Intra-tumor heterogeneity and transcriptional profiling in glioblastoma: translational opportunities

Sara G.M. Piccirillo[1]*, Inmaculada Spiteri[2]

[1]John van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0PY, tel. 01223-334121, email: sp577@cam.ac.uk, sara.piccirillo@gmail.com

[2] The Institute of Cancer Research, Centre for Evolution and Cancer, 15 Cotswold Road, Sutton SM2 5NG

*Author for correspondence

Abstract

The study of phenotypic and genetic intratumor heterogeneity in glioblastoma is attracting a lot of attention. Recent studies have demonstrated that transcriptional profiling analysis can help interpret the complexity of this disease. Previously proposed molecular classifiers have been recently challenged due to the unexpected degree of intratumor heterogeneity that has been described spatially and at single cell level. Different computational methods have been employed to analyze this huge amount of data, but new experimental designs including multisampling from individual patients and single-cell experiments require new specific approaches. In light of these results, there is hope that integration of genetic, phenotypic and transcriptional data coupled with functional experiments might help define new therapeutic strategies and classify patients according to key pathways and molecular targets that can be further investigated to develop personalized and combinatorial treatment strategies.

Keywords

cancer stem cells, glioblastoma, intratumor heterogeneity, molecular classification, transcriptional

analysis.

Tumor heterogeneity

Cancer is characterized by heterogeneity that is based on extensive phenotypic and genetic diversity between tumors of the same type (Intertumor heterogeneity) as well as among cells within the same tumor (intratumor heterogeneity) [1].

Recently, technological advancements have allowed interrogation of cancer genomes at high resolution [2]. Deep genome-wide analyses of mutations and complex computational approaches have allowed the reconstruction of the genomic history of tumors [3, 4]. This has revealed that tumors of the same histological subtype share only a minority of genetic aberrations thus explaining the differential patterns of response to therapies among patients with the same tumor.

Such complexity is further exacerbated by increasing evidence of genetic diversity among areas of the same tumor or between the tumor and its metastases (spatial intratumor heterogeneity) [5-17], serial sampling taken during tumor evolution and relapse (temporal intratumor heterogeneity) [18-25], and single cells of the same tumor [26-31].

This scenario is even more complicated due to the stochastic and unpredictable behavioral diversity of subclones with similar genotypes as well as the presence of aberrations that occur through post-translational and epigenetic modifications. The tumor microenvironment (including endothelial cells, pericytes, immune inflammatory cells, cancer-activated fibroblasts and stromal cells) adds an extra level of heterogeneity which affects tumor growth, progression and obviously treatment response [32, 33], thus therapeutic targeting should be aimed at several levels to achieve disease eradication.

Intratumor heterogeneity in glioblastoma

Glioblastoma (GBM) is the commonest brain tumor in adults and is characterized by a high degree of heterogeneity both at cellular and molecular level [34, 35]. At clinical presentation it can be classified either as primary or secondary with the former being characterized by absence of clinical or histological evidences of a less malignant tumor and the latter being the result of progression from a low-grade glioma [36].

Historically, primary and secondary GBMs have been considered conceptually different but from a clinical and histopathological perspective they remain largely identical [36]. In the last decade, it has become increasingly clear that primary and secondary GBMs represent different entities characterized by distinct genetic pathways. In this text we refer to primary GBM unless specified otherwise; yet, here we summarize the genetic alterations and pathways involved in the development of both primary and secondary GBMs as a service to the readers.

Primary and secondary GBMs show genetic alterations that are significantly different in frequency: LOH 10q is frequent in both primary and secondary GBMs, whereas epidermal growth factor receptor (*EGFR*) amplification, $p16^{INK4a}$ deletion and phosphatase and tensin homolog (*PTEN*) mutations are more frequent in primary GBMs. Isocitrate dehydrogenase 1 (*IDH1*) mutations occur in approximately 10% of all GBMs [37] and are enriched in secondary GBMs [38]. *p53* mutations are early common genetic events in secondary GBMs [36] but are present also in approximately 30% of primary GBMs. Despite this, more than 70% of primary GBMs show alterations in the p53 pathway [35] .

In addition to the amplification of EGFR (present in almost 50% of GBM [39-41]) and the presence of extrachromosomal double minutes copies of this gene [42], EGFR amplicons are often mutated in several variants. Among these, the most common is a deletion of exon 2-7 which generates a

truncated receptor that is unable to bind its ligand and is constitutively active (EGFRvIII) [36].

Amplification of *EGFR* is frequently associated with deletion of the *INK4a/Arf* gene locus that encodes two tumor suppressors, $p16^{INK4a}$ and $p19^{Arf}$ ($p14^{Arf}$ in humans), translated from alternatively spliced mRNAs [43]. $p16^{INK4a}$ prevents the formation of the complex between cyclin-dependent kinase 4 (CDK4) and cyclin D1, thus preventing G1 entry [44], while $p19^{Arf}$ stabilizes p53 through binding to mouse double minute 2 homolog (MDM2).

The handling of GBM is challenging due to the heterogeneous nature of the disease, invasive potential and poor response to chemo- and radio-therapy. Median life expectancy in optimally managed patients is only 17-62 weeks with only 25% surviving 2 years [45]. The current clinical management of patients diagnosed with a GBM involves a combination of surgery, radiotherapy and chemotherapy. Radiotherapy has been the principal therapeutic modality since the late 1970s and the addition of targeted chemotherapy has only provided modest benefit [45, 46].

The term 'multiforme' that is commonly used to describe GBM clearly summarizes the various morphological characteristics of the tissue and of the cell populations showing heterogeneous tracts in their phenotype and genotype [47]. Despite extensive brain tumor biology studies, the definitive identification of the cell type(s) that contribute to tumor growth has yet to come [48]. This is extremely relevant in order to provide insights into the cellular target(s) of the initial mutational events and to dissect the phenotypic heterogeneity emerging as a result of tumor growth.

Evidences from the cancer stem cell (CSC) hypothesis suggest that tumors hijack the functional properties of somatic stem cells and phenotypic heterogeneity is the result of a hierarchical organization of the tumor with its cell population showing various degrees of tumorigenic

potential regardless of their genotype and epigenetic status [49]. However, the CSC hypothesis does not address the issue of cell of origin [50]. More importantly, the lack of identification of cellular targets and the phenotypic complexity of the disease at the time of clinical presentation have hampered the development of innovative therapeutic approaches aimed at improving patient survival.

