Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal

Gene Expression Alterations in Central Nervous System Neoplasms With *EGFR* Amplification

Marco Filipe Semião Moedas

Mestrado em Biologia Molecular Humana

Lisboa 2009

Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal

Gene Expression Alterations in Central Nervous System Neoplasms With *EGFR* Amplification

Marco Filipe Semião Moedas

Mestrado em Biologia Molecular Humana

Lisboa 2009

Dissertação orientada por:

Orientadora Externa: Doutora Lúcia Roque, IPOLFG EPE, CIPM

Orientadora Interna: Professora Doutora Maria do Céu Correia, FCUL, DBV

Index.

Agradecimentos

À Doutora Lúcia Roque, pela oportunidade que me deu de realizar este trabalho, pelo apoio e confiança, pelo que me ensinou, pela ajuda e amizade que sempre demonstrou.

À Doutora Carmo Martins, ao Dr. Vasco e à Joaquina pelo que me ensinaram e ajudaram, e sobretudo pela amizade e apoio que me deram.

Aos membros do CIPM, pela ajuda, pelo inestimável apoio e bom ambiente que criaram, e sem o qual este trabalho também não teria sido possível.

Ao Doutor Jorg Becker e à Júlia Lobato, pelos Arrays e pelo Partek, e à Isabel Marques pelo IPA.

À Professora Doutora Maria do Céu Correia, por me ter ajudado em tudo o que precisei.

A todos os SIATAs, pela incondicional amizade e apoio.

À Sofia, por tudo. Pelo apoio, ajuda, confiança e amor, por me ter tornado melhor.

À minha família, em particular os meus pais, irmão e cunhada, por tudo. Por terem contribuído com tudo para que eu pudesse estar aqui, pelo apoio, pelo amor e carinho e pela confiança que sempre me transmitiram.

A todos, um MUITO OBRIGADO

Marco Moedas

Glossary:

AKT	v-akt Murine Thymoma Viral oncogene homolog
ANOVA	Analysis Of Variance
AO	Anaplasic Oligodendroglioma
BCL2	B-cell CLL/lymphoma 2
BCLX	BCL2-like 1
cCGH	Chromosomal Comparative Genomic Hybridization
CNS	Central Nervous System
DNA	Deoxyribonucleic Acid
cDNA	Complementary Deoxyribonucleic Acid
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
EGFRvIII	Epidermal Growth Factor Receptor variant III
FC	Fold Change
FOS	FBJ murine osteosarcoma viral oncogene homolog
GBM	Glioblastoma Multiforme
pGBM	Primary Glioblastoma Multiforme
sGBM	Secondary Glioblastoma Multiforme
HB-EGF	Heparin-Binding EGF like factor
IFNγ	Interferon gamma
IL6	Interleukin 6
IPA	Ingenuity Pathway Analysis
ITPR	Inositol 1,4,5-Triphosphate Receptor
JAK	Janus Kinase
JUN	Jun oncogene
LOH	Loss of Heterozygosity
LS Mean	Least Squares Mean
MAPK	Mitogen Activated Protein Kinase
MEKK1	Mitogen-Activated Protein Kinase Kinase Kinase 1
MDM2	Murine Double Minute 2
MGMT	O-6-Methylguanine-DNA Methyltransferase
MLPA	Multiplex Ligation-dependent Probe Amplification
MTIC	Methyltriazeno-Imidazole-Carboxamide
MTOR	Mechanistic Target of Rapamycin (serine/threonine kinase)
MYC	v-myc Myelocytomatosis Viral Oncogene homolog
ODN	Decoy Oligonucleotides
OMS	Organização Mundial de Saúde

OSMR	Oncostatin M receptor
$p16^{INK4A}$	CDKN2A, Cyclin-Dependent Kinase inhibitor 2A
$p14^{ARF}$	Alternative reading frame of CDKN2A
PDGFRα	Platelet-Derived Growth Factor Receptor, alpha
PFS	Progression Free Survival
PI3K	Phosphoinositide 3-Kinase
PIAS	Protein Inhibitors of Activated STATs
PIP	Phosphatidylinositol
PIP3	Phosphatidylinositol-(3,4,5)-triphosphate
PLCG1	Phospholipase C, Gamma 1
PTEN	Phosphatase and Tensin Homologue
PTPRD	Protein Tyrosine Phosphatase Receptor-type D
RAS	v-Ha-ras Harvey Rat Sarcoma Viral Oncogene homolog
Rb	Retinoblastoma
RTK	Receptor Tyrosine Kinase
RNA	Ribonucleic Acid
RAF	v-raf-1 Murine Leukemia Viral Oncogene homolog 1
cRNA	Complementary Ribonucleic Acid
siRNA	Small Interfering RNA
shRNA	Small Hairpin RNA
SNC	Sistema Nervoso Central
SOCS	Supressor of Cytokine Signaling
SRC	v-src Sarcoma (Schmidt-Ruppin A-2) Viral Oncogene homolog (avian)
STAT	Signal Transducers and Activators of Transcription
TCGA	The Cancer Genome Atlas consortium
TGFα	Transforming Growth Factor, alpha
TKI	Tyrosine Kinase Inhibitor
TP53	Tumor Protein p53
WHO	World Health Organization

Abstract:

Central Nervous System (CNS) Neoplasms are characterized by their cell of origin and their histopathological features. Tumors of glial cell origin (Gliomas) are the most frequent, with Glioblastoma Multiforme (GBM) rising as the most common. GBM tumors of grade IV accordingly with the World Health Organization (WHO), are generally lethal, with a median survival time of 4.9 months, and their most striking histopathological features are the high degree of vascularization and necrosis.

The most common genetic alterations in GBM are the amplification, overexpression and mutation of the *EGFR* gene, and the deletion of the long arm of chromosome 10, where, among others, the *PTEN* gene is located. These genes are related, respectively, with the activation and inhibition of pathways like the MAPK cascade, the PIP-mediated signaling and STAT signaling which are involved in cellular proliferation and inhibition of apoptosis. Deregulation of these pathways renders them the logical target for inhibition of growth and proliferation of tumoral cells, and some anti-EGFR therapies have been tried, but with relatively poor success.

The goal of this work is to analyse the genetic expression of the genes that make up the EGFR-activated signaling pathways in gliomas, and identify those molecules where targeted intervention would make sense in such a way that cellular proliferation would cease, and differentiation and apoptosis would be induced.

Tumor samples (n=100) were characterized by Multiplex Ligation-dependent Amplification (MLPA) and Chromosomal Comparative Genomic Hybridization (cGGH). Further analysis of tumors samples (n=15) was done by using Gene Expression Arrays GeneChip® HuGene 1.0ST and data analysis softwares Partek Genomics Suite and Ingenuity Pathway Analysis, in order to determine the expression values of the various genes that make up the EGFR-activated signaling pathways. This allowed us to identify a particular pathway that appears to have its components constantly overexpressed, the STAT signaling pathway, mainly through the *STAT3* gene.

The STAT3 protein is activated by various receptors and is implicated in tumorigenesis and immune evasion, and as such, may be a suitable target for anti-neoplasic therapies.

Key-words: Glioma, EGFR, STAT3.

Resumo

Os tumores do Sistema Nervoso Central (SNC) são caracterizados pelo tipo de célula original e características histopatológicas. Representam cerca de 1,2% das neoplasias diagnosticadas em 2002 (Parkin *et al.*, 2005). Os tumores de origem glial (Gliomas) são os mais comuns, destacando-se entre eles o Glioblastoma Multiforme (GBM) que representa cerca de 15% das neoplasias intracranianas.

Os tumores GBM de grau IV segundo a Organização Mundial de Saúde (OMS), são geralmente letais, com um tempo médio de sobrevida de 4,9 meses, e são caracterizados histopatologicamente pelo seu elevado grau de vascularização e necrose (Louis *et al.*, 2007). Apesar de indistinguíveis a nível histopatológico, estas neoplasias podem ser subdivididas em GBM primário (pGBM) e secundário (sGBM). Estes grupos distinguem-se pelo processo de progressão e pelas características genéticas. Os pGBMs apresentam as características histopatológicas de GBM, aquando do seu diagnóstico inicial, enquanto que os sGBMs são caracterizados por uma progressão a partir de neoplasias de grau inferior (Ohgaki *et al.*, 2005).

As alterações genéticas mais comuns em tumores GBM (com diferentes incidências entre os subtipos) são o ganho de cópias do cromossoma 7 com amplificação em 7p12 [onde se situa, entre outros, o gene *EGFR* (Epidermal Growth Factor Receptor)], a deleção do braço longo do cromossoma 10 [onde se situa, entre outros, o gene *PTEN* (Phosphatase and Tensin Homologue)] e a mutação no gene *TP53* (Nicholas *et al.*, 2006).

Este tipo de alterações tem impacto directo na activação e regulação de várias vias de sinalização envolvidas no controlo da proliferação celular, no controlo do ciclo celular e na indução da apoptose. Destas vias, as mais relevantes na caracterização dos GBM são as vias relacionadas com os genes *Rb, TP53* e *EGFR* (The Cancer Genome Atlas (TCGA), 2008).

A via TP53/MDM2/p14^{ARF} é responsável pelo controlo da integridade do DNA antes da divisão celular. Em tumores GBM é comum haver alteração do gene *TP53* (com inibição da actividade da proteína TP53), o que permite a proliferação celular neoplásica. Além da alteração directa da proteína TP53, também é comum ocorrer a sobreexpressão de *MDM2* (inibidor de TP53) e a subexpressão de *p14^{ARF}* (inibidor de MDM2) (Ohgaki *et al.*, 2007).

A via associada ao gene *Rb* é responsável pelo controlo da progressão do ciclo celular de G1 para S. Em tumores GBM é comum a subexpressão do gene *Rb* e de *p16^{INK4A}* (ambos inibidores da progressão G1-S), o que favorece a progressão desregulada do ciclo celular (Ohgaki *et al.*, 2007).

O gene *EGFR* codifica um receptor com actividade cinásica de tirosinas (RTK – *Receptor Tyrosine Kinase*) da família ErbB, que é responsável pela transdução de sinal de vários factores de crescimento. A activação deste receptor pelos seus ligandos (e.g. EGF, HB-EGF e TGF-α, entre outros), provoca a homodimerização de EGFR ou a sua heterodimerização com outro membro da família ErbB, levando a uma autofosforilação de resíduos na extremidade –COOH da proteína, o que a torna capaz de fosforilar moléculas das vias de sinalização das MAP cinases, do Fosfatidilinositol-3,4,5-Fosfato (PIP3) e das STATs (Sebastian *et al.*, 2006). Quando activadas, estas vias levam a um aumento generalizado da capacidade de proliferação das células neoplásicas e, consequentemente, da tumorigénese. Em tumores GBM, é comum observar desregulação destas vias por sobreexpressão dos seus activadores (e.g. *EGFR*) e/ou por subexpressão dos seus reguladores (e.g. *PTEN*) (Ohgaki *et al.*, 2007).

O tratamento de tumores GBM consiste na remoção cirúrgica da área afectada, seguida de radioterapia. A temozolamida (um agente alquilante) (Stupp *et al.*, 2009), pode ser adicionada e, em certos casos, aumenta a sobrevida para 14,6 meses.

O reconhecimento da importância do EGFR no processo tumorigénico em gliomas, tornou-o o alvo óbvio na busca de novas terapias anti-neoplásicas e no aumento da sobrevida (Kuan *et al.*, 2001; Kari *et al.*, 2003). Com esse intuito, foram tentadas várias estratégias para impedir a activação das vias de sinalização activadas por EGFR (Halatsch *et al.*, 2006). Uma das estratégias baseia-se no impedimento físico do reconhecimento de ligandos e de dimerização do receptor e é representada pelo anticorpo monoclonal Cetuximab (Li *et al.*, 2005; Halatsch *et al.*, 2006), que tem, no entanto, demonstrado fracos resultados em ensaios clínicos (Neyns *et al.*, 2009). Uma outra estratégia baseia-se no bloqueio da actividade cinásica do receptor, utilizando inibidores moleculares da cinase de tirosinas (TKIs) (Halatsch *et al.*, 2006; Omuro *et al.*, 2007), que têm demonstrado igualmente fracos resultados na inibição da sinalização mediada pelo EGFR (Lassman *et al.*, 2005).

O facto de as vias de sinalização activadas pelo EGFR serem partilhadas por outras RTKs pode explicar, em parte, o fraco desempenho destas estratégias terapêuticas. O objectivo deste trabalho consiste em analisar as vias de sinalização activadas pelo EGFR com base na expressão génica dos seus constituintes, e identificar potenciais alvos nos quais uma intervenção dirigida faça sentido de modo a bloquear a proliferação celular e induzir apoptose e diferenciação nestas neoplasias.

Para alcançar este objectivo analisámos cerca de 100 gliomas com diferentes alterações genéticas previamente identificadas por *Chromosomal Comparative Genomic Hybridization* (cCGH) e avaliámos nestes casos a presença de possíveis mutações no gene *EGFR* por *Multiplex Ligation-dependent Probe Amplification*

(MLPA). Dentro deste grupo de gliomas, determinámos os perfis de expressão de 15 casos, usando os *Arrays* de expressão génica GeneChip® HuGene 1.0ST da Affymetrix, a partir de extractos de RNA total. Os perfis de expressão dos diferentes grupos de amostras foram depois comparados com um controlo de córtex cerebral normal usando o *software* Partek Genomics Suite. A utilização deste *software* permitiu-nos quantificar as diferenças de expressão génica entre os diferentes subgrupos de tumores e o córtex cerebral normal. Estes valores foram depois inseridos no *software* Ingenuity Pathway Analysis (IPA), que origina vias de sinalização canónicas e redes de interacções entre genes sobrepondo-lhes os valores de expressão de cada gene envolvido.

A partir dos dados obtidos pelo *software* IPA, foi possível determinar o padrão de expressão dos vários componentes das vias de sinalização activadas pelo EGFR nos vários subgrupos de tumores:

- 5 GBM com ganho de cromossoma 7 e amplificação em 7p12;
- 3 GBM com ganho de cromossoma 7;
- 1 Ganglioglioma Anaplásico com ganho de cromossoma 7 e respectiva linha celular;
- 2 Gliomas (um GBM e um Oligodendroglioma Anaplásico) com amplificação em 8q24;
- 1 GBM com ganho de cromossoma 7 e amplificação de 4q12;
- 3 Linhas celulares com diferentes alterações genéticas.

À excepção dos dois casos com amplificação em 8q24 (onde se situa, entre outros, o gene *c-Myc*), todos os outros tumores e linhas celulares demonstram uma sobreexpressão dos componentes da via STAT (*Signal Transducers and Activators of Transcription*). Nas vias de sinalização PIP e das MAP cinases não se verificam alterações substanciais que nos indiquem que estas possam ser preferencialmente utilizadas na sinalização activada por EGFR.

A via de sinalização STAT, particularmente a molécula STAT3, é directamente activada pelo receptor EGFR, por fosforilação de um resíduo de tirosina na posição 705, levando à homo- ou heterodimerização (com STAT1) e, consequentemente, a um aumento da actividade transcricional, mediada por STAT, de vários genes (Dauer *et al.*, 2005), entre os quais o gene *c-Myc*. Propomos que a activação preferencial da via STAT permita à célula tumoral aumentar a expressão de c-Myc, pelo que quando a sua expressão está aumentada por outras razões (como a que resulta da amplificação do seu *locus*) diminua a pressão selectiva para que a via STAT tenha os seus componentes sobreexpressos (como se verifica nos casos com amplificação de 8q24).

A presença de STAT3 activado, em gliomas, foi demonstrada (Abou-Ghazal *et al.*, 2008) e o seu papel na tumorigénese (Abou-Ghazal *et al.*, 2008) e na

imunosupressão (Kortyleswki *et al.*, 2008) foi realçado, tornando-a um alvo terapêutico óbvio, tendo já sido tentadas várias abordagens de inibição desta molécula, que incluem o uso de RNA de interferência (Konnikova *et al.*, 2003; Li *et al.*, 2009), oligonucleótidos inibitórios (Gu *et al.*, 2008), fosfopéptidos (Shao *et al.*, 2003) e inibidores moleculares (Fuh *et al.*, 2009).

O facto de STAT3 estar sobreexpresso na maioria dos casos analisados, da sua activação estar comprovada em gliomas de alto grau, de estar envolvido no processo tumorigénico e de imunossupressão, e de também poder ser activado por vários outras RTKs e não-RTKs que não o EGFR, leva-nos a propor que em gliomas a via STAT3 é a via de sinalização preferencialmente activada por EGFR, entre outros.

Palavras-chave: Glioma, EGFR, STAT3

Introduction

Central Nervous System (CNS) Neoplasms are a result of an abnormal growth of cells within the CNS, and can be of various types, as determined by their cell of origin and histopathological features. The World Health Organization (WHO) classifies and grades CNS neoplasms into four grades, Grade 1 being the less severe and Grade 4 the most severe (Louis *et al.,* 2007).

CNS Neoplasms account for 1.7% of all neoplasms that were diagnosed in 2002, and around 2.1% of cancer related fatalities (Parkin *et al.,* 2005). One of the most distinctive features of CNS neoplasms is their lethality and short survival period. The most common CNS tumor, Glioblastoma Multiforme (GBM), accounting for 12-15% of intracranial neoplasms, and 60-75% of astrocytic neoplasms, is a WHO Grade 4 glioma, with a median survival time of 4.9 months (Ohgaki *et al.*, 2005). This neoplasm is defined by a typically astrocytic differentiation, and its histopathological features include nuclear atypia, cellular pleomorphism, mitotic activity, vascular thrombosis, microvascular proliferation and necrosis (Louis *et al.,* 2007). It affects mainly adults and it emerges mostly in the white subcortical matter of the brain hemispheres (Louis *et al.,* 2007) (Figure 1).

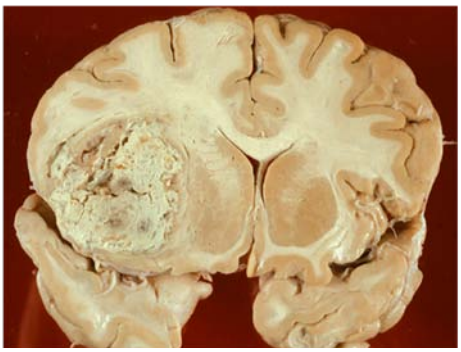

Figure 1. Glioblastoma Multiforme in a transverse cross section of the adult brain. (http://www.neuropathologyweb.org/chapter7/chapter7bGliomas.html)

GBM can be divided into two classes, Primary Glioblastoma (pGBM), accounting for roughly 95% of GBM cases, and Secondary Glioblastoma (sGBM), around 5% of GBM cases. Although they are histopathologically indistinguishable, they do have several differences. The median age of diagnosis of pGBM is 62 years, with a median survival time of 4.7 months, unlike sGBM, in which the median age of diagnosis is 45 years, with a median survival time of 7.8 months (mostly due to age difference between the groups) (Ohgaki *et al.,* 2005). Another distinction that can be made between pGBM and sGBM is the way they both progress. sGBM progresses from less severe lesions (like a low-grade astrocytoma), unlike pGBM, which is primarily diagnosed as such, without evidence of previous lesions. Also, their genetic modifications are somewhat different, *EGFR* (Epidermal Growth Factor Receptor) amplification and overexpression being most common in pGBM (40% and >60% in pGBM, 8% and <10% in sGBM), and *TP53* mutation appearing mostly in sGBM (28% in pGBM, 65% in sGBM). One feature that seems to be similar in both types of GBM is Loss of Heterozygosity (LOH) in 10q (70% of pGBM and 63% of sGBM) (Nicholas *et al.*, 2006) (Figure 2).

