LEBANESE AMERICAN UNIVERSITY

Characterizing the Extent of Genetic Aberrations in Glioblastoma Multiforme using Bioinformatics Tools

By

Marcelino A. Al Ghafari

A thesis
submitted in partial fulfillment of the requirements

for the degree of Master of Science in Biological Sciences

School of Arts and Sciences

May 2022

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

THESIS APPROVAL FORM

Student Name: Marcelino Al ghafari I.D. #: 201201540

Thesis Title: Characterizing the Extent of Genetic Aberrations in Glioblastoma Multiforme using Bioinformatics tools.

Program: MS in Biological Sciences

Department: Natural Sciences

School: Arts & Sciences

The undersigned certify that they have examined the final electronic copy of this thesis and approved it in Partial Fulfillment of the requirements for the degree of:

MS in the major of Biologcal Sciences

Thesis Advisor's Name: Ralph Abi Habib

Signature: [redacted] Date: 09 / 05 / 2022 (Day / Month / Year)

Committee Member's Name: Roy Khalaf

Signature: [redacted] Date: 09 / 05 / 2022 (Day / Month / Year)

Committee Member's Name: Michella Ghassibe Sabbagh

Signature: [redacted] Date: 09 / 05 / 2022 (Day / Month / Year)

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

THESIS COPYRIGHT RELEASE FORM

LEBANESE AMERICAN UNIVERSITY NON-EXCLUSIVE DISTRIBUTION LICENSE

By signing and submitting this license, you (the author(s) or copyright owner) grants the Lebanese American University (LAU) the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic formats and in any medium, including but not limited to audio or video. You agree that LAU may, without changing the content, translate the submission to any medium or format for the purpose of preservation. You also agree that LAU may keep more than one copy of this submission for purposes of security, backup and preservation. You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant LAU the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission. IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN LAU, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT. LAU will clearly identify your name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

Name: **Marcelino Al Ghafari**

Signature: [redacted]

Date: **09** / **05** / **2022**
Day Month Year

LAU
الجامعة اللبنانية الأميركية
Lebanese American University

PLAGIARISM POLICY COMPLIANCE STATEMENT

I certify that:

1. I have read and understood LAU's Plagiarism Policy.
2. I understand that failure to comply with this Policy can lead to academic and disciplinary actions against me.
3. This work is substantially my own, and to the extent that any part of this work is not my own I have indicated that by acknowledging its sources.

Name: **Marcelino Al Ghafari**

Signature: [redacted]

Date: **09** / **05** / **2022**
Day / Month / Year

ACKNOWLEDGEMENT

I want to take the opportunity here to extend my sincere acknowledgement to my doctor, supervisor, and mentor Dr. Ralph Abi Habib who influenced my journey at LAU since my bacholer's degree. Dr. Ralph. supported and guided me through my academic path to fulfill my requirements to complete my thesis. Thank you, Dr, for your understanding of my situation and allowing me to impose changes on the thesis in the light of COVID-19. I would like also to thank my committee members, Dr. Michella Ghassibe Sabbagh and Dr. Roy Khalaf for taking the time to participate in my thesis evaluation and for being part of my thesis committee.

I am also very grateful to my University (LAU) and my colleagues and friends whom without their encouragement and assistance I wouldn't have done this. Thank you Yara Salem, Carla Abrahamian, Sahar Saab, Dalia El Hadi and Hiba Maalouf.

I dedicate this work to my beloved family, Dad, Mom and Sister.

Characterizing the Extent of Genetic Aberrations in Glioblastoma Multiforme using Bioinformatics Tools

By

Marcelino A. Al Ghafari

ABSTRACT

Glioblastoma multiforme (GBM) are considered the most common type of brain cancer. They are known for their high mortality and morbidity. Advances in genetics and molecular understanding of GBM can explain the failure of many targeted therapies. As a result of this, the World Health Organization (WHO) have updated the classification to provide more accurate diagnosis and treatment. As a result of this WHO classification, GBM emerged to be a group of complex disease with several mutations and different manifestations. Thus, it is important to study the nature of each tumor to apply the right therapy for it. For this reason, bioinformatics is becoming a fundamental pillar in the field of cancer biology. Bioinformatics yields and generates massive data that, through the right tools, can be analyzed, interpreted and then transferred to be applied to cancer therapeutics. Many tools are available online. In this study, cBioPortal for Cancer Genomics (http://cbioportal.org) has been used. This platform is a web resource for studying, viewing, and interpreting cancer data. The data collected from cBioPortal were analyzed using Cytoscape, which generated a network showing the mutual exclusivity of the genes of interest. In addition to this, single nucleotide polymorphisms (SNPs) were

studied using Ensembl and MutationTaster to predict their oncogenicity. All of this provide a general overview of how the study of genetic aberrations in GBM can be accessible through bioinformatics tools. It emphasizes on the importance of such tools coupled with precises analysis to determine the best therapy for cancer. In this paper, genes related to GBM are highlighted and analyzed using cBioPortal, Cytoscape, and MutationTaster. The results highlight the importance for knowing the oncogenic nature of mutations which helps in targeted therapy. In addition to this, a comparison between Lebanese GBM patient samples and cBioPortal samples. This aims to give new understanding of the pathways altered in GBM, methods of therapy, chemoresistance, and autophagy.

Key words: Glioblastoma Multiforme (GBM), cBioPortal, Cytoscape, MutationTaster, SNPs, Bioinformatics.

Table of Contents

List of Tables

List of Figures

Chapter One

Introduction

1.1. Gliomas: Brain Tumors

Gliomas are the most common brain tumors that affect adults. They are characterized by malignancy and being the primary tumors that affect the nervous system. Gliomas can occur in the nervous system specifically in the brain (Ostrom et al., 2014). The histological aspect of these cells is like those of normal glial cells, and their name originates from the similarities between the normal cells and the cancerous cells. However, what remains uncertain is where do glioma cells originate from, whether from the normal glial cells, stems cells, or other type of cells. The standard criterion that was used by pathologists for classifying, diagnosing, and grading gliomas was determined by the 2007 version of the World Health Organization (WHO). The diagnosis and the grading in this classification depend on the malignancy levels that was the "gold standard". The tumors are graded according to WHO grade I to grade IV based on their increasing malignancy as well as the presence of atypia and mitotic activity. In addition to this, some subtypes might demonstrate microvascular proliferation along with pseudopalisading necrosis (Chen et al., 2017).

1.2. Classification and epidemiology

According to WHO grade II and the International Classification of Diseases-Oncology version 3 (ICDO-3) gliomas can be astrocytic, oliogodendrocytic or a combination of these two cell types. Gliomas can be WHO grades I – IV based on malignancy. The most

common types of gliomas in adults histologically are astrocytoma (grades I – IV), oligodendroglioma (grades II – III), and oligoastrocytoma (grades II– III). Grade I tumors include pilocytic astrocytomas, pleomorphic xanthoastrocytomas, and subependymal giant cell astrocytomas, as well as the more common infiltrating gliomas, including grade II oligodendrogliomas and astrocytomas, and grade III anaplastic oligodendrogliomas, anaplastic astrocytomas, anaplastic oligoastrocytomas, anaplastic ependymomas, and grade IV glioblastomas multiforme (GBM) (Wernsch et al., 2002).

The definitions and the taxonomy of gliomas have changed over time, using immunohistochemistry and electron microscopy as tools for lineage determinations. The recent additions of molecular surrogates allowed the use of these tools to provide prognostic, diagnostic, and predictive help in the field of gliomas classifications. Most gliomas in adult patients are diffuse gliomas. These gliomas are characterized by the infiltrative growth that diffuses the central nervous system (CNS) parenchyma leading to the formation of a network of cells throughout the neuropil. The growth of these cells is accompanied with the gathering and the aggregation of cancerous neoplastic cells around the neurons, the membrane of the pia matter, as well as the blood vessels. Diffuse glioma usually invade and span over large areas on myelinated fibers, and frequently cross the corpus callosum to the opposite hemisphere leading to a pattern known as “butterfly glioma”. These gliomas are classified as oligodendroglial, astrocytic, or mix of both, and according to WHO they are grade II, which is low-grade, grade III which is anaplastic, or grade IV which is glioblastoma. Most glioblastoma patients die within 1-2 years after their diagnosis. On the other hand, patients with grade II can survive for over than 10 years, however, in most cases, these tumors will progress to a higher grade of malignancy leading

to death. As mentioned above, most diffuse gliomas as astrocytic, oligodendroglial, or mixed oligodendroglial are astrocytic. The subtypes are assessed based on the similarities between the normal non-neoplastic glial and the tumor cells. Oligodendrogliomas are characterized with having a uniformly rounded nuclei, while astrocytomas have irregular nuclei and hyperchromasia (Louis et al., 2007; Perry et al., 2016).

The last several decades had witnessed an increase in the discovery process that led to the uncovering of many important genetics and molecular foundations of the gliomas. These findings led to changes in the classifications of gliomas, and provided new understanding regarding the tumor initiation, ontogeny, and progression. Many limitations in the histological-based classification were found. One of the main problems in this is "interobserver variability". Interobserver as definition is the agreement between two or more independent observers in the clinical settings. There was a lack of precision in prognosis for patients in the traditional classification. Even when factors such as age, patient's response to treatment, and resection, the differences were accounted. These differences suggested that there are many underlying factors in these tumors that are not considered in the WHO classification. Thus, there was a need to have a detailed understanding of the molecular levels of the gliomas. This understanding was crucial to get a better diagnostic criterion and prognostic biomarkers that lead to developing effective targeted therapies (Chen et al., 2017).

To track the incidence of gliomas, many organizations collect data through cancer surveillance that is done by many governments, or through health system records. Factors such as age, gender, race, and country have a significant influence on the incidence of cancer. Some gliomas such as anaplastic astrocytoma and glioblastoma are more likely to

happen to people between 75 -84 years old. The most conclusive prognostic factors for glioblastoma are the tumor resection, age at diagnosis, and the karnofsky performance status. The survival of the patients is affected by the grade across the different glioma subtypes. From population-based studies, it was estimated that the survival for glioma patients is about 5 years. The survival rates of patients vary depending on the type of gliomas: pilocyctic astrocytoma has the highest survival being grade I, while glioblastoma has the poorest survival. Patients with glioblastoma have a survival rate between 0.05% to 4.7% over 5 years post diagnosis (Ostrom et al., 2014).

Glioblastoma multiforme is the most common and most lethal type of brain cancer, and it will be the topic under study.

Many studies have studied the molecular differences between tumors that share similar features. In 2008, the cancer genome atlas (TCGA) published a classification of GBM based on molecular markers which included Mesenchymal, Classical, Neural and Proneural. Those molecular markers were Isocitrate Dehydrogenase IDH, *MGMT* methylation, 1p/19q co-deletion, and EGFR (Ludwig et al.,2017). Other molecular markers are important to be studied and will be shed the lights on.

In 2015, two groups of researchers published papers that described chromosomal and molecular subtypes of glioma that clarified the classification of these tumors. In the first study, more than 1000 patients with WHO grade II–IV glioma were studied for the presence or absence of three tumor biomarkers. The first of these mutations are mutations in the promoter of *TERT* which encodes telomerase reverse transcriptase. Interestingly, mutations in the *TRET* promoter are observed in the least aggressive glioma, which is grade II oligodendroglioma, and at the same time the most aggressive types; grade IV

astrocytoma. The second was co-deletions of chromosome arms 1p and 19q which are observed in oligodendrogliomas and are linked with the sensitivity to alkylating agents during chemotherapy, as well as mutations in *IDH1* or *IDH2* which encode isocitrate dehydrogenase NADP. The IDH enzyme facilitates the transformation of isocitrate to alphaketoglutarate (KG). It has three isoforms: IDH1 located in the cytoplasm, IDH2 located in the mitochondria, and IDH3 which is part of the tricarboxylic acid cycle. The third group of mutations are the most common in glioma that are associated with distinctive tumor-cell metabolism and survival (Eckel-Passo et al., 2015; Molinro et al., 2019; Bendahou et al., 2020).

The second study had almost 300 patients with WHO grade II and III gliomas. The researchers found three different and distinctive subtypes of glioma based on TP53, IDH mutations, and 1p-19 q co-deletions. All these findings, along with other data from different studies, and WHO classifications led in 2016 to a new classification system for adult glioma. The new classification system is made up of five categories of adult glioma which glioblastoma is one of them.

Mutations in TP53 were found to occur commonly in secondary gliomas, while epidermal growth factor receptor gene EGFR amplifications and PTEN mutations are more frequent in primary gliomas. The improvements in bioinformatics and sequencing after the determination of the human genome sequence have played a major role in allowing the genome sequence analysis in human cancers.

1.3. Glioblastoma Multiforme (GBM)

Glioblastoma Multiforme (GBM) is a class IV neoplasm, and most common tumor of the CNS that is considered malignant. GBM has a vast range of alterations that can have genetics or epigenetics basis. This wide set of alterations lead to the existence of many mutation subgroups (MSG). These MSG require independent treatment and have their own patient survival. Thus, the term multiforme was introduced due to genotypical and phenotypical multiformity of this tumor (Stoyanov, 2017). GBM can be primary or secondary tumor. Primary GBM is the very most common form and occur in older patients, while secondary arises in younger patients. Both subtypes show difference at the molecular levels, but at the end, the same pathways are affected. Primary GBM has amplified, mutated EGFR. This mutation yields in increase in the activity of tyrosine kinase receptor that will activate RAS and PI3K pathways. In addition, alterations in MDM2, PTEN, and IDH1 exist in GBM (Alifieris,2015).

1.4. Biomarkers

1.4.1. Isocitrate dehydrogenase (IDH)

In a study by Yan et al. (2009), 20661 genes in 22 samples from human were analyzed. This study found mutations in IDH*1* and *IDH2* to be playing a role in the malignancy of GBM. Gliomas having *IDH* mutations were distinctive from gliomas with wild-type *IDH* genes. Two subtypes of gliomas of WHO grade II or III such oligodendrogliomas and astrocytomas demonstrated *IDH* mutations but no other genetic mutations that were detectable during early stages of the gliomas' progression. This suggests that *IDH* mutations occur during the early stages of the gliomas development which can give rise

to oligodendrocytes and astrocytes. *IDH1* mutations were identified in 10 out of 10 anaplastic oligoastrocytomas. On the other hand, mutations in IDH1 or IDH2 were not found in pilocytic astrocytomas of WHO grade I which suggests that there is a different mechanism of development for these tumors. This finding is verified by the clinical observation that these tumors rarely undergo transformation to become malignant (Yan et al.,2007).