Spatial and temporal genetic intratumor heterogeneity in glioblastoma

The most extensive and well-characterized type of heterogeneity in GBM occurs at the genetic level. Initial studies in the 1980s revealed intratumor heterogeneity in GBM tissues and cells [51, 52] and were complemented by more sophisticated analyses using comparative genomic hybridization and laser scanning cytometry of microdissected GBM tissues [53, 54]. Seminal molecular studies in the 1990s shed light on the genetic heterogeneity across all glioma grades [55-57] and identified critical alterations in pathways that were mutually exclusive in GBM (CDKN2A ($p16^{INK4A}$, MTS1) and CDK4 versus RB) suggesting genetic heterogeneity among patients [58, 59].

In the last years, large-screening data have elucidated the role of key signaling pathways in GBM by performing comprehensive analysis and integrating genomic, epigenomic and transcriptional data [60-62]. Focal somatic copy-number alterations found in The Cancer Genome Atlas (TCGA) study are amplification of *EGFR, CDK4, CDK6, PDGFRA, MDM2, MDM4, MET, MYCN, CCND2* and *PIK2CA*. In addition to these well-know aberrations in GBM, less common alterations were also detected, such as the amplification of the serine/threonine protein kinase *AKT3* and the homozygous deletions of *NF1* and *PARK2* [62]. Mutations were also found in *PTEN*, *TP53*, *EGFR*, *PIK3CA*, *PIK3R1*, *NF1*, *RB1*, *IDH1* and *PDGFRA* [60] and an additional 61 mutated genes were

identified at low frequency [60].

This scenario was further complicated by the elegant observation that amplifications of the receptor tyrosine kinases (RTKs; EGFR, PDGFRA and MET) show a mosaic pattern in distinct population of cells coexisting in the same GBMs [10, 13]. In particular, a striking example of RTK genetic intratumor heterogeneity is found in studies focusing on the expression of EGFR wild type and that of its mutant oncogenic variant vIII [63]: while EGFR over-expression is widespread in the majority of GBM, immunostaining for EGFRvIII typically shows that only a minority of the tumor cells are positive for this variant. Interestingly, GBM characterized by EGFR overexpression and presence of EGFRvIII are likely to show ependymal spread of the disease [64]. The interaction between EGFR wild type and vIII as well as the downstream signaling pathway activated by this functional cooperation has been described in a study by Fan and colleagues [65] who showed that EGFR phosphorylates the mutant variant vIII and causes its nuclear translocation. In the nucleus, EGFRvIII forms a nuclear complex with STAT3 and is able to drive transformation [65]. The coexpression of EGFR wild type and vIII drives the phosphorylation of STAT *in vitro* and *in vivo* and enhances tumorigenicity providing further insights into the oncogenic cooperation among variants of the same genetic lesion. A more recent study has suggested that an additional mechanism involving the activation of MET operates between EGFR wild type and vIII [66]. All multiple somatic events and the heterogeneous expression of EGFR and its variants in GBM have been comprehensively resolved at single-cell level [67]. Mutations in a single RTK, like *EGFR*, contribute to enhance clonal diversification which further expands upon the recent observation of intratumor heterogeneity of multiple concurrent RTK amplifications [10, 13].

Most recent studies have revealed spatial and temporal intratumor heterogeneity [5, 8, 26]. Using a fluorescence-guided resection method based on 5-ALA [68, 69] we have been able to

dissect distinct regions of the same GBM, starting from the residual disease (margin) that is left behind during surgery [68] to the resected disease (tumor mass) [8]. This has allowed us to describe the clonal evolutionary trajectories of GBM development and to show that *EGFR* amplification and *CDKN2A/B/p14ARF* deletion represent early events whereas *PDGFRA* amplification and *PTEN* deletion emerge later during GBM evolution.

An even more interesting observation of spatial intratumor heterogeneity came from the histological and genomic analysis of the subependymal zone (SEZ), the most well characterized neurogenic region of the human brain. Our recent study revealed that this region contributes to the genetic intratumor heterogeneity of GBM and is characterized by patterns of drug response that differ from the tumor mass [5]. Phylogenetic reconstruction showed different GBM evolution patterns in the analyzed patients where tumor cells from the SEZ correspond to clones that are generated early during tumor growth [5], in agreement with what proposed in mouse model studies [70-75].

The use of patient-derived xenografts (PDX) has become increasingly popular in recent years, as they are believed to faithfully mimic the original human disease. In GBM, it has been shown that the 'avatar' model based on PDX recapitulates the pathological and genomic features of the original tumor and can be used to identify key regulatory signatures of clinical aggressiveness in patients [76] and to test novel therapeutic agents [77]. Very recently, we have exploited this model to reveal the genetic complexity during GBM evolution by identifying the clones that possess the tumor-initiating capacity and are able to transfer the disease to immunosuppressed animals [26]. This study represents an initial step into the complexity of GBM temporal evolution that is further complicated by the influence of the tumor microenvironment and the application of radical treatments, i.e. chemotherapy and radiotherapy, which impact on the clonal architecture

of the tumor. There are many examples of drug resistance conferred by the emergence of subclones harbouring specific somatic mutations [78] that might be therapy-driven.

This has been proven true for GBM that arises as a result of progression from a low-grade disease [19] but evidences in primary GBM are yet to come.

The influence of epigenetics

The most challenging type of heterogeneity occurs at the environmental level as this is extrinsically dynamic and multifactorial depending on different types of stromal cells, their interactions and different niches types (vascular, hypoxic, etc.) [79, 80]. The cellular microenvironment impacts the epigenetic status of a cell and the downstream epigenetic marks (methylation of cytosine nucleotides, histone protein modifications, chromatin remodeling complexes and noncoding RNA interference) will influence the transcriptional potential of the cell. Genomic hyper- and hypomethylation has been shown in many cancer types and is associated to tumor formation and progression. As DNA methylation occurs primarily at CG dinucleotides in mammals, many cancer studies have targeted CG-rich regions and analyzed their aberrant 'methylomes' [81-83]. However, in recent years, non-CpG methylation has been found to regulate gene expression in embryonic stem cells and new research has focused on analysing the distribution and regulation of non-CpG methylation traits [84-86]. Genomic methylation patterns can vary among individual cells within the same tissue and are subject to dynamic changes during development [87, 88]. This epigenetic heterogeneity has also been described within tumors where only a subset of aberrant DNA methylation events are recurrently found across different tumor regions [89-91].