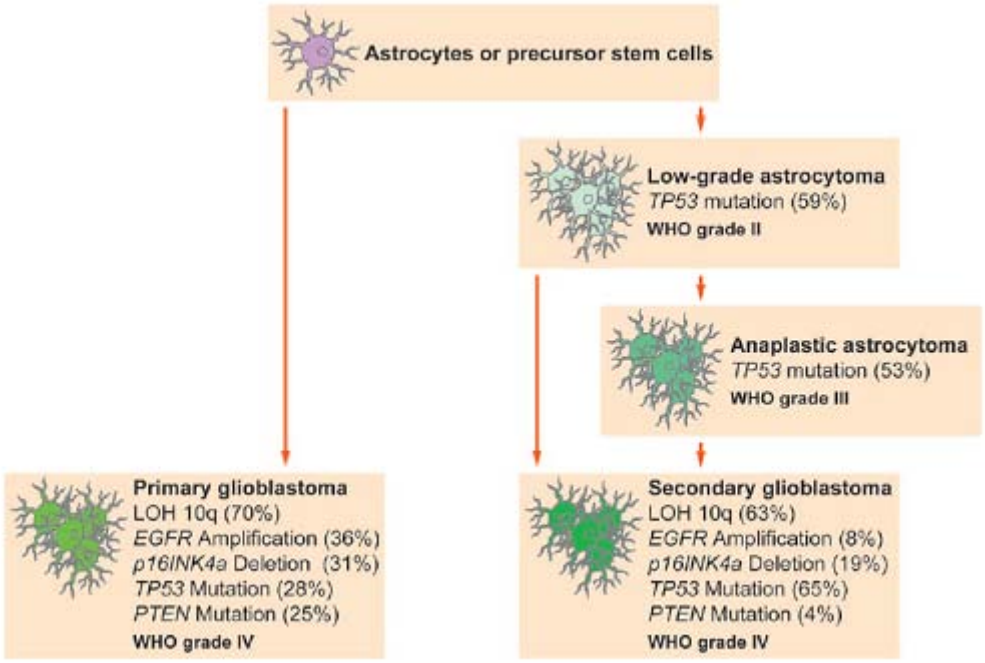

Figure 2. Differences between pGBM and sGBM (adapted from Ohgaki *et al.*, 2007).

Some of the most defining genetic modifications in GBM have direct impact on various signaling pathways. The most studied signaling pathways in GBM are related with the *TP53, Rb* and *EGFR* genes (Louis *et al.*, 2007; The Cancer Genome Atlas Research Network (TCGA), 2008).

The TP53/MDM2/p14^{ARF} pathway is responsible for maintaining DNA integrity prior to cell division. *TP53* (a tumor suppressor gene, coding for the p53 protein) is one of the most important and tightly regulated genes of the cell, having the ability to block cell division if there is any type of damage in the genome (Lodish *et al.*, 2008). In most sGBM (65%), *TP53* gene is altered (57% of mutations located in codons 248 and 273), this event occurring less frequently in pGBM (30%). Direct modifications of the *TP53* gene that disable its activity are enabling of neoplastic cell proliferation. Beside the direct modifications on the *TP53* gene, GBM tumors also have other mechanisms of disabling its activity. For instance, *MDM2* (Murine Double Minute 2, an inhibitor of p53) is amplified in 31% of pGBM, which leads to increased capacity of the cell to inhibit p53, and hence, increased capacity of cell division. There is yet another way of the cell to decrease p53 activity, and this is achieved by deregulation of the MDM2 protein inhibitor, p14^{ARF}. The *p14^{ARF}* gene is underexpressed in 76% of GBM (either by homozygous deletion or promoter methylation) (Ohgaki *et al.*, 2007).

Thus, the neoplastic cell has various features enabling blockage p53 regulation, either by modification of the *TP53* gene (mainly in sGBM) or by overexpressing of its inhibitor *MDM2* (mainly pGBM), or even by underexpressing MDM2's inhibitor, p14^{ARF}.

The Rb1 (Retinoblastoma) protein regulates progression through the cell cycle, mainly the G1-S transition. In a normal cell, a CDK4/ciclinD1 complex phosphorylates the

Rb1 protein, allowing the release of the E2F transcription factor which regulates several genes responsible for the G1-S transition. $p16^{INK4A}$ is the inhibitor of CDK4, and therefore, the inhibitor of the progression through the cell cycle (Lodish *et al.*, 2008). In GBM this pathway is usually deregulated by either the deregulation of the *p16*^{INK4A} gene, the underexpression of the *RB1* gene, or by both events. The *p16*^{INK4A} gene is deregulated by either homozygous deletion, or by promoter methylation, and is most common in pGBM. *RB1* gene underexpression is achieved in GBM by promoter methylation, an event which occurs in 43% of sGBM and 14% of pGBM (Ohgaki *et al.*, 2007).

In a similar way to the p53 deregulation described above, Rb1 deregulation may come by either direct inhibition of the *RB1* gene, or by modification of other members of the pathway that inhibit cell cycle progression.

One of the most defining features of GBM is the amplification, overexpression and mutation of the *EGFR* gene, and the prevalent activation of the *EGFR* family (TCGA, 2008). This gene occupies roughly 200 kb of the 7p12 region in chromosome 7 [which is amplified in approximately 40% of GBM (Olson *et al.,* 2009)] and is composed by 28 exons (Reiter *et al.*, 2001). A common feature in GBM is that *EGFR* amplification is commonly accompanied by various mutations. The most common mutation is *EGFRvIII* [accounting for roughly 60% of mutations (Nicholas *et al.*, 2006)] present in 54% of cases with amplification of the *EGFR* gene (Olson *et al.,* 2009). This mutation is defined by an *inframe* deletion of exons 2 through 7, leading to a truncated protein lacking the extracellular binding domain. This truncated version of the EGFR is constitutively active, leading to increased activation of downstream signaling pathways (Ohgaki *et al.*, 2007).

The EGFR protein is a receptor tyrosine kinase (RTK) of the ErbB family. This receptor can directly translate extracellular stimulus (by binding of the EGFR ligands, like EGF, HB-EGF, TGFα and others, and dimerization with another member of the ErbB family), into phosphorylation of downstream targets of the Ras/MAPK cascade signaling pathway, the PIP signaling pathway and the STAT signaling pathway (Oda *et al.*, 2005; Sebastian *et al.,* 2006) (Figure 3).

The mitogen-activated protein kinase (MAPK) cascade is involved in a variety of cellular functions such as growth, proliferation, differentiation, migration and apoptosis (Lodish *et al.*, 2008; Zohrabian *et al.*, 2009). These pathways rely on the sequential phosphorylation of several proteins, beginning in the activation of Ras [a known oncogene, mutated in 2% of GBM (TCGA, 2008)] and ending in the activation of various transcription factors (c-Jun and c-Fos) that regulate expression of cell-cycle progression and differentiation related proteins (elegantly reviewed in Lodish *et al.*, 2008 Chapters 14.3 and 14.4).

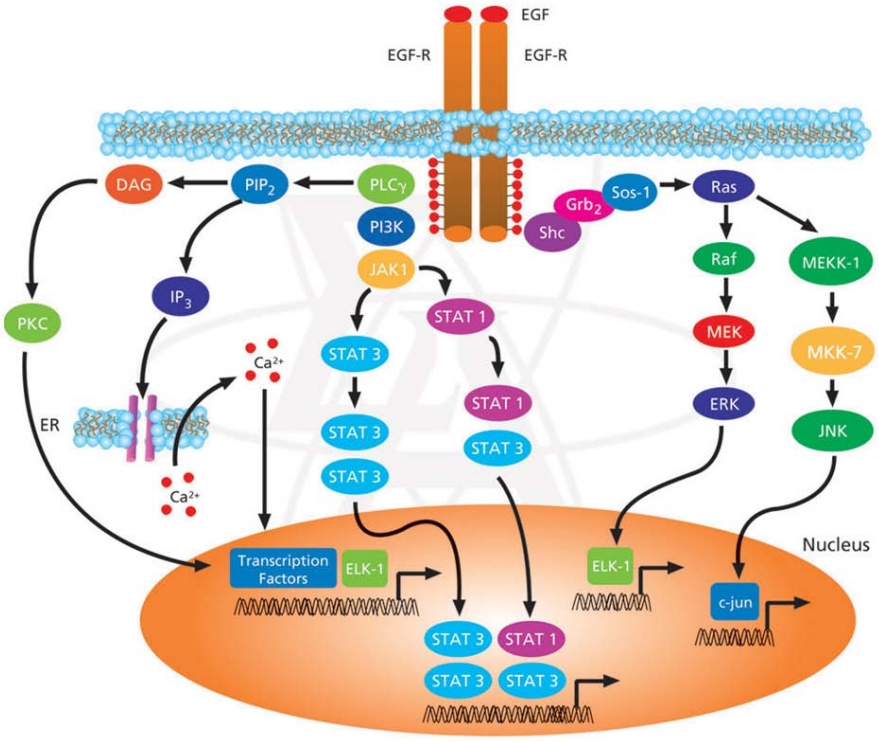

Figure 3. EGF receptor mediated signaling pathway overview (www.sigmaaldrich.com).

Phosphatidylinositol (PIP) signaling is one of the most important pathways in the cell. It regulates multiple biological events such as apoptosis, metabolism, cell proliferation and cell growth (Blanco-Aparicio *et al.*, 2007). RTKs can signal through PIP either by activation of the PIP3/DAG or PI3K pathways. Both these pathways lead to activation of Akt or mTOR proteins and thus to cell proliferation (as reviewed in Lodish *et al.*, 2008 Chapter 14.5, and in Blanco-Aparicio *et al.*, 2007). Regulation of the PI3K pathway is accomplished by the PTEN (Phosphatase and Tensin homolog) protein. The *PTEN* gene is located on the long arm of chromosome 10, which loss is a common event in GBM (70% of pGBM and 63% of sGBM). In a normal cell PTEN removes phosphate from PIP3 (a key molecule in PIP signaling), inhibiting PIP downstream signaling. If PTEN is lost, PIP signaling can occur almost constitutively, promoting cell survival and proliferation (Lodish *et al.*, 2008).

STAT (Signal Transducer and Activator of Transcription) proteins are a family of transcription factors that are activated (phosphorylated) as a response to extracellular stimulus of various receptors like RTKs, cytokine receptors (like IL6R, OSMR) and others (Caló *et al.*, 2003; Lodish *et al.*, 2008 Chapter 14.2). Activation of the STAT transcription factors can lead to events such as differentiation, proliferation, cell survival, apoptosis and angiogenesis. STAT proteins can be divided into two groups, according to their specific functions. One group, made up of STAT2, STAT4 and STAT6, is activated by cytokines and plays a distinct role in T-Cell differentiation and IFNγ signaling (Caló *et al.*, 2003). The other group, made up of STAT1, STAT3 and STAT5, is activated by various ligands, and plays a role in controlling cell-cycle progression and apoptosis (Bromberg, 2002).

STAT phosphorylation can be mediated either by JAK (Janus Kinase) (Wilks, 2008), Non-RTK proteins like Src (Silva, 2004) or directly by RTKs like EGFR (Coffer *et al.*, 1995; Ihle, 1996). STAT signaling is regulated by the SOCS proteins (Supressor of cytokine signaling), PIAS proteins (Protein Inhibitor of Activated STATs) and Protein tyrosine phosphatases (Brantley *et al.*, 2008a). SOCS proteins attenuate STAT signaling through inhibition of upstream JAK activation, in a classical feedback loop. PIAS are constitutively expressed proteins that mediate transcriptional repression by interfering with STAT transcription factors ability to bind DNA. Protein tyrosine phosphatases, like PTPRD (Protein Tyrosine Phophatase Receptor-type D) (Veeriah *et al.*, 2009) inhibit STAT signaling by removing phosphate from STAT proteins, rendering them inactive.

STAT3 activation has been linked to GBM (Brantley *et al.*, 2008a), either by constitutive activation of STAT3 protein or by loss of its inhibitors, like PIAS-3 (Brantley *et al.*, 2008b). STAT3 regulates transcription of several genes, like *c-Myc* (Dauer *et al.*, 2005; Bowman *et al*, 2009), *EGFR* (Dauer *et al.*, 2005) and other genes like *Bcl-2, Bcl-x* and *mcl-1,* which are related to inhibition of apoptosis (Rahaman *et al.*, 2002).

The treatment of GBM is based heavily upon radiation therapy and adjuvant temozolomide followed by surgical resection of the tumor mass increasing progression-free survival (PFS) to 14.6 months (Stupp *et al.*, 2005; Olson *et al.*, 2009).

Temozolomide is a cytostatic prodrug that is rapidly absorbed by oral administration, is able to cross the blood-brain barrier and is spontaneously hydrolyzed to the active metabolite methyltriazeno-imidazole-carboxamide (MTIC). In the cell, MTIC methylates DNA at several positions, of which methylation at the O6 posititon of a guanine is regarded as fatal for the cell (Friedman *et al*., 2000). The methyl group at O6 can be removed by the DNA repair enzyme MGMT. Methylation of the *MGMT* gene promoter in GBM has been linked with an increase of PFS to 21.7 months (Hegi *et al.*, 2005).

Altough improved with the use of chemotherapy and adjuvant temozolomide, prognosis of GBM still remains dismal. Alternative therapies have been attempted with various results (Omuro *et al.,* 2007). Some of these therapies target the EGFR protein. As mentioned previously, *EGFR* overexpression and amplification, in league with the EGFR protein's ability to activate various pathways related to oncogenesis, make it a natural target for cancer therapy (Kuan *et al.*, 2001; Kari *et al.*, 2003).

Attempts of targeting EGFR as a means to induce tumor regression have been made, mainly by two approaches (Halatsch *et al.*, 2006).

The first approach is based upon the inhibition of the receptors ability to either bind its ligands or to dimerize. This type of inhibition may be accomplished by the use of antibodies like Cetuximab and other EGFR specific antibodies (Halatsch *et al.*, 2006). Cetuximab physically inhibits binding of EGFR ligands and also prevents EGFR conformational

modifications that allow it to dimerize. (Li *et al.*, 2005). Cetuximab has had promising results in animal models (Martens *et al.*, 2008) but as shown limited potential in phase-II clinical trials (Neyns *et al.*,2009).

The second approach is based upon the inhibition of the receptors ability to phosphorylate downstream targets, that is, inhibition of its Tyrosine Kinase domain. This may be accomplished by Tyrosine Kinase Inhibitors (TKIs) such as Erlotinib and Gefitinib (Halatsch *et al.*, 2006). Both Erlotinib and Gefitinib inhibit the TK domain of EGFR by selectively binding its ATP-binding position, thus rendering it inactive (Omuro *et al.,* 2007). Regrettably, these compounds have failed at showing significant inhibition of EGFR-mediated signaling in GBM (Lassman *et al.*, 2005). An effort has been made to explain why TKIs are unsuccessful. Conflicting theories have come up in regard to EGFR copy number and EGFRvIII status as markers of sensitivity (Mellinghoff *et al.*, 2005) or resistance (Learn *et al.*, 2004). Also, successful inhibition of EGFR may only diminish downstream signaling pathways, because EGFR shares its signaling pathways with many other RTKs and non-RTKs. Also, *PTEN* inhibition as been suggested as a marker of increased tumor resistance to TKIs (Mellinghoff *et al.*, 2005; Guillamo *et al.*, 2009).

The dismal prognosis of GBM, allied with the yet relative failure at therapies based upon EGFR direct targeting, lead us to suggest that GBM treatment should by targeted at intra-cellular points in the EGFR activated pathways. Therefore, the goal of this project is to study the signaling pathways activated by EGFR, and by means of gene expression analysis, understand if there is any genetic preference to a particular pathway or molecule of a pathway, in which targeted therapy would be a valid option.

Materials and Methods

DNA Extraction for Multiplex Ligation-dependent Probe Amplification (MLPA) and Chromosomal Comparative Genomic Hybridization (cCGH)

DNA extraction from frozen glioma tumor samples (n=100) was performed with the Proteinase K and Phenol protocol for isolation of High-molecular-weight DNA (Sambrook *et al.*, 2001), concentration and purity were determined with NanoDrop ND-1000 spectrophotometer (*NanoDrop* Technologies, Wilmington, DE, USA). For MLPA, DNA was diluted in Tris-EDTA 1X buffer pH 8.0. For cCGH, DNA was diluted in Tris buffer 1X pH 8.0.

Chromosomal Comparative Genomic Hybridization

cCGH analysis was performed from frozen material at the Cytogenetics laboratory – CIPM of the Lisbon Portuguese Cancer Institute. cCGH was performed according to the method of Kallioniemi (Kallioniemi *et al*., 1994). Briefly, tumor DNA was labeled with biotin 16dUTP (Enzo-Roche), and normal reference DNA with digoxigenin-11-dUTP (Enzo-Roche), in a standard nick translation reaction. Equal amounts (400ng) of labeled tumor DNA and labeled reference DNA were coprecipitated with 15 µg of Cot-1-DNA (Invitrogen) in ethanol. After a 3-day hybridization period, fluorescent detection ot the biotin- and digoxigenin-labeled DNAs was accomplished by using avidin-FITC (Jackson Immunoresearch) and antidigoxigenin rhodamine (Enzo-Roche) antibodies, respectively. Samples were counterstained in DAPI in antifade solution (Vector). For image acquisition, an epifluorescent microscope (Zeiss Axioplan II) equipped with a cooled CCD Camera (Photomic Science) and a triple-band beam splitter and emission filters (Chroma Technology, USA) were used. For each tumor, three color images (blue, red and green) were acquired from at least 10 metaphases. Image analysis was performed using the cCGH analysis software from CytoVision System (version 2.51 Applied Imaging, Sunderland Tyne & Wear, UK). A simplified overview of this protocol is given in Supplementary Figure S1.