In another study done on *IDH* mutations, Parson et al. (2008) showed that most of the tumors that were analyzed demonstrated alterations in genes found in the *PI3K, TP53, RB1* pathways. Beside these pathways, this study showed that *IDH* gene is mutated in glioblastoma multiforme (GBM). The mutation in this gene is usually an amino acid substitution at position 132. This residue is conserved and located in the binding site of the isocitrate. The authors speculated that this mutation is an activating one due to the absence of frameshift or stop mutations which are usually inactivating mutations, as well as the absence of mutations in active sites of important residues. This study revealed that *IDH1* mutations can be viewed as a specific marker for secondary GBM. Patients with *IDH1* mutations had a very high frequency of *TP53* mutations, and better prognosis. In addition to this, they suggested that *IDH1* can be used as a target for treatments, because in a study done previously the inhibition of IDH2 –different IDH enzyme- increased the sensitivity of cancer cells to chemotherapy.

Mutations in the active site of *IDH1* at position R132 and IDH2 at position R172 is very common in patients with astrocytomas and oligodendrogliomas. Several studies done by computational approach identified the mutations in IDH1 such as I147S, V444A, and D375Y, as well as in IDH2 such as N439D and R140G. In a recent paper published by

Bendahou et al (2020) different amino acids substitutions were identified in IDH1, IDH2, and p53 proteins. The following substitutions were found: IDH1 R132, IDH2 R173M, and p53 R175H, R158G, and K305N. The IDH mutations that occur in the catalytic site leads to the loss in the activity of the enzyme. All these mutations have been identified in grade II and III glioma, suggesting that they might have a synergistic role in the start and the propagation of the glioma.

1.4.2. Epidermal growth factor receptor (EGFR)

Epidermal growth factor receptor (*EGFR)* is a prototype of the *EGFR* family. It is involved in promoting cell proliferation and opposing apoptosis. It is known to be a proto-oncogene. In normal cells, the expression of *EGFR* is estimated to be around 40,000–100,000 receptors per cell, from these receptors more than 106 receptors are overexpressed in cancer cells. EGF regulates its own receptor by increasing the its expression. This happens by stimulating the expression of ETE which is the EGFR-specific transcription factor (Wee et al., 2017).

EGFR consists of an extracellular ligand binding and dimerization arm, a transmembrane domain, and an intracellular tyrosine kinase domain. The extracellular region is divided into four domains. Domain I along with III participate in ligand binding, while domain II plays a role in the formation of homo or hetero dimers. Domain IV forms disulfide bonds with domain II and links to the transmembrane domain. The transmembrane domain of the *EGFR* plays a role in dimerization, while the tyrosine kinase domain is involved in trans-autophosphorylation in which the N-lobe of one receptor that is mainly made up of β-sheet structure interacts with the C-lobe of another receptor which is mainly a mainly α-helical structure.

The C-terminal tail is made up of many tyrosine residues that can be phosphorylated. When phosphorylated, the receptor will dimerize, then transphosphorylation occurs. *EGFR* activates the extracellular-signal-regulated kinase (ERK), mitogen-activated protein kinase (MAPK), *AKT-PI3K*, and PLC-γ1-PKC pathways, focusing on the molecular interactions between each protein in the pathways. The RAS-RAF-MEK-ERK MAPK pathway might be the most important pathway that is a response of the *EGFR* activation. *RAS* and *RAF* which are considered proto-oncogenes are found in this pathway, along with *MEK* which is a major target for cancer therapy. ERK/MAPK found in this pathway initiate a wide range of responses after interacting with different substrates, these responses include proliferation, growth, migration, inhibition of apoptosis, and differentiation (Wee et al., 2017).

EGFR is found to be overexpressed in around 60% of primary glioblastomas and it is considered as a characteristic of aggressive glioblastoma. Along with overexpression of *EGFR*, another mechanism can cause its faulty induction such as gene amplification, mutation of the receptor, which according to the TCGA dataset occurs in up 57% of glioblastomas, and overexpression of the ligands. While regarding at the subtypes, classical glioblastoma has a rate of 95% of *EGFR* focal amplification, on the other hand, neural, mesenchymal and proneural glioblastomas demonstrate reduced rates of *EGFR* amplification at 67, 29, 17 % respectively. Of all the mutations, the extracellular domain *EGFR* mutations are the most observed in glioblastoma. EGFRΔIII, is the most prevalent *EGFR* mutation. This mutation occurs because of an in-frame deletion of 801 bp in the DNA sequence that encodes the extracellular domain. This deletion leads to the truncation of the receptor; however, this renders it always active. This mutation in the *EGFR* is

exclusive for cancer cells and won't be found in normal cells, making it a therapeutic target. Recent studies have shown that EGFRΔIII is expressed in more than 50% of glioblastomas leading to the amplification the receptor. In addition, TCGA data indicates that this mutation is found in 23% of classical tumors. The constitutive signaling of EGFRΔIII gives the glioblastoma cells the ability to survive *in vivo* through the increase in the production of many mitogenic factors, such as Akt and through the inhibition of apoptosis.

Studies have shown that glioblastoma cells that express EGFRΔIII are usually resilient to both radiations and chemotherapy, and that patients with EGFRΔIII have poor survival prognosis. Besides to EGFRΔIII, analysis of *EGFR* sequence in 151 glioblastoma tumors identified another missense mutation. These mutations were point mutations, and were identified as follows R108K, T263P, A289V, G598V, and L861Q. The presence of these mutations keeps the *EGFR* active through the hyper-phosphorylation of the receptor even without the presence of a ligand (Xu et al.,2017).

After the transphosphorylation of the receptor, the activated *EGFR* binds to the growth factor receptor binding protein 2 (GRB2) by its SH2 domain. In addition, the activated *EGFR* brings Src homology and collagen (SHC). RAS is a protein that is found upstream in the MEK-ERK pathway. RAF-1 is a complex protein but the phosphorylation of its on Tyr341 and Ser338 residues makes them a binding site for MEK1/2. MEK1/2 activates ERK1/2 which in turn phosphorylates different substrates (Wee et al., 2017).

The activated *RAS* activates RAF-1 along with PI3K and *RAS*-like guanine nucleotide-dissociation stimulator (Ra1GDS). These effectors play very important rules in major pathways thus indicating why *RAS* mutations are very common in cancer. Around 30% of all human cancer carry mutations in the *RAS* genes. These isoforms share a common sequence in the region where they interact with GDP/GTP..*RAS* mutations depend on the primary tissue in human cancer and this cause difference between the isoforms. For examples K-RAS is the most frequently mutated isoform and it is found in lung, and colorectal cancers. The second highly mutated isoform is N-RAS that is found in melanoma. The last one with the lowest frequency observed is H-RAS and it can be detected in bladder cancer. *RAS* mutations are found to be very rare in GBM, however a high *RAS* activity is detected in the tumors. In addition, the RAS-GAP *NF1* is mutated in about 18% GBM patients. Tumors having NF1 mutations show activation of *RAS*. (Kano et al.,2016; Wee et al., 2017).

RAF is a protein kinase that is encoded by the *RAF* gene. It is made up of 648 amino acids, and it has a serine-threonine kinase activity after it get activated by *ras*. This kinase can be activated by different ways such as: localization of Raf on the inside of the cell membrane after binding to Ras, phosphorylation and dephosphorylation, and dimerization of Raf protein (Guo et al., 2019).

The activated MAPK leads the activation of MAPK kinase kinase (MAPKKK) and a MAPK kinase (MAPKK). The ERK family is subdivide into two groups: the ERK1 and ERK2 that are the classic ERKs and ERK5 that is the larger ERKs. The classic ERKs have a kinase domain while the large ERKs have a carboxy-terminal domain in addition to their kinase domain. ERK1/2 are mainly activated by growth factors and mitogens that lead to

cell growth and differentiation. Many important upstream regulators are found in this pathway such as cell surface receptors, for example tyrosine kinases (RTKs), as well as G-protein-coupled receptors (GPCRs), in addition to small GTPases *RAS* and *RAF*. The MAPKKs in this pathway are MEK1 and MEK2, and the MAPKKKs are Mos, Tpl2, and Raf family. In addition to its involvement in regulating cell proliferation, differentiation, cell cycle, tissue formation, and apoptosis, the ERK/MAPK pathway also plays a role in tumorigenesis. In glioblastoma, *EGFR* and *RAS* were found to be overly expressed in multiple cases (Guo et al., 2019).

1.4.3. The PI3K/AKT pathway

The phosphatidylinositol 3 kinases (PI3Ks) are a family of kinases that phosphorylate lipids. These kinases get signals from cytokines, growth factor, and environmental cues then integrate them into signals that affect and regulate many other signaling pathways. The activated pathway feed into cell proliferation and metabolism (Thrope et al., 2015). Based on their different structures and specificity to substrates they are divided into three different classes.

The PI3K-AKT pathway is very conserved, and it is activated and regulated through a multistep process. When *EGFR* is activated, it associates with p85 through the dimerization with human epidermal growth factor receptor 3 (HER3), or through binding to the docking protein GRB2-associated binder 1 (GAB1), both will relieve the inhibition caused by p85. When this happens, p110 phosphorylates phosphotidylinositol 4,5-Bisphosphate (PIP2) to phosphatidylinositol (3,4,5)-trisphosphate (PIP3), this leads to the binding of *AKT* on the membrane through PIP3. Then, phosphoinositide-dependent kinase 1 (PDK1) activates *AKT* partially by phosphorylating the T308 site of *AKT*. The substrates

of mTORC1 are translation initiation factor 4E binding protein (4EBP1) that plays a role in eukaryotic translation, in addition the ribosomal protein S6 kinase (S6K1). S6k1 phosphorylates the ribosomal protein (S6/RPS6) and this promotes protein synthesis and cellular proliferation. The full activation of *AKT* happens through phosphorylating it at S473 in the carboxy-terminal hydrophobic motif by the mammalian target of rapamycin complex 2 (mTORC2). This full activation of *AKT* leads to the activation of different substrates through phosphorylation. These substrates can carry plenty of cellular functions such as angiogenesis, metabolism, growth, proliferation, survival, protein synthesis, transcription, and apoptosis. This pathway is downregulated by the phosphatase and tensin homolog (*PTEN*) which is a tumor suppressor that is mutated in glioblastoma. *PTEN* dephosphorylates PIP3 to PIP2, thus negatively regulating this pathway (An et al., 2018; Hemmings et al., 2012). High levels of phosphorylated *AKT* have been correlated with poor prognosis for GBM patients. Even though a dominated mutation of the genes coding for *AKT* hasn't been found yet in GBM and in human tumors, thus the activation of *AKT* in cancer models is suggested to be the result of alterations of its molecules upstream. For example, *EGFR* amplification that is very common in glioblastoma leads to the activation of PI3K/AKT pathway. As mentioned as well, PDK1 and *mTOR* which are positive regulators of *AKT* are upregulated in GBM. It is found in GBM that the main negative regulator of *AKT* and *PTEN* is inactive because of a mutation in the gene or a methylation. The most recent studies suggest that an *AKT* phosphorylation decrease through *PTEN* can be obtained through the suppression of two miRNAs which are miR-92b or miR-494-3p. When these two miRNAs decrease their expression, the expression of *PTEN* increases leading to the decrease in phosphorylation levels of *AKT*. The expression of both miRNAs

is significantly higher in GBM tissues compared to normal brain tissues (Majewska et al., 2017).

Mutations that activate PI3K are found in most of glioblastoma patients. As mentioned before, the *EGFR* amplification is very common in patients with glioblastoma, this amplification leads to the constant activation of PI3K pathway. The EGFRVIII mutation leads to the constitutive signaling that affects the PI3K pathway. In addition to this, loss of *PTEN* occurs in more than 50% of glioblastomas and drive the continuous activation of PI3K pathway. *NF1* loss can contribute to the activation of *PI3K* in glioblastoma patients either through genetic inactivation or enhanced proteasomal degradation. In mouse models, it was shown that the continuous activation of *PI3K* pathway is needed for the glioblastoma formation and its malignancy along its progression. In these models the PI3K altered pathway interacts with other commonly altered components of different pathways such as RAS/MEK/ERK, pRB, and p53 (Akhavan et al.,2010).

mTOR is a 289 kDa threonine and serine protein kinase that is found into two forms and protein complexes that are distinct and they are known as mTORC1 and mTORC2. mTORC1 is made up of a regulatory associated protein of mTOR (RAPTOR), mammalian lethal with Sec-13 protein 8 (mLST8), proline-rich AKT substrate 40 kDa (PRAS40), and DEP domain TOR-binding protein (DEPTOR). mTOR is inhibited by rapamycin that is produced by *streptomyces hygroscopicus.* Rapamycin inhibits mTORC1 through binding to the intracellular FK506-binding protein (FKBP12). This protein interacts directly with mTORC1 but not with mTORC2. Different cues activate mTORC1 such as stress, oxygen, growth factors, amino acid concentration, and energy status (Mecca et al.,2018).

When mTORC1 is activated, it activates cell growth by the phosphorylation of 4EBP1 and S6K. The phosphorylation drive protein synthesis by triggering S6K to phosphorylates the 40S ribosomal subunit, this will initiate mRNA translation, and phosphorylate 4EBP1 on serine 422. When phosphorylated, the 4EBP1 will bind to eukaryotic translation initiation 3 (EIF3) and form the complex EIF4F. In addition, mTORC1 activates the transcription intermediary factor 1-alpha (TIF-1A) which induces the rRNA to be transcribed by RNA polymerase and inhibits MAF1 which is a polymerase III repressor, therefore enabling 5sRNA and tRNA transcription.

In addition to its role in protein synthesis, mTORC1 plays an important role in lipid and nucleotide synthesis. In glioblastoma the rapid turnover of lipid and nucleotide is a hallmark of the cancer. mTORC1 phosphorylates LIPIN-1, thus preventing it to enter the nucleus, this will cause the suppression of the transcription factor involved in cholesterol and fatty acid synthesis. This transcription factor is called the sterol regulatory element-binding protein 1/2 (SREBP1/ 2). On the other hand, mTORC1 is involved in purine synthesis through the triggering of the activating transcription factor 4 (ATF4).

