

HUMAN TANDEM REPEATS IN BREAST CANCER
PROGRESSION

By
Krista Kinnard

A Thesis Submitted to the Honors College


In Partial Completion of the Bachelors Degree With
Honors in

Molecular and Cellular Biology

The University of Arizona

May 2010

Approved by:



Dr. Julie Lang
Arizona Cancer Center, Bio5 Institute

STATEMENT BY AUTHOR

I hereby grant to the University of Arizona Library the nonexclusive worldwide right to reproduce and distribute my thesis and abstract (herein, the "licensed materials"), in whole or in part, in any and all media of distribution and in any format in existence now or developed in the future. I represent and warrant to the University of Arizona that the licensed materials are my original work, that I am the sole owner of all rights in and to the licensed materials, and that none of the licensed materials infringe or violate the rights of others. I further represent that I have obtained all necessary rights to permit the University of Arizona Library to reproduce and distribute any nonpublic third party software necessary to access, display, run, or print my thesis. I acknowledge that University of Arizona Library may elect not to distribute my thesis in digital format if, in its reasonable judgment, it believes all such rights have not been secured.

SIGNED: 

Abstract

Introduction: A tandemly-repeated sequence of DNA located approximately 1kb upstream of HIC1 has been identified which appears to regulate the expression of this important tumor suppressor gene. Loss of HIC1 expression in tumors, either by deletion or hypermethylation, has previously been shown to correlate with a more severe prognosis in multiple cancers. Initial data show that larger alleles of this tandem repeat do not influence incidence of disease but do appear to correspond with a heritable predisposition to more aggressive cancers, represented by earlier onset and increased metastasis. This study hypothesizes that there may be a connection between these more aggressive types of cancer but also with the deleterious BRCA mutation.

Methods: Using 39 breast cancer and DNA samples from the Tumor Bank, this study evaluated the effect of the tandem repeat upon HIC1 function and correlated the results with the disease stage and prognosis in those individuals. In this study, each sample was genotyped for the tandem repeat alleles by PCR and evaluated chromatin methylation status using Sequenom. An annotated database was created from the medical records of each of these cancer cases highlighting specific markers of aggressive cancers. These annotations include breast cancer incidence, BRCA status, nodal status, ER, PR and Her2 status and the genotype of the tandem repeat.

Results: Based on statistical analysis of these data, no significant correlations could be made between tandem repeat length and these specific cancer characteristics. Additionally, Sequenome testing revealed that no sample had more than 10% methylation of the HIC1 gene, thus indicating that there is no statistical correlation between

methylation of HIC1 and repeat length.

Conclusion: The results of the statistical analysis indicate that the sample size for all of the proposed correlations is too small. For several of these hypotheses, over one thousand more patient samples are necessary to produce statistically significant results. Therefore, no commentary can be made about the status of a large tandem repeat allele as a possible breast cancer marker.

Introduction

Over the past several decades, great strides have been made to better understand the causes of cancer and how it progresses on a molecular level. Cancer experts estimate that approximately 90% of cancers are sporadically occurring, meaning that they are not caused by a specific inherited mutation. The remaining 10% are thought to be caused by a heritable mutation in a given gene (4). Tests have been developed to test for these specific heritable markers, for example BRCA in breast cancer, to help physicians make better diagnoses and provide better treatment for these specific types of cancers. Recent studies have suggested that HIC1, a tumor suppressor gene, plays an important role in the development of many cancers (9, 18, 20). A tandem repeat has been identified in the promoter region, which is highly polymorphic in humans (17). Initial studies of a small number breast cancer samples showed a correlation with increased number of repeats (>5) and earlier onset of cancer. In addition, studies on 95 colon cancers showed a correlation with increased number of repeats (>5) and a poorer prognosis in patients (Tim Helentjaris, Nina Sun and Vicki Chandler, unpublished data). This project aims to further investigate the relationship between the numbers of tandem repeats at HIC1 and cancer prognosis. Further studies to explore whether there is a strong relationship between the numbers of tandem repeats and cancer prognosis, tests could be developed to screen patients early allowing for more specific treatment of disease.

HIC1 function

Cell cycle regulation is critical to the proper function of a cell. Thousands of proteins have been found to be involved in this regulatory process at several different key points in the cell cycle. The protein p53 has been shown to be essential in the regulation

of the cell cycle between G1 and S phase. This protein has several different responsibilities, the most important of which is suspending the cell cycle in the presence of DNA damage and determining whether the damage can be repaired or whether the damage is beyond repair and thus signals apoptosis. p53, is one of the final steps of several different signaling cascades (20). This study focuses on the cascade involving HIC1 and SIRT 1 (See Figure 1).

