FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

The role of Epigenetics in Glioblastoma - A Meta Analysis

Maria Beatriz Vieira Andrade

Mestrado em Engenharia Biomédica

Supervisor: João Oliveira Ventura

Second Supervisor: Peng Weng Kung

September, 2020

The role of Epigenetics in Glioblastoma - A Meta Analysis

Maria Beatriz Vieira Andrade

Mestrado em Engenharia Biomédica

Faculdade de Engenharia da Universidade do Porto

September, 2020

Abstract

Advances in the study and analysis of patients with glioblastoma (GB) have made it possible to determine the main challenges imposed on physicians to obtain a good prognosis. The lack of a 100% effective treatment in patients with GB is due to the fact that this tumor is very heterogeneous, making GB to be considered one of the most deadly tumors. Given the high urgency to implement new measures to combat these challenges, there has recently been an increase in the use of computational epigenetics as a crucial tool for the identification of potential biomarkers and for the advent of new personalized therapeutic approaches.

The purpose of this work is to respond to the need for the evolution of this field in the field of medicine, more precisely in the area of molecular oncology, through the development of an integrative meta-analysis of DNA methylation and gene expression to identify the main genes epigenetics in glioblastoma, using bioinformatics.

The developed algorithm integrates, in a first phase, the download of DNA methylation and gene expression data from the TCGA database, selecting only primary samples of glioblastoma and normal samples. After pre-processing the data, it was possible to identify about 12427 differentially expressed genes (DEGs) among the different samples, including 5783 positively regulated genes and 6644 negatively regulated genes. Then, all samples were randomly divided into two, sets, training set (n = 69) and test set (n = 69). Subsequently, the survival analysis of the training set was carried out using the Kaplan-Meie estimate. In the future, performing the analysis of Pearson's correlations between positively regulated DEGs and negatively regulated methylated survival genes, as well as negatively regulated DEGs and positively regulated methylated survival genes, will allow the construction of a prognostic risk model based on methylation genes , providing health professionals with the identification of biomarkers for a better prognosis and treatment, in order to increase the survival rate of patients with GB.

Resumo

Avanços no estudo e na análise de pacientes com glioblastoma (GB) têm permitido determinar os principais desafios impostos aos médicos para obtenção de um bom prognóstico. A falta de um tratamento 100% eficaz em pacientes com GB deve-se ao facto deste tumor ser muito heterogéneo, fazendo com que o GB seja considerado um dos tumores mais mortíferos. Dada à elevada urgência de implementação de novas medidas para combater estes desafios, recentemente, verificou-se um aumento da utilização da epigenética computacional como uma ferramenta crucial para a identificação de potenciais biomarcadores e para o advento de novas abordagens terapêuticas personalizadas.

O propósito deste trabalho vem no sentido de dar resposta à necessidade de evolução deste campo na área da medicina, mais precisamente na área da oncologia molecular, através do desenvolvimento de uma meta-análise integrativa de metilação de ADN e expressão gênica para identificar os principais genes epigenéticos no glioblastoma, recorrendo à bioinformática.

O algoritmo desenvolvido integra numa primeira fase o download dos dados de metilação de ADN e de expressão genética do banco de dados TCGA, selecionando apenas amostras de glioblastoma primário e amostras de tecido sólido normal. Após o pré-processamento dos dados, foi possível identificar cerca de 12427 genes diferencialmente expressos (DEGs) entre as diferentes amostras, incluindo 5783 genes regulados positivamente e 6644 genes regulados negativamente. Em seguida, recorreu-se à divisão aleatória de todas as amostras em dois conjuntos, conjunto de treino (n = 69) e conjunto de teste (n = 69). Posteriormente, procedeu-se à realização da análise de sobrevivência do conjunto de treino utilizando a estimativa de Kaplan-Meie. Futuramente, a realização da análise das correlações de Pearson entre DEGs regulados positivamente e genes metilados para sobrevivência regulados negativamente, bem como DEGs regulados negativamente e genes metilados para sobrevivência regulados positivamente, permitirá a construção de um modelo de risco de prognóstico baseado em genes de metilação, proporcionando aos profissionais de saúde a identificação de biomarcadores para um melhor prognóstico e tratamento, de forma a aumentar a taxa de sobrevivência de pacientes com GB.

Acknowledgements

After a long journey of 5 years, here comes the completion of yet another chapter so difficult. As such, I would like to thank not only everyone who helped me to grow as a professional person in the world of work during the course of this dissertation, but I would also like to thank everyone who was present throughout my journey school.

First, I would like to thank Professor João Oliveira Ventura, my supervisor, for all his help, patience, dedication and availability above all. I would also like to thank Dr. Peng Weng Kung, my co-supervisor, for all the availability, teaching and all the constructive advice that made me grow a lot as an independent person.

On October 1st, I started a two-month internship at INL - International Iberian Nanotechnology Laboratory, under the guidance of Peng Weng Kung. Through this internship, I had the opportunity to learn microfluidics and precision medicine. No less important, I also want to thank you for this occasion, as it was one more evaluation for me to develop my English skills. Also, I would like to many thanks Murali precious help for his knowledge in 3D cell culture and for having accompanied at INL, during first semester. I also thank José Fernandes, responsible at INL for the manufacture and printing of mi-crofluidic devices for having taught me how to work with the ArtCam and grbl software and with the CNC micro-milling machine. Also, I want to thank Diana Pinho for all her knowledge in the microfluidics and cell culture laboratory. Finally, I would also like to thank Jéssica Pinto, Frederico Tremoço and Tiago Sardo who accompanied me during the short time I had in the second semester at INL.

I also want to thank Tiago Chedraoui Silva, for his availability, patience and for all the knowledge he gave me about the construction of the algorithm and epigenetics, despite being on the other side of the world.

Next, I would like to thank all my friends who accompanied me on this journey, especially Rute Costa, who has always supported me in the bad and good times, as well as Diana Duarte and Catarina Sousa, who have always been present at a distance of one call. I also thank my twins Márcia Mendes and Diana Mendes, who have always accompanied me and are like sisters to me.

Finally, I would like to thank my whole family, especially my parents and my brothers for all their attention, patience, love and education, because without their help I would never have been able to get here. In addition to the thanks, I also wanted to apologize mainly to my mother for the moments of worry, bad mood, crying and for always being at my side to support me conditionally.

Maria Beatriz Vieira Andrade

"Maybe we not have been able to do the best, but we strive for the best to be done. We are not what we should be, we are not what we will be .. but thank God, we are not what we were."

Martin Luther King

Contents

List of Figures

List of Tables

Nomenclature

GB	Glioblastoma
BBB	Blood Brain Barrier
CAT	Computed Axial Temography
cfDNA	cell-free Nucleic Acids
CNS	Central Nerve System
CsC	Cancer Stem Cells
CTC	Circularity Tumour Cells
ctDNA	Circularity Tumour DNA
ctRNA	Circularity Tumour RNA
DNA	Deoxyribonucleic acid
EGFR	Growth Factor Receptor
EMT	Epithelial mesenchymal
EVs	Extracellular Vesicles
FISH	Fluorescent probes in situ
MEC	Extracellular Matrix
NMR	Nuclear Magnetic Resonance
PM	Personalized Medicine
PMMA	Poly-methyl Methacrylate
RNA	Ribonucleic acid
VEGF	Vascular Endothelial Growth Factor
WHO	World Health Organization
MRI	Magnetic Resonance Imaging
TCGA	The Cancer Genome Atlas
GEO	Gene Expression Omnibus
GDC	Genomic Data Commons
DEGs	Differentially Expressed Genes
DNMT	DNA methyltransferase
Chr	Chromosome
EMP3	Epithelial Membrane Protein 3
TGF	Transforming Growth Factor
PDGF	Platelet Derived Growth Factor
HDAC	Histone Deacetylase
TF	Transcription Factor
SAM	Sadenosylmethionine
HAT	Histones Acetyl Transferases
NAD	Nicotinamide adenine dinucleotide
DMRs	Differentially Methylated Regions
FDA	Federal Drug Administration
TSS	Transcription Initiation Sites
KNN	Nearest k-Neighbors
LFC	Logarithmic Fold Change
FDR	False Discovery Rate
VST	Variance Stabilization Transformations
DMS	Differentially Methylated Sites

Chapter 1

Introduction

Glioblastoma (GB) is one of the most deadly tumors worldwide. This corresponds to the most malignant form of gliomas (grade IV) and is the most common malignant brain tumor in adults. GB patients have a very short survival period (an average of only 15 months after surgery), as this is a very heterogeneous and treatment-resistant type of tumor. Another significant reason for patients to have a reduced time of passage, after the surgery, is the fact that the diagnosis is very late. This is due to the fact that the images obtained by magnetic resonance imaging (MRI) reveal the presence of GB only when it is already at a very advanced stage. Therefore, this cause also influences the treatment of patients, since, if the tumor is already very advanced, it will hinder the development and application of an appropriate therapy to the phenotype of that tumor. As all tumors are genetically, epigenetically and phenotypically different from others, GB is no exception. Therefore, there is an urgent need to create new strategies for a clinical diagnosis and personalized treatment of this tumor. However, a major challenge for researchers in the discovery of new therapies is to determine the cellular origin of GB, since it is extremely difficult to create a GB treatment that is successful without first finding out what type of cell is responsible for the appearance of this tumor. As epigenetic modifications, enzymes and non-coding RNAs are often unique to each cell type, it makes these cellular elements the first to be explored to identify the cell of origin. However, determining these changes becomes a very complex and difficult process, as the cells that acquire a mutation (mutation cell) may not be the same as the source cell. That said, recently, biomedical engineers are relying on the use of personalized epigenetics as a potential weapon in the identification of epigenetic biomarkers in the treatment of cancer. In addition to this being a relatively new area, its use allowed to observe a rapid evolution in cancer research. Personalized epigenetics is the combination of epigenetics with personalized medicine. Epigenetics refers to the hereditary changes in gene expression that result from changes in the chromatin structure, without any involvement of changes in the genetic information stored in the DNA [18], on the other hand personalized medicine consists of a medical model that suggests the customization of the services of health, that is, the appropriate choice of medical decisions, practices, and products for the body of each patient. One way to learn more about this concept is to carry out integrative analyzes of disparate heterogeneous data to generate global interpretations and biological knowledge

[19]. The appearance of the epigenome is due to interactions between different epigenetic mechanisms, including discrete biomolecules and chemical modifications, thus forming an elaborate combinatory complexity [18]. These epigenetic marks are crucial to understand the diverse nature of the data resulting from epigenomic studies on many diseases. The meta-analysis of each of these brands requires the involvement of specific techniques and workflows, resulting in different types of data. Throughout the studies, the data present a high heterogeneity, thus emerging a set of challenges for computational analysis and information management, similar to those produced by other complex systems. Great advances in the field of bioinformatics provide a new panorama of a set of possibilities for the massive generation of biological knowledge. The development of databases, such as TCGA (The Cancer Genome Atlas) and GEO (Gene Expression Omnibus), and the use of specialized computer software for the creation of dynamic algorithms and simulations allows to obtain integrative and more effective analyzes of clinical data, allowing to expand the final frontiers for a better understanding of the epigenetic changes involved not only in cancer, but also in other diseases. Through these meta-analyzes, the identification of fundamental units, with genes or proteins, responsible for the appearance or progression of the tumor, allow the development of crucial biomarkers in preventing the prognosis and in increasing the survival rate in infected patients, thus contributing to a more personalized medicine.

1.1 Objectives

Different subtypes of GB have an enormous diversity of gene pathways, genomic aberrations and gene expression profiles, contributing to an increase in intratumoral heterogeneity. So far, no genuine biomarker has been developed, which would significantly contribute to the increased survival of patients with GB. Recently, epigenetic changes have emerged as a common feature of cancer, including GB. There are several epigenetic marks, but DNA methylation is considered to be one of the central elements in epigenetic alteration. DNA methylation has the function of regulating genomic function and is one of the most important characteristics that interfere in carcinogenesis. DNA methylation plays an important role in several cellular processes, such as cell differentiation, genome stability and gene printing. The observation of biological processes, such as changes in DNA methylation, can allow the identification of potential biomarkers for the early diagnosis and prognosis of GB, thus constituting an important basis for the development of new ideas for future clinical applications. Therefore, the main objective of this work is to carry out an integrative meta-analysis of DNA methylation and gene expression to identify the main epigenetic genes in glioblastoma, using bioinformatics. In other words, we intend to build an algorithm using the Rstudio software [20], which is a free software from an integrated development environment for R, and open data from the GDC (Genomic Data Commons) of DNA methylation and gene expression of patients with GB in the TCGA database [21]. Tasks such as determining differentially expressed genes (DEGs) and methylated sites, univariate and multivariate Cox regression analysis of methylation sites obtained from GB, as well as correlation analysis of DEGs and methylated survival genes are some of the final objectives of this dissertation. However, through the discovery

of the main epigenetic genes or tumor suppressor genes of GB, it is possible in a future work to develop personalized therapies in order to obtain a better prognosis and personalized treatment in the near future.

1.2 Methodology and Structure

Regarding the application used in the present dissertation, its theoretical demonstration goes through a discerning and exhaustive search obtained from information and materials that allow the determination and construction of the different chapters and the respective state of the art surrounding each one, consisting mainly of the GB tumor, epigenetics, including DNA methylation and gene expression, that is, the specific concepts, the problems that are imposed and, finally, an idealization of possible measures to combat these problems. As a form of action, these steps will be the subject of careful reflection, the results of which will be translated into a descriptive and comparative meta-analysis, which deals with facts from past work as the foundation of the current performance of the different actors. The concepts approached and the implementation of new solutions defended by several authors, and reproduced in documents of varying order, were processed according to a conjunctural view, described and reflected throughout the dissertation. As for the documentation listed above, this was obtained through access to different sources, namely through an extensive online search of books and articles that report the works produced. The relevant bibliography related to the theme is explored and classified as of added value for research.

That said, the following document is divided into 6 chapters. The second chapter, **Personalized Medicine**, refers to the use of a more personalized medicine for oncological diseases, including the use of biomarkers, technologies based on OMICs and liquid biopsies for a better monitoring of a certain pathology in a patient. The third chapter, **Glioblastoma Multiforme**, exposes the basic principles of GB tumor, including statistics (its incidence and mortality), how heterogeneity influences the diagnosis, how the diagnosis is made, including MRI, tumor and liquid biopsies, how to collect information about the cells surrounding this tumor, as well as the epigenetic changes in GB and the evolution of personalized epigenetics in it. This chapter describes part of the problem addressed in this paper in the current reality. The fourth chapter, **Epigenetics and molecular biology of cancer**, covers the concept of epigenetics and its epigenetic changes, that is, aberrant DNA methylation and modifications in histones, as well as epigenetic processes, including inactivation of the X chromosome and genomic imprinting . In addition, it also includes epigenetics in cancer and personalized epigenetics linked to bioinformatics, these topics being very recent and with a huge impact in the biomedical area. The fifth chapter, **Integrative meta-analysis of an epigenetic study**, reveals the entire construction step by step of the meta-analysis algorithm used to identify the main epigenetic genes. Subsequently, all the experimental results obtained in this work and their respective analysis and discussion are observed.

Finally, the sixth chapter, the **conclusions** of this dissertation are highlighted with a summary and critical discussion of the most important results obtained. A vision is also presented for future work to be carried out in order to create a fully customized system.

Chapter 2

Personalized medicine

2.1 Overview

Personalized medicine (PM) consists of a medical model that suggests better personalization of health services, that is, the practice of better medical decisions and better products suited to the needs and the organism of each patient. This makes use of the individual's phenotypic and genotypic characterization, based on sociodemographic, environmental and lifestyle data, clinical and medical image information and genetic profiles, in the assessment of individual capacity for a given pathology [22]. This contrasts with traditional medicine, in which care is only based on a patient's family history, social circumstances, environment and lifestyle. This type of medicine has as main objective to use the right medicine in the right dose, with minimal or no toxicity, for the right patient at the right time, thus presenting a better cost-benefit ratio, a better treatment efficiency and less toxicity [23]. The growing investment of a new generation of technological innovations in PM seeks to customize medical practice with the focus on the individual, based on the use of genetic tests, identification of biomarkers and development of new medications [24]. Investments in the implementation of PM have been made mainly in the area of oncology and rare diseases. Personalized oncology appeared with the emergence of the main technologies and consists of performing specific tests and decentralizing the care stage (PoC) (for example, self-monitoring), in contrast to the traditional clinical diagnosis, Figure 2.1 [1]. However, these advances, especially at the cellular and molecular levels, allow for advanced diagnostics and monitoring, reduce the clinical burden using minimally invasive body fluids and provide results to the electronic records of patients, and subsequently improve the patient's well being [25].

2.2 Role of biomarkers in Personalized Medicine

A biomarker allows to obtain indications about the diagnosis of the patient, through measurements, based on precision and reproducibility [26]. In 1998, a Working Group on Biomarker Definitions from the National Institutes of Health defined a biomarker or biological marker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes,

Figure 2.1: Comparison of traditional clinical oncology with personalized oncology [1].

pathogenic processes or pharmacological responses to a therapeutic intervention [27]. Biomarkers can be of diagnosis, prognosis, prediction, pharmacokinetics/pharmacodynamics, stratification and toxicity/efficacy. Diagnostic biomarkers are used to discriminate different pathological stages of the patient's condition, in order to critically determine whether a patient has a specific medical condition for which treatment can be indicated [28]. Through a predictive diagnosis we can obtain information about the probability that a treatment will work on a specific patient or on the probability that this treatment will cause an unwanted side effect [29]. However, these three types of biomarkers are the most used and when applied in clinical practice they have different functions. After the identification of biomarkers and properly validated during preclinical tests, it is necessary to investigate whether they have diagnostic or prognostic capacity to indicate the presence of a certain pathology or the progression of it in a patient. To this end, we must resort to specific and rigorous measures, in order to assess the behavior of biomarkers before they are adopted in practice. Figure 2.2 shows the impact of biomarkers on PM. These impacts relate to the development of a molecular diagnosis associated with the drug which tests or directs treatment to a sub-set of the population with a specific genetic mutation. Such a change implies greater assertiveness of the research, in order to assess whether the use of these biomarkers alters the clinical results, in order to satisfy the needs of several interested parties [30]. However, biomarkers play a crucial role in improving the process of early diagnosis, drug development, as well as in the field of biomedical research. Advances in these molecular diagnostic technologies have been used mainly to detect biomarkers of various diseases, such as cancer, metabolic disorders, infections and diseases of the central nervous system.

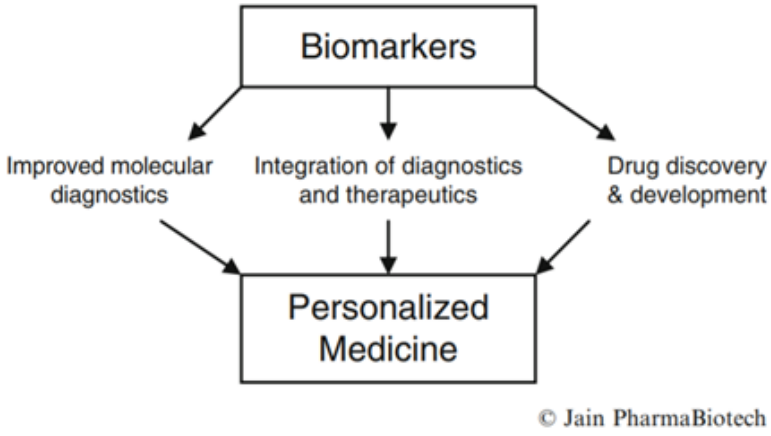

Figure 2.2: Impact of biomarkers in personalized medicine [2].

2.3 OMICS

To better understand the pathophysiology of most diseases, recently the main focus of research has been on omics-based approaches. The word "omics" refers to the area of biological sciences that ends with -omics, such as genomics, transcriptomics, proteomics or metabolomics [12]. This set of terms allows obtaining specific data from dynamic interactions between the numerous components of a biological system. The interactions are in-tended to analyze a set of networks, paths and inter-active relationships that exist between the various components and processes of life (for example genes, transcripts, proteins, metabolites and cells) [31]. Thus, the main objective of "omics" is to identify, characterize and quantify all biological molecules involved in the structure, function and dynamics of a cell, tissue or organism [32]. According to the World Health Organization (WHO), genomics is defined as the science that studies the structure, function, evolution and mapping of genomes and the respective characterization and quantification of genes [33, 34]. A genome corresponds to a complete set of DNA in an organism, including all of its genes [35]. Regarding transcriptomics, it is defined as the set of all transcribed RNA molecules (mRNA, non-coding RNA, rRNA and tRNA) produced by the genome [36]. This term appeared, for the first time, when Schena et al. (1995) developed the "microarray" technology using the inkjet DNA synthesizer, thus allowing to analyze a vast set of mRNA cells [37]. Epigenetics is the science that addresses hereditary changes in gene expression or phenotype. These changes are made at the level of histones that involve chromatin and DNA methylation [38]. As for proteomics, it studies all proteins existing in human biology [39]. The research related to this term allows to obtain important information about its biochemical properties and respective functions, thus helping in the discovery of new biomarkers to identify and locate possible post-translational modifications, as well as to understand the role of protein-protein interactions [40]. Finally, the metabolic allows studying the cellular metabolism of living organisms, in real time [41]. This allows collecting quantitative data on endogenous enzymes, cellular biochemical reactions and synthesis of cellular

metabolites within a biological system [42]. That said, an important factor in systems biology, which we must take into account, is data mining. The coordination and correct access of data can become a somewhat complicated task, given the diversity of data managed by current high-performance technologies [43]. With appropriate algorithms, this database can be extracted by doc-tors, in order to make medical decisions based on personalized “omics” data with an appropriate interpretation, thus benefiting the understanding of all the physiological stages of health and disease [44]. Therefore, we conclude that the "Omics" area is increasingly linked to cancer studies. These studies reveal the molecular mechanisms that are involved behind various types of cancer and help to identify biomarkers for early diagnosis and treatment. However, with the rapid reduction in the cost of omics profiles, it is possible to anticipate a high number of personalized drug applications in many aspects of healthcare, in order to contribute to a better future of precision medicine, thus providing personalized monitoring of health [45].

2.4 Liquid Biopsy

Liquid biopsies arecrucial tools for early diagnosis of the presence of a certain cancer. Liquid biopsies are personalized blood tests that aim to extract the tumor and study its heterogeneity and evolution in real time [3]. These are minimalinvasive processes, when compared with tissue biopsies, and allow the identification of the presence of biomarkers (Figure 2.3), such as, for example, circulating tumor cells (CTCs), circulating cell-free tumor DNA (ctDNA) and RNA (ctRNA), circulating microRNAs, proteins, tumor forming platelets (TEPs) and extracellular vesicles (EVs) [46]. CTCs are circular tumor cells that are eliminated from primary tumors or metastases (homotypic or heterotypic clusters) and therefore spread in the bloodstream [47]. These present themselves as metastatic precursor cells [48], which undergo transitions at the epithelial mesenchymal (EMT) level. The appearance of these cells is due to the increase in tissue hypoxia that derives from tumor growth [49]. After collecting the blood sample, using liquid biopsy, it is necessary to isolate the CTCs. The isolation of these cells can be done through different processes, such as filtration, chip technology, gradient or density centrifugation, electric field, sound waves and microfluidic technology [50]. However, isolation of CTCs is quite difficult, since there is a large number of cells in the blood of patients, which corresponds to approximately 1 in 109 cells [3]. However, its extraction is very important, as it allows to obtain crucial information for the understanding of the biological processes of the tumor and its molecular characterization also contributes to explain how CTCs have a high capacity for resistance to treatments. With the advances in the re-search of these cells, the cells of DNA and RNA in CTCs contribute to the determination of the degree of heterogeneity and the degree of similarity of the cancer cells. Regarding cell-free nucleic acids (cfDNAs), these are valuable markers in different diagnostic protocols, such as in the early diagnosis of genetic diseases and in the detection of tumors [51]. cfDNAs are fragmented molecules that contain signals that can be of tissue or cellular origin. In patients with cancer, there is an enormous amount of cfDNA formed by apoptosis and necrosis of tumor cells, where the release of DNA from necrotic cells occurs only by phagocytosis [52]. However, according to the circulation

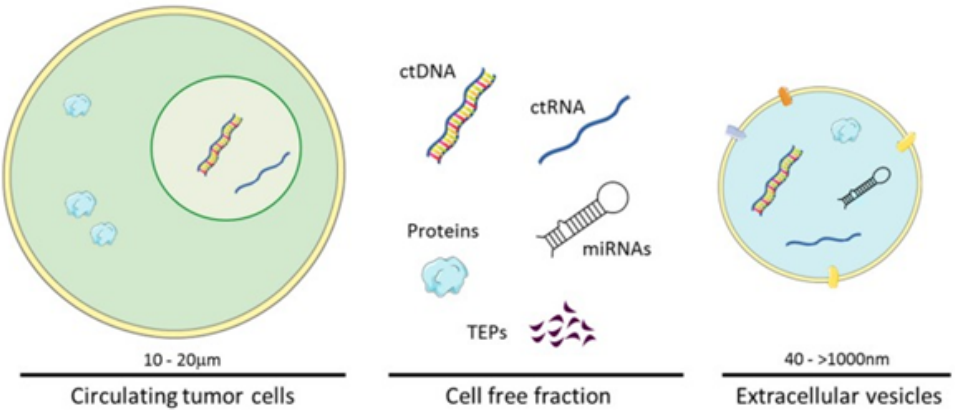

Figure 2.3: The different components that can be found in a liquid biopsy (blood or cerebrospinal fluid) of patients with glioblastoma. [3]

of the cells, the cfDNA molecules can be divided into: fragments of free DNA, DNA bound to the vesicle and DNA-macromolecular complexes [53]. The fragments of free DNA present themselves with bare sequences, as they are not linked to any other molecule or surface. As for RNAs, they are minimally unstable and susceptible to degradation induced by ribonucleases [54]. Circulating cell-free tumor RNA (ctRNA) includes mRNAs, long noncoding RNAs (lncRNAs) and mainly small ncRNAs, cellless mRNA (cf-mRNA) and tRNAs. However, all of these compounds can be released into the extracellular medium with the function of regulating various signaling pathways through the repression of multiple genes at the translation level [55]. Recently, extracellular vesicles (EVs) have contributed to the development of important biomarkers for non-invasive or minimal- invasive diagnosis and monitoring of cancer [56]. EVs are small closed spheres composed of an external hydrophobic lipid bilayer around a hydrophilic aqueous nucleus [57]. These have a diameter be-tween 40 and 1000 nm and are formed and secreted from different cells resulting from complex and strongly regulated mechanisms [3]. Extra-cellular vesicles are classified according to their biogenesis, in exosomes (40–200 nm), membrane-derived vesicles (40 to >1000 nm) and apoptotic bodies (100–5000 nm) [58]. EVs are present in various body fluids, namely blood, CSF or urine, and their biomolecular composition is quite diverse [59]. These compounds can transport portions of mRNA, miRNA, DNA, cellular proteins, lipids and metabolites, which can be transferred to near-by or distant cells by direct contact with the membrane of EV cells, fusion or internalization [3]. However, EVs are associated with tumor progression, interfering with the transport of factors capable of controlling and deregulating processes, with proliferation, drug resistance, migration, angiogenesis induction, and invasion.

Chapter 3

Glioblastoma

3.1 Overview

The brain tumor is a rare tumor and is inserted in the type of tumors where the cure is practically nonexistent [60]. It can be classified as primary or as metastatic [61]. Primary brain tumors originate in the brain, whereas metastatic brain tumors originate from other organs. Gliomas are a group of brain tumors whose origin is glial cells. These cells are part of the central nerve system (CNS) and act as support cells for neurons [62]. However, gliomas can be classified as astrocytomas, oligodendrogliomas and ependiomas. Within the astrocytoma we have four grades (I, II, III, IV), where grade IV is called Glioblastoma (GB). GB is an aggressive malignant tumor, and it is the brain tumor that occurs most frequently. It originates from primary brain tumors, but over time it becomes a metastatic tumor. The survival rate is very low (an average of only 15 months after surgery), because this is a very heterogeneous and treatment-resistant type of tumor [63]. The age groups most affected with this type of tumor are male groups in adult-hood (between 65-69 years old) [64].

The initial diagnosis of GB is made through neuroimaging, followed by resection or biopsy of tumor tissue to diagnose, classify and characterize the tumor [65]. Currently, tissue biopsies are the most used technique for the diagnosis of GB. However, these biopsies are highly invasive, putting patients at risk for life due to consequences such as the development of brain swellings or due to consequences in terms of neurological functions [66]. As an alternative to tissue biopsies, liquid biopsies are minimal-invasive and allow the detection of circulating biomarkers [67]. These facilitate sampling and can monitor possible dynamic changes in the tumor throughout personalized therapy. In addition to liquid biopsy, diagnosis based on “Omics” is also becoming a crucial tool in the personalized diagnosis of GB.

3.2 Statistics

Globally, more than 241,000 people with brain or CNS cancer die each year, with glioblastoma being the most common form of the disease [4]. According to the data obtained from Globocan (an

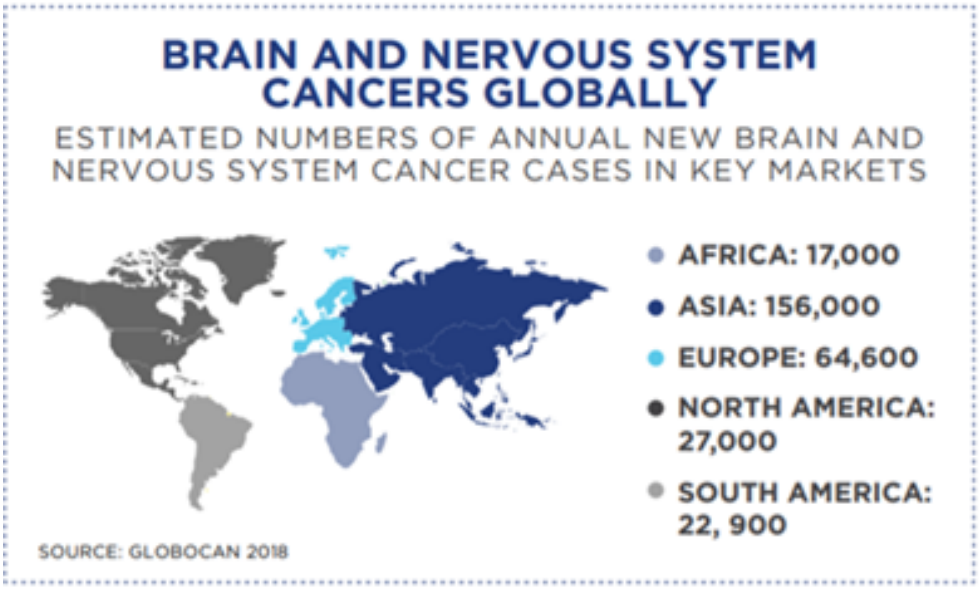

Figure 3.1: Worldwide distribution of GB tumor in 2019 [4]

interactive platform that presents statistical data on all types of cancer existing worldwide) [4], it was found that in 2018 in Europe there were about 64600 cases of GB (Figure 3.1). The incidence rate of GB is estimated at 4.6 per 100,000 people [15], and this estimate varies according to the region, with the development of the country and with the increase of age, being verified that as the age of the person increases the probability of developing GB is also higher (the average age of diagnosis is found at 64 years). As previously mentioned, the survival of patients with GB is extremely reduced, although there is already a huge advance in diagnosis and treatment. Through the studies, a survival rate of approximately 34% has been observed for patients who survive 1 year, 12% for those who survive 2 years, and less than 5% for those who survive 5 years, since the date of diagnosis (Viegas, 2018) [64].

3.3 Classification and Morphology

According to the World Health Organization (WHO), tumors associated with the brain and the CNS are grouped according to their morphological similarity, immunohistological profile, differentiated brain cells and cytoarchitecture. In 2016, WHO updated the classification criteria for these tumors [68], where they are now based on the tumor's genotype and phenotype. Glioblastomas, according to this classification, are inserted in diffuse atrocitary and oligodendroglial tumors, and are subdivided into three types (Table 3.1):

- Glioblastoma HDI-wildtype - These are called primary glioblastomas and represent about 90% of all glioblastomas. They are characterized by their high mitotic activity, vast cellularity and the presence of necrosis or microvascular proliferation [69].

• Glioblastoma HDI-mutant - Corresponds to approximately 10% of cases and is called secondary glioblastomas. GB associated with point mutations in isocitrate dehydrogenase 1 or 2 (HDI1/2), are linked to most glioblastomas resulting from a diffuse astrocytoma (WHO grade II) or anaplastic astrocytoma (WHO grade III) [70].