On the other hand, mTORC2 is made up of the rapamycin-insensitive companion of mTOR (RICTOR), DEPTOR, mLST8, protein observed with RICTOR (PROTOR), and stress-activated map kinase-interacting protein 1 (mSIN1). mTORC2 upstream activators aren't very well known, and it is activated by growth factors, but it doesn't show an activation to nutrients. Once it is activated, mTORC2 leads to cell proliferation, survival, and motility. mTORC2 modulates all these processes through activation of many AGC protein kinases. Gulati el al. treated glioblastoma cell line with rapamycin, and noticed a time dependent decrease of S6K phosphorylation, in addition to an unexpected increase

in AKT on serine 473 which is known to play a role in cell proliferation. mTORC2 also plays a role in induction of Warburg effect. Warburg effect is a metabolic process in which tumor cells metabolize glucose through glycolysis in the presence of oxygen levels to supply the need of the growing cells. In addition to this, it has been recently shown that mTORC2 grant resistance to glioblastoma cells towards alkylating agent "cisplatin" (Mecca et al.,2018).

As mentioned above class I PI3K plays a major role in cancer. It was first demonstrated in the late 1990s that the fusion between p110α and the viral sequences or the SRC myristoylation sequence was activating to tumor progression and highly oncogenic. In 2004, scientists discovered that PIK3CA mutations are very frequent and common in human cancers. Since then, it was established that PIK3CA mutations are causative in many types of cancers. Most of the mutations that occur in PIK3CA are missense mutations. Analysis of cancer cells confirmed that these mutations confer the transformation of normal cells into oncogenic through the activation of p110α. In many studies on genetically engineered mouse models (GEMMs), the role of PIK3CA has been demonstrated in tumor initiation and progression. Mutations that occur in the helical domain lead to the reduction in the inhibition of p110α caused by p85 (Thorpe et al.,2015).

PIK3CA mutations were found in 27% of glioblastoma tumor, and it has been reported that the deregulation of this isoform is frequently detected in glioblastoma. Most of the mutations of PIK3CA show a PI3K gain of function which lead to the recruitment of p110α to membrane phospholipids. This recruitment activates p110α and Akt. The knockdown of PIK3CA in GBM cells lead to the inhibition of cell viability, invasion, and migration through the decrease in the activation of FAK and *AKT*. In addition to this, a

suppression in the cell survival, growth, and migration was observed in GBM cells after being treated with p110α isoform-selective inhibitors A66 or PIK-75 in vitro (Zhao et al.,2017).

In a study conducted by Samojedny et al. (2014), AKT3 and PIK3CA genes were silenced using siRNA. The expression levels of both mRNAs were downregulated significantly in the transfected cells, this was proven through western blot and flow cytometry. It is known that AKT plays a role in promoting the transition of cells from G1 phase to S phase in the cell cycle. In addition, it was shown that the knockdown of these two genes initiated a decrease in the proliferation of GBM. Many pathways are involved in the cell proliferation such as kinase activation, and signal transduction. It was confirmed by results from Samojedny et al. (2014) and Weber et al. (2014) that PI3K/AKT pathway is very crucial and important in cellular proliferation.

When AKT3 and PI3KCA were silenced by siRNAs, this led to the reduction of the viability of GBM cells. This reduction was shown with the increase in the percentage of GBM cells that were in the subG1 phase. The subG1 phase is characterized by apoptotic cells, and it was suggested that the AKT3 and PIK3CA siRNAs induce an apoptotic pathway in T98G cell line. AKT stops apoptosis through the activation of BCL-2, IκB kinase (IKK) and HDM2 which are antiapoptotic proteins. Inhibition of apoptosis depends on the intracellular balance between BAX and BCL-2 along with other proteins from the BCL-2 family such as BAD, BAG1, BCL-XL, and BCL-XS. When the ratio of BAX/BCL-2 increases, cytochrome c is released from the mitochondria into the cytosol. In addition to this, the knockdown of AKT3 and PI3KCA gene was related to higher apoptotic index. The findings from this study were consistent with the results from other

studies. Opel et al. showed that inhibiting the PI3K can be an efficient way to make glioblastoma sensitive to the induction of apoptosis. In another study, Koseoglu et al. (2016) studied the effect of AKT1, AKT2, and AKT3 knockdown on the induction of apoptosis in twenty different types of human tumor cell lines. It was found that these knockdowns lead to apoptosis. Mure et al. noticed that an induction of apoptosis happened in GBM cells were exposed to AKT3 knockdown. This apoptosis was meditated through a pathway in mitochondria where BAD was dephosphorylated and caspases 3 and 9 were activated (Samojedny et al., 2014).

Other studies have related the mitochondrial dysfunction to cancer inhibition. When the mitochondria are damaged, the levels of ATP will be depleted, thus there will be a depression in the metabolism and the growth of the cancer cells. In addition, malfunctions in the mitochondria overloads the cytoplasm with calcium leading to impairment in cancer metastasis and migration. According to all the pervious information about the mitochondria, many arguments have been established regarding the decisive role that mitochondria play in determining the cancer cell fate. Recent studies have found that division of the mitochondria is an upstream mediator of the dysfunction in the mitochondria. For example, in many cancers such as gastric, liver, lung, breast, tongue cancer, cardiac reperfusion injury, and endothelial oxidative stress model, the activation of mitochondria division leads to mitochondrial dysfunction. This dysfunction was manifested by ATP depression, mitochondrial respiratory dysfunction, calcium overload, and apoptosis activation. Furthermore, some recent studies have investigated the mechanism and the role of the mitochondrial division in the progression and viability of glioblastoma cells. It was found in several studies that PTEN might be a promising and a

potential target to control the progression of glioblastoma. Earlier reports demonstrated that mutations in the PTEN gene are significantly correlated with bad prognosis in glioblastoma patients. PTEN has an important role at the molecular level of glioblastoma. It regulated the oncogenesis through the chromatin-associated histone H3.3. Also, PTEN affects the metabolic reprogramming of glioblastoma, and it is associated to the increase of glioblastoma sensitivity to radiation thus playing a role in cancer cell death. Moreover, it controls the mitochondrial energy metabolism and signaling in human glioblastoma cells. All these suggest a very tight relationship between PTEN and homeostasis in the mitochondria. On the other hand, there were no studies conducted in regards of whether PTEN is responsible for the activation of mitochondrial division in glioblastoma. However, because Akt pathway is related to the energy status in cancer, and it was found that the same pathway modulates the mitochondrial division in many disease models, it might be suggested that PTEN activates mitochondrial division and thus induced cancer cell death in a manner dependent on the Akt pathway.

Bao et al. (2019) showed that PTEN affected the mitochondrial division thus it played a role in modulation the viability of glioblastoma. At a molecular level, when PTEN is overexpressed, mitochondrial division is activated. This process is moderated by Drp1 upregulation, when Drp1 is knocked down the effects of PTEN vanish.

1.4.4. The P53 pathway

P53 is a transcription factor that is made up of different domains. A nucleotide-binding domain that interacts with the consensus DNA sequence for p53 and the other domains that interacts with other regulatory pathways and thus regulate transcription. When the conditions are normal in a cell, the p5er activity is low and it is controlled by MDM2 and

MDM4 (MDMX) through ubiquitination and degradation. While under stress conditions the interaction between p53 and MDM2 is disrupted. This disruption leads to p53 induction. The main role of p53 is maintaining homeostasis in the cell, and it is usually downregulated in cancer. It is at the center of many cellular networks that controls cell survival, proliferation as well the genome integrity. Summing all these up P53 is called the "Guardian of the Genome". P53 is important in the regulation of cellular processes. According to The Cancer Genome Atlas (TCGA), GBM project shows that the p53 pathway (including CDKN2A, MDM2 and TP53) was deregulated in ~85% of tumors. When further analysis was held, TP53 was found to be mutated in around 28 % of lower grade gliomas exhibiting more than 90% TP53 mutations. Most of the mutations is a missense occurring between exons five and eight. The mutant TP53 gene product leads to the upregulation of p53 leading to either loss or gain of p53 function. This mutation halts the correction of DNA errors thus leading to gliomagenesis. Nowadays scientists are using the p53 immunohistochemistry to assess the presence of mutant TP53 in gliomas. P53 is activated when there is DNA damage, genotoxicity, activation of an oncogene, aberrant growth signals, and hypoxia. All these events that can be encountered or faced during carcinogenesis. Inactivation of p53 in glioblastoma cells is associated with more proliferative, less apoptotic, more invasive and stem-like phenotype. Glioblastoma cells that exhibit this inactivation were found to be more resistant to cisplatin which is a DNA-damaging drug that is used in cancer therapy. Even though p534 inactivation is found in high mutation it was correlated to survival in glioblastoma patients (Zhang et al.,2018).

P53 was found to be a tumor suppressor and a transcription factor that is mutated in about 50% of cancers where its mutations led to constitutive expression. The inactivation of p53

can also happen when MDM2 gene undergo multiplication. P53 mutations have been categorized into three major patterns of disruption: loss-of-function, gain-of-function, and dominant-negative function. However, such patterns and their individual effects in GBM have not been fully defined. Mutation in p53 can lead to loss of function which can lead to the loss in the growth inhibitory effects that are usually demonstrated by wild type p53. With this mutation, cell cycle and apoptosis genes are lost and thus cells undergo uncontrolled proliferation. In addition, mutant p53 can cause gain of function effects which include the expression of distinct subsets of genes. Mutant p53 demonstrate as well dominant negative function where it can tetramerize with wt-p53 and lead to the downregulation of the anti-tumorigenic activity of the wt –p53 (England et al., 2013).

In the p53 pathway, MDM2 and MDM4 perform the role of oncogenic inhibitors of p53, by this they inhibit the suppressive activity of p53. MDM2 act as a negative regulator by targeting p53 for degradation. MDM2 transcription at the same time is activated by p53, thus creating a negative feedback loop that regulates activity of p53 and the expression of MDM2 (England et al., 2013). When MDM4 and MDM2 are amplified they can activate p53. This activation led to many tumor suppressing functions such as apoptosis, growth arrest, DNA repair and senescence. In glioblastoma upregulation of p53 expression has been demonstrated due to mutations in p53 conversed DNA-binding domain which prevents the regulation of p53. Wild type p53 is usually kept at levels by polyubiquitylated proteosomal degradation as well as ubiquitin-independent proteosomal degradation. However, mutant p53 is more prone to degradation, which is mediated by MDM2, but this fails to activate MDM2. When MDM2 is not activated, p53 can escape degradation and thus lead to all its effects after being mutated. In addition to this, mutation of *PTEN* which

is very common in glioblastoma, cause tumor formation by preventing MDM2 from degrading p53 (England et al., 2013; Zhang et al.,2018).

1.5. Bioinformatics

Studying all these genes is essential to get a well-founded knowledge of glioblastoma multiforme and get a bigger image of the genomics behind this cancer. Cancer treatment is a complex process, in which many challenges are faced. Unfortunately, there is no unique solution for all cases. Even though surgery, radiation, and chemotherapy are the primary treatments against cancer, the advances in the molecular mechanisms underlying tumor growth are generating new targeted approaches to the diagnosis and treatment of the disease. Thus, there is a need for bioinformatics tools. Bioinformatics is the field that deals with biology, mathematics, medicine, statistics, computer science, and information technology. This new field led to the introduction of many tools and databases that cause groundbreaking changes in the cancer field. In the past decade, an increase in the biological data was witnessed related to genes, proteins, CpG islands and many other entities. This increase led to the increase in the number of bioinformatics tools. Data can be analyzed used different tools, and as soon as they were developed, new data sources emerged such as TCGA, TARGET, CGCI, etc (Canzoneri et al.,2019). In this paper, TCCA and cBioPortal along with Ensembl and mutation taster will be used to study the pathways involved in glioblastoma, through the analysis of different genes that can lead to new or modified cancer therapeutics. These tools provide a basic but an inevitable understanding of genes interaction, mutations, and the nature of the mutation.

Chapter Two

Material and Methods

2.1. Computer

A computer with an internet browser and with Javascript enabled. All the following browsers are supported: Google Chrome, Firefox 3.0 and above, Safari, and Internet Explorer 9.0 and above. Adobe Flash player Note is required in order to view the networks on the analysis tab. Java Runtime Environment Note application is needed for launching the Integrative Genomics Viewer (IGV). It can be downloaded from http://www.java.com/getjava/ (Gao et al., 2014).

2.2. cBioPortal

2.2.1. Introduction

cBioPortal has more than 200 studies, with all the TCGA projects and studies from different literatures. This portal allows the study and the integration of many gene databases. By looking and mapping the mutations, copy number, and interpreting the gene expression; cBioportal supports the study of different types of gene mutations in tumors as well as other functions. Considering the complexity of glioblastoma along with all the genes involved in tumorigenesis, a tool should be used to study and compare all the factors in this cancer type.. The cBioPortal for Cancer Genomics (http://cbioportal.org) was designed to allow scientists to access the complex data with ease and facilitate translating of genomic data into biological perceptions, clinical trials, and therapies (Gao et al., 2014).

One of the most important features of cBioPortal is the ease of use. All the features on the portal are available through a streamlined 4-step web interface. Users can be guided to choose:

1) A cancer study of interest which can be for example TCGA Glioblastoma Multiforme (GBM)

2) Genomic profiles such as mutations and copy number alterations.

3) A patient case that can be all TCGA patients with glioblastoma mutation

4) A gene set of interest, users can write the HUGO gene symbols gene aliases, or Entrez Gene IDs and can enter arbitrary gene sets or pathways of interest.

In addition to this, users could calculate the mutual exclusivity and co-occurrence between all genes. The tool offers the option of performing cross-cancer queries (Cermani et al., 2012).