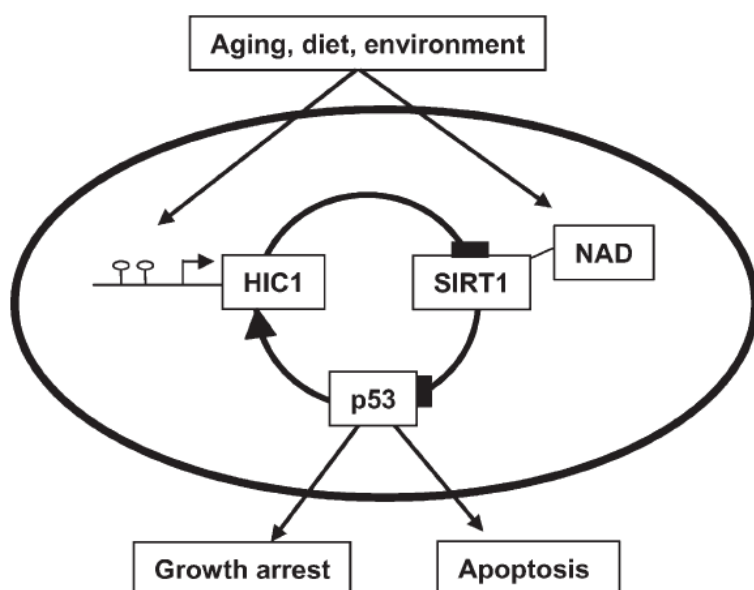


Figure 1. A model for the role of HIC1 in Tumor Suppression.

A circular regulation of HIC1, SIRT1, and p53 is proposed for modulation of cellular responses to DNA damage. HIC1 represses the transcription of SIRT1, SIRT1 deacetylates p53 posttranscriptionally, and p53 *trans*-activates HIC1. (2)

Previous studies have shown that HIC1 is a regulator protein of p53 through the action of another protein, SIRT1 (8). According to the Chen et al study, mutation of the HIC1 protein upregulates SIRT1, a negative regulator of p53, which causes p53 to be upregulated (2). Unregulated p53 does not properly function thus allowing damaged DNA to pass through the cell cycle and propagate in the next cell generation. The growth of cells with damaged DNA leads to cancer. Thus, HIC1 is an important player in the

regulation of cell cycle and cancer development (2).

HIC1 and Breast Cancer

In 1994, a Brazilian team of scientists noticed that an increased number of women with poor prognosis breast cancer had similar mutations of chromosome 17p (11). Later studies showed that the location of this mutation is in fact the tumor suppressor gene HIC1 and linked the gene to several different types of cancer (9, 16). In addition, a 1998 study and several subsequent studies have shown that methylation of HIC1, which silences the gene, is specifically linked to progression of human breast cancers (5, 6, 13).

Tandem Repeat

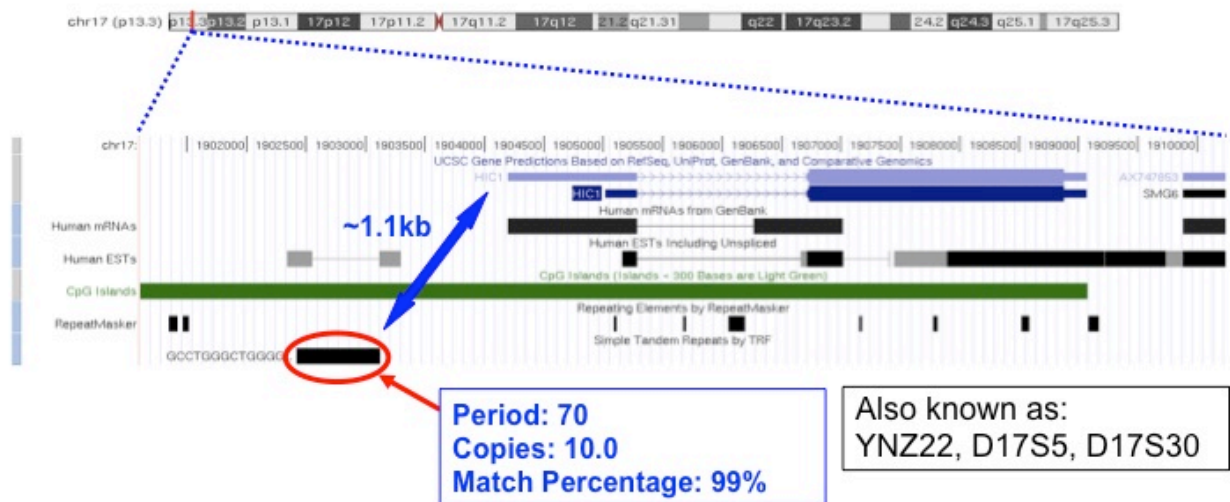


Figure 2. Map of Tandem Repeat.

The tandem repeat of period 70 bp is mapped in relation to HIC1 on chromosome 17p.

A tandem repeat upstream of the HIC1 gene in the promoter region has been investigated in Dr. Vicki Chandler's lab at The University of Arizona's Bio5 Institute (Figure 2). The number of repeats is highly polymorphic in different individuals with the numbers of the 70bp sequence ranging from 1 to 12 repeats in the most common alleles

(1, 17). Initial studies of breast and colon cancer suggested that there is a correlation between the length of the repeat and the age of onset or aggressiveness of cancer (Tim Helentjaris and Vicki Chandler, unpublished data). This study is a collaboration with Dr. Julie Lang, a breast surgical oncologist at the Arizona Cancer Center, to examine additional breast cancers to explore the interaction between this tandem repeat, HIC1 and the effects of the repeat size on breast cancer progression and prognosis. Understanding the effects of this tandem repeat allele on breast cancer could potentially help physicians in the diagnosis and treatment of certain breast cancers.