• NOS glioblastoma - These glioblastomas are so named, since there is no information on the presence of some type of diagnosed mutation [64].

In addition to the morphological division, brain tumors, according to WHO, are also subdivided according to their grade-I, II, III and IV (Table 4.2), this subdivision being associated with the rising order of malignancy [71]. As previously mentioned, GBs correspond to the highest grade (IV) gliomas, and can be classified as primary or secondary, according to their origin. Primary or "de novo" glioblastomas correspond to approximately 85 to 90% of glioblastomas and develop without precursor damage [72]. As for secondary glioblastomas, they arise from a pre-existing glioma of lesser degree [64] and correspond to approximately 10% to 15% of all glioblastomas.

Table 3.1: Classification of glioblastoma in relation to tumor diffusion, according to the adaptation of the WHO table [14]

Diffuse astrocytic and oligodendroglial tumours	
Difusse astrocytoma, IDH-mutant	9400/3
Difusse astrocytoma, IDH-mutant Gemistocytic astrocytoma, IDH-mutant	9411/3
Diffuse astrocytoma, IDH-wildtype	9400/3
Diffuse astrocytoma, NOS	9400/3
Anaplastic astrocytoma, IDH-mutant	9401/3
Anaplastic astrocytoma, IDH-wildtype	9401/3
Anaplastic astrocytoma, NOS	9401/3
Glioblastoma, IDH-wildtype	9440/3
Giant cell glioblastoma	9441/3
Gliosarcoma	9442/3
Epithelioid glioblastoma	9440/3
Glioblastoma IDH-mutant	9445/3*
Glioblastoma NOS	9440/3
Diffuse midline glioma, H3 K27M-mutant	9385/3*
Oligodendroglioma, IDH-mutant and 1p/19q codeleted	9450/3
Oligodendroglioma, NOS	9450/3
Anaplastic oligodendroglioma, IDH-mutante and 1p/19q codeleted	9451/3
Anaplastic oligodendroglioma, NOS	9451/3
Oligoastrocytoma, NOS	9383/3
Anaplastic oligoastrocytoma, NOS	9382/3

Usually, the origin of GB occurs in the brain, and these tumors are located mainly in the subcortical white matter of the cerebral hemispheres [73]. The sites that are most affected are the temporal (31%), parietal (24%), frontal (23%) and occipital (16%) lobes, being the most common

Table 3.2: Classification of glioblastoma according to the degree of glioma of the tumor, according to the adaptation of the WHO table [15]

Gliomas Degrees	
Diffuse astrocytomas and oligodendrogliomas	
Diffuse astrocytoma, IDH-mutant	II
Anaplastic astrocytoma, IDH-mutant	III
Glioblastoma, IDH-wildtype	IV
Glioblastoma, IDH-mutant	IV
Oligodendroglioma, IDH-mutant and 1p / 19q-codeleted	II
Anaplastic oligodendroglioma, IDH-mutant and 1p / 19q-codeleted	III
Other Astrocytomas	
Pilocytic Astrocytoma	I
Subependymal giant cell astrocytoma	I
Pleomorphic Xanthoastrocytoma	II
Anaplastic pleomorphic xanthoastrocytoma	III
Ependymomas	
Subependymoma	I
Mixopapillary ependymoma	I
Ependymoma	II
Anaplastic Ependymoma	III

frontotemporal region, Figure 3.2 [74]. The progression of this tumor is very fast, except when the development of GB occurs within an astrocytoma, which is the most common tumor of the central nervous system [75]. Regarding its morphology, GB corresponds to 50% of all gliomas, and represents a set of variations in the tumor cell and tissue structure. The formation of this tumor is based on small cells, characterized by polymorphism, anaplasia and significant anisocariosis [14]. The cells that are part of its constitution are polygonal and fusiform and have an acidophilic cytoplasm and common cell borders. The increase in cytoplasm induces the development of nuclear polymorphism. Binuclear and multinucleated cells can also be present in GB, namely necrotic cells, lymphocytes, macrophages and neutrophils [5].

Glioblastoma multiforme is one of the tumors with the highest vascularization rate. The newly formed vessels are morphologically composed of endothelial cells (phenotypically different from normal endothelial cells), and their surface is covered by a layer of pericytes (cells that normally line blood vessels) [76]. In GB, it was also verified the presence of regions of necrosis with neoplastic cells in pseudo palsy [77]. Histologically, there are two types of necrosis present in the tumor, which depend on the location and size of the necrotic area. The first type results from insufficient blood supply in all primary glioblastomas, and the second type is more common in

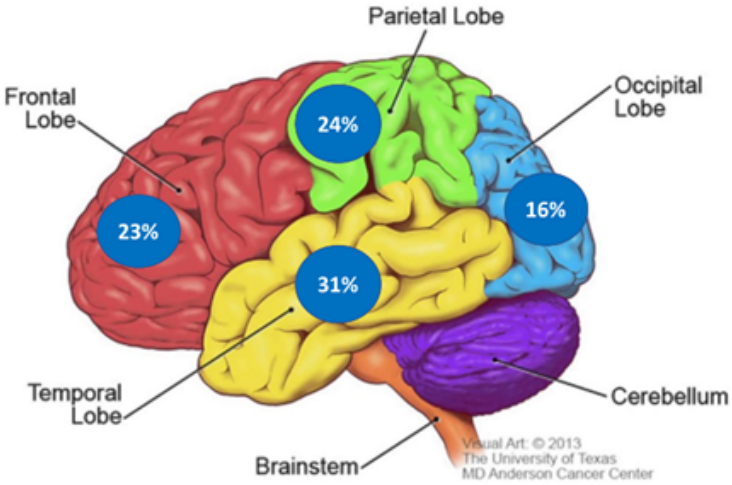

Figure 3.2: Geographic representation of the tumor percentage in each brain region [5]

secondary glioblastomas, where they have necrotic foci surrounded by pseudopalisated regions formed by glial cells. However, when we look at this tumor microscopically, the center of the tumor's necrotic tissue is usually encircled by a gray border and a yellowish gray texture, which is surrounded by the white substance. In microscopic view, it is also possible to observe black hemorrhagic lines and thrombosed vein [78].

3.4 Etiology

The development process of GB is very fast, and the manifestations of the presence of this tumor depend a lot on the location, the tumor volume and the ability to spread. Approximately 1% to 5% of GB are of hereditary origin. GB presents itself as a spontaneous tumor and its cause is unknown in most cases. Genetics, epigenetics, bacterial infections, and many other factors influence GB oncogenesis [60]. Regarding genetics, it contributes to glioblastoma as follows: tuberous sclerosis [79], Turcot syndrome [80], multiple endocrine neoplasia type IIA [81] and neurofibromatosis type I, NF1 [82]. The development of GB is related to the dysregulation of the G1/S checkpoint in the cell cycle and the occurrence of many genetic disorders in glioma cells, such as the amplification of the EGFR (7p12) gene, mutations in the TP53 gene and the loss of chromosome 10 [83]. Traumatic brain injuries can also be the cause of glioblastoma, as well as exposure to organic solvents and pesticides [84]. Nevertheless, signs such as hemiparesis, loss of sensory signals and loss of vision, are often linked to the location of the tumor. One of the very common symptoms in patients with GB is the presence of headache [85], at the time of diagnosis. Another symptom that can also be associated with this type of pathology is the presence of cognitive difficulties and the constant changes in personality, which are sometimes confused with psychiatric illnesses, especially because it manifests more in the elderly [85]. In

addition, about 10% of all gliomas are caused by irradiation, these occurrences being rarer. These, usually, appear because complex genetic abnormalities combined with unknown environmental factors, dispose individuals to the development of gliomas [86].

3.5 Heterogeneity

The heterogeneity of a tumor corresponds to the diversity of tumor cells present in a given tumor. These cells induce distinct morphological and phenotypic lesions, with different cell morphologies, gene expression, metabolism, microenvironment, proliferation and the possibility of metastatic lesions [87]. Heterogeneity can be called intertumoral or intratumoral heterogeneity [88]. Intertumoral heterogeneity arises when there is any type of genetic alteration in the different tumors located in the same organ. While the intratumoral heterogeneity corresponds to the diversity of existing cells within a given tumor. The existence of different population groups, which have the same genetic characteristic (called clones), in the same tumor are the main cause of the failure of therapies that are applied in the treatment. This is mainly due to the characteristics of tumorigenesis responsible for progression, resistance, metastatic potential and tumor renewal.

The evolution of the origin of intratumoral heterogeneity can be explained in four stages. The first stage (Figure 3.3, a)) concerns the model of clonal evolution, which is based on the division of tumor cells under selective pressure and exhibits genetic or epigenetic mutations [87]. This stage occurs when there are environmental changes, that is, when there is an alteration induced by chemotherapy or radiotherapy, thus observing, again, the growth of a group of dominant cells. More recently, the second stage (Figure 3.3, b)) corresponds to the theory of cancer stem cells (CsC). This theory consists of a hierarchical model that is based on the existence of groups of biologically distinct cells, with a high tumorigenic capacity [6]. This group of cells is called CSCs and they are grouped in specific anatomical-functional sites, where they are in direct contact with the different types of cells and the respective extracellular matrix. When asymmetric division occurs, CSCs originate a cell like itself (self-renewing capacity), while the most differentiated cells undergo successive divisions and give rise to daughter cells with non tumorigenic properties [89]. The marked manifestation of specific carrier proteins [90] (for example P glycoprotein and BCRP) correspond to yet another characteristic of CSCs. The third step (Figure 3.3, c)) consists of the heterogeneity formed by cellular plasticity in response to microenvironment signals [91]. As the microenvironment is not homogeneous and the differences in oxygen pressure, density of blood vessels, factors of growth and composition of the extracellular matrix are the main consequences of the phenotypic and genetic differences observed in tumor cells as a new source of CSCs. In addition, they also promote angiogenesis through pro-angiogenic growth factors, such as vascular endothelial growth factor (VEGF), endothelial migration and tube development. Tumors with CD133 + have more necrosis, hemorrhages and have a high rate of vascularization, whereas tumors C133- do not [92]. Regarding the last stage (Figure 3.3, d)), this corresponds to a very recent perspective and depends on the clonal evolution of the tumors through the growth of a tree [93, 94]. As the tree grows, the various regions of the tree play a crucial role in the development of the

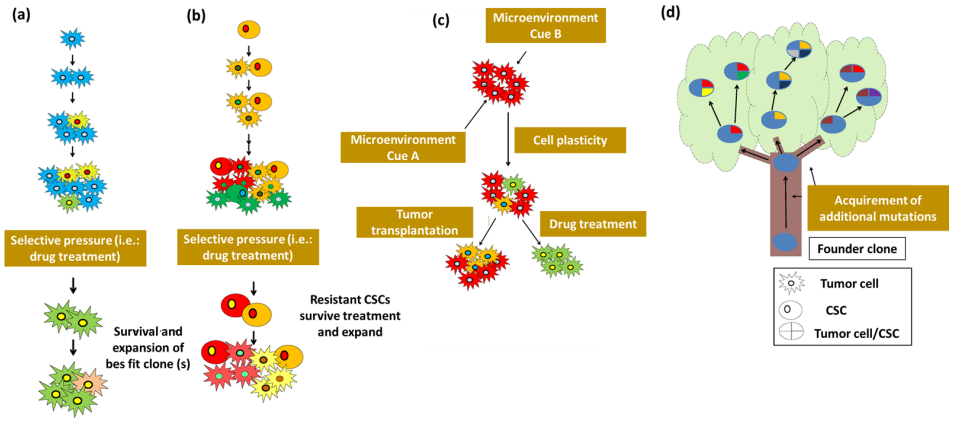

Figure 3.3: Schematic representation of the four stages of the evolution of the heterogeneity of a given tumor: (a) clonal evolution, (b) hierarchical model, (c) the role of CSCs in the evolution of heterogeneity and (d) model of the "tree structure" of a tumor [6].

tumor. The stem corresponds to the original clone that carries "controlled" mutations, and which are responsible for the initiation of malignant transformation and tumor growth. The branches that leave the trunk are the clones that originate from the accumulation of new mutations, these representing the heterogeneity that exists in a tumor) [6].

3.5.1 GB heterogeneity

GB is characterized by its high tumor heterogeneity, hence it is called glioblastoma multiforme. It presents an enormous genomic and phenotypic intratumor heterogeneity, which contains characteristics with invasive infiltrative properties, high proliferation capacity, microvascular hyperplasia, nuclear atypia and necrotic spots surrounded by pseudo palisated cells [95]. The primary GB can be subdivided into four subgroups according to their order of heterogeneity, that is, it can be mesenchymal, classic, neural and proneural [96, 7] (Figure 3.4), with the secondary GB it is also inserted in the latter. The presence of CSCs in the GB has a fundamental role in understanding this heterogeneity, namely in explaining how the GB's CSCs exhibit an iron elimination mechanism normally seen in liver cells, and how this iron uptake promotes tumor proliferation[97]. The increased expression of the gene that encodes transferrin (Tf) in the CSCs of glioblastoma, during RNA sequencing, positively influences the regulation of the ferritin stored iron transporter. Tf is a protein typically secreted by the liver that aims to eliminate iron from the blood and maintain normal brain activity. However, if CSCs capture more iron than non-CSCs, there is a significant loss of iron, thus contributing positively to tumor differentiation and progression [97]. Nevertheless, cellular plasticity also plays an important role in the heterogeneity of GB, since it corresponds to the process where brain tumor cells acquire different phenotypes. This process occurs due to factors such as the presence of acidity and hypoxia microenvironments, as well as the stress caused

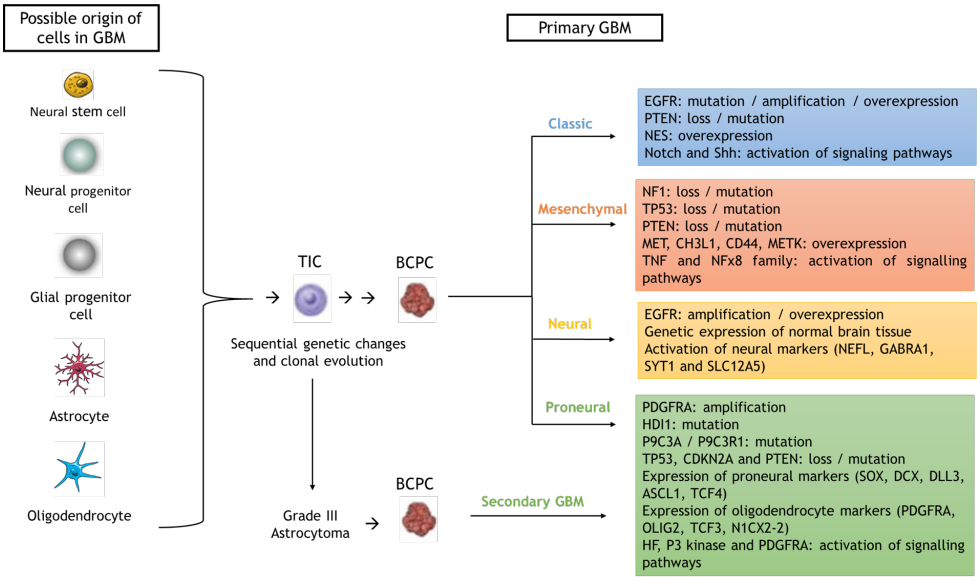

Figure 3.4: Schematic representation of the sequencing of the successive genetic changes observed in the pathogenesis of different GB subtypes. Mutations in normal brain cells cause the growth of a population of tumor-initiating cells (TICs). The agglomeration of TICs with genetic and epigenetic alterations results in the appearance of brain cancer spreading cells (BCPC). [7]

by chemotherapy and radiation, thus influencing the state of GB cells [97]. This process occurs in the presence of certain conditions, such as the existence of acidic and hypoxia microenvironments, as well as the stress propitiated by chemotherapy and radiation, thus influencing the state of GB cells [97]. However, there are other examples that explain the high heterogeneity of the tumor. For example, the appearance of particular changes, such as the p53 mutation in low-grade gliomas or the MGMT expression in GB [98], or the amplification of the EGFR gene [99], which is usually surrounded by a mutation in the similar protein, and gives rise to a frequently constitutive receptor called EGFRvIII, ΔEGFR or EGFR [83].

3.6 Diagnosis

The main diagnostic methods used today to identify the presence of GB necessarily depend on neurological tests and neuroimaging methods [100, 101]. Usually, these tests are done when glioblastoma is already in an advanced stage, since the late diagnosis is due to the slowness of the dissemination processes typical of brain tumors and the absence of clinical manifestations [102]. In 2007, Kleihues et al. revealed that the diagnosis of GB consisted of a tissue pattern with anaplastic cells, which present mitotic activities and vascular proliferations [103]. Currently, computed axial tomography (CAT) and magnetic resonance imaging (MRI), with or without contrast, are the main diagnostic tools used in the detection of GB. The use of contrast in MRI (for example gadolinium particles) facilitates the diagnosis, thus making it possible to differentiate the different peripheral hypercellular areas of the tumor (Figure 3.5) [7]. Nevertheless, during diagnosis,

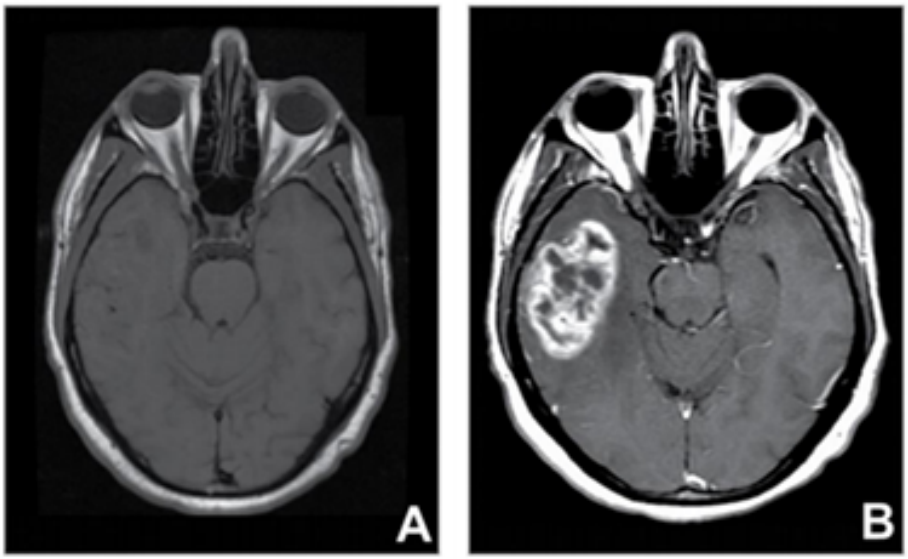

Figure 3.5: Brain nuclear magnetic resonance of a patient with GB: (A) without contrast, with poor lesion delimitation, and (B) with contrast, where the heterogeneity of the tumor can be seen, with the necrotic lesion being well marked [8]

tumors have different diameters, with a tumor diameter of approximately 4 cm being the most common [104]. In 1993, studies by Simpson et al. demonstrated that in 56% of the tumors can have diameters between 5 to 10 cm, 38% correspond to diameters smaller than 5 cm, and with a smaller percentage (6%) the tumors can acquire diameters larger than 10 cm [105].

Furthermore, doctors also take samples of abnormal tissue, called biopsies. Depending on the location of the glioma, a biopsy can be performed with a needle prior to assigning treatment to the patient or as part of an operation to remove the brain tumor. Normally, for biopsies they use stereotactic needles to detect gliomas in difficult to reach areas or very sensitive areas inside the brain that can be injured by a more extensive operation. Usually, this type of surgery is accompanied by computed tomography or magnetic resonance imaging. After surgery, the biopsy sample is analyzed under a microscope to determine whether it is malignant or benign. That said, biopsy is the only way to definitively diagnose a brain tumor and provide a prognosis to guide treatment decisions. Based on the characteristics of the tissue, the cancer doctor can determine the degree of the glioma as well as what stage of the brain tumor the patient is in.

Several laboratory techniques can also be used to differentiate primary tumor from secondary tumor. These techniques can be through immunohistochemistry, hybridization with fluorescent probes in situ (FISH), pyrosequencing, and direct sequencing by methods based on the polymerase chain reaction (PCR). Typically, the FISH, pyro-sequencing and PCR techniques are very complex and time-consuming and, therefore, are not as used as immunohistochemistry, which is fast and more economical [7]. However, immunohistochemistry is more suitable for screening tests. Tissue biopsies are invasive and ineffective processes [106], since they only remove small

regions of the tumor and, therefore, there is an enormous probability that GB will reproduce. In addition, there is still a great difficulty in having access to brain tumors, thus making it difficult to extract tumor samples, in addition to the fact that the invasion of the process confers high risks to patients [9]. Often, patients with glioblastoma after tumor removal by surgery are subjected to successive chemotherapy and radiation therapy sessions. Nevertheless, these treatments have a very low efficiency rate because this tumor is very heterogeneous, has a high recurrence rate and the diagnosis is late. To this end, recently, doctors have been engaging on using more personalized diagnoses and treatments, namely those based on omic properties and liquid biopsy to identify possible effective biomarkers. With an early diagnosis, patients are able to obtain a higher survival rate.

3.6.1 Conventional ways of Diagnosis

3.6.1.1 OMICS

To better understand the pathophysiology of GB, recently the focus of the research has been on omics-based approaches (Figure 3.6) [107]. Through these approaches we can understand and study the biodiversity and behavior that complements GB, namely, mutations that occur at the level of DNA and RNA, changes in proteins, epigenetic changes such as DNA methylation, transcriptome analysis and DNA sequencing/RNA of the complete genome [107]. However, advances in the era of omics in GB are still very few, due to the late diagnosis of the disease in the patient, the genetic heterogeneity and the intrinsic molecular complexity of the tumor [108]. However, progress has been observed in this area, namely studies on genomics, which have allowed the development of bioinformatics techniques. These techniques have provided new opportunities in molecular analysis and, therefore, have contributed positively to the understanding of gene expression, identification of CTCs and regulatory RNA, metabolomic changes and immunomodulation approaches in GB. Cancer research based on omics will play a key role in the early diagnosis, treatment and monitoring of patients with GB, thus providing greater stability to patients.

3.6.1.2 Liquid biopsy

Recently, advances in precision oncology and liquid biopsy have made it possible to identify important biomarkers genetic and epigenetic for diagnostic and prognostic purposes. The success of the liquid biopsies used in GB depends on its location. GB cells are usually located in the blood-brain barrier (BBB). This barrier separates the blood that circulates in the CNS from all the rest of the blood that circulates in the bloodstream, and its function is to regulate the access and exchange of nutrients, vitamins and other molecules in the brain. The proteins that have the scaffold function (claudin-3, claudin-5, claudin-12) [109] are the main ones for maintaining brain stability, so a mutation at the level of these proteins contributes to explain the GB tumor progression [110]. GB produces pro-angiogenic factors (angiogenesis), thus inducing an inflamed environment [111]. This induction leads to a decrease in tight junctions between adjacent endothelial cells, where they are responsible for forming a barrier that prevents the passage of most

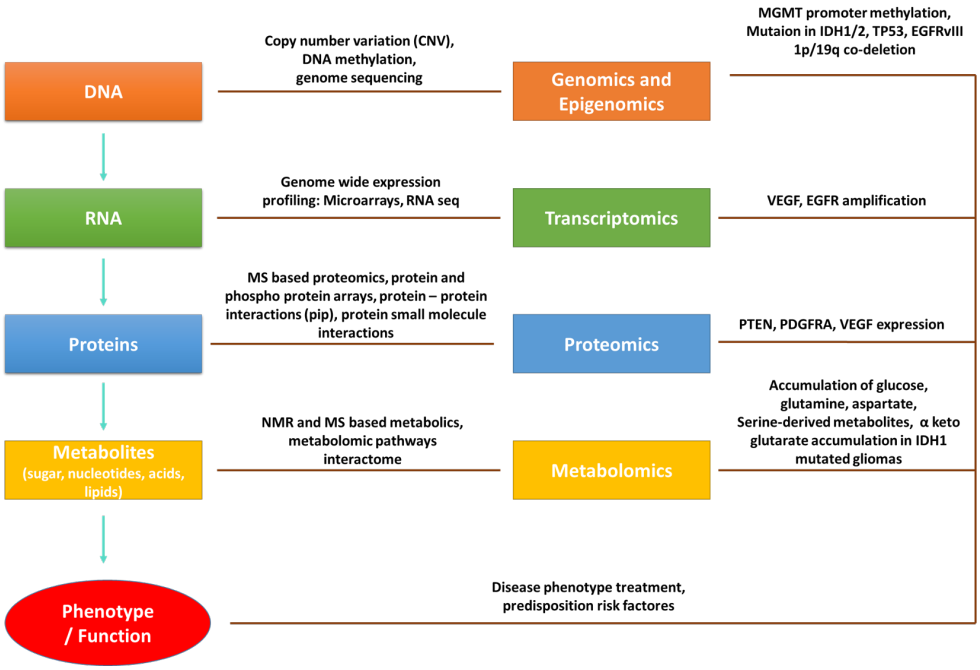

Figure 3.6: Schematic representation of OMICs applied to GB tumor cells [9]

toxic blood ions and molecules to the brain. However, as previously mentioned, through liquid biopsy we can extract biological components (Figure 3.7) [10], such as CTCs, circulating cell-free DNA (cfDNA), circulating tumor DNA (ctDNA), circulating tumor RNA (ctRNA), mRNAs, extracellular vesicles (EVs), proteins, metabolic and platelets [106]. These cells are developed from the tumor tissue and promote the real time manifestation of glioblastoma. In addition, the rapid renewal of GB tumor cells is due to the constant release of nucleic acids and vesicles derived from CTCs, which are separated by neoplasia and meet the bloodstream [112, 113]. The CTCs present in GB are mostly formed from primary tumors, and CTCs developed from metastases are rare due to the high rate of tumor renewal or the possibility of tumor growth outside the CNS by the system immunological or by the BBB, since the presence of the BBB pre-vents the entry of cells into the circulation [114, 115, 116]. Studies done to explain the role of CTCs in GB, have shown that the possibility of using genomic analyzes of single cells to detect the presence of collective mutations found in CTCs and tumor tissue, in order to verify that CTCs arose from GB. After analyzing these studies, it was concluded that GB CTCs had more mesenchymal phenotype. This phenomenon explains the rare events of unusual cases of extracranial metastases in GB. However, there is still very little information about the presence of CTCs in GB, which also makes it difficult to understand the influence of these cells on glioblastoma.

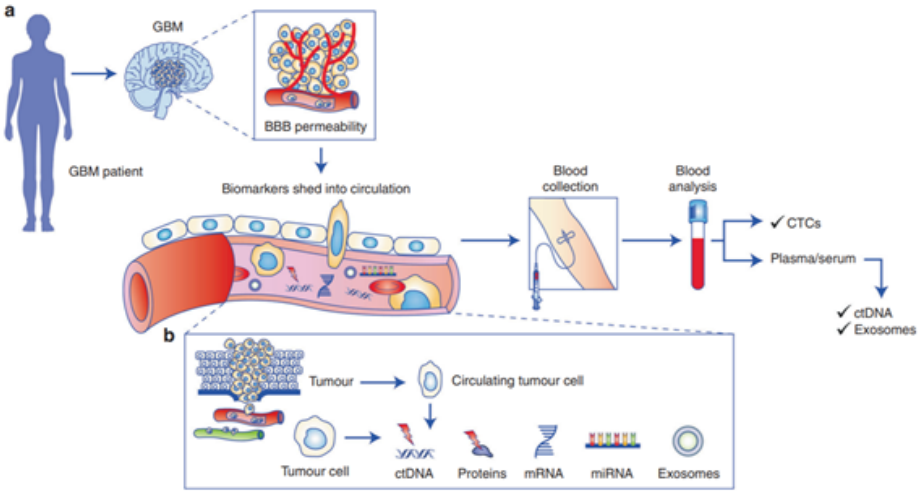

Figure 3.7: Schematic representation of GB tumor cells present in the bloodstream. (a) A schematic representation of the biomolecular transport of a tumor through the BBB to the circulation. (b) Representation of a tumor cell leak [10].

3.7 Epigenetics in GB

Recently, epigenetic studies have shown that epigenetic changes in the genome, including DNA methylation, histone modifications and microRNA synthesis are essential mechanisms for various physiological and pathological processes in the brain. DNA methylation is a fundamental process during brain development, and in some genomic sites it is possible to verify the variation of methylation levels in different regions of the brain [117, 118]. In addition, DNA methyltransferase enzymes are also essential in the development and function of the CNS. Methyltransferase DNMT1, a maintenance enzyme, has a high rate of expression in the mammalian brain [119, 120], including in post-mitotic neurons. The methylation of DNA and the trimethylation of histone H3 K27 mediated by polycomb (H3K27me3) have the function of regulating neuronal differentiation, where DNA methylation contributes to suppress pluripotency in neural progenitors compromised with the strain [121].

In humans, the manifestation of several neurodevelopmental disorders is due to the appearance of mutations in genes that encode proteins involved in epigenetic mechanisms. An example of this is Rett's syndrome. Doctors consider this pathology to be one of the most serious neurodevelopmental disorders. It arises from mutations in the MECP2 gene that encodes a protein that can bind to methylated DNA and regulate the expression of the gene [122]. Nevertheless, the dependence of the CNS on epigenetic regulation depends not only on DNA methylation and DNA methyltransferases, but also on mutations in genes that encode other epigenetic regulatory proteins and, as a consequence, trigger neurodevelopmental disorders. Another example is when mutations of the JARID1/SMCX genes occur, which encode a histone demethylase containing the JmjC domain, and the mutation of these causes a form of mental retardation linked to the X chromosome [122]. Therefore, these examples not only allow us to realize the importance of epi-

genetic control of gene expression in the development and function of the CNS, but also contribute to a better understanding of the role of epigenetic regulation in tumors arising from CNS cells.

As previously mentioned, the origin of glioblastoma multiforme (GB) occurs at the level of astrocytes. The treatment of patients with GB is multimodal aggressive and consists of surgery, radiation and chemotherapy. However, the prognosis of GB is extremely poor [123, 124], since health researchers characterize GB as a tumor that presents a diversity of genetic alterations, affecting genes that control cell growth, apoptosis, angiogenesis and invasion. The identification of a set of procedures underlying the development of GB has become one of the main objectives for the development of new treatments. The explanation for the causes and consequences of the different types of epigenetic dysregulation in GB and cancer in general continue to be a strongly investigated area, where there is already some progress in understanding the consequences. Epigenetic changes affect the expression of genes associated with the tumor alone or in combination with genetic mechanisms [125]. In the GB cell lines there was also a marked decrease in the expression of DNMT1 and a decrease in the expression of DNMT3a, with the decrease in DNMT3a inducing hypomethylation of Sat2 in the subpopulation of tumor stem cells [126]. It is assumed that an initial genetic or epigenetic abnormality can induce tumorigenesis and cause subsequent genetic and epigenetic changes. In the glioma, there is an association of the combined deletion of chromosomes 1p and 19q with the epigenetic silencing of the tumor suppressor gene WNK, lysine deficient protein kinase 2 (WNK2) [127]. This protein has the function of indirectly inhibiting MEK1, thus increasing the signs of growth promotion through EGFR [128]. That said, it is plausible that there is a huge involvement between the epigenetic silencing of WNK2 and the genetic alteration of EGFR signalling, a common abnormality in glioblastomas. Recently, a GB has been observed with a point mutation only in WNK2 [129], and commonly occurring cancers usually show point mutations in all four WNK genes. However, of the diverse epigenetic changes, aberrant methylation of gene promoters is the most widely studied alteration and occurs during oncogenesis.

DNA methylation has become a crucial tool in the classification and subgroup of tumors with distinct differences in prognosis and potentially contributes to improving diagnosis. Given the enormous variety of tumor clones and transcriptomic expression profiles previously described in GB tumors, there is an urgent need to understand and characterize the heterogeneity of intratumoral DNA methylation and how it affects biomarkers and classification based on methylation.

Heterogeneity increases the complexity of GB treatment, as the response to treatment using radiation and chemotherapy varies depending on the number of gliomas as well as the degree of them in the tumor. Currently, treatment for patients with GB is surgical resection followed by concomitant and adjuvant temozolomide (TMZ) and radiotherapy, which still results in only an average survival of 15 months. The 6-methylguanine-DNA methyltransferase (MGMT) is a DNA repair enzyme that neutralizes the damage induced by TMZ. The MGMT gene can be silenced by methylation at the promoter, and a survival benefit has been demonstrated for patients with methylated MGMT. However, this gene has been considered the main mechanism of patients' resistance to treatment. DNA methylation can also be used to determine the age of tissue methylation, which

is accelerated in tumors and to predict mortality in healthy individuals. The heterogeneity of intertumoral DNA methylation in GBM has been demonstrated by the existence of different subclasses of GBM with a different prognosis, as well as for gliomas of different degrees based on the isocitrate dehydrogenase 1 (IDH1) mutation and on the glioma with CpG island methylating phenotype (G-CIMP). Knowledge of the intratumoral heterogeneity in GBM at the level of DNA methylation is essential to obtain future improvements in tumor classification and biomarkers based on methylation, in order to avoid diagnostic errors due to tumor heterogeneity and sampling bias.