The portal allows the studying of multidimensional cancer genomics data through the analysis across genes, samples, and data types. Users can compare gene alteration frequencies and study the genomic alteration in a tumor sample. In addition to this, users can explore all the biological pathways, survival analysis, and analysis of mutual exclusivity. Users can download the data, and publication-quality summary visualization. cBioportal facilitates the access to cancer genomic data. This portal provides a simple way to integrate data, interpret graphs, and visualize results all which help scientists and researchers to use the cancer genomic sets for developing new insights in cancer biology (Cermani et al., 2012; Gao et al., 2014).

Users can visualize and study the alterations and mutations on the genomic level of a certain set of genes. In order to perform this, there are four steps to do it. Users can choose from more than 25 cancer studies and types such as CNS/Brain, kidney, breast etc... When the cancer type is selected, all the genomes, alterations, pathways, and mutations specific for this cancer are selected by default. In this paper CNS/Brain was selected. Once CNS/Brain was selected, twenty matching studies were shown. These studies are divided according to the glioma type, then glioblastoma is selected. Glioblastoma Multiforme (TCGA, PanCancer Atlas, 592 samples) was chosen to be under study. Once the samples are selected, the HUGO genes symbols were entered manually for analysis by pressing Query by Gene option. The portal as well gives the option of selecting genes from a list of genes by pressing Explore Selected studies. The selected genes were: IDH1, EGFR, KRAS, MAP3K1, PI3KA, AKT1, mTOR, PTEN, TP53, MDM2, MDM4. Based on the query, cBioPortal classifies the analysis and visualizations into separate tabs. Each result tab shows different set of information whether related to mutations, pathways, and alterations. The first result tab is the OncoPrint (Gao et al., 2014).

2.2.2. Oncoprint

An Oncoprint (Figure 1) is a graphical representation and summary of all the genomics alterations in the sample across the query genes. The rows represent the genes, and the columns represent the samples. Color coding and glyphs are used to show mutations, amplifications, deletions, and changes in the gene expressions. Additional information is available by moving the mouse over the column which gives you the patient ID or cell line and allows to visualize the clinical attributes. For better view the data, heatmaps can be added as well to the Oncoprint. To export the Oncoprint, choose to download the

Oncoprint as an XML file in scalable vector graphic (SVG) format by pressing the SVG button (Gao et al., 2014; Wu et al., 2019). Graphical representation of the Oncoprint was done on Prism 9.

2.2.3. Mutual Exclusivity

cBioPortal can identify patterns between genes in order to tell whether they are mutual exclusive or co-occur together. The portal uses Fisher's exact test to analyze these relationships (Gao et al., 2014; Wu et al., 2019).

2.2.4. Mutations

This tab provides a graphical summary of all the mutations that are non-synonymous and are identified in the query gene. In the summary, all the mutations' positions are shown along with the sample ID, protein change, annotations, type (missense, nonsense, splice site, frameshift insertion or deletion, in-frame insertion or deletion, nonstop, nonstart), copy number, number of mutations in sample, number of mutations at this position in COSMIC (Catalogue of Somatic Mutations in Cancer), and frequency. In addition to this, a link to the 3D structure with the mutations are highlighted. All the DNA canonical mutations are standardized to the RefSeq isoform (Gao et al.,2014).

2.2.5. Comparison/ Survival

Overall survival and disease-free survival differences between the alerted group and unaltered group could be studied. In addition to this, plots to compare between tumor samples that have one mutation in a query gene and samples that don't have this mutation are available. The results are displayed in a Kaplan Meier plots (Gao et al.,2014).

2.3. Cytoscape

Cytoscape allows biologist to explore and study biological networks, such as pathways and genes regulations, based on experimental data. For example, data about gene expression and mutual exclusivity score were obtained and uploaded to Cystoscape which produced a network with nodes. This network can be imported. Cytoscape could be downloaded following the protocols in doi:10.1002/0471250953.bi0813s47.

When a file is uploaded to Cytoscape, it has no layout so by default a layout will be done by Cytoscape.

1) Go to menu 'Layout → Apply Preferred Layout'

2) Click on the Styles tab to view the styles.

3) Click on the 'Edge' tab on the bottom to reveal edge attribute mappings.

4) Scroll down to reveal 'Stroke Color'

5) In the Column field, click and select 'interaction'.

Cytoscape can hide labels, node graphics and other information to improve visualization speed. You can see these through 'View → Show Graphics Details' (Su et al., 2015)

2.4. MutationTaster

MutationTaster evaluates the pathogenicity of DNA sequence alterations. It predicts the consequences of amino acid substitutions on the function of the protein. In addition, it shows the effects of intronic and synonymous alterations, and deletion. A single analysis now takes less than 0.10 seconds on average. MutationTaster automatically maps the

specific mutation to all the genes and the transcripts, it also puts all the details about this mutation such as chromosomal position and predicted effect (Schwarz et al., 2014).

Chapter Three

Results

3.1. Oncoprint

Samples from Glioblastoma Multiforme (TCGA, PanCancer Atlas):

Oncoprint of the query genes in a sample of 592 glioblastoma multiforme shows the types of mutations taking place in each gene along their frequencies. The Oncoprint below is showing the genes of interest along with percentage of alterations. Each dash is a patient and showing the type of alteration that this patient is harboring.Data from the Oncoprint can be downloaded in different versions, and it is occasionally being updated to provide the most up-to-date data possible.

Figure 1. Oncoprint of TP53, PTEN, EGFR, PIK3CA, MAPK3K1, IDH1, mTOR, KRAS, AKT1, MDM2, MDM4 in 592 samples of Glioblastoma Multiforme showing the alteration type as well as the heat map.

RNA seq V2 RSEM: RSEM (RNA-seq by Experimentation Maximization) determines the total RNA transcript. Next next-generation sequencing can be downloaded.

Cerami et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery. May 2012 2; 401. Gao et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. 6, pl1 (2013).

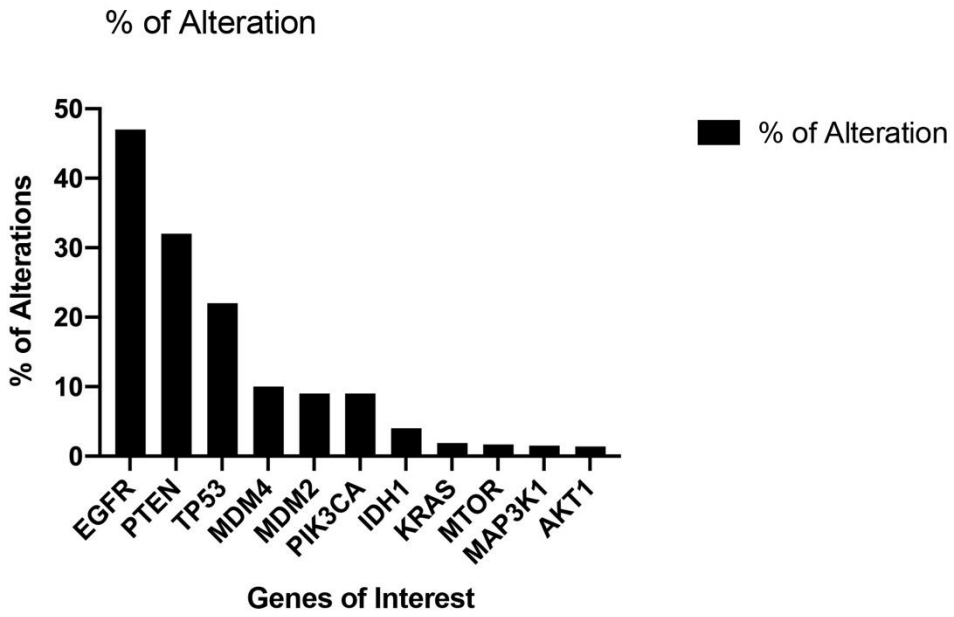


Figure 2. Percentage of alteration in each of the gene of interest according to the Oncoprint.

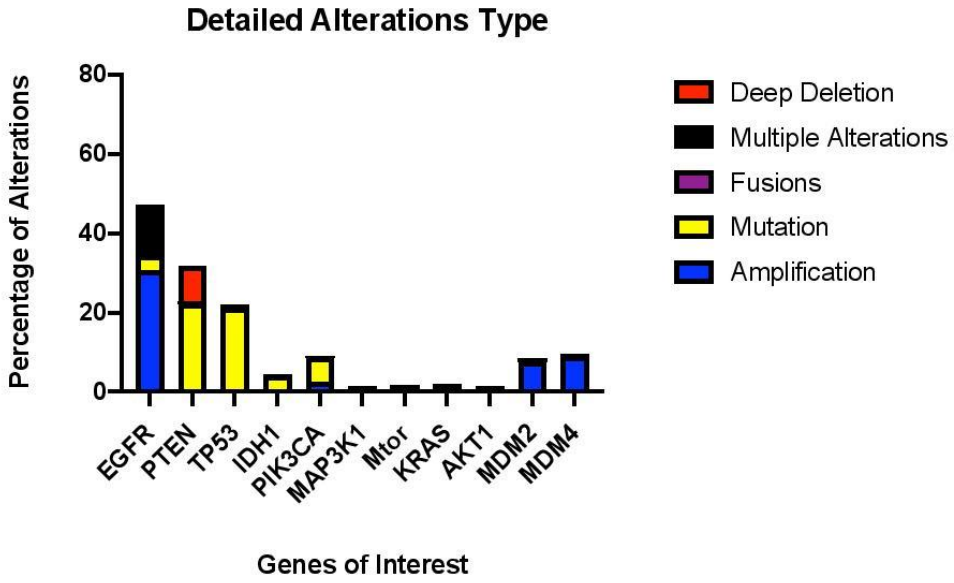


Figure 3. Detailed alterations type for each of the genes of interests according to the Oncoprint.

3.2. Mutual Exclusivity

The query genes were analyzed for mutual exclusivity and co-occurrence.

A	B	Neither	A Not B	B Not A	Both	Log2 Odds Ratio	p-Value	q-Value	Tendency
TP53	IDH1	252	103	1	22	>3	<0.001	<0.001	Co-occurrence
TP53	EGFR	91	87	162	38	-2.027	<0.001	<0.001	Mutual exclusivity
EGFR	IDH1	157	198	21	2	<-3	<0.001	<0.001	Mutual exclusivity
PTEN	IDH1	188	167	22	1	<-3	<0.001	<0.001	Mutual exclusivity
TP53	MDM2	223	122	30	3	-2.452	<0.001	0.010	Mutual exclusivity
PTEN	PIK3CA	177	156	33	12	-1.277	0.007	0.068	Mutual exclusivity
KRAS	MDM2	340	5	30	3	2.766	0.025	0.174	Co-occurrence
EGFR	MAP3K1	177	192	1	8	2.883	0.028	0.174	Co-occurrence
PTEN	AKT1	204	168	6	0	<-3	0.028	0.174	Mutual exclusivity
MAP3K1	MDM2	339	6	30	3	2.498	0.035	0.195	Co-occurrence
TP53	MAP3K1	250	119	3	6	2.071	0.039	0.195	Co-occurrence
PTEN	EGFR	91	87	119	81	-0.490	0.063	0.287	Mutual exclusivity
KRAS	IDH1	349	6	21	2	2.470	0.079	0.319	Co-occurrence
TP53	MDM4	219	115	34	10	-0.836	0.081	0.319	Mutual exclusivity
PTEN	MDM4	190	144	20	24	0.663	0.102	0.370	Co-occurrence
TP53	PIK3CA	227	106	26	19	0.646	0.112	0.370	Co-occurrence
MDM2	IDH1	322	33	23	0	<-3	0.114	0.370	Mutual exclusivity
PIK3CA	MDM4	297	37	36	8	0.835	0.133	0.397	Co-occurrence
EGFR	AKT1	177	195	1	5	2.182	0.137	0.397	Co-occurrence
PTEN	MAP3K1	207	162	3	6	1.354	0.154	0.410	Co-occurrence
MAP3K1	MTOR	363	8	6	1	2.919	0.156	0.410	Co-occurrence
PIK3CA	MTOR	328	43	5	2	1.609	0.198	0.494	Co-occurrence
PIK3CA	IDH1	311	44	22	1	-1.638	0.213	0.510	Mutual exclusivity
MDM4	IDH1	312	43	22	1	-1.600	0.225	0.515	Mutual exclusivity
KRAS	MDM4	328	6	42	2	1.380	0.236	0.519	Co-occurrence
TP53	PTEN	144	66	109	59	0.240	0.258	0.532	Co-occurrence
TP53	MTOR	247	124	6	1	-1.591	0.266	0.532	Mutual exclusivity
EGFR	MTOR	176	195	2	5	1.174	0.275	0.532	Co-occurrence
MAP3K1	MDM4	327	7	42	2	1.153	0.282	0.532	Co-occurrence
PIK3CA	MAP3K1	326	43	7	2	1.115	0.292	0.532	Co-occurrence
EGFR	KRAS	173	197	5	3	-0.924	0.300	0.532	Mutual exclusivity
TP53	AKT1	250	122	3	3	1.035	0.313	0.539	Co-occurrence
EGFR	PIK3CA	155	178	23	22	-0.264	0.338	0.563	Mutual exclusivity
EGFR	MDM4	159	175	19	25	0.258	0.349	0.564	Co-occurrence
PIK3CA	KRAS	325	45	8	0	<-3	0.359	0.564	Mutual exclusivity
PIK3CA	MDM2	303	42	30	3	-0.471	0.428	0.619	Mutual exclusivity
MAP3K1	IDH1	347	8	22	1	0.979	0.435	0.619	Co-occurrence
MDM2	MDM4	304	30	41	3	-0.432	0.447	0.619	Mutual exclusivity
PIK3CA	AKT1	327	45	6	0	<-3	0.465	0.619	Mutual exclusivity
AKT1	MDM4	328	6	44	0	<-3	0.473	0.619	Mutual exclusivity
MTOR	MDM2	339	6	32	1	0.820	0.475	0.619	Co-occurrence
PTEN	MDM2	191	154	19	14	-0.130	0.478	0.619	Mutual exclusivity
PTEN	KRAS	205	165	5	3	-0.424	0.490	0.619	Mutual exclusivity
EGFR	MDM2	163	182	15	18	0.104	0.495	0.619	Co-occurrence
TP53	KRAS	248	122	5	3	0.286	0.526	0.643	Co-occurrence
AKT1	MDM2	339	6	33	0	<-3	0.576	0.682	Mutual exclusivity
MTOR	MDM4	328	6	43	1	0.346	0.583	0.682	Co-occurrence
PTEN	MTOR	206	165	4	3	-0.095	0.621	0.712	Mutual exclusivity
MTOR	IDH1	348	7	23	0	<-3	0.642	0.721	Mutual exclusivity
AKT1	IDH1	349	6	23	0	<-3	0.684	0.753	Mutual exclusivity
MAP3K1	KRAS	361	9	8	0	<-3	0.823	0.888	Mutual exclusivity
MTOR	KRAS	363	7	8	0	<-3	0.860	0.893	Mutual exclusivity
MAP3K1	AKT1	363	9	6	0	<-3	0.865	0.893	Mutual exclusivity
KRAS	AKT1	364	8	6	0	<-3	0.879	0.893	Mutual exclusivity
MTOR	AKT1	365	7	6	0	<-3	0.893	0.893	Mutual exclusivity

Table 1. Mutual exclusivity of TP53, PTEN, EGFR, PIK3CA, MAPK3K1, IDH1, mTOR, KRAS, AKT1, MDM2, MDM4 in 592 samples of Glioblastoma Multiforme.