Research Hypotheses: Correlations between tandem repeat length and breast cancer characteristics

This study intends to explore potential correlations between patients with large alleles of the tandem repeat and particularly aggressive markers of breast cancer. The following hypotheses have been developed:

1. There is a correlation between BRCA positive status and patients containing at least one large allele.

BRCA is another tumor suppressor gene that has already been widely studied and identified to be linked to very specific types of cancer. Like HIC1, the BRCA protein is involved in cell cycle regulation, though by a much different mechanism that will not be discussed here. Though BRCA and HIC1 appear to function separately in the cell, they are both involved in regulation that could result in breast cancer improperly managed (i.e. by a mutation in the gene). Therefore, this study is exploring a potential connection between breast cancer patients that have a known BRCA mutation (known as BRCA positive status).

2. There is a correlation between node positive cancers and patients containing at least one large allele.

Any type of cancer is considered to be more aggressive if the cancer cells metastasize from the local tissue to the surrounding axillary and sentinal nodes. During surgery, a pathologist identifies this type of metastasis (12, 21). This study explores the relationship between cancers that include nodal metastasis (known as node positive cancers) and tandem repeat length.

3. There is a correlation between triple negative cancers and patients containing at least one large allele.

Pathological work up of cancer tissue identifies if a specific breast cancer is estrogen receptor (ER) positive, progesterone receptor (PR) positive and Her2 (an epidermal growth factor that is responsible for signal transduction) positive. Being positive for any or all of these three characteristics means that hormonal or other specific treatments may aid in the management and treatment of cancer. Patients who test negative for all three (known as triple negative cancers) are at a severe disadvantage as many of the current treatment options are not effective in managing the disease (7). Thus, triple negative status is no of the indicators of a more aggressive type of breast cancer. This study intends to identify a potential correlation between triple negative cancers and tandem repeat length.

4. There is a correlation between age of cancers onset and patients containing at least one large allele.

Cancers that occur at a younger age (under the age of 50) are generally considered to be linked to more aggressive types of cancer (7). Thus, this study explores the relationship

between age of onset of the cancer and tandem repeat length.

5. There is a correlation between patient ethnicity and patients containing at least one large allele.

Studies have shown that patients of certain ethnic backgrounds are at greater risk for specific types of cancer than others. Breast cancer incidence has been shown to be a particularly high risk for African-Americans (3). This study explores the potential relationship between different ethnic groups and tandem repeat length.

6. There is a correlation between methylation of the HIC1 tumor suppressor gene and patients containing at least one large allele.

As stated earlier, HIC1 becomes hypermethylated in breast cancer. This study explores the potential relationship between the extent of this hypermethylation and tandem repeat length.

Materials and Methods

Patient Sample Collection

All patients were undergoing treatment for a diagnosis or suspected diagnosis of breast cancer at the Arizona Cancer Center. The patient population was limited to adults over the age of 21 who have consented to having their tissue collected and banked. Before surgery, each patient signed an informed consent allowing her blood and tissue to be used for scientific study. Patients were allowed to ask questions to Dr. Lang about the process or the resulting scientific studies. Both Dr. Lang and the Arizona Cancer Center's Tumor Bank/TACMASS laboratory have copies of the consent. During surgery, the Tumor Bank collected whole blood and other tissues. All tissues collected were given a T-number to deidentify the patient from the sample. As this study used

whole blood from patients, collected blood samples were separated into 350µl aliquots then stored at -80°C. This study focused on 39 confirmed breast cancer cases with whole blood collected between 2006 and 2009. These were the only samples available for this molecular study.

Genomic DNA purification

To obtain genomic DNA, each of the 39 samples of whole blood was purified using BioRad BioRobot EZ1 worktable (Hercules, CA). This automated machine purified genomic DNA from 6 whole blood samples at a time. Using a premade kit, samples were placed in the EZ1 worktable producing pure genomic DNA. The resulting genomic DNA for each sample was stored at -80°C with minimal melting/re-freezing to preserve the integrity of the DNA.

DNA Quality

Before analysis could be done on the DNA, it first had to be examined to ensure that the DNA remained in large fragments and was not degraded. Agarose gels are designed to separate small fragments of DNA, thus running an entire genome will reveal only a thick band at the top of the gel if there has been no degradation. Each of the 39 samples was run on an agarose gel to ensure that only a thick band appeared near the top of the gel, any smaller bands that appeared would indicate that the DNA had been partially degraded.

PCR

Each of the 39 samples was subjected to PCR with specific primers designed to flank the tandem repeat region of the DNA. The PCR results produce a different sized fragment for each of the series of repeat numbers, thus this method was useful in

determining the number of tandem repeats in each sample. Fragments <5 repeats are indicated by the genotype S (small) and fragments with >5 repeats are given the genotype L (large). The results for each sample show either two fragments of different sizes indicating a heterozygote, or one strong fragment indicating a homozygote. As indicated in Figure 3, each sample was assigned a genotype for the tandem repeat of SS, SL or LL.

