ADRIÀ-JAUME ROURA CANALDA

A multi-omics evaluation of somatic mutations, transcriptomic dysregulation, chromatin accessibility and remodeling in High-Grade Gliomas

PhD thesis
Completed at the Laboratory of Molecular Neurobiology of the Nencki Institute of Experimental Biology
Polish Academy of Sciences

SUPERVISOR:
Dr hab. Bartosz Wojtas

SUPERVISOR:
Prof. dr hab. Bozena Kaminska-Kaczmarek

NARODOWE CENTRUM NAUKI

European Funds Smart Growth

FNP Foundation for Polish Science

European Union European Regional Development Fund

Warsaw, Poland 2021

Author's declaration

I, the undersigned(s) *Adria-Jaume Roura Canalda* consent to the storage and release of my dissertation, titled: *"A multi-omics evaluation of somatic mutations, transcriptomic dysregulation, chromatin accessibility and remodeling in High-Grade Gliomas"* by the M. Nencki Institute of Experimental Biology PAS Library in a printed form, in a reading room and as a part of interlibrary loans, on the basis of free use.

At the same time, I grant the M. Nencki Institute of Experimental Biology PAS a free non-exclusive license to use the above work without time and territorial limitation in the following fields of exploitation:

1) placing the content of the thesis as a pdf file together with metadata in the digital repository RCIN (*Digital Repository of Scientific Institutes,* collection: *Institute of Experimental Biology PAS/Dissertations*) located at the following address https://rcin.org.pl/dlibra/collectiondescription/121
2) digital multiplication of the work (digitalisation of the work if it is necessary to scan the printed version)

Warsaw, Poland 23/11/21 Signature

This study was supported by the Foundation for Polish Science TEAM-TECH Core Facility project: “NGS platform for comprehensive diagnostics and personalized therapy in neuro-oncology” and by Polish National Science Centre grant [DEC-2015/16/W/NZ2/00314].

Acknowledgment

First and foremost, I want to express my deepest gratitude to Prof. Bożena Kamińska-Kaczmarek for allowing me to join the laboratory in 2017, where I have grown as a research scientist and learned a lot in this fascinating field. Your insightful feedback and assistance in developing the research questions and methodology pushed me to improve my thinking and raise the quality of my work.
I would like to extend my sincere thanks to Dr Bartosz Wojtaś. His immense assistance, support, guidance, dedication and involvement in every step throughout the process, as well as his enthusiasm for science, have made this research a fantastic adventure.

Aside from my supervisors, I'd like to thank my lab mates for the stimulating discussions and assistance I've received over the years; their knowledge and expertise have greatly aided my research. I would like to recognize the assistance and effort that I received from different people who contributed to the research presented here; they played decisive roles and the completion of my dissertation would not have been possible otherwise: Dr Bartomiej Gielniewski for performing library preparation and sequencing on all of the samples; Paulina Pilanc for her excellent contribution with immunostaining on glioblastoma slices; Paulina Szadkowska, Dr Katarzyna Poleszak, and Kamil Wojnicki for their invaluable contributions to the EMSA experiments and their optimization; Dr Karolina Stępniak for performing ATAC-seq and ChIP-seq on glioma specimens as well as the crucial description of glioma enhancers; Dr Chinchu Jayaprakash for carrying out all SMARCA-related wet lab experiments; Dr Michał J. Dąbrowski for his guidance, enthusiasm and huge contribution to the study of DNA methylation in gliomas.

My special thanks go to Dr Aleksandra Ellert and Dr María Banqueri for their significant contributions, guidance, discussion and support during the development of this thesis. Many thanks go to Salwador Cyranowski for insightful discussions and support and to Beata Kaza for assistance in many aspects of laboratory work.

In addition, none of this could have happened without my family. To my mother, who inculcated in me the importance of hard work, respect and humbleness. Regardless of the path I would take, you were always supportive. To my aunt *Pepi*, a great pillar of support in my life, always gratefully for your love. To my siblings, nephews and nieces for their unconditional love and encouragement; you are always there for me. To the family of *Lloret*, who is always present. To my love, her support over the years has been essential. Always willing to contribute and help; sharing this part of the journey with you has been incredible.

Finally, I would like to thank my friends, especially the group from *Blanes* and the group from *Girona/Barcelona Titus Tecs*; they are an indispensable part of my life.

