Spring 5-1-2009

Molecular Analysis of the Genetic Heterogeneity Between Primary and Recurrent Glioblastoma

Anushi Shah

Molecular Analysis of the Genetic Heterogeneity Between Primary and Recurrent Glioblastoma

A Capstone Project Submitted in Partial Fulfillment of the Requirements of the Renée Crown University Honors Program at Syracuse University

Anushi Shah

Candidate for B.S. Biology & Psychology and B.A. Anthropology Degree
and Renée Crown University Honors

May 2009

Honors Capstone Project in: _____Biology_____

Capstone Project Advisor: _____

(Dr. Frank Middleton)

Honors Reader: _____

(Dr. Shannon Novak)

Honors Director: __ _____

Samuel Gorovitz

Date: _____April 21st, 2009

Abstract

Introduction: Glioblastoma multiforme (GBM) is one of the deadliest forms of brain cancer, and affects more than 18,000 new cases each year in the United States alone. The current standard of treatment for GBM includes surgical removal of the tumor, along with radiation and chemotherapy. Despite these treatments, recurrence of GBM is extremely common, and once it recurs, the life expectancy is measured in weeks or months. One of the reasons for the deadly nature of the recurrent GBM is thought to be selection for therapy-resistant tumor cells. In this project, we sought to characterize the molecular changes in recurrent GBM specimens compared to primary GBM specimens from the same subjects.

Methods: Whole-genome DNA microarrays were used to identify genes changed in mRNA expression in seven recurrent GBM samples compared to seven primary GBM samples from the same subjects. Real-time quantitative RT-PCR was used in an attempt to validate changes seen by microarray for 18 genes of interest chosen from the microarray screen.

Results: The microarray experiments identified several dozen mRNA transcripts with evidence of significant differences in expression. From these genes, we chose 18 for PCR validation. Overall, the PCR experiments validated the microarray findings quite well. There was a very high correlation for the magnitude of expression changes seen for the 18 genes (Pearson's $R = 0.852$, $P < 0.001$). Individually, 13 of the 18 genes showed statistically significant changes by PCR in the recurrent versus primary tumor pairs. Of the 5 genes that did not

validate at the $P<0.05$ level, 4 showed trends in the direction predicted by the microarray, while 1 gene did not.

Conclusion: Real time PCR has proven useful for validating changes in recurrent GBMs that could have important clinical applications.

Table of Contents

Total number of Tables: 4

Total number of Figures: 15

Acknowledgements

I would like to take this time to thank a few people that have helped me learn and grow throughout the past four years of my research experience.

First, I would like to thank my mentor, Dr. Frank Middleton, for giving me the opportunity to be part of such an amazing lab the past four years. If it wasn't for his trust in having an amateur freshman in his lab, I would have never been able to grow as a scientist.

Second, I would like to thank Dr. Peter Kim for giving me the chance to become a part of his research project. He has not only been a great educator; but has also helped me understand the importance of enjoying and making the best out of every experience.

Third, I would like to thank Karen Gentile for teaching me the small details of working in a lab such as pipetting and always willing to answer all my questions no matter how busy she might be.

Fourth, I would like to thank my colleagues-- Tanya, Yanli, Julie, and Aditya – for always creating a fun environment in lab!

Last, I would like to thank the Upstate Medical University for giving me the chance to become part of such stimulating research facility.

Abbreviations:

ANTXR1	anthrax toxin receptor 1
CCND2	cyclin D2
cDNA	complementary DNA
CLDN2	claudin 2
Cp	cycles to pass threshold
CSPG4	chondroitin sulfate proteoglycan 4
CSPG5	chondroitin sulfate proteoglycan 5 (neuroglycan C)
EDIL3	EGF-like repeats and discoidin I-like domains 3
EGFR	epidermal growth factor receptor (erythroblastic leukemia viral (v-erb-b) oncogene homolog, avian)
ENPP2	ectonucleotide pyrophosphatase/phosphodiesterase 2 (autotaxin)
GALNT6	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6 (GalNAc-T6)
GBM	glioblastoma multiforme
NCAN	neurocan
NPTX2	neuronal pentraxin II
PCR	polymerase chain reaction

PTPRZ1	protein tyrosine phosphatase, receptor-type, Z polypeptide 1
qPCR	quantitative real-time PCR
RNA	ribonucleic acid
RT	reverse transcription
SNCA	synuclein, alpha (non A4 component of amyloid precursor)
SNCAIP	synuclein, alpha interacting protein (synphilin)
SNP	single nucleotide polymorphism
ST18	suppression of tumorigenicity 18 (breast carcinoma) (zinc finger protein)
STK17A	serine/threonine kinase 17a
SYN2	synapsin II
TPPP	tubulin polymerization promoting protein

Introduction:

Tumors of the brain are one of the toughest types of cancers to study and treat. "Over 34,000 (1.5%) of the 1.1 million cancers diagnosed annually in the United States are found to involve the brain. Approximately 17,000 of these are primary intracranial origin, with the remainder resulting from intracranial metastasis," (Grossman and Loftus, 1999). "Approximately 50% of all tumors occurring within the intracranial cavity fall under the heading 'glioma.' (Grossman and Loftus, 1999). Gliomas are defined as "a hetereogenous collection of neoplasms unified by the fact that they arise from glial tissues," (Grossman and Loftus, 1999). Glioblastoma Multiforme (GBM) is the deadliest form of brain cancer. GBM is a type of astrocytic primary intracranial tumor that

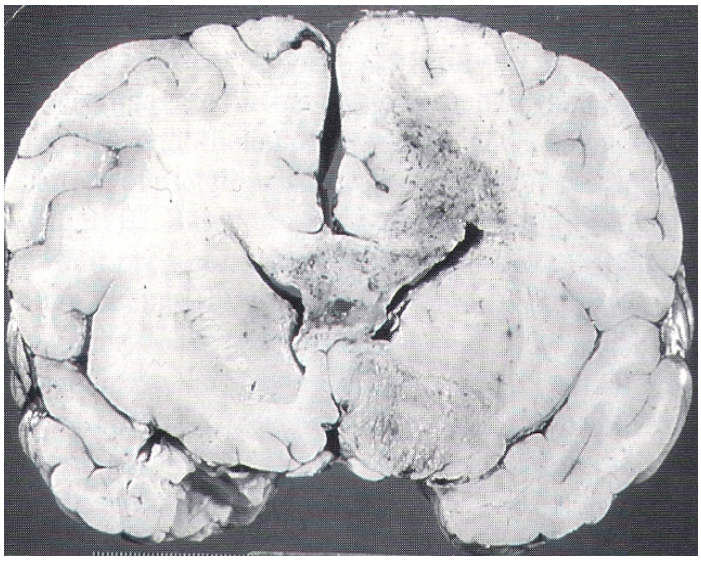

Figure 1: Coronal section anterior view revealing a GBM. Extracted from Grossman and Loftus, 1999.

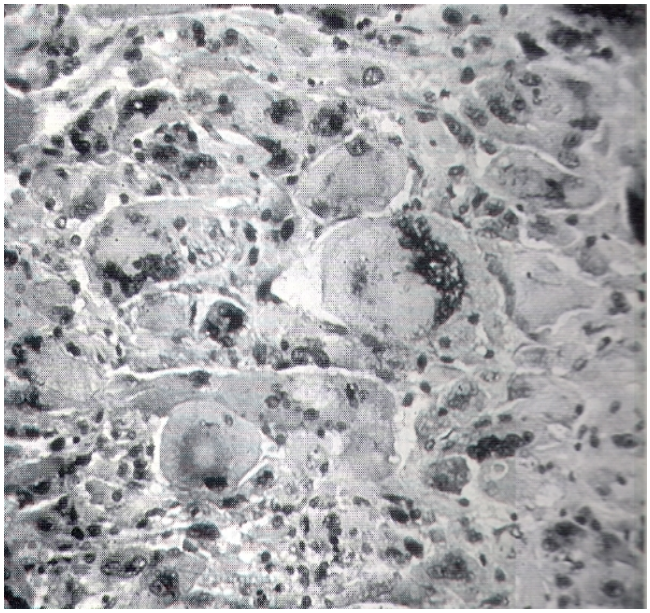

Figure 2: Giant Cell GBM showing enlarged neoplastic glial cells. Extracted from Grossman and Loftus, 1999.

accounts for 15-23% of cases. 35% of these cases are gliomas (Grossman and Loftus, 1999) (***see Figure 1***). As stated before, gliomas are tumors of glial origin (***see Figure 2***). Over 80% of gliomas are derived from astrocytic components of the glia. These tumors, termed astrocytomas, are the most common form of primary intracranial malignancy," (Grossman and Loftus, 1999).Glial cells are non neuronal cells in the brain that are responsible for balance, destruction of pathogens, removal of dead neurons, and provision of insulation and oxygen for other neurons. The World Health Organization (WHO) characterizes GBM as a grade IV astrocytic tumor (Louis et. al). It is the most common and aggressive brain tumor.