In patients with GBM the oral alkylating agent temozolomide is the current standard of

care. This drug is more effective on tumors where there is epigenetic silencing of the O-6-methylguanine-DNA methyltransferase (*MGMT)* gene encoding a methyltransferase that inhibits the cytotoxic effect of temozolomide. Two regions in the *MGMT* promoter have been recently identified as critical for transcriptional silencing and prediction of the response to alkylating agents [92]. Interestingly, hypermethylation of the *MGMT* promoter has been observed in all the subtypes of the Verhaak classification [34]. Recently, we have shown that *MGMT* methylation status seems to be quite homogeneous in different GBM tumor areas [8] and generally predicts response to temozolomide [5]. Notwithstanding, beneficial response to this drug has also been seen in *MGMT*-unmethylated tumors, which suggests that *MGMT* methylation changes over time [93] and is influenced by treatment [94] reflecting a temporal rather than spatial heterogeneous status.

Several groups have described additional promoter hypermethylation in GBM, often influencing the expression of GBM tumor suppressor genes like *PTEN* and *P53*, as well as of regulatory genes involved in cell proliferation, invasion and cytokine signaling [95] . Also the expression of EGFRvIII has been shown to be epigenetically modulated through methylation and hystone acetylation [96]. Epigenomic profiling in GBM has revealed a distinct subgroup called glioma CpG island methylator phenotype (G-CIMP) [97] that is enriched for the proneural expression group. If analyzed in the context of the proneural subtype, the G-CIMP group is clinically characterized by longer survival and it is associated with *IDH1* mutations [95, 97]. Interestingly, the *MGMT* promoter methylation is independent from the CpG island methylator phenotype.

These findings were followed by a comprehensive epigenetic and biological classification of pediatric and adult GBM [98] that defined six GBM subtypes based on global DNA methylation patterns, *IDH1* mutations and two mutations in aminoacids K27 and G34 of the histone variant

H3.3. Interestingly, the authors found that the anatomical localization of K27 and G34 mutant tumors was different (midline locations and embryonic regions of neocortex and striatum, respectively) and that these mutations occurred mainly in pediatric GBM [98]. The molecular mechanisms underlying the transcriptional signature of K27 mutant tumors were described in a subsequent study showing that the reduction of K27 trimethylation on histone H3 interferes with the enzymatic activity of histone methyltransferase EZH2. This causes alterations in H3K27me3 occupancy at the transcription start site of promoters across the whole genome of K27 mutant tumors. As a result, a pronounced global DNA hypomethylation resetting drives the emergence of a K27 mutant-specific gene expression profile [99].

Epigenetic modulators are currently being targeted as druggable candidates in several cancers. Epigenetic therapy has been applied with reasonable success in hematologic malignancies using DNA methyltransferase (decitabine, 5-azacitidine) and histone deacetylase inhibitors (vorinostat, romidepsin). In contrast, relatively low efficacies have been achieved on solid tumors; this is likely due to their intrinsic heterogeneous nature which allows tumors to develop treatment resistance [100]. A fundamental characteristic of epigenetics is that the same genomic sequence can show alternative phenotypes by reprogramming gene expression. Thus, intratumor heterogeneity at the epigenetic level will be reflected into transcriptional variability within tumors which will affect its evolutionary potential and is of vital importance on future clinical decision making.

Transcriptional analysis

In a critical review of the pathology of cerebral gliomas in 1940, Scherer H.J. described that primary GBMs are constituted of at least two or more “distinct pathological entities” [101]. Transcriptional analysis in GBM has been extensively used with the aim to dissect the phenotypic

complexity well known already by pathologists almost 80 years ago. Following from gene expression profiling studies on oligodendrogiomas, astrocytomas and normal brain tissues, in 2005 complementary DNA (cDNA) microarrays were used to analyze GBM specimens and identify genes associated with immune cells, hypoxia, extracellular matrix and cellular proliferation [102].

With the initial evidences of the CSC hypothesis applied to solid tumors, transcriptional profiling analysis has helped to identify pathways which might suggest normal stem cell parallels for cancer cells. In GBM, seminal studies have shown that tumor cells retain a stem cell signature [68, 103-107] and a lot of effort has been put in optimizing culture conditions that preserve the genetic and transcriptional profile of the original tumor. The ability of GBM tumor cells to mimic the functional properties of stem cells has been definitively described in a recent study by Suvà and colleagues [108] where transcriptional analysis has been fundamental to identify a set of critical neurodevelopmental transcription factors for GBM propagation (POU3F2 [BRN2], SOX2, SALL2, and OLIG2).

In an initial study including all grades of astrocytic tumors, an artificial neural network algorithm was developed to define survival prognostic subtypes [109] (denoted ANGIO, INTER, and LOWER) that had a perfect correspondence with the histological grade of these tumors.

More recently, several groups have focused on developing accurate molecular prognostic classification systems to specifically stratify GBM patients. To this aim, large cohort genome-wide profiling analyses [110-112] have attempted to use microarray gene expression data from hundreds of GBM samples. The largest analysis to date [110] identified four GBM subtypes: Proneural, Neural, Classical and Mesenchymal, each characterized by a distinct gene expression profile associated with prognosis and response to therapy. The Proneural subtype is characterized by aberrations in *PDGFRA*, *p53* and *IDH1* mutations. The Classical subtype presents all common

alterations of GBM (chromosome 7 amplification, chromosome 10 deletion, *CDKN2A/B/p14ARF* deletion and *EGFR* amplification). The Mesenchymal subtype shows high expression of *CHI3L1* (also known as *YKL40*) and *MET* and is characterized by *NF1* deletion (that is sufficient to convert a Proneural GBM into a Mesenchymal GBM both *in vitro* and *in vivo* [113]). The Neural subtype has the highest expression of neuronal markers and genes associated with axon and synaptic transmission and it is typical of samples characterized mainly by normal tissue. Interestingly, the Neural subtype has been found to be characteristic of the margin area of GBM which can be identified by using 5-ALA during surgical tumor debulking [68] (Piccirillo SGM et al., unpublished data).

The clinical relevance of this classification relies on its association to survival and response to treatment that differs among subtypes: the Proneural subtype is associated with a trend toward longer survival but the better response to treatment is observed in the Classical subtype [110]. More recently, a different approach based on gene co-expression modules of key pathways, such as EGFR and PDGFR, has been explored to develop a clinically relevant classifier for both adult low-grade and high-grade gliomas [114].