The addition of methylation on the CpG Island causes hypermethylation and promotes carcinogenesis, silencing tumor suppressor genes. In turn, the loss of methylation causes hypomethylation and increases the transcriptional activation of oncogenes, inducing chromosomal instability [130, 131]. These two mechanisms in GB will be better explained below.

1. **Hypomethylation of DNA in GB**

   In 1980, pioneering studies made it possible to observe a subset of GBs that showed a global attenuation of 5-methylcytosine [132]. Follow-up studies have recently shown that primary glioblastomas have a high frequency of hypomethylation (approximately 80%) [133]. In the different GBs, a variation in the level of hypomethylation can be seen, which causes a variation in brain levels and a possible demethylation of approximately 10 million CpG sites per tumor cell. The most proliferative primary GBs are those that have a high rate of hypomethylation, and these arise due to the demethylation and transcriptional activation of the putative oncogene MAGEA1 [133]. A satDNA is defined as a repeated sequence in tandem, that is, it presents in its constitution several units of the monomeric sequence, which in turn, is organized in the genomes in a matrix tandem, that is, generally in heterochromatic regions, highly repeated, and / or interchanged. The tandem repetition satellite 2 (Sat2) that is present in the justacentromeric regions of Chr 1, 9 and 16, and the D4Z4 repetition in the subtelomeric regions of Chr4q35 and Chr10q26, show a high rate of hypomethylation in a variety of cancers [134, 135]. In addition, hypomethylation of these sequences can also be observed in patients with immunodeficiency, centromere instability and facial anomaly syndrome (ICF), which occurs due to mutations in the new DNA methyltransferase DNMT3b [136]. Thus, in primary GBs with severe hypomethylation, dramatic Sat2 hypomethylation of a small percentage of the normal brain, and moderate D4Z4 hypomethylation of approximately 70% to 85% of the brain [133] are observed. Samples of GB with hypomethylated Sat2 that encompass changes in the number of copies of adjacent euchromatic sequences, more specifically located in the pericentromeric region of Chr1, revealed that a consequence of repetitive hypomethylated sequences in GB corresponds to the predisposition of chromosomal break age. On the other hand, hypomethylation of DNA in specific single copy genes can be observed. An example of this event is the specific hypomethylation of loci in gliomas in the testicular cancer antigen gene, MAGEA1. Hypomethylation was also observed in astrocytoma and GB [133, 137]. The MAGEA1 gene belongs to the MAGE gene family, and corresponds to a group of germline-specific genes

that are transcriptionally activated in various tumors, causing hypomethylation of the entire genome and an increase in cell proliferation [133, 138]. Expression of the MAGEA1 gene in tumors causes recognition of tumor-specific genes on the cell surface by cytolytic T lymphocytes and inhibits the function of p53 and the response to chemotherapy [139].

2. **Hypermethylation of CpG promoter island DNA in GB**

In GB, locus specific hypermethylation is observed, mainly in CpG island promoters (CGI). CGIs correspond to regions of CpG nucleotides of approximately 500bp to 1kb. In GB, hypermethylation of the CGI promoter is observed in genes that have several functions associated with tumorigenesis and tumor progression, which involve regulation of the cell cycle, DNA repair, apoptosis, angiogenesis, invasion and drug resistance. An example of this phenomenon is the retinoblastoma (RB), PI3K and p53 pathways that occur by hypermethylation of CGI promoters, including the CDKN2 / p16, RB, PTEN, TP53 and p14ARF gene promoters [140, 141, 142]. In addition, in GB, hypermethylation of the CpG Island can also be verified in genes that are not expressed in the brain. The treatment of glioma cells with a DNA demethylation agent and that exhibit cell expression, allowed researchers to discover a candidate tumor suppressor called epithelial membrane protein 3 (EMP3). EMP3 is a gene associated with myelin and causes cell proliferation and cell-cell interactions [143]. When primary gliomas provide hypermethylation, the EMP3 gene is silenced. Furthermore, in mice it was observed that the reintroduction of EMP3 in neuroblastoma cell lines with EMP3 silencing caused an attenuation of colony formation in vitro and a decrease in xenograft growth, suggesting tumor suppressor function [143]. In GBs, the regulation of the transforming growth factor (TGF)-beta oncogenic signalling pathway and the proliferation promoter is due to the hypermethylation of the promoter. Normally, a high concentration of TGF-beta signalling levels is associated with a poor prognosis. TGF-beta signalling induces proliferation, a consequence of the induction of platelet-derived growth factor (PDGF)-B. However, the proliferative effects, which increase with TGF-beta signalling, can be eliminated by PDGF-B epigenetic silencing. Over the years, it has been found that hypermethylation of the promoter can cause a sensitivity in the drugs used in the treatment and the respective radiotherapy in patients with GB. The methylation of the O-6-methylguanine-DNA methyltransferase (MGMT) promoter and the response to DNA alkylating agents are the best known examples where this phenomenon is observed. MGMT encodes a DNA repair protein whose function is to remove adults from alkyl at the O6 position of guanine and, in some cases, at the O4 position of thymine [144]. During the expression of MGMT the protection of normal cells from carcinogens is verified. In addition, the protection of cancer cells from chemotherapeutic alkylating agents can also be verified. Methylation of the MGMT promoter occurs for two reasons, the first of which has to do with expression and decreased transcription factor binding in GB cell lines [145].

Transcriptional silencing is the main source of several primary human tumors, including glioma, lymphoma, breast cancer, prostate cancer and retinoblastoma, silencing being the

consequence of hypermethylation of the MGMT promoter [146]. In human colorectal cancer, methylation of the MGMT promoter occurs due to the addition of the G to A transition mutations in TP53 and K-RAS [147]. In GB, methylation of the MGMT promoter is associated with the inability to repair incompatibility and a hypermutator phenotype [148]. Normally, MGMT hypermethylation is associated with the survival time of patients with low-grade GBs and gliomas treated with radiation and alkylating agents, including temozolomide [149, 150]. This hypermethylation is duo to increased proliferation and pseudo progression of tumor and reduction of MGMT expression [151]. Pseudo progression is defined as progressive and enhanced lesions that can be seen during MRI immediately after the end of treatment, thus concluding that this result is not due to the actual progression of the tumor, but rather to an effect of radiotherapy and treatment with temozolomide [151].

Nevertheless, these two mechanisms play a fundamental role in tumor progression. It is observed that different subtypes of glioma and tumor degrees can exhibit various aberrant DNA methylation profiles [152]. Secondary GBs typically exhibit a higher frequency of promoter methylation when compared to primary GBs, mainly for p14ARF, p16INK4a, RB1, MGMT and TIMP-3 [153] promoters. In low-grade gliomas and secondary GBs, methylation of the PTEN promoter and activation of the PI3K pathway occurs due to phosphorylation of protein kinase B (PKB / AKT). nevertheless, the origin of methylation of the PTEN promoter is rare in primary GBs [154]. However, it is possible to verify the sharing of the same epigenetic changes in different types of brain tumors, even in different species. For example, inactivation of the SLC5A8 gene can occur in human oligodendrogliomas and astrocytomas and in mouse models of oligodendroglial tumors by an epigenetic mechanism, whereas aberrant epigenetic silencing of this locus is common in mammalian gliomas [155]. In addition to gliomas, other CNS cancers may also exhibit various methylation profiles, such as hMLH1, TIMP3, MGMT, p73 and THBS1, which are regularly hypermethylated in schwannomas [156], and NF2 [157]. However, over time, the progression of these tumors is due to several distinct epigenetic patterns.

Hypermethylation of DNA may also be one of the consequences of altering the differentiation properties of the cell fraction in gliomas. These properties can be considered putative cancer stem cells or also called tumor-initiating cells (TICs). The regulation process of normal astroglial differentiation is carried out by bone morphogenetic protein (BMP) and by the ciliary neurotrophic factor (CNTF), via Jak / STAT mediated. In a subset of glioblastoma ICTs, differentiation is inhibited by epigenetic silencing of the BMP receptor 1B gene (BMPR1B) [158]. This silencing comes from the zest 2 histone methyltransferase intensifier (EZH2) and DNA methylation, and may arise due to treatment with a DNA demethylation agent, thus allowing the observation of the blocking of a subset of human GB cells in differentiation by default in the BMP signalling pathway. In addition, forced expression of BMPR1B can cause a potential differentiation of these cells and, as a consequence, decrease tumorigenicity. The gene that encodes the cell surface marker CD133 used in the enrichment and identification of GB TICs can be subject to epigenetic regulation. In GB cell lines, a negative link is observed between methylation of the CG13 DNA of the CD133 promoter and the expression of CD133 [159]. In addition, it is not possible to verify

CD133 methylation in the normal brain. This process can only be observed in culture systems and primary tumors, becoming a tumor-specific epigenetic marker.

Unlike DNA methylation, changes in histones in GB are still a poorly studied area. An aberrant standard feature of histone modifications is epigenetically silenced loci. The addition of methylation and loss of acetylation of histone H3K9 silences promoters on the CpG Island. It is said that in embryonic stem cells (ES), which present a double presence of methylation in inactivated histone H3K27 and a methylation in histone H3K4 associated with activation, called bivalent domains, there is the formation of a chromatin state " balanced "for genes regulated by development. The formation of this phenomenon causes silencing in ES cells and subsequent transcriptional activation or repression in differentiated cells [160]. Hypotheses about the changes that take place at the level of histones suggest that these changes occur due to tumor suppressor genes to DNA hypermethylation and silencing of hereditary genes [161]. The gene that encodes BMI-1, an element that is part of the complex of the polycomb group that regulates the methylation of histone H3K27, can frequently undergo changes in the number of copies in low and high grade gliomas, and usually the deletions of BMI- 1 are associated with a poor prognosis in patients [162]. GB shows epigenetic changes in histones associated with the expression of some histone deacetylase (HDAC) proteins. HDACs have the function of catalysing the deacetylation of lysine residues within the N-terminal tails of central histones and in non-histone proteins. Typically, HDACs provide a closed chromatin structure, inducing transcription. There are a total of about 18 HDACs known in humans. These are divided into 5 main classes, with different target principles [163]. The set of Class I HDACs (HDACs 1, 2, 3 and 8), Class IIA (HDACs 4, 5, 7 and 9), Class IIB (HDACs 6 and 10) and Class IV (HDAC 11), have in place active zinc and can be inhibited by HDAC (HDACi) TSA inhibitors and suberoylanilide hydroxamic acid (SAHA; Vorinostat). Unlike the previous set, Class III HDACs (sirtuins) neither have zinc nor suffer from inhibition by TSA or SAHA. In GB, class II and IV HDACs show a decrease in mRNA expression, when compared to low-grade astrocytomas and normal brain. A high rate of histone H3 acetylation [164] is also observed in GBs. The cause of the appearance of the involvement of many genes in mutations in epigenetic regulation, including histone deacetylases HDAC2 and HDAC9, histone demethylases JMJD1A and JMJD1B, histone methyltransferases SET7, SETD7, MLL, MLLD1 and protein binding domain 1 MLLD1 (MBD1 methyl-CpGG) is due to the large-scale sequencing of genes encoding proteins in GBs [165]. These intriguing initial studies have led to the conclusion that changes in epigenetic mechanisms in histones may be an important defect in GB.

In addition to the epigenetic changes that play an important role in explaining some early stages of GB development, they can also function as biomarkers during the treatment of this tumor, as is the case with histone deacetylase inhibitors (HDAC). The use of HDACi in cancer therapy can be done in two ways. First, HDACi induces the formation of a more open or "relaxed" chromatin, making the gene more available for transcription, and can therefore more easily allow DNA damaging agents to chromatin, thereby increasing the sensitivity to death by these agents. Second, HDACi prevents silencing of the aberrant epigenetic gene in GBs, providing an increase in the cell cycle and apoptosis of agents that damage DNA [166]. As previously mentioned,

HDACi have in their constitution several classes of compounds, including hydroxymates (SAHA, TSA), cyclic peptides (depsipeptide), aliphatic acids (valproic acid, butyrate) and benzamides. No HDACi is highly effective against all HDACs. Normally, HDACi cause an increase in acetylation in histone and non-histone proteins, and p21 reactivation may occur, thus contributing to the rest of the stomach cell cycle [167, 168]. As a rule, there is a greater resistance of cancer cells to the effects of HDACi, however the reasons for this selective sensitivity are not yet clear [133]. HDAC inhibitors modify the expression levels of only a subset of genes expressed in transformed cells [169, 170]. Recently, in a testing phase, SAHA was used as monotherapy or combined therapy in 5 phase I or I - II clinical trials for gliomas. Pre-clinical studies of SAHA have found that it sensitizes glioma cells in vitro, ex vivo and in vivo to chemotherapy and radiation [171]. Treatment with SAHA induces increased acetylation of histone H3 in the p21 promoter in U87 glioma samples [167]. Recently, they tested valproic acid against GB in combination with Temozolomide and then applied radiation. The second test involves, again, valproic acid against neuronal tumors and brain metastases in combination with etoposide. Valproic acid plays an active role against Class I and IIA HDACs in millimolar concentrations, while the depsipeptide only plays an active role against Class I HDACs in nanomolar concentrations. There are several HDACi that are successful against cancer cells, but have not yet been tested for gliomas, including Panobinostat (LBH589) [172] and Belinostat (PXD101) [173]. Nevertheless, the discovery and development of new epigenetic compounds mainly in enzymes has proved to be a very active area of research in the pharmaceutical industry, and this is mainly due to two reasons. First, some genes that require DNA methylation or deacetylation of histones for silencing in normal cells can cause involuntary activation by agents that inhibit DNMTs or HDACs. Second, cancer genomes can be characterized by hyper and hypomethylation of DNA. Thus, the use of drugs that cause reactivation of silenced tumor suppressors may suffer as a consequence of the unwanted activation of oncogenes through hypomethylation. That said, there is an enormous need to solve these problems in order to obtain a more complete understanding of the molecular events resulting from epigenetic based therapy.

3.8 Three-dimensional (3D) cell culture systems of GB

Cell culture is an essential in vitro tool to maintain live cells used in the laboratory, regardless of the organism that formed them [30]. This technique helps us to better understand the biology of cells, the development of engineering and tissue morphology, disease mechanisms and the action of drugs. There are two types of cell cultures, the culture of two-dimensional (2D) or monolayer cells and the culture of 3D or spheroid cells. 2D cell cultures represent the most widely used drug discovery culture and is grown in monolayer for cell fixation and growth [11] (figure 3.8). However, this has limitations mainly in the area of oncology, the main one being the fact that they grow in monolayer or in petri dishes, thus not allowing to observe, in most cases, the heterogeneity of tumors. In order to overcome these limitations, 3-dimensional cultures were developed (3D) (Figure 3.8) [11].

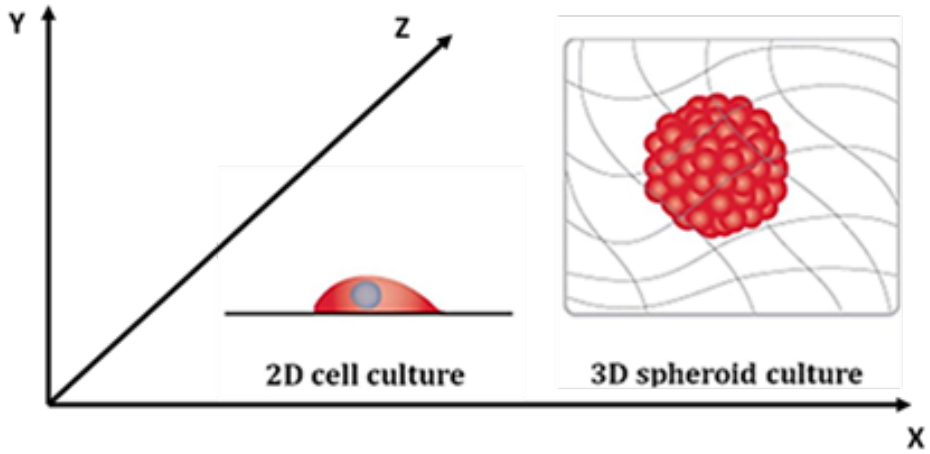

Figure 3.8: Simplified representation of 2D and 3D cultures [11].

The purpose of this type of cultures is to allow cells to explore the three dimensions of space (x, y and z), thus increasing the cell-environment and cell-cell interactions. The culture of 3D cells can be performed within a scaffold [174]. Typically, the types of scaffold used are hydrogels or inert matrices. Regarding cultures without scaffolds, these only depend on the clustering of the cells and can be made in low-adhesion plates, surfaces and in suspended drops. The liquid overlap technique (LOT) represents the growth of 3D cells in low adhesion plates [175]. This is one of the most used techniques in 3D cell culture. Here, cells are forced to group together due to the limited adhesion of certain polymers or biomaterials. The use of these polymers allows obtaining artificial matrices used for the formation of spheroids. An example of a polymer used is agarose, which is extracted from seaweed that contains agar [176], has non-adhesive properties and provides cells with adequate moisture and nutrition to keep cells alive. After obtaining the artificial matrix, cells involved in culture medium are added to obtain a cluster of cells. In this process, it appears that the aggregation of cells is due to the adhesive forces between the cells being stronger than the forces that are established between the cells and the non-adhesive biomaterial [177]. Nevertheless, the spheroids formed can be loose with irregular or firm surfaces, and have different dimensions from culture to culture, since they vary depending on the number of cells that are cultured and the cell type. These present a cellular heterogeneity inside, differential exposure to several factors, such as oxygen and nutrients, and microenvironment conditions, such as hypoxia [177]. Therefore, the development of spheroid cultures has allowed great advances to be made in the area of oncology, since they have characteristics almost identical to cancer cells, thus allowing the development of therapies for the medical screening of patients.

GB is a very resistant tumor and, in order to develop therapies to combat this problem, experiments and research have been carried out over the last few years with 3D spheroid cultures of GB cells. In 2015, Maria Vinci et al. developed a protocol to grow 3D glioblastoma cells [178], since this tumor presents itself as one of the most deadly tumors in the world. They cultured a

human glioblastoma cell line in a 96-well plate to check the level of resistance it has against epidermal growth factor receptor (EGFR) inhibitors, when associated in a matrix. More recently, in 2019, M.E. Oraiopoulou et al. demonstrated by culturing 3D cells from the T98 GB secondary cell line, that they exhibit 3D invasive and proliferative features and therefore can serve as control cell lines for the in vitro 3D study of primary GB cell cultures [178]. In the same year, D. Sood et al. developed, in the scope of bioengineering, a 3D brain tissue platform, which integrates microenvironmental clues from native ECMs derived from brain cancer cells, namely, pediatric ependymoma cells and GB [179]. This platform also visualizes the cells live, thus allowing to systematically evaluate the responses to brain tumors derived from patients. Therefore, nowadays, the development of GB 3D spheroid cultures becomes a crucial tool for the biological monitoring of GB patients, to reduce the mortality rate through the development of new personalized therapies and high-throughput screening.

Chapter 4

Epigenetics and molecular biology of cancer

4.1 Epigenetics

The authors James Watson and Francis Crick discovered the DNA double helix model [179], allowing researchers, in the field of biology, to make great strides in explaining the “Central dogma of molecular biology” in the transcription of genetic information inserted into DNA and RNA, and their translation into proteins [180]. After half a century has passed, the term "central dogma of molecular biology" remains the main interest of researchers, especially in the field of gene expression. Gene expression is essential to elucidate the interactions between the environment and the genome. Genomic plasticity defines the ability of a given genotype to confer different phenotypes, that is, different changes in the presence of different factors, such as DNA gain, loss or rearrangement. Epigenetic mechanisms control all of these changes through gene expression. Thus, epigenetics refers to inheritable changes in gene expression that occur without modifying genetic information, that is, the DNA sequence [181]. The word epigenetics first appeared in 1942 by Conrad Waddington, referring to all the mechanisms necessary for the unfolding and development of genetic processes [182]. In 1958, Nanney decides to extend this concept beyond developmental biology, transmitting that the phenotype would be the product of two systems, one genetic and the other epigenetic, the latter being structured by constituents other than nucleic acids [183]. That said, the epigenetic designation has presented several definitions, and in general, all are exposed to inheritable factors that do not harm the sequence of nucleotides and result in the regulation of gene expression. According to the aforementioned concept, two types of mechanisms are established, a trans action mechanism of self-propagation and another by molecular signatures of cis action physically associated with the DNA sequence that they regulate (Figure 4.1) [184]. The most common type of trans epigenetic situations corresponds to the transcriptional states of self-propagation that are preserved through feedback loops and transcription factor (TFs) networks [185]. When a TF activates its own transcription (or represses antagonistic networks) it creates an epigenetic state that is self-sustaining after the elimination of the original stimulus. After each cell division, the

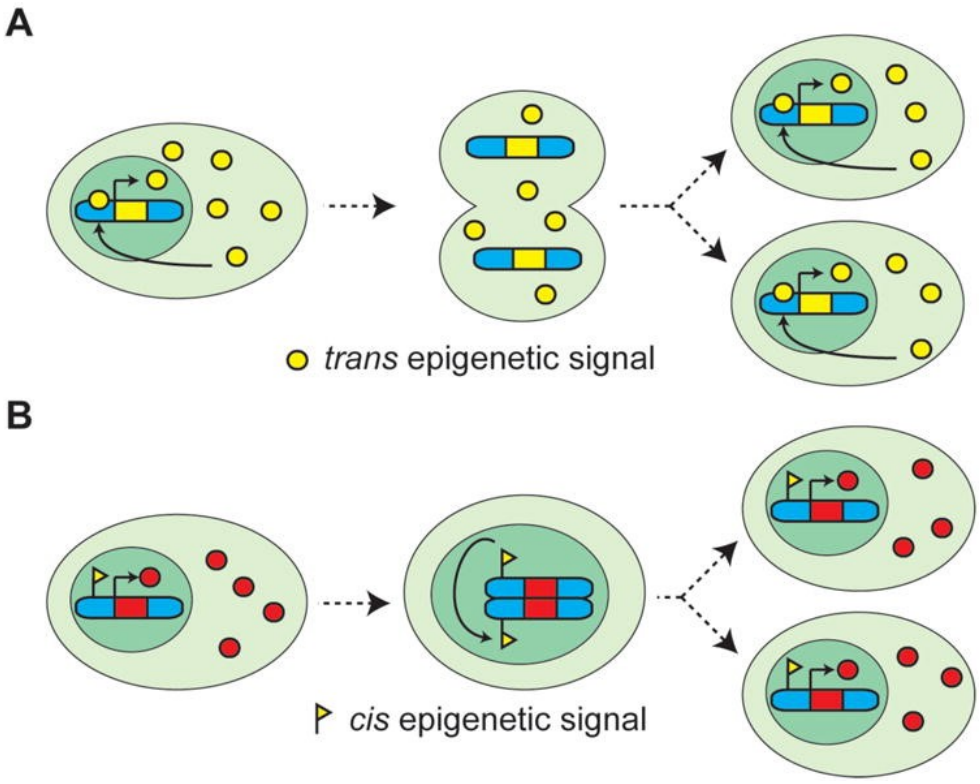

Figure 4.1: Trans and cis epigenetic signals. (A) The yellow circles correspond to trans epigenetic signals that are transported by dividing the cytoplasm during cell division and maintained by feedback cycles. (B) The yellow flags represent the cis signals, where they are molecular signatures physically linked to DNA and inherited via chromosomal segregation during cell division [12].

inherited TFs regain their trans function in the regulatory DNA sequences. However, cis epigenetic signals are physically associated with the chromosome, which are then conducted to the next generation through the segregation of chromosomes during cell division, such as, for example, covalent modification of the DNA itself, DNA methylation, or through changes in the structure of histones [184]. Histones are chromatin proteins and can transmit information from their primary sequence in post-translational modifications which are often present in the terminal tails N and C [186, 187]. In addition to the chromosome, cis epigenetic information can also be encoded in chromatin through the stable association of non-histone proteins (higher order chromatin structure and nuclear localization).

The most classic epigenetic marks are: DNA methylation, post-translational modifications in histones, non-coding RNAs and positioning of nucleosomes. In eukaryotic beings, the nuclear genetic material is compacted in a structure formed by DNA and proteins called chromatin, which is composed of repetitive units called nucleosomes. Each nucleosome consists of eight histone discs (octameter) H2A, H2B, H3 and H4 surrounded by 146bp of DNA [188]. The ordering of nucleosome positioning is directly related to the accessibility of chromatin to regulatory elements of gene expression [189]. This ordering can be controlled by other epigenetic marks, such as chro-

matin remodeling factors and specific DNA sequences [190]. In 1998, Fire, Xu et al. described the role of non-coding RNAs (ncRNA) in the regulation of gene expression [191]. Some steps are proposed for these elements, among them, the recruitment of epigenetic marks linked to transcriptional silencing, the blocking of transcription by preventing the binding of RNA polymerase II and the decrease in the amount of messenger RNA [192]. NcRNAs act mainly on gene silencing, however they are also associated with the activation of some genes [193]. Nowadays, the epigenetic marks that are most studied are DNA methylation and post-translational modifications of histones. Both act in cooperation with the other epigenetic brands, but they are responsible for ultimately forming all the essential steps for the establishment of epigenetic control.

4.2 DNA methylation

In 1975, Holliday and Pugh suggested that methylation in specific adjacent DNA sequences was the main epigenetic marker responsible for controlling gene transcription [194]. Nowadays, it is said that this process is responsible for silencing genes by inhibiting the binding of transcription factors [195]. The DNA methylation process takes place at the level of cytosines (C), more precisely in three different nucleotide sequences: CG, CHG and CHH (where H = C, T or A), the last two being common in plants, but also found in mammalian stem cells [196]. This mark occurs when a methyl group, consisting of a carbon atom attached to three hydrogen atoms, is added covalently to the 5-carbon position of the cytosine ring in the genomic DNA by DNA methyltransferase enzymes (DNMTs) [197]. Nevertheless, in humans, a large part of the DNA methylation process occurs in the CPG islands (the "p" refers to the phosphate between nucleotide bases), namely at the level of the CpG dinucleotides in somatic cells, with only a small part of the methylation appearing in a non-CpG context in embryonic stem cells [198]. Typically, CpG dinucleotides are within CpG clusters or islands, which are not methylated and are associated with gene promoters. During the formation of the zygote, the DNA methylation process does not take place, since it is removed. In mammals, DNA methylation performs several important functions that are essential in normal development. These functions are genomic imprinting, inactivation of the X chromosome and suppression of transcription and repetitive transposition of elements that, when deregulated, contributes to diseases such as cancer [197, 199, 200, 201]. Normally for these functions, the CpG islands present differential methylation, that is, one of the two chromosomes is completely methylated in the sequence of the CpG island, and in opposition there is no methylation of the same sequence in the other chromosome. CpG islands can have hundreds to tens of thousands of base pairs along their length. The regions of the genome that are associated with 5mC (5-Methylcytosine), especially the promoter regions, present themselves in a transcriptionally repressive state and are referred to as heterochromatin, when they are found together with histones marked with repressive changes [202]. However, euchromatin refers to the process involving regions that are associated with unmethylated cytosines and the respective activation of successive changes in histones when they are in a permissive state of transcription [202]. When DNA methylation occurs in the promoter region of a gene, there is an inversion in the correlation

of the methylation level with the gene transcription, by means of an uncontrolled impediment, and the chromatin is compacted, which physically prevents some proteins from DNA binding to reach access to their recognition sequence (s). CG dinucleotides and CpG islands exhibit transcriptional activity and can be located within genes [203]. The choice of the union site among CG dinucleotides, present in the vicinity or in the sequences of the union site, may be influenced by the methylation status, leading to the production of alternative isoforms in the transcript [204, 205]. Intergenic DNA methylation is a common process and occurs in the presence of retrotransposon elements, whose sequences tend to be rich in CG. Retrotransposon elements are genetic elements that amplify themselves throughout the genome [206]. It is believed that the constant methylation of CG-rich sequences eliminates the activity and its respective ability to transpose to other regions of the genome, thus becoming catastrophic for a cell [207]. It is possible to verify the occurrence of another intergenic methylation along the stimulating or insulating elements [208]. In that case, methylation may be a signal to recruit proteins, such as proteins from the methyl binding domain (MBD) that bind to DNA containing methylated CGs inserted in the recognition sequence [209]. Nevertheless, the methylation of CGs that is incorporated into the binding sequence can block the binding process of proteins that have a more sensitive binding to methylation. If this process described above takes place in a repressive element, greater transcription than normal can be seen, thus representing a rarer phenomenon of DNA methylation which is positively correlated with gene expression [210]. The DNA methylation process is coordinated by the enzymes DNMTs (Figure 4.2): DNMT1, DNMT2, DNMT3A, DNMT3B and DNMT3L [211]. DNMTs eliminate the methyl group from donor SAM (Sadenosylmethionine) and bind it to cytosine [212]. DNMT1 is the enzyme responsible for maintaining DNA methylation in cell division, more precisely in the S phase, where this methyl CpG dinucleotides in the new DNA sequence after replication [213]. The DNMT3A, DNMT3B and DNMT3L enzymes belong to the DNMT3 family. This family differs from DNMT1, since they have N-terminal domains that are shorter than those of DNMT1 and need regulatory elements that give specificity to hemi-methylated DNA. Regarding DNMT3A and DNMT3B, these are called new methyltransferases and are essential for the occurrence of new methylation in the formation of gametes, in the acquisition of the methylation patterns of imprinted genes (genes that are expressed only by one allele, while the other allele is methylated) and during early embryonic development [214]. DNMT3L is catalytically inactive, but has the function of co-regulating DNMT3A and DNMT3B. DNMT3A is located in the pericentromeric heterochromatin, while DNMT3B is diffusely located in the nucleus [215]. As for the expression of DNMT3A and DNMT3B, there is a greater expression of these enzymes in undifferentiated embryonic cells, in relation to differentiated and adult tissues [216]. DNMT3L is more expressed during gametogenesis [217]. However, DNMT2 corresponds to a methyltransferase homologue that methylates cytosine at position 38 in the anti-codon loop of aspartic acid transfer RNA instead of DNA.

Throughout the mammalian life cycle, there are two stages of demethylation, passive or active, followed by remethylation, where the first occurs during cell division, that is, during the formation of gametes, and the second, after fertilization. As one of the functions of the DNMT1 enzyme is

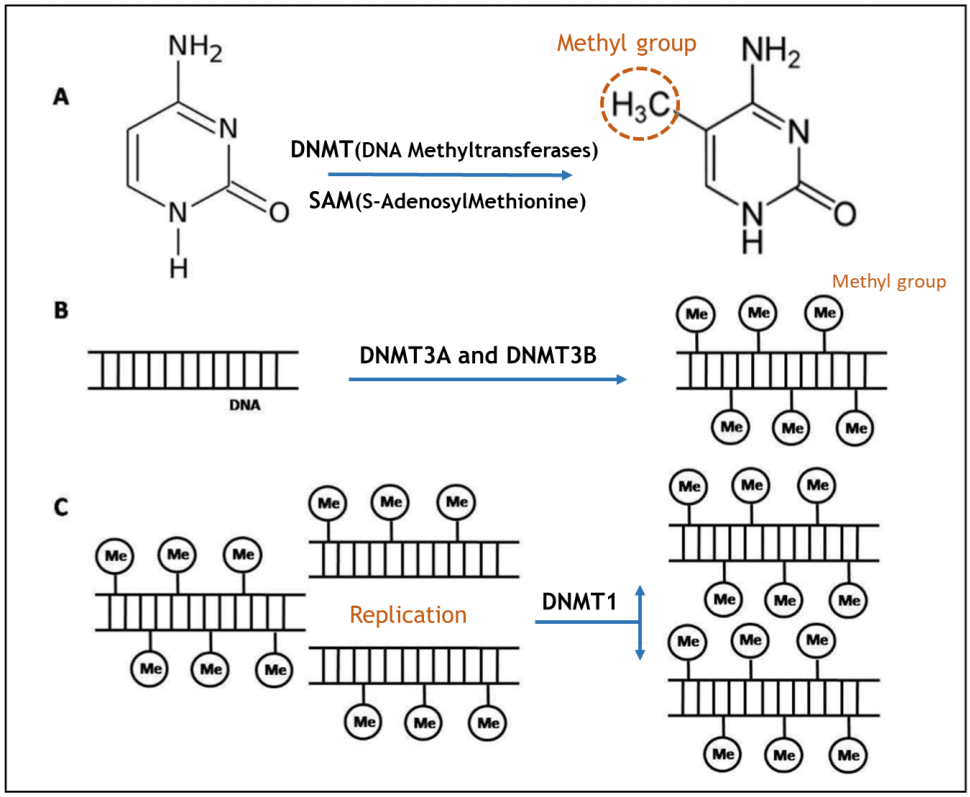

Figure 4.2: DNA methylation. (A) Cytosine carbon-5 receives a methyl group by the action of DNMTs in the presence of SAM. (B) The enzymes DNMT3A and 3B are responsible for methylation back into DNA. (C) The maintenance of methylation is guaranteed by DNMT1, which acts on hemi-methylated DNA (adapted from Turek-Plewa and Jagodzinski, 2005).

to actively maintain DNA methylation during cell replication, its inhibition or dysfunction causes the newly incorporated cytosine to remain unmethylated and, thus, the overall methylation level is reduced after each cell division. In addition to observing the demethylation of active DNA in cells that have undergone division, this process can also occur in cells that have not undergone division. However, this last process requires enzymatic reactions that modify 5mC, by deamination or oxidation reactions, in order to revert it to a single Cytosine [218]. Regarding remethylation, it starts first in the male germline, before birth, while in females it starts after birth [219]. After the fertilization period, successive modifications to the chromatin occur in the zygote, mainly in the genome. This phase, demethylation occurs by an active mechanism before DNA replication, namely through the oxidation of methylcytosine [220]. On the other hand, the maternal genome occurs through passive demethylation, through the absence of maintenance of methylation [219].