Methylation Study

Since HIC1 is known to become hypermethylated in many types of cancer, there could be a potential link between tandem repeat size and the extent of methylation of HIC1. This was done using Sequenome (San Diego, CA) methylation analysis. This process of identifying the percent methylation of a given DNA sequence (in this case the HIC1 gene) first by sequencing the section of the genome then identifying the methylated CpGs (methylation is the result of a methyl group being added to the Cytosine nucleotide of the CpG dinucleotide). A percentage of methylation is then automatically generated (22).

Annotated Database

Once the genotype for the tandem repeat of each breast cancer sample was determined, an annotated database detailing the unique type of cancer for each patient sample was created using the surgical notes and chart information on the University Medical Center's Sunrise Clinical Manager patient management system. Specific fields of interest included Age of onset, ethnicity, BRCA status, nodal status and ER, PR and Her 2 status. Age of onset indicates the age at which the patient was first diagnosed with cancer. Ethnicity identifies the cultural background of the patient. This database correlates all of these specific effects of breast cancer with the tandem repeat genotype.

Results

DNA Quality Results

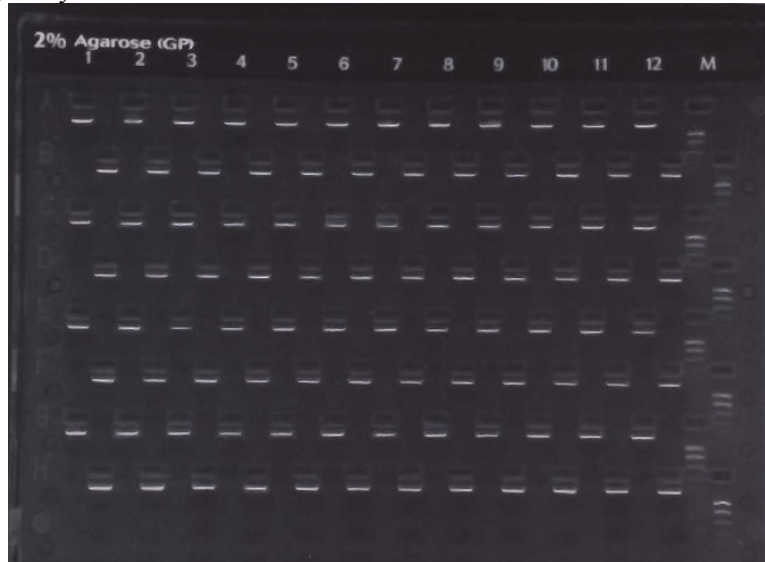


Figure 3. 2% Agarose gel indicating genomic integrity.

This is a representative gel illustrating that all samples are in fact genomic DNA.

Since this particular gel has multiple lines of wells, a large band near the well indicates quality genomic DNA. Each of the breast cancer samples illustrated in Figure 3 depicts a large band near the well. This indicates that the quality of the genomic DNA is high, and thus usable for this study, as there are no stray bands appearing in the gel.

PCR results

Phenotype	Repeat #	Fragment size (bp)
S	1	359
	2	429
	3	499
	4	569
	5	639
L	6	709
	7	779
	8	849
	9	919
	10	989
	11	1059
	12	1129

Figure 4. Matching fragment size to number of repeats and genotype

This chart illustrates the indicated genotype and repeat number for the fragment size of each sample as determined by PCR.

Figure 4 shows the number of repeats that correspond to the fragment size (which is inferred from Figure 5) which represent an allele of the tandem repeat.