These intracranial tumors are the least differentiated and the most aggressive form of astrocytomas (Grossman and Loftus, 1999). Research has shown that the "intratumoral signal heterogeneity is apparent in the vast majority of cases and is thought to represent cystic, hemorrhage, and variability in cellular density," (Grossman and Loftus, 1999). Furthermore, GBMs are characterized by "the presence of an irregular rim of high signal intensity, which probably represents tumor mantle, where hemosiderin deposits, breakdown of the blood brain barrier (BBB), and a high nuclear to cytoplasmic ratio are present," (Grossman and Loftus, 1999).

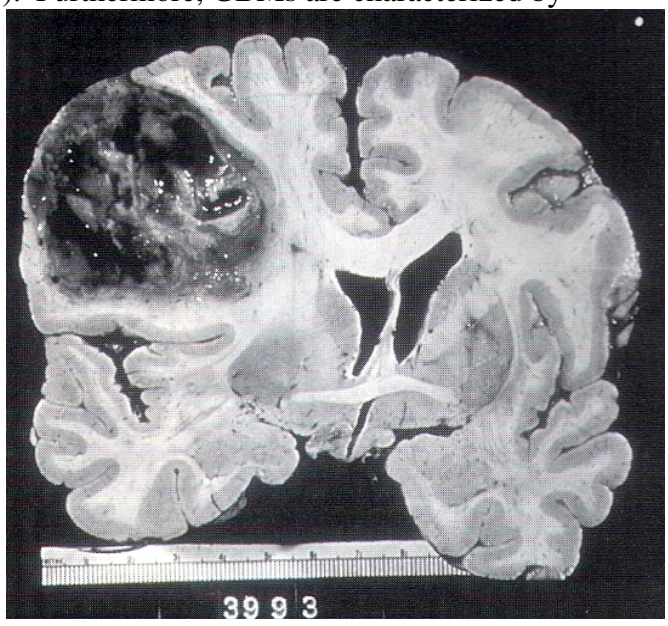

Figure 3: Necrosis due to GBM Extracted from Grossman and Loftus, 1999.

These rapidly dividing cells in the brain derived from glial cells characterized by a centrally located necrotic area, which is also extremely hypoxic (***see Figure 3***). The hypoxic area develops as the tumor outgrows its blood supply and spreads throughout the neighboring cells. The tumor tissue is hypercellular and if examined under a microscope, it would show high cellular density and mitotic activity of the rapidly dividing cells (Grossman and Loftus, 1999). The rapid division gives the tumor the ability to diffuse and infiltrate nearby areas resulting in inferior survival and short time to recurrence (Showalter *et al*., 2007). The necrotic center and the endothelial proliferation are characteristics commonly used to differentiate this tumor from another grade of astrocytoma (Grossman and Loftus, 1999).

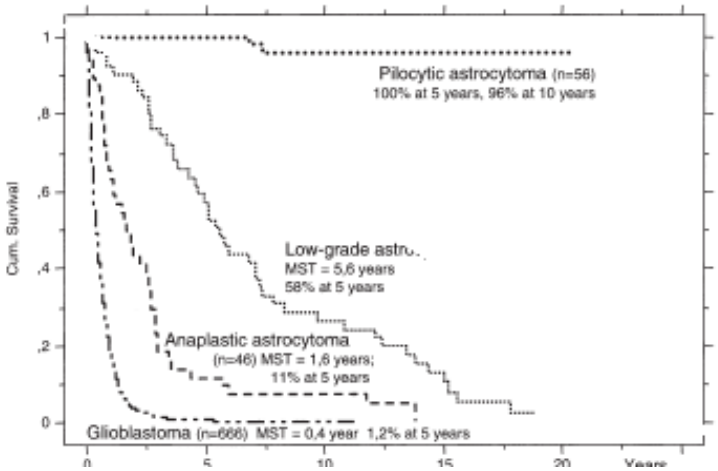

Figure 4: Survival rate of patients with GBMs is poor. Extracted from Ohgaki and Kleihues, 2005).

Research has also indicated that "necrosis is associated with significantly worse prognosis in anaplastic gliomas with both oligdendroglial and astrocytic components," (Louis *et al.*, 2007). The overall median survival rate is lower in patients with a necrotic center than the patients without (Louis *et al.*, 2007).

Despite it being the most common type of brain cancer, there is no definite cure for GBM and research is ongoing but limited by the low life expectancy of

the patients diagnosed with GBM (***see Figure 4***). Treatment of the tumor involves a combination of surgery, radiation, and chemotherapy. The patients are usually given more than one of these treatments at a time. Average survival after the primary GBM treatment is about 12 months (Loftus and Grossman, 1999). "The most common cause of mortality among patients with glioblastoma, regardless of age, is recurrence at the primary tumor site," (Loftus and Grossman, 1999). The average survival after diagnosis of recurrent tumor is only couple months maximum. Researchers and physicians currently recommend aggressive and postoperative radiation because improved survival rates have been correlated with aggressive resection (Grossman and Loftus, 1999). It is important to recognize, though, that combinations of treatment are highly dependent on the health status of the person. With age, the treatment available for the patient is minimal. Younger patients (usually less than 40) are able to withstand aggressive treatment but older patients are not (Ohgaki and Kleihues, 2005) (***see Figure 5***). Furthermore, the effectiveness of chemotherapy treatments in conjunction to radiation and total resection are reduced due to the heterogeneity of

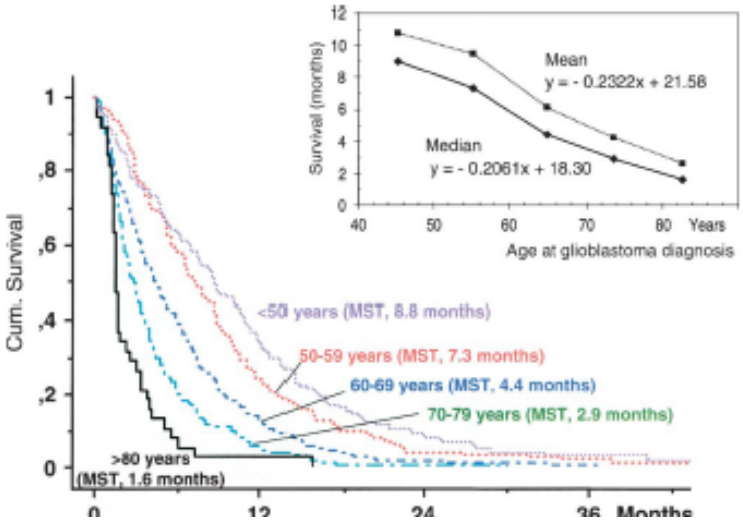

Figure 5: Age at diagnosis plays an important role in the life expectancy o f a GBMpatient. Table shows that the older the patient, the lower the life expectancy after diagnosis. Extracted from Ohgaki and Kleihues, 2005).

the GBM tumor cells. “As a result of this heterogeneity, each subpopulation of cells within a neoplasm may manifest variability in its sensitivity to treatments such as chemotherapy. Additionally, glioblastomas may be heterogeneous with respect to their physical structure, often containing differing microenvironments within the same tumor mass,” (Loftus and Grossman, 1999). Marginally viable but extremely resistant tumor cells can be spotted in the necrotic region of the GBM tumor which could have gotten past the drug penetration presented by radiation treatments or any other type of intervention, thus, lengthening the cell cycle, while using the hypoxia as a cover (Loftus and Grossman, 1999). Current research focuses on what makes the primary vs. recurrent tumor so heterogeneous even though they have essentially originated from the same group of cells.

There are two types of GBMs: Primary and Recurrent. Primary GBMs are tumors that stay in the region of the origin of the malignant mass (***see Figure 6a & 7***). Primary tumors are usually treated with a combination of surgery and radiation. Adjuvant chemotherapy treatment varies on the patient’s profile and health status.