Neither: Number of samples with alterations in neither A nor B

A not B: Number of samples with alterations in A but not in B

B not A: Number of samples with alterations in B but not in A

Both: Number of samples with alterations in both A and B

Log2 odds ratio: Quantifies how strongly the presence or absence of alterations in A are associated with the presence or absence of alterations in B in the selected samples. OR = (Neither * Both) / (A Not B * B Not A)

p-value: Derived from one-sided Fisher Exact Test

q-value: Derived from Benjamini-Hochberg FDR correction procedure

Tendency: log2 ratio > 0: Tendency towards co-occurrence
Log2 ratio <= 0 : Tendency towards mutual exclusivity
q-Value < 0.05 : Significant association

Cerami et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery. May 2012 2; 401. PubMed. Gao et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. 6, pl1 (2013). PubMed

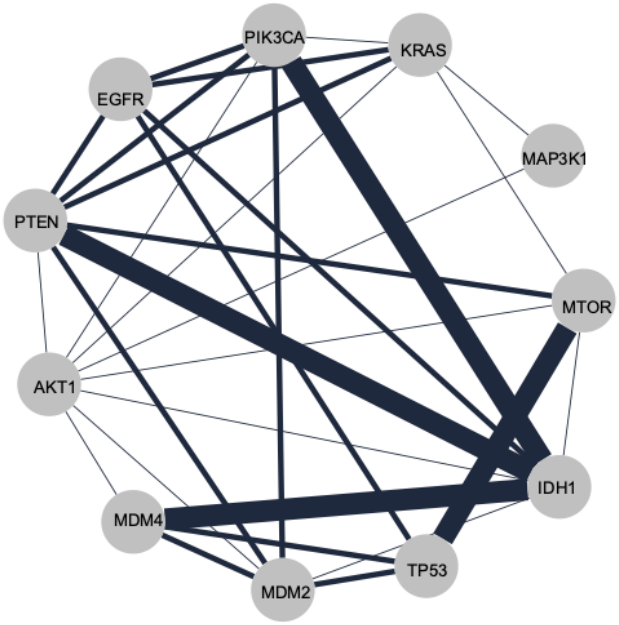


Figure 4 (a). Network of the genes of interest showing their mutual exclusivity relationship. The network is generated by cytoscape.

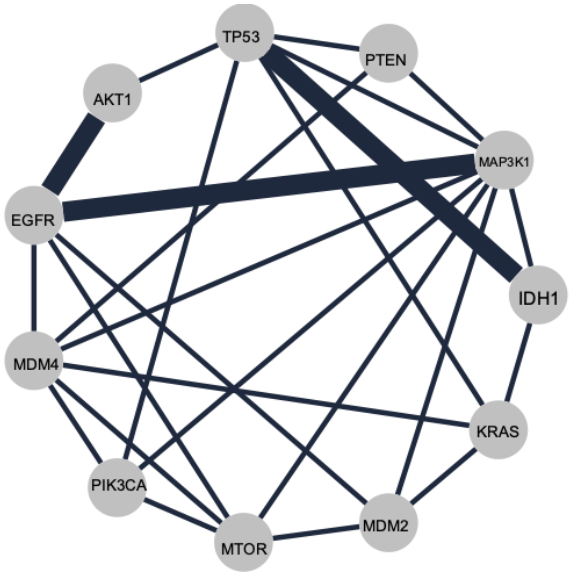


Figure 4 (b). Network of the genes of interest showing their co-occurrence relationship. The network is generated by cystoscape.

Figures above were using the log2 ration and the p-values that were uploaded to cytoscape to generate the networks.

3.3. Mutations

The mutations in each of the query genes were visualized with the cBioPortal through the mutations tab which shows the mutations occurring in every domain of the gene, their frequencies, along with the mutation type and annotations. The annotation shows various information regarding a specific type of protein change such as the clinical implications, hotspots, and general information. Each of the following figures shows the mutations that are happening in the genes of interest, and the type of mutations happening on their domains along with the frequency of each mutation. The higher the dash the more frequent is the mutation. Mutations with green circle are missense and those with black are nonsense. With for example, EGFR has two common missense mutations. One of these mutations is A289 V/T/D.

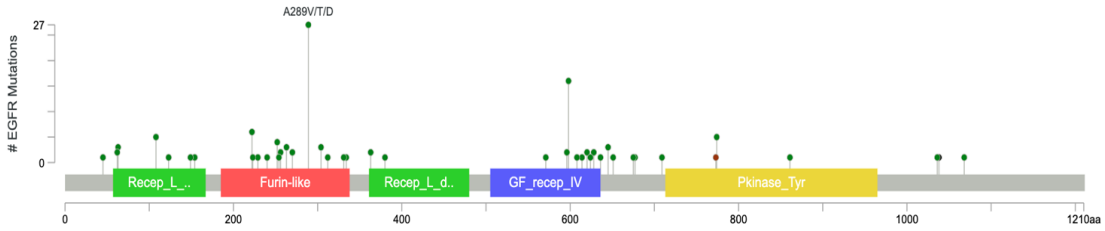


Figure 5 (a)

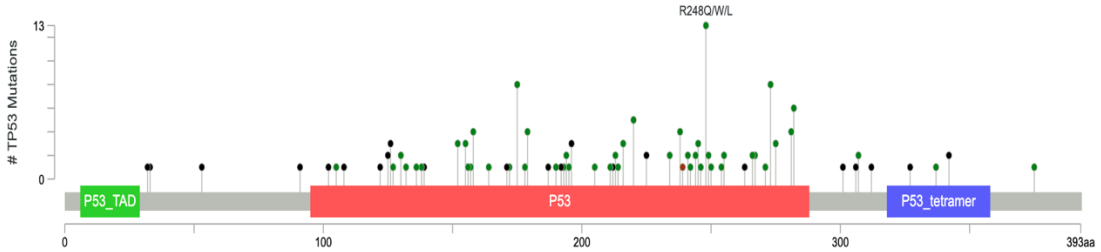


Figure 5 (b)

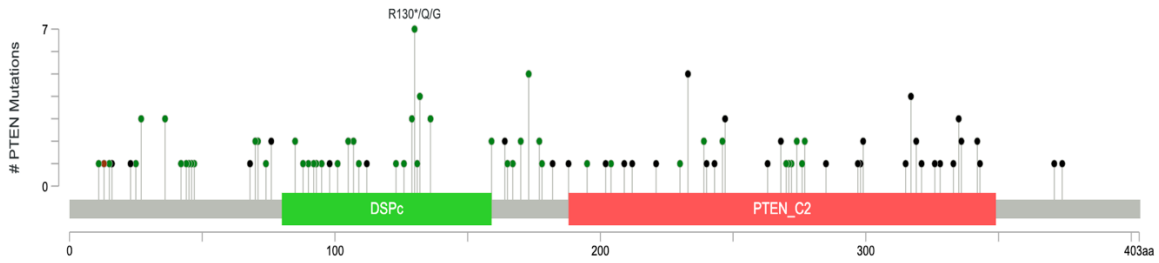


Figure 5 (c)

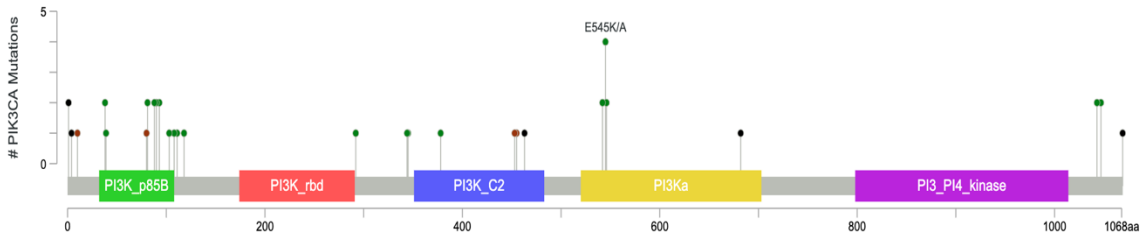


Figure 5 (d)

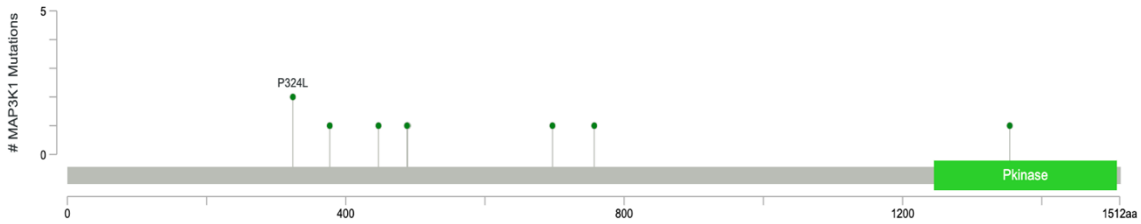


Figure 5 (e)

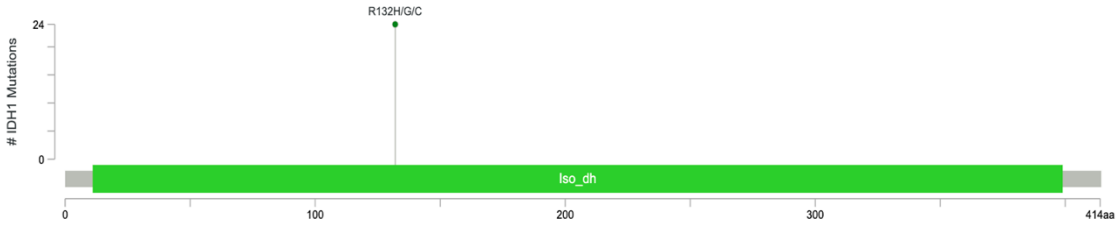


Figure 5 (f)

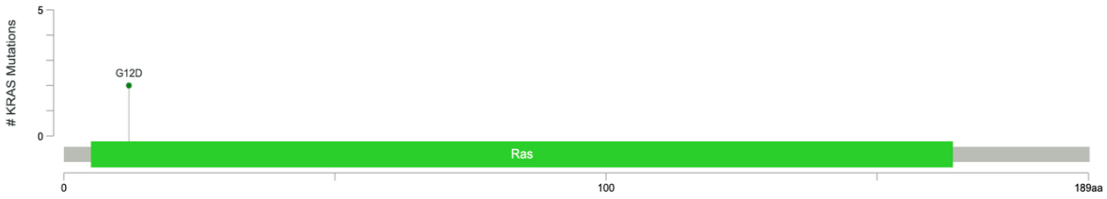


Figure 5 (g)

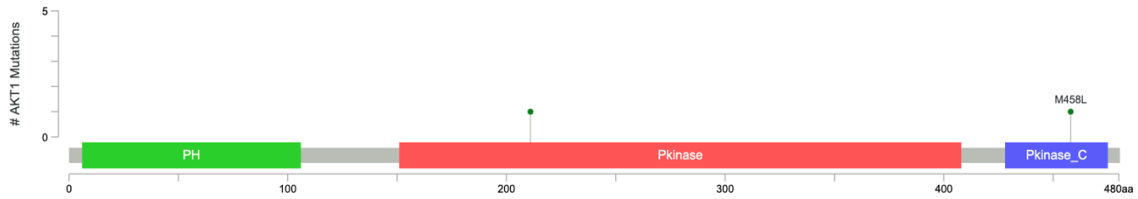


Figure 5 (h)

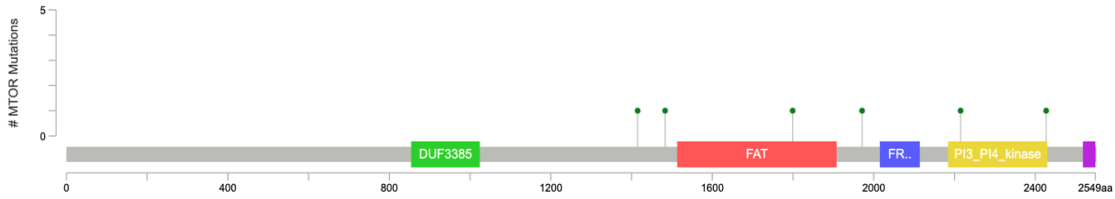


Figure 5 (i)

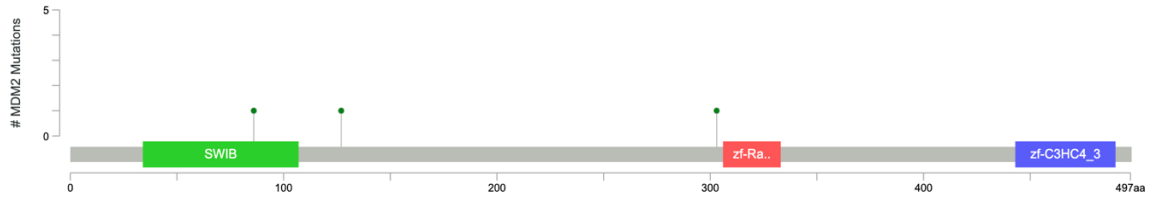


Figure 5 (j)

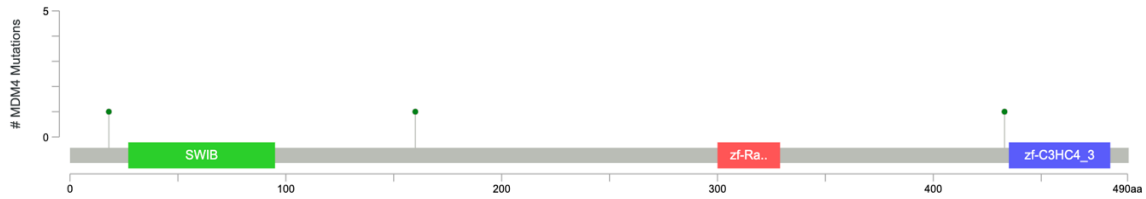


Figure 5 (k)

Figure 5 (a-k). The lollipop models of the query genes showing respectively *EGFR, TP53, PTEN, PIK3CA, MAPK3K1, IDH1, mTOR, KRAS, AKT1, MDM2, MDM*4.

