression levels are given in terms of EST’s. An EST, or expressed sequenced tag, is a tag of the messenger RNA, made by the gene. By counting the number of EST’s in a tissue, this is equivalent to counting the number of messages transcribed from the gene, and this is what we mean by expression level. The EST’s are counted by sequencing a cDNA library made from a tissue. cDNA stands for copy DNA and is made by mRNA from reverse transcription. For each gene, four numbers are given by the Virtual Northern program. (Note: Northern refers to a gel analysis of mRNA and the name is a play on the term Southern, which refers to the last name of the man who first analyzed DNA by gels.) The numbers are the counts for the EST’s specific to the gene in normal and cancer tissue, and the total number of ESTs for all genes expressed in normal and cancer tissue.

Third, I downloaded the expression level data (EST counts) for each gene into Excel to analyze the data. Differential expression is when the number of ESTs per gene is statistically different between normal tissue and cancerous tissue. A single gene might express just one message or a few hundred. A gene in normal tissue that has a significantly different level of expression than the same gene in cancer tissue is a gene that I am going to analyze further for a functional role in cancer. Because this gene is different in cancer than normal tissue, the gene may be a factor in the development of the cancer. After I have identified the statistically significant differentially expressed genes in breast tissue, I study each gene from a website called Locus Link, which provides extensive information about individual genes, such as their functions and whether they can be targeted by drugs.

I have selected this topic because I have an interest in the study of pharmacology and cancer, so this project exposes me to the frontier of drug discovery and I am able to use the same resources that professional cancer researchers use. My project is only examining a small part of this new and exciting part of cancer research. There is still an overwhelming amount of research that needs to be done. Cancer researchers know that cancer is still the second most common cause of death for people of all ages. One in nine women will develop breast cancer at some point in her life. This is why cancer researchers have a tremendous drive to continue working hard to find answers.