Jackson Poulos

Gene Expression Profiling in Cancer Diagnosis

Cancer is a disease that manifests itself through uncontrolled growth of cells. These cells proliferate throughout the human body and oftentimes form tumors and invade parts of the body. Cancer has a tendency to spread to other parts of the body from where it begins in a process known as metastasis. It is presently unclear as to the exact cause(s) of cancer, although several risk factors have been discovered. The more common risk factors include tobacco use, alcohol use, high body mass index, low fruit and vegetable intake, and lack of physical activity (WHO). However, there have been suggestions that 5-10% of cancers may occur as a result of genetics ("Heredity…"). Deaths as a result of cancer have also been increasing. In 2008, cancer accounted for 7.6 million deaths, and is project to account for 13.1 million deaths in 2030.

Gene expression profiling is a technique that is used to monitor the expression of thousands of genes at the same time. Performing gene expression profiling on cancer cells has led into a more efficient ability to diagnose cancer. One of the more common forms of gene expression profiling being used in cancer research involves the use of a DNA microarray (Profiling). The DNA microarray allows the identification of genes that are "turned on," or genes that are producing messenger RNA (also known as mRNA). From a generic standpoint, the process works by first collecting samples of the desired cell comparisons, such as fat cells from an obese study participant compared to fat cells from a non-obese participant (Profiling). The mRNA is isolated from the cells and then made into DNA copies, each usually identified with different colors. This is done through a process known as PCR, or polymerase chain reaction. The labelled DNA samples are put onto the microarray, where the colors are then analyzed to

determine which genes are expressed in different cases. For example, in comparing obesity among patients, genes of note might only be expressed in the obese patient's cells and not in the non-obese patient, or conversely, expressed in the non-obese patient yet not expressed for the obese patient. The goal of gene expression profiling is to establish trends for the relative expression of certain genes.

Gene expression profiling has potential to be used in cancer diagnosis for several reasons. A key concept that links gene expression profiling with cancer is oncogenesis, or the formation and creation of cancerous tumor cells (Delaval). At its most basic level, oncogenesis involved a cell's internal mechanisms becoming altered in order to allow for uninhibited cell division, which forms a malignant mass, known as a tumor. These cells continue to replicate, and as they are better at reproducing and dividing than normal cells, form a cancerous mass where many of those cancer cells share a common gene expression (Delaval). It is suggested that cancer may arise due to the presence of multiple cancer-causing genes known as oncogenes (Delaval). Oftentimes these oncogenes work in conjunction with each other or are expressed in pairs or groups to transform health human cells into cancer cells. This paper will examine several cases of how gene expression profiling can be used in order to improve cancer diagnosis.

Traditional classification of cancer is done by use of the WHO's classification system involving many signs, symptoms, and chemical staining of tissue. However, historically there have been issues with classical classification systems. For example, in large B-cell lymphoma, which is the most common subtype of non-Hodgkin's lymphoma, 40% of patients respond to

current therapy and enjoy surviving for an extended period of time past the disease. The remaining 60%, however, often die as a direct result of the disease (Alizadeh). It is hypothesized that there are differences in cancer cells indicating different types of this lymphoma, which, if it were identifiable information, could be communicated to doctors and treated differently. Gene expression provides a method of differentiating these types of tumor cells from each other and identifying them for different types of treatment. This acts as an example (which will be discussed in more detail later), of how gene expression can lead to enhanced treatment for cancer patients, as it provides an avenue by which differentiations of cancers may be distinguished where they would be indecipherable otherwise.

Another example that will be discussed is breast cancer. Breast cancer has been historically very difficult to classify and predict, with patients in the same stage of disease reacting drastically different to treatment options and having different outcomes to treatment (van't Veer). Additionally, the currently used predictors for breast cancer fail to accurate classify the cancer according to their clinical behavior (van't Veer). This leads to less desirable outcomes such as chemotherapy treatment for those who might survive without it. (van't Veer) One of the current goals for breast cancer treatment is to tailor treatment on a patient by patient basis, and gene expression profiling has been effective in working towards that goal.

In short, gene expression profiling has impacted cancer diagnosis include classifying tumors based on molecular structure and determining tumor sub classes. Due to the gene expression pattern for each cell or tissue type being unique, gene expressions of malignant tumor tissue can be compared to known tumors' gene expression "signatures" in order to

identify cancer or tumor types. After the type of tumor has been classified, additional tumor subclasses can be found through gene expression profiling, as the different types of tumor cells will also express different genes. This provides a new level of information for cancer diagnosis that would be unattainable without gene expression profiling.