Multiplex Ligation-dependent Probe Amplification (MLPA)

For EGFR mutational analysis SALSA P315 Kit for EGFR from MRC-Holland was used (MRC-Holland, Amsterdam, The Netherlands). This kit includes probes for each of the 28 exons of the EGFR gene. As a control, this kit also includes 9 probes for different chromosomal positions (2q37.3, 3q11.2, 5p13.2, 5q35.3, 7q21.1, 9q31.1, 10q23.3, 12q23 and 17p13.1). MLPA Protocol was carried out as previously described (Schouten *et al.*, 2002), using 100ng of DNA for each normal control and tumor sample. 3 µL of the amplified sample product was analyzed using ABI 3130 Genetic Analyzer and as an internal size standard the Genescan 500 LIZ (ABI 4322682). Successful ligation reaction and identification of samples with insufficient amount of DNA were verified using MLPA's internal

ligation-independent probes. Data analysis was carried out with the MRC-Coffalyser version 8.0. Ratios were calculated using the Tumor Analysis LS method in MRC-Coffalyser V8. Tumor Analysis LS is a direct analysis method which performs slope correction in all probes and then compares each individual probe peak height in a sample (which is related to the relative quantity of the probe target in the sample) to its counterpart in a normal reference control, establishing ratios between them: Homozygous loss of a particular probe was considered when ratios were 0; Heterozygous loss was considered when ratios were between 0,3 and 0,7; Normal status was considered when ratios were between 0,7 and 1,3; Gain was considered when ratios were between 1,3 and 1,7; Amplification was considered if ratios were above 2,5. A simplified overview of this protocol is given in Supplementary Figure S2.

Microarray Analysis

Human cerebral cortex total RNA was purchased from Clontech (Mountain View, CA, USA), consists of a pool of cerebral cortex RNA obtained from 10 individuals who died from sudden death, and was used as normal baseline reference for microarray experiments.

Total RNA was extracted from tumor samples, using the RNeasy Lipid Tissue Mini Kit (Qiagen, Hilden, Germany). Concentration and purity was determined with NanoDrop ND-1000 spectrophotometer (*NanoDrop* Technologies, Wilmington, DE, USA), and integrity was confirmed using an Agilent 2100 Bioanalyzer with a RNA 6000 Nano Assay (Agilent Technologies, Palo Alto, CA, USA). Samples were selected if they had more than 100 ng of RNA and a RNA Integrity Number (RIN) larger than 7. RNA was processed for use on Affymetrix (Santa Clara, CA, USA) GeneChip® HuGene 1.0 ST Arrays, according to the manufacturer's Whole Transcript Sense Target Labeling Assay, at the Affymetrix Core Facility located in the Gulbenkian Institute of Science. Briefly, 100 ng of total RNA containing spiked in Poly-A RNA controls (GeneChip® Expression GeneChip® Eukaryotic Poly-A RNA Control Kit; Affymetrix) were used in a reverse transcription reaction (GeneChip® WT cDNA Synthesis Kit; Affymetrix) to generate first-strand cDNA. After second-strand synthesis, double-stranded cDNA was used in an in vitro transcription (IVT) reaction to generate cRNA (GeneChip® WT cDNA Amplification Kit; Affymetrix). 10 µg of this cRNA were used for a second cycle of first-strand cDNA synthesis (GeneChip® WT cDNA Synthesis Kit; Affymetrix). 5.5 µg of single stranded cDNA were fragmented and end-labeled (GeneChip® WT Terminal Labeling Kit; Affymetrix). Size distribution of the fragmented and end-labeled cDNA, respectively, was assessed using an Agilent 2100 Bioanalyzer with a RNA 6000 Nano Assay. 5 µg of end-labeled, fragmented cDNA were used in a 100-µl hybridization cocktail containing hybridization controls. 80 µl of mixture were hybridized on arrays for 17h at 45°C. Standard post hybridization wash and double-stain protocols (FS450_0007; GeneChip®

HWS kit, Affymetrix) were used on an Affymetrix GeneChip® Fluidics Station 450. Arrays were scanned on an Affymetrix GeneChip® scanner 3000 7G. A simplified overview of this protocol is given in Supplementary Figure S3.

Gene Expression Analysis

RMA Background Correction, Quantile Normalization, Log2 Transformation and Median Polish Summarization were performed on the initial microarray data with the Partek Genomics Suite (GS) software (Partek Inc. St. Louis, MO). To determine gene expression ratios and fold-change between the various samples (Table 2) and the normal cerebral brain cortex control, one-way ANOVA statistical analysis was performed with Partek GS. ANOVA, or Analysis of Variance, is a parametric test which is used to test differences in means of a response variable between different groups. Also, in the context of the ANOVA analysis, linear contrasts between the different groups were performed, yielding gene expression Ratios and Fold-Change values (FC). Ratios were calculated by the ratio between the Least Squares Mean (LS Mean, calculated as the linear combination of the estimated means generated by the ANOVA model) of the different groups. Fold-change was determined in a similar fashion. Genes identified with a FC over 2 (e.g. overexpressed in sample A *vs* B) and genes with a FC below -2 (e.g. underexpressed in sample A *vs* B) were then selected for gene ontology analysis and EGFR pathway construction

Gene ontology, canonical pathways, and functional network analysis

Gene ontology, canonical pathways and functional network analysis were performed with the Ingenuity Pathway Analysis software (Ingenuity Systems, Redwood City, CA), which enables the discovery, visualization and exploration of molecular interaction networks in gene expression data. The gene lists identified by ANOVA, containing Genebank accession numbers as clone identifiers as well as FC values, were uploaded into the Ingenuity Pathway Analysis software. Each clone identifier was mapped to its corresponding gene object in the Ingenuity pathway knowledge base. These so-called focus genes were then used as a starting point for generating biological networks. A score was computed for each network according to the fit of the original set of significant genes. This score reflects the negative logarithm of the *P* that indicates the likelihood of focus genes in a network being found together due to random chance. Significance for biological functions were then assigned to each network by determining a *P* for the enrichment of the genes in the network for such functions or pathways compared with the whole Ingenuity pathway knowledge base as a reference set. Right-tailed Fisher's exact test was used with $\alpha=0.05$. The same statistical approach was used for gene ontology analysis of the initial gene lists. Molecules in networks and canonical pathways were given a color code directly related with their expression value.

A molecule is light-green if its expression value is two-fold smaller than in the control (which means that particular molecule is underexpressed in the sample set). A molecule is light-red if its expression value is two-fold larger than in the control (which means that molecule is overexpressed in the sample set). The darker the tone of green or red, the more that molecule is respectively under- or overexpressed. A molecule is colored grey if its expression value is between these two limits. A molecule is colored white if no information is available in regard to its expression. A list with Network shapes (Supplementary Figure S4) and Relationship types (Supplementary Figure S5) is given in the Supplementary Data section.

EGFR Signaling Pathway:

Along with the IPA EGFR signaling pathways, an EGFR Signaling Pathway adapted from the work of Oda *et al.* (Oda *et al.*, 2005) was also constructed. This group constructed a comprehensive pathway for EGFR-mediated signaling based on published scientific papers, namely EGFR endocytosis followed by its degradation or recycling, small guanosine triphosphatase (GTP-ase)-mediated signal transduction such as mitogen-activated protein kinase (MAPK) cascade, phosphatidylinositol polyphosphate (PIP) signaling, cell cycle, and G Protein-coupled receptor (GPCR)-mediated transactivation via intracellular Ca^{2+} signaling. Oda's *et al.* map was created using CellDesigner (http://celldesigner.org), a software package that enables users to describe molecular interactions using a well-defined and consistent graphical notation. The data of molecular interactions are stored in Systems Biology Markup Language (SBML; http://sbml.org/). They are based on the molecular interactions documented in 242 papers accessible from PubMed (http://www.ncbi.nlm.nih.gov). It comprises 211 reactions (131 state transitions, 34 transportations, 32 associations, 11 dissociations, 2 truncations and 1 unknown transition) and 322 species (202 proteins, 3 ions, 21 simple molecules, 73 oligomers, 7 genes and 7 mRNAs). A 'species' is defined by SBML as 'an entity that takes part in reactions' and it is used to distinguish the different states that are caused by enzymatic modification, association, dissociation, and translocation. In our EGFR Signaling Pathway adapted from Oda's *et al.* work, we included the GTP-ase mediated signal transduction (MAPK cascade), the PIP signaling and also STAT signaling. Like in the Ingenuity Pathway Analysis, we also colored the individual molecules in a direct relation with their expression values. So, molecules colored red had a two-fold expression value above the normal control (overexpressed), green colored molecules had a two-fold expression value below the normal control (underexpressed), and as in the IPA pathways the darker the tone of green or red, the more that molecule is respectively under- or overexpressed. White colored molecules had expression values between these limits. Molecules colored white and with *italic* lettering have no associated gene expression data.

Results

Multiplex Ligation-dependent Probe Amplification (MLPA)

To determine the status of the *EGFR* gene, in regard to its copy number and possible presence of the *EGFRvIII* mutation, we analyzed 98 cases with previously diagnosed CNS tumors with the MLPA Salsa P315 Kit for *EGFR*. MLPA was theoretically suitable for this task because it can be both a qualitative and quantitative method (Schouten *et al.*, 2002). The SALSA P315 Kit for *EGFR* has 28 probes (one for each exon of the gene) and 9 control probes for other genomic regions. This kit enabled us to quantify each individual exon in a given sample, when compared to a set of normal control samples. The resulting ratios for each sample would then allow us to determine if a particular exon is over- or underrepresented, and if the complete set had a particular pattern of representation.

By MLPA and as show in Table 1 and Figure 4 (page 12), *EGFR* amplification was found in 43% of GBM, 42% of high-grade Oligodendrogliomas (Oligodendrogliomas grades IV and III, and Astrocytoma grade IV) and 9% of low-grade Oligodendroglioma (Oligodendroglioma grade II) (this is a particular case, in which the only amplification detected was also the only case of *EGFRvIII* detection in the low-grade Oligodendroglioma group). *EGFRvIII* mutation was detected in 8% of GBM, 5% of high-grade Oligodendroglioma and 9% of low-grade Oligodendroglioma. *EGFR* gain was detected in 10% of GBM, 24% of high-grade Oligodendroglioma and 9% of low-grade Oligodendroglioma. No change of *EGFR* status was detected in 39% of GBM, 29% of high-grade Oligodendroglioma and 82% of low-grade Oligodendroglioma.

Table 1. *EGFR* gene status analysis in CNS Tumors by MLPA.

Hystological Type	***EGFR* Amplification**	***EGFR* Amplification and *EGFRvIII* Mutation**	***EGFR* Gain**	**No change**	**Total**
Glioblastoma Multiforme (GBM)	21 cases (43%)	4 cases (8%)	5 cases (10%)	19 cases (39%)	49 cases
Oligodendroglioma Grades IV and III + Astrocytoma Grade IV	16 cases (42%)	2 cases (5%)	9 cases(24%)	11 cases (29%)	38 cases
Oligodendroglioma Grade II	0 cases (0%)	1 case (9%)	1 case (9%)	9 cases (82%)	11 cases
Total	37 cases (38%)	7 cases (7%)	15 cases (15%)	39 cases (40%)	98 cases

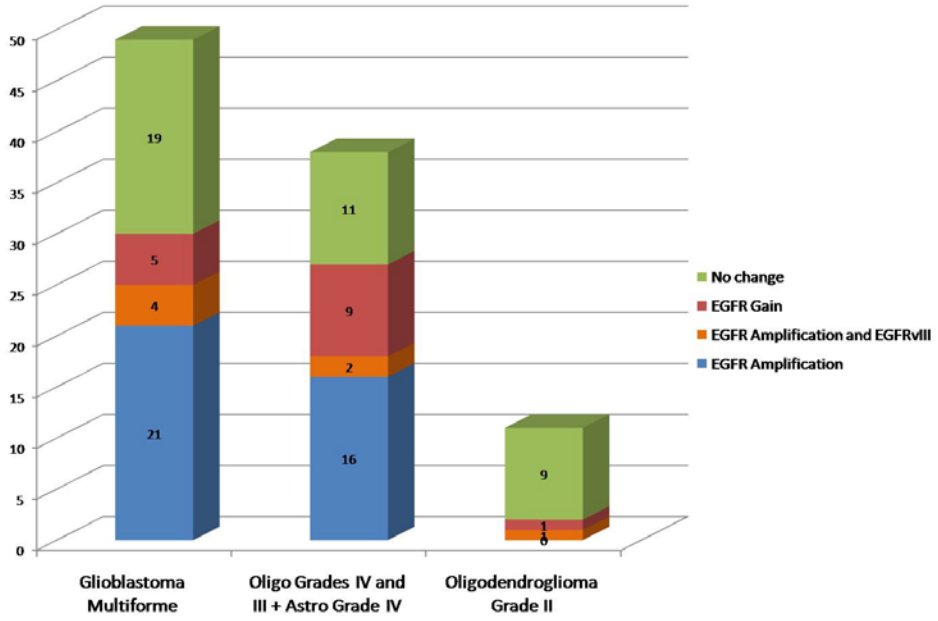

Figure 4. *EGFR* gene status analysis in CNS Tumors by MLPA. Blue bars represent *EGFR* Amplification, Orange bars represent *EGFR* Amplification and simultaneous detection of *EGFRvIII* mutation, Red bars represent *EGFR* Gain and Green bars represent no change in *EGFR* status. Numbers indicate the number of samples related with the particular *EGFR* status.

EGFR signaling pathways gene expression analysis

To understand how the EGFR signaling pathways behave in GBM, we selected a group of 15 samples (previously analyzed by cCGH) that was broad enough to describe the most common genetic events in GBM (Table 2, page 13). The first group of samples (cases 1, 2, 3 ,4 and 5; Group 1) was composed by cases with GBM that were characterized by cCGH as having (among other changes) gain of chromosome 7 and amplification of 7p12 (amp7p12). The second group (cases 6, 7 and 8; Group 2) was composed by cases with GBM that were characterized by cCGH as having (among other changes) gain of chromosome 7. The third group (cases 9 and 10; Group 3) was composed by one case of GBM and another case of AO that were characterized by cCGH as having (among other changes) amplification of 8q24. In one case (case 11) we obtained RNA for both the cell line and tumor, which allowed us to determine if the EGFR signaling pathway profile was maintained between the tumor and the cell line. One case (case 12), was characterized by cCGH as having (among other changes) gain of chromosome 7 and amplification in 4q12. There were also in our group of samples different individual cell lines derived from primary tumors that had unique changes in their cCGH profiles (cases 13, 14 and 15). These samples were included to determine if different genomic modifications would produce different EGFR signaling pathways profiles. Also, due to sample size limitation, we sought to statistically validate our results by comparing them with the publicly available data of Group 1 (79 cases) and Group 2 (106 cases) tumor types derived from the work of "The Cancer Genome Atlas" (TCGA, 2008), of which we used a total of 185 tumor cases.

Table 2. Samples used in Gene Expression Analysis.

Case	Type	Sex	Hystological Type	Chromosomal Gains (determined by cCGH)
1	Tumor	F	Glioblastoma Multiforme	**Total Gain of Chromosome 7 with Amplification of 7p12**
2	Tumor	F	Glioblastoma Multiforme	**Total Gain of Chromosome 7 with Amplification of 7p12**
3	Tumor	M	Glioblastoma Multiforme	**Total Gain of Chromosome 7 with Amplification of 7p12**
4	Tumor	F	Glioblastoma Multiforme	**Total Gain of Chromosome 7 with Amplification of 7p12**
5	Tumor	M	Glioblastoma Multiforme	**Total Gain of Chromosome 7 with Amplification of 7p12**
6	Tumor	M	Glioblastoma Multiforme	**Total Gain of Chromosome 7**
7	Tumor	M	Glioblastoma Multiforme	**Total Gain of Chromosome 7**
8	Tumor	M	Glioblastoma Multiforme	**Total Gain of Chromosome 7**
9	Tumor	M	Glioblastoma Multiforme	**DNA Amplification of 8q24.2**
10	Tumor	M	Anaplastic Oligodendroglioma	**DNA Amplification of 8q24.1**
11	Cell Line	M	Anaplastic Ganglioglioma	**Total Gain of Chromosome 7**
	Tumor			**Total Gain of Chromosome 7**
12	Tumor	M	Glioblastoma Multiforme	**Total Gain of Chromosome 7 and Amplification of 4q12**
13	Cell Line	M	Glioblastoma Multiforme	**Total Gain of Chromosome 7**
14	Cell Line	F	Glioblastoma Multiforme	**DNA Amplification of 3q, 4p, 18q and Xq**
15	Cell Line	F	Glioblastoma Multiforme	**Total Gain of Chromosome 7 and Amplification of 3q, 4q and 12q**

Note: In one case (11), RNA for tumor and cell line was available for microarray analysis.

To understand how the molecules in the EGFR signaling pathways were expressed in our samples, we used Affymetrix's GeneChip® HuGene 1.0 ST gene expression microarrays that allow us to determine the expression values of 28,869 genes. Having the data for each individual sample, we then compared the subgroups with the normal brain cerebral cortex control. This comparison was made with Partek Genomics Suite's ANOVA, which made it possible to generate gene expression Ratios and FC between the subgroups and the normal control. Gene lists for each comparison were made, and only genes with a FC over 2 or below -2 were selected for further analysis. Gene lists filtered by FC were then inputted into the IPA software, generating networks and canonical pathways, of which the EGFR Signaling pathway was of particular interest (Table 3, page 14).

Table 3. EGFR Canonical Pathway in the various samples groups (Ingenuity Pathway Analysis).

Cases	**Genes Overexpressed (FC ≥ 2)**	**Genes Underexpressed (FC ≤ -2)**	**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Group 1 *vs* N	7/48(14.5%)	4/48 (8.4%)	37/46 (77.1%)
Case 3 (*EGFRvIII vs* N)	8/48 (16.6%)	4/48 (8.4%)	36/48 (75%)
Group 2 *vs* N	7/48 (14.5%)	6/48 (12.5%)	35/46 (73%)
Group 1 *vs* Group 2	2/48 (4%)	0/48 (0%)	46/48 (96%)
Group 1 *vs* Group 2 (TCGA data)	1/48 (2%)	0/48 (0%)	47/48 (98%)
Group 3 *vs* N	5/48 (10.5%)	5/48 (10.5%)	38/48 (79%)
Case 11 - Cell Line *vs* N	6/48 (12.5%)	11/48 (23%)	31/48 (64.5%)
Case 11 – Tumor *vs* N	5/48 (10.5%)	7/48 (14.5%)	36/48 (75%)
Case 12 *vs* N	7/48 (14.5%)	8/48 (16.6%)	33/48 (68.9%)
Case 13 *vs* N	6/48 (12.5%)	11/48 (23%)	31/48 (64.5%)
Case 14 *vs* N	5/48 (10.5%)	13/48 (27%)	30/48 (62.5%)
Case 15 *vs* N	8/48 (16.6%)	11/48 (23%)	29/48 (60.4%)

N = Normal Cerebral Cortex

Tumors with Gain of Chromosome 7 and Amplification of 7p12 *vs* N

When comparing expression data of Group 1 *vs* N in our series as depicted in Figure 5 (page 15) and Table 4 (page 16), in what concerns EGFR canonical pathway, several findings drew our attention. The most striking was the degree of overexpression of the *EGFR* gene with a fold-change of 10.66 over the normal control (p-value 0.111). Analysis of EGFR signaling revealed different profiles in the various intracellular transducer branches. Accordingly, in the MAPK signaling cascade, we verified that transducers molecules did not have a distinctive uniform profile of gene expression. That is, molecules in this pathway were either underexpressed (e.g. *MKK4,* FC = -3.07), had no change in expression (e.g. *hRas,* FC = -1.07) or were overexpressed (e.g. *MEKK1,* FC = 2.27) without an evident pattern that could indicate a preferential activation of this pathway branch. In the PIP signaling pathway we saw no changes in expression, except for the underexpression of one of the genes that constitute the ITPR (e.g. *ITPR1,* FC = -5.54).