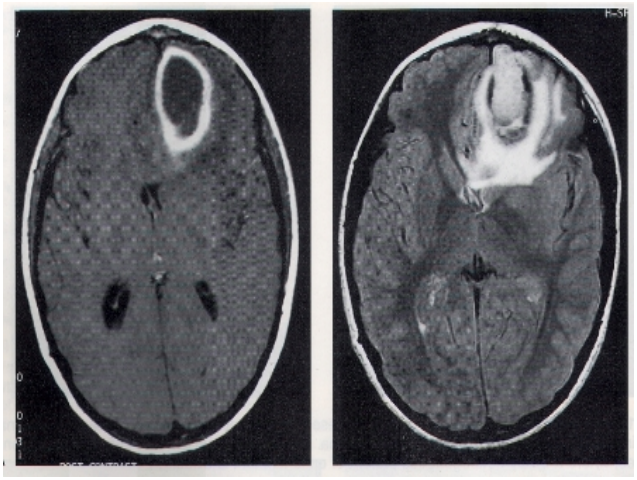

Figure 6a: Primary GBM tumor with a clear site of origin. Extracted from Loftus and Grossman, 1999).

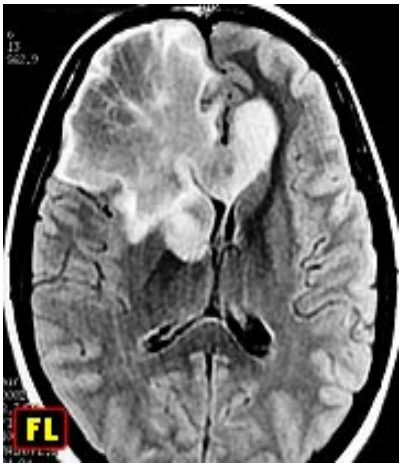

Figure 6b: Recurrent GBM tumor spreading out to different parts of the brain. Original site of tumor can still be seen.

As stated before, aggressive therapy of the primary tumor increases life expectancy to about 8-12 months. On the other hand, recurrent GBMs are tumors that have originated in the brain but then have metastasized to different parts of the brain and possibly other organs as well (***see Figure 6b & 8***). This type of tumor is even harder to treat than the primary tumor because of its pathogenic characteristics. These tumors cells were able to resist primary aggressive therapy and have the ability to proliferate and invade at a higher rate than the primary tumor. Due to this reason, the life expectancy after the treatment or even diagnosis of the recurrent tumor is only about 2 months (at the most). There is a clear heterogeneous difference between the primary and the recurrent tumors, but there is not enough research to pinpoint the difference on a molecular level. Research indicates that there are certain genetic markers that differ from one grade to next grade of GBM. Certain genes end up being over or under expressed as needed for the cancerous cells to proliferate and spread throughout the brain (Ohgaki and Kleihues, 2005).

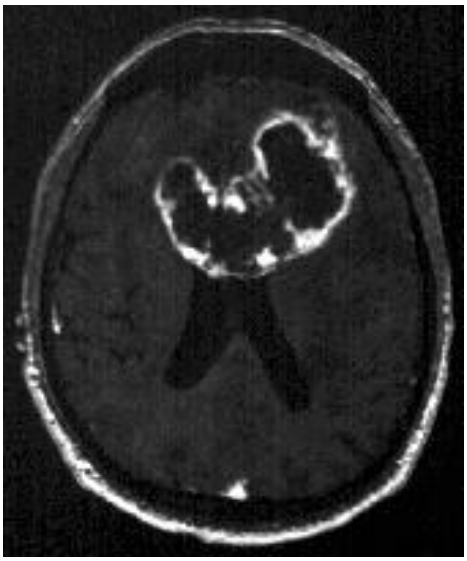

Figure 7:Primary GBM with necrotic center.

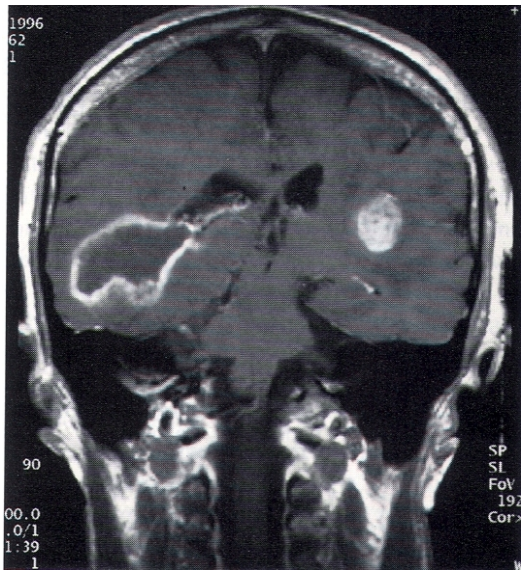

Figure 8: Metastatic primary tumor showing multiple foci. Extracted from Loftus and Grossman, 1999).

Presentation of the tumor and patient profile is very different for GBMs compared to any other type of tumor. Presentation of the tumor in adults shows with severe, unbearable headaches due to the edema and pushing of the tumor against the cranial chamber, seizures, bleeding, decomposition, reduction of the brain's ability to handle stressful situations or stress in general, weak focal vision. In adults, the tumor is mostly supra-tentorial. The tentorium separates the cerebral hemisphere from the brain stem and cerebellum and supra means above. Therefore, in adults the tumor is mostly in the cerebral hemisphere. Moreover, the tumor is mostly large in adults and presents in grade 2, 3, 4 stage. On the other hand, in children, this tumor is usually grade 1and infra-tentorial tumor with cerebellar signs, nausea, and vomiting (Loftus and Grossman, 1999).

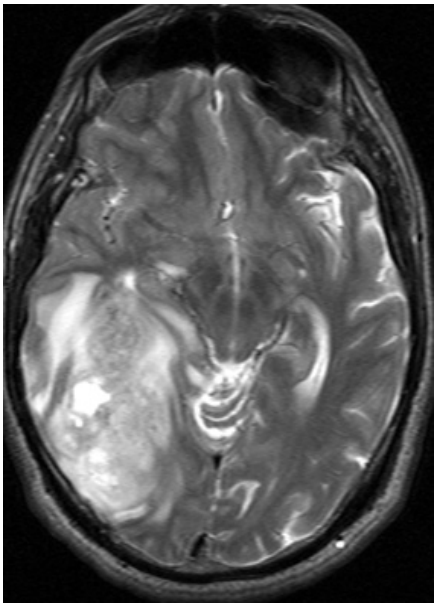

Figure 9: Reccurent GBM taking over the brain. Notice that it is impossible to determine where the brain starts and the tumor ends.

When a patient is presented with a case of GBM, if he/she is in decent health, then the tumor is resected. On average, MOST of the time there is recurrence (***see Figure 9***). The second time, if the person is healthy enough, the physician tries to resect that tumor as well. In either case, the goal is to buy time for the patient. In primary tumors, the person gets about a year, in recurrent

Table 1: Univariate analysis for the Effect of Genetic Alterations on Survival and Mean Age of GBM Patients. Extracted from Ohgaki and Kleihues, 2005).

	Median Survival Time	Mean Age of Patients
TP53 mutation	8.2 months† (n = 126)	52.5 ± 15.8 years§ (n = 126)
Wild-type	7.2 months (n = 259)	59.0 ± 11.9 years (n = 271)
EGFR amplification	6.8 months (n = 127)	59.5 ± 10.8 years (n = 133)
No amplification	7.6 months (n = 244)	57.9 ± 15.0 years (n = 260)
$p16^{INK4a}$ deletion	8.5 months (n = 102)	56.9 ± 10.9 years (n = 102)
No deletion	8.0 months (n = 226)	55.0 ± 13.3 years (n = 228)
PTEN mutation	8.8 months (n = 77)	58.6 ± 11.4 years‖ (n = 77)
Wild-type	8 months (n = 247)	54.2 ± 13.3 years (n = 250)
LOH 10q	7.7 months‡ (n = 185)	56.4 ± 12.1 years¶ (n = 185)
No LOH 10q	9.3 months (n = 84)	52.7 ± 14.6 years (n = 84)

Significantly different in log-rank test (†, p = 0.0173; ‡, p = 0.0561).
Significantly different in Student *t*-test (§, p < 0.001; ‖, p = 0.0091; ¶, p = 0.0313).

treatment, the person gets a few weeks. There is a lot of debate over the palliative treatment because if there is such a small time frame for survival, then why go through the aggressive treatment once or even twice? A GBM case is presented differently than any other form of cancer. Usually, a healthy patient is rushed into the hospital due to severe headache or any other symptom, diagnosed with brain tumor, taken to surgery usually at that visit, and then told that they only have about one year left. The time bought from the surgery is very much appreciated by the patients and it is also helpful for research purposes.