4.3 Changes in histones

The concept of changes in histones first appeared in 1964, by Allfrey, who described the role of histone acetylation and methylation in the regulation of gene expression [221]. From this

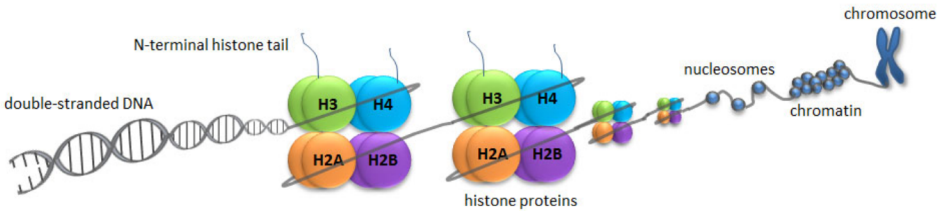

Figure 4.3: Schematic representation of the organization of genetic material. Nucleosomes are represented by DNA (gray) wrapped around eight histone proteins, H2A, H2B, H3 and H4 (colored circles). In histones H3 and H4, the tails of N-terminal histone (blue) can be seen [13]

moment on, new modifications were found in histones, also involved in the control of gene expression, giving rise to what became known as the "Hypothesis of the histone code" [222]. This hypothesis proposes that the involvement of different combinations of histone modifications can regulate the chromatin structure and the state of transcription [221, 223]. These modifications are post-translational and aim to restructure chromatin in several ways, including phosphorylation, ubiquitination, acetylation and methylation [224]. Nucleosomes are basic units of chromatin and correspond to the junction of eight histones plus approximately 150 base pairs of DNA (Figure 4.3) [13]. Nevertheless, information such as the location of the nucleosomes in relation to the place of the beginning of the transcription of a gene, together with specific combinations of sites, types and extensions of histone modifications, allow to obtain an enormous complexity of the histone code [225].

Acetylation and methylation are the modifications in histones that are most studied in epigenetic studies. The role of methylation in relation to the control of genetic transcription may vary according to the addition of residues and methyl groups. The trimethylation of lysine 4 from histone H3 (H3K4me3) is an example of methylation that is associated with eukromatin, while the trimethylation of lysine 27 from histone H3 (H3K27me3) is associated with heterochromatin. The chromatic decondensation is an acetylation mechanism [226]. The addition of these epigenetic marks alter the biophysical properties of the nucleosome through neutralization or addition of charge, inducing an increase in its mobility, modulating the contact between nucleosomes and thus allowing an action by the chromatin remodeling proteins [227]. Over the years, several researchers have identified different histone-modifying proteins and "readers" of the histone code. Histones methyltransferases and histones demethylases are catalyst enzymes responsible for the addition and removal of methyl groups from the amino acids lysine (K) and arginine (R). Methylation in a histone is read by proteins with chromium-like domains, MBT (Malignant Brain Tumor), Tudor and PhD (Plant Homeodomain) [224]. Another group of histone-modifying proteins are histones acetyl transferases (HAT) and histones deacetylases (HDAC), which are responsible for verifying a balance between acetylation and lysine deacetylation, respectively. These enzymes are not as specific as those involved in methylation and may have substrates other than histones [224]. During the histone acetylation process, the ε-amino group of the lysine residues at the H3 and H4 terminals is associated with the promotion of transcription. Newly synthesized histones can

receive acetylation either in the cytoplasm, by the action of a type B HAT (HAT B) as a newly synthesized protein signal, or in the nucleus, which is associated with transcription and is performed by type A HAT (HAT A) [228]. However, this process is too complex, since the chromatin structure when acetylated can also allow access to transcriptional repressors. For example, some factors containing bromodomains, such as BRG1 and Brd4, which are defined as protein domains of approximately 110 amino acids that recognize monoacetylated lysine residues, and are present in the histone N-terminal tails. Bromodomains target acetylated histones, where they can mediate the formation of repressor complexes (or activators) [229]. Deacetylation of histones is associated with the process of CpG methylation and the inactivated state of chromatin. According to yeast homology, HDACs are divided into four classes (I to IV). Classes I, II and IV are classified as zinc-dependent, while class III is classified as NAD+ dependent. HDACI is present in the nucleus and is widely expressed, while classes II, which transits between nucleus and cytoplasm, and IV, which transits in the nucleus, have a tissue-dependent relationship [230]. In the inhibition of HDACs, several molecules of natural or synthetic origin are related to this phenomenon, such as hydroxamic acid (SAHA) and tricostatin (TSA). However, they are not able to inhibit all classes of HDACs [231].

4.4 Epigenetic processes

4.4.1 X chromosome inactivation

Over the years, the inactivation of the X chromosome has become an important research model and tool do study epigenetics. This is an essential mechanism that occurs in all somatic cells of females. In organisms that have a chromosomal sexual determination, there is an imbalance of genetic products between the sexes, allowing this phenomenon to establish a selective pressure for the formation of a dosage compensation mechanism [232]. In humans, the X and Y sex chromosomes indicate the sex of an individual. Women have two X (XX) chromosomes, while men have one X and one Y (XY) chromosome. All Y chromosome genes are necessary for male development, whereas X chromosome genes are necessary for both male and female development. Women inherit two copies of many of the genes needed for normal functioning, and this is because they are given two X chromosomes. However, extra copies of genes or chromosomes can trigger abnormal development. An example of this is Down syndrome, which is caused by an extra copy of part or all of chromosome 21. In female mammals, a process called inactivation X has evolved to compensate for the extra X chromosome. In X inactivation, each cell 'turns off' one of its X chromosomes [233]. Inactivation of the X chromosome can be described in two forms: imprinted and random. In imprinted form, the inactive X chromosome always has the same parental origin. This occurs in all tissues of marsupials [234], and in extraembryonic tissues of mice [235] and cattle [236], in which the paternal X chromosome is inactivated. On the other hand, random inactivation occurs in embryonic tissues of eutheric mammals and also in extraembryonic tissues in humans [237], where one of the two X chromosomes can be inactivated [238]. The X chromosome

inactivation process can be divided into steps such as initiation, which includes pairing, counting and choosing the X chromosome to be inactivated, propagating the inactivation signal, and finally maintaining silencing in the descending cells [239]. In the initial phase of the inactivation process, the relative number of X chromosomes is counted by autosomes in each cell, so that only one X chromosome remains active per diploid genome. These counting and choosing steps occur simultaneously and involve elements of the X chromosome in the imprinted form [240]. To protect the dosage of the gene, in women with two X chromosomes, from the formation of a potentially harmful duplication, it is necessary to intervene with a silencing mechanism, which involves the production of a long non-coding RNA (17 kb) called transcript specific inactive X (XIST), along with DNA methylation. XIST has a positive regulation on the inactive X chromosome of a region called the X inactivation center (XIC). It then lines the cis chromosome to initiate gene silencing. During epigenetic reprogramming in early life development, DNA methylation plays an important role in inactivating the X chromosome, since it occurs in the CpG islands of the promoter on the inactive X chromosome [241]. Although there are multiple mechanisms for silencing the X chromosome gene, only 25% of the genes in the inactive X are able to escape inactivation and are expressed bialelically [242]. The promoter regions of the escaped genes are not methylated [243]. It is said that there is double the amount of methylation in the active X chromosome than in the inactive X, and most of this methylation is inserted in the genes, promoting an increase in the expression of these genes [244].

4.4.2 Genomic Imprinting

In DNA methylation, the expression of a distinct group of genes depends on the DNA methylation profiles and the gender of the parent from which each allele was inherited. These genes are assigned the term genomically printed genes. Printed genes are responsible for regulating growth and development, particularly during pregnancy [245]. The normal inactivation of an allele in a group of genes implies that that group is vulnerable to the occurrence of possible genetic or epigenetic changes that silence the active copy or activate the silenced copy. This process is called "loss of impression" and is associated with many malignant diseases [246]. In humans, there are about 100 confirmed printed genes, although studies related to bioinformatics show that there are many more [247, 248]. The determination of DNA methylation samples that establish and maintain the impression state of the genes occurs during gametogenesis [249]. During this process, the DNA methylation profile of the previous generation is eliminated in the primordial germ cells. Next, these profiles are repaired so that all subsequent spermatozoids and oocytes produced have methylation profiles at these sites. These profiles are stable and, in some cases, are transmitted by all subsequent cell divisions. Nevertheless, these profiles are also referred to as differentially methylated regions (DMRs), since these in the same sequences differ between the chromosomes derived from the parents. A large proportion of DMRs serve as print control regions (ICRs), which exercise control over a particular group of printed genes. After fertilization, the new methylation profiles inserted in the gametes are able to resist reprogramming, since it removes methylation

marks throughout the genome. These profiles are then transmitted along the division of the somatic cells into the three germ layers and differentiated tissues during development.

4.5 Epigenetic in Cancer

The change in the population's lifestyle and the adoption of unhealthy lifestyles have become relevant factors in the advancement of public health problems for both developed and developing countries. Cancer is a multifactorial disease that encompasses both genetic and epigenetic changes, being responsible for a large percentage of all causes of death in the world. The epigenetics of cancer are genetic modifications that play an important role in the initiation and progression of cancer [250]. These epimutations occur due to global changes in DNA methylation and histone modification patterns, as well as altered expression profiles of chromatin-modifying enzymes. The development and progression of these malignant diseases are the consequence of these epigenetic changes, which result from the global deregulation of the gene expression profiles [251]. One consequence of epimutations is the silencing of tumor suppressor genes independently. Epimutations can inactivate tumor suppressors and can also promote tumorigenesis by activating oncogenes. The events that trigger these epigenetic abnormalities are not yet fully understood. However, as epigenetic changes and genetic mutations are mitotically hereditary, they can be selected from a population of cancer cells that have rapid and uncontrolled growth.

4.5.1 DNA methylation in cancer

The first identified epigenetic changes, which induce cancer initiation and progression, occur during DNA methylation [252]. Hypomethylation occurs throughout the genome, more precisely in several genomic sequences, including repetitive elements, retrotransposons, poor CpG promoters, introns and gene deserts [253]. During hypomethylation of DNA in repeated sequences, there is an increase in genomic instability, due to the promotion of chromosomal rearrangements [254]. However, the hypomethylation of retrotransposons also triggers genomic instability, due to its activation and translocation to other genomic regions [255]. An induction of genomic instability by hypomethylation can be observed in patients with immunodeficiency, centromeric region instability and facial anomaly syndrome, where it presents a germline mutation in the DNMT3B enzyme, resulting in hypomethylation and later chromosomal instability [256]. In addition, DNA hypomethylation can also trigger the activation of growth-promoting genes, such as R-Ras and MAPSIN in gastric cancer, S-100 in colon cancer and MAGE in melanoma [257], and a loss of impression in tumors [258]. Unlike hypomethylation, which increases genomic instability and activates proto-oncogenes, specific local hypermethylation contributes to the silencing of tumor suppressor genes. The discovery of hypermethylation of the CpG island of the Rb promoter, a tumor suppressor gene associated with retinoblastoma [259], allowed researchers to discover several other tumor suppressor genes, including p16, MLH1 and BRCA1, which suffer from tumor-specific silencing by hypermethylation [260]. Nevertheless, the involvement of these genes in cellular processes such as DNA repair, cell cycle, cell adhesion, apoptosis and angiogenesis, can

trigger the development and progression of cancer. In addition to the direct inactivation of tumor suppressor genes, DNA hypermethylation can also silence genes, transcription factors and DNA repair genes. The methylation process of tumor-specific CpG islands takes place through a sequence-specific instructional mechanism, where DNMTs are directed to specific genes associated with oncogenic transcription factors. An example of this process is aberrant hypermethylation and silencing of specific gene promoters by the PML-RAR fusion protein in acute promyelocytic leukemia [261]. Normally, hypermethylated regions in cancer are pre-marked with the polycomb mark H3K27me3 on stem cells [262], where this process results in a link between regulation of development and tumorgenesis. This phenomenon allows to partially explain the theory of the 'CpG island methylation phenotype' or CIMP, which assumes that there is a coordination of the methylation of a subset of CpG islands in tumors [263].

4.5.2 Modification of histones in cancer

Recent studies in high-throughput sequencing have enabled researchers to make major strides in mapping the genome with respect to chromatin changes that occur during tumorigenesis. These studies made it possible to verify an overall loss of lysine 16 in histone H4 acetylated (H4K16ac) and of trimethylation of lysine 20 of histone H4 (H4K20me3) [264]. The loss of histone acetylation is a consequence of genetic repression. The operation of HATs in conjunction with HDACs makes it possible to maintain histone acetylation levels. This functioning can in certain circumstances be altered by cancer. In leukemia, it is possible to verify the uncontrolled formation of fusion proteins by means of HAT chromosomal translocations and genes related to HAT, such as, for example, MOZ, MORF, CBP and p300 [265]. In addition to changes in histone acetylation, generalized changes in histone methylation patterns can also be seen in cancer cells. The mechanism for silencing aberrant genes in various forms of cancer is due to changes in the methylation patterns of histones H3K9 and H3K27 [266]. This phenomenon is due to the deregulation of HMTs responsible for repressive brands. For example, the enhancer of homologous zeste 2 (EZH2) is a histone-lysine N-methyltransferase enzyme (H3K27 HMT) encoded by the EZH2 gene, it is overexpressed in breast and prostate cancer [267]. On the other hand, lysine-specific demethylases also play a crucial role in cancer progression. These work in coordination with HMTs to maintain global histone methylation patterns [268]. Lysine-specific demethylase 1A (LSD1) can eliminate the activation and repression marks of H3K4 and H3K9 methylation, respectively. It is dependent on specific binding elements [269] and can act as a co-repressor or a co-activator. After the discovery of LSD1, several other HDMs were discovered, including Jumonji C domain proteins [270].

4.6 Personalized Epigenetics

In-depth investigations of epigenetic diseases allow us to provide not only possible solutions for the etiology of the disease, but also to discover biomarkers for is diagnosis [271]. On the other hand, there has been an increasing interest in the production of drugs with modes of epigenetic

action used mainly for the treatment of these diseases [272]. This is due in particular to the fact that these epigenetic marks have reversible mechanisms. Nowadays, cancer remains the biggest challenge for doctors and researchers in the health field, with regard to the epigenetics of diseases. In 1980, the first links between DNA methylation and cancer were established. However, over the years this number of connections has been increasing. Recently, it has been found that most current applications of epigenetic biomarkers predominantly involve DNA methylation [273]. In the United States, clinics use nucleic acid-based tests, which are regulated and approved by the Federal Drug Administration (FDA) as medical devices. Just as epigenetic biomarkers have been also approved by the FDA. ColoGuard® is one of these tests, usually used to screen for colorectal cancer in adults over 50 years of age. It uses DNA methylation levels in BMP3 and NDRG4, when bound to mutated KRAS and in an immunochemical assay for hemoglobin (Table 4.1). During the implementation of this test in the screening, it can be seen that it has a higher sensitivity and a slightly lower specificity for colorectal cancer, when compared to the traditional screening method, the fecal immunochemical test (FIT) [274]. Thus, more recent results have shown that FIT is more effective and less expensive than ColoGuard®, since the latter is very expensive for patients [275]. This demonstrates the existence of economic barriers that these tests must overcome in addition to demonstrating effectiveness and reproducibility. In addition to ColoGuard®, two more epigenetic tests are currently available in the United States. These are classified as Laboratory Developed Tests (LDTs) and follow the protocol of the Clinical Laboratory Improvement Amendments (CLIA). This means that these tests are only carried out inside the laboratory where they were developed, and they must follow CLIA standards. These tests are ConfirmMDx, for prostate cancer, and AssureMDx, for bladder cancer. In 1990, researchers first demonstrated the hypermethylation of the glutathione S-transferase (GSTP1) gene promoter in prostate cancer [276]. GSTP1 together with APC and RASSF1 are part of the ConfirmMDx test composition (Table 4.1). This test is used to treat false negative prostate biopsy problems [277]. The AssureMDx test for bladder cancer involves a set of analyzes of the DNA methylation levels of three genes (TWIST1, ONECUT2 and OTX1) with the mutation of three others [278].

Table 4.1: Classification of US FDA-approved epigenetic drug classes according to mechanism of action [16]

Product	Proprietor/Launch year	Disease	Specimen	Epigenetic Targets	Regulation
Cologuard	Exact sciences/2014	Colorectal cancer	Stool	DNA methylation of NDRG4 and BMP3 (plus other genetic markers)	FDA
ConfirmMDx	MDxHealth/2012	Prostate cancer	Tissue	DNA methylation of GSTP1, RASSF1 and APC	LDT/CLIA
AssureMDx	MDxHealth/2016	Bladder cancer	Urine	DNA methylation of TWIST, ONECUT2 and OTX1 (plus other genetic markers)	LDT/CLIA

has not yet been approved. In both breast and ovarian cancer, hypermethylation of the BRCA1 promoter region is observed [279]. Normally, BRCA1 is used to epigenetically repress the expression of the oncogenic microRNA miR-155 through a process involving histone deacetylase 2 (HDAC2) [280]. Recently, Anjum et al. (2014) carried out a study, where they identified a DNA methylation signature of blood cells in BRCA1, thus allowing to predict the risk of breast

cancer several years before diagnosis [281]. That said, there is a great advance in the epigenetic studies of non-cancer diseases. These studies have shown promising results, particularly with regard to DNA methylation, including neurological disorders such as Alzheimer's disease [282] and Parkinson's disease [283], autoimmune disorders such as systemic lupus erythematosus [284] and psychiatric disorders such as schizophrenia and autism [285]. The heterogeneity of epigenetic diseases has become the main challenge, with regard to monitoring reliable changes in epigenetic patterns among individuals and the response to medical interventions. Recently, computational epigenetics has become a crucial tool for identifying solutions to this challenge. Recently, there has been a growing increase in computational analysis for DNA methylation patterns, modifications of histones and ncRNAs, as well as quantitative analysis of proteins, in order to identify specific differences in epigenetic profiles between individuals, which are then applied to personalized medicine. These analyzes allowed researchers to discover potential individual epigenetic biomarkers based on disease progression. In the basic nucleus of personalized epigenetics, attention should be paid to inter-individual variability in a set of epigenetic marks, such as DNA methylation, histone modifications and ncRNA. Variations in epigenetic signatures between cells and asymmetric or allele-specific DNA methylation changes in an individual are some of the limitations in personalized epigenetics. Despite these limitations, the characterization of differences between individuals in epigenetics has made significant progress and continues to reveal the importance of inter-individual epigenetic variability in medicine. On the other hand, computational epigenetics has enabled researchers to analyze and understand the difficulties inherent in the use of epigenomic information and how important these are in the diagnosis, prognosis and individual therapy. The increase in the appearance of new epigenetic biomarkers and the respective knowledge of the personalized epigenetic responses caused by these to drugs and environmental toxic factors will continue to be the main challenges for the application of epigenetics in personalized medicine.

Chapter 5

Integrative Meta-Analysis of an Epigenetic Study in GB

5.1 Methodology

5.1.1 Download and separation of data from the TCGA database and Preprocessing

To carry out the practical part of this dissertation, GB data on DNA methylation and gene expression from the TCGA database were used (https://cancergenome.nih.gov/) [21]. The selection and collection of data from the TCGA database followed certain criteria. In the beginning, we started by selecting all the DNA methylation and gene expression data corresponding to the brain as the primary site and the TCGA-GBM project. Then, regarding DNA methylation data, data from the "DNA Methylation" category and from the "Illumina Human Methylation 450" platform were selected, obtaining 155 samples at the end. Regarding the gene expression data, only the data belonging to the category "Transcriptome Profiling", data type "Gene Expression Quantification" and the workflow type "HTSeq-Counts", resulting in the end of approximately 174 samples. Later, it was also decided to use DNA methylation data from GEO (https://www.ncbi.nlm.nih.gov/geo/) [286], where GEO accession is GSE41826. In each CpG, the beta value (β) represents the DNA methylation level, which can be calculated using the expression (M/(M + U)), where M corresponds to the intensity of the methylated allele and U to the intensity of the allele not methylated. Generally, beta values are observed in the range of 0 to 1, reflecting the fraction of methylated alleles in each CpG in each tumor, so beta values close to 0 indicate low levels of methylation and beta values close to 1 indicate high levels of methylation levels [287]. As a first step in the pre-processing of the data, we started by selecting from the 155 samples only primary samples of tumor and normal samples, and this step is also common in the pre-processing of gene expression data. As a next step, it was decided to remove CpGs on sex chromosomes, X and Y, and keep only cg probes, in order to avoid artificial false positives that are independent of the normalization stage. Then, CpG sites with a missing value> 70% of normal samples and tumor samples were

removed. In order to make the imputation of the missing values of the methylation data, the R Impute package [288] was used by the method of the nearest k-neighbors (KNN). As a last step, CpGs were selected in promoter regions using the R ELMER package [289], defined as 2 kb upstream to 0.5 kb downstream of the transcription initiation sites (TSS) [290]. Regarding the pre-processing of gene expression data, the R DESEq2 [291] package was used to perform a pre-filtering of the genes. In order to obtain the total clinical samples of DNA methylation data and the respective clinical information, colData was used, where 138 samples were obtained at the end. As a next step, these samples were divided equally into two groups, the training set (n = 69) and the test set (n = 69). For the division of these two sets, the following criteria were followed: 1) the samples were randomly assigned to the training and test sets and (b) the age distribution, follow-up period and proportion of cases of death in the two groups must be similar , more specifically "sampleID", "vital-status", "age-at-initial-pathologic-diagnosis", "days-to-death" and "days-to-last-followup".

5.1.2 Determination of differentially expressed genes (DEGs) from GBM and methylated CpGs sites

Typically, one way to analyze RNA-Seq count data is to determine which genes are differentially expressed. The counting data, in the form of a table, allows to inform, for each sample, the distribution of the number of sequence fragments corresponding to each gene. These counts occur due to HTSeq [292] and are calculated at the gene level. During the analysis of these data, in addition to the reduced number of samples being analyzed, attention should be paid to the specificities of the counting data, especially NA data (not available). However there is still a greater challenge, which is the small number of samples in high-throughput screening experiments (HTS), where the repetition of samples due to disease is often observed. Due to the high uncertainty rate of the variance estimates within the group, the inferential methods responsible for each gene suffer from an absence of power. However, this limitation can be solved through high-throughput assays, grouping information between genes, more specifically, by discovering possible hypotheses about the similarity of the variances of different genes measured in the same experiment [293]. The analysis of differential expression of RNA-seq data can be done through several methods, the most common being the empirical analysis of digital gene expression data in R (edgeR) [294] and the analysis of differential expression of genes based on negative binomial distribution (DESeq2) [291]. Through the edgeR method, it is possible to observe the sharing of information between genes to estimate variation or dispersion [295, 296], using a weighted conditional probability. Nevertheless, in order to carry out this step, we chose to use the DESeq2 method, which is a successor of the DESeq method [297] and allows to integrate methodological advances through several new resources, allowing to obtain a more quantitative analysis of comparative data easily of RNA-seq using shrinkage estimators for dispersion and fold change [298].

In the analysis of transcriptomic data, it is observed that the most common approach in this type of analysis is to test the null hypothesis that the logarithmic change (LFC) between two different groups, in this case between samples of primary tumor and normal samples, for the expression

of a gene it is exactly zero, that is, the gene is not affected in any way by one of these groups. Most of the time, the objective of the differential analysis is to generate a list of genes that pass the multiple test adjustment, classified by the P value. However, the existence of small, highly significant changes makes candidates no longer interesting for further investigation. On the other hand, sorting by fold change becomes complicated due to the noise of the LFC estimates for genes with low counts. In addition, the number of genes designated by "differentially expressed" depends on both the sample size and other aspects of the experimental design and the biology of the experiment [299]. Nevertheless, to reduce this noise, we decided to use the lfcShrink function of the R apeglm package [300]. This shrinkage method uses a Bayesian procedure to moderate gene folding changes with very low and highly variable counts. This serves to attenuate noise in differential transcription counts. To visualize the differentially expressed genes, the R EnhancedVolcano [301] and complexHeatmap [302] packages were used. In high-performance studies, DNA methylation can be used to identify functional changes in transcription enhancers and other cis-regulatory modules (CRMs) in tumors and other tissues of primary diseases [303]. Initially for the identification of differentially methylated CpGs sites between primary tumor samples and normal solid tissue, we decided to use the R ELMER package (Enhancer Linking by Methylation/Expression Relationships) [289], which produces a systematic approach for the reconstruction of regulatory networks of genes combining methylation and gene expression data derived from the same set of samples. However, as the use of ELMER to identify methylated sites was not possible, later it was decided to identify methylated sites using the Wilcoxon classification sum test (μ) and the t test. The Wilcoxon classification sum test or also called the Mann-Whitney (μ) test, is a non-parametric test based on classification and used in the R methyAnalysis package as a method of differential methylation analysis [304], using DNA methylation data TCGA and GEO. This test is normally used as an alternative to the t test of two independent samples when an assumption of normal data distribution is observed, and this is violated the t test. Suppose that the methylation level is represented by β values or M values for the locus (i), the group (j) and the subject (k) by y ijk. Now suppose that j = 1 represents the group of normal solid tissue samples and j = 2 represents the group of primary tumor samples. For each DNA methylation locus, it is observed that the null hypothesis of the Wilcoxon rank sum test of the distribution of y i 1 k is equal to the distribution of y i 2 k for i = 1,2, ..., m. The gross P values of the Wilcoxon rank sum test are usually adjusted using the procedure by Benjamini and Hochberg to supervise the FDR at level α [305] using the p.adjust function in R. On the other hand, uses the t test method is in the analysis of genomic data to test the equivalence of means between two groups [306]. For two independent sample t tests, there are two t-test procedures, depending on whether there is equality between the variances of these two groups or not. In the procedure of the t test of unequal variation, that is, the Welch t test, the same variation between groups is not normally used. The calculation of the gross p-values of the t-tests is based on the t-distribution [307], although the adjusted p-values are obtained using the procedure of Benjamini and Hochberg through the same function p.adjust in R. Finally, the genes with rate of false discovery (FDR) <0.01 were selected as differentially expressed genes or DEGs and significant methylated sites.

5.1.3 Survival analysis of training set

From the middle of the twentieth century, the implementation of the analysis of survival of a given disease, as a statistical procedure in the field of medicine, was more frequently performed. Survival analysis aims to analyze the life span of individuals from the moment they enter the study, until the moment when the event of interest occurs, an event that is defined at the outset [308]. Usually, this event is defined as a failure, which may be death, a relapse of a disease or even when a certain treatment begins to take effect on the patient. These analyzes have the particularity of dealing with censored data, that is, in some of the individuals, it may not be possible to observe the event of interest during the observation period. Many times, it is verified the existence of individuals who are still alive after the end of the study or who abandoned the treatment. In these cases, it is said that individuals have a censored life span. That said, the life span of a given individual, of a homogeneous population, is usually represented by a random variable T, non-negative and absolutely continuous [308]. In this way, the survival function of a certain individual is defined as the probability of surviving beyond an instant t and is represented by:

$$(S) = P(T > t), t \geqslant 0 \tag{5.1}$$

which is characterized as a monotonous, non-increasing and continuous function and which has the following properties:

1. S(0)=1;
2. $S(+\infty) = \lim_{h\to\infty}(t) = 0$;

In 1959, Kaplan and Meier, proposed a generalization of the empirical survival function, known as the Kaplan-Meier (KM) estimator or "product-limit" estimator [309]. This function reflects the probability of cumulative survival over time. The observation of a horizontal curve indicates that over a certain period there was no event. Then, if there is a decrease vertically it means that there has been a change in the survival function. On the other hand, censorship corresponds to a type of lost data problem exclusively for survival analysis. Therefore, the main objective of survival analysis is to compare survival functions in different groups. In monitoring a given set of patients until death, corresponding to two groups, it is expected that the survival curve ends at 0%, but there are cases where one group may have survived on average much longer than the other group. Survival analysis does this by comparing the hazard at different times during the observation period. Usually, Kaplan-Meier curves are used to visualize differences in survival between two groups [310], but there are situations in which the use of this method does not work well to assess the effect of quantitative variables such as age, gene expression, DNA methylation, etc. Cox PH or Cox proportional hazard regression can evaluate the effect of categorical and continuous variables and can model the effect of several variables. On the other hand, Cox PH regression models the natural log of the hazard rate at time t, called (t), as a function of the baseline risk ($h_0(t)$) and multiple exposure variables x_1, x_2,..., xp [310]. The form of the Cox PH model is represented by:

$$log(h(t)) = log(h_0) + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \tag{5.2}$$

In order to expose both sides of the equation and limit the right side to just a single categorical exposure variable (x_1) with two groups ($x_1 = 1$ for exposed and $x_1 = 0$ for unexposed), the equation is now represented this way:

$$h_1(t) = h_0(t) \times e^{(\beta_1 x_1)} \tag{5.3}$$

The reorganization of this equation makes it possible to estimate the risk ratio, that is, it allows comparing individuals exposed to those not exposed at time t:

$$HR(t) = \frac{h_1(t))}{h_0(t)} = e^{(\beta_1)} \tag{5.4}$$

where $e^{(\beta_1)}$ corresponds to the risk ratio and remains constant over time t. The values are the regression coefficients and can be estimated from the model and represent the log(HazardRatio) for each unit increase in the corresponding prediction variable [310].

That said, for the realization of this stage in this dissertation, a survival analysis of clinical training data was started, using the R survival and survminer packages [311, 312]. Before building the training set survival curves, it was converted the "death" and "alive" variables in the "vital-status" column for 0 and 1, in order to facilitate the next steps. Then, the survfit() function was used to calculate the kaplan-Meier survival estimate. However, due to the time imposed it was not possible to carry out the stages of survival analysis of the set of methylation profiles and clinical training data, including analysis of the univariate Cox proportional hazard regression model based on each methylation site and data of survival.

5.2 Results and Discussion

5.2.1 Separation of DNA methylation data and respective characteristics

In this dissertation, it was decided to produce an integrative analysis of DNA methylation and gene expression in order to identify the main epigenetic genes in GB, following the procedure presented in [17]. It was decided, due to time constrains, to carry selected tasks and to do so, we used the workflow shown in Figure 5.1.

Nevertheless, carrying out this replication study become difficult since there was no information about the base functions used in the construction of the algorithm. As such, the results showed some discrepancies, as shown in Table 5.1.