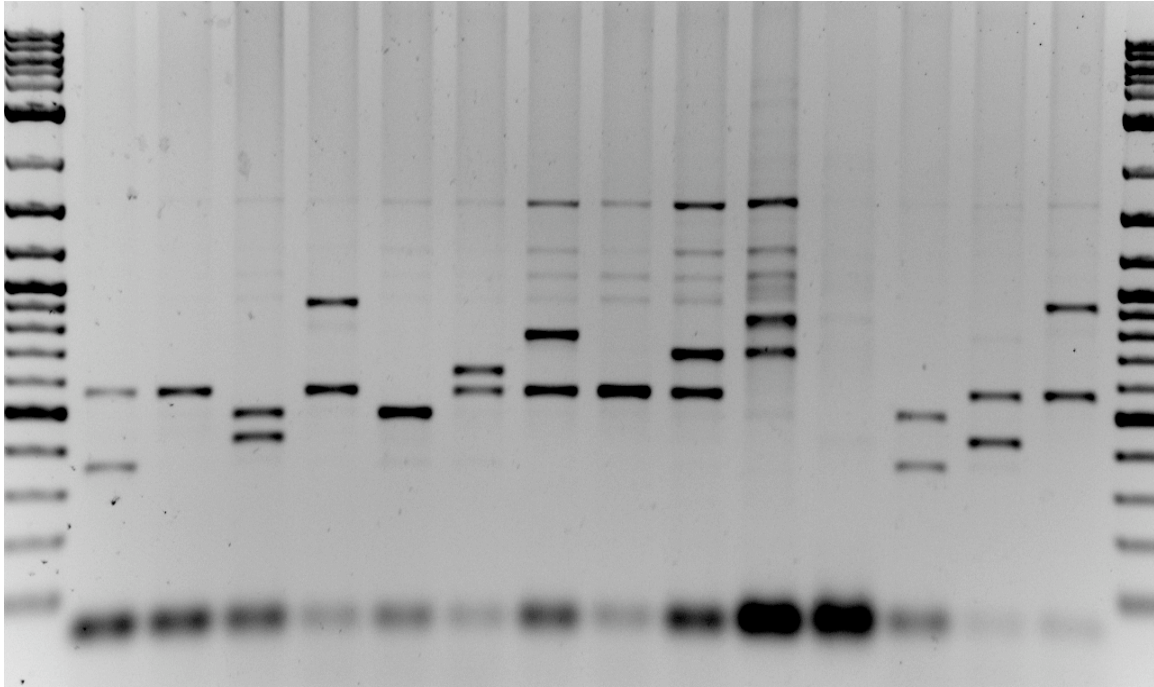


Figure 5. Representative PCR gel

A representative gel representing the PCR results.

The 39 breast cancer samples were subjected to PCR analyses as described in the materials and methods. Figure 5 is a representative gel illustrating what the results look like and confirming previous analyses that indicate this repeat is quite polymorphic in humans. The exact length of each band is identified based on the 1000 bp ladder at the far left and right of the image. The blurred bands at the bottom of the image represent primer dimerization.

Tandem Repeat Genotype

Table 1 summarizes the results. Of the 39 breast cancer cases, 21 were genotype SS, 17 were genotype SL and 1 was genotype LL.

T #	Fragment 1 Size (bp)	Repeat #	Fragment 2 size (bp)	Repeat #	Genotype
T07-897A	500	3	950	9	SL
T07-1033A	450	2	650	5	SS
T07-2164	850	8	1000	10	LL
T07-2826	350	1	600	4	SS
T07-3159	500	3	850	8	SL
T07-3365	350	1	500	3	SS
T07-3378	600	4	850	8	SL
T08-0122	500	3	775	6	SL
T08-0213	600	4	homo	4	SS
T08-0352	475	2	900	8	SL
T08-0408	575	4	725	6	SL
T08-0506	450	2	900	7	SL
T08-0511	450	2	500	3	SS
T08-0588	550	3	950	9	SL
T08-0726	500	3	950	9	SL
T08-0737	550	3	700	5	SS
T08-0777	350	1	500	3	SS
T08-0792	450	2	550	3	SS
T08-0996A	425	1	homo	1	SS
T08-1018A	450	2	homo	2	SS
T08-1033A	600	4	homo	4	SS
T08-1037B	450	2	800	7	SL
T08-1386	400	1	450	2	SS
T08-1414	500	3	homo	3	SS
T08-1423	450	2	550	3	SS
T08-1428	500	3	950	9	SL
T07-2130	600	5	850	8	SL
T07-2135	450	2	600	5	SS
T07-2709	550	3	homo	3	SS
T08-0122	500	3	775	6	SL
T08-0204	600	4	1000	10	SL
T08-0223	500	3	homo	3	SS
T08-0380	450	2	950	9	SL
T08-0602	500	3	500	3	SS
T08-0711	550	3	homo	3	SS
T08-0845	400	1	550	3	SS
T08-0873	550	3	850	8	SL
T08-0912A	550	3	850	8	SL
T07-191A	650	5	homo	5	SS

Table 1 Genotype of each sample

This table shows the size in bp of both alleles of the tandem repeat as well as the number of repeats and the genotype.