Survival rate of GBM patients is 0.1%; therefore, any type of molecular change that can be determined could be helpful in the understanding of the heterogeneity of the different grades of GBM tumors. Previous research has shown genetic alternation association with GBM tumors. "… LOH 10q was the most frequent genetic alteration (69%), followed by EGFR amplification (34%), TP53 mutations (31%), $p16^{INK4a}$ deletion (31%), and P*TEN* mutations (24%),"(Ohgaki and Kleihues, 2005). Although, "studies on genetic alterations and how they influence response to therapy and survival are usually based on small number of patients, often contradictory and difficult to validate. In recent years, it has been established that primary glioblastomas and secondary glioblastomas derived from low-grade or

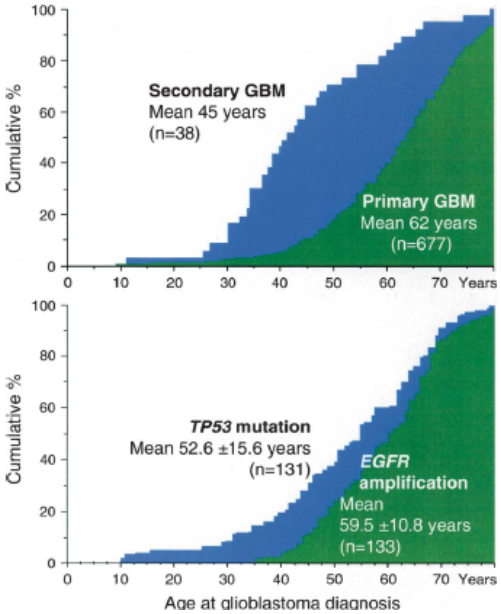

Figure 10: Extracted from Ohgaki and Kleihues, 2005. The image shows the age difference of primary and secondary GBM incidence (top). The image also describes specific gene changes occurance correlating age as a factor during incidence.

anaplastic gliomas develop though different genetic pathways," (Ohgaki and Kleihues, 2005) (***see Figure 10 & Table 1***). The variation differs on many levels, including the age of the patient presenting the tumor. Nevertheless, poor prognosis cannot be defined by genetic alteration especially in older patients. Studies have shown that, "EGFR amplification, TP53 mutations, $p16^{INK4a}$ homozygous deletion, and P*TEN* mutations are considered key genetic events in the evolution of glioblastomas, but the presence or absence of any of these changes does not affect survival," (Ohgaki and Kleihues, 2005)(***see Figure 11 & 12***). "Although there may be as yet unidentified transformation-associated genes that are more frequently altered in glioblastomas of older patients and may affect the susceptibility to therapy, it is also possible that the sum of all changes, i.e. the level of genetic instability, is more relevant," (Ohgaki and Kleihues, 2005). Other molecular research has indicated that hypermethylation of the MGMT promoter region could be associated with long term survival (long term GBM patients are patients who survive more than 36 months) (Krex *et al.*, 2007). The purpose of this research project focuses on the molecular differences of the primary vs. recurrent tumors. Why does a full resection for the recurrent tumors less effective in terms of life expectancy than primary tumor resection?

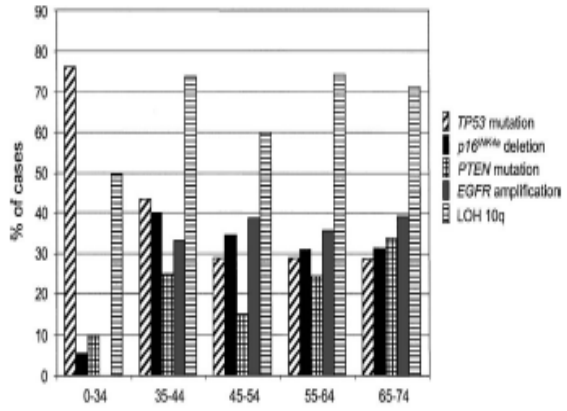

Figure 11: Gene expressional changes according to age of the patient at the time of diagnosis. Extracted from Ohgaki and Kleihues, 2005).

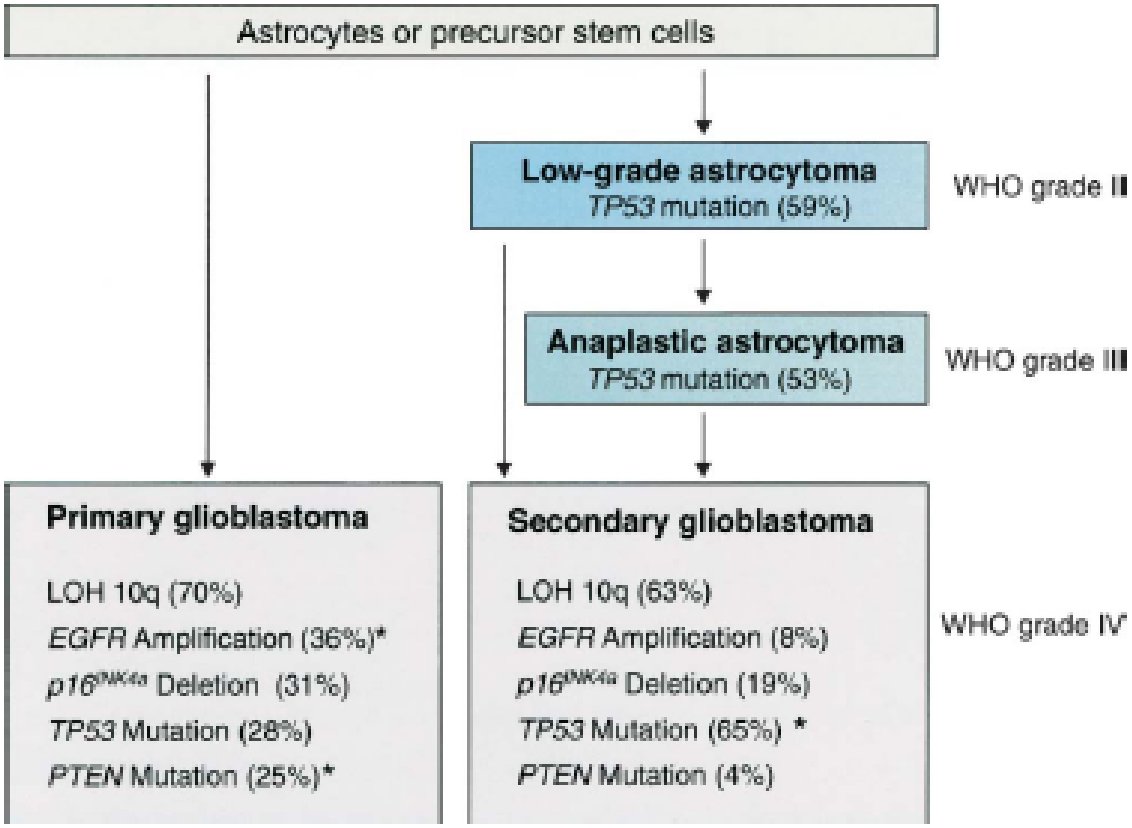

Figure 12: Known gene expressional changes according to grade of astrocytoma. Extracted from Ohgaki and Kleihues, 2005

Since both primary and recurrent tumors are essentially from the same group of cells, the research work focuses on looking at molecular changes in the evolution of cells, DNA copy number changes, and mRNA expression changes. Copy number changes are a marker of many diseases including cancer. Combining the genotype and copy number analyses gives greater insight into the underlying genetic alterations in cancer cells with identification of complex events including loss and reduplication of loci. This paper specifically looks at cancer related gene expression changes that showed consistent change in the microarray analysis in

the recurrent vs. primary tumors of DNA tissue from seven patients with both primary and recurrent GBMs at Upstate Medical University. A total of 18 candidate genes (many with known roles in cancer proliferation) were selected for the study from the microarray results for validation: ANTXR1, CCND2, CLDN2, CSPG3, CSPG4, CSPG5, EDIL3, EGFR, ENPP2, GALNT6, NPTX2, PTPRZ1, SNCA, SNCAIP, ST18, STK17A, SYN2, and TPPP.