In the STAT signaling pathway we saw an overexpression of both *STAT3* and *STAT1* genes (with an average FC of 2.1 for *STAT1* and *STAT3* when compared with normal cerebral cortex, with a p-value of 0.168 for *STAT1* and 0.015 for *STAT3*). Although through analysis of figure 5 there is seemingly no direct interaction between the STAT proteins and

the EGFR protein, it is known that EGFR is able to directly activate these proteins (Coffer *et al.*, 1995) without need of interaction with the JAK proteins.

The combined results from the three different EGFR activated signaling transducers branches lead us to suggest that there is an apparent genetic preference for EGFR signaling to proceed through the STAT signaling pathway in this group of cases.

In one case, *EGFRvIII* mutation was detected by the MLPA technique (case 3), and we also proceeded to analysis of the EGFR signaling pathway in this case. In similarity with other cases with amp7p12, an apparent genetic expression preference for EGFR signaling to proceed through STAT signaling was also found (Supplementary Figure S6 and Table S1).

Although the IPA canonical pathways were very useful in ascertaining if there was a genetic preference for signaling in GBM with amp7p12, it was clear that the IPA canonical pathway for EGFR signaling was incomplete. We therefore used a more detailed EGFR pathway (Oda *et al.*, 2005) depicted in Figure 6 (page 16). The results obtained were similar to those obtained with the IPA pathway (Figure 5), with evidence of a genetic preference for EGFR signaling to proceed through the STAT proteins becoming apparent.

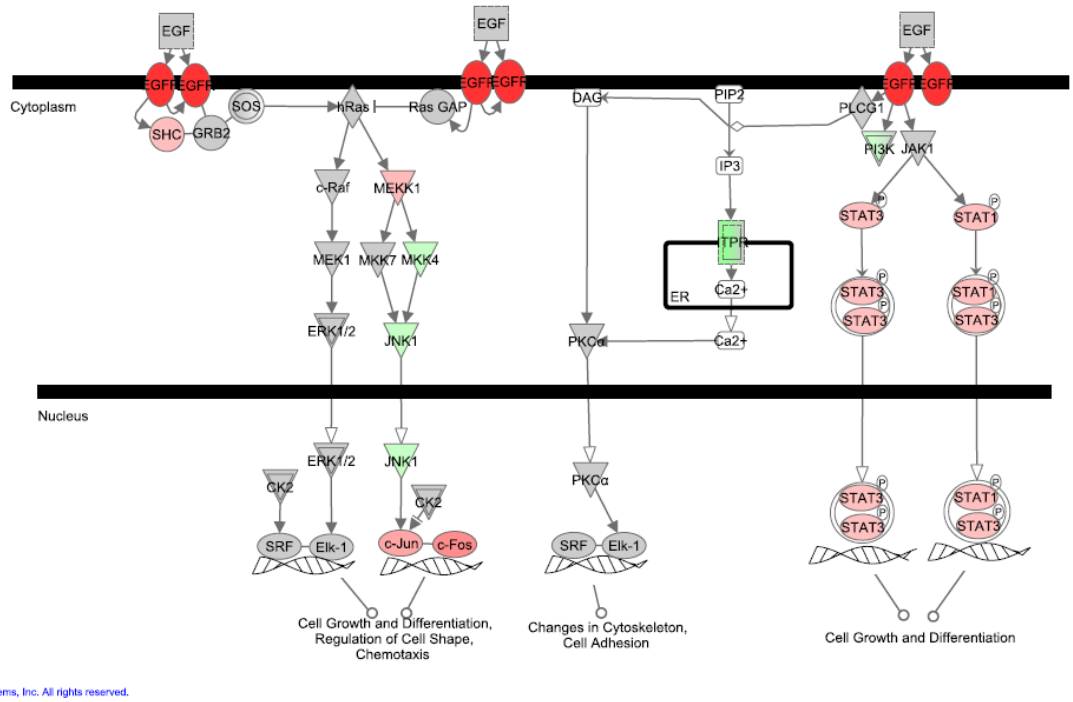

Figure 5. IPA EGFR Canonical Pathway in Tumors with Gain of Chromosome 7 and Amplification in the Chromosomal Region 7p12 *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table 4. EGFR Canonical Pathway characterization in Group 1 *vs* Normal Cerebral Cortex.

Cases	Genes Overexpressed (FC ≥ 2)		Genes Underexpressed (FC ≤ -2)		Genes with no Change in Expression (-2 ≤ FC ≤ 2)
Group 1 *vs* N	7/48 (14.5%)	*c-Fos, c-Jun, EGFR, MEKK1, SHC, STAT1, STAT3*	4/48 (8.4%)	*IPTR1, JNK1, MKK4, PIK3CB*	37/48 (77.1%)

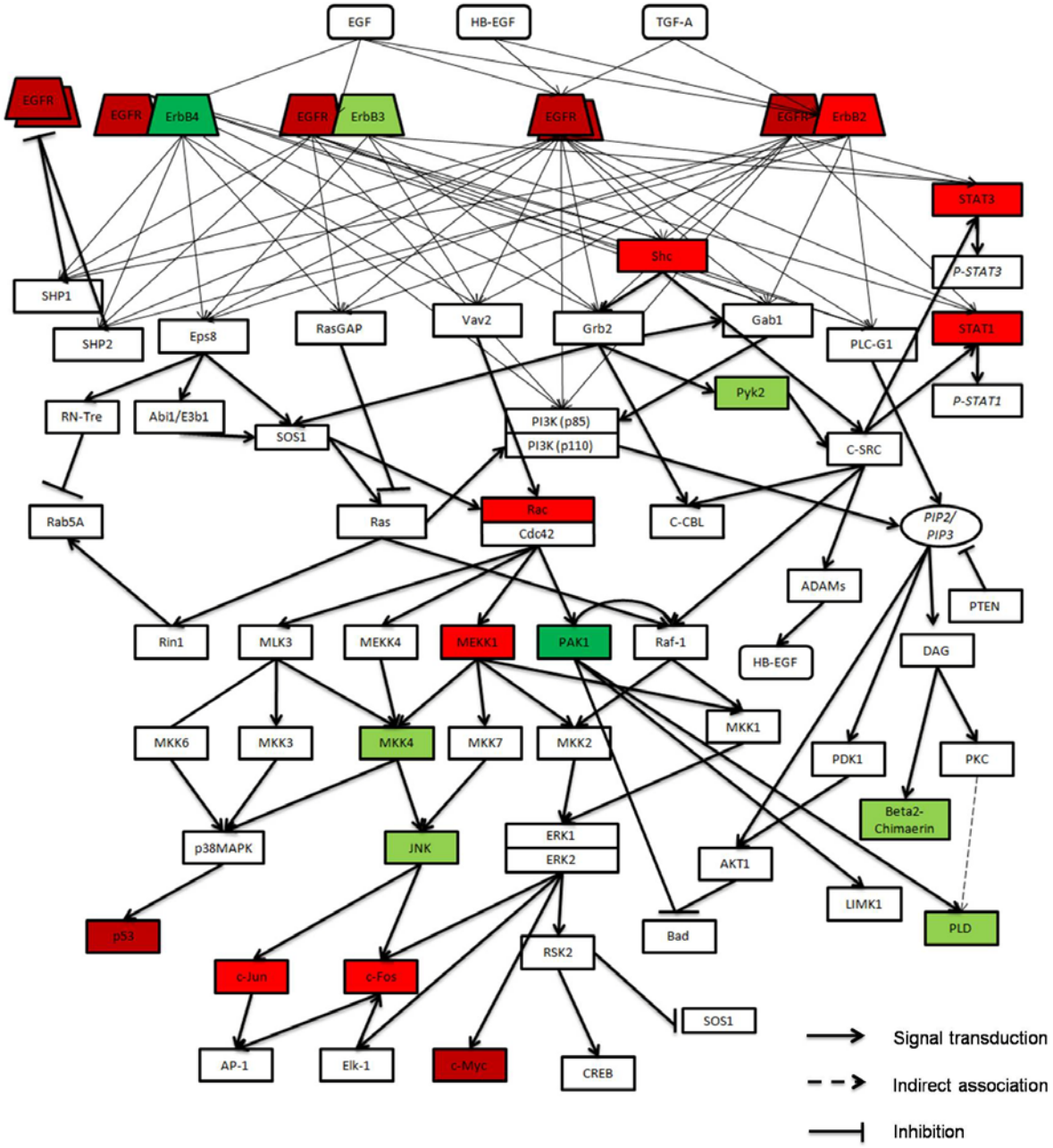

Figure 6. EGFR Pathway (adapted from Oda *et al.*, 2005) in Tumors with Gain of Chromosome 7 and Amplification in the Chromosomal Region 7p12 (Group 1) *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored white have no change in gene expression and molecules colored white with italic lettering have no data of expression Trapezoids represent receptor tyrosine kinases, rectangles represent generic proteins, circles represent simple molecules.

Tumors with Gain of Chromosome 7 *vs* N

When comparing Group 2 *vs* N we could observe that the EGFR canonical pathway depicted in Figure 7 and Table 5 (page 18) was very similar to the one obtained for the previous cases (Group 1). The most striking difference between them was the degree of overexpression of the *EGFR* gene, which was 2.4 times more expressed in these samples than in the normal cerebral cortex control (p-value = 0.108). In regard to the signaling pathway branches, results were similar of those obtained in Group 1 vs N, where no distinctive pattern of gene expression in MAPK or PIP signaling pathways was observed. Overexpression of STAT, more specifically, the *STAT3* gene with a 2.3 fold-change over the normal cerebral cortex control (p-value = 0.065) was a distinctive feature in this cases.

As in the previous samples group, a more detailed EGFR signaling pathway was built as depicted in Figure 8 (page 18) and again, the results were similar of those obtained with the IPA EGF canonical pathway.

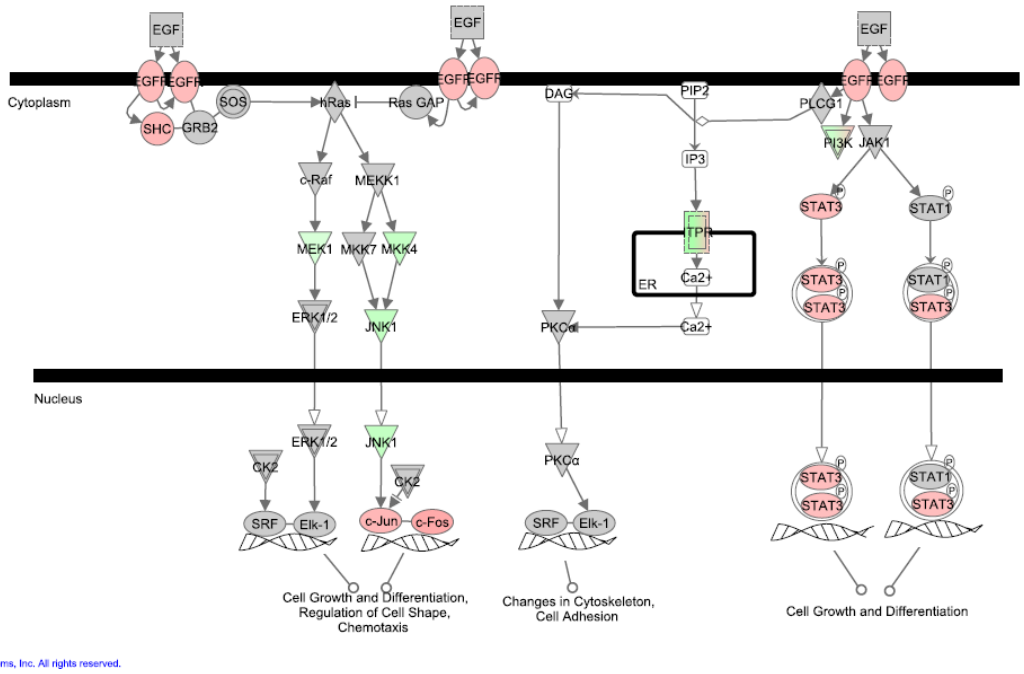

Figure 7. IPA EGFR Canonical Pathway in Tumors with Gain of Chromosome 7 *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table 5. EGFR Canonical Pathway characterization in Group 2 *vs* Normal Cerebral Cortex.

Cases	Genes Overexpressed (FC ≥ 2)		Genes Underexpressed (FC ≤ -2)		Genes with no Change in Expression (-2 ≤ FC ≤ 2)
Group 2 *vs* N	7/48 (14.5%)	*c-Fos, c-Jun, EGFR, ITPR3, PIK3CG, SHC, STAT3*	6/48 (12.5%)	*IPTR1, JNK1, MEK1, MKK4, PIK3CB, PIK3C2B*	35/48 (73%)

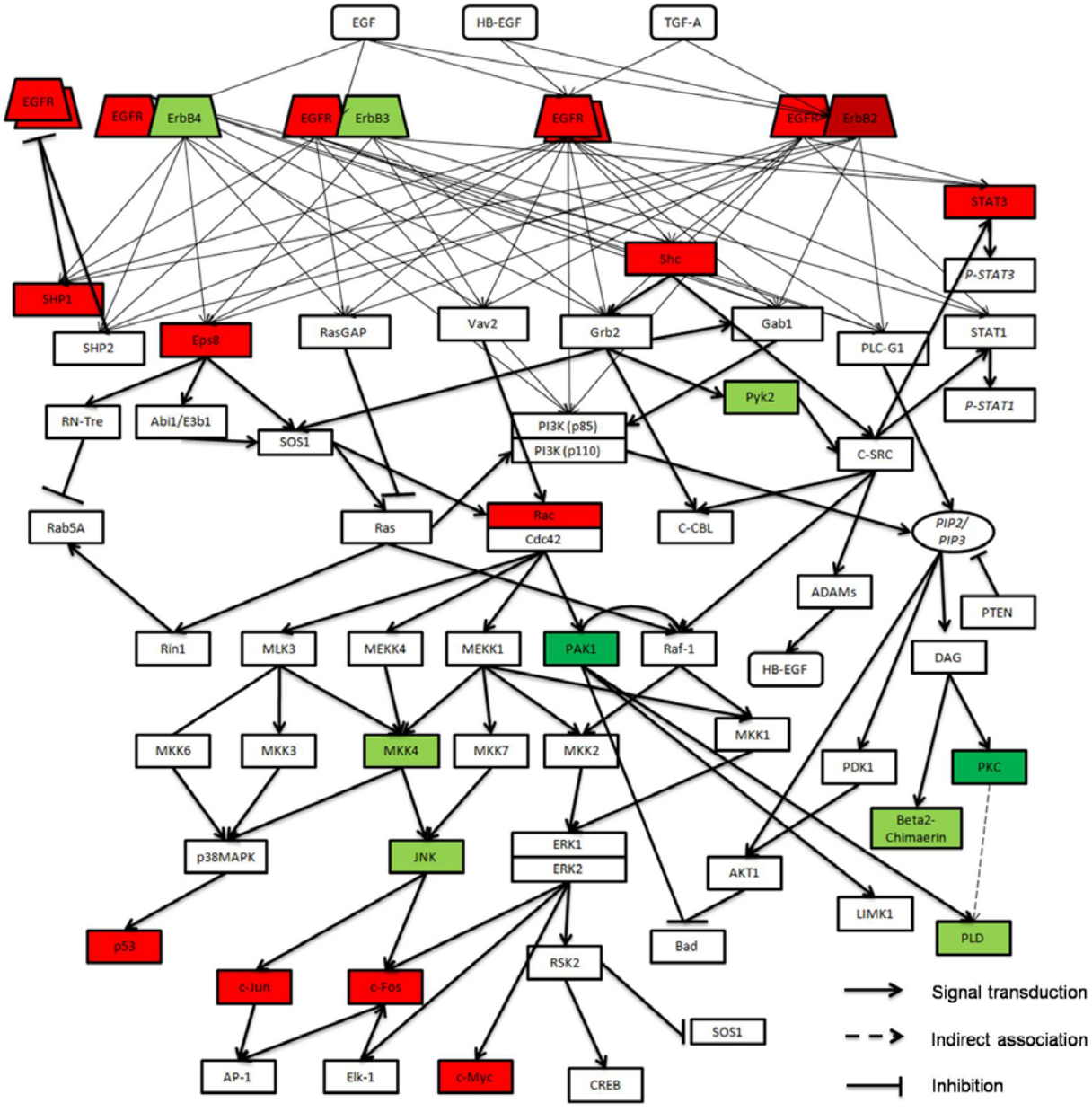

Figure 8. EGFR Pathway (adapted from Oda *et al*, 2005) in Tumors with Gain of Chromosome 7 (Group 2) *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored white have no change in gene expression and molecules colored white with italic lettering have no data of expression. Trapezoids represent receptor tyrosine kinases, rectangles represent generic proteins, circles represent simple molecules.

Tumors with Gain of Chromosome 7 and Amplification in 7p12 *vs* Tumors with Gain of Chromosome 7

Comparison between Group 1 and Group 2 of our series as depicted in Figure 9 and Table 6, revealed that there were no significant differences between the gene expression of most of the individual molecules in the EGFR mediated signaling pathways between these two groups. However, there was a striking difference between them in the overexpression of the *EGFR* gene, which was 5.42 fold overexpressed (p-value of 0.0387), what can be related to the lack of amplification of the *EGFR* gene in Group 2 samples.

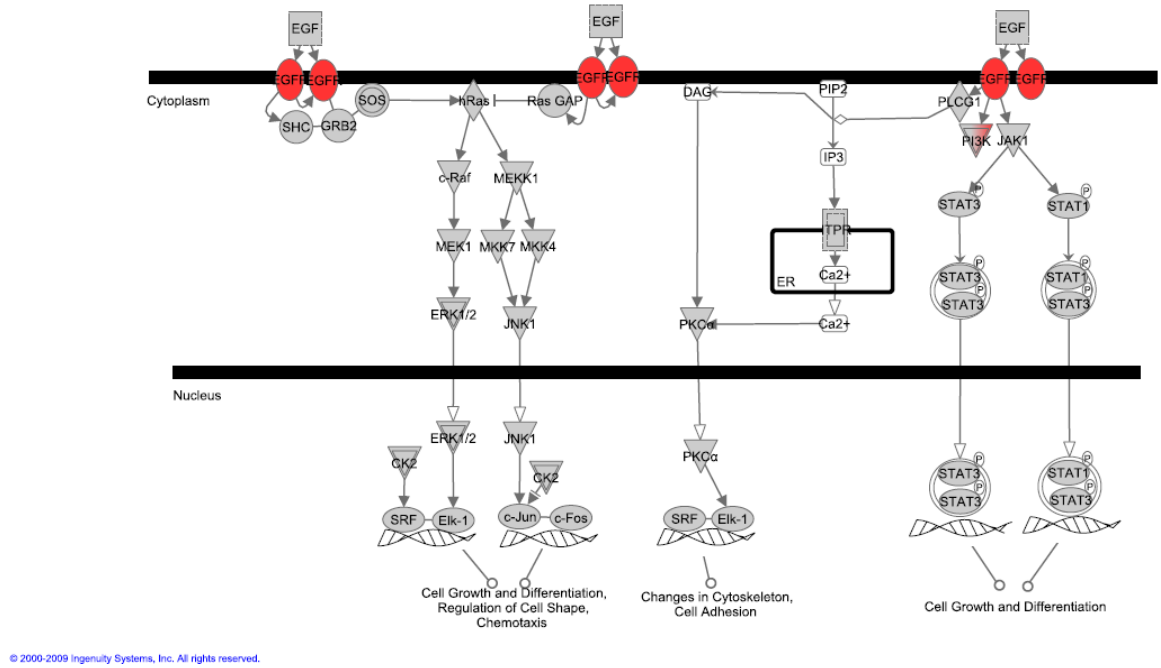

Figure 9. IPA EGFR Canonical Pathway in Tumors with Gain of Chromosome 7 and Amplification in 7p12 *vs* Tumors with Gain of Chromosome 7. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table 6. EGFR Canonical Pathway characterization in Group 1 *vs* Group 2.