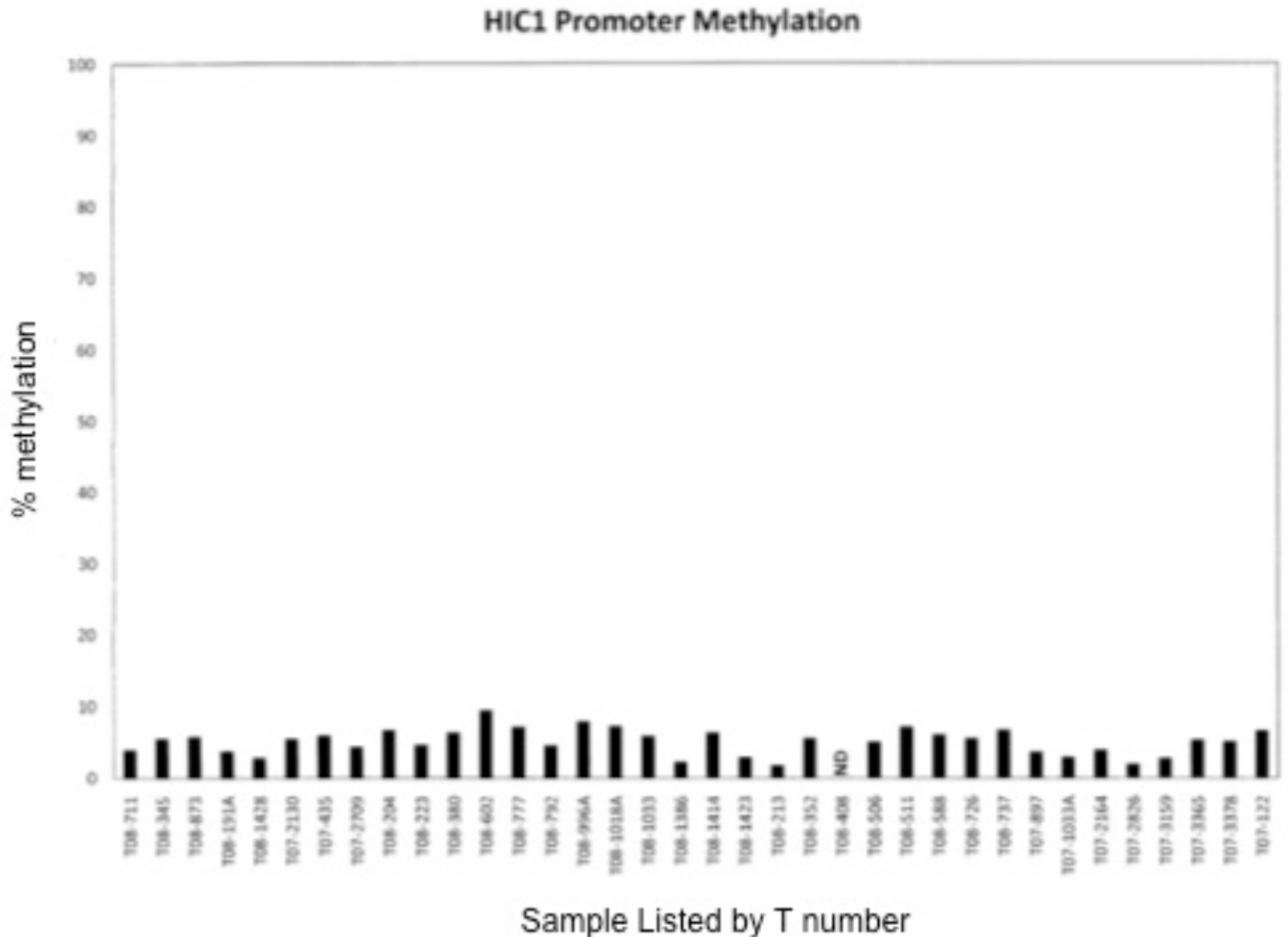
Methylation Study**Figure 6. Percent methylation**

Figure 6 shows the percent methylation for the HIC1 gene for each sample (listed by T number, which represents the individual, deidentified patient number). No sample has more than 10% methylation, thus there was no hypermethylation of HIC1 in any of the blood samples from the patients examined.

Database

Table 2, contains all of the annotation data on regarding the type of cancers, ethnicity and the tandem repeat genotype.

T number	BRCA	Node Positive	Triple Negative	Age of Onset	Ethnicity	Genotype
T07-897A	no	yes	yes	40	Hispanic	SL
T07-1033A	no	no	no	46	Caucasian	SS
T07-2164	no	no	no	58	Caucasian	LL
T07-2826	yes	no	yes	60	Caucasian	SS
T07-3159	no	no	no	52	Caucasian	SL
T07-3365	no	no	no	51	Caucasian	SS
T07-3378	no	no	yes	56	Caucasian	SL
T08-0122	no	no	no	66	Caucasian	SL
T08-0213	no	yes	no	35	Hispanic	SS
T08-0352	no	no	no	78	Caucasian	SL
T08-0408	no	no	no	54	Caucasian	SL
T08-0506	no	no	no	70	Caucasian	SL
T08-0511	no	yes	no	50	Caucasian	SS
T08-0588	no	yes	no	42	Hispanic	SL
T08-0726	no	no	no	58	Other	SL
T08-0737	no	no	no	44	Caucasian	SS
T08-0777	no	no	no	55	Caucasian	SS
T08-0792	no	no	no	63	Caucasian	SS
T08-0996A	no	yes	no	50	Caucasian	SS
T08-1018A	no	no	no	64	Caucasian	SS
T08-1033A	no	no	no	78	Caucasian	SS
T08-1037B	no	no	no	58	Caucasian	SL
T08-1386	no	yes	no	62	Caucasian	SS
T08-1414	no	no	no	77	Caucasian	SS
T08-1423	no	no	no	51	Caucasian	SS
T08-1428	no	yes	no	47	Caucasian	SL
T07-2130	no	no	no	50	Caucasian	SL
T07-2135	no	no	no	41	Hispanic	SS
T07-2709	no	no	no	46	Other	SS
T08-0122	no	no	no	70	Caucasian	SL
T08-0204	no	no	no	54	Caucasian	SL
T08-0223	no	no	no	69	Caucasian	SS
T08-0380	no	no	no	83	Caucasian	SL
T08-0602	no	no	no	44	Caucasian	SS
T08-0711	no	no	no	66	Caucasian	SS
T08-0845	yes	yes	no	58	Other	SS
T08-0873	no	yes	no	33	Hispanic	SL
T08-0912A	no	yes	no	55	Caucasian	SL
T07-191A	no	yes	no	59	Caucasian	SS