Methods:

The Upstate Medical University tumor bank has tissues of seven patients with primary and recurrent GBMs. Three different approaches were used to determine the molecular changes. All three approaches were initiated with microarray GeneChip analysis method. Microarray GeneChip analysis uses probes and targets for different applications such as mutations and differences in expression on a chromosomal level (***see Figure 13***). Screening shows markers that are consistently changing, either increasing or decreasing. The assay mathematically averages out the noise and finds strong, consistent change. A downside of this method is that gene expression varies from person to person even in normal individuals. Therefore, the findings of the microarray analysis need to be validated to make sure that the expressional changes are due to the cancer and not just because normal biological genetic variation. As a result, all

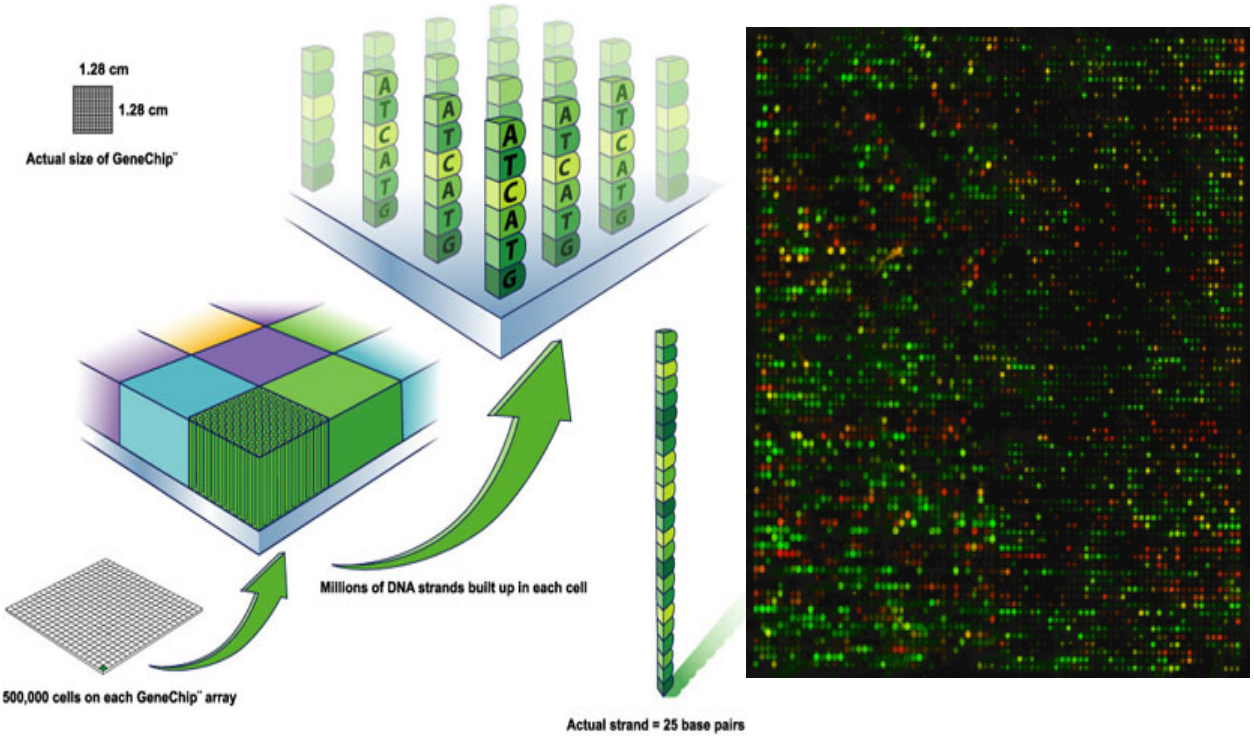

Figure 13: Microarray (GeneChip) analysis method

three approaches were validated with real time quantitative RT-PCR (Real Time qPCR). Real Time qPCR provides highly sensitive quantitative gene transcription measurements by amplifying and quantifying DNA simultaneously. Just like regular polymerase chain reactions, the DNA is denatured, annealed, and extended in Real Time qPCR method. Moreover, Real Time qPCR involves attachment of a SYBR green dye to generate signal values. SYBR green dye specifically binds to double-stranded DNA (dsDNA), but emits no signal when unbound (in the presence of single stranded DNA). The increase in SYBR green fluorescence is measured once during each PCR cycle, and the numbers of cycles it takes for this signal to surpass background is determined (as the delta Cp value).

The real time qPCR studies that were performed were part of the validation for a multilevel analysis of recurrent GBMs that involved three complementary approaches. These three types of approaches are characterized on a convergent flow chart below. RNA expressional changes, DNA copy number changes, and the promoter methylation changes were focused upon in order to see which genes play an essential role that could help explain the recurrent vs. primary tumor changes and ultimately help identify new treatments for GBM (***see Figure 14***).

Once the reactions were run for the three approaches, we asked two pertinent questions: a) Are there any common genes implicated from the approach? and b) Are these genes implicated in cancer?

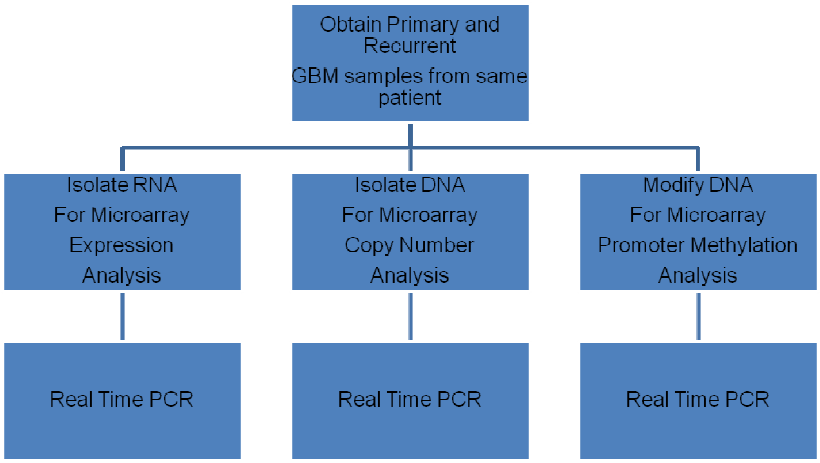

Figure 14: 3-way GBM research method for this study

Microarray Methods

Genomic DNA and RNA were purified using standard protocols (Qiagen RNeasy kit for RNA and the Epicentre DNA purification kit for DNA). The yield, purity, and integrity of RNA and DNA were confirmed by UV spectrophotometry and the Agilent Bioanalyzer NanoChip. Changes in RNA expression were determined using the Affymetrix U133 plus 2.0 Gene Chip which assays for hybridization for over 47,000 fragments representing over 38,000 known genes. A single microarray (Human Genome Xba SNP Array) containing probes for ~58,000 single nucleotide polymorphisms (SNPs) was also run on the DNA from each tumor sample to probe for changes in DNA copy number, and chromosomal rearrangements. Lastly, because previous studies have shown that methylation of DNA at CpG islands in the promoter sites of certain genes correlates with sensitivity of glioma cells to chemotherapeutics (including temozolamide, Everhard 2006), we also mapped changes in genome-wide promoter methylation status using a novel assay developed in the SUNY Microarray Core Facility that is analyzed with the Human Promoter 1.0 Array (Affymetrix).

Real Time RT-PCR Methods

Quantitative real time RT-PCR (qRT-PCR) was performed to validate the changes seen by microarray screens. Briefly, 500ng of total RNA (in 12 µL PCR grade water) from each RNA preparation was used in an integrated reverse transcription (RT) first and second strand cDNA synthesis procedure that incorporates removal of genomic DNA contamination (QiantiTect, Qiagen, Valencia, CA). The RT reaction was terminated by heating to 95^{o}C for 3 min, and diluted to a volume of 100µL for use in qRT-PCR. For quantification of transcript differences, 1.0 µL of the RT reaction from each of the samples was evaluated in duplicate PCR reactions for each gene of interest on 384-well plates in a Roche LightCycler 480 instrument (Roche, Indianapolis, IN).

Each qRT-PCR reaction was performed in duplicate using the Roche SYBR Green Master Mix, in 10 µL volumes as follows: Activation at 95^{o}C for 5 min, followed by 40 cycles of 95^{o}C for 10 sec, 58^{o}C for 15 sec, and 72^{o}C for 15 sec. Amplification in the absence of cDNA template was used to verify the absence of signal that would have occurred due to primer dimerization and extension or DNA contamination and end point melt-curve analysis was used to confirm the presence of single amplicons in each reaction well.

Analyses of the data from the qRT–PCR studies was performed by first determining the number of PCR cycles that it took for each reaction to pass the detection threshold (Cp). Then, to control for differences in starting material, these values were normalized for each gene from a specific sample by subtracting

the Cp value of 18S rRNA to generate a ΔCp value. Statistical significance of the PCR based differences in expression was determined using a 1-tailed pairwise Student's T test. Pairwise relative differences in expression between the recurrent and primary tumors from the same subject were then determined by subtracting the ΔCp value of each gene for the recurrent tumor from the ΔCp value of the primary tumor. These differences were termed ΔΔCp values, and represent the Log2 difference in expression in recurrent versus primary GBMs. For interpretive purposes, the Log2 differences can be converted to fold changes using the formula Fold Change = $2^{\Delta\Delta Cp}$. In addition to examining the data for evidence of validating the overall change in expression seen by microarray, we also calculated the overall correlation in the pairwise differences observed between the recurrent and primary tumor using a Pearson's R. All of these values are displayed in the results that follow. qPCR validation of the genes CLDN2, EDIL3, ENPP2, NPTX2, PTPRZ1, and STK17A was performed by Dr. Peter Kim. Validations of the remaining 12 genes were the focus of my research work: ANTXR1, CCND2, CSPG3, CSPG4, CSPG5, EGFR, GALNT6, SNCA, SNCAIP, ST18, SYN2, and TPPP.