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Group 1 *vs* Group 2	2/48 (4%)	*EGFR, PIK3C2B*	0/48 (0%)	-	46/48 (94%)

Tumors with Gain of Chromosome 7 and Amplification in 7p12 *vs* Tumors with Gain of Chromosome 7 (TCGA data)

Statistical validation of our results was performed by comparing them with the publicly available gene expression array data of GBM derived from the work of "The Cancer Genome Atlas" (TCGA) consortium (http://cancergenome.nih.gov/dataportal/data/about/). TCGA used a different microarray set (GeneChip® Human Genome U133 Plus 2.0 Array from Affymetrix) to analyze gene expression in 79 cases of GBM with amp7p12 and 106 cases of GBM with gain of chromosome 7. Data was analyzed with the same protocol used in our series of data (Partek one-way ANOVA analysis with FC filtering, followed by Ingenuity Pathway Analysis). The EGFR canonical pathway retrieved from IPA depicted in Figure 10 and Table 7 (page 21) markedly resembled the one obtained with our data series. Concurrently, using the TCGA database analysis evidenced that the most significant difference between GBM with amp7p12 and GBM with gain of chromosome 7 was the expression level of the *EGFR* gene. This gene presented an average 13.29 fold-change (p-value = 1.73E-34) in GBM with amp7p12 over GBM with gain of chromosome 7 in the TCGA data series.

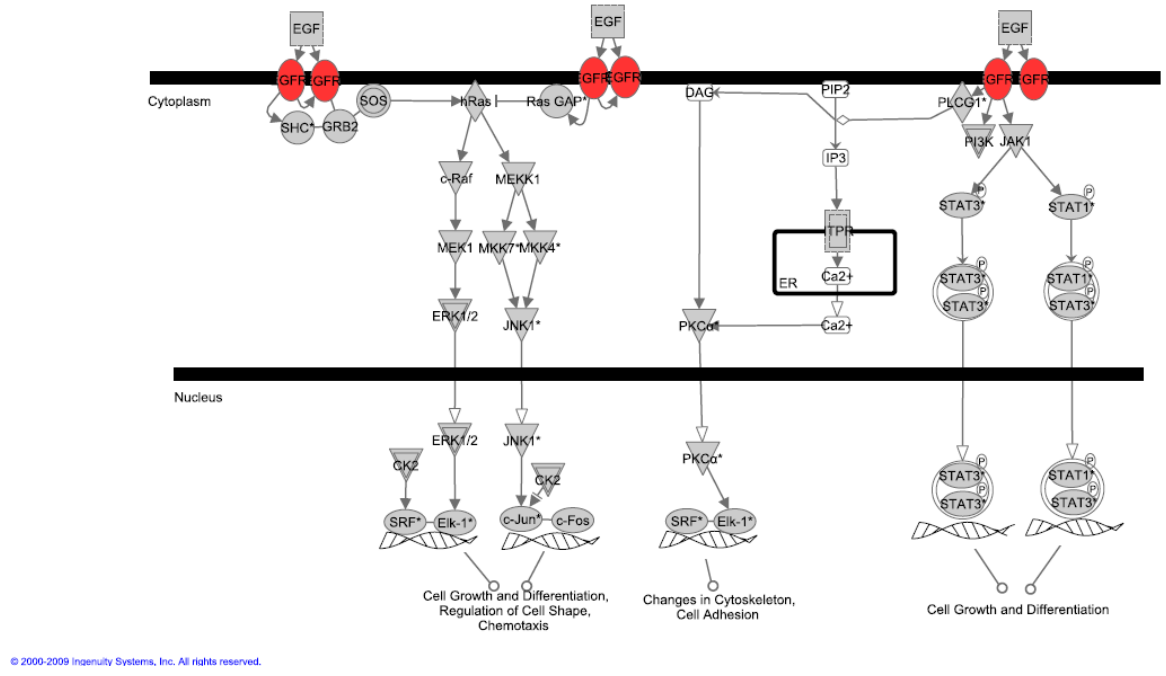

Figure 10. IPA EGF Canonical Pathway in Tumors with Gain of Chromosome 7 and Amplification in 7p12 *vs* Tumors with Gain of Chromosome 7 (TCGA data). Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table 7. EGFR Canonical Pathway characterization in Group 1 *vs* Group 2 (TCGA data).

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Group 1 *vs* Group 2 (TCGA data)	1/48 (2%)	*EGFR*	0/48 (0%)	-	47/48 (98%)

Tumors with Amplification in 8q24 *vs* N

Analysis of Group 3 *vs* N data revealed a different pattern for the EGFR signaling pathway branches as depicted in Figure 11 and Table 8 (page 22). Although overexpression of the *EGFR* gene was evident with a 2.3 fold-change over the normal cerebral cortex control in this sample group, there was no ascertainable genetic preference for any of the signaling pathways, in contrast with previous sample groups (Groups 1 and 2) in which a genetic preference for STAT signaling was apparent.

In these cases, *STAT3* gene had a FC of 1.7 over the normal control, which falls below our selected limit. This result was also evident in the more detailed EGFR signaling pathway (Supplementary Figure S7).

Data analysis of these samples also revealed a striking overexpression of the *c-Myc* gene (which is amplified in these samples), with a FC of 6.16. Recently c-Myc overexpression was reported in glioma (Faria *et al.*, 2008) and it has been implicated as a key gene in gliomagenesis (Lassman *et al.*, 2004; Bredel *et al.*, 2005) and maintenance of glioma cancer stem cells (Wang *et al.*, 2008).

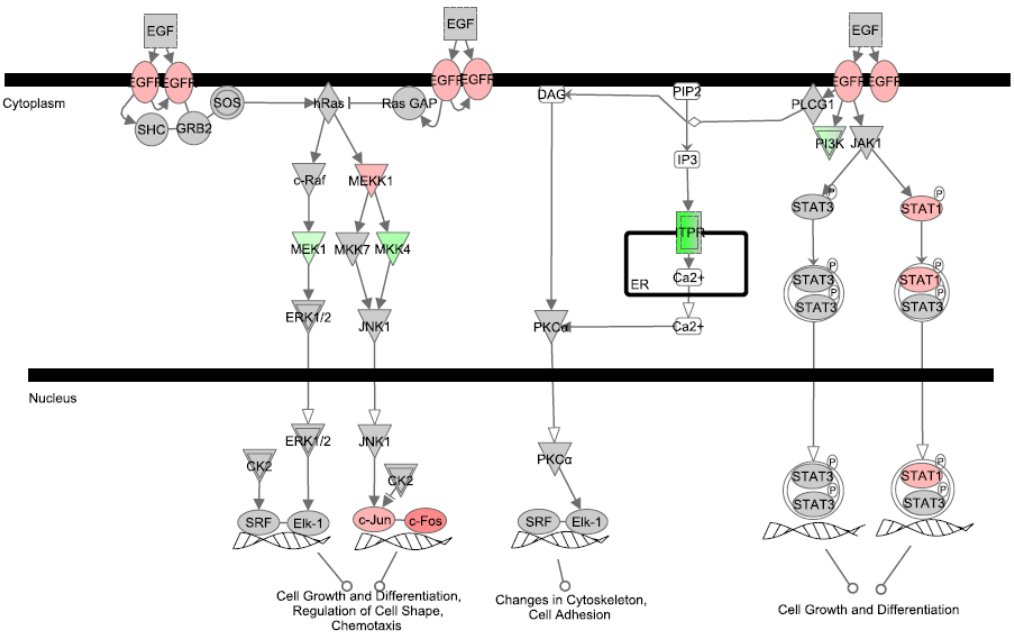

Figure 11. IPA EGFR Canonical Pathway in Tumors with Amplification in 8q24 *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table 8. EGFR Canonical Pathway characterization in Group 3 *vs* Normal Cerebral Cortex.

Cases	Genes Overexpressed (FC ≥ 2)		Genes Underexpressed (FC ≤ -2)		Genes with no Change in Expression (-2 ≤ FC ≤ 2)
Group 3 *vs* N	5/48 (10.5%)	*c-Fos, c-Jun, EGFR, MEKK1, STAT1*	5/48 (10.5%)	*ITPR1, MEK1, MKK4 PIK3CB, PIK3C2B*	38/48 (79%)

Anaplastic Ganglioglioma Cell Line and Tumor *vs* N

Another goal of this work was to determine if cell lines derived from primary tumors maintained the EGFR signaling pattern of the original tumors. For one of the cases (case 11) of our group of samples, RNA quality and quantity for both tumor and cell line were adequate for microarray analysis, enabling us to investigate the EGFR signaling pathway in both samples. As depicted in Figure 12 and Table 9 (cell line) and Figure 13 and Table 10 (tumor, page 23), the EGFR signaling pathway is essentially equal in both cell line and tumor, with, again, an apparent preference toward STAT signaling, in a very similar result to those of the other samples analyzed. The comparison of the cell line sample *vs* the tumor sample yielded a totally unchanged EGFR pathway (not shown). The more detailed EGFR signaling pathway confirmed these results (not shown).

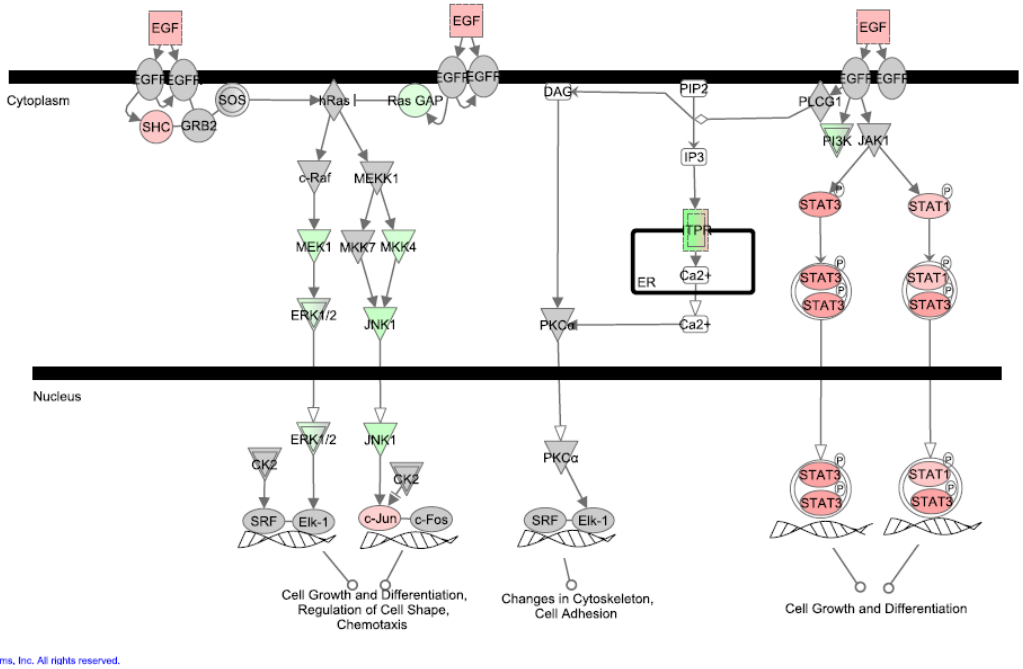

Figure 12. IPA EGFR Canonical Pathway in a Tumor Cell Line with derived from an Anaplastic Ganglioglioma of case 11 *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table 9. EGFR Canonical Pathway characterization in Case 11 (Cell Line) *vs* Normal Cerebral Cortex.

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Case 1 – Cell Line *vs* N	6/48 (12.5%)	*c-Jun, EGF, ITPR3, SHC1, STAT1, STAT3*	11/48 (23%)	*ERK 3,ITPR1, ITPR2, JNK1, MEK1, MKK4, PIK3CB, PIK3C2B, PIK3C3, PIK3R1, RasGAP,*	31/48 (64.5%)

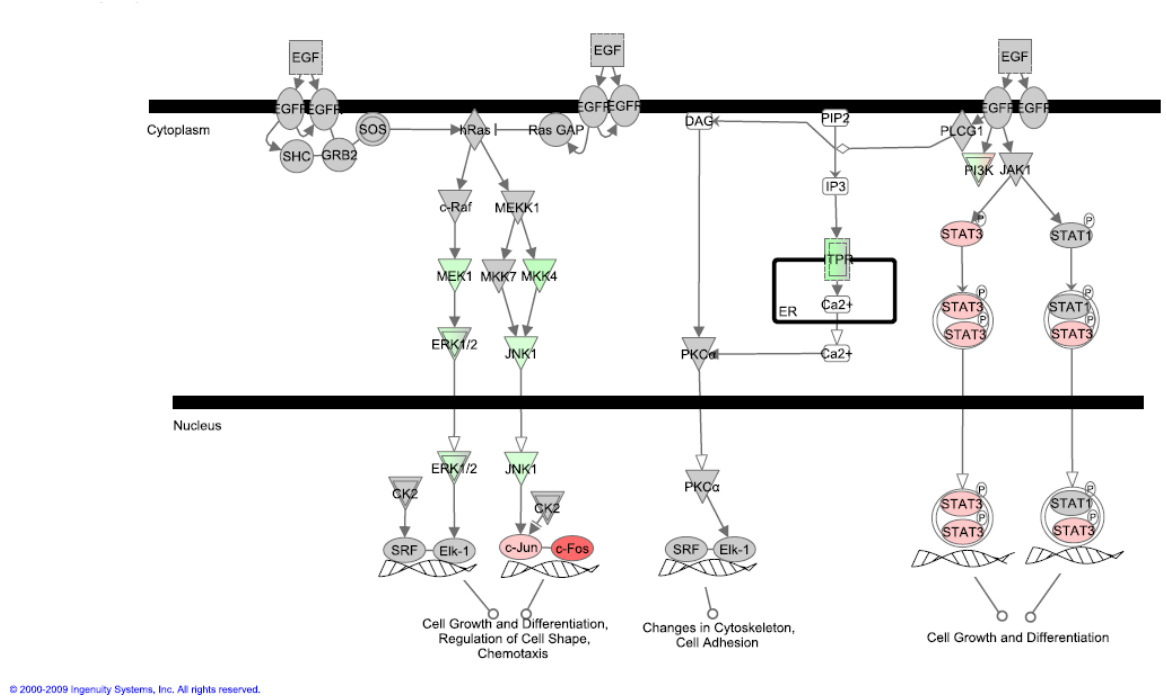

Figure 13. IPA EGFR Canonical Pathway in an Anaplastic Ganglioglioma of case 11, with gain of chromosome 7 *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table 10. EGFR Canonical Pathway characterization in Case 11 (Tumor) *vs* Normal Cerebral Cortex.

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Case 11 - Tumor *vs* N	5/48 (10.5%)	*c-Fos, c-Jun, PIK3CG, PIK3R5, STAT3*	7/48 (14.5%)	*ERK3, ITPR1, JNK1, MEK1, MKK4, PIK3C2B, PIK3CB*	36/48 (75%)

Tumor with Gain of Chromosome 7 and Amplification in 4q12 *vs* N

The comparison of case 12 *vs* normal cerebral cortex, as represented in Figure 14 and Table 11 (page 24), revealed an apparent preference towards STAT signaling, as in our previous results for Groups 1 and 2. In this case, *STAT3* gene had a 2.5 fold overexpression over the normal control. Data analysis also revealed overexpression of *PDGFRα* (whose

locus is amplified in this sample). This protein is an indirect activator of STAT3 (activation is mediated by JAK) (Vignais *et al.*, 1999) and in our samples had a 7.85 fold gene overexpression over the normal control. This event may constitute an alternative activating pathway for STAT3 in this sample. Results were confirmed by the more detailed EGFR signaling pathway (not shown).

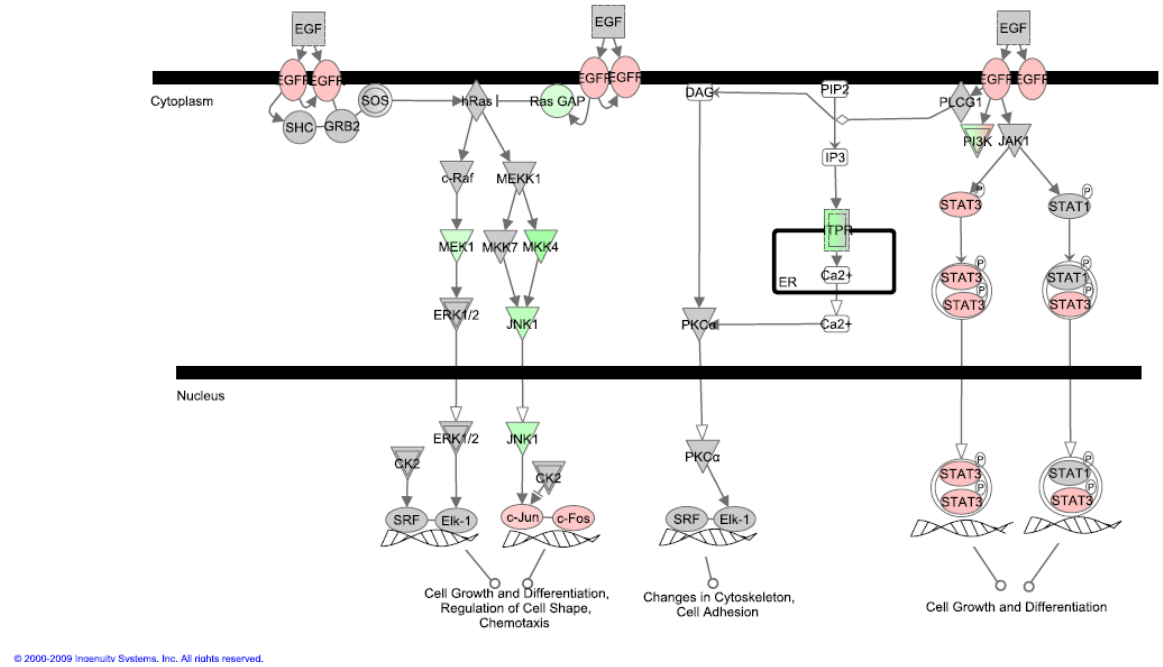

Figure 14. IPA EGFR Canonical Pathway in a Tumor with Gain of Chromosome 7 and Amplification in 4q12 (case 12) *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table 11. EGFR Canonical Pathway characterization in Case 12 *vs* Normal Cerebral Cortex.