However, in the last years it became clear that given the extensive spatial genetic intratumor heterogeneity in GBM it is difficult to make sense of these data based on a single GBM sample per patient (Fig.1). Also, such classification does not segregate all the GBM key aberrations, in fact some are found in more than one subtype, in other words *CDKN2A* and *CDKN2B*, and are shared with low-grade gliomas [34] suggesting that they are essential for the stepwise malignant growth.

More recently, transcriptome sequencing has been employed to reveal gene fusions in GBM [61] some of which occur in noncoding genes, resulting in the expression of noncoding RNAs that are not expressed in normal cells [115].

In colorectal cancer, single-cell PCR gene expression analysis has been used to dissect the cellular composition of normal and cancer tissue [116]. In GBM, single-cell transcriptome analysis by RNA sequencing has revealed extensive intratumor heterogeneity and a “stem cell-differentiation gradient” that is not fully recapitulated in *in vitro* models [27]. More importantly, single-cell analysis has confirmed that more than one sub-type coexists in the same tumor as previously reported [8] (Fig.1) and suggested that hybrid phenotypic states reminiscent of progenitor and differentiated cells are present in individual cells of the same GBM [27]. These data highlight the dynamics of cellular states in GBM that can have implications for developing new therapeutic approaches. An example of this dynamic regulation has been recently provided in a study about extrachromosomal EGFRvIII showing that GBM cells suppress the expression of EGFRvIII following treatment with tyrosine kinase inhibitors but is upregulated after drug removal [117].

In the last decade, thousands of computational methods and bioinformatics tools have been developed to analyze gene expression patterns initially derived from gene expression microarray measurements and more recently from RNA-sequencing data. Most of these bioinformatics tools have used empirical Bayes methods [118-120] or estimates of the mean and variance [121, 122] to analyze differential expression levels. These tools have been applied to quantify transcription levels among single biopsies from several cancer patients or between cancer patients and controls.

New studies that consider the variability in intratumor expression levels include multiple samples from a single patient and/or multiple single cells from the same tumor. These studies aim at gaining a better understanding of regional and temporal transcriptional profiles within a tumor. However, new dependencies emerge from the measurements (mRNA expression levels for each

patient/tumor) taken on this type of experiments as neither the profiled genes nor the tissue samples/single cells are likely to be independent from a statistical point of view. These high-dimensional, complex gene expression datasets lead to important statistical challenges. Firstly, the higher number of genes compared to the number of samples limits the application of multivariate tests. Secondly, the dependence among the tissue samples for each subject might restrict the use of practical approaches that rely on mixing univariate standard testing procedures such as ANOVA and multiple testing correction methods. An important factor for accurate identification of differentially expressed genes is the number of sample replicates (biological or technical) as most methods model the variability in gene expression measurements. However, in light of the extensive intratumor heterogeneity, biological replicates are not really adequate to study cancer datasets and running technical replicates of tumor samples is often difficult due to limiting starting material or elevated cost. Few new methods have been developed to overcome these limitations but an increasingly number of multisampling experiments are currently being processed. If not addressed, these issues will limit the potential of the multisampling and single-cell experiments and ultimately compromise the identification of potential prognostic/predictive signatures to develop better therapeutic approaches. Moreover, these problems would also affect the downstream functional analyses that investigate the pathways in which dis-regulated genes are involved. A recent publication by Touloumis and colleagues [123] describes a suitable method for both, studying the heterogeneous transcriptional levels within tumors and identifying those pathways or biological processes affected by genes that are heterogeneously expressed in different tumor regions. In this study, the authors found differential expression levels between three spatially separated regions of GBM tumors: the tumor mass, the SEZ and the tumor margin [5, 68]. Furthermore, when genes were classified into one of the Gene Ontology biological

processes they found marked functional heterogeneity as most of the analyzed groups harbor differentially expressed genes.

Computational approaches that allow integration of transcriptomic, genomic and epigenomic data will be key to achieve a comprehensive understanding of intratumor heterogeneity. These methods will lead to new 'heterogeneity signatures' that will aid with clinical decision making to apply individualised treatments to cancer patients.

Conclusion

Intratumor heterogeneity highlights the complexity of tumor genomic landscape and may present major challenges to biomarker identification and overcoming treatment resistance. The existence of treatment resistant clones and the influence of microenvironment favour tumor relapse. In GBM initial insights into clonal distribution [5, 8, 26, 27] and cooperativity among EGFR variants [65, 124] have shed new light into the mechanisms underlying tumor heterogeneity. At trascriptional level, recent studies have shown a previously unrecognised heterogeneous profile in distinct region of the same GBMs as well as at single-cell level [5, 8, 27] (Fig. 1). These findings support the concept that a detailed molecular inter- and intra-patient analysis is needed to develop personalized therapies and eventually target regions of the same GBM with different therapeutic approaches.

By comparing the SEZ and the tumor mass of the same patient, we have observed differential pattern of response to the standard treatment based on the alkylating agent temozolomide, cisplatin and cediranib [5]. Similar analyses on primary cells derived from different regions of the same GBM might inform therapeutic decision.

Molecular tumor subtypes identified through transcriptional analysis have been proposed to be

clinically relevant and important for prognostic value, however these classifications did not take into account sample heterogeneity. Analysis of primary tumor cells and single-cell technologies might help refining the existing GBM subtype classification and improving patient stratification. Unfortunately, there is currently no approved therapy besides the Stupp protocol [45, 46], thus clinicians have no alternatives to offer after exhausting the traditional treatments (surgery, radiation, alkylating chemotherapy and eventually anti-angiogenetic therapy).

Hence, performing exhaustive genetic, transcriptomic and epigenetic testing seem invaluable to provide insights into more personalized treatments and/or further therapeutic alternatives for GBM patients. Nevertheless, these molecular-based tests are certainly time consuming and the integration of the data is not trivial which makes the interpretation of the results quite complex.

Future perspective

With the emergence of next-generation sequencing technologies, we have changed our understanding of tumor heterogeneity. These technologies have revealed an unexpected molecular complexity from bulk population to single-cell level and have challenged the way by which we think of cancer development and treatment.