Results:

The results of this paper focus on the microarray expression analysis and their validation using real time qPCR. On the microarray screen, some of the genes are repeated because there was more than one probe for that particular gene. T-test results show that most of the genes with expressional differences are statistically significant. The values that are not are the genes with more than one probe. Majority of the probes for that specific gene are still statistically significant. Expressional changes were averaged according to primary minus recurrent tumors from each of the seven patients on a log 2 scale. A negative value in the average log 2 column indicated a decrease in expression while a positive value indicates an increase in gene expression.

Among the initial genes that were validated by Dr. Peter Kim, PTPRZ1 and STK17A showed a decrease in expression while CLDN2, EDIL3, ENPP2, NPTX2, and SYN2 and showed an increase in expression in the recurrent tumor vs. primary tumor (***see Table 2***). Round 2 candidate genes ANTXR1, CCND2, CSPG3, CSPG4, CSPG5, EGFR, and SNCAIP showed a decrease in expression whereas GALNT6, SNCA, ST18, SYN2, and TPPP showed an increase in expression in the recurrent tumor vs. primary tumor (***see Table 3***).

Overall, there was a high correlation for the magnitude of expression changes seen for the 18 genes by both approaches (Pearson's $R = 0.852$; $P < 0.001$; ***see Figure 15***). Individually, 13 of the 18 genes showed statistically significant changes by PCR in the recurrent versus primary tumor pairs. Of the 5

genes that did not validate at the P<0.05 level, 4 showed trends in the direction predicted by the microarray, while 1 gene did not.

Table 2: Preliminary (round 1) microarray screen expressional changes and qPCR expressional changes. qPCR t – test analysis indicates whether the expressional changes from microarray screen were validated.

Preliminary Validation (Fall 2007, performed by Dr. P. Kim)

	Pairwise	Pairwise	Pairwise Log2 Differences by Microarray Screen						
Gene	T Test P	Log2 Diff	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Pair 7
Claudin	0.000005	0.18	0.06	1.10	0.77	0.49	-0.64	0.12	-0.65
ENPP2	0.032863	1.99	0.37	0.00	1.44	0.49	5.88	3.47	2.25
EDIL3	0.054454	1.61	0.66	-0.05	1.17	0.76	4.45	2.77	1.50
NPTX2	0.076730	1.51	0.03	1.02	0.67	0.90	4.58	1.62	1.75
PTPRZ1	0.009154	-1.81	-2.06	-0.82	-4.80	-2.36	-1.40	-0.61	-0.62
STK17A	0.007560	-1.57	-2.41	-2.17	-2.89	-1.15	-1.11	-1.08	-0.21

			Pairwise Log2 Differences by PCR						
Gene	T Test P	Log2 Diff	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Pair 7
Claudin	0.005773	2.61	1.63	1.33	2.76	2.05	6.37	3.60	0.55
ENPP2	0.005117	2.11	1.85	0.72	2.50	1.33	4.43	3.62	0.30
EDIL3	0.005230	2.70	1.03	1.01	2.98	1.79	6.64	3.34	2.13
NPTX2	0.005251	2.70	-0.25	2.97	3.61	2.91	5.98	2.54	1.15
PTPRZ1	0.008861	-1.88	-1.07	-0.15	-4.79	-1.86	-2.40	-0.61	-2.30
STK17A	0.021679	-1.03	-0.41	-0.88	-2.08	0.97	-1.67	-1.18	-1.98

Table 3: Round 2 microarray screen expressional changes and qPCR expressional changes. qPCR t – test analysis indicates whether the expressional changes from microarray screen were validated.

Round 2 Validation

			Pairwise Log2 Differences by Microarray Screen						
Gene	T Test P	Log2 Diff	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Pair 7
ANTXR1	0.003806	-1.31	-2.00	-1.70	-2.28	-2.41	-0.54	-0.40	0.18
CCND2	0.022083	-1.01	-1.91	-0.64	-2.15	-0.89	-0.50	-0.80	-0.19
CSPG3	0.036635	-1.44	-2.00	-2.05	-2.12	-2.33	-1.24	0.04	-0.37
CSPG4	0.002523	-1.53	-3.10	-1.91	-2.34	-0.69	-0.46	-2.61	0.40
CSPG5	0.001982	-1.17	-1.87	-0.62	-2.82	-0.78	-0.81	-1.31	0.00
EGFR	0.037536	-0.86	-3.53	0.68	-0.63	-0.47	-1.22	-1.01	0.15
GALNT6	0.010349	1.69	2.82	1.42	1.24	3.14	2.17	1.72	-0.66
SNCA	0.001518	1.17	1.47	0.55	2.74	0.02	2.25	1.53	-0.35
SNCAIP	0.104891	-1.01	-1.56	-0.47	-2.94	-0.53	-0.48	-0.58	-0.47
ST18	0.004882	1.67	0.26	0.92	2.25	0.37	4.36	2.60	0.91
SYN2	0.000004	0.41	0.63	0.92	0.91	0.26	-0.22	0.51	-0.11
TPPP	0.018529	1.56	0.04	0.42	1.46	0.34	4.19	2.06	2.41

			Pairwise Log2 Differences by PCR						
Gene	T Test P	Log2 Diff	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Pair 7
ANTXR1	0.063626	-1.30	0.11	-1.02	0.22	-1.05	-5.32	-2.05	-0.01
CCND2	0.021110	-1.97	-0.41	-0.20	-3.43	-0.64	-5.80	-2.03	-1.29
CSPG3	0.029636	-2.05	0.68	-0.87	-2.11	-3.79	-6.35	-0.98	-0.93
CSPG4	0.020126	-2.32	-0.85	-1.42	-1.79	-0.74	-6.94	-4.00	-0.51
CSPG5	0.018768	-2.15	-0.07	-1.30	-3.07	-1.62	-6.25	-2.67	-0.08
EGFR	0.122273	-0.98	-4.50	-1.87	-0.24	0.02	-0.90	-1.45	2.08
GALNT6	0.041500	0.98	3.54	0.93	0.88	1.04	0.79	-0.03	-0.31
SNCA	0.228274	0.47	1.76	-0.45	-1.55	1.01	0.14	-0.61	2.98
SNCAIP	0.032785	-3.27	-0.84	-2.17	-2.16	-10.76	-5.79	-2.02	0.81
ST18	0.237864	0.35	1.13	0.12	1.05	1.49	0.97	-0.38	-1.95
SYN2	0.468723	-0.03	0.10	-0.37	0.60	1.84	-0.20	-0.81	-1.39
TPPP	0.033173	2.04	-2.23	1.07	5.23	0.90	2.78	2.71	3.80

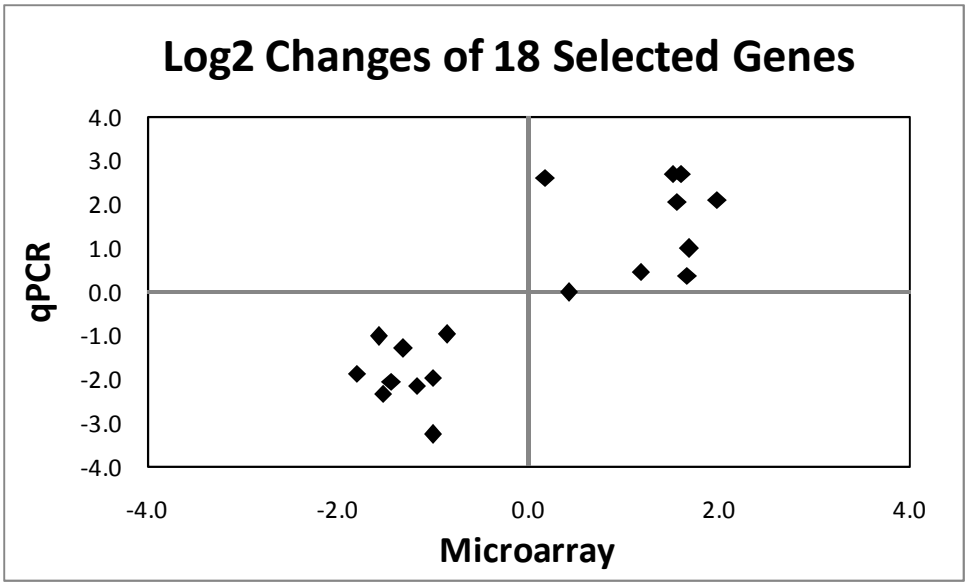

Figure 15: Comparison of changes seen in recurrent versus primary GBM for 18 genes in initial microarray screen, and real time qPCR. Overall, there was a very high correlation in the difference seen for the two techniques (Pearson's R = 0.852, P<0.001).