*"**Tot està per fer i tot és posible"** ~* Miquel Martí i Pol, Catalan poet (1929-2003)

Table of contents

List of figures

3.18. Transcriptomic differences between high- and low-grade gliomas

3.19. Gene set enrichment analysis

3.20. Pathway enrichment analysis

3.21. *JUN* expression in pan-cancer and across glioma grades

3.22. Chromatin accessibility profiling and TFBS prediction in cis-regulatory elements of over-expressed genes in gliomas of a given grade

3.23. Supervised hierarchical clustering of WHO grade II and IV gliomas samples based on selected genes

3.24. c-Jun binding prediction in cis-regulatory elements of targeted genes

3.25. Deregulation of expression of selected genes coding for TFs across gliomas of various grades

3.26. Deregulation of TF expression in gliomas of different grades and in cultured glioma cells

3.27. Putative targets of GIV glioma-specific transcription factors

3.28. Correlation between *JUN* expression and target gene expression at the mRNA level

3.29. Correlation between c-Jun phosphorylation and target gene expression at the mRNA level

3.30. Expression of c-Jun targets and GBM patient survival (TCGA)

3.31. Expression of c-Jun targets and LGG patient survival (TCGA)

3.32. Glioma enhancers and their intersection with TFBS

3.33. TFBS prediction in glioma enhancers in the context of c-Jun binding sites

3.34. Long-range distance c-Jun targets in TCGA

3.35. Median DNA methylation in the promoters of c-Jun targets

3.36. DNA methylation in the promoter of c-Jun targets and *JUN* coding gene

3.37. Methylation of DNA in C-rich regions of c-Jun targeted promoters

3.38. Cross-validation of DNA methylation and gene expression interrelationship for the predicted c-Jun targets in TCGA

3.39. EMSA and EMSA supershift experiments confirm c-Jun binding to the *VIM* promoter

3.40. *SMARCA2* and *SMARCA4* expression in glioma samples and gene silencing in LN18 glioma cells

3.41. Western blots detecting SMARCA2 and SMARCA4 protein levels in LN18 glioma cells

3.42. The effect of SMARCA2 and/or SMARCA4 knockdown on cell proliferation and cell viability

3.43. Landscape of chromatin accessibility in control and SMARCA2/4 depleted cells

3.44. Peak distribution and differences in ATAC-seq peaks in SMARCA depleted cells

3.45. Biological implications of SMARCA4/2 silencing in glioma cells

3.46. Gene Ontology analysis and genes involved in specific pathways in SMARCA4/2 depleted cells

List of tables

Abbreviations

2-HG 2-hydroxyglutarate
5-ALA 5-aminolevulinic acid
5mC 5-methylcytosine
AP-1 Activator protein-1
ATAC-seq Assay for Transposase Accessible Chromatin with high-throughput sequencing
ATCC American Type Culture Collection
BH Benjamini-Hochberg
bHLH Basic helix-loop-helix
BM Bone marrow
BP Biological processes
bp Base pair
BRD Bromodomain
bZIP Basic leucine zipper
CBS Circular Binary Segmentation
CCL2 C-C motif ligand 2
ChIP-seq Chromatin immunoprecipitation followed by sequencing
CNA Copy number aberration
CNS Central nervous system
CRC Chromatin remodeling complex
CSC Cancer stem cells
CX3CL1 C-X3-C motif ligand 1
DA Diffuse astrocytoma
DC-SIGN/CD209 Dendritic cell-specific ICAM-3-Grabbing non-integrin 1
DEG Differentially expressed gene
DNA-seq DNA-sequencing
DMEM Dulbecco's Modified Eagle Medium
ECM Extracellular matrix
EMSA Electrophoretic mobility shift assay
ETS Erythroblast transformation specific
FC Fold change
FDR False discovery rate
FPKM Fragments Per Kilobase of transcript per Million mapped reads
GII Grade II, Grade 2, World Health Organization grade II glioma
GIII Grade III, Grade 3, World Health Organization grade III glioma