A total of 142 samples of primary GBMs and normal samples were initially obtained with data from clinical information. However, when analyzing the data, it was found that four samples did not contain any type of clinical information and it was then decided to remove these thus obtaining a total of 138 samples.In addition, when downloading the data, 56457 genes and 485577 CpGs were obtained from the TCGA database for further analysis. Then, a set of pre-processing steps

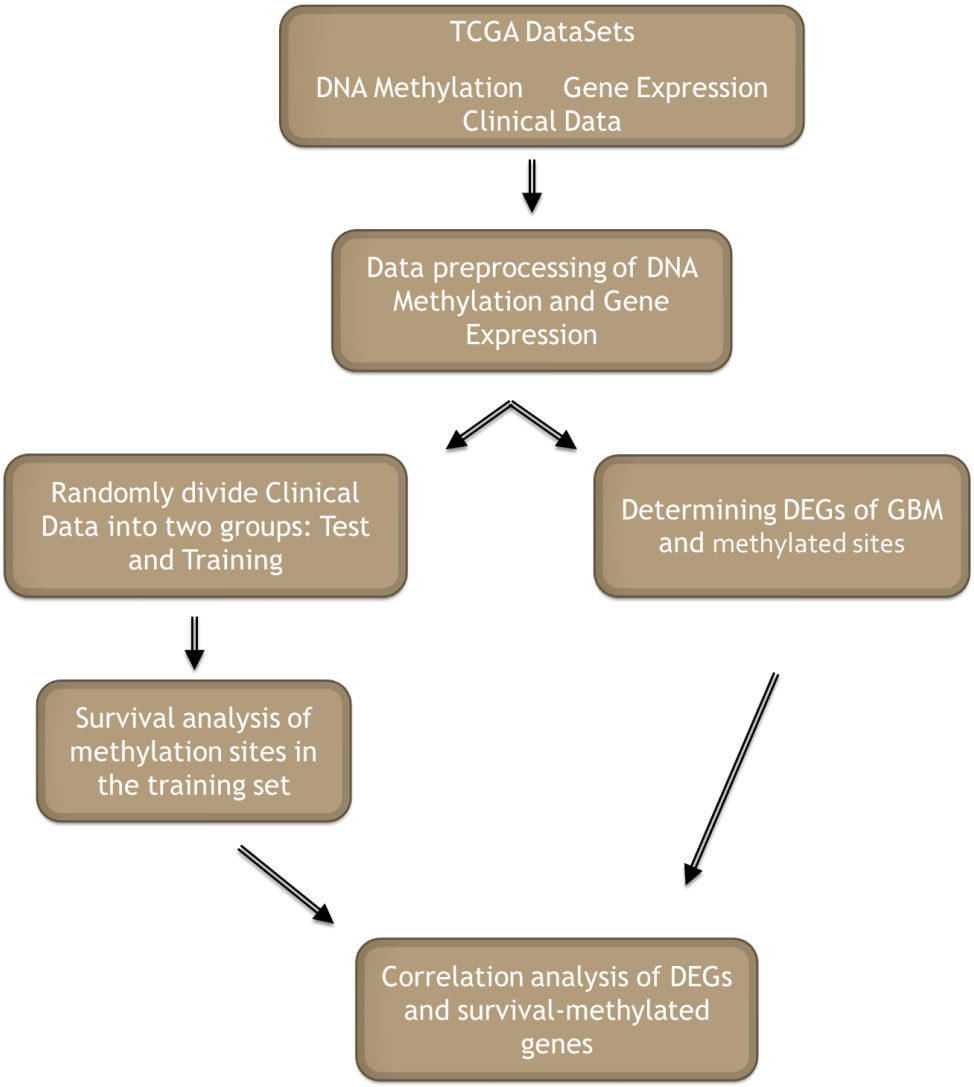

Figure 5.1: Flowchart describing the schematic overview of the project under study.

were carried out, including adaptation of missing values, removal of batch effects, removal of sex chromosomes and single nucleotide polymorphisms and extraction of CpGs in promoter regions. As DNA methylation in promoter regions strongly influences gene expression, CpGs were selected in promoter regions, which were defined as 2 kb upstream to 0.5 kb downstream of TSS. After pre-processing the data, 161318 methylation sites were finally obtained for downstream analysis (**Supplementary table 1**) and 46752 genes (**Supplementary table 2**).

5.2.2 Clinical patient characteristics

From the total of 138 samples, clinical information was obtained, including sample identification, vital status, age at initial pathological diagnosis, days until death, days until the last follow-up and the respective degree. All samples were randomly divided into two groups: the training set, where

Table 5.1: Results obtained vs D. Jia et al. [17].

	My results	**D. Jia et al. [17]**
Download: DNA Methylation/ CpGs	155 samples/ 485577 CpGs	155 samples/ 485577 CpGs
Download: Gene expression	174 samples/ 56457 genes	n.a/ n.a
Preprocessing: DNA Methylation	161318 CpGs	145907 CpGs
Preprocessing: Gene expression	46752 genes	20530 genes
Match between the clinical data and the methylation profiles	138 samples	138 samples
Train and Test set	69 samples in both sets	69 samples in both sets
Determining DEGs of GBM	12427 genes	4881 genes
Genes upregulated/ genes downregulated	5783 genes upregulated/ 6644 genes downregulated	1111 genes upregulated/ 3770 genes downregulated
Determining of the methylated sites of GBM using Wilcoxon's classification sum test and t-test	342607 CpGs in both tests	n.a

69 samples were obtained, and the test set, where the same number of samples was obtained. The division of these two groups respected the following criteria: first, the samples are randomly assigned to the training set and the test set and secondly, the age distribution, follow-up time and the patient's mortality rate must be similar in these two groups. The clinical information of the training set and the test set can be seen in **Supplementary tables 3** and **4**, respectively.

5.3 Determination of differentially expressed genes (DEGs) from GBM and methylated CpGs sites

According to the screening criteria, a total of 12427 DEGREES were obtained, all of which were originals of primary tumors and normal samples. It was found that there are about 5783 positively regulated genes and 6644 negatively regulated genes. In order to visualize these data, it was decided to start using the MA plot, Figure 5.2. A MA plot is a two-dimensional (2D) scatterplot of log-fold change (M values, that is, the log of the ratio of level counts for each gene between two a) against a logarithmic average (A values, that is, as mid-level counts for each gene in the two) [1]. In this graph, it is observed that genes with similar levels of expression between primary tumor samples and normal samples appear around the horizontal line M = 0. With respect to points distant from the line M = 0, these indicate genes with significant expression, where a gene that is up-regulated is above the M=0 line and a gene that is down-regulated is below the M=0 line.

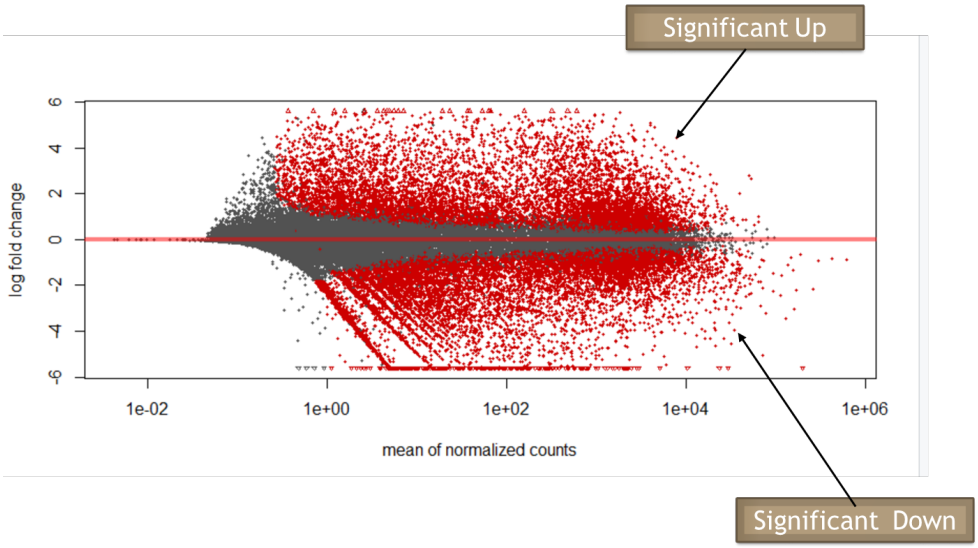

Figure 5.2: MA plot of differentially expressed genes.

However, the MA graph does not consider the statistical measures (P values or adjusted P values) and, therefore, it is not possible to indicate which are the genes with statistically significant differences between tumor and normal samples in the graph. For this, it was decided to use the Vulcano graph, using the R EnhancedVolcano package, in order to visualize which genes have the greatest significant differences, Figure 5.3. Through the graph it was verified that the negatively regulated genes ENSG00000213553 (RPLP0P6), ENSG00000171848 (RRM2) and ENSG00000235655 (H3F3AP4) present greater significant differences. Regarding the positively regulated genes, it was observed that the ENSG00000050748 (MAPK9), ENSG00000155744 (FAM126B) and ENSG00000073969 (NSF) genes have greater significant differences.

In addition to using these two graphs, it was also decided to view the expression profiles of the 100 most significant genes, Figure 5.4. Before building the Heatmap, variance stabilization transformations (VST) were used to obtain data transformed on the log2 scale and remove the dependence of the variance with the mean. After this step, 100 most differentially expressed genes were selected and the R ComplexHeatmap package was used to build the heatmap. In this heat map, the Z scores are calculated for each gene, subtracting the mean and dividing by the standard deviation, and are then plotted instead of the normalized expression values, ensuring that the expression patterns / trends you want to visualize are not overloaded by expression values.

Regarding the determination of differentially methylated sites (DMS), initially we tried to use the get.diff.meth function of the R ELMER package to identify distal probes with significant differential methylation of DNA, but we obtained zero relevant probes as output. This was due to having only 1 sample of normal solid tissue and 50 samples of primary tumor. To this end, it was decided to add the GEO data to the DNA Methylation data of the TCGA. Subsequently, it was decided to create two functions, one parametric and the other non-parametric, for the identification of these methylated sites. The non-parametric function corresponds to the Wilcoxon test and the

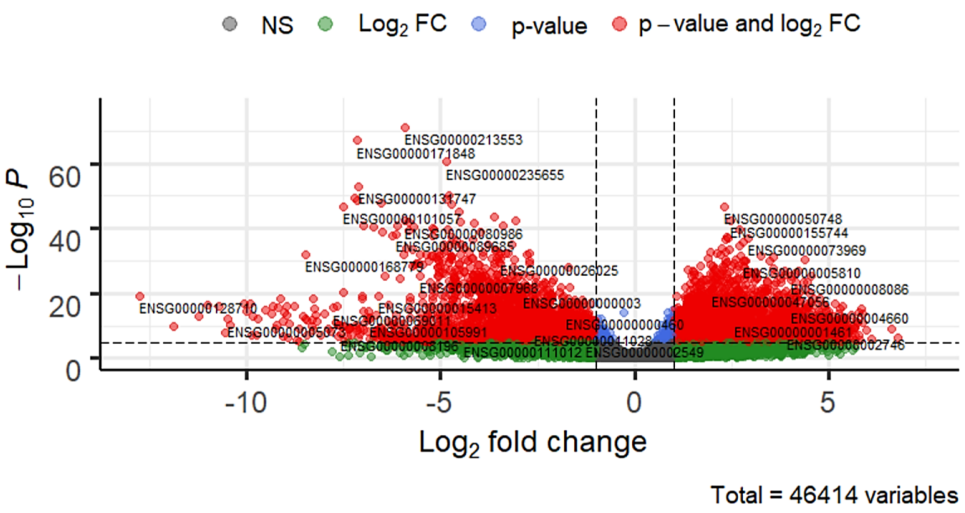

Figure 5.3: Vulcan graph of differentially expressed genes.

parametric function corresponds to the t test. In the end, 342607 differentially methylated sites were obtained in both tests in **Supplementary tables 5** and **6**. Nevertheless, there is a difference between the results of the t test and the Wilcoxon test, this is due to the t test testing the significant differences between the means and the Wilcoxon test testing the significant differences between the medians. However, the Wilcoxon test takes about 4 hours to perform while the t test only takes 30 minutes. Furthermore, according to this article [304], the t test has a greater resolving power than the Wilcoxon test. Therefore, for the next analysis it was decided to use the t test, taking into account what was said earlier.

5.4 Survival analysis of training set

In order to analyze the data obtained in this stage, a statistical analysis of the data was carried out first and only then proceeded for the survival analysis. According to the literature, it appears that the appearance of GB occurs more frequently in people aged between 40 and 70 years, with a slight increase between 65 and 70 years, considering this age group as a risk group. To prove this veracity, it was decided to draw a histogram of the age of the patients for the training set after diagnosis, where it can be concluded that in fact there is a larger increase in cases of GB between 40 years and 70 years, Figure 5.5.

After this analysis we proceeded to the survival analysis of the training set data, including clinical information ("sample", "vital-status", "age-at-diagnosis", "days-to-death", "days-to-last-follow-up"), where later it was added also clinical information “gender”. Usually, most patients die within 2 years and the survival time is generally less than one year from the date of diagnosis [313, 314], with patients with a maximum survival time of 5 years, but this percentage is very reduced. To this end, it was decided to calculate the probability of patients in the training set to

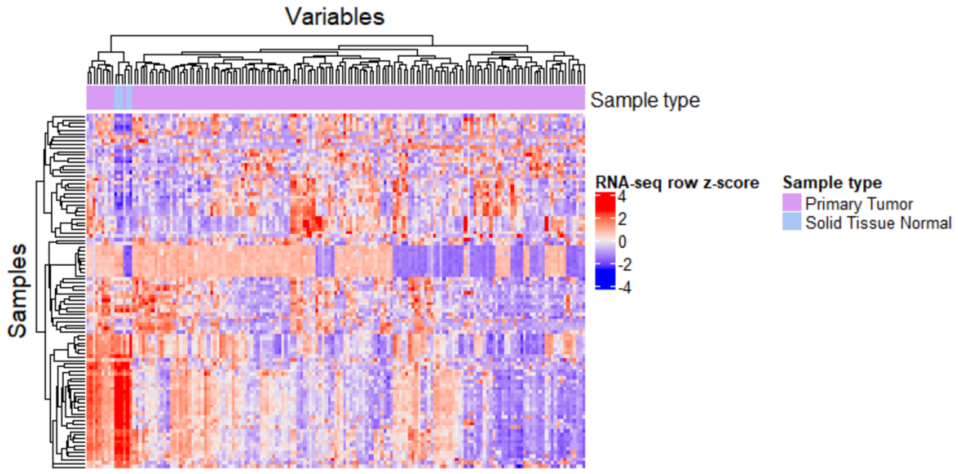

Figure 5.4: The heatmap expression profiles of the 100 most significant genes.

survive after 1 year and after 5 years, with and without the gender variable. After 1 year and 5 years, it is observed that in general the probability of patients with GB to survive is approximately 74.6% and 45.3%, respectively. Normally, GB affects more men than women, so it is expected that women are more likely to survive than men. In fact, it was observed that in the graph with the gender variable, for example, after 5 years the probability of the man surviving is very low (24.8%), when compared with the probability of surviving the women (71,1%).

Then the Kaplan-Meier estimate was calculated for the survival function without and with the variable "gender", using the **Supplementary tables 3** and **7**. To visualize these two sets, Figure 5.6, the plot and ggsurvplot function were used. Through the analysis of the first graph, he concluded that there is in fact a sharp decrease in the survival rate of these patients, since this tumor is very aggressive and heterogeneous. Another reason why the survival rate is also very low is that there is still no effective treatment. Regarding the second graph, it can be concluded that the survival rate can also be affected by the gender difference, with a higher survival rate for women than for men. This is in line with what is said in the literature, that is, although there are few insights that distinguish male and female GB at the molecular level or allow the specific targeting of these biological differences, there are already studies that prove that women generally have a higher survival rate than men, as well as there are already studies that prove that they also present a better response to treatment, thus increasing their survival time [315].

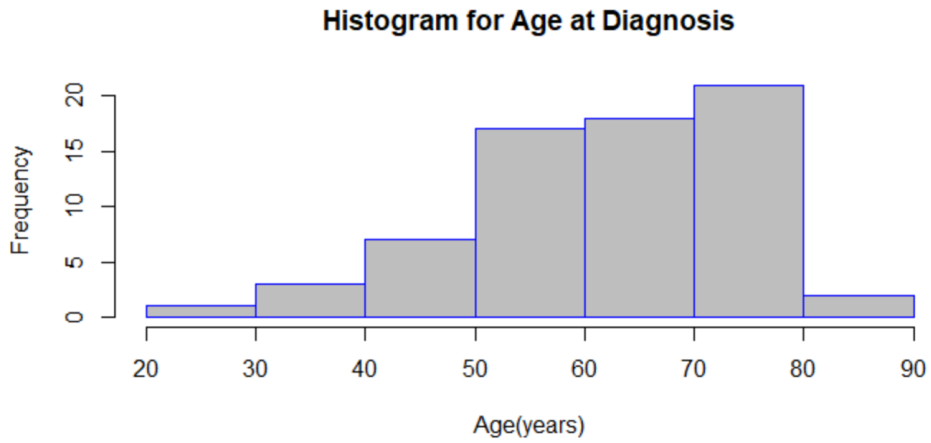

Figure 5.5: Histogram of the number of GB cases by age group.

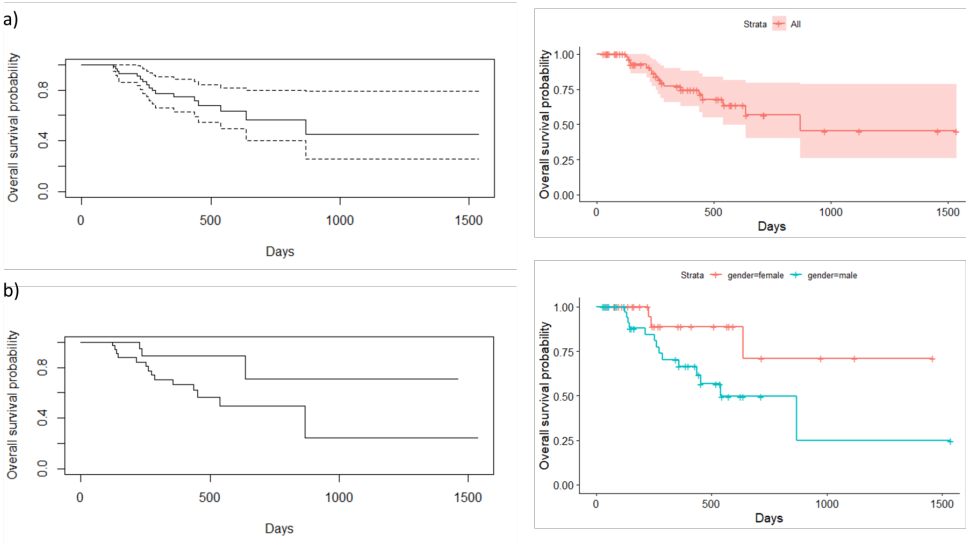

Figure 5.6: The survival analysis of the training set data using the Kaplan-Meier plot of a) the whole set b) of men and women.

Chapter 6

Conclusion

Throughout the document, the advantages and challenges of epigenetic studies in Glioblastoma were discussed, as well as the realization of an integrative meta-analysis of an epigenetic study. Epigenetic therapy has been shown to be a successful approach to the treatment of different malignancies. Future studies may be able to identify methods for recognizing response mechanisms by combining genomic sequencing and gene expression profiles. An important challenge in epigenetic therapy is to know which are the conducting genes and which are the stimulated genes. The increasing development in the sequencing of the entire genome together with RNA data profiles allowed to acquire a large amount of information crucial for the precise identification of epigenetic changes. Comparing and reconciling this enormous amount of information will help to discover epigenetic changes that occur as a cause and effect or are totally dependent on each other. That said and with a focus on epigenetics, the algorithm to be developed aims to identify the main epigenetic genes. In other words, the intention is to use DNA methylation and TCGA gene expression data to build a prognostic risk model based on methylation genes, allowing the identification of potential biomarkers for a better prognosis and a better survival rate of GB patients.

First, a global analysis of all the work carried out will be presented.. Finally, the future work to be carried out in order to finish the proposed objectives is discussed, as well as to develop a system of GB cell spheroid cultures, in order to analyze the evolution and study the behavior of these epigenetic genes, thus allowing to implement a personal and unique therapeutic approach to the treatment of each patient.

6.1 Conclusion of work

The work proposed for the dissertation addresses the realization of a meta-analysis of an epigenetic study through the construction of an algorithm, using R language, capable of identifying the main epigenetic genes in GB, in order to build a risk model for prognosis based on methylation genes. As far as we know, there is still no treatment for patients that is 100% effective. Nevertheless, there is already a diversity of epigenetic studies in the area of oncology, including glioblastoma,

which are undergoing clinical trials, to be subsequently implemented. However, this process is very time-consuming and rigorous.

The dissertation started with an in-depth explanation, including a state of the art, about what glioblastoma is, as well as the concept of epigenetics and what is its role in GB. Glioma is a common malignant primary brain tumor that has high recurrence rates, short survival times, high mortality rates and treatment difficulties. Previously, treatments for patients with GB were based on conventional surgeries and chemo-radiotherapy protocols, which minimally improved quality of life and slightly prolonged the survival of some patients. Nevertheless, there is an urgent need for further investigations and reviews of treatment strategies for malignant glioma. The genetic instability and heterogeneity of the glioma are the biggest challenges in investigating new solutions for the treatment of these patients. In general, epigenetic changes are completely associated with proliferation, metastasis, invasion and prognosis of the glioma. Several epigenetic changes appear in order to participate in the occurrence and progression of the glioma. In the basic nucleus of personalized epigenetics, attention should be paid to inter-individual variability in a set of epigenetic marks, such as DNA methylation, histone modifications and ncRNA. Variations in epigenetic signatures between cells and asymmetric or allele-specific DNA methylation changes in an individual are some of the limitations in personalized epigenetics. Despite these limitations, the characterization of differences between individuals in epigenetics has made significant progress and continues to reveal the importance of inter-individual epigenetic variability in medicine. On the other hand, computational epigenetics allowed researchers to analyze and understand the difficulties inherent in the use of epigenomic information and how important these are in the diagnosis, prognosis and individual therapy.

After the bibliographic review, it was decided to identify the main epigenetic genes in GB, which was a topic that captivated me due to the fact that patients have a poor prognosis and the identification of these genes is crucial for obtaining a better prognosis and diagnosis, contributing to the increase in the survival rate. For the integrative meta-analysis of the epigenetic study, the Rstudio software was used. However, it was not possible to complete all the proposed objectives and, therefore, it was also not possible to identify which epigenetic genes are most expressed in samples of normal solid tissue and in samples of primary tumor tissue. Only tasks such as downloading and processing of data, identification of differentially expressed genes, a crucial step for future steps, and survival analyzes were performed. Through the analysis of the results obtained, it was only possible to confirm some predictions. By comparing the results obtained with the results of D. Jia et al. [17], it was observed that slightly different results were obtained, due to the lack of important information on the basic functions used in that reference. In section 6.2, we will discuss the continuation of the remaining steps as future work, as well as the production of spheroids to evaluate the evolution of these genes in a tumor environment.

That said, my biggest difficulties in carrying out this meta-analysis were the adaptation of new objectives that involved totally new concepts, such as epigenetics and its involvement in GB. A big challenge for me was learning a new programming language, Rstudio, which I dedicated most of my time to. In addition, another problem arises when the available information is very little in

relation to the part of the analysis of a data set of gene expression and DNA methylation using R, raising a lot of doubts and thus imposing a larger amount of time in the theoretical analysis of these concepts, in order to identify solutions to certain problems.

However, it is concluded that the use of computational algorithms for epigenetic analysis in certain diseases will allow the study of epigenetics as one of the main factors that contributes to the formation of normal and tumor cells, and will open new views for the identification of potential biomarkers and for the advent of new personalized therapeutic approaches. Furthermore, the combination of epigenetic therapies with traditional therapies will allow to obtain certain treatments for the reversal of drug-resistant tumors.

6.2 Future work

Due to the imposed time it was not possible to complete all the steps that were in the workflow. As future work, the continuation of this work involves the determination of methylation profiles and survival data in the training set, and each methylation site must be analyzed according to the steps: (a) analysis of the proportional risk regression model Cox analysis based on each methylation site and survival data, and (b) adding age as a covariant for CpG survival analysis. The construction of a regression model is a very important step, since it is from it that we learn which factors (covariates) influence the life span of an individual. Depending on the distribution we are going to use for the life span of individuals, we will have a parametric or non-parametric regression model. Subsequently, a correlation must be made between differentially methylated genes and methylated survival genes. The level of DNA methylation can affect gene expression, just as high methylation expression often inhibits expression of the downstream gene, and the low level of methylation tends to increase expression of the downstream gene. In a first phase, carrying out the steps of correlation analysis and calculation of differentially expressed genes and differentially methylated genes, involves calculating the intersection of differentially methylated genes and DEGs. Then proceed to the identification of the number of genes whose differential expression is regulated positively and the methylated expression differentially is regulated negatively, and the identification of the number of genes whose differential expression is regulated negatively and the differential methylation is regulated positively. After this step, an analysis of Pearson's correlations between positively regulated DEGs and negatively regulated methylated genes should be performed, as well as negatively regulated DEGs and positively regulated methylated survival genes. Through this last stage, it is possible to determine the central objective of the dissertation, that is, to identify the main epigenetic genes. On the other hand, performing this step also allows us to observe that highly expressed genes in primary tumor samples show less promoter methylation in normal solid tissue samples, indicating a negative correlation between promoter DNA methylation and gene expression in normal and tumor tissues. The completion of all the proposed steps then allows the construction of a prognostic risk model based on methylation genes, allowing the identification of potential biomarkers for a better prognosis and a better survival rate for patients with GB.

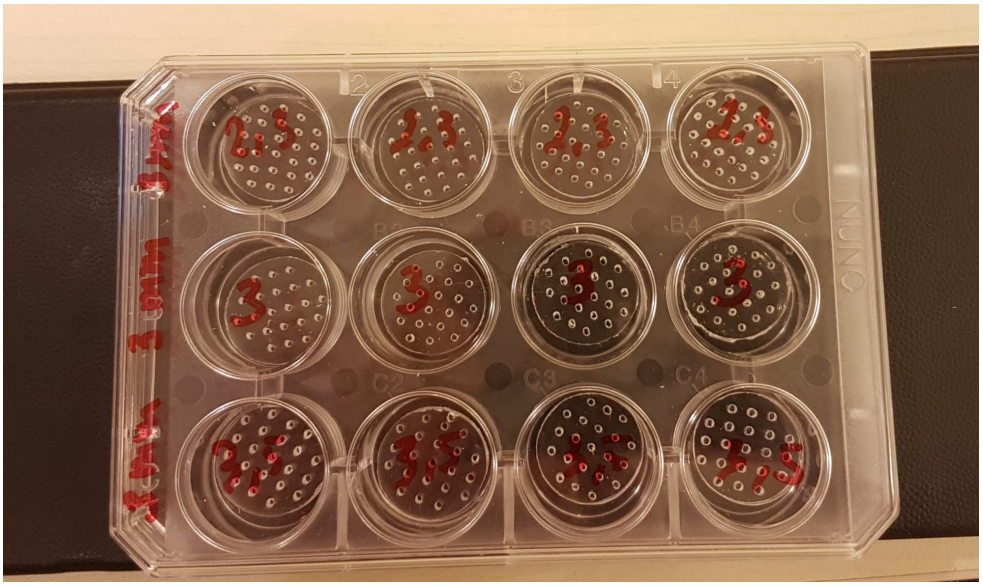

Figure 6.1: PMMA well plates.

After completing all the steps previously proposed, another future work would involve continuing the work initially proposed. The initial objectives of this dissertation were based on the development of 2D and 3D cell cultures, but due to the measures that were imposed by the pandemic, the realization of these was not possible. Therefore, a future vision would be to implement these cell cultures to assess the evolution of epigenetic genes in glioblastoma U87 cell lines. The entire procedure for the development of cell cultures is carried out in well plates, Figure 6.1. For the construction of the twelve plates, PMMA(Polymethylmethacrylate or acrylic) was used as the manufacturing material, ArtCAM as design and simulation software [316] and, finally, for the manufacture, a Micro Milling CNC machine was used. The diameter of all plaques are the same, varying only the distance between the wells and the number of wells in each plate, in order to try to mimic the tumor environment as much as possible. After the production of the plates, it was sterilized in order to remove all the dirt obtained by the production, Figure 6.2. Sterilization is a very important step, as the removal of dirt prevents contamination of the cells.

Cell lines either grow as adherent or suspended cultures, and all cell culture preparation must involve strict and detailed monitoring of a Protocol. In an initial phase, a protocol was developed (in annex), which is still subject to changes. One of the phases of the development of the protocol is the choice of the culture medium, which is one of the most important ingredients for the growth of the cultures, since a good choice of the culture medium will determine the possibility of obtaining of good results. Therefore, the continuation of this work involves steps such as the appropriate choice of culture medium, the production of cell culture and the analysis of its evolution. Once the cell cultures are ready, determining the spheroid detachment method is an important step, as poor detachment can cause the spheroids to rupture. However, the analysis of spheroid growth allows at the same time to analyze the evolution of certain epigenetic genes. This analysis makes it possible

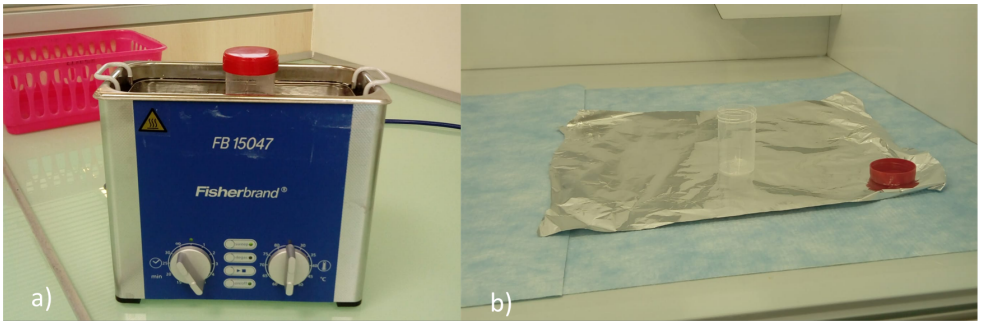

Figure 6.2: UV sterilization of PMMA plates with 70% ethanol. a) Corresponds to the stage of the ultrasonic bath. b) Corresponds to the drying and sterilization stage by UV (ultraviolet radiation).

to evaluate the behavior of genes, allowing the tumor to be classified according to its degree of heterogeneity and to identify the main causes of tumor onset and progression, thus providing the identification of crucial biomarkers to obtain a better prognosis, diagnosis and personalized treatment of patients with GB.

Appendix A

Annex

Supplementary tables: https://www.dropbox.com/sh/fo0bar928peacle/AAAogBQZ1Qo_UxAjTlCSwdjZa?dl=0

Protocol: : https://www.dropbox.com/s/ntozm9ys25ln0oh/Protocol.pdf?dl=0

References

[1] Andrea Cruz and Weng Kung Peng. Perspective: Cellular and molecular profiling technologies in personalized oncology. *Journal of Personalized Medicine*, 9(3), 2019. doi:10.3390/jpm9030044.

[2] Kewal K. Jain. Textbook of Personalized Medicine. *Textbook of Personalized Medicine*, pages 91–97, 2015. doi:10.1007/978-1-4939-2553-7.

[3] Ander Saenz-Antoñanzas, Jaione Auzmendi-Iriarte, Estefania Carrasco-Garcia, Leire Moreno-Cugnon, Irune Ruiz, Jorge Villanua, Larraitz Egaña, David Otaegui, Nicolás Samprón, and Ander Matheu. Liquid biopsy in glioblastoma: Opportunities, applications and challenges. *Cancers*, 11(7):1–20, 2019. doi:10.3390/cancers11070950.

[4] GLOBALLY. MOST COMMON AND MOST AGGRESSIVE FORM OF PRIMARY BRAIN CANCER . page 1, 2019.

[5] Gabriela Sousa, Armando Rocha, Teresa Alfaiate, Teresa Carvalho, Antonio Veiga e Moura, and Mario Rui Ferreira. Glioblastoma multiforme... com apresentaçao multifocal. *Acta Medica Portuguesa*, 15(4):321–324, 2002.

[6] Maria-del-Mar Inda, Rudy Bonavia, and Joan Seoane. Glioblastoma Multiforme: A Look Inside Its Heterogeneous Nature. *Cancers*, 6(1):226–239, jan 2014. doi:10.3390/cancers6010226.

[7] J. P. Posti, M. Bori, T. Kauko, M. Sankinen, J. Nordberg, M. Rahi, J. Frantzén, V. Vuorinen, and J. O.T. Sipilä. Presenting symptoms of glioma in adults. *Acta Neurologica Scandinavica*, 131(2):88–93, feb 2015. doi:10.1111/ane.12285.

[8] Antonio Omuro and Lisa M. DeAngelis. Glioblastoma and other malignant gliomas: A clinical review, 2013. doi:10.1001/jama.2013.280319.