Conclusion:

The overall high correlation between the microarray expression analysis and the qPCR validation results indicates that Real time PCR has proven useful for validating changes in recurrent GBMs that could have important clinical applications. Of the round 2 gene validation, 7 out of 12 genes were statistically significant for expression change. The gene functions are presented in ***table 4***. The biology of these seven genes can be related to tumor functions.

CCND2 is a cell cycle regulator (Lossos 2004). The average expression of CCND2 is decreased in the recurrent tumor compared to the primary tumor, thus, indicating that the high proliferation rate of the recurrent tumor may be the result of improper function of cell cycle regulator.

CSPG3 (also known as NCAN), CSPG4, and CSPG5 belong to the same family of genes. They are known to be responsible for the regulation of the extracellular matrix (Gladson 1999 & Zhang 2003). Real time validation showed that the average expressions in all three genes are decreased in the recurrent tumor. This shows that since the matrix is not regulated, the tumor cells can easily spread to the neighboring cells because there is not traffic in the extracellular matrix to slow down the proliferation of the tumor cells.

GALNT6 is known to be expressed in low levels in the brain. It is associated with fibronectin glycosylation kinetics. Furthermore, members of this family are responsible for the transfer of N-acetylgalactoamine to serine and threonine residues (a type of protein glycosylation) (Bennett 1999).The

expression of this gene in the recurrent tumor is shown to increase. Higher than normal levels of this gene in the brain could lead to increased glycosylation and mutations in the cell cycle, thus, promoting abnormal cell proliferation.

SNCAIP is a synuclein interacting protein (Chung 2001). The validation of this gene was chosen due to significant change of SNCA gene in the microarray analysis. We were not able to validate the SNCA expression change and therefore cannot make any conclusions about the SNCAIP expression change.

TPPP is known to promote aggregation of SNCA gene (Lindersson 2005). Even though expression of this gene is significantly increased in the recurrent tumor compared to the primary tumor, we cannot draw any conclusions about the role of this gene in tumor proliferation since the SNCA gene validation was not statistically significant. Nevertheless, the SNCA gene expression change from microarray analysis and real time PCR did show a trend in the same direction. SNCA has been known to play a role in neurodegenerative disorders; but more research needs to be conducted to analyze its correlation with TPPP gene and their role in proliferation of brain tumors.

Even though we were not able to validate the expressional change for ANTXR1 gene, it is worth mentioning that latest research has shown some correlation between ANTXR1 gene and other types of cancers. ANTXR1 has been known to be involved in vasculogenesis (Nanda 2004). Vasculogenesis involves the recruitment of blood supply for the tumor by forming new blood

vessels. If the blood supply growth can be shut down, then the tumor can be eliminated. This could be useful for future therapeutic interventions.

This study shows that Microarray analysis and real time PCR can be used together to understand the gene expression change in GBMs. The findings can be used to improve upon and/or invent new treatment options for GBM patients which could, if not cure, increase the life expectancy of GBM patients.

Table 4: 18 gene functions with references that evaluate their relevance to cancer proliferation. The highlighted genes were validated by qPCR analysis. Red indicates a decrease in expression change and green indicates an increase in expression change in the recurrent GBM versus primary GBM

Gene	Function	Reference
ANTXR1	vasculogenesis	Nanda 2004
CCND2	cell cycle regulator	Lossos 2004
CLDN2	Tight junction	Thakur 2007
CSPG 3	regulation of extracellular matrix	Zhang 2003
CSPG4	regulation of extracellular matrix	Gladson 1999
CSPG5	regulation of extracellular matrix	Gladson 1999
EDIL3	angiogenesis	Aoka 2002
EGFR	cell signaling molecule	Wang 2004
ENPP2	adherence, motility	Kishi 2006
GALNT6	galactosaminotranserase	Bennett 1999
NPTX2	neuronal development	Carlson 2007
PTPRZ1	cell signaling	Lu 2005
SNCA	neurodegenerative disorders	Beyer 2008
SNCAIP	SNCA interacting protein	Chung 2001
ST18	transcriptional regulation	Steinbach 2006
STK17A	tumor suppressor	Wittig 2002
SYN2	synaptogenesis	Lee 2005
TPPP	promotes aggregation of SNCA	Lindersson 2005

References

Aoka *et al*. The embryonic angiogenic factor Del1 accelerates tumor growth by enhancing vascular formation. *Microvasc Res*. 2002 Jul;64(1):148-61.

Bennett *et al*. Cloning and characterization of a close homologue of human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy. *J Biol Chem.* 1999 Sep 3;274(36):25362-70.

Beyer *et al*. Identification and characterization of a new alpha-synuclein isoform and its role in Lewy body diseases. *Neurogenetics* 9: 15-23, 2008.

Carlson *et al*. Relationship between survival and edema in malignant gliomas: role of vascular endothelial growth factor and neuronal pentraxin 2. *Clin Cancer Res.*2007 May 1;13(9):2592-8.

Chung *et al*. Parkin ubiquitinates the alpha-synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease. *Nat Med.*2001 Oct;7(10):1144-50

Everhard et. al. MGMT methylation: a marker of response to temozolomide in low-grade gliomas. Ann Neurol.2006 Dec;60(6):740-3.

Gladson, C. L. The extracellular matrix of gliomas: modulation of cell function. *J Neuropathol Exp Neurol.* 1999 Oct;58(10):1029-40.

Grossman, R., and Loftus, C. Principles of Neurosurgery. Philadelphia: Lippincott-Raven, 1999. 469-520.

Hoelzinger *et al*. Gene expression profile of glioblastoma multiforme invasive phenotype points to new therapeutic targets. *Neoplasia*. 2005 Jan;7(1):7-16.

Hoelzinger *et al.* Autotaxin: a secreted autocrine/paracrine factor that promotes glioma invasion. *J Neurooncol.* 2008 Feb;86(3):297-309.

Kishi *et al.* Autotaxin is overexpressed in glioblastoma multiforme and contributes to cell motility of glioblastoma by converting lysophosphatidylcholine to lysophosphatidic acid. *J Biol Chem*. 281 (25) 17492-500 2006

Krex *et al*. Long-term survival with glioblatoma multiforme. *Brain*. 2007 Oct;130(Pt 10):2596-606

Lee *et al*. Association study of polymorphisms in synaptic vesicle-associated genes, SYN2 and CPLX2, with schizophrenia.*Behav Brain Funct.* 2005 Aug 31;1:15.

Lindersson *et al*. p25alpha Stimulates alpha-synuclein aggregation and is co-localized with aggregated alpha-synuclein in alpha-synucleinopathies. *Behav Brain Funct.* 2005 Aug 31;1:15.

Lossos, *et al*. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *N Engl J Med*. 2004 Apr 29;350(18):1814-6.

Louis *et al*. The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropathol*. 2007 Aug;114(2):97-109

Lu *et al.* Differential induction of glioblastoma migration and growth by two forms of pleiotrophin. *J Biol Chem*. 2005 Jul 22;280(29):26953-64

Luo *et al*. Transcriptomic and genomic analysis of human hepatocellular carcinomas and hepatoblastomas. *Hepatology*. 2006 Oct;44(4):1012-24.

Manning *et al*. Role of lysophosphatidic acid and rho in glioma cell motility.*Cell Motil Cytoskeleton.* 2000 Mar;45(3):185-99.

Nanda *et al*. TEM8 interacts with the cleaved C5 domain of collagen alpha 3(VI). *Cancer Res.*2004 Feb 1;64(3):817-20

Ohgaki, H. and Kleihues, P. Population – Based Studies on Incidence, Survival Rates, and Genetic Alterations in Astrocytic and Oligodendroglial Gliomas. *J Neuropathol Exp Neurol*. 2005 Jun;64(6):479-89

Sanjo *et al*. DRAKs, novel serine/threonine kinases related to death-associated protein kinase that trigger apoptosis. *J Biol Chem.* 1998 Oct 30;273(44):29066-71.