GBM is a remarkable example of a phenotypic and genetic heterogeneous tumor and despite the challenges associated with improving current treatments, there is hope that the integration of accurate data reflecting the genetic, transcriptional and epigenetic intratumor heterogeneity in addition to functional studies may lead to the definitive identification of tumor drivers and druggable therapeutic targets (Fig. 2).

Executive summary

- Glioblastoma (GBM) is the most common and aggressive brain cancer in adults.
- The prognosis of GBM remains dismal despite the current treatments based on surgical resection and combination of chemo- and radio-therapy.
- Cancer stem cells have been isolated in GBM and their identification has helped define a model of phenotypic intratumor heterogeneity based on a hierarchical organization of cells with different tumorigenic potential.
- Genetic and phenotypic intratumor heterogeneity has been described in GBM in seminal studies on RTK expression, phylogenetic reconstruction, high-throughput single-cell analysis and differential response to drug treatment.
- Multiple-sampling and single-cell analysis have revealed different patterns of transcriptional profiling in the same GBM.
- Many computational methods have been developed to interpret tumor transcriptional profiles yet, robust methods that are suitable for multi-sampling/single-cell experimental designs are much needed.
- Bioinformatics tools that allow integration of transcriptomic, genomic and epigenetic data will be key to achieve a comprehensive understanding of intratumor heterogeneity that will lead to better classification of patients and aid to define novel therapeutic strategies important for clinical decision making.

Conclusion

- Transcriptional profiling needs to be further investigated in light of recent data of intratumor heterogeneity in GBM.
- A comprehensive analysis of distinct tumor areas and single cells can shed new light on the molecular mechanisms sustaining tumor growth and develop more refined molecular

classifiers of GBM patients.

Future perspective

- Integration of genetic, phenotypic and transcriptional data with functional experiments can help design new molecular patient-specific therapeutic approaches.

Figure legends

Fig. 1. Transcriptional profiling of human GBM can be performed using: a single tumor sample, multiple samples taken from distinct tumor areas (5 samples from the tumor mass in the example shown consistently with Sottoriva&Spiteri and colleagues [8]) or at single-cell level (430 cells/GBM as described by Patel and colleagues [27]). The data generated by these 3 different approaches can be used to investigate intra-tumor heterogeneity at transcriptional level and potentially gain new insights into GBM biology and prognosis. By applying a previously published classifier, the multi-sampling scheme and single-cell analysis have revealed that multiple subtypes coexist in the same tumor with implications for patients stratification and response to standard therapy. Pron=Proneural, Mes= Mesenchymal, Cla= Classical.

Fig. 2. Integration of spatial genetic, phenotypic, epigenetic and transcriptional intra-tumor heterogeneity with temporal analysis before and after treatment and functional *in vitro* and *in vivo* assays might lead to an in-depth molecular classification and to the development of personalized therapies aimed at improving patients survival.

Financial disclosure

The authors have no financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript.

No writing assistance was used in the production of this manuscript.

References

***** = of interest

****** = of considerable interest

1. Marusyk A, Almendro V, Polyak K: Intra-tumour heterogeneity: a looking glass for cancer? *Nat Rev Cancer* 12(5), 323-334 (2012).

2. Yates LR, Campbell PJ: Evolution of the cancer genome. *Nat Rev Genet* 13(11), 795-806 (2012).

3. Nik-Zainal S, Van Loo P, Wedge DC *et al.*: The life history of 21 breast cancers. *Cell* 149(5), 994-1007 (2012).

4. Campbell PJ, Pleasance ED, Stephens PJ *et al.*: Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc Natl Acad Sci U S A* 105(35), 13081-13086 (2008).

5. Piccirillo SGM, Spiteri I, Sottoriva A *et al.*: Contributions to drug resistance in glioblastoma derived from malignant cells in the sub-ependymal zone. *Cancer Res* 75(1), 194-202 (2015).

   **** Shows that a neurogenic region in the brain, the subependymal zone, is a reservoir of tumor cells in glioblastoma patients and emphasizes that these cells play a role into tumor growth dynamics and drug resistance.**

6. De Bruin EC, Mcgranahan N, Mitter R *et al.*: Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 346(6206), 251-256 (2014).

7. Gerlinger M, Horswell S, Larkin J *et al.*: Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet* 46(3), 225-233 (2014).

8. Sottoriva A, Spiteri I, Piccirillo SGM *et al.*: Intratumor heterogeneity in human glioblastoma

reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A* 110(10), 4009-4014 (2013).

**** Describes an integrated genomic analysis and phylogenetic reconstruction of 11 glioblastoma using a multiple-sampling approach and shows that multiple subtypes are present in the same tumor.**

9. Gerlinger M, Quezada SA, Peggs KS *et al.*: Ultra-deep T cell receptor sequencing reveals the complexity and intratumour heterogeneity of T cell clones in renal cell carcinomas. *J Pathol* 231(4), 424-432 (2013).
10. Szerlip NJ, Pedraza A, Chakravarty D *et al.*: Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. *Proc Natl Acad Sci U S A* 109(8), 3041-3046 (2012).
11. Gerlinger M, Rowan AJ, Horswell S *et al.*: Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366(10), 883-892 (2012).
12. Hernandez L, Wilkerson PM, Lambros MB *et al.*: Genomic and mutational profiling of ductal carcinomas in situ and matched adjacent invasive breast cancers reveals intra-tumour genetic heterogeneity and clonal selection. *J Pathol* 227(1), 42-52 (2012).
13. Snuderl M, Fazlollahi L, Le LP *et al.*: Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. *Cancer Cell* 20(6), 810-817 (2011).
14. Campbell PJ, Yachida S, Mudie LJ *et al.*: The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* 467(7319), 1109-1113 (2010).
15. Yachida S, Jones S, Bozic I *et al.*: Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* 467(7319), 1114-1117 (2010).
16. Ding L, Ellis MJ, Li S *et al.*: Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 464(7291), 999-1005 (2010).

17. Khalique L, Ayhan A, Weale ME, Jacobs IJ, Ramus SJ, Gayther SA: Genetic intra-tumour heterogeneity in epithelial ovarian cancer and its implications for molecular diagnosis of tumours. *J Pathol* 211(3), 286-295 (2007).

18. Bolli N, Avet-Loiseau H, Wedge DC *et al.*: Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun* 5, 2997 (2014).

19. Johnson BE, Mazor T, Hong C *et al.*: Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* 343(6167), 189-193 (2014).