[9] UdayB Maachani, Uma Shankavaram, Kevin Camphausen, and Anita Tandle. Advances in Omics Technologies in GBM. *Biomedical Research Journal*, 2(1):6, 2015. doi:10.4103/2349-3666.240621.

[10] Jônatas Bussador Do Amaral, Estudo De, Câncer De, Mama Humano, Orientadora : Prof, and Gláucia Maria Machado-Santelli. CÉLULAS MCF-7 COMO MODELO 3D NO. Technical report, 2010.

[11] Jordan Villanueva. The Promises and Challenges of 3D Models. URL: https://worldwide.promega.com/resources/pubhub/features/the-promises-and-challenges-of-3d-models/.

[12] Rania Ibrahim, Maria Pasic, and George M. Yousef. Omics for personalized medicine: defining the current we swim in. *Expert Review of Molecular Diagnostics*, 16(7):719–722, 2016. doi:10.1586/14737159.2016.1164601.

[13] Tony Kouzarides. Chromatin Modifications and Their Function, feb 2007. doi:10.1016/j.cell.2007.02.005.

[14] Chang Hyuk Kwon, Dawen Zhao, Jian Chen, Sheila Alcantara, Yanjiao Li, Dennis K. Burns, Ralph P. Mason, Eva Y.H.P. Lee, Hong Wu, and Luis F. Parada. Pten haploinsufficiency accelerates formation of high-grade astrocytomas. *Cancer Research*, 68(9):3286–3294, 2008. doi:10.1158/0008-5472.CAN-07-6867.

[15] Hiroko Ohgaki and Paul Kleihues. 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, 2005. doi:10.1093/jnen/64.6.479.

[16] Bernd Schmidt, Volker Liebenberg, Dimo Dietrich, Thomas Schlegel, Christoph Kneip, Anke Seegebarth, Nadja Flemming, Stefanie Seemann, Jürgen Distler, Jörn Lewin, Reimo Tetzner, Sabine Weickmann, Ulrike Wille, Triantafillos Liloglou, Olaide Raji, Martin Walshaw, Michael Fleischhacker, Christian Witt, and John K. Field. SHOX2 DNA Methylation is a Biomarker for the diagnosis of lung cancer based on bronchial aspirates. *BMC Cancer*, 10(1):600, nov 2010. doi:10.1186/1471-2407-10-600.

[17] Danyun Jia, Wei Lin, Hongli Tang, Yifan Cheng, Kaiwei Xu, Yanshu He, Wujun Geng, and Qinxue Dai. Integrative analysis of DNA methylation and gene expression to identify key epigenetic genes in glioblastoma. *Aging*, 11(15):5579–5592, aug 2019. doi:10.18632/aging.102139.

[18] Aaron D. Goldberg, C. David Allis, and Emily Bernstein. Epigenetics: A Landscape Takes Shape, feb 2007. doi:10.1016/j.cell.2007.02.006.

[19] David Gomez-Cabrero, Imad Abugessaisa, Dieter Maier, Andrew Teschendorff, Matthias Merkenschlager, Andreas Gisel, Esteban Ballestar, Erik Bongcam-Rudloff, Ana Conesa, and Jesper Tegnér. Data integration in the era of omics: current and future challenges, 2014. doi:10.1186/1752-0509-8-S2-I1.

[20] Download RStudio - RStudio. URL: https://rstudio.com/products/rstudio/download/.

[21] The Cancer Genome Atlas Program - National Cancer Institute. URL: https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga.

[22] Lisboa Insa, Astrid M Vicente, Ana João Santos, Sónia Namorado, Liliana Antunes, Ana Paula Gil, José Manuel Boavida, Ana Clara Silva, Patrícia Vargas, Alexandra Costa, Rui Salvador, and Luciana Costa. Observ ações. 2017.

[23] Quinn T. Ostrom, Haley Gittleman, Paul Farah, Annie Ondracek, Yanwen Chen, Yingli Wolinsky, Nancy E. Stroup, Carol Kruchko, and Jill S. Barnholtz-Sloan. CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006-2010. *Neuro-Oncology*, 15(SUPPL.2), 2013. doi:10.1093/neuonc/not151.

[24] Jorge Alberto Bernstein Iriart. Medicina de precisão/medicina personalizada: análise crítica dos movimentos de transformação da biomedicina no início do século XXI. *Cadernos de saude publica*, 35(3):e00153118, 2019. doi:10.1590/0102-311X00153118.

[25] Machine learning: applications of artificial intelligence to imaging and diagnosis. *Biophysical Reviews*, 11(1):111–118, 2019. doi:10.1007/s12551-018-0449-9.

[26] Kyle Strimbu and Jorge Tavel. The maturometer - Instrumental test and redesign. *Current Opinion HIV AIDS*, 5(6):463–466, 2010. doi:10.1097/COH.0b013e32833ed177.What.

[27] A. J. Atkinson, W. A. Colburn, V. G. DeGruttola, D. L. DeMets, G. J. Downing, D. F. Hoth, J. A. Oates, C. C. Peck, R. T. Schooley, B. A. Spilker, J. Woodcock, and S. L. Zeger. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology and Therapeutics*, 69(3):89–95, 2001. doi:10.1067/mcp.2001.113989.

[28] Last Updated. BEST (Biomarkers , EndpointS , and other Tools) Resource. (Md), 2020.

[29] PMC Members. Predictive Biomarkers. URL: http://www.personalizedmedicinecoalition.org/Education/Predictive{_}Biomarkers.

[30] Document Version. *Towards optimal decision making in personalized medicine Potential value assessment of biomarkers in heart failure exemplars*. 2016.

[31] Farhana R. Pinu, David J. Beale, Amy M. Paten, Konstantinos Kouremenos, Sanjay Swarup, Horst J. Schirra, and David Wishart. Systems biology and multi-omics integration: Viewpoints from the metabolomics research community. *Metabolites*, 9(4):1–31, 2019. doi:10.3390/metabo9040076.

[32] Manikhandan Mudaliar, Funmilola Clara Thomas, and Peter David Eckersall. *Omic approaches to a better understanding of mastitis in dairy cows*. 2017. doi:10.1007/978-3-319-43033-1_8.

[33] World Health. Report of the advisory committee on health research. *Revista Panamericana de Salud Publica/Pan American Journal of Public Health*, 2(6):428–434, 1997. doi:10.1590/S1020-49891997001200014.

[34] World Health Assembly. Resolutions and Decisions: Genomics and world health WHA57.13. *World Health Assembly*, pages 21–22, 2004.

[35] The Human Genome Project | NHGRI, 2020.

[36] Zhi Cheng Dong and Yan Chen. Transcriptomics: Advances and approaches. *Science China Life Sciences*, 56(10):960–967, 2013. doi:10.1007/s11427-013-4557-2.

[37] Mark Schena, Dari Shalon, Ronald W. Davis, and Patrick O. Brown. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270(5235):467–470, 1995. doi:10.1126/science.270.5235.467.

[38] Elsevier. Volume 3 Histone Phosphorylation. *ScienceDirect*, 2016. URL: https://www.sciencedirect.com/topics/neuroscience/histone-phosphorylation.

[39] Bensu Karahalil. Overview of Systems Biology and Omics Technologies. *Current Medicinal Chemistry*, 23(37):4221–4230, 2016. doi:10.2174/0929867323666160926150617.

[40] Kondethimmanahalli Chandramouli and Pei-Yuan Qian. Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity. *Human Genomics and Proteomics*, 1(1), 2009. doi:10.4061/2009/239204.

[41] László G. Boros, Megan R. Lerner, Daniel L. Morgan, Stephanie L. Taylor, Brenda J. Smith, Russell G. Postier, and Daniel J. Brackett. [1,2-13C2]-D-glucose profiles of the serum, liver, pancreas, and DMBA-induced pancreatic tumors of rats. *Pancreas*, 31(4):337–343, 2005. doi:10.1097/01.mpa.0000186524.53253.fb.

[42] Julian L. Griffin and John P. Shockcor. Metabolic profiles of cancer cells. *Nature Reviews Cancer*, 4(7):551–561, 2004. doi:10.1038/nrc1390.

[43] Anton Nekrutenko and James Taylor. Next-generation sequencing data interpretation: Enhancing reproducibility and accessibility. *Nature Reviews Genetics*, 13(9):667–672, 2012. doi:10.1038/nrg3305.

[44] Guy Haskin Fernald, Emidio Capriotti, Roxana Daneshjou, Konrad J. Karczewski, and Russ B. Altman. Bioinformatics challenges for personalized medicine. *Bioinformatics*, 27(13):1741–1748, 2011. doi:10.1093/bioinformatics/btr295.

[45] Abdellah Tebani, Carlos Afonso, Stéphane Marret, and Soumeya Bekri. Omics-based strategies in precision medicine: Toward a paradigm shift in inborn errors of metabolism investigations. *International Journal of Molecular Sciences*, 17(9), 2016. doi:10.3390/ijms17091555.

[46] NEURO-ONCOLOGY. *Oxford University Press on behalf of the Society for Neuro-Oncology*, page 2014, 2014. doi:10.1093.

[47] Evi S. Lianidou, Athina Markou, and Areti Strati. Molecular characterization of circulating tumor cells in breast cancer: Challenges and promises for individualized cancer treatment. *Cancer and Metastasis Reviews*, 31(3-4):663–671, 2012. doi:10.1007/s10555-012-9366-8.

[48] Miguel López-Lázaro. The stem cell division theory of cancer. *Critical Reviews in Oncology/Hematology*, 123:95–113, 2018. URL: http://dx.doi.org/10.1016/j.critrevonc.2018.01.010, doi:10.1016/j.critrevonc.2018.01.010.

[49] K. Ameri, R. Luong, H. Zhang, A. A. Powell, K. D. Montgomery, I. Espinosa, D. M. Bouley, A. L. Harris, and S. S. Jeffrey. Circulating tumour cells demonstrate an altered response to hypoxia and an aggressive phenotype. *British Journal of Cancer*, 102(3):561–569, 2010. URL: http://dx.doi.org/10.1038/sj.bjc.6605491, doi:10.1038/sj.bjc.6605491.

[50] J. Zhang, K. Chen, and Z. H. Fan. *Circulating Tumor Cell Isolation and Analysis*, volume 75. Elsevier Inc., 1 edition, 2016. URL: http://dx.doi.org/10.1016/bs.acc.2016.03.003, doi:10.1016/bs.acc.2016.03.003.

[51] Frank Diehl, Meng Li, Devin Dressman, Yiping He, Dong Shen, Steve Szabo, Luis A. Diaz, Steven N. Goodman, Kerstin A. David, Hartmut Juhl, Kenneth W. Kinzler, and Bert Vogelstein. Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proceedings of the National Academy of Sciences of the United States of America*, 102(45):16368–16373, 2005. doi:10.1073/pnas.0507904102.

[52] Sabine Jahr, Hannes Hentze, Sabine Englisch, Dieter Hardt, Frank O. Fackelmayer, Rolf Dieter Hesch, and Rolf Knippers. DNA fragments in the blood plasma of cancer patients: Quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Research*, 61(4):1659–1665, 2001.

[53] Marilesia Ferreira de Souza, Hellen Kuasne, Mateus de Camargo Barros-Filho, Heloísa Lizotti Cilião, Fabio Albuquerque Marchi, Paulo Emilio Fuganti, Alexandre Rossi Paschoal, Silvia Regina Rogatto, and Ilce Mara de Syllos Có lus. Circulating mRNAs and miRNAs as candidate markers for the diagnosis and prognosis of prostate cancer. *PLoS ONE*, pages 1–16, 2017. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5598937/pdf/pone.0184094.pdf, doi:10.1371/journal.pone.01840940184094.

[54] Heidi Schwarzenbach, Naohiro Nishida, George A. Calin, and Klaus Pantel. Clinical relevance of circulating cell-free microRNAs in cancer. *Nature Reviews Clinical Oncology*, 11(3):145–156, 2014. URL: http://dx.doi.org/10.1038/nrclinonc.2014.5, doi:10.1038/nrclinonc.2014.5.

[55] Ondrej Pös, Orsolya Biró, Tomas Szemes, and Bálint Nagy. Circulating cell-free nucleic acids: Characteristics and applications. *European Journal of Human Genetics*, 26(7):937–945, 2018. doi:10.1038/s41431-018-0132-4.

[56] Mercedes Tkach and Clotilde Théry. Communication by Extracellular Vesicles: Where We Are and Where We Need to Go. *Cell*, 164(6):1226–1232, 2016. doi:10.1016/j.cell.2016.01.043.

[57] Mikołaj P. Zaborowski, Leonora Balaj, Xandra O. Breakefield, and Charles P. Lai. Extracellular Vesicles: Composition, Biological Relevance, and Methods of Study. *BioScience*, 65(8):783–797, 2015. doi:10.1093/biosci/biv084.

[58] María Yáñez-Mó, Pia R.M. Siljander, Zoraida Andreu, Apolonija Bedina Zavec, Francesc E. Borràs, Edit I. Buzas, Krisztina Buzas, Enriqueta Casal, Francesco Cappello, Joana Carvalho, Eva Colás, Anabela Cordeiro-Da Silva, Stefano Fais, Juan M. Falcon-Perez, Irene M. Ghobrial, Bernd Giebel, Mario Gimona, Michael Graner, Ihsan Gursel, Mayda Gursel, Niels H.H. Heegaard, An Hendrix, Peter Kierulf, Katsutoshi Kokubun, Maja Kosanovic, Veronika Kralj-Iglic, Eva Maria Krämer-Albers, Saara Laitinen, Cecilia Lässer, Thomas Lener, Erzsébet Ligeti, Aija Line, Georg Lipps, Alicia Llorente, Jan Lötvall, Mateja Manček-Keber, Antonio Marcilla, Maria Mittelbrunn, Irina Nazarenko, Esther N.M. Nolte-'t Hoen, Tuula A. Nyman, Lorraine O'Driscoll, Mireia Olivan, Carla Oliveira, Éva Pállinger, Hernando A. Del Portillo, Jaume Reventós, Marina Rigau, Eva Rohde, Marei Sammar, Francisco Sánchez-Madrid, N. Santarém, Katharina Schallmoser, Marie Stampe Ostenfeld, Willem Stoorvogel, Roman Stukelj, Susanne G. Van Der Grein, M. Helena Vasconcelos, Marca H.M. Wauben, and Olivier De Wever. Biological properties of extracellular vesicles and their physiological functions. *Journal of Extracellular Vesicles*, 4(2015):1–60, 2015. doi:10.3402/jev.v4.27066.

[59] David R. Santiago-Dieppa, Jeffrey Steinberg, David Gonda, Vincent J. Cheung, Bob S. Carter, and Clark C. Chen. Extracellular vesicles as a platform for 'liquid biopsy' in glioblastoma patients. *Expert Review of Molecular Diagnostics*, 14(7):819–825, 2014. doi:10.1586/14737159.2014.943193.

[60] Susana Filipa and Jordão Viegas. Tumores Cerebrais-Histopatologia e Tratamento de Gliomas. 2018.

[61] NIH-National Cancer Institute. Adult Central Nervous System Tumors Treatment (PDQ®)–Patient Version - National Cancer Institute, mar 2020. URL: https://www.cancer.gov/types/brain/patient/adult-brain-treatment-pdq.

[62] Helivania Sardinha dos Santos. Células da glia - Biologia Net. URL: https://www.biologianet.com/histologia-animal/celulas-glia.htm.

[63] Seamus Caragher, Anthony J. Chalmers, and Natividad Gomez-Roman. Glioblastoma's next top model: Novel culture systems for brain cancer radiotherapy research. *Cancers*, 11(1), 2019. doi:10.3390/cancers11010044.

[64] SESARAM(Serviço de Saúde da RAM). Registo Oncológico na Região Autónoma da Madeira, 2017. 2019.

[65] Juliana Müller Bark, Arutha Kulasinghe, Benjamin Chua, Bryan W. Day, and Chamindie Punyadeera. Circulating biomarkers in patients with glioblastoma. *British Journal of Cancer*, 122(3):295–305, 2020. URL: http://dx.doi.org/10.1038/s41416-019-0603-6, doi:10.1038/s41416-019-0603-6.

[66] Ganesh M. Shankar, Leonora Balaj, Shannon L. Stott, Brian Nahed, and Bob S. Carter. Liquid biopsy for brain tumors. *Expert Review of Molecular Diagnostics*, 17(10):943–947, 2017. URL: https://doi.org/10.1080/14737159.2017.1374854, doi:10.1080/14737159.2017.1374854.

[67] Manfred Westphal and Katrin Lamszus. Circulating biomarkers for gliomas. *Nature Reviews Neurology*, 11(10):556–566, 2015. URL: http://dx.doi.org/10.1038/nrneurol.2015.171, doi:10.1038/nrneurol.2015.171.

[68] David N. Louis, Arie Perry, Guido Reifenberger, Andreas von Deimling, Dominique Figarella-Branger, Webster K. Cavenee, Hiroko Ohgaki, Otmar D. Wiestler, Paul Kleihues, and David W. Ellison. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica*, 131(6):803–820, 2016. doi:10.1007/s00401-016-1545-1.

[69] William Kupsky M.D. Momal Tara Chand, M.D. CNS tumor Diffuse astrocytic and oligodendroglial tumors Glioblastoma, IDH wildtype. URL: http://www.pathologyoutlines.com/topic/cnstumortumorglioIDHwild.html.

[70] Ph.D. Fausto J. Rodriguez M.D. Meaghan Morris, M.D. Pathology Outlines - Glioblastoma, IDH mutant, 2019. URL: http://www.pathologyoutlines.com/topic/cnstumorglioblastomaidhmutant.html.

[71] David N. Louis, Hiroko Ohgaki, Otmar D. Wiestler, Webster K. Cavenee, Peter C. Burger, Anne Jouvet, Bernd W. Scheithauer, and Paul Kleihues. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathologica*, 114(2):97–109, 2007. doi:10.1007/s00401-007-0243-4.

[72] Hiroko Ohgaki and Paul Kleihues. Genetic pathways to primary and secondary glioblastoma. *American Journal of Pathology*, 170(5):1445–1453, 2007. doi:10.2353/ajpath.2007.070011.

[73] Dietmar Krex, Barbara Klink, Christian Hartmann, Andreas Von Deimling, Torsten Pietsch, Matthias Simon, Michael Sabel, Joachim P. Steinbach, Oliver Heese, Guido Reifenberger, Michael Weller, and Gabriele Schackert. Long-term survival with glioblastoma multiforme. *Brain*, 130(10):2596–2606, 2007. doi:10.1093/brain/awm204.

[74] MD Anderson Cancer Center. Q&A: Understanding glioblastoma - YouTube, apr. URL: https://www.youtube.com/watch?v=8d0B5R3HLFY.

[75] Kaja Urbanska, Justyna Sokolowska, Maciej Szmidt, and Pawel Sysa. Glioblastoma multiforme - An overview. *Wspolczesna Onkologia*, 18(5):307–312, 2014. doi:10.5114/wo.2014.40559.

[76] Peter C Burger and Bernd W Scheithauer. Tumors of the Central Nervous System (AFIP Atlas of Tumor Pathology, Series 4. Fascicle 7). *Neuropathology and Applied Neurobiology*, 34(4):473–474, 2008. doi:10.1111/j.1365-2990.2008.00962.x.

[77] Hiroko Ohgaki and Paul Kleihues. Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas, 2005. URL: https://pubmed.ncbi.nlm.nih.gov/15977639/, doi:10.1093/jnen/64.6.479.

[78] G B Hafez C Padmalatha, R C Harruff, D Ganick. Glioblastoma multiforme with tuberous sclerosis. Report of a case - PubMed. URL: https://pubmed.ncbi.nlm.nih.gov/6254467/.

[79] E. Grips, N. Wentzensen, C. Sutter, O. Sedlaczek, J. Gebert, R. Weigel, A. Schwartz, M. von Knebel-Doeberitz, and M. Hennerici. Glioblastoma multiforme als Manifestation des Turcot-Syndroms. *Der Nervenarzt*, 73(2):177–182, feb 2002. doi:10.1007/s00115-001-1233-8.

[80] Ruth Sánchez-Ortiga, Evangelina Boix Carreño, Oscar Moreno-Pérez, and Antonio Picó Alfonso. Glioblastoma multiforme y neoplasia endocrina múltiple tipo 2 A, jul 2009. doi:10.1016/j.medcli.2008.06.021.

[81] Marike L. D. Broekman, Roelof Risselada, JooYeon Engelen-Lee, Wim G. M. Spliet, and Bon H. Verweij. Glioblastoma Multiforme in the Posterior Cranial Fossa in a Patient with Neurofibromatosis Type I. *Case Reports in Medicine*, 2009:1–4, 2009. doi:10.1155/2009/757898.

[82] Dr Jacques GRILL. Orphanet: Glioblastoma, nov 2007. URL: https://www.orpha.net/consor/cgi-bin/OC{_}Exp.php?Lng=PT{&}Expert=360.

[83] Patrícia Trevisan. UNIVERSIDADE FEDERAL DE CIÊNCIAS DA SAÚDE DE PORTO ALEGRE-UFCSPA PROGRAMA DE PÓS-GRADUAÇÃO EM PATOLOGIA. Technical report, jun 2018.

[84] Telmo Belsuzarri, Maick Neves, Otávio Costa, Diego Soares, Fernando Melo Filho, Mariana Guimarães, Tiago Gonçales, Wolnei Zeviani, and João Araújo. Heterogeneidade dos tumores cerebrais. *Arquivos Brasileiros de Neurocirurgia: Brazilian Neurosurgery*, (June), 2018. doi:10.1055/s-0038-1625962.

[85] Silvia Hofer, Elisabeth Rushing, Matthias Preusser, and Christine Marosi. Molecular biology of high-grade gliomas: What should the clinician know?, 2014. doi:10.5732/cjc.013.10218.

[86] Dong Hoon Lee, Hyun Wook Ryu, Hye Rim Won, and So Hee Kwon. Advances in epigenetic glioblastoma therapy, 2017. doi:10.18632/oncotarget.14612.

[87] Maria del Mar Inda, Rudy Bonavia, and Joan Seoane. Glioblastoma multiforme:A look inside its heterogeneous nature, mar 2014. doi:10.3390/cancers6010226.

[88] Peter C. Nowell. The clonal evolution of tumor cell populations. *Science*, 194(4260):23–28, 1976. doi:10.1126/science.959840.

[89] Romana Rea Begicevic and Marco Falasca. ABC transporters in cancer stem cells: Beyond chemoresistance, nov 2017. doi:10.3390/ijms18112362.

[90] Andriy Marusyk and Kornelia Polyak, jan 2010. doi:10.1016/j.bbcan.2009.11.002.

[91] Qinghai Liu, David H. Nguyen, Qinghua Dong, Peter Shitaku, Kenneth Chung, On Ying Liu, Jonathan L. Tso, Jason Y. Liu, Veerauo Konkankit, Timothy F. Cloughesy, Paul S. Mischel, Timothy F. Lane, Linda M. Liau, Stanley F. Nelson, and Cho Lea Tso. Molecular properties of CD133+ glioblastoma stem cells derived from treatment-refractory recurrent brain tumors. *Journal of Neuro-Oncology*, 94(1):1–19, 2009. doi:10.1007/s11060-009-9919-z.

[92] Marco Gerlinger, Andrew J. Rowan, Stuart Horswell, James Larkin, David Endesfelder, Eva Gronroos, Pierre Martinez, Nicholas Matthews, Aengus Stewart, Patrick Tarpey, Ignacio Varela, Benjamin Phillimore, Sharmin Begum, Neil Q. McDonald, Adam Butler, David Jones, Keiran Raine, Calli Latimer, Claudio R. Santos, Mahrokh Nohadani, Aron C. Eklund, Bradley Spencer-Dene, Graham Clark, Lisa Pickering, Gordon Stamp, Martin Gore, Zoltan Szallasi, Julian Downward, P. Andrew Futreal, and Charles Swanton. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *New England Journal of Medicine*, 366(10):883–892, mar 2012. doi:10.1056/NEJMoa1113205.

[93] Timothy A. Yap, Marco Gerlinger, P. Andrew Futreal, Lajos Pusztai, and Charles Swanton. Intratumor heterogeneity: Seeing the wood for the trees, mar 2012. doi:10.1126/scitranslmed.3003854.

[94] Jonathon J. Parker, Peter Canoll, Lee Niswander, B. K. Kleinschmidt-DeMasters, Kara Foshay, and Allen Waziri. Intratumoral heterogeneity of endogenous tumor cell invasive behavior in human glioblastoma. *Scientific Reports*, 8(1):1–10, 2018. doi:10.1038/s41598-018-36280-9.

[95] Cameron W. Brennan, Roel G.W. Verhaak, Aaron McKenna, Benito Campos, Houtan Noushmehr, Sofie R. Salama, Siyuan Zheng, Debyani Chakravarty, J. Zachary Sanborn, Samuel H. Berman, Rameen Beroukhim, Brady Bernard, Chang Jiun Wu, Giannicola Genovese, Ilya Shmulevich, Jill Barnholtz-Sloan, Lihua Zou, Rahulsimham Vegesna, Sachet A. Shukla, Giovanni Ciriello, W. K. Yung, Wei Zhang, Carrie Sougnez, Tom Mikkelsen, Kenneth Aldape, Darell D. Bigner, Erwin G. Van Meir, Michael Prados, Andrew Sloan, Keith L. Black, Jennifer Eschbacher, Gaetano Finocchiaro, William Friedman, David W. Andrews,

Abhijit Guha, Mary Iacocca, Brian P. O'Neill, Greg Foltz, Jerome Myers, Daniel J. Weisenberger, Robert Penny, Raju Kucherlapati, Charles M. Perou, D. Neil Hayes, Richard Gibbs, Marco Marra, Gordon B. Mills, Eric S. Lander, Paul Spellman, Richard Wilson, Chris Sander, John Weinstein, Matthew Meyerson, Stacey Gabriel, Peter W. Laird, David Haussler, Gad Getz, Lynda Chin, Christopher Benz, Wendi Barrett, Quinn Ostrom, Yingli Wolinsky, Bikash Bose, Paul T. Boulos, Madgy Boulos, Jenn Brown, Christine Czerinski, Matthew Eppley, Thelma Kempista, Teresa Kitko, Yakov Koyfman, Brenda Rabeno, Pawan Rastogi, Michael Sugarman, Patricia Swanson, Kennedy Yalamanchii, Ilana P. Otey, Yingchun Spring Liu, Yonghong Xiao, J. Todd Auman, Peng Chieh Chen, Angela Hadjipanayis, Eunjung Lee, Semin Lee, Peter J. Park, Jonathan Seidman, Lixing Yang, Steven Kalkanis, Laila M. Poisson, Aditya Raghunathan, Lisa Scarpace, Ryan Bressler, Andrea Eakin, Lisa Iype, Richard B. Kreisberg, Kalle Leinonen, Sheila Reynolds, Hector Rovira, Vesteinn Thorsson, Matti J. Annala, Joseph Paulauskis, Erin Curley, Martha Hatfield, David Mallery, Scott Morris, Troy Shelton, Candace Shelton, Mark Sherman, Peggy Yena, Lucia Cuppini, Francesco DiMeco, Marica Eoli, Emanuela Maderna, Bianca Pollo, Marco Saini, Saianand Balu, Katherine A. Hoadley, Ling Li, C. Ryan Miller, Yan Shi, Michael D. Topal, Junyuan Wu, Gavin Dunn, Caterina Giannini, B. Arman Aksoy, Yevgeniy Antipin, Laetitia Borsu, Ethan Cerami, Jianjiong Gao, Benjamin Gross, Anders Jacobsen, Marc Ladanyi, Alex Lash, Yupu Liang, Boris Reva, Nikolaus Schultz, Ronglai Shen, Nicholas D. Socci, Agnes Viale, Martin L. Ferguson, Qing Rong Chen, John A. Demchok, Laura A.L. Dillon, Kenna R. Mills Shaw, Margi Sheth, Roy Tarnuzzer, Zhining Wang, Liming Yang, Tanja Davidsen, Mark S. Guyer, Bradley A. Ozenberger, Heidi J. Sofia, Julie Bergsten, John Eckman, Jodi Harr, Christine Smith, Kelly Tucker, Cindy Winemiller, Leigh Anne Zach, Julia Y. Ljubimova, Greg Eley, Brenda Ayala, Mark A. Jensen, Ari Kahn, Todd D. Pihl, David A. Pot, Yunhu Wan, Nathan Hansen, Parvi Hothi, Biaoyang Lin, Nameeta Shah, Jae Geun Yoon, Ching Lau, Michael Berens, Kristin Ardlie, Scott L. Carter, Andrew D. Cherniack, Mike Noble, Juok Cho, Kristian Cibulskis, Daniel DiCara, Scott Frazer, Stacey B. Gabriel, Nils Gehlenborg, Jeff Gentry, David Heiman, Jaegil Kim, Rui Jing, Michael Lawrence, Pei Lin, Will Mallard, Robert C. Onofrio, Gordon Saksena, Steve Schumacher, Petar Stojanov, Barbara Tabak, Doug Voet, Hailei Zhang, Nathan N. Dees, Li Ding, Lucinda L. Fulton, Robert S. Fulton, Krishna Latha Kanchi, Elaine R. Mardis, Richard K. Wilson, Stephen B. Baylin, Larry Harshyne, Mark L. Cohen, Karen Devine, Andrew E. Sloan, Scott R. Van Den Berg, Mitchel S. Berger, Daniel Carlin, Brian Craft, Kyle Ellrott, Mary Goldman, Theodore Goldstein, Mia Grifford, Singer Ma, Sam Ng, Joshua Stuart, Teresa Swatloski, Peter Waltman, Jing Zhu, Robin Foss, Barbara Frentzen, Raquel McTiernan, Anthony Yachnis, Yong Mao, Rehan Akbani, Oliver Bogler, Gregory N. Fuller, Wenbin Liu, Yuexin Liu, Yiling Lu, Alexei Protopopov, Xiaojia Ren, Youting Sun, W. K.Alfred Yung, Jianhua Zhang, Ken Chen, John N. Weinstein, Moiz S. Bootwalla, Phillip H. Lai, Timothy J. Triche, David J. Van Den Berg, David H. Gutmann, Norman L. Lehman, Daniel Brat, Jeffrey J. Olson, Gena M. Mastrogianakis, Narra S. Devi, Zhaobin Zhang, Eric Lipp, and Roger McLendon. The somatic genomic landscape of glioblastoma. *Cell*, 155(2):462, oct 2013. doi:10.1016/j.cell.2013.09.034.

[96] Preferential Iron Trafficking Characterizes Glioblastoma Stem-like Cells. *Cancer Cell*, 28(4):441–455, oct 2015. doi:10.1016/j.ccell.2015.09.002.

[97] Monika E. Hegi, Lili Liu, James G. Herman, Roger Stupp, Wolfgang Wick, Michael Weller, Minesh P. Mehta, and Mark R. Gilbert. Correlation of O6-methylguanine methyltransferase

(MGMT) promoter methylation with clinical outcomes in glioblastoma and clinical strategies to modulate MGMT activity, 2008. doi:10.1200/JCO.2007.11.5964.

[98] Nobuaki Ishii, Mitsuhiro Tada, Marie France Hamou, Robert C. Janzer, Kathleen Meagher-Villemure, Otmar D. Wiestler, Nicolas De Tribolet, and Erwin G. Van Meir. Cells with TP53 mutations in low grade astrocytic tumors evolve clonally to malignancy and are an unfavorable prognostic factor. *Oncogene*, 18(43):5870–5878, oct 1999. doi:10.1038/sj.onc.1203241.

[99] H. J. Su Huang, Motoo Nagane, Candice K. Klingbeil, Hong Lin, Ryo Nishikawa, Xiang Dong Ji, Chun Ming Huang, Gordon N. Gill, H. Steven Wiley, and Webster K. Cavenee. The enhanced tumorigenic activity of a mutant epidermal growth factor receptor common in human cancers is mediated by threshold levels of constitutive tyrosine phosphorylation and unattenuated signaling. *Journal of Biological Chemistry*, 272(5):2927–2935, 1997. doi:10.1074/jbc.272.5.2927.

[100] Abir Mondal, Divya Kumari Singh, Suchismita Panda, and Anjali Shiras. Extracellular vesicles as modulators of tumor microenvironment and disease progression in glioma, jul 2017. doi:10.3389/fonc.2017.00144.

[101] Eefje M. Sizoo, Lies Braam, Tjeerd J. Postma, H. Roeline W. Pasman, Jan J. Heimans, Martin Klein, Jaap C. Reijneveld, and Martin J.B. Taphoorn. Symptoms and problems in the end-of-life phase of high-grade glioma patients. *Neuro-Oncology*, 12(11):1162–1166, nov 2010. doi:10.1093/neuonc/nop045.