Showalter *et al*. Multifocal Glioblastoma Multiforme: Prognostic Factors and Patterns of Progression. *Int J Radiat Oncol Biol Phys*. 2007 Nov 1;69(3):820-4

Staub *et al*. A novel repeat in the melanoma-associated chondroitin sulfate proteoglycan defines a new protein family. *FEBS Lett.*2002 Sep 11;527(1-3):114-8

Steinbach *et al*. Identification of a set of seven genes for the monitoring of minimal residual disease in pediatric acute myeloid leukemia.*Clin Cancer*

*Res.*2006 Apr 15;12(8):2434-41

Thakur *et al*. Aberrant expression of X-linked genes RbAp46, Rsk4, and Cldn2 in breast cancer. *Mol Cancer Res.* 2007 Feb;5(2):171-81

Wang *et al*. Epidermal growth factor receptor-deficient mice have delayed primary endochondral ossification because of defective osteoclast recruitment. *J Biol Chem*. 2004 Dec 17;279(51):53848-56

Wittig *et al*. Candidate genes for cross-resistance against DNA-damaging drugs. *Cancer Res.* 2002 Nov 15;62(22):6698-705

Zhang *et al*. Accumulation of neurocan, a brain chondroitin sulfate proteoglycan, in association with the retinal vasculature in RCS rats. *Invest Ophthalmol Vis Sci.*2003 Mar;44(3):1252-61.

Summary:

Molecular Analysis of the Genetic Heterogeneity Between Primary and Recurrent Glioblastoma

Glioblastoma multiforme (GBM) is one of the deadliest forms of brain cancer, and affects more than 18,000 new cases each year in the United States alone. Gliomas are defined as "a hetereogenous collection of neoplasms unified by the fact that they arise from glial tissues," (Grossman and Loftus, 1999). Glial cells are non neuronal cells in the brain that are responsible for maintenance of homeostasis, destruction of pathogens, removal of dead neurons, and provision of insulation and oxygen for other neurons. There are two types of GBMs: Primary and Recurrent. Primary GBMs are when the tumor remains at the site of origin. On the other hand, in recurrent GBMs, the tumor has spread out to other parts of the brain from the site of origin. Presentation of the tumor and patient profile is very different for GBMs compared to any other type of tumor. Presentation of the tumor in adults shows with severe, unbearable headaches due to the edema and pushing of the tumor against the cranial chamber, seizures, bleeding, decomposition, reduction of the brain's ability to handle stressful situations or stress in general, weak focal vision (Loftus and Grossman, 1999).

The current standard of treatment for GBM includes surgical removal of the tumor, along with radiation and chemotherapy. Despite these treatments, recurrence of GBM is extremely common, and once it recurs, the life expectancy is measured in weeks or months. The treatment plan for recurrent tumor is the same as the primary tumor depending on the health status of the patient. GBMs

are more common in older age individuals; therefore, the patient's health status is usually not strong enough to handle another surgical operation with radiation and chemotherapy treatments. There is a lot of debate over the palliative treatment because if there is such a small time frame for survival, then why go through the aggressive treatment once or even twice? A GBM case is presented differently than any other form of cancer. Usually, a healthy patient is rushed into the hospital due to severe headache or any other symptom, diagnosed with brain tumor, taken to surgery usually at that visit, and then told that they only have about one year left. The time bought from the surgery is very much appreciated by the patients and it is also helpful for research purposes.

One of the reasons for the deadly nature of the recurrent GBM is thought to be selection for therapy-resistant tumor cells. Research indicates that there are certain genetic markers that differ from one grade to next grade of GBM. Certain genes end up being over or under expressed as needed for the cancerous cells to proliferate and spread throughout the brain (Ohgaki and Kleihues, 2005). In this project, we specifically looked at cancer related gene expression changes that showed consistent change in the microarray analysis in the recurrent vs. primary tumors of DNA tissue from seven patients with both primary and recurrent GBMs at Upstate Medical University. A total of 18 candidate genes (many with known roles in cancer proliferation) were selected for the study from the microarray results for validation: ANTXR1, CCND2, CLDN2, CSPG3, CSPG4, CSPG5, EDIL3, EGFR, ENPP2, GALNT6, NPTX2, PTPRZ1, SNCA, SNCAIP, ST18, STK17A, SYN2, and TPPP.

Whole-genome DNA microarrays were used to identify genes changed in mRNA expression in seven recurrent GBM samples compared to seven primary GBM samples from the same subjects. Microarray GeneChip analysis uses probes and targets for different applications such as mutations and differences in expression on a chromosomal level. Screening shows markers that are consistently changing, either increasing or decreasing. The assay mathematically averages out the noise and finds strong, consistent change. A downside of this method is that gene expression varies from person to person even in normal individuals. Therefore, the findings of the Microarray analysis need to be validated to make sure that the expressional changes are due to the cancer and not just because normal biological genetic variation. As a result, Real-time quantitative RT-PCR was used in an attempt to validate changes seen by microarray for 18 genes of interest chosen from the microarray screen. Real Time qPCR provides highly sensitive quantitative gene transcription measurements by amplifying and quantifying DNA simultaneously. Once the reactions were run, we asked two pertinent questions: a) Are there any common genes implicated from the approach? and b) Are these genes implicated in cancer?

The microarray experiments identified several dozen mRNA transcripts with evidence of significant differences in expression. From these genes, we chose 18 for PCR validation. The results of this paper focus on the microarray expression analysis and their validation using real time qPCR. On the microarray screen, some of the genes are repeated because there was more than one probe for that particular gene. T-test results show that most of the genes with expressional

differences are statistically significant. The values that are not are the genes with more than one probe. Majority of the probes for that specific gene are still statistically significant. Expressional changes were averaged according to primary minus recurrent tumors from each of the seven patients on a log 2 scale. A negative value in the average log 2 column indicated a decrease in expression while a positive value indicates an increase in gene expression. Overall, the PCR experiments validated the microarray findings quite well. There was a very high correlation for the magnitude of expression changes seen for the 18 genes (Pearson's $R = 0.852$, $P < 0.001$). Individually, 13 of the 18 genes showed statistically significant changes by PCR in the recurrent versus primary tumor pairs. Of the 5 genes that did not validate at the $P<0.05$ level, 4 showed trends in the direction predicted by the microarray, while 1 gene did not.

The overall high correlation between the microarray expression analysis and the qPCR validation results indicates that Real time PCR has proven useful for validating changes in recurrent GBMs that could have important clinical applications. Of the round 2 gene validation, 7 out of 12 genes were statistically significant for expression change. The biology of these seven genes can be related to tumor functions.

- CCND2 is a cell cycle regulator (Lossos 2004). The average expression of CCND2 is decreased in the recurrent tumor compared to the primary tumor, thus, indicating that the high proliferation rate of the recurrent tumor may be the result of improper function of cell cycle regulator.

- CSPG3 (also known as NCAN), CSPG4, and CSPG5 belong to the same family of genes. They are known to be responsible for the regulation of the extracellular matrix (Gladson 1999 & Zhang 2003). Real time validation showed that the average expressions in all three genes are decreased in the recurrent tumor. This shows that since the matrix is not regulated, the tumor cells can easily spread to the neighboring cells because there is not traffic in the extracellular matrix to slow down the proliferation of the tumor cells.
- GALNT6 is known to be expressed in low levels in the brain. It is associated with fibronectin glycosylation kinetics. Furthermore, members of this family are responsible for the transfer of N-acetylgalactoamine to serine and threonine residues (a type of protein glycosylation) (Bennett 1999).The expression of this gene in the recurrent tumor is shown to increase. Higher than normal levels of this gene in the brain could lead to increased glycosylation and mutations in the cell cycle, thus, promoting abnormal cell proliferation.
- SNCAIP is a synuclein interacting protein (Chung 2001). The validation of this gene was chosen due to significant change of SNCA gene in the microarray analysis. We were not able to validate the SNCA expression change and therefore cannot make any conclusions about the SNCAIP expression change.
- TPPP is known to promote aggregation of SNCA gene (Lindersson 2005). Even though expression of this gene is significantly increased in the

recurrent tumor compared to the primary tumor, we cannot draw any conclusions about the role of this gene in tumor proliferation since the SNCA gene validation was not statistically significant. Nevertheless, the SNCA gene expression change from microarray analysis and real time PCR did show a trend in the same direction. SNCA has been known to play a role in neurodegenerative disorders; but more research needs to be conducted to analyze its correlation with TPPP gene and their role in proliferation of brain tumors.

- Even though we were not able to validate the expressional change for ANTXR1 gene, it is worth mentioning that latest research has shown some correlation between ANTXR1 gene and other types of cancers. ANTXR1 has been known to be involved in vasculogenesis (Nanda 2004). Vasculogenesis involves the recruitment of blood supply for the tumor by forming new blood vessels. If the blood supply growth can be shut down, then the tumor can be eliminated. This could be useful for future therapeutic interventions.

This study shows that Microarray analysis and real time PCR can be used together to understand the gene expression change in GBMs. The findings can be used to improve upon and/or invent new treatment options for GBM patients which could, if not cure, increase the life expectancy of GBM patients.