GIV Grade IV, Grade 4, World Health Organization grade IV glioma

G-CIMP Glioma CpG island methylator phenotype

GAM Glioma-associated microglia and macrophages

GBM Glioblastoma

GEPIA Gene Expression Profiling Interactive Analysis

GSC Glioma stem cell

GSEA Gene Set Enrichment Analysis

GTEx Genotype-Tissue Expression

HAT Histone acetyltransferases

HDAC Histone deacetylases

HGG High-grade glioma

iDC Immature dendritic cells

IDH Isocitrate dehydrogenase

IF Immunofluorescence

IHC Immunohistochemistry

KDM4 Lysine demethylase 4A

KDM5 Lysine demethylase 5B

KEGG Kyoto Encyclopaedia of Genes and Genomes

LGG Low-Grade Glioma

MAPK Mitogen-activated protein kinase

MEP Megakaryocyte–erythroid progenitor

MGMT O6-methylguanine DNA methyltransferase

MMP14 Matrix metalloproteinase-14

MMP2 Matrix metalloproteinase-2

MRI Magnetic resonance imaging

MSC Mesenchymal stem cell

NFAT Nuclear factor of activated T-cells

NHA Normal human astrocytes

OS Overall survival

PA Pilocytic astrocytoma

PBZ Peripheral brain zone

PcG Polycomb group

PI Phosphatidylinositol

POU Pit-Oct-Unc

PRC1 Polycomb Repressive Complex 1

PRC2 Polycomb Repressive Complex 2

PWM Position weight matrix

RANO Response Assessment in Neuro-Oncology
RNA-seq RNA-sequencing
RT Room temperature
RTK Receptor tyrosine kinase
SIFT Sorting Intolerant From Tolerant
SNP Single nucleotide polymorphism
SRCR/CD163 Scavenger receptor cysteine-rich
SWI/SFN SWItch/Sucrose Non-Fermentable
TAD Topologically associated domain
TCGA The Cancer Genome Atlas
TET Ten-eleven translocation
TF Transcription factor
TFBS Transcription actor binding site
Thc T helper cells
THYM Thymoma
TLR Toll-like receptors
TLR2 Toll-like receptors 2
TME Tumour microenvironment
TMZ Temozolomide
Tn5 Hyperactive DNA transposase
TPM Transcripts per million
TSS Transcription start sites
VAF Variant allele frequency
WG12 Primary glioblastoma cell line (GII)
WHO World Health Organisation
α-KG α-ketoglutarate

Abstract

High-grade gliomas (HGGs), the most frequent and severe primary brain tumours in adults, invariably recur due to incomplete surgery or therapeutic resistance. Therapy resistance, rapid recurrence and poor clinical outcome are all linked to the intra-tumoral genetic and cellular heterogeneity in HGGs. The major checkpoint in regulation of gene expression is the initiation of transcription, which is mostly regulated by a class of DNA-binding proteins known as transcription factors (TFs). TFs are involved in a variety of human disorders, including cancer, and can bind to -specific DNA regulatory sequences called DNA motifs to activate or inhibit transcription, ultimately influencing mRNA levels. To do so, TFs rely on two unique interaction surfaces: a sequence-specific DNA-binding domain and an activation/repression domain that interacts with a variety of cofactors. Because TFs generally bind to DNA in nucleosome-depleted regions, their interactions are impacted by chromatin environment and chromatin remodelers. The expression of essential TFs is required by cancer cells to carry on a variety of biological processes in cancer cells such as cellular transformation, oncogenesis and progression, cell proliferation, metastasis, and chemo-resistance. TFs activities in cancer may be influenced by a number of direct and indirect processes, including gene amplifications, point mutations, and changes in their expression levels, as well as DNA methylation and histone modifications. Epigenetic modifications such as methylation of cytosine at CpG in DNA, in particular, regulate gene expression and can directly suppress transcription by blocking the binding of specific TFs to their recognition sites. **In HGGs, several interconnected biological components such as somatic mutations, transcriptomic and TF dysregulations, as well as alterations in histone modifications, DNA methylation and chromatin remodeling contribute to the disease aggressiveness. Transcriptomic profiles of HGGs at recurrence have not been thoroughly investigated yet. Moreover, despite significant efforts, the specific regulation of genes overexpressed in HGGs by TFs remains largely unknown. A better understanding of events occurring in open chromatin regions in HGGs is crucial to comprehend routes of brain cancer progression.**

In this thesis, we employed targeted sequencing of cancer-related genes (DNA-sequencing) and transcriptomics (RNA-sequencing) to identify single nucleotide variants, small insertions and deletions, copy number aberrations (CNAs), gene expression alterations and pathway dysregulations in 16 matched pairs of primary and recurrent HGGs. The majority of somatic mutations found in primary HGGs were not found in relapsed tumours, implying a sub-clone substitution during tumour progression. A novel frame-shift insertion in the *ZNF384* gene was discovered, which may play a role in extracellular matrix remodeling. The presence of focal CNAs in the *EGFR* and *PTEN* genes was found to be inversely correlated. Transcriptomic analysis revealed that genes involved in mRNA splicing, cell cycle, and DNA repair are down-regulated, while genes involved in interferon signalling

and phosphatidylinositol (PI) metabolism are up-regulated in recurrent HGGs when compared to primary HGGs. *In silico* analysis of the tumour microenvironment demonstrated that tumour supportive (M2) macrophages and immature dendritic cells are enriched in recurrent HGGs indicating a prominent immunosuppressive signature in those tumours. Immunohistochemistry staining of tumour sections confirmed the accumulation of immunosuppressive cells in recurrent HGGs.