Table 2. Annotated database

From this database (Table 2), the nature of the cancer and correlations tandem repeat genotype can be examined. Of these samples, only two are BRCA positive and both are genotype SS. Additionally, eleven of the samples are node positive with six being of genotype SS, five with genotype SL and zero with genotype LL. Only three of these samples are triple negative, one is SS, two are SL and zero are LL. Cancers that occur before the age of 50 are considered more aggressive. Of these 39 breast cancer cases, 10 were under the age of 50. Six were genotype SS, four were SL and zero were LL. Looking at the ethnicity of the patient for each of these samples shows 31 are Caucasian with 17 as SS, 13 as SL and 1 as LL. A total of 5 of the samples are from Hispanic patients. Two of these samples were SS, three were SL and zero are LL. There are three samples from patients who listed their ethnicity as other. Two of these were SS, one was SL and zero were LL.

Discussion

Statistical Analysis

Based on the data presented in Table 2, potential correlations can be made between the three different genotypes of the tandem repeat (SS, SL and LL) and each of the five cancer characteristics. Statistical correlation is performed through the usage of two hypotheses: the research hypothesis, which states that there is a correlation and the null hypothesis, which states that there is no correlation. For this study there were several research hypotheses:

To support any of these hypotheses, there must be enough data to reject the null hypothesis (that there is zero correlation) 80% of the time. Simple calculations give the sample size necessary to achieve this power for each hypothesis and unfortunately this

study is yet too small to be able to test any of the above hypotheses.

Based on the number of samples that are either SL or LL (18 samples), a sample size has been calculated for each of these hypotheses. To show correlation with BRCA status, 152 samples of SL and LL plus an additional 152 SS samples are necessary. To show correlation with node status, over 1000 samples of both SL/LL samples and SS are required. To show correlation between large alleles and triple negative cancers, 240 each of SL/LL and SS are necessary. To show correlation with both age of onset and ethnicity, over 1000 samples are necessary for both SL/LL and SS genotypes.

Unfortunately, the sample size of 21 SS and 18 SL/LL samples was not nearly large enough to reveal any correlation for any of these hypotheses. However, the data collected thus far provides a start for continuation of this study.

Clinical Significance

Through these correlations between large alleles of the tandem repeat and particular cancer characteristics, this study aimed to identify cancer markers that could potentially act as an aid in cancer diagnostics. Since the location and excellent primers are known, simple DNA tests could be performed on blood cells to identify the genotype of this allele in a given breast cancer patient.

Cancer treatment is a continuously developing area of study and medical professionals already have specified treatments for certain cancer characteristics. For example, breast cancers that are triple negative are treated differently than cancers that are estrogen receptor and/or progesterone receptor positive. The same is true for BRCA positive cancers and node positive cancers. If this study later shows a correlation between one of these cancer characteristics and large genotypes of this tandem repeat,

this tandem repeat could serve as a marker for physicians, providing additional insight as to the best, most specific treatment for a patient's unique form of cancer.

REMARK criteria for identifying cancer markers

Historically, defining a cancer marker and establishing agreement among the scientific and medical communities has been controversial. Many studies have claimed to have found new cancer markers only to be disproven by others. This can be attributed to many different causes such as poor study design, lack of assay reproducibility or misleading statistics (10). In 2006, a group of scientists associated with the National Cancer Institute published guidelines for the reporting of cancer markers in an effort to standardize the process and eliminate the confusion and debate that has existed previously. In the publication *Reporting recommendations for tumor MARKer prognostic studies (REMARK)*, the committee gives specific recommendations in an 18 step outline, for description of the marker in question, study design, power studies for sample size calculation, statistical analysis, presentation of data, discussion of significance of results as well as the implications and clinical value of the results.

Since tandem repeat study aims to introduce a new cancer marker for potential clinical use, it is only appropriate that these guidelines be followed. In terms of outlining a clear direction for the study and clearness of the final goal, this study has followed the REMARK guidelines. However, many more samples need to be analyzed before we will know if there is any correlation. Originally, the Tumor Bank supplied over 250 whole blood samples for genotype analysis. The specific type of cancer for each of the samples, however, was unknown. Only later, after the genotyping was done, was it discovered that only 39 of these 250 samples were breast cancer. This is simply not a large enough

sample size to produce any significant data.

Because this study's design and organization has already been established in accordance with the REMARK criteria, it will be straightforward to collect additional samples to produce statistically significant results. Once the necessary sample sizes are attained, statistical analysis and data reporting will be presented according to the criteria.

Future Directions

The most immediate steps for continuing this study are to collect enough samples to obtain statistically significant results for large allele correlations with BRCA status, node status, triple negative cancers, age of onset and ethnicity. This tandem repeat and its interaction with HIC1 have been shown to be involved in several different types of cancers including ovarian and colorectal cancer. In the future, collecting whole blood samples from patients with these cancers and analysis between tandem repeat length and each of these cancers specific characteristics could provide even more insight to the function and implications of this tandem repeat. This could potentially show that the HIC1 tandem repeat is particularly important in one specific type of cancer with very specific implications (i.e. likelihood for nodal metastasis) or could prove to be part of a greater function in many different types of cancer progression. The current data reported at this time are hypothesis generating; additional studies are necessary to determine the significance of tandem repeats of HIC1 as a biomarker in breast cancer.