Diffuse B-Cell Lymphoma

One example of tumor sub-classification through gene expression profiles is the identification of different types of diffuse B-cell lymphoma (DLBCL) identified by gene expression profiling. Classification of lymphomas began in 1832 when human lymphomas were first discovered by Thomas Hodgkin (Alizadeh). Since then morphologic (form and function) and molecular parameters have helped identify and differentiate different types of lymphoma (Alizadeh). There has been considerable trouble in defining subgroups of diffuse large B-cell lymphoma using those methods. This particular type of lymphoma accounts for 40% of cases of non-Hodgkin's Lymphoma, and currently, fewer than half of patients receive treatment that results in a sustainable remission (Alizadeh).

As mentioned before, oncogenesis results from an ordinary cell display mutations that alter or remove its ability to regulate cell division. As such, these malignant B-cells that result in DLBCL were traced to a certain stage of differentiation in the B-cells where the malignant cells acquired differences from the normal, healthy B-cells. A study conducted in 2000 attempted to determine whether gene expression profiling could identify distinct types of B-Cell malignancy, determine new classifications of tumors, and relate the cells back to the normal, health B-cells (Alizadeh). DLBCL was chosen specifically for the diverse amount of clinical behaviors present in

patients with the same diagnosis. Essentially, the researchers hoped to prove that gene expression profiling could be used to differentiate tumor types in a way that would potentially help in clinical treatment of the disease.

This particular DLBCL study was also noteworthy for its use of a specialized DNA microarray. It provides an example of how a DNA microarray may be used in a research setting for cancer diagnosis. The researchers for this study elected to use a microarray specialized for their disease, which they dubbed the 'Lymphochip' (Alizadeh). Due to the suspected involvement of mutated or altered B-cells with the oncogenesis of these non-Hodgkin's Lymphomas, the DNA placed on the microchip were from B-cell libraries including B-cells in those from multiple types of lymphomas, including DLBCL, follicular lymphoma (FL), and chronic lymphocytic leukemia (CLL) (Alizadeh). These microarrays were used to develop and identify gene expression patterns for DLBCL, FL and CLL. Additionally, gene expression profiles were created for normal lymphocytes in normal human tissue. These multiple microarrays and gene expression profiles allows for comparisons across the different cells in the hopes of finding defining characteristics for each disease.

A total of 1.8 million measurements of gene expression were made in 96 lymphocyte samples of both types (malignant and normal) for a total of 128 microarrays. The researchers also developed a classifying algorithm for analyzing the data and identifying certain gene expression proteins as certain diseases (Alizadeh). After collecting the data, the researchers needed to decide how to identify and cluster the groups of genes and their expressions in order to run the algorithm on and therefore analyze and make conclusions about the data. The

algorithm first resulted in identification of DLBCL tumors that would also have been identified through traditional methods. DLBCL tumors were able to be differentiated based on gene expression profiles; however there was no discernibility between the different types of DLBCL that the researchers hoped to find (Alizadeh). This was a result of hierarchical clustering, where genes were grouped based on similarities in their patterns. However, modification of the algorithms resulted in an altogether more promising verdict. By grouping the tumors by the expression exhibited by germinal center B cells (which was seen to be a set of cells differently expressed among DLBCL tumors), a different result emerged. The DLBCL cluster of tumors previously identified was now able to be split into two groups defined as germinal center B-like DLBCL and activated B-like DLBCL (Alizadeh). This was especially important for patient prognosis, as the germinal center B-like DLBCLs have an 80% survival rate over 10 years while the activated B-like DLBCLs has a 40% survival rate over a 10-year period.

The results of this study had important ramifications for treatment of DLBCL and demonstrate some of the power and ability gene expression profiling has in cancer diagnosis. For DLBCL, it showed that there may be two different types that could not be previously differentiable using the traditional methods. It is important to note that the treatment being given at the time was the same for all DLBCL patients, and the survival rates were wildly different between the two variations (Alizadeh). This may lead to development of future treatments that can specifically target one of the two different versions of the disease, as it is now possible to identify which type of lymphoma a patient to a degree beyond the reach of previous diagnostic options. Additionally, patients can now receive treatment and priority based on the type of this disease they have.

Breast Cancer

Another example of gene expression profiling in cancer diagnosis is its use in breast cancer diagnosis. Breast cancer, similar to diffuse B-cell lymphoma is difficult to classify using traditional methods. Clinical behavior fluctuates wildly among patients who have been diagnosed with the same types of cancer, and while treatment can be effective (chemotherapy and hormonal therapy can reduce the risk of metastases), 70-80% of patients receiving this type of treatment would have survived without it (van't Veer). Furthermore, it has been impossible to tailor therapy to patients on an individual level with previous levels of classification (van't Veer). A study similar to the DLBCL study was conducted in 2001 that looked at the ability of gene expression to provide prognosis for breast cancer patients.