20. Ding L, Ley TJ, Larson DE *et al.*: Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481(7382), 506-510 (2012).

21. Keats JJ, Chesi M, Egan JB *et al.*: Clonal competition with alternating dominance in multiple myeloma. *Blood* 120(5), 1067-1076 (2012).

22. Cooke SL, Temple J, Macarthur S *et al.*: Intra-tumour genetic heterogeneity and poor chemoradiotherapy response in cervical cancer. *Br J Cancer* 104(2), 361-368 (2011).

23. Shah SP, Morin RD, Khattra J *et al.*: Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* 461(7265), 809-813 (2009).

24. Mullighan CG, Phillips LA, Su X *et al.*: Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* 322(5906), 1377-1380 (2008).

25. Shah NP, Nicoll JM, Nagar B *et al.*: Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2(2), 117-125 (2002).

26. Piccirillo SGM, Colman S, Potter NE *et al.*: Genetic and functional diversity of propagating cells in glioblastoma. *Stem Cell Reports* 4(1), 7-15 (2015).

27. Patel AP, Tirosh I, Trombetta JJ *et al.*: Single-cell RNA-seq highlights intratumoral

heterogeneity in primary glioblastoma. *Science* 344(6190), 1396-1401 (2014).

**** Describes intratumor heterogeneity in glioblastoma at the highest possible resolution to date by using single-cell RNA sequencing to profile approximately 450 tumor cells from five human glioblastoma.**

28. Kreso A, O'brien CA, Van Galen P *et al.*: Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* 339(6119), 543-548 (2013).
29. Navin N, Kendall J, Troge J *et al.*: Tumour evolution inferred by single-cell sequencing. *Nature* 472(7341), 90-94 (2011).
30. Anderson K, Lutz C, Van Delft FW *et al.*: Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* 469(7330), 356-361 (2011).
31. Notta F, Mullighan CG, Wang JC *et al.*: Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. *Nature* 469(7330), 362-367 (2011).
32. Bourzac K: Biology: Three known unknowns. *Nature* 509(7502), S69-71 (2014).
33. Junttila MR, De Sauvage FJ: Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* 501(7467), 346-354 (2013).
34. Goodenberger ML, Jenkins RB: Genetics of adult glioma. *Cancer Genet* 205(12), 613-621 (2012).
35. Collins VP: Mechanisms of disease: genetic predictors of response to treatment in brain tumors. *Nat Clin Pract Oncol* 4(6), 362-374 (2007).
36. Ohgaki H, Kleihues P: Genetic pathways to primary and secondary glioblastoma. *Am J Pathol* 170(5), 1445-1453 (2007).
37. Parsons DW, Jones S, Zhang X *et al.*: An integrated genomic analysis of human glioblastoma multiforme. *Science* 321(5897), 1807-1812 (2008).

38. Yan H, Parsons DW, Jin G *et al.*: IDH1 and IDH2 mutations in gliomas. *N Engl J Med* 360(8), 765-773 (2009).

39. Schlegel J, Merdes A, Stumm G *et al.*: Amplification of the epidermal-growth-factor-receptor gene correlates with different growth behaviour in human glioblastoma. *Int J Cancer* 56(1), 72-77 (1994).

40. Hurtt MR, Moossy J, Donovan-Peluso M, Locker J: Amplification of epidermal growth factor receptor gene in gliomas: histopathology and prognosis. *J Neuropathol Exp Neurol* 51(1), 84-90 (1992).

41. Jaros E, Perry RH, Adam L *et al.*: Prognostic implications of p53 protein, epidermal growth factor receptor, and Ki-67 labelling in brain tumours. *Br J Cancer* 66(2), 373-385 (1992).

42. Vogt N, Lefevre SH, Apiou F *et al.*: Molecular structure of double-minute chromosomes bearing amplified copies of the epidermal growth factor receptor gene in gliomas. *Proc Natl Acad Sci U S A* 101(31), 11368-11373 (2004).

43. Quelle DE, Zindy F, Ashmun RA, Sherr CJ: Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. *Cell* 83(6), 993-1000 (1995).

44. Serrano M, Hannon GJ, Beach D: A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature* 366(6456), 704-707 (1993).

45. Stupp R, Mason WP, Van Den Bent MJ *et al.*: Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* 352(10), 987-996 (2005).

46. Stewart LA: Chemotherapy in adult high-grade glioma: a systematic review and meta-analysis of individual patient data from 12 randomised trials. *Lancet* 359(9311), 1011-1018 (2002).

47. Kleihues P, Sobin LH: World Health Organization classification of tumors. *Cancer* 88(12), 2887 (2000).

48. Nabors LB: Targeted molecular therapy for malignant gliomas. *Curr Treat Options Oncol* 5(6), 519-526 (2004).

49. Magee JA, Piskounova E, Morrison SJ: Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer Cell* 21(3), 283-296 (2012).

50. Visvader JE: Cells of origin in cancer. *Nature* 469(7330), 314-322 (2011).

51. Wikstrand CJ, Bigner SH, Bigner DD: Demonstration of complex antigenic heterogeneity in a human glioma cell line and eight derived clones by specific monoclonal antibodies. *Cancer Res* 43(7), 3327-3334 (1983).

52. Shapiro JR, Yung WK, Shapiro WR: Isolation, karyotype, and clonal growth of heterogeneous subpopulations of human malignant gliomas. *Cancer Res* 41(6), 2349-2359 (1981).

53. Jung V, Romeike BF, Henn W *et al.*: Evidence of focal genetic microheterogeneity in glioblastoma multiforme by area-specific CGH on microdissected tumor cells. *J Neuropathol Exp Neurol* 58(9), 993-999 (1999).

54. Harada K, Nishizaki T, Ozaki S, Kubota H, Ito H, Sasaki K: Intratumoral cytogenetic heterogeneity detected by comparative genomic hybridization and laser scanning cytometry in human gliomas. *Cancer Res* 58(20), 4694-4700 (1998).

55. Ichimura K, Bolin MB, Goike HM, Schmidt EE, Moshref A, Collins VP: Deregulation of the p14ARF/MDM2/p53 pathway is a prerequisite for human astrocytic gliomas with G1-S transition control gene abnormalities. *Cancer Res* 60(2), 417-424 (2000).

56. Schmidt EE, Ichimura K, Messerle KR, Goike HM, Collins VP: Infrequent methylation of

CDKN2A(MTS1/p16) and rare mutation of both CDKN2A and CDKN2B(MTS2/p15) in primary astrocytic tumours. *Br J Cancer* 75(1), 2-8 (1997).