[102] Aldape KD Brat DJ Biernat W Bigner DD Nakazato Y Plate KH Giangaspero F von Deimling A Ohgaki H Cavenee WK Kleihues P, Burger PC. Glioblastoma. *World Health Organization Classification of Tumors of the Central Nervous System.*, pages 46–47, 2007.

[103] Gorkem Aksu Oguz Cetinayak Okan Kuzhan Fatih Ors Murat Beyzadeoglu Cüneyt Ulutin 1, Merdan Fayda. Primary glioblastoma multiforme in younger patients: a single-institution experience - PubMed, sep 2006.

[104] J. R. Simpson, J. Horton, C. Scott, W. J. Curran, P. Rubin, J. Fischbach, S. Isaacson, M. Rotman, S. O. Asbell, J. S. Nelson, A. S. Weinstein, and D. F. Nelson. Influence of location and extent of surgical resection on survival of patients with glioblastoma multiforme: Results of three consecutive radiation therapy oncology group (RTOG) clinical trials. *International Journal of Radiation Oncology, Biology, Physics*, (2):239–244, may 1993. doi:10.1016/0360-3016(93)90203-8.

[105] Antonio Omuro and Lisa M. DeAngelis. Glioblastoma and other malignant gliomas: A clinical review, 2013. doi:10.1001/jama.2013.280319.

[106] Ganesh M. Shankar, Leonora Balaj, Shannon L. Stott, Brian Nahed, and Bob S. Carter. Liquid biopsy for brain tumors, oct 2017. doi:10.1080/14737159.2017.1374854.

[107] Veronique Frattini, Vladimir Trifonov, Joseph Minhow Chan, Angelica Castano, Marie Lia, Francesco Abate, Stephen T. Keir, Alan X. Ji, Pietro Zoppoli, Francesco Niola, Carla Danussi, Igor Dolgalev, Paola Porrati, Serena Pellegatta, Adriana Heguy, Gaurav Gupta, David J. Pisapia, Peter Canoll, Jeffrey N. Bruce, Roger E. McLendon, Hai Yan, Ken Aldape, Gaetano Finocchiaro, Tom Mikkelsen, Gilbert G. Privé, Darell D. Bigner, Anna

Lasorella, Raul Rabadan, and Antonio Iavarone. The integrated landscape of driver genomic alterations in glioblastoma. *Nature Genetics*, 45(10):1141–1149, oct 2013. doi:10.1038/ng.2734.

[108] Hartwig Wolburg, Susan Noell, Petra Fallier-Becker, Andreas F. MacK, and Karen Wolburg-Buchholz. The disturbed blood-brain barrier in human glioblastoma, oct 2012. doi:10.1016/j.mam.2012.02.003.

[109] Mohammed M.A. Almutairi, Chen Gong, Yuexian G. Xu, Yanzhong Chang, and Honglian Shi. Factors controlling permeability of the blood-brain barrier, jan 2016. doi:10.1007/s00018-015-2050-8.

[110] Zhihong Chen and Dolores Hambardzumyan. Immune microenvironment in glioblastoma subtypes, may 2018. doi:10.3389/fimmu.2018.01004.

[111] Peter Ulz, Gerhard G. Thallinger, Martina Auer, Ricarda Graf, Karl Kashofer, Stephan W. Jahn, Luca Abete, Gunda Pristauz, Edgar Petru, Jochen B. Geigl, Ellen Heitzer, and Michael R. Speicher. Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nature Genetics*, 48(10):1273–1278, oct 2016. doi:10.1038/ng.3648.

[112] Safia El Messaoudi, Fanny Rolet, Florent Mouliere, and Alain R. Thierry. Circulating cell free DNA: Preanalytical considerations, sep 2013. doi:10.1016/j.cca.2013.05.022.

[113] Jackson D. Hamilton, Marion Rapp, Timo Marcel Schneiderhan, Michael Sabel, Anne Hayman, Axel Scherer, Patric Kröpil, Wilfried Budach, Usha Kretschmar, Peter Arne Gerber, Sujit Prabhu, Lawrence E. Ginsberg, Edwin Bölke, and Christiane Matuschek. Glioblastoma multiforme metastasis outside the CNS: Three case reports and possible mechanisms of escape. *Journal of Clinical Oncology*, 32(22), aug 2014. doi:10.1200/JCO.2013.48.7546.

[114] James P. Sullivan, Brian V. Nahed, Marissa W. Madden, Samantha M. Oliveira, Simeon Springer, Deepak Bhere, Andrew S. Chi, Hiroaki Wakimoto, S. Michael Rothenberg, Lecia V. Sequist, Ravi Kapur, Khalid Shah, A. John Iafrate, William T. Curry, Jay S. Loeffler, Tracy T. Batchelor, David N. Louis, Mehmet Toner, Shyamala Maheswaran, and Daniel A. Haber. Brain tumor cells in circulation are enriched for mesenchymal gene expression. *Cancer Discovery*, 4(11):1299–1309, nov 2014. doi:10.1158/2159-8290.CD-14-0471.

[115] Lidija Orlic, Branka Sladoje-Martinovic, Ivana Mikolasevic, Zeljko Zupan, and Sanjin Racki. Patients with primary brain tumors as organ donors. *BANTAO Journal*, 13(1):34–38, dec 2015. doi:10.1515/bj-2015-0007.

[116] Juliana Müller Bark, Arutha Kulasinghe, Benjamin Chua, Bryan W. Day, and Chamindie Punyadeera. Circulating biomarkers in patients with glioblastoma, feb 2020. doi:10.1038/s41416-019-0603-6.

[117] Riichi Tawa, Tetsuya Ono, Akihiro Kurishita, Shigefumi Okada, and Shingo Hirose. Changes of DNA methylation level during pre-and postnatal periods in mice. *Differentiation*, 45(1):44–48, oct 1990. doi:10.1111/j.1432-0436.1990.tb00455.x.

[118] Christine Ladd-Acosta, Jonathan Pevsner, Sarven Sabunciyan, Robert H. Yolken, Maree J. Webster, Tiffany Dinkins, Pauline A. Callinan, Jian Bing Fan, James B. Potash, and Andrew P. Feinberg. DNA methylation signatures within the human brain. *American Journal of Human Genetics*, 81(6):1304–1315, 2007. doi:10.1086/524110.

[119] Kaoru Goto, Masayuki Numata, Jun Ichiro Komura, Tetsuya Ono, Timothy H. Bestor, and Hisatake Kondo. Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. *Differentiation*, 56(1-2):39–44, 1994. doi:10.1046/j.1432-0436.1994.56120039.x.

[120] Jacquetta M. Trasler, Daphne G. Trasler, Timothy H. Bestor, En Li, and Felicia Ghibu. DNA methyltransferase in normal and Dnmt(n)/Dnmt(n) mouse embryos. *Developmental Dynamics*, 206(3):239–247, jul 1996. doi:10.1002/(SICI)1097-0177(199607)206:3<239::AID-AJA2>3.0.CO;2-J.

[121] Fabio Mohn, Michael Weber, Michael Rebhan, Tim C. Roloff, Jens Richter, Michael B. Stadler, Miriam Bibel, and Dirk Schübeler. Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. *Molecular Cell*, 30(6):755–766, jun 2008. doi:10.1016/j.molcel.2008.05.007.

[122] Ruthie E. Amir, Ignatia B. Van Den Veyver, Mimi Wan, Charles Q. Tran, Uta Francke, and Huda Y. Zoghbi. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl- CpG-binding protein 2. *Nature Genetics*, 23(2):185–188, oct 1999. doi:10.1038/13810.

[123] Lars Riff Jensen, Marion Amende, Ulf Gurok, Bettina Moser, Verena Gimmel, Andreas Tzschach, Andreas R. Janecke, Gholamali Tariverdian, Jamel Chelly, Jean Pierre Fryns, Hilde Van Esch, Tjitske Kleefstra, Ben Hamel, Claude Moraine, Jozef Gécz, Gillian Turner, Richard Reinhardt, Vera M. Kalscheuer, Hans Hilger Ropers, and Steffen Lenzner. Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. *American Journal of Human Genetics*, 76(2):227–236, 2005. doi:10.1086/427563.

[124] Patrick Y. Wen and Santosh Kesari. Malignant gliomas in adults, jul 2008. doi:10.1056/NEJMra0708126.

[125] Kathy Rock, O. McArdle, P. Forde, M. Dunne, D. Fitzpatrick, B. O'Neill, and C. Faul. A clinical review of treatment outcomes in glioblastoma multiforme - The validation in a non-trial population of the results of a randomised Phase III clinical trial: Has a more radical approach improved survival? *British Journal of Radiology*, 85(1017), sep 2012. doi:10.1259/bjr/83796755.

[126] Raman P. Nagarajan and Joseph F. Costello. Epigenetic mechanisms in glioblastoma multiforme, jun 2009. doi:10.1016/j.semcancer.2009.02.005.

[127] Chibo Hong, K. Scott Moorefield, Peter Jun, Kenneth D. Aldape, Samir Kharbanda, Heidi S. Phillips, and Joseph F. Costello. Epigenome scans and cancer genome sequencing converge on WNK2, a kinase-independent suppressor of cell growth. *Proceedings of the National Academy of Sciences of the United States of America*, 104(26):10974–10979, jun 2007. doi:10.1073/pnas.0700683104.

[128] S. Moniz, F. Veríssimo, P. Matos, R. Brazão, E. Silva, L. Kotevelets, E. Chastre, C. Gespach, and P. Jordan. Protein kinase WNK2 inhibits cell proliferation by negatively modulating the activation of MEK1/ERK1/2. *Oncogene*, 26(41):6071–6081, sep 2007. doi:10.1038/sj.onc.1210706.

[129] D. Williams Parsons, Siân Jones, Xiaosong Zhang, Jimmy Cheng Ho Lin, Rebecca J. Leary, Philipp Angenendt, Parminder Mankoo, Hannah Carter, I. Mei Siu, Gary L. Gallia, Alessandro Olivi, Roger McLendon, B. Ahmed Rasheed, Stephen Keir, Tatiana Nikolskaya, Yuri Nikolsky, Dana A. Busam, Hanna Tekleab, Luis A. Diaz, James Hartigan, Doug R. Smith, Robert L. Strausberg, Suely Kazue Nagahashi Marie, Sueli Mieko Oba Shinjo, Hai Yan, Gregory J. Riggins, Darell D. Bigner, Rachel Karchin, Nick Papadopoulos, Giovanni Parmigiani, Bert Vogelstein, Victor E. Velculescu, and Kenneth W. Kinzler. An integrated genomic analysis of human glioblastoma multiforme. *Science*, 321(5897):1807–1812, sep 2008. doi:10.1126/science.1164382.

[130] James G. Herman and Stephen B. Baylin. Gene Silencing in Cancer in Association with Promoter Hypermethylation, nov 2003. doi:10.1056/NEJMra023075.

[131] Adam R. Karpf and Sei Ichi Matsui. Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. *Cancer Research*, 65(19):8635–8639, oct 2005. doi:10.1158/0008-5472.CAN-05-1961.

[132] Miguel A. Gama-sosa, Valerie A. Slagel, Ronald W. Trewyn, Ronald Oxenhandler, Kenneth C. Kuo, Charles W. Gehrke, and Melanie Ehrlich. The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Research*, 11(19):6883–6894, oct 1983. doi:10.1093/nar/11.19.6883.

[133] Kallur Saraswathy, Lovejeet Kaur, Seerat Talwar, Jyoti Mishra, Suraj Huidrom, M. Sachdeva, and Manju Puri. Methylenetetrahydrofolate reductase gene-specific methylation and recurrent miscarriages: A Case-Control Study from North India. *Journal of Human Reproductive Sciences*, 11(2):142–147, apr 2018. doi:10.4103/jhrs.JHRS_145_17.

[134] Mario F. Fraga, Esteban Ballestar, Ana Villar-Garea, Manuel Boix-Chornet, Jesus Espada, Gunnar Schotta, Tiziana Bonaldi, Claire Haydon, Santiago Ropero, Kevin Petrie, N. Gopalakrishna Iyer, Alberto Pérez-Rosado, Enrique Calvo, Juan A. Lopez, Amparo Cano, Maria J. Calasanz, Dolors Colomer, Miguel Ángel Piris, Natalie Ahn, Axel Imhof, Carlos Caldas, Thomas Jenuwein, and Manel Esteller. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature Genetics*, 37(4):391–400, apr 2005. doi:10.1038/ng1531.

[135] Martin Widschwendter, Guanchao Jiang, Christian Woods, Hannes M. Müller, Heidi Fiegl, Georg Goebel, Christian Marth, Elisabeth Müller-Holzner, Alain G. Zeimet, Peter W. Laird, and Melanie Ehrlich. DNA hypomethylation and ovarian cancer biology. *Cancer Research*, 64(13):4472–4480, jul 2004. doi:10.1158/0008-5472.CAN-04-0238.

[136] Melanie Ehrlich. The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clinical Immunology*, 109(1):17–28, 2003. doi:10.1016/S1521-6616(03)00201-8.

[137] Jian Yu, Hongyu Zhang, Jun Gu, Song Lin, Junhua Li, Wei Lu, Yifei Wang, and Jingde Zhu. Methylation profiles of thirty four promoter-CpG islands and concordant methylation behaviours of sixteen genes that may contribute to carcinogenesis of astrocytoma. *BMC Cancer*, 4(1):65, sep 2004. doi:10.1186/1471-2407-4-65.

[138] Charles De Smet, Olivier De Backer, Isabella Faraoni, Christophe Lurquin, Francis Brasseur, and Thierry Boon. The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. *Proceedings of the National Academy of Sciences*

of the United States of America, 93(14):7149–7153, jul 1996. doi:10.1073/pnas.93.14.7149.

[139] Pierre Van der Bruggen, Yi Zhang, Pascal Chaux, Vincent Stroobant, Christophe Panichelli, Erwin S. Schultz, Jacques Chapiro, Benoît J. Van den Eynde, Francis Brasseur, and Thierry Boon. Tumor-specific shared antigenic peptides recognized by human T cells. *Immunological Reviews*, 188:51–64, oct 2002. doi:10.1034/j.1600-065X.2002.18806.x.

[140] Silencing of p16/CDKN2 expression in human gliomas by methylation and chromatin condensation - PubMed. URL: https://pubmed.ncbi.nlm.nih.gov/8625319/.

[141] Nakamura M, Yonekawa Y, Kleihues P, and Ohgaki H. Promoter hypermethylation of the RB1 gene in glioblastomas. *Laboratory investigation; a journal of technical methods and pathology*, 81(1), 2001. doi:10.1038/LABINVEST.3780213.

[142] The p53/Mdm2/p14ARF cell cycle control pathway genes may be inactivated by genetic and epigenetic mechanisms in gliomas [1]. URL: https://www.researchgate.net/publication/7339835{_}The{_}p53Mdm2p14ARF{_}cell{_}cycle{_}control{_}pathway{_}genes{_}may{_

[143] Alaminos M, Dávalos V, Ropero S, Setién F, Paz MF, Herranz M, Fraga MF, Mora J, Cheung NK, Gerald WL, and Esteller M. EMP3, a myelin-related gene located in the critical 19q13.3 region, is epigenetically silenced and exhibits features of a candidate tumor suppressor in glioma and neuroblastoma. *Cancer Research*, 65(7):2565–2571, apr 2005. URL: https://europepmc.org/article/med/15805250, doi:10.1158/0008-5472.CAN-04-4283.

[144] Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. *Nature reviews. Cancer*, 4(4), 2004. doi:10.1038/NRC1319.

[145] Graded methylation in the promoter and body of the O6-methylguanine DNA methyltransferase (MGMT) gene correlates with MGMT expression in human glioma cells - PubMed. URL: https://pubmed.ncbi.nlm.nih.gov/8006031/.

[146] Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia - PubMed. URL: https://pubmed.ncbi.nlm.nih.gov/10029064/.

[147] Promoter hypermethylation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis - PubMed. URL: https://pubmed.ncbi.nlm.nih.gov/11406538/.

[148] Bigner D Van Meir EG Brat DJ Mastrogianakis M et al McLendon R, Friedman A. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature 2008 455:7216*, 455(7216):1061–1068, sep 2008. doi:10.1038/nature07385.

[149] Monika E. Hegi, Annie-Claire Diserens, Thierry Gorlia, Marie-France Hamou, Nicolas de Tribolet, Michael Weller, Johan M. Kros, Johannes A. Hainfellner, Warren Mason, Luigi Mariani, Jacoline E.C. Bromberg, Peter Hau, René O. Mirimanoff, J. Gregory Cairncross, Robert C. Janzer, and Roger Stupp. MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma. *http://dx.doi.org/10.1056/NEJMoa043331*, oct 2009. doi:10.1056/NEJMOA043331.

[150] Everhard S, Kaloshi G, Crinière E, Benouaich-Amiel A, Lejeune J, Marie Y, Sanson M, Kujas M, Mokhtari K, Hoang-Xuan K, Delattre JY, and Thillet J. MGMT methylation: a marker of response to temozolomide in low-grade gliomas. *Annals of neurology*, 60(6), 2006. doi:10.1002/ANA.21044.

[151] Brandes AA, Franceschi E, Tosoni A, Blatt V, Pession A, Tallini G, Bertorelle R, Bartolini S, Calbucci F, Andreoli A, Frezza G, Leonardi M, Spagnolli F, and Ermani M. MGMT promoter methylation status can predict the incidence and outcome of pseudoprogression after concomitant radiochemotherapy in newly diagnosed glioblastoma patients. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 26(13), 2008. doi:10.1200/JCO.2007.14.8163.

[152] Uhlmann K, Rohde K, Zeller C, Szymas J, Vogel S, Marczinek K, Thiel G, Nürnberg P, and Laird PW. Distinct methylation profiles of glioma subtypes. *International journal of cancer*, 106(1), 2003. doi:10.1002/IJC.11175.

[153] Hiroko Ohgaki and Paul Kleihues. Genetic Pathways to Primary and Secondary Glioblastoma. *The American Journal of Pathology*, 170(5):1445, 2007. doi:10.2353/AJPATH.2007.070011.

[154] John K. Wiencke, Shichun Zheng, Nanette Jelluma, Tarik Tihan, Scott Vandenberg, Tanja Tamgüney, Rachel Baumber, Ramon Parsons, Kathleen R. Lamborn, Mitchel S. Berger, Margaret R. Wrensch, Daphne Adele Haas-Kogan, and David Stokoe. Methylation of the PTEN promoter defines low-grade gliomas and secondary glioblastoma. *Neuro-Oncology*, 9(3):271, 2007. doi:10.1215/15228517-2007-003.

[155] Chibo Hong, Alika Maunakea, Peter Jun, Andrew W. Bollen, J. Graeme Hodgson, David D. Goldenberg, William A. Weiss, and Joseph F. Costello. Shared epigenetic mechanisms in human and mouse gliomas inactivate expression of the growth suppressor SLC5A8. *Cancer Research*, 65(9):3617–3623, may 2005. doi:10.1158/0008-5472.CAN-05-0048.

[156] M. Josefa Bello, Victor Martinez-Glez, Carmen Franco-Hernandez, Carolina Pefla-Granero, Jose M. de Campos, Alberto Isla, Luis Lassaletta, Jesus Vaquero, and Juan A. Rey. DNA methylation pattern in 16 tumor-related genes in schwannomas, jan 2007. doi:10.1016/j.cancergencyto.2006.02.022.

[157] Jesus Lomas, M. Josefa Bello, Dolores Arjona, M. Eva Alonso, Victor Martinez-Glez, Isabel Lopez-Marin, Cinthia Amiñoso, Jose M. De Campos, Alberto Isla, Jesus Vaquero, and Juan A. Rey. Genetic and epigenetic alteration of the NF2 gene in sporadic meningiomas. *Genes Chromosomes and Cancer*, 42(3):314–319, mar 2005. doi:10.1002/gcc.20141.

[158] Jeongwu Lee, Myung Jin Son, Kevin Woolard, Nicholas M. Donin, Aiguo Li, Chui H. Cheng, Svetlana Kotliarova, Yuri Kotliarov, Jennifer Walling, Susie Ahn, Misuk Kim, Mariam Totonchy, Thomas Cusack, Chibawanye Ene, Hilary Ma, Qin Su, Jean Claude Zenklusen, Wei Zhang, Dragan Maric, and Howard A. Fine. Epigenetic-Mediated Dysfunction of the Bone Morphogenetic Protein Pathway Inhibits Differentiation of Glioblastoma-Initiating Cells. *Cancer Cell*, 13(1):69–80, jan 2008. doi:10.1016/j.ccr.2007.12.005.

[159] Mi Yi Joo, Hsing Chen Tsai, Sabine C. Glöckner, Steven Lin, Joyce E. Ohm, Hari Easwaran, C. David James, Joseph F. Costello, Gregory Riggins, Charles G. Eberhart,

John Laterra, Angelo L. Vescovi, Nita Ahuja, James G. Herman, Kornel E. Schuebel, and Stephen B. Baylin. Abnormal DNA methylation of CD133 in colorectal and glioblastoma tumors. *Cancer Research*, 68(19):8094–8103, oct 2008. doi:10.1158/0008-5472.CAN-07-6208.

[160] Bradley E. Bernstein, Tarjei S. Mikkelsen, Xiaohui Xie, Michael Kamal, Dana J. Huebert, James Cuff, Ben Fry, Alex Meissner, Marius Wernig, Kathrin Plath, Rudolf Jaenisch, Alexandre Wagschal, Robert Feil, Stuart L. Schreiber, and Eric S. Lander. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell*, 125(2):315–326, apr 2006. doi:10.1016/j.cell.2006.02.041.

[161] Joyce E. Ohm, Kelly M. McGarvey, Xiaobing Yu, Linzhao Cheng, Kornel E. Schuebel, Leslie Cope, Helai P. Mohammad, Wei Chen, Vincent C. Daniel, Wayne Yu, David M. Berman, Thomas Jenuwein, Kevin Pruitt, Saul J. Sharkis, D. Neil Watkins, James G. Herman, and Stephen B. Baylin. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nature Genetics*, 39(2):237–242, feb 2007. doi:10.1038/ng1972.

[162] Valtteri Häyry, Minna Tanner, Tea Blom, Olli Tynninen, Annariikka Roselli, Miina Ollikainen, Hannu Sariola, Kirmo Wartiovaara, and Nina N. Nupponen. Copy number alterations of the polycomb gene BMI1 in gliomas. *Acta Neuropathologica*, 116(1):97–102, jul 2008. doi:10.1007/s00401-008-0376-0.

[163] Milos Dokmanovic, Cathy Clarke, and Paul A. Marks. Histone deacetylase inhibitors: Overview and perspectives, oct 2007. doi:10.1158/1541-7786.MCR-07-0324.

[164] Agda K.B. Lucio-Eterovic, Maria A.A. Cortez, Elvis T. Valera, Fabio J.N. Motta, Rosane G.P. Queiroz, Helio R. Machado, Carlos G. Carlotti, Luciano Neder, Carlos A. Scrideli, and Luiz G. Tone. Differential expression of 12 histone deacetylase (HDAC) genes in astrocytomas and normal brain tissue: Class II and IV are hypoexpressed in glioblastomas. *BMC Cancer*, 8, aug 2008. doi:10.1186/1471-2407-8-243.

[165] D. Williams Parsons, Siân Jones, Xiaosong Zhang, Jimmy Cheng Ho Lin, Rebecca J. Leary, Philipp Angenendt, Parminder Mankoo, Hannah Carter, I. Mei Siu, Gary L. Gallia, Alessandro Olivi, Roger McLendon, B. Ahmed Rasheed, Stephen Keir, Tatiana Nikolskaya, Yuri Nikolsky, Dana A. Busam, Hanna Tekleab, Luis A. Diaz, James Hartigan, Doug R. Smith, Robert L. Strausberg, Suely Kazue Nagahashi Marie, Sueli Mieko Oba Shinjo, Hai Yan, Gregory J. Riggins, Darell D. Bigner, Rachel Karchin, Nick Papadopoulos, Giovanni Parmigiani, Bert Vogelstein, Victor E. Velculescu, and Kenneth W. Kinzler. An integrated genomic analysis of human glioblastoma multiforme. *Science*, 321(5897):1807–1812, sep 2008. doi:10.1126/science.1164382.

[166] Sith Sathornsumetee, David A. Reardon, Annick Desjardins, Jennifer A. Quinn, James J. Vredenburgh, and Jeremy N. Rich. Molecularly targeted therapy for malignant glioma, jul 2007. doi:10.1002/cncr.22741.

[167] Dong Yin, John M. Ong, Jinwei Hu, Julian C. Desmond, Norihiko Kawamata, Bindu M. Konda, Keith L. Black, and H. Phillip Koeffler. Suberoylanilide hydroxamic acid, a histone deacetylase inhibitor: Effects on gene expression and growth of glioma cells in vitro and in vivo. *Clinical Cancer Research*, 13(3):1045–1052, feb 2007. doi:10.1158/1078-0432.CCR-06-1261.

[168] C. Y. Gui, L. Ngo, W. S. Xu, V. M. Richon, and P. A. Marks. Histone deacetylase (HDAC) inhibitor activation of p21WAF1 involves changes in promoter-associated proteins, including HDAC1. *Proceedings of the National Academy of Sciences of the United States of America*, 101(5):1241–1246, feb 2004. doi:10.1073/pnas.0307708100.

[169] S. G. Gray, C. N. Qian, K. Furge, X. Guo, and B. T. Teh. Microarray profiling of the effects of histone deacetylase inhibitors on gene expression in cancer cell lines. *International journal of oncology*, 24(4):773–795, 2004. doi:10.3892/ijo.24.4.773.

[170] Melissa J. Peart, Gordon K. Smyth, Ryan K. Van Laar, David D. Bowtell, Victoria M. Richon, Paul A. Marks, Andrew J. Holloway, and Ricky W. Johnstone. Identification and functional significance of genes regulated by structurally different histone deacetylase inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 102(10):3697–3702, mar 2005. doi:10.1073/pnas.0500369102.

[171] Ilker Y. Eyüpoglu, Eric Hahnen, Rolf Buslei, Florian A. Siebzehnrübl, Nicolai E. Savaskan, Mike Lüders, Christian Tränkle, Wolfgang Wick, Michael Weller, Rudolf Fahlbusch, and Ingmar Blümcke. Suberoylanilide hydroxamic acid (SAHA) has potent anti-glioma properties in vitro, ex vivo and in vivo. *Journal of Neurochemistry*, 93(4):992–999, may 2005. doi:10.1111/j.1471-4159.2005.03098.x.

[172] Leigh Ellis, Yan Pan, Gordon K. Smyth, Daniel J. George, Chris McCormack, Roxanne Williams-Truax, Monica Mita, Joachim Beck, Howard Burris, Gail Ryan, Peter Atadja, Dale Butterfoss, Margaret Dugan, Kenneth Culver, Ricky W. Johnstone, and H. Miles Prince. Histone deacetylase inhibitor panobinostat induces clinical responses with associated alterations in gene expression profiles in cutaneous T-cell lymphoma. *Clinical Cancer Research*, 14(14):4500–4510, jul 2008. doi:10.1158/1078-0432.CCR-07-4262.

[173] Peter Gimsing, Mads Hansen, Lene M. Knudsen, P. Knoblauch, Ib Jarle Christensen, Chean Eng Ooi, and Peter Buhl-Jensen. A phase I clinical trial of the histone deacetylase inhibitor belinostat in patients with advanced hematological neoplasia. In *European Journal of Haematology*, volume 81, pages 170–176. Eur J Haematol, sep 2008. doi:10.1111/j.1600-0609.2008.01102.x.

[174] Ciências Da Saúde, Marco António, Paulo De Carvalho, and Ciências Biomédicas. UNIVERSIDADE DA BEIRA INTERIOR Modelos de cultura celular para rastreio de fármacos. Technical report, oct 2016.

[175] Agarose - an overview | ScienceDirect Topics. URL: https://www.sciencedirect.com/topics/chemistry/agarose.

[176] Elisabete C. Costa, Duarte de Melo-Diogo, André F. Moreira, Marco P. Carvalho, and Ilídio J. Correia. Spheroids Formation on Non-Adhesive Surfaces by Liquid Overlay Technique: Considerations and Practical Approaches, 2018. doi:10.1002/biot.201700417.

[177] Maria Vinci, Carol Box, and Suzanne A. Eccles. Three-dimensional (3D) tumor spheroid invasion assay. *Journal of Visualized Experiments*, 2015(99):52686, may 2015. doi:10.3791/52686.

[178] Mariam Eleni Oraiopoulou, Maria Tampakaki, Eleftheria Tzamali, Theodoros Tamiolakis, Venediktos Makatounakis, Antonios F. Vakis, Giannis Zacharakis, Vangelis Sakkalis, and

Joseph Papamatheakis. A 3D tumor spheroid model for the T98G Glioblastoma cell line phenotypic characterization. *Tissue and Cell*, 59:39–43, aug 2019. doi:10.1016/j.tice.2019.05.007.

[179] Roger C. Lo. Application of Microfluidics in Chemical Engineering, aug 2013.

[180] J. D. Watson and F. H.C. Crick. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–738, 1953. doi:10.1038/171737a0.

[181] Francis Crick. Central dogma of molecular biology. *Nature*, 227(5258):561–563, 1970. doi:10.1038/227561a0.

[182] Arthur Henrique, Fernandes Rodrigues, Carolina Pimentel Canales De Albuquerque, Camylla Duarte Cavalcante, Amanda Da, and Silva Peixoto. MECANISMOS EPIGENÉTICOS NO CÂNCER DE MAMA: O PAPEL DOS BIOMARCADORES E DA MEDICINA PERSONALIZADA EPIGENETIC MECHANISMS IN BREAST CANCER: THE ROLE OF BIOMARKERS AND PERSONALIZED MEDICINE. Technical report.

[183] Robin Holliday. Epigenetics: A historical overview, 2006. doi:10.4161/epi.1.2.2762.

[184] D. L. Nanney. EPIGENETIC CONTROL SYSTEMS. *Proceedings of the National Academy of Sciences*, 44(7):712–717, jul 1958. doi:10.1073/pnas.44.7.712.

[185] Roberto Bonasio, Shengjiang Tu, and Danny Reinberg. Molecular signals of epigenetic states, oct 2010. doi:10.1126/science.1191078.

[186] An Introduction to Systems Biology: Design Principles of Biological Circuits (Chapman & Hall/CRC Mathematical and Computational Biology): Alon, Uri: 8601200639971: Amazon.com: Books. URL: https://www.amazon.com/Introduction-Systems-Biology-Mathematical-Computational/dp/1584886420.

[187] Bradley R. Cairns. The logic of chromatin architecture and remodelling at promoters, sep 2009. doi:10.1038/nature08450.

[188] Paul B. Talbert and Steven Henikoff. Histone variants ancient wrap artists of the epigenome, apr 2010. doi:10.1038/nrm2861.

[189] Gaurav Arya, Arijit Maitra, and Sergei A. Grigoryev. *Journal of Biomolecular Structure and Dynamics*, (6):803–820. doi:10.1080/07391102.2010.10508585.

[190] Ryoiti Kiyama and Edward N. Trifonov. What positions nucleosomes?-A model, jul 2002. doi:10.1016/S0014-5793(02)02937-X.

[191] Eran Segal and Jonathan Widom. What controls nucleosome positions?, aug 2009. doi:10.1016/j.tig.2009.06.002.

[192] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello. Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. *Nature*, 391(6669):806–811, feb 1998. doi:10.1038/35888.

[193] Kevin V. Morris. RNA-directed transcriptional gene silencing and activation in human cells, dec 2009. doi:10.1089/oli.2009.0212.

[194] Anne Marie W. Turner and Kevin V. Morris. Controlling transcription with noncoding RNAs in mammalian cells, 2010. doi:10.2144/000113442.

[195] R Holliday and JE Pugh. DNA modification mechanisms and gene activity during development. *Science*, 187(4173):226–232, jan 1975. doi:10.1126/science.187.4173.226.

[196] Robert J. Klose and Adrian P. Bird. Genomic DNA methylation: The mark and its mediators, feb 2006. doi:10.1016/j.tibs.2005.12.008.

[197] Suhua Feng, Steven E. Jacobsen, and Wolf Reik. Epigenetic reprogramming in plant and animal development, oct 2010. doi:10.1126/science.1190614.

[198] Keith D. Robertson. DNA methylation and human disease, aug 2005. doi:10.1038/nrg1655.

[199] Ryan Lister, Mattia Pelizzola, Robert H. Dowen, R. David Hawkins, Gary Hon, Julian Tonti-Filippini, Joseph R. Nery, Leonard Lee, Zhen Ye, Que Minh Ngo, Lee Edsall, Jessica Antosiewicz-Bourget, Ron Stewart, Victor Ruotti, A. Harvey Millar, James A. Thomson, Bing Ren, and Joseph R. Ecker. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462(7271):315–322, nov 2009. doi:10.1038/nature08514.