The researchers looked at several groups of breast cancer patients. Sample were collected from patients who developed distant metastases within 5 years, those who were disease free after 5 years, patients with BRCA1 germline mutations and patients who were BRCA2 carriers (van't Veer). Gene expression profiles were created using RNA from each of the patient's tumor material and microarrays were created. Immediately, using unsupervised testing through a hierarchical algorithm, the researchers were able to distinguish between 'good' and 'poor' prognosis patients (van't Veer). Further analysis revealed that breast cancers can differ in their ER (estrogen receptor) status and lymphocytic infiltration (van't Veer).

The results from this study agree with another study of gene expression profiling in breast cancer diagnosis. Stanford researchers conducted a similar study in 200 examining how breast cancers may be differentiated through the use of gene expression profiles. They

examined patterns across 8,102 genes and looked to identify patters that could indicate

diversity in breast cancers (Perou). Similar to the previous study, breast cancers were divided

into groups based on whether they were ER-positive or ER-negative (Perou). The ER-positive

cancer had high expression of genes generally found in breast luminal cells. The ER-negative

groups were further diversified through the identification of expression of Erb-B2 and keratin 5-

and 17- genes (Perou).

Conclusion and Discussion

Similar to the results of the DLBCL study, these different types of breast cancers have

effects on the prognosis and expected outcomes of treatment. Patients belonging to these

groups exhibited similar outcomes to other patients in their groups and differed from patients

in other groups (Perou). The results of both of these studies have impacts on how cancer

diagnosis may be treated. They provide examples of how gene expression profiling may be used

to change how we think of cancer diagnosis, and will allow medicinal professionals to apply

even more specialized treatment to those who need it. With the two cases examined above,

one can see how treatment was not being properly allocated. With no knowledge of the

differences between the disease in patients, doctors simply had to apply the same treatment

for everyone. However, with the advent of gene expression profiling, doctors will be able to

assign the treatment to those who need it. Patients who do not need chemotherapy will not

receive it, and those with a more negative prognosis will receive the necessary treatment now

that the information is available.

Gene expression profiling is not limited to just these two cancer types. More recently, there have been indications that it may be a diagnostic tool for other cancers such as thyroid cancer (Vriens), prostate cancer (Lapointe), and others. The goal in all cases is to identify cancers with a better degree of precision in order to create better treatments that are individually tailored to the patient's exact cancer types. The technology has the potential to ensure patients get the right amount and type of treatment that is correct for them, as well as perhaps lead to the development of more types of treatment that are specified to these newly discovered cancer types. If the trend continues and more and more cancers are revealed to have subtypes identifiable by gene expression profiling, then cancer treatment will continue to refine and become more and more individualized for the patient, which will more than likely lead to better and more efficient treatment for everyone.

Works Cited

Alizadeh, Ash A., et al. "Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling." *Nature* 403.6769 (2000): 503-511. http://www.nature.com/nature/journal/v403/n6769/pdf/403503a0.pdf

"Cancer" World Health Organization. Retrieved 2013. http://www.who.int/mediacentre/factsheets/fs297/en/

Delaval, Bénédicte, and Daniel Birnbaum. "A cell cycle hypothesis of cooperative oncogenesis (Review)." *International journal of oncology* 30.5 (2007): 1051. http://www.ncbi.nlm.nih.gov/pubmed/17390006

"Heredity and Cancer". American Cancer Society. Retrieved 2013. http://www.cancer.org/cancer/cancercauses/geneticsandcancer/heredity-and-cancer

Lapointe, Jacques, et al. "Gene expression profiling identifies clinically relevant subtypes of prostate cancer." *Proceedings of the National Academy of Sciences of the United States of America* 101.3 (2004): 811-816. http://www.pnas.org/content/101/3/811.short

Perou C.M. et al. (2000). "Molecular portraits of human breast tumors". Nature 406 (6797): 747–752. http://www.nature.com/nature/journal/v406/n6797/full/406747a0.html

"Profiling Technique: microarray analysis." The University of Utah. http://learn.genetics.utah.edu/content/labs/microarray/analysis/

van't Veer, Laura J., et al. "Gene expression profiling predicts clinical outcome of breast cancer." *nature* 415.6871 (2002): 530-536. http://www.nature.com/nature/journal/v415/n6871/pdf/415530a.pdf

Vriens, Menno R., et al. "MicroRNA expression profiling is a potential diagnostic tool for thyroid cancer." Cancer 118.13 (2012): 3426-3432. http://onlinelibrary.wiley.com/doi/10.1002/cncr.26587/full