57. Schmidt EE, Ichimura K, Reifenberger G, Collins VP: CDKN2 (p16/MTS1) gene deletion or CDK4 amplification occurs in the majority of glioblastomas. *Cancer Res* 54(24), 6321-6324 (1994).

58. Ueki K, Ono Y, Henson JW, Efird JT, Von Deimling A, Louis DN: CDKN2/p16 or RB alterations occur in the majority of glioblastomas and are inversely correlated. *Cancer Res* 56(1), 150-153 (1996).

59. Ichimura K, Schmidt EE, Goike HM, Collins VP: Human glioblastomas with no alterations of the CDKN2A (p16INK4A, MTS1) and CDK4 genes have frequent mutations of the retinoblastoma gene. *Oncogene* 13(5), 1065-1072 (1996).

60. Brennan CW, Verhaak RG, Mckenna A *et al.*: The somatic genomic landscape of glioblastoma. *Cell* 155(2), 462-477 (2013).

**** Presents the most recent TCGA data comprising 578 glioblastoma patients from 17 centers. These data are based on integration of exome-seq, RNA-seq, copy-number, DNA methylation, protein, mRNA and miRNA profiles.**

61. Frattini V, Trifonov V, Chan JM *et al.*: The integrated landscape of driver genomic alterations in glioblastoma. *Nat Genet* 45(10), 1141-1149 (2013).

62. Cancer Genome Atlas Research Network: Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455(7216), 1061-1068 (2008).

63. Inda MM, Bonavia R, Mukasa A *et al.*: Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. *Genes Dev* 24(16), 1731-1745 (2010).

64. Heimberger AB, Suki D, Yang D, Shi W, Aldape K: The natural history of EGFR and EGFRvIII in glioblastoma patients. *J Transl Med* 3, 38 (2005).

65. Fan QW, Cheng CK, Gustafson WC *et al.*: EGFR phosphorylates tumor-derived EGFRvIII driving STAT3/5 and progression in glioblastoma. *Cancer Cell* 24(4), 438-449 (2013).

*** Shows the cell-intrinsic interaction between two key drivers of glioblastoma growth, EGFR wild type and its variant vIII, and highlights the molecular mechanism downstream of their oncogenic cooperativity.**

66. Li L, Puliyappadamba VT, Chakraborty S *et al.*: EGFR wild type antagonizes EGFRvIII-mediated activation of Met in glioblastoma. *Oncogene*, (2013).

67. Francis JM, Zhang CZ, Maire CL *et al.*: EGFR variant heterogeneity in glioblastoma resolved through single-nucleus sequencing. *Cancer Discov* 4(8), 956-971 (2014).

68. Piccirillo SGM, Dietz S, Madhu B *et al.*: Fluorescence-guided surgical sampling of glioblastoma identifies phenotypically distinct tumour-initiating cell populations in the tumour mass and margin. *Br J Cancer* 107(3), 462-468 (2012).

**** Highlights the use of 5-ALA technology to interrogate distinct tumor regions of glioblastoma.**

69. Stummer W, Pichlmeier U, Meinel T, Wiestler OD, Zanella F, Reulen HJ: Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. *The lancet oncology* 7(5), 392-401 (2006).

70. Chen J, Mckay RM, Parada LF: Malignant glioma: lessons from genomics, mouse models, and stem cells. *Cell* 149(1), 36-47 (2012).

71. Chen J, Li Y, Yu TS *et al.*: A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* 488(7412), 522-526 (2012).

72. Alcantara Llaguno S, Chen J, Kwon CH *et al.*: Malignant astrocytomas originate from neural stem/progenitor cells in a somatic tumor suppressor mouse model. *Cancer Cell* 15(1), 45-56 (2009).

73. Wang Y, Yang J, Zheng H *et al.*: Expression of mutant p53 proteins implicates a lineage relationship between neural stem cells and malignant astrocytic glioma in a murine model. *Cancer Cell* 15(6), 514-526 (2009).

74. Zheng H, Ying H, Yan H *et al.*: p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. *Nature* 455(7216), 1129-1133 (2008).

75. Holland EC, Celestino J, Dai C, Schaefer L, Sawaya RE, Fuller GN: Combined activation of Ras and Akt in neural progenitors induces glioblastoma formation in mice. *Nat Genet* 25(1), 55-57 (2000).

76. Joo KM, Kim J, Jin J *et al.*: Patient-specific orthotopic glioblastoma xenograft models recapitulate the histopathology and biology of human glioblastomas in situ. *Cell Rep* 3(1), 260-273 (2013).

77. Piccirillo SG, Reynolds BA, Zanetti N *et al.*: Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. *Nature* 444(7120), 761-765 (2006).

78. Fisher R, Pusztai L, Swanton C: Cancer heterogeneity: implications for targeted therapeutics. *Br J Cancer* 108(3), 479-485 (2013).

79. Hanahan D, Coussens LM: Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* 21(3), 309-322 (2012).

80. Hanahan D, Weinberg RA: Hallmarks of cancer: the next generation. *Cell* 144(5), 646-674 (2011).

81. Kostareli E, Holzinger D, Bogatyrova O *et al.*: HPV-related methylation signature predicts survival in oropharyngeal squamous cell carcinomas. *J Clin Invest* 123(6), 2488-2501 (2013).

82. Szyf M: DNA methylation signatures for breast cancer classification and prognosis. *Genome Med* 4(3), 26 (2012).

83. Rauch TA, Wang Z, Wu X, Kernstine KH, Riggs AD, Pfeifer GP: DNA methylation biomarkers for lung cancer. *Tumour Biol* 33(2), 287-296 (2012).

84. Guo JU, Su Y, Shin JH *et al.*: Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci* 17(2), 215-222 (2014).

85. Truong M, Yang B, Wagner J, Desotelle J, Jarrard DF: Analysis of promoter non-CG methylation in prostate cancer. *Epigenomics* 5(1), 65-71 (2013).

86. Ziller MJ, Muller F, Liao J *et al.*: Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet* 7(12), e1002389 (2011).

87. Tanasijevic B, Dai B, Ezashi T, Livingston K, Roberts RM, Rasmussen TP: Progressive accumulation of epigenetic heterogeneity during human ES cell culture. *Epigenetics* 4(5), 330-338 (2009).

88. Stoger R, Kajimura TM, Brown WT, Laird CD: Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene FMR1. *Hum Mol Genet* 6(11), 1791-1801 (1997).