We identified glioma grade-specific TFs binding sites in glioblastoma tissues as well as in human LN18 and LN229 glioma cells, using chromatin accessibility (ATAC-seq) data and confirmed their roles in controlling gene regulatory networks in HGGs. We explored different datasets that comprise DNA methylation profiles (targeted bisulfite sequencing), histone acetylation (H3K27ac) profiles, glioblastoma cell transcriptomics profiles (RNA-seq) and TCGA (the Cancer Genome Atlas) datasets (RNA-seq and lllumina 450K array DNA methylation). The comparative analyses of those profiles in gliomas of different malignancy grades revealed the importance of the c-Jun TF for the disease progression. c-Jun may play a role in the regulation of genes overexpressed in glioblastoma such as *VIM, FOSL2, UPP1, TRIB1* or *GPR3* by binding to the gene promoters. Furthermore, we found that in the majority of c-Jun gene targets, DNA methylation plays an important role in the c-Jun dependent regulation. We also found a significant positive correlation between c-Jun mRNA/protein expression and target gene expression in TCGA datasets, indicating that c-Jun likely regulates the expression of a number of invasion-related genes in glioblastomas. The bioinformatic predictions have been validated experimentally by testing c-Jun binding to various probes in the electrophoretic mobility shift assay (EMSA).

Chromatin remodeling proteins SMARCA2 and SMARCA4 are frequently mutated in high-grade gliomas. To determine the role of those proteins, we performed knockdown of genes coding them in human LN18 glioma cells and tested the impact of *SMARCA2* and *SMARCA4* deficiencies on chromatin accessibility using the Assay for Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq). We discovered an increase in chromatin openness in SMARCA2/4 deficient cells, which affected expression of genes critical for signal transduction, including those from the transforming growth factor beta pathway: *SMAD1*, *SMAD3*, *BMPR1A*, and *TGFBR2*, implying the interdependence of chromatin remodelers and specific signalling pathways.

Overall, this PhD dissertation provides novel insights pointing to significant transcriptomic deregulation in glioma cells. It reveals both the clonal evolution and the changed immune microenvironment of recurrent HGGs, which could have important implications when considering frontline immunotherapies in glioblastoma (GBM) is considered. The presented comprehensive identification of key TFs driving tumorigenesis in HGGs may pave the way to the potential future strategies in the treatment of malignant gliomas.

Streszczenie

Glejaki o wysokim stopniu złośliwości (według Światowej Organizacji Zdrowia High Grade Gliomas, HGGs) są najczęstszymi, pierwotnymi guzami mózgu u dorosłych. Z powodu niekompletnego usunięcia komórek nowotworowych (rozprzestrzeniających się w parenchymie mózgu) lub oporności na radio- i chemioterapię guzy odrastają w kilka miesięcy od diagnozy. Oporność na leczenie i szybka wznowa nowotworu są związane z genetyczną i komórkową różnorodnością tych guzów. Poznanie zmian genomicznych i transkrypcyjnych specyficznych dla HGGs lub dla określonego stadium progresji guza może poszerzyć wiedzę o patogenezie tych guzów i ujawnić nowe cele terapeutyczne. Inicjacja transkrypcji jest kluczowym etapem w regulacji ekspresji genów i jest zależna od białek wiążących DNA zwanych czynnikami transkrypcyjnymi (Transcription factors, TFs). Czynniki transkrypcyjne mogą wiązać się ze specyficznymi sekwencjami regulatorowymi w DNA zwanymi motywami DNA, aby aktywować lub hamować transkrypcję, ostatecznie wpływając na poziom mRNA. Czynniki transkrypcyjne mają domeny rozpoznające i wiążące specyficzne sekwencje w DNA oraz domenę aktywacji/represji, które oddziałują z podstawowymi białkami kompleksu transkrypcyjnego i polimerazy RNA II. Wiązanie TFs z DNA występuje w regionach otwartej chromatyny, której struktura jest regulowana przez białka kształtujące przestrzenne formowanie chromatyny (chromatin remodelers). Zmieniona ekspresja lub aktywność czynników transkrypcyjnych i deregulacja transkrypcji prowadzi do transformacji nowotworowej, skutkując zwiększoną proliferacją komórek, nabywaniem odporności komórek nowotworowych na chemioterapię, ich wzmożoną migracją i inwazyjnością oraz w konsekwencji progresją guza powiązaną ze zmianami mikrośrodowiska i unieczynnieniem odpowiedzi przeciwnowotworowej układu odpornościowego. Na aktywność czynników transkrypcyjnych w nowotworach mogą wpływać różne mechanizmy: amplifikacja genów, mutacje punktowe i zmiany w ekspresji, a także metylacja DNA i modyfikacje histonów zmieniające dostępność chromatyny. Modyfikacje epigenetyczne, takie jak np. metylacja cytozyny w miejscach CpG (cytozyna w dinukleotydzie z guaniną -CG) w genomie regulują ekspresję genów i mogą bezpośrednio hamować transkrypcję poprzez blokowanie wiązania określonych czynników transkrypcyjnych z miejscami regulatorowymi.