Cerami et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery. May 2012 2; 401. PubMed. Gao et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. 6, pl1 (2013). PubMed

3.4. Mutations Oncogenicity

Oncogenic	Likely oncogenic	Neutral	Inconclusive
A289V(EGFR), G598V(EGFR), R130*(PTEN), E545K/A(PIK3CA) R173H(PTEN), R1329H/G/C(IDH1), G12D(KRAS), C1483R(MTOR)	A289D(EGFR), A289T(EGFR), R130Q(PTEN), R233*(PTEN), R173C(PTEN), R248Q/WL (TP53)	G598A(EGFE)	P324L(MAP3K1), T160S(MDM4), D86Y(MDM2), T211I(AKT1), M458L(AKT1), A19171T(MTOR) (However can be oncogenic).

Table 2. Oncogenicity status for specific mutations in GBM. Data from cBioPortal. Genes are indicated in the parenthesis.

Cerami et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery. May 2012 2; 401. PubMed. Gao et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. 6, pl1 (2013). PubMed

Oncogenic: Strong evidence shows that the alteration is established in the literature as promoting cell proliferation or other hallmark of cancer as defined by Douglas Hanahan and Robert Weinberg (Hanahan and Weinberg, 2011).

Likely oncogenic: Evidence suggests the alteration likely promotes cell proliferation or other hallmarks of cancer as defined by Douglas Hanahan and Robert Weinberg (Hanahan and Weinberg, 2011).

Neutral: Evidence suggests the alteration does not alter protein activity or does not confer growth or survival advantage when expressed in cells.

Inconclusive: There is conflicting and/or weak data describing the oncogenic effect of the mutant alteration

The hallmarks of cancer according to Douglas Hanahan and Robert Weinberg:

1. Continuous proliferation signaling
2. Resistance to cell death
3. Escape growth suppressors
4. Metastasis
5. Angiogenesis

3.5. Therapies for GBM

Mutation	Therapy
A289V/T/D. (EGFR)	Erlotinib, afatinib,gefitinib, and lapatinib
G598V/A (EGFR)	Erlotinib, afatinib, and gefitinib
R130*/Q (PTEN)	GSK2636771 and AZD8186
R173H/C (PTEN)	GSK2636771 and AZD8186
R233* (PTEN)	GSK2636771 and AZD8186
R248 Q/W/L (TP53)	NA
R132H/G/C (IDH1)	Ivosidenib
E545K/A (PIK3CA)	Aleplisib (Effect unknown)
P324L (MAP3K1)	NA
A19171T (MTOR)	MLN0128 and GDC-0980
C1483R (MTOR)	NA
G12D (KRAS)	MEK & ERK inhibitors
T211I (AKT1)	NA
D86Y (MDM2)	NA
T160S (MDM4)	NA

Table 3. Therapies for specific mutations in GBM. Data from cBioPortal.

Cerami et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery. May 2012 2; 401. PubMed. Gao et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. 6, pl1 (2013). PubMed

NA: Not Applicable

3.6. Comparison and survival

The overall patient survival graphs were viewed through the comparison and survival tab. The unaltered group is the group of people in the sample that don't have any alteration in any of the query genes.

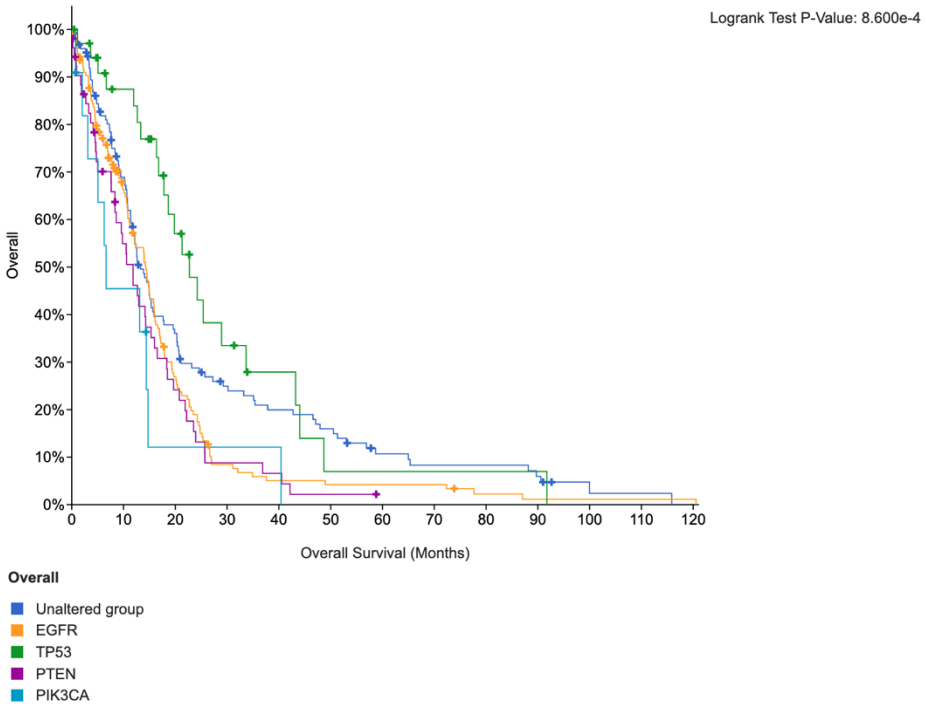


Figure 6. The overall survival of the unaltered group in comparison to the groups of the query altered genes.

The unaltered group contains all the patients with GBM that aren't having any alterations in the genes of interest, however they might be harboring alterations in different genes which aren't in the scope of the study. Some of the data for the patients in blue were showing that they were undergoing therapy while others were without therapy. Patients with TP53 showed better survival than unaltered group for the first 42.75 months, this can be contributed to harboring ID1 mutations which gives patients better prognosis.

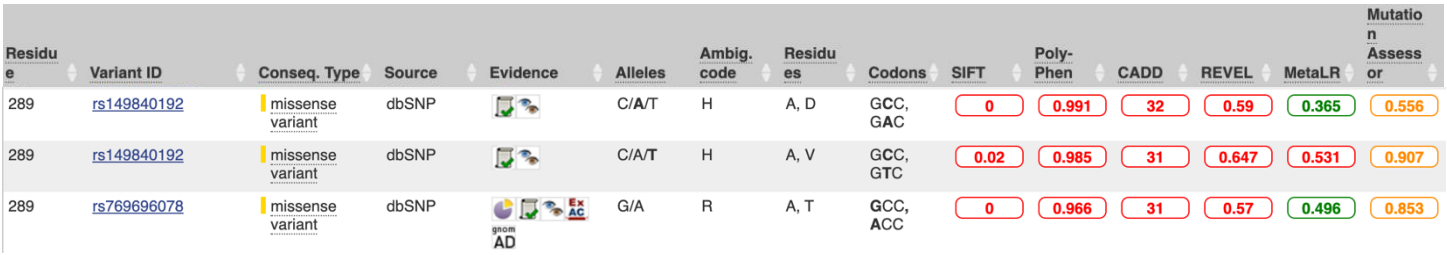
Chapter Four

Discussion

This study presents a genomic evaluation of eleven mutated genes in GBM. These genes are EGFR, TP53, PTEN, PIK3CA, MAP3K1, IDH1, mTOR, KRAS, AKT1, MDM2, MDM4. The reason behind the choice of those genes is that they are part of very important pathways that play major roles in the glioblastoma multiforme tumorigenesis and progression. Glioblastoma multiforme is a very complex and aggressive type of cancer, the reason behind this complexity is the mutations and the alterations that occur in different genes. In the 592 samples of glioblastoma multiforme chosen, EGFR was mutated the most with a percentage of 47% (figure 1,2). Most of mutations in this gene were amplifications, shown in red in the Oncoprint in figure (1) and in blue in the graph in figure (3) that coincided with missense mutations. Some of these missense mutations were putative drivers while others were of unknown significance.

According to Binder et al. (2018) EGFR is known in GBM for its increased transcriptional activity leading to its overexpression. This increase happens in early stages of the GBM, so it is related to the malignancy. EGFR controls lots of other genes, such as PI3K, mTOR, AKT, RAS, as well as PTEN which plays a role in inhibiting AKT. According to table 2, and to the network produced by cytoscape in (figure 4a) EGFG, according to the results from cBioPortal was found to be mutually exclusive with TP53 and IDH1 (p-value <0.01) which means that whenever there was a mutational event in EGFR overrode the mutations in TP53 and IDH1. EGFR co-occur with AKT1 and MAP3K1. *(*Figure 5 a) shows all the mutation that were found in this sample. The two most prevalent mutations were at

position 289 and position 598. According to the lollipop model that is generated by cBioportal it showed that 26 out of 126 mutations were A289V/T/D and 16 mutations were G598 V/A. According to cBioPortal and Binder et al. (2018), these two missense mutations are highly oncogenic. This was shown using Mutationtaster which is a software that predicts the effect of different mutations on a certain gene. In order to verify the results for the specific mutation are valid, a transcript used for this search is from Ensembl and it is the same transcript that is used for EGFR in the cBioPortal (ENST00000275493). From Ensembl a single nucleotide polymorphism (SNP), rs149840192 was identified related to A289V/T/D. Similarly, for G598V/A SNP, rs139236063 was found. SNPs are the most common type of genetic variation. They can be silent, harmless, or harmful. SNPs can be used as a genetic marker for specific genes if they occur in high frequencies as well as cancer markers (Mahdi et al., 2013). The change in the base pair at 289 and 598 was found in Ensembl that shows the substitution that took place. After the search, it was found by Mutationtaster that both missense mutations and SNPs are disease causing which validates the data from cBioportal (Table 2). Because mutation taster gives information about the DNA change and not the amino acid change, additional input is needed to know if the amino acid that is mentioned in cBioportal is validated. A closer look at the rs149840192 when there is a nucleotide change from A to T at 289, revelead that the amino acid changes from A (Alanine) to V (Valine) (Figure 7).



Residue	Variant ID	Conseq. Type	Source	Evidence	Alleles	Ambig. code	Residues	Codons	SIFT	Poly-Phen	CADD	REVEL	MetaLR	Mutation Assessor
289	rs149840192	missense variant	dbSNP		C/A/T	H	A, D	GCC, GAC	0	0.991	32	0.59	0.365	0.556
289	rs149840192	missense variant	dbSNP		C/A/T	H	A, V	GCC, GTC	0.02	0.985	31	0.647	0.531	0.907
289	rs769696078	missense variant	dbSNP	gnomAD	G/A	R	A, T	GCC, ACC	0	0.966	31	0.57	0.496	0.853

Figure 7. REVEL, MetalR and Mutation Assessor values for mutation at residue 289 of EGFR

REVEL: The Rare Exome Variant Ensemble Learner REVEL is an ensemble method for predicting the pathogenicity of missense variants. Score range from 0 to 1 and variants with higher scores are predicted to be more likely to be pathogenic.

MetaLR: MetaLR uses logistic regression to predict the deleteriousness of missense variants. Variants are classified as 'tolerated' or 'damaging'; a score between 0 and 1 is also provided and variants with higher scores are more likely to be deleterious.

Mutation Assessor: Mutation Assessor predicts the functional impact of amino-acid substitutions in proteins. The predictions are 'neutral', 'low', 'medium' and 'high', and the rank score, which is between 0 and 1 where variants with higher scores are more likely to be deleterious.

The color coded scores are the results of a number of algorithms that evaluate the effect of mutations on the genes function. The scores are calculated by the dbNSFP project.

As shown in figure 7 the change of A to V has the highest REVEL, MetaLR, and mutation assessor scores compared to the change of A to D, or to T. This shows that A to V amino acid change is more oncogenic and damaging than the other mutations which validates the results shown in table 2. Similarly, for G598V/A the amino acid change from G to V more oncogenic and damaging than the G to A change (figure 7 and table 2).

Also, according to figure 6 patients with EGFR mutations showed low survival rates which suggests the aggressivity of such mutations. In addition to this mutation taster detected that these mutations happened on chromosome 7, as well the PhastCons values. The PhastCons values can be between 0 and 1 and they show the probability of a

nucleotide to belong to a conserved element, the closer the value is to 1 the more the nucleotide is conserved. A289V/T/D and G598V/A demonstrated values of 0.996 and 0.997 respectively, showing that they are part of a conserved element cBioPortal as well shows that these two mutations are cancer hotspot, meaning that the change of amino acids is recurrent and statistically significant on these residues. In addition, cBioPortal identified Lapatinib as an inhibitor for A289V/T/D, however no drugs were identified for G598V/A (Table 3).