89. Easwaran H, Tsai HC, Baylin SB: Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. *Mol Cell* 54(5), 716-727 (2014).

90. Smallwood SA, Lee HJ, Angermueller C *et al.*: Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods* 11(8), 817-820 (2014).

91. Brocks D, Assenov Y, Minner S *et al.*: Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Rep* 8(3), 798-806 (2014).

92. Malley DS, Hamoudi RA, Kocialkowski S, Pearson DM, Collins VP, Ichimura K: A distinct region of the MGMT CpG island critical for transcriptional regulation is preferentially methylated in glioblastoma cells and xenografts. *Acta Neuropathol* 121(5), 651-661 (2011).

93. Brandes AA, Franceschi E, Tosoni A *et al.*: O(6)-methylguanine DNA-methyltransferase methylation status can change between first surgery for newly diagnosed glioblastoma and second surgery for recurrence: clinical implications. *Neuro Oncol* 12(3), 283-288 (2010).

94. Parkinson JF, Wheeler HR, Clarkson A *et al.*: Variation of O(6)-methylguanine-DNA methyltransferase (MGMT) promoter methylation in serial samples in glioblastoma. *J Neurooncol* 87(1), 71-78 (2008).

95. Sturm D, Bender S, Jones DT *et al.*: Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. *Nat Rev Cancer* 14(2), 92-107 (2014).

*** Highlights the genetic and epigenetic hallmarks of pediatric and adult glioblastoma.**

96. Del Vecchio CA, Giacomini CP, Vogel H *et al.*: EGFRvIII gene rearrangement is an early event in glioblastoma tumorigenesis and expression defines a hierarchy modulated by epigenetic mechanisms. *Oncogene* 32(21), 2670-2681 (2013).

97. Noushmehr H, Weisenberger DJ, Diefes K *et al.*: Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* 17(5), 510-522 (2010).

98. Sturm D, Witt H, Hovestadt V *et al.*: Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* 22(4), 425-437 (2012).

99. Bender S, Tang Y, Lindroth AM *et al.*: Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. *Cancer Cell* 24(5), 660-672 (2013).

100. Ho AS, Turcan S, Chan TA: Epigenetic therapy: use of agents targeting deacetylation and

methylation in cancer management. *Onco Targets Ther* 6, 223-232 (2013).

101. Scherer HJ: A Critical Review: The Pathology of Cerebral Gliomas. *J Neurol Psychiatry* 3(2), 147-177 (1940).

102. Liang Y, Diehn M, Watson N *et al.*: Gene expression profiling reveals molecularly and clinically distinct subtypes of glioblastoma multiforme. *Proc Natl Acad Sci U S A* 102(16), 5814-5819 (2005).

103. Sandberg CJ, Altschuler G, Jeong J *et al.*: Comparison of glioma stem cells to neural stem cells from the adult human brain identifies dysregulated Wnt- signaling and a fingerprint associated with clinical outcome. *Exp Cell Res* 319(14), 2230-2243 (2013).

104. Engstrom PG, Tommei D, Stricker SH, Ender C, Pollard SM, Bertone P: Digital transcriptome profiling of normal and glioblastoma-derived neural stem cells identifies genes associated with patient survival. *Genome Med* 4(10), 76 (2012).

105. Pollard SM, Yoshikawa K, Clarke ID *et al.*: Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. *Cell Stem Cell* 4(6), 568-580 (2009).

106. Lee J, Kotliarova S, Kotliarov Y *et al.*: Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell* 9(5), 391-403 (2006).

107. Galli R, Binda E, Orfanelli U *et al.*: Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. *Cancer Res* 64(19), 7011-7021 (2004).

108. Suvà ML, Rheinbay E, Gillespie SM *et al.*: Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. *Cell* 157(3), 580-594 (2014).

109. Petalidis LP, Oulas A, Backlund M *et al.*: Improved grading and survival prediction of human

astrocytic brain tumors by artificial neural network analysis of gene expression microarray data. *Mol Cancer Ther* 7(5), 1013-1024 (2008).

110. Verhaak RG, Hoadley KA, Purdom E *et al.*: Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* 17(1), 98-110 (2010).

**** Describes a gene-expression based classification of glioblastoma into four clinically relevant subtypes and their functional annotation using somatic mutations and DNA copy number data of the TCGA cohort.**

111. Phillips HS, Kharbanda S, Chen R *et al.*: Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* 9(3), 157-173 (2006).

112. Freije WA, Castro-Vargas FE, Fang Z *et al.*: Gene expression profiling of gliomas strongly predicts survival. *Cancer Res* 64(18), 6503-6510 (2004).

113. Ozawa T, Riester M, Cheng YK *et al.*: Most human non-GCIMP glioblastoma subtypes evolve from a common proneural-like precursor glioma. *Cancer Cell* 26(2), 288-300 (2014).

114. Sun Y, Zhang W, Chen D *et al.*: A glioma classification scheme based on coexpression modules of EGFR and PDGFRA. *Proc Natl Acad Sci U S A* 111(9), 3538-3543 (2014).

115. Shah N, Lankerovich M, Lee H, Yoon JG, Schroeder B, Foltz G: Exploration of the gene fusion landscape of glioblastoma using transcriptome sequencing and copy number data. *BMC Genomics* 14, 818 (2013).

116. Dalerba P, Kalisky T, Sahoo D *et al.*: Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol* 29(12), 1120-1127 (2011).

117. Nathanson DA, Gini B, Mottahedeh J *et al.*: Targeted therapy resistance mediated by

dynamic regulation of extrachromosomal mutant EGFR DNA. *Science* 343(6166), 72-76 (2014).

118. Law CW, Chen Y, Shi W, Smyth GK: voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15(2), R29 (2014).

119. Robinson MD, Mccarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1), 139-140 (2010).

120. Hardcastle TJ, Kelly KA: baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* 11, 422 (2010).

121. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L: Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 31(1), 46-53 (2013).

122. Anders S, Huber W: Differential expression analysis for sequence count data. *Genome Biol* 11(10), R106 (2010).

123. Touloumis A, S. T, J.C. M: Testing the Mean Matrix in High-Dimensional Transposable Data. *Biometrics* arXiv:1404.7683 (2015).

124. Bonavia R, Inda MM, Cavenee WK, Furnari FB: Heterogeneity maintenance in glioblastoma: a social network. *Cancer Res* 71(12), 4055-4060 (2011).