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Case 12 *vs* N	7/48 (14.5%)	*c-Fos, c-Jun, EGFR, PIK3C2G, PIK3CG, PIK3R5, STAT3*	8/48 (16.6%)	*JNK1, ITPR1, MEK1, MKK4, JNK1, PIK3C3, PIK3C2B, PIK3CB, RasGAP*	33/48 (68.9%)

Other Cases *vs* N

In other comparative analysis using samples that represent unique cell lines derived from primary GBM (depicted in Supplementary Figures S8, S9 and S10 and Tables S2, S3 and S4), the EGFR signaling pathway continued to reveal an apparent preference towards STAT signaling independently of the cytogenetic changes the cell lines possess, what was also confirmed by the more detailed EGFR signaling pathways (not shown).

Discussion

EGFRvIII mutation detection by MLPA

The results of our MLPA experiments were different from those previously documented in the Portuguese population (Viana-Pereira *et al.*, 2008) in regard to the *EGFRvIII* mutation prevalence. In their work, Viana-Pereira *et al.* studied the incidence of *EGFR* overexpression, amplification and *EGFRvIII* mutation in a Portuguese group of high-grade gliomas. In this group's cases *EGFR* amplification was found in 54% of GBM and 67% of Anaplastic Oligodendrogliomas (AO) [*versus* 43% and 42% (AO are similar to high-grade Oligodendrogliomas) respectively, in our cases]. *EGFRvIII* mutation was detected in 22% of GBM and 8% of AO (*versus* 8% and 5% respectively, in our cases).

The difference between our results and Viana-Pereira's group may be explained by the different methods used (Viana-Pereira *et al*. used immunohistochemistry in their attempt to identify the prevalence of *EGFRvIII* mutation). To our knowledge, no work has been published, using the MLPA kit P315 for EGFR, for investigation of *EGFRvIII* prevalence in any population.

To address the issue of different prevalence of the *EGFRvIII* mutation between our technique and Viana-Pereira's, we propose that a third different technique should be used to evaluate the mutation's prevalence in the same group of individuals, mainly a mRNA based evaluation like the one proposed by Mellinghoff *et al.* (Mellinghoff *et al.*, 2005) and Yoshimoto *et al.* (Yoshimoto *et al.*, 2008).

What our results also reflect is that amplification of EGFR is a current event in high-grade gliomas (both GBM and high-grade Oligodendrogliomas) (Figure 4, page 12) being relatively rare in low-grade gliomas, as previously stated by Ohgaki *et al.* (Ohgaki *et al.*, 2005).

EGFR Pathway Analysis in Glioma

Some of the hallmarks of high-grade gliomas are total gain of chromosome 7, amplification of the *EGFR* gene and total loss of chromosome 10 or partial loss of chromosome arm 10q (Ohgaki *et al.*, 2007). EGFR signaling can be deregulated by these two events, at the activation level by *EGFR*, and at the inhibition level by *PTEN* (located in 10q23). EGFR signaling deregulation has been confirmed as a key event in high-grade gliomas (TCGA, 2008). With this work, we sought to understand the pattern of expression of the genes that compose the EGFR pathway and if there was any preference for a particular "branch" of the pathway.

In the majority of our samples, the *STAT3* and *STAT1* genes have, on average, a two-fold increment over the normal control. On the other hand, genes in the PIP and MAPK branches of the EGFR signaling pathway have different levels of expression, that do not

clearly indicate any type of preference toward one or the other. In the PIP pathway there was no ascertainable change in its component's gene expression in the majority of our samples, what lead us to believe that there was no particular preference for EGFR signaling to proceed through PIP signaling. In what concerns the MAPK cascade in the EGFR signaling, there was an apparent antagonistic level of expression in its transducer molecules. As depicted in Figure 5 (page 15 of the Results section), overexpression (e.g. *MEKK1*), underexpression (e.g. *MKK4*) and no change in expression (e.g. *MKK7*) coexisted in the same branch of the pathway. This seemingly incompatible gene expression between the constituents of the MAPK cascade may lead to a decreased signaling ability by this pathway, what lead us to believe that, like for PIP signaling, there was no particular preference for EGFR signaling to proceed through the MAPK cascade.

Of the three possible EGFR transducer intracellular signaling cascades (PIP, MAPK and STAT), we could recognize that gene expression changes support a preferential signaling through STAT1/3.

However, in one sample group (Group 3, amp8q24), we did not observe this event. As depicted in Figure 11 and Table 8 (pages 21 and 22 of the Results section), the Group 3 *vs* N analysis did not show a clear preference for STAT3, but data analysis of gene expression revealed an increased expression of the *c-Myc* gene. In the comparisons of Group 1 *vs* N and Group 2 *vs* N, *c-Myc* overexpression was also evident, with a median FC of 4.05. It is known that STAT3 is a mediator of *c-Myc* transcription (Dauer *et al.*, 2005) and inhibition of STAT3 in glioma cells leads to decreased expression of *c-Myc* (Gu *et al.*, 2008). Thus, the STAT3 signaling preference revealed in Group 1 and 2 cases can be interpreted as a glioma cell strategy to induce increased expression of *c-Myc* when no other possibility (like amplification of the gene locus) exists. If so, amplification of the *c-Myc* locus, which occurs in Group 3, would relieve selective pressure for EGFR signaling to proceed through STAT3, leading to decreased overexpression of this gene. This hypothesis could therefore explain why STAT3 signaling preference is not so evident in Group 3 cases when compared with Group 1 and 2 cases.

STAT3 activation has been encountered in roughly 51% of GBM (Abou-Ghazal *et al.*, 2008), and has been identified as a key regulator of both immune suppression and tumorigenesis. STAT3 activity in natural killer cells and neutrophils leads to decrease of their citotoxic activity. In dendritic cells STAT3 reduces the expression of MHC II, CD80, CD86 and IL12, which renders these cells unable to stimulate T cells and generate antitumor immune responses (Kortyleswki *et al.*, 2008). STAT3 participation in tumorigenesis lies in its ability to induce transcription of several genes involved in preventing apoptosis and enhancing proliferation, angiogenesis, invasion and metastasis (Dauer *et al.*, 2005; Abou-Ghazal *et al.*, 2008).

As previously stated, the STAT3 protein is a direct target of EGFR-mediated phosphorylation, what occurs when STAT3 binds phosphorylated residues in the C-terminus of the EGFR protein (Y1068 and Y1086) (Xia *et al.*, 2002), which in turn phosphorylates the Tyrosine 705 residue of STAT3, leading to STAT3 homo- or heterodimerization, nuclear translocation and gene activation, making this the most straight-forward pathway for EGFR signaling. STAT3 activity is regulated by various proteins, mainly PIAS, SOCS and protein phosphatases like PTPRD. In the majority of our samples, the genes encoding these proteins are either underexpressed or have no change in gene expression. The *PTPRD* gene for instance has an average five-fold gene expression reduction in Group 1 and 2 samples *vs* N.

Knowing the importance of STAT3 activation in glioma, several attempts at promoting apoptosis and decreased proliferation of glioma by inhibition of STAT3 activity have been made.

One of the first attempts of inhibiting STAT3 was made by Konnikova *et al.* (Konnikova *et al.*, 2003) with siRNA. In this work, Konnikova *et al.* inhibited STAT3 in astrocytoma cell lines and normal human astrocytes (NHA) and observed a decreased viability and induction of apoptosis in the astrocytoma cell lines but not in the NHA, proposing siRNA should be used as a possible therapy.

Another effort, by Shao *et al.* (Shao *et al.*, 2003), used phosphopeptides in squamous carcinoma cells. Phosphopeptides, mimicking the phosphorylated residues of EGFR where STAT3 binds for activation (Y1068 and 1086), resulted in a destabilization of STAT3 homodimers, and thus, lead to a decrease of STAT3-DNA binding and cell proliferation.

One attempt at inhibiting STAT3 by blocking its capacity to bind DNA was made by Gu *et al.* (Gu *et al.*, 2008). Gu *et al.* used decoy oligonucleotides (ODN) mimicking STAT3 specific cis-elements in two glioma cell lines. These ODN blocked STAT3 signaling and decreased its capacity to mediate transcription of several genes including *c-Myc*, *Cyclin D1* and *Bcl-xl*, resulting in decreased cell proliferation by inducing of apoptosis and cell-cycle arrest.

Use of small molecule inhibitors of STAT3 was proposed by Fuh *et al.* (Fuh *et al.*, 2009). Fuh *et al*. used LLL-3, a polyphenol that binds the SH2 domain of STAT3 (the binding domain), leading to a decreased ability for STAT3 to dimerize. Applying LLL-3 to various glioma cell lines led to the inhibition of STAT3-dependent transcription and to induction of apoptosis in the cell lines. In mice with injection of GBM cell lines, the use of LLL-3 increased viability from 16 days to 28.5 days.

Li *et al.* (Li *et al.*, 2009) used a strategy based upon delivery of a lentiviral vector with a shRNA specific for STAT3, to induce knockdown of STAT3 expression in GBM cell lines. STAT3 inhibition with this approach led to suppression of growth and increase in apoptosis in the cell lines and also an increase in differentiation of some cells.

Conclusion:

Deregulation of the EGFR signaling pathway is one of the most important features of GBM. When deregulated, this pathway confers upon the GBM cells an increased ability to proliferate and evade both apoptosis and immune responses from the host. Suppressing the activity of this pathway should, in theory, lead to a decrease in some of these abilities, and hence, to a decrease in tumor mass and an increase in survival of the host. Efforts in the targeting of this pathway have been made, mostly through the inhibition of the EGFR protein. Although some progress has been made, progression free survival of GBM patients still remains dismal. This probably occurs because EGFR shares most of its pathways with other RTKs and non-RTKs, giving the glioma cells the ability to bypass inhibition of EGFR, through activation of its core signaling pathways via other RTKs and non-RTKs. Clearly, another target, downstream from activators like EGFR, must be selected to enable blockage of GBM core pathways. Several attempts at targeting key components of the PIP and MAPK signaling cascade have been made in this regard, but median survival time remains dismal. Our work reveals STAT3 as another potential molecule that can fit this description. STAT3 overexpression and activation have been proven in GBM, but, to our knowledge, no work has been published on the preferential activation of the STAT3 pathway in GBM (Figure 15).

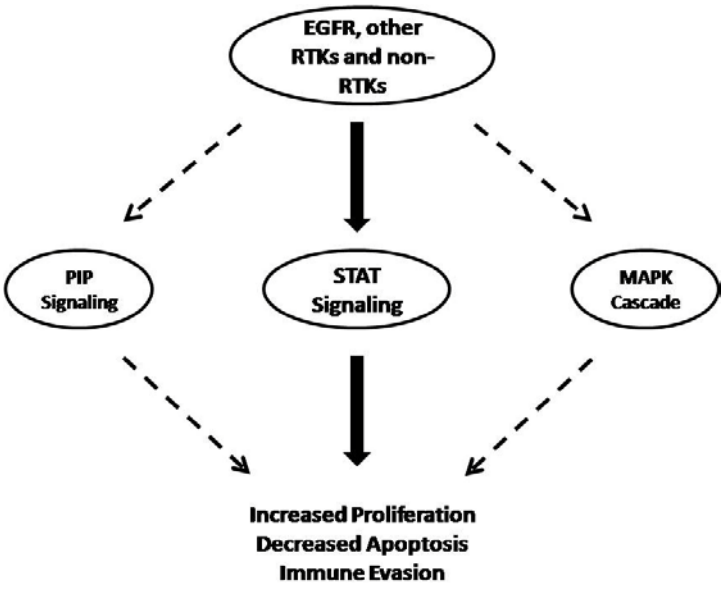

Figure 15. EGFR Signaling pathway model in GBM.

We can look at STAT3 as a "molecular hub" because it is activated by various RTKs and non-RTKs, and it can lead to increased transcription of several genes related to proliferation, immune evasion and inhibition of apoptosis. In theory, inhibition of the signaling pathways at this level should overcome the obstacles found in the inhibition of EGFR. And in fact, some recent work on inhibition of STAT3 in glioma cells has yielded promising results, although, to our knowledge, no clinical trials involving any of the strategies mentioned in the discussion are ongoing.

We are aware that sample size is a limitation of our work, so we propose as a future work to scale-up our results, with different approaches directed at STAT3: evaluation of the mRNA levels through real-time RT-PCR and validation of STAT3 activation by evaluation of Tyrosine 705 Phosphorylated STAT3 and Tyrosine 1068 Phosphorylated EGFR presence through immunohistochemistry in a larger set of samples.

References:

Abou-Ghazal M, Yang DS, Qiao W, Reina-Ortiz C, Wei J, Kong L-Y, Fuller GN, Hiraoka N, Priebe W, Sawaya R, Heimberger AB (2008) The Incidence, Correlation with Tumor-Infiltration, and Prognosis of Phosphorylated STAT3 Expression in human Gliomas. Clinical Cancer Research 14(24):8228-8235;

Blanco-Aparicio C, Oliver R, Leal JMF, Carnero A (2007) PTEN, more than the Akt pathway. Carcinogenesis 28(7)1379-1386;

Bowman T, Broome MA, Sinibaldi D, Wharton W, Pledger WJ, Sedivy JM, Irby R, Yeatman T, Courtneidge SA, Jove R (2001) Stat3-mediated Myc expression is required for Src transformation and PDGF-induced mitogenesis. Proceedings of the Natural Academy of Sciences of the USA 98(13):7319-7324;

Brantley EC, Benveniste EN (2008a) Signal Transducer and Activator of Transcription-3: A molecular hub for signaling pathways in Gliomas. Molecular Cancer Research 6(5)675-684;

Brantley EC, Nabors LB, Gillespie GY, Choi YH, Palmer CA, Harrison K, Roarty K, Benveniste EN (2008b) Loss of protein inhibitors of activated STAT-3 Expression in Glioblastoma Multiforme Tumors: Implications for STAT-3 Activation and gene Expression; Clinical Cancer Research 14(15)4694-4704;

Bredel M, Bredel C, Juric D, Harsh GR, Vogel H, Recht LD, Sikic BI (2005) Functional Network Analysis Reveals Extended Gliomagenesis Pathway Maps and Three Novel MYC-Interacting Genes in Human Gliomas. Cancer Research 65(19):8679-8689;

Bromberg J (2002) Stat proteins and oncogenesis. Journal of Clinical Investigation 109(9):1139-1142;

Caló V, Migliavacca M, Bazan V, Macaluso M, Buscemi M, Gebbia N, Russo A (2003) STAT Proteins: From Normal Control of Cellular Events to Tumorigenesis. Journal of Cellular Physiology 195:157-168;

Coffer PJ, Kruijer W (1995) EGF Receptor deletions define a region specifically mediating STAT transcription factor activation. Biochemical and Biophysical Research Communications 210(1)74-81;

Dauer DJ, Ferraro B, Song L, Yu B, Mora L, Buettner R, Enkemann S, Jove R, Haura EB (2005) Stat3 regulate genes common to both wound healing and cancer. Oncogene 24:3397-3408;

Faria MH, Khayat AS, Burbano RR, Rabenhorst SH (2008) c-Myc Amplification and expression in astrocytic tumors. Acta Neuropathologica 116(1):87-95;

Friedman HS, Kerby T, Calvert H (2008) Temozolomide and Treatment of Malignant Glioma. Clinical Cancer Research 6:2585-2597;

Fuh B, Sobo M, Cen L, Hutzen B, Cisek K, Bhasin D, Regan N, Lin L, Chan C, Caldas H, DeAngelis S, Li C, Li P-K, Lin J (2009) LLL-3 inhibits STAT3 activity, suppresses Glioblastoma cell growth and prolongs survival in a mouse Glioblastoma model. British Journal of Cancer 100:106-112;

Gu J, Li G, Sun T, Su Y, Zhang X, Shen J, Tian Z, Zhang J (2008) Blockage of the STAT3 signaling pathway with decoy oligonucleotides suppresses growth of human malignant glioma cells. Journal of Neurooncology 89:9-17;

Guillamo JS, de Boüard S, Valable S, MArteau L, Leuraud P, Marie Y, Poupon MF, Parienti JJ, Raymond E, Peschanksi M (2009) Molecular mechanisms underlying effects of epidermal growth factor receptor inhibition on invasion, proliferation, and angiogenesis in experimental glioma. Clinical Cancer Research 15(11):3697-3704;

Halatsch ME, Schmidt U, Behnke-Mursch J, Unterberg A, Wirtz CR (2006) Epidermal growth factor receptor inhibition for the treatment of glioblastoma multiforme and other malignant brain tumours. Cancer Treatment Reviews 32:74-89

Hegi ME, Diserens A-C, Gorlia T, Hamou M-F, de Tribolet N, Weller M, Kros JM, Hainfellner JA, Mason W, Mariani L, Bromberg JEC, Hau P, Mirimanoff RO, Cairncross JG, Janzer RC, Stupp R (2005) MGMT Silencing and benefit from Temozolomide in Glioblastoma. New England Journal of Medicine 352:997-1003;

Ihle JN (1996) STATs: Signal Transducers and Activators of Transcription. Cell 84:331-334;

Kallioniemi OP, Kallioniemi A, Piper J, Isola J, Waldman FM, Gray JW, Pinkel D (1994) Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors. Genes, Chromosomes and Cancer 10(4):231-243;

Kari C, Chan TO, Quadros MR, Rodeck U (2003) Targeting the Epidermal Growth Factor Receptor in Cancer. Cancer Research 63(1):1-5;

Konnikova L, Kotecki M, Kruger MM, Cochran BH (2003) Knockdown of STAT3 expression by RNAi induces apoptosis in Astrocytoma cells. BMC Cancer 3:23;

Kortylewski M, Yu H (2008) Role of Stat3 in suppressing anti-tumor immunity. Current Opinion in Immunology 20:228-233;

Kuan CT, Wikstrand CJ, Bigner DD (2001) EGF mutant receptor vIII as a molecular target in cancer therapy. Endocrine-Related Cancer 8:83-96;

Lassman AB, Dai C, Fuller GN, Vickers AJ, Holland EC (2004) Overexpression of c-Myc promotes an undifferentiated phenotype in cultured astrocytes an allows elevated Ras and Akt signaling to induce gliomas from GFAP-expressing cells in mice. Neuron Glia Biology 1:157-163;

Lassman AB, Rossi MR, Raizer JJ, Abrey LE, Lieberman FS, Grefe CN, Lamborn K, Pao W, Shih AH, Kuhn JG, Wilson R, Nowak NJ, Cowell JK, DeAngelis LM, Wen P, Gilbert MR, Chang S, Yung WA, Prados M, Holland EC (2005) Molecular study of malignant gliomas treated with epidermal growth factor receptor inhibitors: tissue analysis from North America Brain Tumor Consortium Trials 01-03 and 00-01. Clinical Cancer Research 11(21):7841-50;