References

1. Breen G, Collier D, Craig I, Quinn J (2008) Variable number tandem repeats as agents of functional regulation in the genome. *IEEE Eng Med Biol Mag* 27:103-4, 108.
2. Chen WY, Wang DH, Yen RWC, et al. Tumor suppressor HIC1 directly regulates SIRT1 to modulate p53-dependent DNA-damage responses. *Cell* 2005; 123(3):437-448.
3. Demicheli R, Retsky MW, Hrushesky WJM, et al. Racial disparities in breast cancer outcome - Insights into host-tumor interactions. *Cancer* 2007; 110(9):1880-1888.
4. Deng CX. BRCA1: Celly Cycle Checkpoint, Genetic Instability, DNA Damage Response and Cancer Evolution. *Nucleic Acids Research* 2006; 34(5): 1416-1426
5. Esteller M. Epigenetic gene silencing in cancer: the DNA hypermethylome. *Human Molecular Genetics* 2007; 16:R50-R59.
6. Fujii H, Biel MA, Zhou WB, et al. Methylation of the HIC-1 candidate tumor suppressor gene in human breast cancer. *Oncogene* 1998; 16(16):2159-2164.
7. Haffy, BG, et al. Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *Journal of Clinical Oncology* 2006; 24(36):5652-5657.
8. Hoff C, Mollenhauer J, Waldau B, et al. Allelic imbalance and fine mapping of the 17p13.3 subregion in sporadic breast carcinomas. *Cancer Genetics and Cytogenetics* 2001; 129(2):145-149.
9. Konishi H, Sugiyama M, Mizuno K, et al. Detailed characterization of a homozygously deleted region corresponding to a candidate tumor suppressor locus at distal 17p13.3 in human lung cancer. *Oncogene* 2003; 22(12):1892-1905.
10. McShane, Lisa M., Altman, Douglas S., et al. Reporting recommendations for tumor MARKer prognostic studies (REMARK). *Breast Cancer Res Treat* 2006; 100:229-235.
11. Nagai MA, Pacheco MM, Brentani MM, et al. ALLELIC LOSS ON DISTAL CHROMOSOME-17P IS ASSOCIATED WITH POOR-PROGNOSIS IN A GROUP OF BRAZILIAN BREAST-CANCER PATIENTS. *British Journal of Cancer* 1994; 69(4):754-758.
12. Nguyen DX, Massague J. Genetic determinants of cancer metastasis. *Nature Reviews Genetics* 2007; 8(5):341-352.
13. Nicoll G, Crichton DN, McDowell HE, et al. Expression of the hypermethylated in cancer gene (HIC-1) is associated with good outcome in human breast cancer. *British Journal of Cancer* 2001; 85(12):1878-1882.
14. Nicoll G, Crichton DN, Thompson AM. Loss of expression of the HIC-1 gene (hypermethylated in cancer) from 17P13.3 in breast cancer. *British Journal of Cancer* 2000; 83:P40.
15. O'Connell P. Genetic and cytogenetic analyses of breast cancer yield different perspectives of a complex disease. *Breast Cancer Research and Treatment* 2003; 78(3):347-357.

16. Oldenburg RA, Meijers-Heijboer H, Cornelisse CJ, Devilee P. Genetic susceptibility for breast cancer: How many more genes to be found? *Critical Reviews in Oncology Hematology* 2007; 63(2):125-149.
17. Odelberg SJ, *et al.* (1989) Characterization of eight VNTR loci by agarose gel electrophoresis. *Genomics* 5:915-24.
18. Sjoblom T, Jones S, Wood LD, *et al.* The consensus coding sequences of human breast and colorectal cancers. *Science* 2006; 314(5797):268-274.
19. Thompson AM, Crichton DN, Elton RA, *et al.* Allelic imbalance at chromosome 17p13.3 (YNZ22) in breast cancer is independent of p53 mutation or p53 overexpression and is associated with poor prognosis at medium-term follow-up. *British Journal of Cancer* 1998; 77(5):797-800.
20. Wales MM, Biel MA, Eldeiry W, *et al.* P53 ACTIVATES EXPRESSION OF HIC-1, A NEW CANDIDATE TUMOR-SUPPRESSOR GENE ON 17P13.3. *Nature Medicine* 1995; 1(6):570-577.
21. Welch DR, Steeg PS, Rinker-Schaeffer CW. Molecular biology of breast cancer metastasis - Genetic regulation of human breast carcinoma metastasis. *Breast Cancer Research* 2000; 2(6):408-416.
22. Sequenom, Inc. - DNA Methylation Analysis. Sequenom, Inc. - Genetic Analysis Systems for SNP Genotyping, DNA Methylation And Oncogene Mutation Detection And Profiling. Web. 05 May 2010. <<http://www.sequenom.com/Genetic-Analysis/Applications/EpiTYPER-DNA-Methylation-Analysis/EpiTYPER-Overview.aspx>>.