Badania nad HGGs doprowadziły do wykrycia licznych mutacji somatycznych, zaburzeń procesów transkrypcji oraz działania czynników transkrypcyjnych, modyfikacji histonów, zmian metylacji DNA i przebudowy chromatyny zachodzących równocześnie i przyczyniających się do agresywności tych guzów. Niewiele jednak wiadomo, jak zmieniają się genomy i profile transkryptomiczne HGGs w trakcie wznowy glejaka. Co więcej, pomimo wielu badań, specyficzna regulacja genów ulegających zwiększonej ekspresji w HGGs przez

czynniki transkrypcyjne pozostaje w dużej mierze nieznana. Lepsze zrozumienie zdarzeń zachodzących w otwartych regionach chromatyny w HGGs ma kluczowe znaczenie dla zrozumienia ścieżek progresji glejaka.

W niniejszej pracy zastosowaliśmy sekwencjonowanie DNA ze wzbogaceniem w sekwencje 700 genów związanych z nowotworzeniem (targeted DNA-sequencing) oraz sekwencjonowanie RNA (RNA-sequencing), aby zidentyfikować zmiany pojedynczych nukleotydów, małe insercje i delecje (indele), zmiany liczby kopii (CNAs), zmiany ekspresji genów i deregulacje szlaków sygnałowych w 16 sparowanych próbkach HGGs pochodzących od tych samych pacjentów, z guzów pierwotnych oraz po wznowach. Większość mutacji somatycznych wykrytych w pierwotnych HGGs nie została znaleziona w odrastających guzach, co sugeruje zamianę pierwotnego dominującego klonu komórek przez inny podczas progresji nowotworu. Zidentyfikowano nową insercję przesuwającą ramkę odczytu w genie *ZNF384*, która może odgrywać rolę w przebudowie macierzy zewnątrzkomórkowej. Stwierdzono, że częstość ogniskowych zmian ilości kopii (focal CNAs) występujących w genach *EGFR* i *PTEN* jest odwrotnie skorelowana. Analiza transkryptomiczna wykazała, że ekspresja genów zaangażowanych w obróbkę mRNA (splicing), cykl komórkowy i naprawę DNA jest podwyższona, podczas gdy geny zaangażowane w ścieżki sygnałowe zależne od interferonu i metabolizm fosfatydyloinozytolu (PI) są obniżone w HGGs po wznowie w porównaniu z guzami pierwotnymi. Analiza *in silico* mikrośrodowiska nowotworu wykazała, że wspierające nowotwór makrofagi (M2 macrophages) i niedojrzałe komórki dendrytyczne (immature dendritic cells, iDCs) są wzbogacone w nawracających HGGs, co wskazuje na obecność sygnatury immunosupresyjnej w tych nowotworach. Wyniki badań immunohistochemicznych na skrawkach guzów potwierdziły akumulację tych komórek immunosupresyjnych w przypadku wznowy.

Wykorzystując dane o dostępności chromatyny (Test chromatyny dostępnej dla transpozazy przy użyciu sekwencjonowania, ATAC-seq) w próbkach glejaka pobranych od pacjentów, a także w komórkach ludzkich glejaków LN18 i LN229 zidentyfikowano miejsca wiązania czynników transkrypcyjnych specyficznych dla HGGs. Potwierdzono rolę wybranych czynników transkrypcyjnych w kontrolowaniu sieci regulatorowych genów w HGGs. Przeanalizowano różne zestawy danych, które obejmowały profile metylacji DNA (ukierunkowane sekwencjonowanie z bisulfitacją), profile acetylacji histonów (H3K27ac ChIPseq) i profile transkryptomiczne komórek glejaka (RNA-seq). Uzyskane przewidywania były zweryfikowane w danych TCGA (The Cancer Genome Atlas) (metylacja DNA z macierzy Illumina 450K oraz RNA-seq) dotyczących glejaków złośliwych i łagodnych. Analizy porównawcze tych profili w glejakach o różnym stopniu złośliwości ujawniły znaczenie czynnika transkrypcyjnego c-Jun dla progresji tych nowotworów. cJun może odgrywać rolę w regulacji genów, które wykazują podwyższoną ekspresję w glejakach złośliwych, takich jak *VIM*, *FOSL2*, *UPP1*, *TRIB1* czy *GPR3,* poprzez wiązanie się z promotorami tych genów. Wykazano, że w większości docelowych