Learn CA, Hartzell TL, Wikstrand CJ, Archer GE, Rich JN, Friedman AH, Bigner DD, Sampson JH (2004) Resistance to Tyrosine Kinase Inhibition by Mutant Epidermal Growth Factor Receptor Variant III Contributes to the Neoplastic Phenotype of Glioblastoma Multiforme. Clinical Cancer Research 10(9):3216-3224;

Li G-h, Wei H, Chen Z-t, Lv S-q, Yin C-l, Wang D-l (2009) STAT3 silencing with lentivirus inhibits growth and induces apoptosis and differentiation od U251 cells. Journal of Neurooncology 91:165-174;

Li S, Schmitz KR, Jeffrey PD, Wiltzius JJ, Kussie P, Ferguson KM (2005) Structural basis for inhibition of the epidermal growth factor receptor by Cetuximab. Cancer Cell 7(4):301-311;

Lodish H, Berk A, Kaiser CA, Krieger M, Scott MP, Breitscher A, Ploegh H (2008) Molecular Cell Biology, Sixth Edition. W.H. Freeman & Company, New York; ISBN-13: 978-0-7167-7601-7;

Louis DN, Ohgaki H, Wiestler OD, Cavenee W (2007) WHO Classification of Tumours of the Central Nervous System, Fourth Edition, WHO Press, ISBN 978-92-832-2430-2;

Martens T, Laabs Y, Gunther HS, Kemming D, Zhu Z, Witte L (2008) Inhibition of Glioblastoma growth in a highly invasive Nude Mouse Model can be achieved by targeting Epidermal Growth Factor Receptor but not Vascular Endothelial Growth Factor Receptor-2. Clinical Cancer Research 14(17):5447-5458;

Mellinghoff IK, Wang MY, Vivanco I, Haas-Koogan DA, Zhu S, Dia EQ, Lu KV, Yoshimoto K, Huang JHY, Chute DJ, Riggs BL, Horvath S, Liau LM, Cavenee WK, Rao N, Beroukhim R, Peck TC, Lee JC, Sellers WR, Stokoe D, Prados M (2005) Molecular Determinants of the Response of Glioblastoma to EGFR Kinase Inhibitors. New England Journal of Medicine 353:2012-2024;

Neyns B, Sadones J, Joosens E, Bouttens F, Verbeke L, Baurain JF, D'Hondt L, Strauven T, Chaskis C, In't Veld P, Michotte A, De Greve J (2009) Stratified phase II trials of cetuximab in patients with recurrent high-grade glioma; Annals of Oncology 20(9):1596-1603;

Nicholas MK, Lukas RV, Jafri NF, Faoro L, Salgia R (2006) Epidermal Growth Factor Receptor-Mediated Signal Transduction in the Development and Therapy of Gliomas. Clinical Cancer Research 12(24):7261-70;

Oda K, Matsuoka Y, Funahashi A, Kitano H (2005) A comprehensive pathway of epidermal growth factor receptor signaling; Molecular Systems Biology 1:2005.0010;

Ohgaki H, Cavenee W (2005) Population-Based Studies on Incidence, Survival Rates, and Genetic alterations in Astrocytic and Oligodendroglial Gliomas. Journal of Neuropathology and Experimental Neurology 64(6)479-489;

Ohgaki H, Kleihues P (2007) Genetic Pathways to Primary and Secondary Glioblastoma. The American Journal of Pathology 170(5):1445-1453;

Olson JJ, Fadul CE, Brat DJ, Mukundan S, Ryken TC (2009) Management of newly diagnosed glioblastoma: guidelines development, value and application. Journal of Neurooncology 93:1-23;

Omuro AMP, Faivre S, Raymond E (2007) Lessons learned in the development of targeted therapy for malignant gliomas. Molecular Cancer Therapy 6(7):1909-1919;

Parkin DM, Bray F, Ferlay J, Pisani P (2005) Global Cancer Statistics 2002. CA: Cancer Journal for Clinicians 55:74-108;

Rahaman SO, Harbor PC, Chernova O, Barnett GH, Vogelbaum MA, Haque SJ (2002) Inhibition of constitutively active STAT3 supresses proliferation and induces apoptosis in glioblastoma multiforme cells. Oncogene 21:8404-8413;

Reiter JL, Threadgil DW, Eley GD, Strunk KE, Danielsen AJ, Sinclair CS, Pearsall RS, Green PJ, Yee D, Lampland AL, Balasubramaniam S, Crossley TD, Magnuson TR, James CD, Maihle NJ (2001) Comparative Genomic Sequence Analysis and Isolation of Human and Mouse Alternative EGFR Transcripts Encoding Truncated Receptor Isoforms. Genomics 71:1–20

Sambrook J, Russel DW (2001) Molecular Cloning: a Laboratory Manual Third Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; ISBN 0-87969-577-3;

Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F and Pals G (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Research 30(12):e57;

Sebastian S, Settleman J, Reshkin SJ, Azzariti A, Bellizi A, Paradiso A (2006) The complexity of targeting EGFR signaling in cancer: From expression to turnover. Biochimica et Biophysica Acta 1766:120-139;

Shao H, Cheng HY, Cook RG, Tweardy DJ (2003) Identification and Characterization of Signal Transducer and Activator of Transcription 3 Recruitment Sites within the Epidermal Growth Factor Receptor. Cancer Research 63:3923-3930;

Silva CM (2004) Role of STATs as downstream signal transducers in Src family kinase-mediated tumorigenesis. Oncogene 23:8017-8023;

Stupp R, Mason WP, Van Den Bent MJ, Weller M, Fisher B, Taphoorn MJB, Belanger K, Brandes AA, Marosi C, Bogdahn U, Curschmann J, Janzer RC, Ludwin SK, Gorlia T, Allgeier A, Lacombe D, Cairncross JG, Eisenhauer E, Mirimanoff RO (2005) Radiotherapy plus concomitant and adjuvant Temozolomide for Glioblastoma. New England Journal of Medicine 352:987-996;

The Cancer Genome Atlas Research Network (2008) Comprehensive genomic characterization defines human Glioblastoma genes and core pathways. Nature 455(23)1061-1068;

Veeriah S, Brennan C, Meng S, Singh B, Fagin JA, Solit DB, Paty PB, Rohle D, Vivanco I, Chmielecki J, Pao W, Ladanyi M, Gerald WL, Liau L, Cloughesy TC, Mischel PS, Sander C, Taylor B, Schultz N, Major J, Heguy A, Fang F, Mellinghoff IK, Chan TA (2009) The tyrosine phosphatase PTPRD is a tumor suppressor that is frequently inactivated and mutated in Glioblastoma and other human cancers. The Proceedings of the Natural Academy of Sciences 106(23):9435-9440;

Viana-Pereira M, Lopes JM, Little S, Milanezi F, Basto D, Pardal F, Jones C, Reis RM (2008) Analysis of EGFR Overexpression, EGFR Gene Amplification and the EGFRvIII Mutation in Portuguese High-grade Gliomas. Anticancer Research 28(2A):913-920;

Vignais ML, Gilman M (1999) Distinct Mechanisms of Activation of Stat1 and Stat3 by Platelet-Derived Growth Factor Receptor in a Cell-Free System. Molecular and Cellular Biology 19(5):3727-3735;

Wang J, Wang H, Li Z, Wu Q, Lathia JD, McLendon RE, Hjelmeland AB, Rich JN (2008) c-Myc is Required for Maintenance of Glioma Cancer Stem Cells. PLoS One 3(11):e3769;

Wilks AF (2008) The JAK kinases: Not just another kinase drug discovery target. Seminars in Cell and Developmental Biology 19:319-328;

Xia L, Wang L, Chung AS, Ivanov SS, Ling MY, Dragoi AM, Platt A, Gilmer TM, Fu X-Y, Chin YE (2002) Identification of Both Positive and Negative Domains within the Epidermal Growth Factor Receptor COOH-terminal Region for Signal Transducer and Activator of Transcription (STAT) Activation. The Journal of Biological Chemistry 277(34):30716-30732;

Yoshimoto K, Dang J, Zhu S, Nathanson D, Huang T, Dumont R, Seligson DB, Yong WH, Xiong Z, Rao N, Winther H, Chakravarti A, Bigner DD, Mellinghoff IK, Horvath S, Cavenee WK, Cloughesy TF, Mischel PS (2008) Development of a Real-time RT-PCR Assay for Detecting EGFRvIII in Glioblastoma Samples. Clinical Cancer Research 14(2):488-493;

Zohrabian VM, Forzani B, Chau Z, Murali R, Jhanwar-Uniyal M (2009) Rho/ROCK and MAPK Signaling pathways are involved in Glioblastoma cell migration and proliferation. Anticancer Research 29:119-124;

http://www.neuropathologyweb.org/chapter7/chapter7bGliomas.html (Acessed July 2009);

http://www.sigmaaldrich.com (Acessed July 2009);

http://cancergenome.nih.gov/dataportal/data/about/ (Acessed September 2009);

http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed (Acessed September 2009);

http://www.mrc-holland.com/ (Acessed September 2009);

http://www.currentprotocols.com/protocol/hg0406 (Acessed September 2009).

Supplementary Data

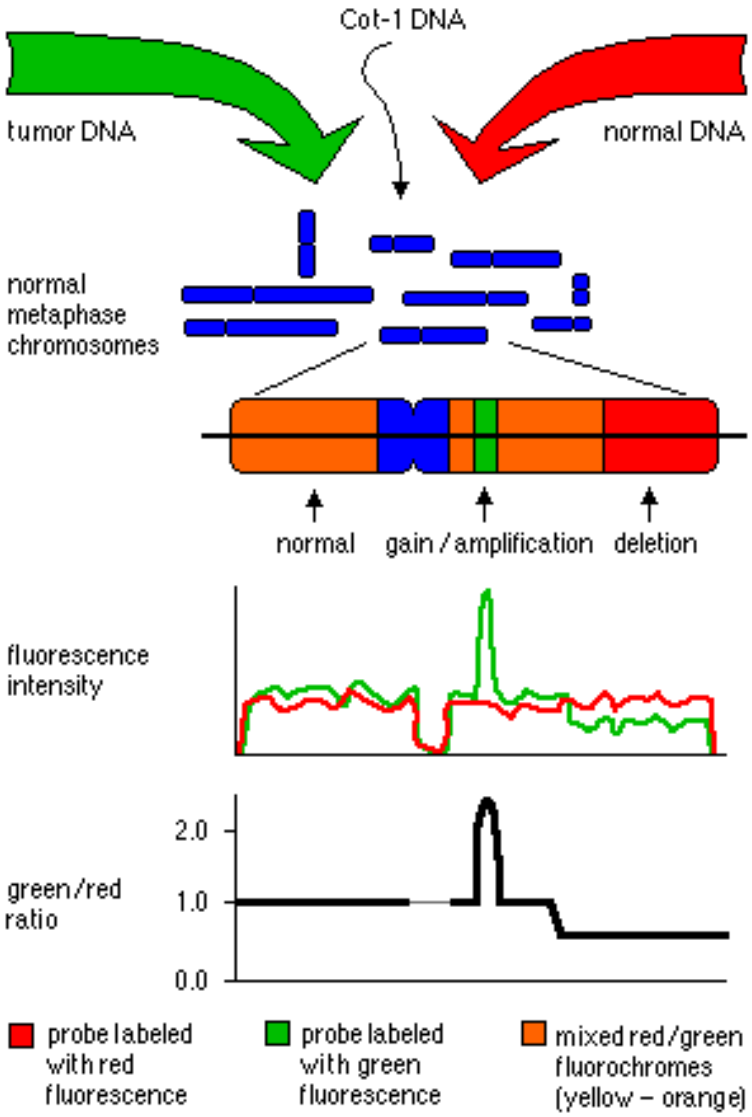

Figure S1. Simplified Overview of the cCGH protocol (*in* http://www.currentprotocols.com/protocol/hg0406).

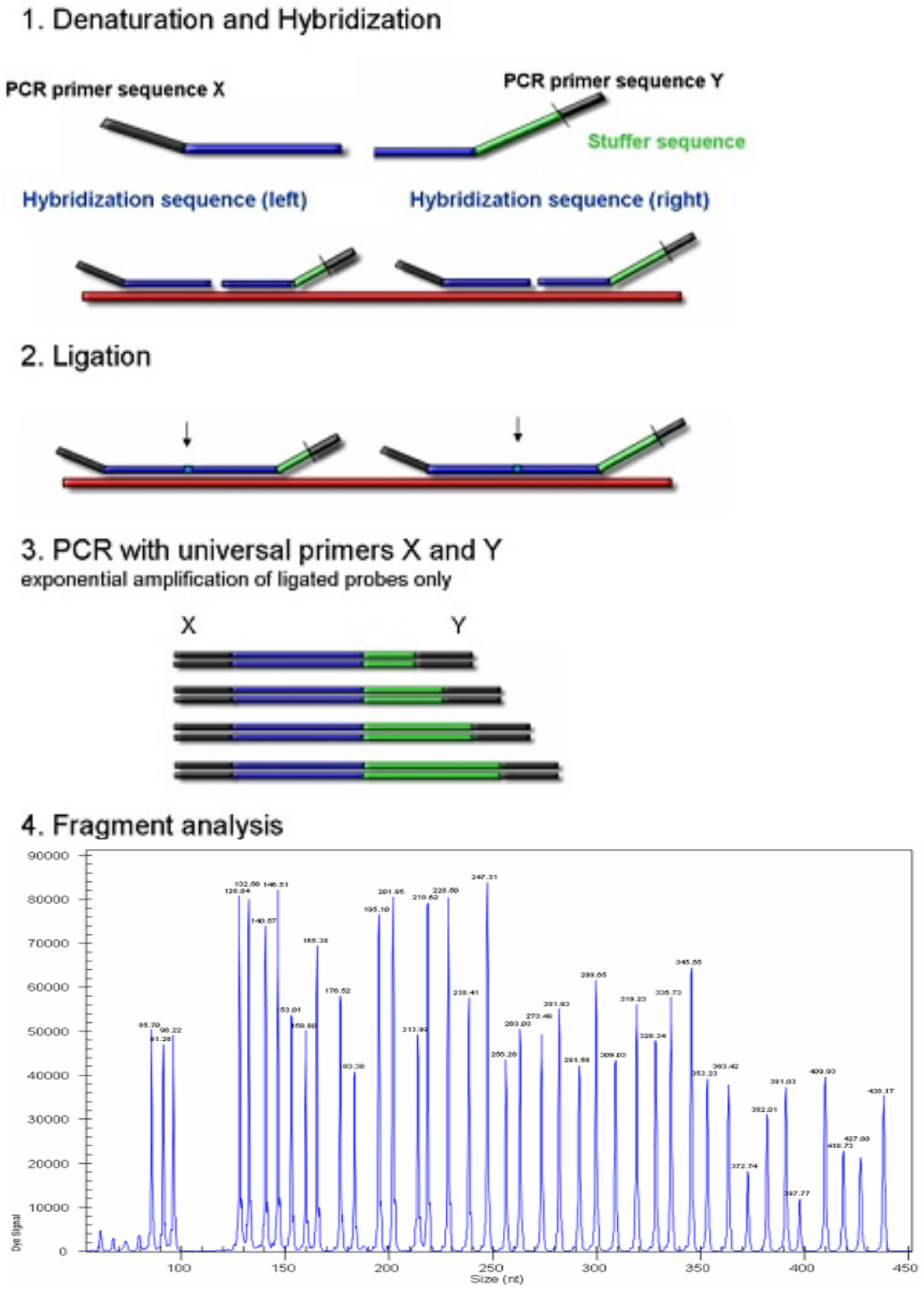

Figure S2. Simplified Overview of the MLPA protocol (*in* http://www.mrc-holland.com/).

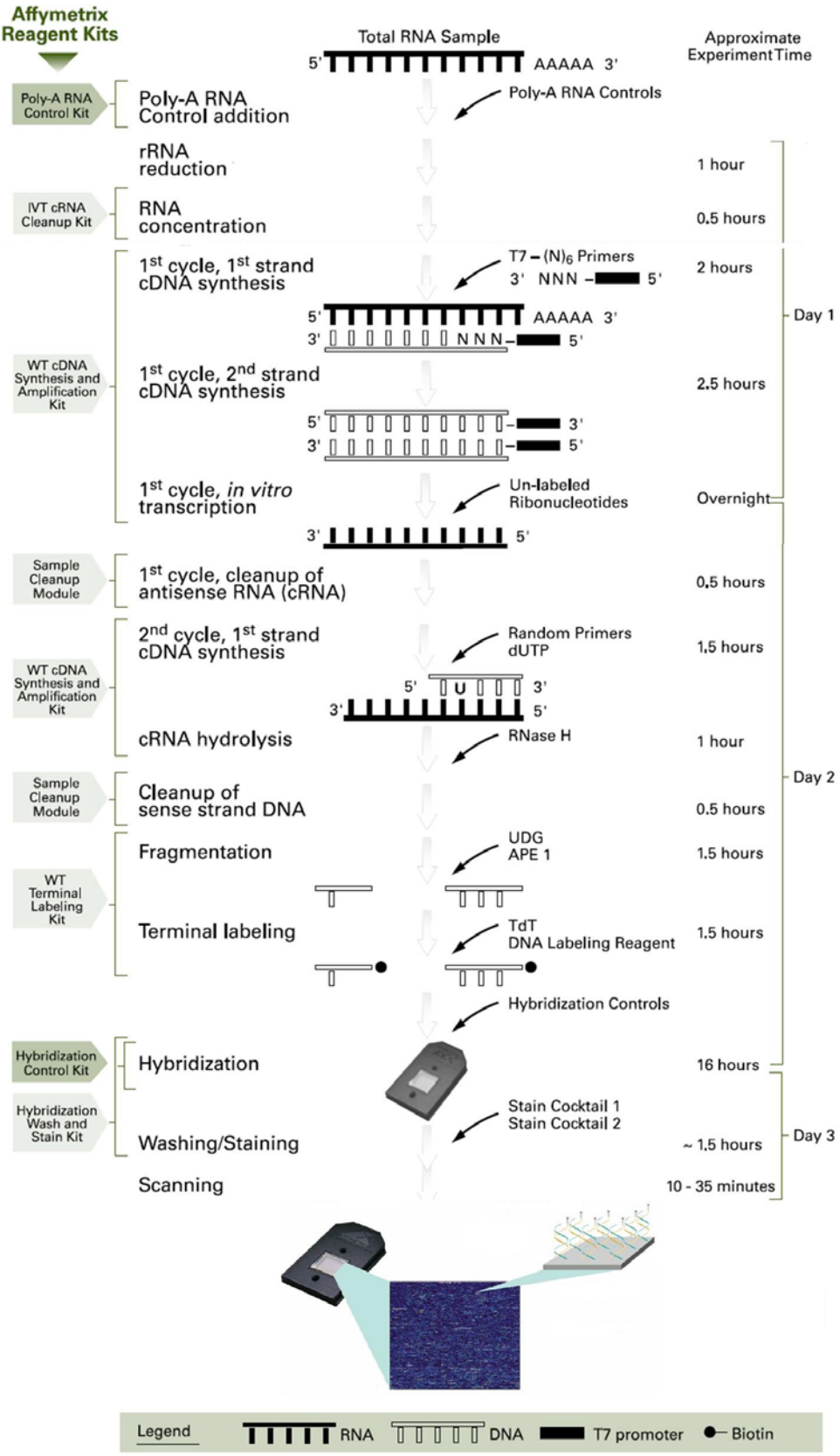

Figure S3. Simplified Overview of the GeneChip Microarray protocol (adapted from the GeneChip Whole Transcript Sense Target Labelling Assay Technical Manual, Affymetrix).