[200] Jaqueline Carvalho de Oliveira. Epigenetics and human diseases Epigenética e doenças humanas. 33(1):21–34, 2012. doi:10.5433/1679-0367.2012v33n1p21.

[201] Bilian Jin, Bing Yao, Jian Liang Li, C. Robert Fields, Amber L. Delmas, Chen Liu, and Keith D. Robertson. DNMT1 and DNMT3B modulate distinct polycomb-mediated histone modifications in colon cancer. *Cancer Research*, 69(18):7412–7421, 2009. doi:10.1158/0008-5472.CAN-09-0116.

[202] Fernando Trigo Marisa Barbieri Roberto Sanchez Vinicius Anelli, Daniel Antunes. Mecanismos Epigenéticos. URL: http://ead.hemocentro.fmrp.usp.br/joomla/index.php/noticias/adotepauta/669-mecanismos-epigeneticos.

[203] Medical Epigenetics (Translational Epigenetics): Amazon.co.uk: Trygve Tollefsbol, Trygve Tollefsbol: 9780128032398: Books. URL: https://www.amazon.co.uk/Medical-Epigenetics-Translational-Trygve-Tollefsbol/dp/0128032391.

[204] Bilian Jin, Jason Ernst, Rochelle L. Tiedemann, Hongyan Xu, Suhas Sureshchandra, Manolis Kellis, Stephen Dalton, Chen Liu, Jeong Hyeon Choi, and Keith D. Robertson. Linking DNA Methyltransferases to Epigenetic Marks and Nucleosome Structure Genome-wide in Human Tumor Cells. *Cell Reports*, 2(5):1411–1424, nov 2012. doi:10.1016/j.celrep.2012.10.017.

[205] Galit Lev Maor, Ahuvi Yearim, and Gil Ast. The alternative role of DNA methylation in splicing regulation, may 2015. doi:10.1016/j.tig.2015.03.002.

[206] Maria Gutierrez-Arcelus, Halit Ongen, Tuuli Lappalainen, Stephen B. Montgomery, Alfonso Buil, Alisa Yurovsky, Julien Bryois, Ismael Padioleau, Luciana Romano, Alexandra Planchon, Emilie Falconnet, Deborah Bielser, Maryline Gagnebin, Thomas Giger, Christelle Borel, Audrey Letourneau, Periklis Makrythanasis, Michel Guipponi, Corinne

Gehrig, Stylianos E. Antonarakis, and Emmanouil T. Dermitzakis. Tissue-Specific Effects of Genetic and Epigenetic Variation on Gene Regulation and Splicing. *PLoS Genetics*, 11(1):1004958, 2015. doi:10.1371/journal.pgen.1004958.

[207] Natasha Zamudio, Joan Barau, Aurélie Teissandier, Marius Walter, Maté Borsos, Nicolas Servant, and Déborah Bourc'his. DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. *Genes and Development*, 29(12):1256–1270, jun 2015. doi:10.1101/gad.257840.114.

[208] Roadmap Epigenomics Consortium, Anshul Kundaje, Wouter Meuleman, Jason Ernst, Misha Bilenky, Angela Yen, Alireza Heravi-Moussavi, Pouya Kheradpour, Zhizhuo Zhang, Jianrong Wang, Michael J. Ziller, Viren Amin, John W. Whitaker, Matthew D. Schultz, Lucas D. Ward, Abhishek Sarkar, Gerald Quon, Richard S. Sandstrom, Matthew L. Eaton, Yi Chieh Wu, Andreas R. Pfenning, Xinchen Wang, Melina Claussnitzer, Yaping Liu, Cristian Coarfa, R. Alan Harris, Noam Shoresh, Charles B. Epstein, Elizabeta Gjoneska, Danny Leung, Wei Xie, R. David Hawkins, Ryan Lister, Chibo Hong, Philippe Gascard, Andrew J. Mungall, Richard Moore, Eric Chuah, Angela Tam, Theresa K. Canfield, R. Scott Hansen, Rajinder Kaul, Peter J. Sabo, Mukul S. Bansal, Annaick Carles, Jesse R. Dixon, Kai How Farh, Soheil Feizi, Rosa Karlic, Ah Ram Kim, Ashwinikumar Kulkarni, Daofeng Li, Rebecca Lowdon, Ginell Elliott, Tim R. Mercer, Shane J. Neph, Vitor Onuchic, Paz Polak, Nisha Rajagopal, Pradipta Ray, Richard C. Sallari, Kyle T. Siebenthall, Nicholas A. Sinnott-Armstrong, Michael Stevens, Robert E. Thurman, Jie Wu, Bo Zhang, Xin Zhou, Arthur E. Beaudet, Laurie A. Boyer, Philip L. De Jager, Peggy J. Farnham, Susan J. Fisher, David Haussler, Steven J.M. Jones, Wei Li, Marco A. Marra, Michael T. McManus, Shamil Sunyaev, James A. Thomson, Thea D. Tlsty, Li Huei Tsai, Wei Wang, Robert A. Waterland, Michael Q. Zhang, Lisa H. Chadwick, Bradley E. Bernstein, Joseph F. Costello, Joseph R. Ecker, Martin Hirst, Alexander Meissner, Aleksandar Milosavljevic, Bing Ren, John A. Stamatoyannopoulos, Ting Wang, and Manolis Kellis. Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539):317–329, feb 2015. doi:10.1038/nature14248.

[209] Pierre Antoine Defossez and Irina Stancheva. Biological functions of methyl-CpG-binding proteins. In *Progress in Molecular Biology and Translational Science*, volume 101, pages 377–398. Elsevier B.V., 2011. doi:10.1016/B978-0-12-387685-0.00012-3.

[210] Kristopher C. Hite, Valerie H. Adams, and Jeffrey C. Hansen. Recent advances in MeCP2 structure and function, feb 2009. doi:10.1139/O08-115.

[211] Keren Bahar Halpern, Tal Vana, and Michael D. Walker. Paradoxical role of DNA methylation in activation of FoxA2 gene expression during endoderm development. *Journal of Biological Chemistry*, 289(34):23882–23892, 2014. doi:10.1074/jbc.M114.573469.

[212] Xiaodong Cheng and Robert M. Blumenthal. Mammalian DNA Methyltransferases: A Structural Perspective, mar 2008. doi:10.1016/j.str.2008.01.004.

[213] Samir Kumar Patra, Aditi Patra, Federica Rizzi, Tapash Chandra Ghosh, and Saverio Bettuzzi. Demethylation of (Cytosine-5-C-methyl) DNA and regulation of transcription in the epigenetic pathways of cancer development, jun 2008. doi:10.1007/s10555-008-9118-y.

[214] F. Xu, C. Mao, Y. Ding, C. Rui, L. Wu, A. Shi, H. Zhang, L. Zhang, and Z. Xu. Molecular and Enzymatic Profiles of Mammalian DNA Methyltransferases: Structures and Targets

for Drugs. *Current Medicinal Chemistry*, 17(33):4052–4071, oct 2010. doi:10.2174/092986710793205372.

[215] Alan M. O'doherty, Lynne C. O'shea, and Trudee Fair. Bovine DNA methylation imprints are established in an oocyte size-specific manner, which are coordinated with the expression of the DNMT3 family proteins. *Biology of Reproduction*, 86(3), mar 2012. doi:10.1095/biolreprod.111.094946.

[216] D. Bourc'his, G. L. Xu, C. S. Lin, B. Bollman, and T. H. Bestor. Dnmt3L and the establishment of maternal genomic imprints. *Science*, 294(5551):2536–2539, dec 2001. doi:10.1126/science.1065848.

[217] Kurtis E. Bachman, Michael R. Rountree, and Stephen B. Baylin. Dnmt3a and Dnmt3b Are Transcriptional Repressors That Exhibit Unique Localization Properties to Heterochromatin. *Journal of Biological Chemistry*, 276(34):32282–32287, aug 2001. doi:10.1074/jbc.M104661200.

[218] M. Okano, S. Xie, and E. Li. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases [1], 1998. doi:10.1038/890.

[219] Wolfgang Mayer, Alain Niveleau, Jörn Walter, Reinald Fundele, and Thomas Haaf. Demethylation of the zygotic paternal genome. *Nature*, 403(6769):501–502, feb 2000. doi:10.1038/35000656.

[220] Sébastien A. Smallwood and Gavin Kelsey. De novo DNA methylation: A germ cell perspective, jan 2012. doi:10.1016/j.tig.2011.09.004.

[221] Khursheed Iqbal, Seung Gi Jin, Gerd P. Pfeifer, and Piroska E. Szabó. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proceedings of the National Academy of Sciences of the United States of America*, 108(9):3642–3647, mar 2011. doi:10.1073/pnas.1014033108.

[222] V. G. ALLFREY, R. FAULKNER, and A. E. MIRSKY. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE. *Proceedings of the National Academy of Sciences of the United States of*, 51(5):786–794, 1964. doi:10.1073/pnas.51.5.786.

[223] Brian D. Strahl and C. David Allis. The language of covalent histone modifications, jan 2000. doi:10.1038/47412.

[224] T. Jenuwein and C. D. Allis. Translating the histone code, aug 2001. doi:10.1126/science.1063127.

[225] Histone Modifications | What is Epigenetics? URL: https://www.whatisepigenetics.com/histone-modifications/.

[226] Eric I. Campos and Danny Reinberg. Histones: Annotating chromatin, dec 2009. doi:10.1146/annurev.genet.032608.103928.

[227] Eva Bártová, Jana Krejcí, Andrea Harničarová, Gabriela Galiová, and Stanislav Kozubek. Histone modifications and nuclear architecture: A review, aug 2008. doi:10.1369/jhc.2008.951251.

[228] Alexander J. Ruthenburg, Haitao Li, Dinshaw J. Patel, and C. David Allis. Multivalent engagement of chromatin modifications by linked binding modules, dec 2007. doi:10.1038/nrm2298.

[229] Kangling Zhang, Katherine E. Williams, Lan Huang, Peter Yau, Joseph S. Siino, E. Morton Bradbury, Patrick R. Jones, Michael J. Minch, and Alma L. Burlingame. Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. *Molecular & cellular proteomics : MCP*, 1(7):500–508, 2002. doi:10.1074/mcp.M200031-MCP200.

[230] Shwu Yuan Wu and Cheng Ming Chiang. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation, may 2007. doi:10.1074/jbc.R700001200.

[231] Saverio Minucci and Pier Giuseppe Pelicci. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer, jan 2006. doi:10.1038/nrc1779.

[232] W. S. Xu, R. B. Parmigiani, and P. A. Marks. Histone deacetylase inhibitors: Molecular mechanisms of action, aug 2007. doi:10.1038/sj.onc.1210620.

[233] Bernhard Payer and Jeannie T. Lee. X chromosome dosage compensation: How mammals keep the balance, 2008. doi:10.1146/annurev.genet.42.110807.091711.

[234] X Inactivation and Epigenetics | Walter and Eliza Hall Institute of Medical Research. URL: https://www.wehi.edu.au/wehi-tv/x-inactivation-and-epigenetics.

[235] G. B. Sharman. Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature*, 230(5291):231–232, 1971. doi:10.1038/230231a0.

[236] Nobuo Takagi and Motomichi Sasaki. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature*, 256(5519):640–642, 1975. doi:10.1038/256640a0.

[237] Fei Xue, X. Cindy Tian, Fuliang Du, Chikara Kubota, Maneesh Taneja, Andras Dinnyes, Yunping Dai, Howard Levine, Lygia V. Pereira, and Xiangzhong Yang. Aberrant patterns of X chromosome inactivation in bovine clones. *Nature Genetics*, 31(2):216–220, may 2002. doi:10.1038/ng900.

[238] Joana Carvalho Moreira de Mello, Érica Sara Souza de Araújo, Raquel Stabellini, Ana Maria Fraga, Jorge Estefano Santana de Souza, Denilce R. Sumita, Anamaria A. Camargo, and Lygia V. Pereira. Random X Inactivation and Extensive Mosaicism in Human Placenta Revealed by Analysis of Allele-Specific Gene Expression along the X Chromosome. *PLoS ONE*, 5(6):e10947, jun 2010. doi:10.1371/journal.pone.0010947.

[239] Jennifer A. Erwin and Jeannie T. Lee. New twists in X-chromosome inactivation, jun 2008. doi:10.1016/j.ceb.2008.04.007.

[240] Edith Heard and Christine M. Disteche. Dosage compensation in mammals: Fine-tuning the expression of the X chromosome, jul 2006. doi:10.1101/gad.1422906.

[241] Céline Morey and Philip Avner. The Demoiselle of X-Inactivation: 50 Years Old and As Trendy and Mesmerising As Ever. *PLoS Genetics*, 7(7):e1002212, jul 2011. doi:10.1371/journal.pgen.1002212.

[242] Anne-Valerie Gendrel and Edith Heard. Noncoding RNAs and Epigenetic Mechanisms During X-Chromosome Inactivation. *Annual Review of Cell and Developmental Biology*, 30(1):561–580, oct 2014. doi:10.1146/annurev-cellbio-101512-122415.

[243] Andrea Slavney, Leonardo Arbiza, Andrew G. Clark, and Alon Keinan. Strong constraint on human genes escaping X-inactivation is modulated by their expression level and breadth in both sexes. *Molecular Biology and Evolution*, 33(2):384–393, feb 2016. doi:10.1093/molbev/msv225.

[244] Erika L. Moen, Edward Litwin, Stephen Arnovitz, Xu Zhang, Wei Zhang, M. Eileen Dolan, and Lucy A. Godley. Characterization of CpG sites that escape methylation on the inactive human X-chromosome. *Epigenetics*, 10(9):810–818, 2015. doi:10.1080/15592294.2015.1069461.

[245] Asaf Hellman and Andrew Chess. Gene body-specific methylation on the active X chromosome. *Science*, 315(5815):1141–1143, feb 2007. doi:10.1126/science.1136352.

[246] Miho Ishida and Gudrun E. Moore. The role of imprinted genes in humans, jul 2013. doi:10.1016/j.mam.2012.06.009.

[247] Katherine S. Pollard, David Serre, Xu Wang, Heng Tao, Elin Grundberg, Thomas J. Hudson, Andrew G. Clark, and Kelly Frazer. A genome-wide approach to identifying novel-imprinted genes. *Human Genetics*, 122(6):625–634, jan 2008. doi:10.1007/s00439-007-0440-1.

[248] Philippe P. Luedi, Fred S. Dietrich, Jennifer R. Weidman, Jason M. Bosko, Randy L. Jirtle, and Alexander J. Hartemink. Computational and experimental identification of novel human imprinted genes. *Genome Research*, 17(12):1723–1730, dec 2007. doi:10.1101/gr.6584707.

[249] Rivka L. Glaser, Joshua P. Ramsay, and Ian M. Morison. The imprinted gene and parent-of-origin effect database now includes parental origin of de novo mutations. *Nucleic acids research*, 34(Database issue):D29, 2006. doi:10.1093/nar/gkj101.

[250] David Monk. Germline-derived DNA methylation and early embryo epigenetic reprogramming: The selected survival of imprints, oct 2015. doi:10.1016/j.biocel.2015.04.014.

[251] Peter A. Jones and Stephen B. Baylin. The fundamental role of epigenetic events in cancer, 2002. doi:10.1038/nrg816.

[252] Gerda Egger, Gangning Liang, Ana Aparicio, and Peter A. Jones. Epigenetics in human disease and prospects for epigenetic therapy, may 2004. doi:10.1038/nature02625.

[253] Jairo Rodriguez, Jordi Frigola, Elisenda Vendrell, Rosa Ana Risques, Mario F. Fraga, Cristina Morales, Victor Moreno, Manel Esteller, Gabriel Capellà, Maria Ribas, and Miguel A. Peinado. Chromosomal instability correlates with genome-wide DNA demethylation in human primary colorectal cancers. *Cancer Research*, 66(17):8462–8468, sep 2006. doi:10.1158/0008-5472.CAN-06-0293.

[254] Amir Eden, François Gaudet, Alpana Waghmare, and Rudolf Jaenisch. Chromosomal instability and tumors promoted by DNA hypomethylation, apr 2003. doi:10.1126/science.1083557.

[255] G. Howard, R. Eiges, F. Gaudet, R. Jaenisch, and A. Eden. Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene*, 27(3):404–408, jan 2008. doi:10.1038/sj.onc.1210631.

[256] Melanie Ehrlich. The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clinical Immunology*, 109(1):17–28, 2003. doi:10.1016/S1521-6616(03)00201-8.

[257] Ann S. Wilson, Barbara E. Power, and Peter L. Molloy. DNA hypomethylation and human diseases, jan 2007. doi:10.1016/j.bbcan.2006.08.007.

[258] Shirley Rainier, Laura A. Johnson, Craig J. Dobry, April J. Ping, Paul E. Grundy, and Andrew P. Feinberg. Relaxation of imprinted genes in human cancer. *Nature*, 362(6422):747–749, 1993. doi:10.1038/362747a0.

[259] Valerie Greger, Eberhard Passarge, Wolfgang Höpping, Elmar Messmer, and Bernhard Horsthemke. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Human Genetics*, 83(2):155–158, sep 1989. doi:10.1007/BF00286709.

[260] Stephen B. Baylin. DNA methylation and gene silencing in cancer, dec 2005. doi:10.1038/ncponc0354.

[261] Luciano Di Croce, Veronica A. Raker, Massimo Corsaro, Francesco Fazi, Mirco Fanelli, Mario Faretta, Francois Fuks, Francesco Lo Coco, Tony Kouzarides, Clara Nervi, Saverio Minucci, and Pier Giuseppe Pelicci. Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science*, 295(5557):1079–1082, feb 2002. doi:10.1126/science.1065173.

[262] Yeshayahu Schlesinger, Ravid Straussman, Ilana Keshet, Shlomit Farkash, Merav Hecht, Joseph Zimmerman, Eran Eden, Zohar Yakhini, Etti Ben-Shushan, Benjamin E. Reubinoff, Yehudit Bergman, Itamar Simon, and Howard Cedar. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nature Genetics*, 39(2):232–236, feb 2007. doi:10.1038/ng1950.

[263] Mario F. Fraga, Esteban Ballestar, Ana Villar-Garea, Manuel Boix-Chornet, Jesus Espada, Gunnar Schotta, Tiziana Bonaldi, Claire Haydon, Santiago Ropero, Kevin Petrie, N. Gopalakrishna Iyer, Alberto Pérez-Rosado, Enrique Calvo, Juan A. Lopez, Amparo Cano, Maria J. Calasanz, Dolors Colomer, Miguel Ángel Piris, Natalie Ahn, Axel Imhof, Carlos Caldas, Thomas Jenuwein, and Manel Esteller. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature Genetics*, 37(4):391–400, apr 2005. doi:10.1038/ng1531.

[264] Xiang Jao Yang. The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases, 2004. doi:10.1093/nar/gkh252.

[265] Mihaela Velicescu Felicidad A Gonzales Joy C Y Lin Gangning Liang Peter A Jones Carvell T Nguyen 1, Daniel J Weisenberger. Histone H3-lysine 9 methylation is associated with aberrant gene silencing in cancer cells and is rapidly reversed by 5-aza-2'-deoxycytidine - PubMed, nov 2002.

[266] Brock C. Reeve. Stem cells and cancer, 2008. doi:10.1016/j.cell.2004.08.005.

[267] Paul A.C. Cloos, Jesper Christensen, Karl Agger, and Kristian Helin. Erasing the methyl mark: Histone demethylases at the center of cellular differentiation and disease, may 2008. doi:10.1101/gad.1652908.

[268] Yujiang Shi, Fei Lan, Caitlin Matson, Peter Mulligan, Johnathan R. Whetstine, Philip A. Cole, Robert A. Casero, and Yang Shi. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, 119(7):941–953, dec 2004. doi:10.1016/j.cell.2004.12.012.

[269] Shikhar Sharma, Theresa K. Kelly, and Peter A. Jones. Epigenetics in cancer, sep 2009. doi:10.1093/carcin/bgp220.

[270] José Luis García-Giménez. *Epigenetic Biomarkers and Diagnostics*. Elsevier Inc., jan 2015. doi:10.1016/c2014-0-01268-3.

[271] Moshe Szyf. Epigenetics, DNA methylation, and chromatin modifying drugs, 2009. doi:10.1146/annurev-pharmtox-061008-103102.

[272] David E. Amacher. A 2015 survey of established or potential epigenetic biomarkers for the accurate detection of human cancers, jul 2016. doi:10.3109/1354750X.2016.1153724.

[273] Thomas F. Imperiale, David F. Ransohoff, Steven H. Itzkowitz, Theodore R. Levin, Philip Lavin, Graham P. Lidgard, David A. Ahlquist, and Barry M. Berger. Multitarget stool DNA testing for colorectal-cancer screening. *New England Journal of Medicine*, 370(14):1287–1297, 2014. doi:10.1056/NEJMoa1311194.

[274] Uri Ladabaum and Ajitha Mannalithara. Comparative Effectiveness and Cost Effectiveness of a Multitarget Stool DNA Test to Screen for Colorectal Neoplasia. *Gastroenterology*, 151(3):427–439.e6, sep 2016. doi:10.1053/j.gastro.2016.06.003.

[275] Wen Hsiang Lee, Ronald A. Morton, Jonathan I. Epstein, James D. Brooks, Pearl A. Campbell, G. Steven Bova, Wen Son Hsieh, William B. Isaacs, and William G. Nelson. Cytidine methylation of regulatory sequences near the π-class glutathione S-transferase gene accompanies human prostatic carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 91(24):11733–11737, nov 1994. doi:10.1073/pnas.91.24.11733.

[276] Alan W. Partin, Leander Van Neste, Eric A. Klein, Leonard S. Marks, Jason R. Gee, Dean A. Troyer, Kimberly Rieger-Christ, J. Stephen Jones, Cristina Magi-Galluzzi, Leslie A. Mangold, Bruce J. Trock, Raymond S. Lance, Joseph W. Bigley, Wim Van Criekinge, and Jonathan I. Epstein. Clinical validation of an epigenetic assay to predict negative histopathological results in repeat prostate biopsies. *Journal of Urology*, 192(4):1081–1087, oct 2014. doi:10.1016/j.juro.2014.04.013.

[277] Mohamad M. Kronfol, Mikhail G. Dozmorov, Rong Huang, Patricia W. Slattum, and Joseph L. McClay. The role of epigenomics in personalized medicine, 2017. doi:10.1080/23808993.2017.1284557.

[278] Kim E.M. van Kessel, Leander Van Neste, Irene Lurkin, Ellen C. Zwarthoff, and Wim Van Criekinge. Evaluation of an Epigenetic Profile for the Detection of Bladder Cancer in Patients with hematuria. *Journal of Urology*, 195(3):601–607, 2016. doi:10.1016/j.juro.2015.08.085.

[279] Olafur A. Stefansson and Manel Esteller. Epigenetic modifications in breast cancer and their role in personalized medicine, oct 2013. doi:10.1016/j.ajpath.2013.04.033.

[280] Shahzia Anjum, Evangelia Ourania Fourkala, Michal Zikan, Andrew Wong, Aleksandra Gentry-Maharaj, Allison Jones, Rebecca Hardy, David Cibula, Diana Kuh, Ian J. Jacobs, Andrew E. Teschendorff, Usha Menon, and Martin Widschwendter. A BRCA1-mutation associated DNA methylation signature in blood cells predicts sporadic breast cancer incidence and survival. *Genome Medicine*, 6(6):47, jun 2014. doi:10.1186/gm567.

[281] Katie Lunnon, Rebecca Smith, Eilis Hannon, Philip L. De Jager, Gyan Srivastava, Manuela Volta, Claire Troakes, Safa Al-Sarraj, Joe Burrage, Ruby Macdonald, Daniel Condliffe, Lorna W. Harries, Pavel Katsel, Vahram Haroutunian, Zachary Kaminsky, Catharine Joachim, John Powell, Simon Lovestone, David A. Bennett, Leonard C. Schalkwyk, and Jonathan Mill. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nature Neuroscience*, 17(9):1164–1170, sep 2014. doi:10.1038/nn.3782.

[282] Ahmad Jowaed, Ina Schmitt, Oliver Kaut, and Ullrich Wüllner. Methylation regulates alpha-synuclein expression and is decreased in Parkinson's disease patients' brains. *Journal of Neuroscience*, 30(18):6355–6359, may 2010. doi:10.1523/JNEUROSCI.6119-09.2010.

[283] Devin M. Absher, Xinrui Li, Lindsay L. Waite, Andrew Gibson, Kevin Roberts, Jeffrey Edberg, W. Winn Chatham, and Robert P. Kimberly. Genome-Wide DNA Methylation Analysis of Systemic Lupus Erythematosus Reveals Persistent Hypomethylation of Interferon Genes and Compositional Changes to CD4+ T-cell Populations. *PLoS Genetics*, 9(8), aug 2013. doi:10.1371/journal.pgen.1003678.

[284] Karolina A. Aberg, Joseph L. McClay, Srilaxmi Nerella, Shaunna Clark, Gaurav Kumar, Wenan Chen, Amit N. Khachane, Linying Xie, Alexandra Hudson, Guimin Gao, Aki Harada, Christina M. Hultman, Patrick F. Sullivan, Patrik K.E. Magnusson, and Edwin J.C.G. Van Den Oord. Methylome-wide association study of schizophrenia: Identifying blood biomarker signatures of environmental insults. *JAMA Psychiatry*, 71(3):255–264, 2014. doi:10.1001/jamapsychiatry.2013.3730.

[285] C. Ladd-Acosta, K. D. Hansen, E. Briem, M. D. Fallin, W. E. Kaufmann, and A. P. Feinberg. Common DNA methylation alterations in multiple brain regions in autism. *Molecular Psychiatry*, 19(8):862–871, 2014. doi:10.1038/mp.2013.114.

[286] Home - GEO - NCBI. URL: https://www.ncbi.nlm.nih.gov/geo/.

[287] Timothy J. Triche, Daniel J. Weisenberger, David Van Den Berg, Peter W. Laird, and Kimberly D. Siegmund. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Research*, 41(7):e90, apr 2013. doi:10.1093/nar/gkt090.

[288] Trevor Hastie, Robert Tibshirani, Balasubramanian Narasimhan, and C Gilbert. impute: Imputation for microarray data. *Bioinformatics*, 17(6):520–525, 2001. doi:10.18129/B9.bioc.impute.

[289] Tiago C. Silva, Simon G. Coetzee, Nicole Gull, Lijing Yao, Dennis J. Hazelett, Houtan Noushmehr, De Chen Lin, Benjamin P. Berman, and Oliver Stegle. ELmer v.2: An r/bioconductor package to reconstruct gene regulatory networks from DNA methylation

and transcriptome profiles. *Bioinformatics*, 35(11):1974–1977, 2019. doi:10.1093/bioinformatics/bty902.

[290] Shumei Zhang, Yihan Wang, Yue Gu, Jiang Zhu, Ce Ci, Zhongfu Guo, Chuangeng Chen, Yanjun Wei, Wenhua Lv, Hongbo Liu, Dongwei Zhang, and Yan Zhang. Specific breast cancer prognosis-subtype distinctions based on DNA methylation patterns. *Molecular Oncology*, 12(7):1047–1060, jun 2018. doi:10.1002/1878-0261.12309.

[291] Michael I. Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), dec 2014. doi:10.1186/s13059-014-0550-8.

[292] HTSeq-Counts - GDC Docs. URL: https://docs.gdc.cancer.gov/Encyclopedia/pages/HTSeq-Counts/.

[293] Simon Anders, Paul Theodor Pyl, and Wolfgang Huber. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2):166–169, jan 2015. doi:10.1093/bioinformatics/btu638.

[294] Mark D. Robinson, Davis J. McCarthy, and Gordon K. Smyth. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140, nov 2009. doi:10.1093/bioinformatics/btp616.

[295] Mark D. Robinson and Gordon K. Smyth. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics*, 23(21):2881–2887, nov 2007. doi:10.1093/bioinformatics/btm453.

[296] Davis J. McCarthy, Yunshun Chen, and Gordon K. Smyth. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, 40(10):4288–4297, may 2012. doi:10.1093/nar/gks042.

[297] Simon Anders and Wolfgang Huber. Differential expression analysis for sequence count data. *Genome Biology*, 11(10):R106, oct 2010. doi:10.1186/gb-2010-11-10-r106.

[298] Judith M. Boer, Wolfgang K. Huber, Holger Sültmann, Friederike Wilmer, Anja Von Heydebreck, Stefan Haas, Bernhard Korn, Bastian Gunawan, Andreas Vente, Laszlo Füzesi, Martin Vingron, and Annemarie Poustka. Identification and classification of differentially expressed genes in renal cell carcinoma by expression profiling on a global human 31,500-element cDNA array. *Genome Research*, 11(11):1861–1870, 2001. doi:10.1101/gr.184501.

[299] Anqi Zhu, Joseph G. Ibrahim, and Michael I. Love. Heavy-Tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics*, 35(12):2084–2092, jun 2019. doi:10.1093/bioinformatics/bty895.

[300] Kevin Blighe. EnhancedVolcano. *Bioconductor*, pages 1–8, 2018. doi:10.18129/b9.bioc.enhancedvolcano.

[301] Zuguang Gu. Bioconductor - ComplexHeatmap, 2016. URL: https://www.bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html.

[302] Cis-Regulatory Module - an overview | ScienceDirect Topics. URL: https://www.sciencedirect.com/topics/neuroscience/cis-regulatory-module.

[303] Frank Wilcoxon. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6):80, dec 1945. doi:10.2307/3001968.

[304] Dongmei Li, Zidian Xie, Marc Le Pape, and Timothy Dye. An evaluation of statistical methods for DNA methylation microarray data analysis. *BMC Bioinformatics*, 16(1):217, jul 2015. doi:10.1186/s12859-015-0641-x.

[305] Hochberg Y. Benjamini Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing on JSTOR, 1995. URL: https://www.jstor.org/stable/2346101.

[306] John A Rice. Mathematical Statistics and Data Analysis. Technical report, 2007. URL: http://www.thomsonrights.com.

[307] Jim Frost. How t-Tests Work: t-Values, t-Distributions, and Probabilities - Statistics By Jim. URL: https://statisticsbyjim.com/hypothesis-testing/t-tests-t-values-t-distributions-probabilities/.

[308] Ana Maria Cortesão, Pais Figueira, Silva Abreu, Alexandra Isabel, Monteiro Borges, and Dissertação De Mestrado. Análise de Sobrevivência com o R. Technical report.

[309] E. L. Kaplan and Paul Meier. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, 53(282):457, jun 1958. doi:10.2307/2281868.

[310] Survival Analysis with R. URL: https://bioconnector.github.io/workshops/r-survival.html{#}other{_}tcga{_}resources.

[311] Maintainer Terry and M Therneau. Package 'survival' Title Survival Analysis. Technical report, 2020.

[312] Drawing Survival Curves using 'ggplot2' [R package survminer version 0.4.8]. URL: https://cran.r-project.org/package=survminer.

[313] Edward R. Laws, Ian F. Parney, Wei Huang, Fred Anderson, Angel M. Morris, Anthony Asher, Kevin O. Lillehei, Mark Bernstein, Henry Brem, Andrew Sloan, Mitchel S. Berger, and Susan Chang. Survival following surgery and prognostic factors for recently diagnosed malignant glioma: Data from the glioma outcomes project. *Journal of Neurosurgery*, 99(3):467–473, sep 2003. doi:10.3171/jns.2003.99.3.0467.

[314] Farina Hanif, Kanza Muzaffar, Kahkashan Perveen, Saima M. Malhi, and Shabana U. Simjee. Glioblastoma multiforme: A review of its epidemiology and pathogenesis through clinical presentation and treatment, 2017. doi:10.22034/APJCP.2017.18.1.3.

[315] Wei Yang, Nicole M. Warrington, Sara J. Taylor, Paula Whitmire, Eduardo Carrasco, Kyle W. Singleton, Ningying Wu, Justin D. Lathia, Michael E. Berens, Albert H. Kim, Jill S. Barnholtz-Sloan, Kristin R. Swanson, Jingqin Luo, and Joshua B. Rubin. Sex differences in GBM revealed by analysis of patient imaging, transcriptome, and survival data. *Science Translational Medicine*, (473), jan. URL: /pmc/articles/PMC6502224/?report=abstractvolume={11},year={2019}, doi:10.1126/scitranslmed.aao5253.

[316] ArtCAM / Autodesk. URL: https://www.autodesk.com/products/artcam/overview.