genów regulowanych przez c-Jun metylacja DNA może odgrywać ważną rolę w ich regulacji. Opisano również istotną pozytywną korelację między ekspresją mRNA/białka c-Jun a ekspresją docelowych genów w zbiorze danych TCGA, co wskazuje, że c-Jun prawdopodobnie reguluje ekspresję wielu genów związanych z inwazyjnością komórek w glejakach. Przewidywania bioinformatyczne zostały zweryfikowane doświadczalnie z wykorzystaniem techniki opóźnienia migracji elekroforetycznej w żelu (EMSA), która potwierdziła wiązanie białka c-Jun do motywu w promotorze genu *VIM* kodującego wimentynę.

Białka SMARCA2 i SMARCA4 są częścią kompleksu zmieniającego strukturę chromatyny. Geny kodujące te białka ulegają mutacjom w złośliwych glejakach. Aby określić rolę tych białek, przeprowadzono wyciszenie ekspresji *SMARCA2* i *SMARCA4* za pomocą specyficznych siRNA w ludzkich komórkach glejaka LN18 i zbadano konsekwencje obniżenia poziomu każdego z białek osobno lub w kombinacji na otwartość chromatyny za pomocą metody ATAC-seq. Wykazano wzrost otwartości chromatyny w komórkach z wyciszeniem SMARCA2/4, co zmieniło profil ekspresji genów. W komórkach z obniżoną ekspresją SMARCA2/4 stwierdzono indukcję ekspresji genów krytycznych dla transdukcji sygnału, w tym ze szlaku transformującego czynnika wzrostu beta (TGFbeta): *SMAD1*, *SMAD3*, *BMPR1A* i *TGFBR2*, co sugeruje współzależność białek zmieniających strukturę chromatyny i specyficznych szlaków sygnałowych.

Podsumowując, niniejsza rozprawa doktorska dostarcza nowych wyników wskazujących na znaczącą deregulację transkrypcji w komórkach glejaków złośliwych. Zidentyfikowano nowe aspekty związane z ewolucją klonalną guza oraz ze zmienionym mikrośrodowiskiem i odpowiedzią przeciwnowotworową w HGGs po wznowie, co może mieć ważne implikacje przy rozważaniu immunoterapii w glejakach. Przedstawiona kompleksowa identyfikacja kluczowych czynników transkrypcyjnych odpowiedzialnych za deregulacje transkrypcji w HGGs może utorować drogę przyszłym strategiom leczenia glejaków złośliwych.