Figure S4. Network shapes in IPA (*in IPA Help Manual).*

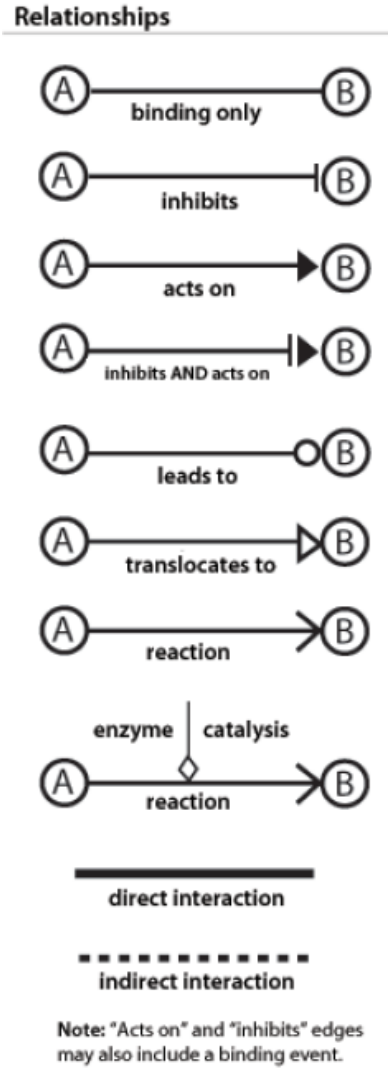

Figure S5. Relationship types in IPA (*in IPA Help Manual*).

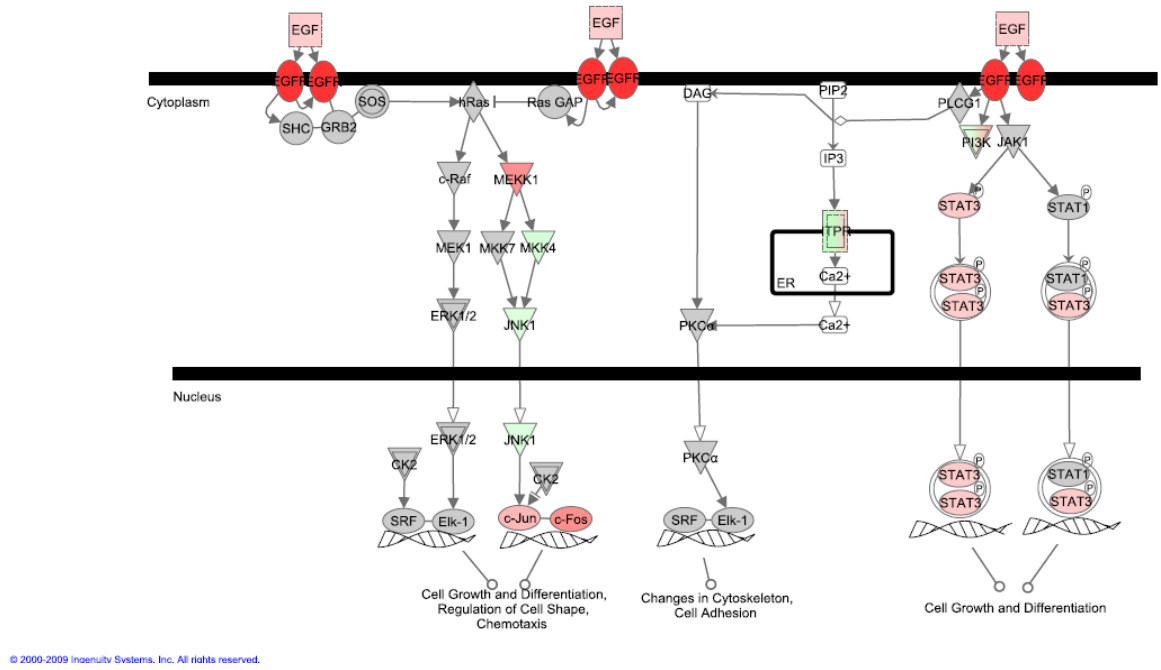

Figure S6. IPA EGFR Canonical Pathway in a Tumor (case 3) with Gain of Chromosome 7, Amplification in 7p12 and *EGFRvIII* mutation *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table S1. EGFR Canonical Pathway characterization in Case 3 *vs* Normal Cerebral Cortex.

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Case 3 *vs* N	8/48 (16.6%)	*c-Fos, c-Jun, EGF, EGFR, ITPR2, MEKK1, PIK3C2B, STAT3*	4/48 (8.4%)	*JNK1, ITPR1, MKK4, PIK3CB*	36/48 (75%)

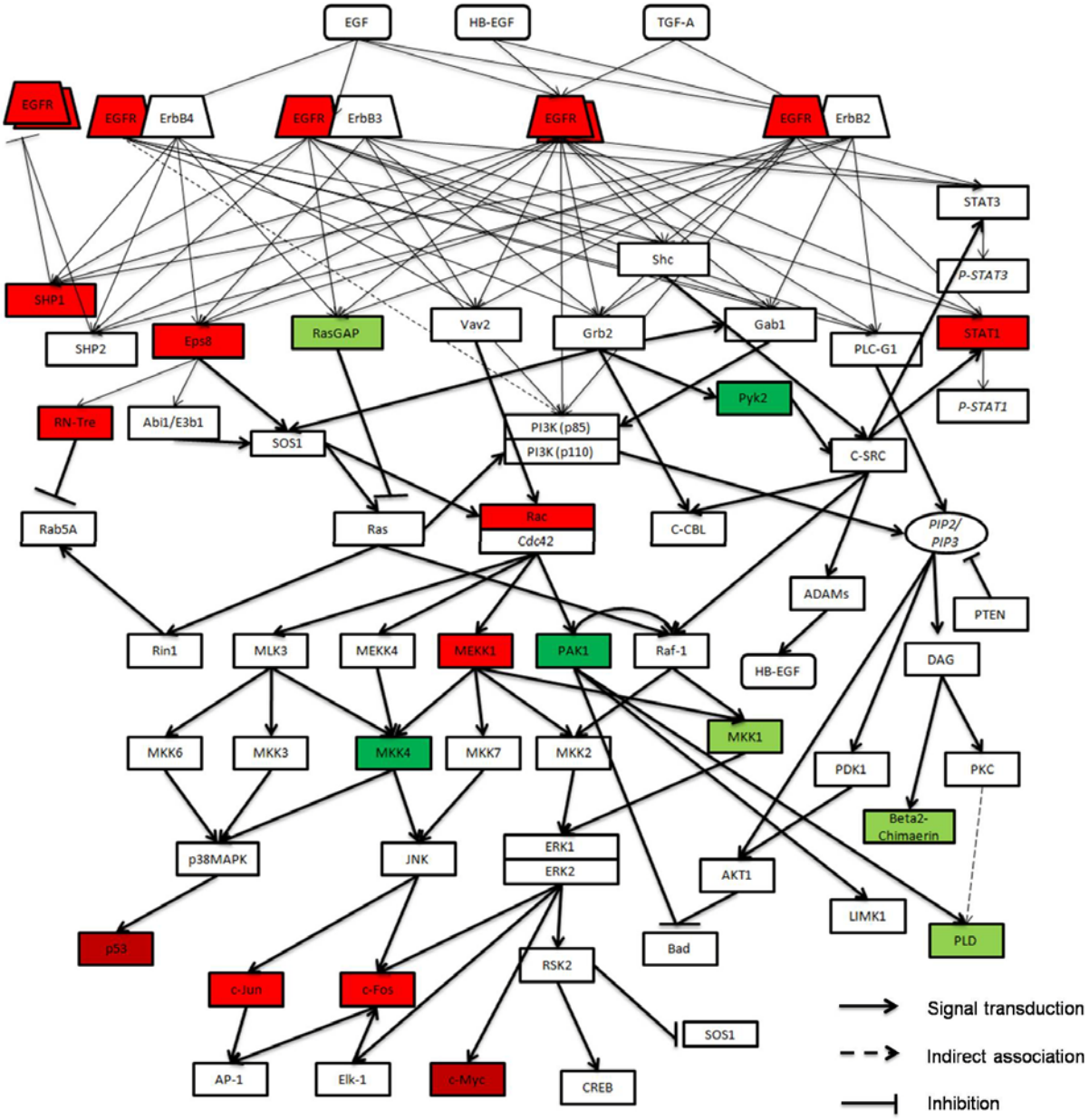

Figure S7. EGFR Pathway (adapted from Oda *et al*, 2005(ref.10)) in Tumors with Amplification in 8q24 (Group 3) *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored white have no change in gene expression and molecules colored white with italic lettering have no data of expression. Trapezoids represent receptor tyrosine kinases, rectangles represent generic proteins, circles represent simple molecules.

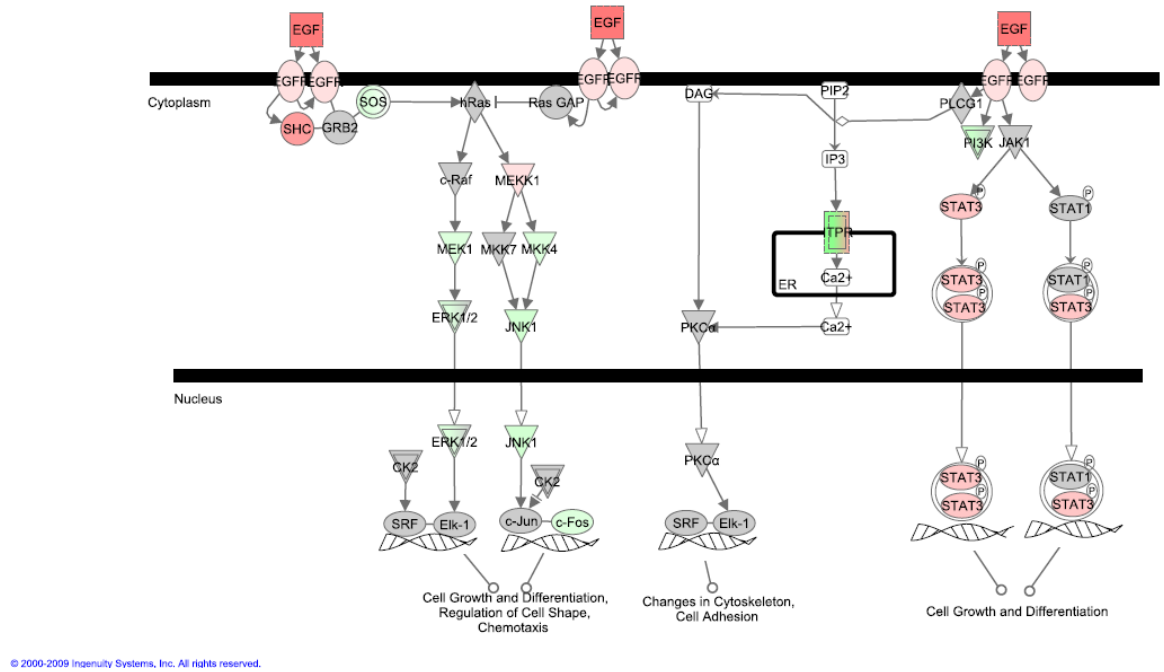

Figure S8. IPA EGFR Canonical Pathway in a Tumor Cell Line with Gain of Chromosomes 7 (case 13) *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table S2. EGFR Canonical Pathway characterization in Case 13 *vs* Normal Cerebral Cortex.

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Case 13 *vs* N	6/48 (12.5%)	*EGF, EGFR, ITPR3, MEKK1, SHC1, STAT3*	11/48 (23%)	*c-Fos, ERK2, ITPR1, ITPR2, JNK1, MEK1, MKK4, PIK3C2B, PIK3CB, PIK3R1, SOS2*	31/48 (64.5%)

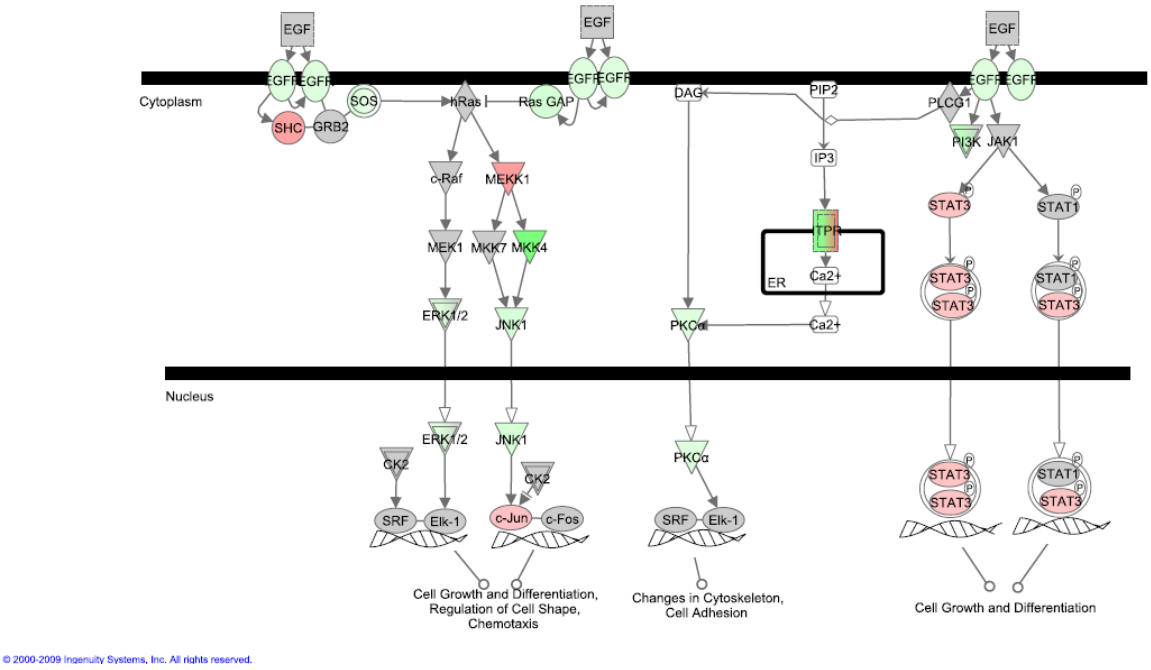

Figure S9. IPA EGFR Canonical Pathway in a Tumor Cell Line with Amplification in 3q, 4p, 18q and Xq (case 14) *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table S3. EGFR Canonical Pathway characterization in Case 14 *vs* Normal Cerebral Cortex.

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Case 14 *vs* N	5/48 (10.5%)	*c-Jun, ITPR3, MEKK1, SHC1, STAT3*	13/48 (27%)	*EGFR, ERK2, ITPR1, ITPR2, JNK1, MKK4, PIK3C2A, PIK3CA, PIK3CB, PIK3R1, PKCα, RasGAP, SOS1*	30/48 (62.5%)

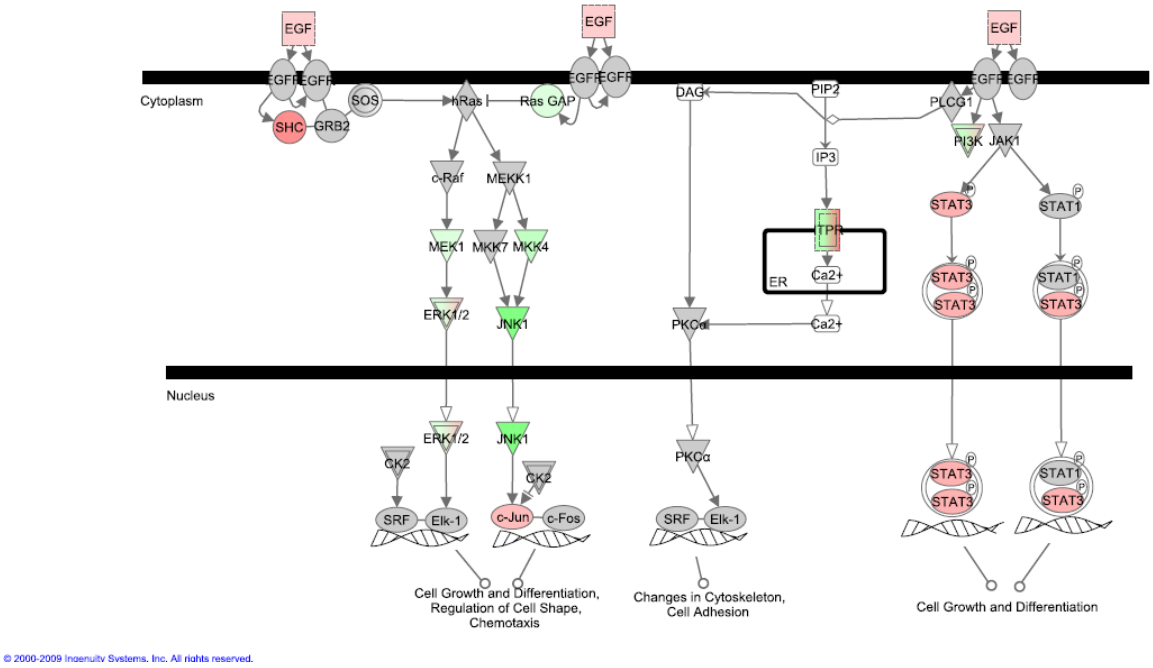

Figure S10. IPA EGFR Canonical Pathway in a Tumor Cell Line with Gain of Chromosome 7 and Amplification in 3q, 4q and 12q (case 15) *vs* Normal Cerebral Cortex. Molecules colored red are overexpressed, molecules colored green are underexpressed, molecules colored grey have no change in gene expression and molecules colored white have no information in regard to gene expression.

Table S4. EGFR Canonical Pathway characterization in Case 15 *vs* Normal Cerebral Cortex.

Cases	**Genes Overexpressed (FC ≥ 2)**		**Genes Underexpressed (FC ≤ -2)**		**Genes with no Change in Expression (-2 ≤ FC ≤ 2)**
Case 15 *vs* N	8/48 (16.6%)	*c-Jun, EGF, ERK5, ITPR3, PIK3CA, PIK3R2, SHC1, STAT3*	11/48 (23%)	*ERK2, ERK3, ITPR1, JNK1, MEK1, MKK4, PIK3C3, PIK3C2B, PIK3CB, PIK3R1, RasGAP*	29/48 (60.4%)