PTEN mutations and loss of function play an important role in the progress from lower grade astrocytomas to glioblastoma multiforme. When PTEN expression is induced in glioma cells that are PTEN-null, this inhibits cell migration, growth, invasion, and focal adhesions. Figure 1, 2, and 3 shows the alterations that are taking place in PTEN according to the sample. According to table 2 and to the network produced by cytoscape in figure 4 (a), PTEN was found to be mutually exclusive with IDH1 which suggests that in this sample whenever there was a mutational event in PTEN, it overrode the mutations in IDH1. Figure 5 (b) shows all the mutations that were found in PTEN in the sample from cBioPotal. The two most prevalent mutation were at position 130 and position 137. According to the lollipop model that is generated by cBioPortal it showed that 6 out of 132 mutations were R130*/Q and 5 mutations were R173H/C. Out of the 6 mutations of R130, 3 mutations were missense (R130Q) where Arginine changed into Glutamine, while the other 3 were nonsense (R130*). In addition to this, a change in the amino acid and a nonsense mutation (R223*) causes a truncation at residue 223.The two-missense mutation, R130Q and R173H occur in the phosphatase domain of PTEN thus affecting it phosphatase activity, while R233 occur in the C2 domain, that plays a vital role in the

binding of PTEN to the lipid plasma membrane (Chang et al., 2019). According to cBioPortal, R130Q and R223* are likely to be oncogenic, while R130* and R173H are known to be oncogenic (table 2). Similarly, to verify this, mutation taster was used. The PTEN transcript ID (ENST00000371953) used in Ensmbl was stated in the cBioPortal. Using Ensmbl the SNPs and the variants on these residues were determined, and the results showed that the mutation at R130Q and R173H are disease causing through amino acid sequence change that affects the protein. The REVEL, MetaLR, and mutation assessor scores of R130Q and R173C/H showed they are both oncogenic and damaging which align with results in table 2. Previous studies showed that these three mutations led to loss of function of the PTEN. Similarly, the truncation at R130* and R233* led to loss of function as well (Carico et al., 2012). Considering that these are hotspots for glioblastoma, drugs have been used to treat patients with these alterations. Laboratory data suggest that cancer cells with loss-of-function PTEN alterations may be sensitive to beta-isoform selective PI3K-targeted inhibitors, such as the investigational agents GSK2636771 and AZD8186 (table 3).

TP53, IDH1, PIK3CA, MAP3K1, mTOR, KRAS, AKT1, MDM2, and MDM4 showed missense mutations according to cBioPortal. mTOR and MDM2 showed alterations in the form of fusions. The most common alteration in TP53, IDH1, PIK3CA, MAP3K1, mTOR, KRAS, AKT1, MDM2, and MDM4 respectively are R248Q/W/L, R132H/C/G, E545K/A, P324L, A19171T C1483R, G12D, T211I, D86Y, and T160S. These results are shown in figure 1, 2, and 3.

According to cBioPortal, TP53 R248 Q/W/L are likely to be oncogenic with high values of REVEL, MetaLR, and mutation assessor (table 2). Data from Ensmbl regarding the

variant rs121912651 shows that the clinical significance of these mutations is pathogenic. These mutations are called contact mutations that lead to the disruptions of the binding to DNA that increase the invasive behavior of glioblastoma cells (Carico et al.,2012).

IDH1 showed 23 missense mutations, which are known to be oncogenic. In glioblastoma, IDH1 mutations have been associated with specific cytogenetic abnormalities, 1p and 19q deletions. There is favorable clinical data in subjects with IDH1 R132-mutant glioma treated with the IDH1-targeted inhibitor ivosidenib (table 3).

PIK3CA, MAP3K1, mTOR, AKT1, and KRAS are all related. PIK3CA showed E545K/A mutations, which is known to be oncogenic (table 2). These mutations most frequently occur in the helical domain (hotspots E545K and E545A) (Ranjbar et al.,2019). When rs121913275 in Ensmbl was tested using mutation taster, it showed that this mutation is oncogenic which validated the results in cBioPortal (table 2). The REVEL scores of these two mutations showed that they are likely oncogenic, however the metaLR and mutation assessors scores suggested that they are tolerated and non-damaging. These mutation assessors scores are affected by homologues chromosomes mutations which might be the reason for the discrepancy.

MAP3K1 had P324L mutation, according to cBioPortal the biological significance for this mutation is unknown (table 2). In order to find a significance, this mutation was tested using mutation taster. Mutation taster predicted that variant rs1336098946 is a disease-causing SNP that is accompanied by change in the protein features. Recent studies showed that MAP3K1 is highly expressed in glioblastoma. This increased expression is highly associated with poor prognosis and therapeutic resistance in glioblastoma (Wang et al., 2019).

The rest of the genes show few mutations in the sample under study. Most of the mutations on mTOR were missense such as L2427Q, A1971T, C1483R, and S2215T. All these mutations are known to be on oncogenic except the A1971T mutation that has an unknown effect (last reviewed 06/20/2018) (table 2). The REVEL, metaLR, and mutation assessor values for this mutation showed that A1971T is damaging and likely to be disease causing. This was also validated by mutation taster that showed that this mutation is a disease-causing mutation. All the mutations are considered hotspots. According to cBioPortal, renal cancer patients with L2427Q mutations respond well to FDA-approved mTOR-targeted inhibitor temsirolimus however its utility in glioblastoma patients is yet to be determined. Patients harboring the other mutations are sensitive to the mTOR-targeted allosteric inhibitors such as sirolimus, everolimus and temsirolimus, or mTOR-targeted ATP-competitive inhibitors such as MLN0128 and GDC-0980 (table 3). Kras showed G12D mutation. This mutation is known to be oncogenic (table 2), and laboratory and preliminary clinical data suggest that KRAS-mutant cancers may be sensitive to MEK- or ERK-targeted inhibitors such as Trametinib, Cobimetinib, and Binimetinib (table 3). Regarding AKT1, MDM2, and MDM4 no information was found for their mutations.

Mutations Oncogenicity 2

Oncogenic	**Likely oncogenic**	**Neutral**	**Inconclusive**
A289V(EGFR), G598V(EGFR), R130*(PTEN), R173H(PTEN), R1329H/G/C(IDH1), G12D(KRAS), C1483R(MTOR), ***R130Q(PTEN), R173C(PTEN),***	A289D(EGFR), A289T(EGFR), R233*(PTEN), R248Q/WL (TP53), ***E545K/A(PIK3CA***)	G598A(EGFR)	T160S(MDM4), D86Y(MDM2), T211I(AKT1), M458L(AKT1)

P324L(MAP3K1), A19171T(MTOR)			

Table 4. Oncogenicity status for specific mutations in GBM after checking the REVEL, metaLR, and mutation assessor values.

The mutations in bold and italics the mutations that their oncogenicity changed from table 3.

Comparing of GBM gene mutations between the Lebanese-GBM group and the TCGA group

In the attempt to compare between an international database and a local database.Samples from 60 GBM patients in Lebanon from June 2015 and June 2016 were collected. Then these samples underwent whole exome sequencing (WES) to highlight the mutations in this cohort. The results from this sample were compared to the results that are on TCGA and cBioPortal (Saadeh et al., 2022). IDH1 R132H mutation showed higher frequency in the Lebanese sample 10% in comparison with the TCGA and cBioPortal which was 3.5% as shown in table 5. This difference most probably is due to the small sample; however, it indicates the importance of IDH1 R132H mutation in the Lebanese population. EGFR A289Dand R222C mutations were 0% in the Lebanese GBM sample while in TCGA and cBioPortal their percentages were 0.8% and 1% respectively.

Gene	Mutation	# Affected Samples (n=60) Lebanese GBM	%Affected Samples(n=60) Lebanese GBM	#TCGA (n=592) cBIOPortal	%TCGA (n=592) cBIOPortal
IDH1	R132H	6	10	21	3.5
EGFR	A289V	1	1.7	16	2.7
EGFR	G598V	3	5	15	2.5
EGFR	A289D	0	0	5	0.8
EGFR	R222C	0	0	6	1
TP53	R248W	1	1.7	5	0.8
TP53	R175H	2	3.3	8	1.3
TP53	R248Q	2	3.3	8	1.3

Table 5. Comparison of common GBM gene mutations between the Lebanese GBM cohort and the TCGA cohort.

Data from "Correlation of genetic alterations by whole-exome sequencing with outcomes of glioblastoma patients"

The standard of care for GBM is possible surgical resection of the cancer directly followed by combining of radiotherapy and chemotherapy. The chemotherapy used for GBM is Temozolomide (TMZ). The sensitivity of TMZ depends on the gene product of MGMT promoter. As the cellular concentration of MGMT increases the glioblastoma multiforme cells sensitivity to TMZ decreases.

A study by Hamed at el. (2015) showed that the drug combination of sorafenib and lapatinib interacted together to kill the CNS tumor cells. Similar results were obtained as well with regorafenib. Data from cBioPortal showed that the clinical utility of afatinib in patients with glioblastoma multiforme harboring the EGFR G598V mutation is unknown. Jensen et al. (2020) demonstrated that combinatorial therapies targeting the signaling

pathways of EGFR and STAT3 hold therapeutic potential for GBM. The *in vitro* data showed that EGFR-mutant glioblastoma cells are sensitive towards afatinib. *In vivo*, in a clinical case study done on a patient with EGFR-mutation, the progression-free survival after treatment with afatinib and protracted TMZ was longer. The differential sensitivities of GBM patients to EGFR inhibition validates the significance of stratifying patients, depending on their molecular profile to use targeted treatments. Thus, a rise in pro-survival STAT3 signaling, both *in vitro* and *in vivo* following EGFR inhibition was observed.

Another potential targeted therapy for glioblastoma is via targeted PI3K/mTOR inhibition through PQR309. It was reported previously that this inhibitor crosses the blood brain barrier and that it inhibits PI3K/mTORC1/2. The proliferation of U87 and U251 cells was inhibited by PQR309. After 24 hours it didn't show any liver toxicity in female rats. This inhibitor led to G1 arrest, reduction in wound-healing, cell migration and cell invasion. In addition, cell apoptosis was promoted. The expression of p-AKT, AKT, Bcl-2, Bcl-xL, Bax, Bad, cleaved caspase-3, MMP-2, MMP-9 and cyclin D1 supported these results. Even though the pharmacokinetic parameter were favorable in rats, mice, and dogs, clinical trials on patients with glioblastoma multiforme is yet to be performed (Yang et al., 2019).

Regarding TP53 mutations, no data is available regarding GMB. However, a study conducted on colorectal cancers in 2018 showed that TP53 R248Q mutations were very common and led to gain of function. Treating the tumor *in vivo* by HSP90 inhibitor suppressed the mutated levels of TP53 and decreased the tumor growth (Schulz-Heddergott et al.,2018). This approach could be used in treatment of glioblastoma,

however the main challenge in GBM is that Hsp90 inhibitors are not able to cross the blood-brain barrier (BBB).

NXD30001, is a novel series of Hsp90 inhibitors. It can cross the BBB and has better pharmacokinetic activity than other Hsp90 inhibitors. This inhibitor doesn't accumulate in the brain and doesn't show any liver toxicity in vivo. NXD30001 induce tumor regression and apoptosis. Chen et al. (2020), showed that the combination between. In addition to this, a better inhibitory effect was witnessed on the GBM growth in vivo. Knowing that GBM exhibit different alternations in EGFR, PTEN, TP53 and Akt as shown in the sample under study, and most of these proteins are Hsp90 clients which means that inhibiting the activity of Hsp90 will lead to the downregulation of these proteins. However, this approach can face lots of challenges. Hsp90 compensation by other co-chaperons might be induced, and studies had shown that cells are less dependent on Hsp90 *in vivo* than in *in vitro*.

Targeting these genes can affect many cellular processes that GBM use in glimoagensis. One of these processes is autophagy. In an experiment done by Gammoth et al. (2018) showed that the inhibition of autophagy impaired the cancerous growth and affected the phosphorylation of AKT and MAPK1/3. In addition to this mTORC1 signaling was altered leading to the induction of senescence. TP53 wasn't affected under those conditions suggesting that senescence might be independent of this protein. These findings come in-line with the results found in figure 1, 2,3 and 4 as AKT, MAPK1/3, mTORC1 were altered at low % and TP53 at higher % showing that autophagy might be playing a role in the tumorigenesis of the GBM in this sample. Autophagy is suggested to play a protective role in GBM through giving the tumor that ability to acquire resistance to the

antiangiogenic therapy. Snice autophagy is regulated with pathways that also regulate tumorigenesis, then the inactivation of these genes will influence inhibiting the autophagy's protective role in GBM.

Chapter five

Conclusions and Perspectives

Studying those various genes gives an insight of how they all play together in the tumorigenesis of GBM. Even though it is known how these interact, it is still a challenge to target them specially in the presence of BBB. Knowing the different mutations that every cancer patient harbors, such as the ones mentioned above can be helpful in starting a personalized treatment for each. The bioinformatics tools used here allowed for the quick and easy access to many genes that are critical in GBM. With these tools it was easy to study the mutations in the genes of interests, validate the nature of mutations found in each, and investigate new therapies. Knowing that the mutations studied here were only missense mutations, future studies can explore new techniques that can study other types of alterations such as truncations and missense mutations. In addition, values such as SIFT and Polyphen can be studied to get a better knowledge of the effect of mutations under study. Finally, comparing international database with Lebanese samples and studying the sample further using different tools will give a better understanding of the tumorigenesis and the dynamics of GBM in certain geographical spots.

References

Ostrom, Q. T., Bauchet, L., Davis, F. G., Deltour, I., Fisher, J. L., Langer, C. E., Pekmezci, M., Schwartzbaum, J. A., Turner, M. C., Walsh, K. M., Wrensch, M. R., & Barnholtz-Sloan, J. S. (2015). Response to "the epidemiology of glioma in adults: A 'state of the Science' Review". *Neuro-Oncology*, *17*(4), 624–626. https://doi.org/10.1093/neuonc/nov022

Alifieris, C., & Trafalis, D. T. (2015). Glioblastoma multiforme: Pathogenesis and treatment. *Pharmacology & Therapeutics*, *152*, 63–82. https://doi.org/10.1016/j.pharmthera.2015.05.005

Stoyanov, G. S., Dzhenkov, D., Ghenev, P., Iliev, B., Enchev, Y., & Tonchev, A. B. (2018). Cell biology of glioblastoma multiforme: From basic science to diagnosis and treatment. *Medical Oncology*, *35*(3). https://doi.org/10.1007/s12032-018-1083-x

Chen, R., Smith-Cohn, M., Cohen, A. L., & Colman, H. (2017). Glioma subclassifications and their clinical significance. *Neurotherapeutics*, *14*(2), 284–297. https://doi.org/10.1007/s13311-017-0519-x

Wrensch, M., Minn, Y., Chew, T., Bondy, M., & Berger, M. S. (2002). Epidemiology of primary brain tumors: Current concepts and review of theliterature. *Neuro-Oncology*, *4*(4), 278–299. https://doi.org/10.1093/neuonc/4.4.278

Louis, D., Ohgaki, H., Wiestler, O., Cavenee, W., Burger, P., Jouvet, A., Scheithauer, B. and Kleihues, P., 2007. The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropathologica*, 114(2), pp.97-109.