Graphical abstract

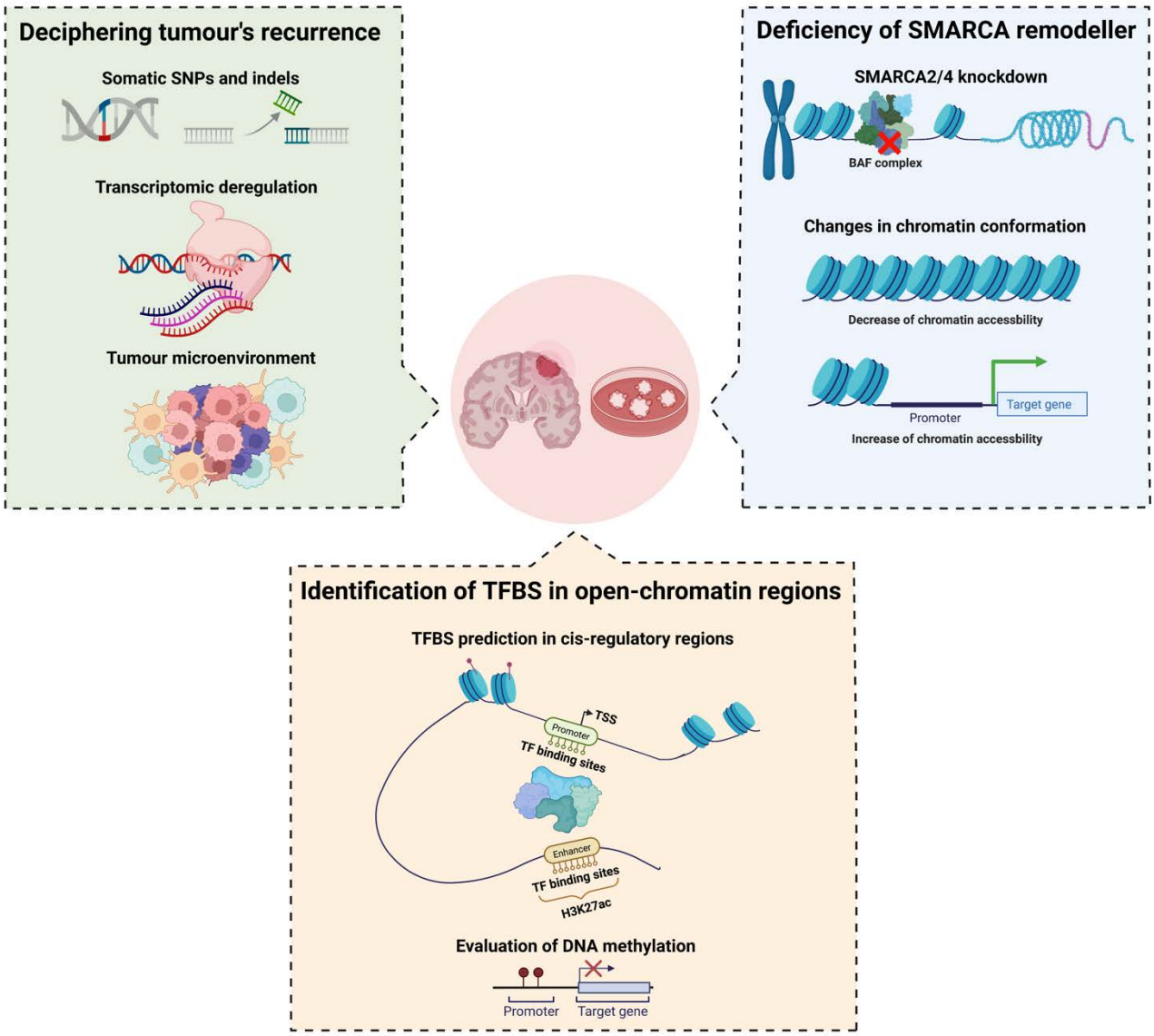


Figure i0. Graphical abstract. Workflow of the main topics covered in this thesis. SNP stands for single nucleotide polymorphism; indels stands for insertions/deletions; TFBS stands for transcription factor binding sites; TSS stands for transcription start site and TF stands for transcription factor. Created with BioRender.

Chapter I

1. Introduction

1.1. Characterisation of gliomas

A variety of benign and malignant tumours may develop in the central nervous system (CNS) (1). Historically, "glioma" is the name of a tumour that develops from a glial cell but recent studies showed that primary brain tumours develop from neural stem cells or progenitor glial cells (2). Adult HGGs are the most common and aggressive primary brain tumours –a tumour starting within the brain–, accounting for more than 80% of CNS tumours (3,4). Low-grade glioma (LGG) arise in young adults and have median survival time of 7 years (5), while survival of patients with glioblastoma (GBM) is about 12 months (6,7). GBMs are highly resistant to the standard treatment, which includes surgical resection followed by radiation and chemotherapy, and are essentially incurable (8,9). The use of temozolomide (TMZ), an alkylating drug, improves median survival of GBM patients by 2.5 months (10). Patient survival slightly improved over the last decades as the knowledge of cancer-specific mechanisms and vulnerabilities has been used to target the disease (11–13).

The application of genomics techniques to neoplasm characterisation has improved the prior classification of brain tumours based on histology and clinical information (14–16). The prominent example is a comprehensive catalogue of genomic aberrations generated thanks to a more available and affordable DNA sequencing technology by various consortia, including the TCGA (The Cancer Genome Atlas) consortium (15,17). Another example is the GLASS (The Glioma Longitudinal Analysis) consortium aiming to molecularly characterise tumour specimens acquired at various time points along the course of glioma progression (18,19).

While those genome-wide efforts have made significant progress in identifying “candidate” cancer genes, they did not fully explain phenotypic and histological disparities across and within tumour types. Simultaneously, gene expression profiling approaches have highlighted the transcriptome variety found in tumour samples (20). GBMs were classified as proneural, neural, classical, or mesenchymal transcriptional subtypes based on gene expression profiles (21). Recent single-cell investigations have discovered that malignant cells in GBM exist in four unique biological states that recapitulate different neural cell types, are regulated by the tumour microenvironment and exhibit plasticity (22). Resolving the transcriptional diversity, intra-tumoural heterogeneity and influence of stromal cells in gliomas and other cancer types would help in the search for tumorigenesis-related mechanisms, designing new therapies, and extracting the prognostic information.