Molinaro, A., Taylor, J., Wiencke, J. and Wrensch, M., 2019. Genetic and molecular epidemiology of adult diffuse glioma. *Nature Reviews Neurology*, 15(7), pp.405-417.

Yan, H., Parsons, D., Jin, G., McLendon, R., Rasheed, B., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G., Friedman, H., Friedman, A., Reardon, D., Herndon, J., Kinzler, K., Velculescu, V., Vogelstein, B. and Bigner, D., 2009. IDH1 and IDH2 Mutations in Gliomas. *New England Journal of Medicine*, 360(8), pp.765-773.

Parsons, D., Jones, S., Zhang, X., Lin, J., Leary, R., Angenendt, P., Mankoo, P., Carter, H., Siu, I., Gallia, G., Olivi, A., McLendon, R., Rasheed, B., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D., Tekleab, H., Diaz, L., Hartigan, J., Smith, D., Strausberg, R., Marie, S., Shinjo, S., Yan, H., Riggins, G., Bigner, D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. and Kinzler, K., 2008. An Integrated Genomic Analysis of Human Glioblastoma Multiforme. *Science*, 321(5897), pp.1807-1812.

Eckel-Passow, J., Lachance, D., Molinaro, A., Walsh, K., Decker, P., Sicotte, H., Pekmezci, M., Rice, T., Kosel, M., Smirnov, I., Sarkar, G., Caron, A., Kollmeyer, T., Praska, C., Chada, A., Halder, C., Hansen, H., McCoy, L., Bracci, P., Marshall, R., Zheng, S., Reis, G., Pico, A., O'Neill, B., Buckner, J., Giannini, C., Huse, J., Perry, A., Tihan, T., Berger, M., Chang, S., Prados, M., Wiemels, J., Wiencke, J., Wrensch, M. and Jenkins, R., 2015. Glioma Groups Based on 1p/19q, IDH, and TERTPromoter Mutations in Tumors. *New England Journal of Medicine*, 372(26), pp.2499-2508.

Bendahou, M., Arrouchi, H., Lakhlili, W., Allam, L., Aanniz, T., Cherradi, N., Ibrahimi, A. and Boutarbouch, M., 2020. Computational Analysis of IDH1, IDH2, and TP53 Mutations in Low-Grade Gliomas Including Oligodendrogliomas and Astrocytomas. *Cancer Informatics*, 19, p.117693512091583.

Ludwig, K. and Kornblum, H., 2017. Molecular markers in glioma. *Journal of Neuro-Oncology*, 134(3), pp.505-512.

Wu, C., Song, H., Fu, X., Li, S. and Jiang, T., 2020. Transcriptomic Analysis of Glioma Based on IDH Status Identifies ACAA2 as a Prognostic Factor in Lower Grade Glioma. *BioMed Research International*, 2020, pp.1-8.

Wee, P. and Wang, Z., 2017. Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers*, 9(12), p.52.

Xu, H., Zong, H., Ma, C., Ming, X., Shang, M., Li, K., He, X., Du, H. and Cao, L., 2017. Epidermal growth factor receptor in glioblastoma. *Oncology Letters*, 14(1), pp.512-516.

Morrison, D., 2012. MAP Kinase Pathways. *Cold Spring Harbor Perspectives in Biology*, 4(11), pp. a011254-a011254.

Kano, Y., Cook, J., Lee, J. and Ohh, M., 2016. New structural and functional insight into the regulation of Ras. *Seminars in Cell & Developmental Biology*, 58, pp.70-78.

An, Z., Aksoy, O., Zheng, T., Fan, Q. and Weiss, W., 2018. Epidermal growth factor receptor and EGFRvIII in glioblastoma: signaling pathways and targeted therapies. *Oncogene*, 37(12), pp.1561-1575.

Guo, Y., Pan, W., Liu, S., Shen, Z., Xu, Y. and Hu, L., 2020. ERK/MAPK signalling pathway and tumorigenesis (Review). *Experimental and Therapeutic Medicine.*

Mao, H., LeBrun, D., Yang, J., Zhu, V. and Li, M., 2012. Deregulated Signaling Pathways in Glioblastoma Multiforme: Molecular Mechanisms and Therapeutic Targets. *Cancer Investigation*, 30(1), pp.48-56.

Thorpe, L., Yuzugullu, H. and Zhao, J., 2014. PI3K in cancer: divergent roles of isoforms, modes of activation and therapeutic targeting. *Nature Reviews Cancer*, 15(1), pp.7-24.

Hemmings, B. and Restuccia, D., 2012. PI3K-PKB/Akt Pathway. *Cold Spring Harbor Perspectives in Biology*, 4(9), pp.a011189-a011189.

Akhavan, D., Cloughesy, T. and Mischel, P., 2010. mTOR signaling in glioblastoma: lessons learned from bench to bedside. *Neuro-Oncology*, 12(8), pp.882-889.

Mecca, C., Giambanco, I., Donato, R. and Arcuri, C., 2018. Targeting mTOR in Glioblastoma: Rationale and Preclinical/Clinical Evidence. *Disease Markers*, 2018, pp.1-10.

Majewska, E. and Szeliga, M., 2016. AKT/GSK3β Signaling in Glioblastoma. *Neurochemical Research*, 42(3), pp.918-924.

Zhao, H., Wang, J., Shao, W., Wu, C., Chen, Z., To, S. and Li, W., 2017. Recent advances in the use of PI3K inhibitors for glioblastoma multiforme: current preclinical and clinical development. *Molecular Cancer*, 16(1).

Paul-Samojedny, M., Pudełko, A., Suchanek-Raif, R., Kowalczyk, M., Fila-Daniłow, A., Borkowska, P. and Kowalski, J., 2014. Knockdown of the AKT3 (PKBγ), PI3KCA, and VEGFR2 genes by RNA interference suppresses glioblastoma multiforme T98G cells invasiveness in vitro. *Tumor Biology*, 36(5), pp.3263-3277.

Bao, L., Li, X. and Lin, Z., 2019. PTEN overexpression promotes glioblastoma death through triggering mitochondrial division and inactivating the Akt pathway. *Journal of Receptors and Signal Transduction*, 39(3), pp.215-225.

Zhang, Y., Dube, C., Gibert, M., Cruickshanks, N., Wang, B., Coughlan, M., Yang, Y., Setiady, I., Deveau, C., Saoud, K., Grello, C., Oxford, M., Yuan, F. and Abounader, R., 2018. The p53 Pathway in Glioblastoma. *Cancers*, 10(9), p.297.

England, B., Huang, T. and Karsy, M., 2013. Current understanding of the role and targeting of tumor suppressor p53 in glioblastoma multiforme. *Tumor Biology*, 34(4), pp.2063-2074.

Gao, J., Aksoy, B., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C. and Schultz, N., 2013. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Science Signaling*, 6(269), pp.pl1-pl1.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B., Sumer, S., Aksoy, B., Jacobsen, A., Byrne, C., Heuer, M., Larsson, E., Antipin, Y., Reva, B., Goldberg, A., Sander, C. and Schultz, N., 2012. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discovery*, 2(5), pp.401-404.

Wu, P., Heins, Z., Muller, J., Katsnelson, L., de Bruijn, I., Abeshouse, A., Schultz, N., Fenyö, D. and Gao, J., 2019. Integration and Analysis of CPTAC Proteomics Data in the Context of Cancer Genomics in the cBioPortal. *Molecular & Cellular Proteomics*, 18(9), pp.1893-1898.

Binder, Z., Thorne, A., Bakas, S., Wileyto, E., Bilello, M., Akbari, H., Rathore, S., Ha, S., Zhang, L., Ferguson, C., Dahiya, S., Bi, W., Reardon, D., Idbaih, A., Felsberg, J., Hentschel, B., Weller, M., Bagley, S., Morrissette, J., Nasrallah, M., Ma, J., Zanca, C., Scott, A., Orellana, L., Davatzikos, C., Furnari, F. and O'Rourke, D., 2018. Epidermal Growth Factor Receptor Extracellular Domain Mutations in Glioblastoma Present Opportunities for Clinical Imaging and Therapeutic Development. *Cancer Cell*, 34(1), pp.163-177.e7.

Mahdi, K., Nassiri, M. and Nasiri, K., 2013. Hereditary Genes and SNPs Associated with Breast Cancer. *Asian Pacific Journal of Cancer Prevention*, 14(6), pp.3403-3409.

Abounader, R., 2009. Interactions between PTEN and receptor tyrosine kinase pathways and their implications for glioma therapy. *Expert Review of Anticancer Therapy*, 9(2), pp.235-245.

Chang, H., Cai, Z. and Roberts, T., 2019. The Mechanisms Underlying PTEN Loss in Human Tumors Suggest Potential Therapeutic Opportunities. *Biomolecules*, 9(11), p.713.

Carico, C., Nuño, M., Mukherjee, D., Elramsisy, A., Dantis, J., Hu, J., Rudnick, J., Yu, J., Black, K., Bannykh, S. and Patil, C., 2012. Loss of PTEN Is Not Associated with Poor Survival in Newly Diagnosed Glioblastoma Patients of the Temozolomide Era. *PLoS ONE*, 7(3), p.e33684.

Shajani-Yi, Z., de Abreu, F., Peterson, J. and Tsongalis, G., 2018. Frequency of Somatic TP53 Mutations in Combination with Known Pathogenic Mutations in Colon Adenocarcinoma, Non–Small Cell Lung Carcinoma, and Gliomas as Identified by Next-Generation Sequencing. *Neoplasia*, 20(3), pp.256-262.

Zhou, Y., Wu, W., Bi, H., Yang, D. and Zhang, C., 2020. Glioblastoma precision therapy: From the bench to the clinic. *Cancer Letters*, 475, pp.79-91.

Ranjbar R, Mohammadpour S, Torshizi Esfahani A, Namazian S, Yaghob-Taleghani M, Baghaei K, et al. Prevalence and prognostic role of PIK3CA E545K mutation in Iranian colorectal cancer patients. Gastroenterol Hepatol Bed Bench 2019;12(Suppl.1):22-29)

McNeill, R., Stroobant, E., Smithberger, E., Canoutas, D., Butler, M., Shelton, A., Patel, S., Limas, J., Skinner, K., Bash, R., Schmid, R. and Miller, C., 2018. PIK3CA missense mutations promote glioblastoma pathogenesis, but do not enhance targeted PI3K inhibition. *PLOS ONE*, 13(7), p.e0200014.

Wang, J., Zuo, J., Wahafu, A., Wang, M., Li, R. and Xie, W., 2019. Combined elevation of TRIB2 and MAP3K1 indicates poor prognosis and chemoresistance to temozolomide in glioblastoma. *CNS Neuroscience & Therapeutics*, 26(3), pp.297-308.

Carlsson, S., Brothers, S. and Wahlestedt, C., 2014. Emerging treatment strategies for glioblastoma multiforme. *EMBO Molecular Medicine*, 6(11), pp.1359-1370.

Hamed, H., Tavallai, S., Grant, S., Poklepovic, A. and Dent, P., 2014. Sorafenib/Regorafenib and Lapatinib Interact to kill CNS Tumor Cells. *Journal of Cellular Physiology*, 230(1), pp.131-139.

Jensen, K., Hao, X., Aman, A., Luchman, H. and Weiss, S., 2020. EGFR blockade in GBM brain tumor stem cells synergize with JAK2/STAT3 pathway inhibition to abrogate compensatory mechanisms in vitro and in vivo. *Neuro-Oncology Advances*, 2(1).

Yang, K., Tang, X., Xu, F., Liu, J., Tan, Y., Gao, L., Sun, Q., Ding, X., Liu, B. and Chen, Q., 2020. PI3K/mTORC1/2 inhibitor PQR309 inhibits proliferation and induces apoptosis in human glioblastoma cells. *Oncology Reports.*

Schulz-Heddergott, R., Stark, N., Edmunds, S., Li, J., Conradi, L., Bohnenberger, H., Ceteci, F., Greten, F., Dobbelstein, M. and Moll, U., 2018. Therapeutic Ablation of Gain-of-Function Mutant p53 in Colorectal Cancer Inhibits Stat3-Mediated Tumor Growth and Invasion. *Cancer Cell*, 34(2), pp.298-314.e7.

Chen, H., Gong, Y., Ma, Y., Thompson, R., Wang, J., Cheng, Z. and Xue, L., 2020. A Brain-Penetrating Hsp90 Inhibitor NXD30001 Inhibits Glioblastoma as a Monotherapy or in Combination with Radiation. *Frontiers in Pharmacology*, 11.

Levy, J., Towers, C. and Thorburn, A., 2017. Targeting autophagy in cancer. *Nature Reviews Cancer*, 17(9), pp.528-542.

Canzoneri, R., Lacunza, E. and C. abba, M., 2019. Genomics and Bioinformatics as Pillars of precision medicine in oncology. *MEDICINA (Buenos Aires) 2019; 79 - N° 6/1 - NÚMERO ESPECIAL 80 ANIVERSARIO: 587-592.*

Taylor, M., Das, B. and Ray, S., 2018. Targeting autophagy for combating chemoresistance and radioresistance in glioblastoma. *Apoptosis*, 23(11-12), pp.563-575.

Pawlowska, E., Szczepanska, J., Szatkowska, M. and Blasiak, J., 2018. An Interplay between Senescence, Apoptosis and Autophagy in Glioblastoma Multiforme—Role in Pathogenesis and Therapeutic Perspective. *International Journal of Molecular Sciences*, 19(3), p.889.

Cerami et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery. May 2012 2; 401. PubMed.

Gao et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. 6, pl1 (2013). PubMed.

Saadeh FS, Morsi RZ, El-Kurdi A, Nemer G, Mahfouz R, Charafeddine M, et al. (2020) Correlation of genetic alterations by whole-exome sequencing with clinical outcomes of glioblastoma patients from the Lebanese population. PLoS ONE 15(11): e0242793. https://doi.org/10.1371/journal. pone.0242793