Importantly, GBMs are considered to be "immunologically cold" tumours because of severe immunosuppression, which influences disease progression and immunotherapy success (23). Furthermore, as it has been extensively documented, glioma-infiltrating microglia and macrophages, the most common immune cells in GBMs, contribute to tumour invasion and generate an immunosuppressive environment (24–28). Glioma stem cells (GSCs), a rare tumour initiating cells, are present in GBMs. GSCs are responsible for tumour resistance to therapy and are thought to contribute to tumour relapse (29).

1.2. Classification of gliomas

The World Health Organisation (WHO) classified gliomas into four malignancy grades, with a WHO grade I being the least aggressive and grade IV being the most aggressive (30) (Fig. i1). Because a patient's diagnosis is based on the most malignant part of the tumour, it is critical to sample the tumour adequately in order to determine its type and assess its malignant potential (31,32). The 5th edition of the WHO Classification of Tumours of the Central Nervous System (CNS), published in 2021, is the most updated international standard for the classification of brain and spinal cord tumours (33). Specifically, a brain tumour's grade is determined by its features and based on five histopathology criteria that are related to the degree of anaplasia: cellular density, nuclear atypia, mitosis, endothelial proliferation and necrosis (34).

Pilocytic astrocytomas (PAs) are benign (WHO grade I), slow growing tumours that may originate in different areas and can be surgically removed (35). Diffuse astrocytomas (DAs) are benign tumours (WHO grade II) but they are more likely to recur in a more malignant manner after treatment (36). Anaplastic astrocytomas (WHO grade III) have rapidly dividing cells and can grow quickly but do not show necrosis (37). Finally, grade IV tumours grow fast, spread quickly, are actively dividing with extensive blood vessel growth and areas of necrosis (38–40). The recent WHO classification of gliomas employs the current knowledge about genetic alterations (Fig. i1).

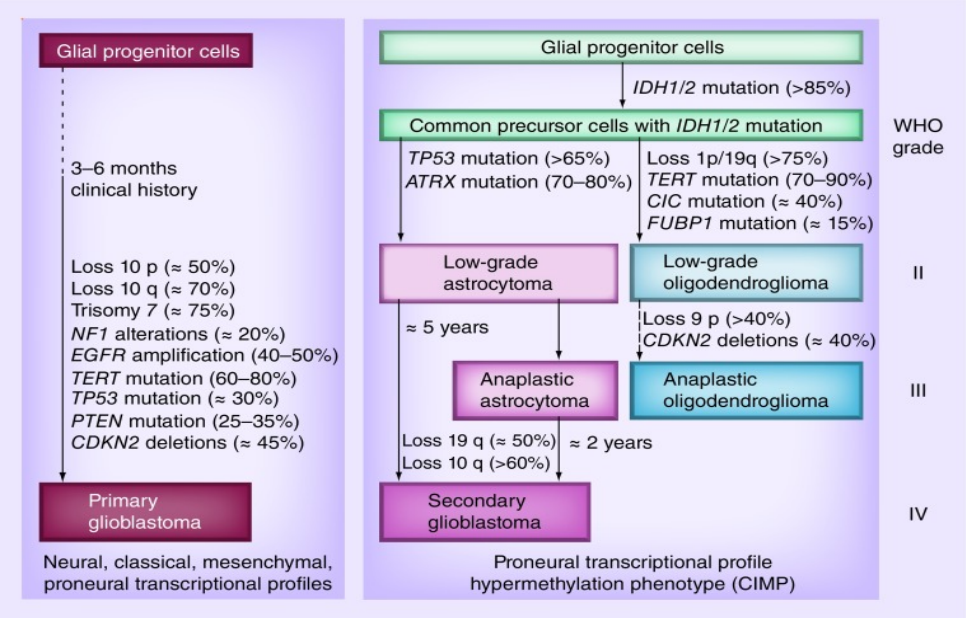


Figure i1. Classification of gliomas and common genetic alterations. Glioblastomas (GBMs, WHO grade IV) are classified as either *de novo* (primary) or secondary (secondary) GBMs, which develop directly from low-grade astrocytomas (WHO grade II) or by malignant transition from anaplastic astrocytomas (WHO grade III). These two GBM types exhibit distinct genomic changes. *IDH1/2* mutations are frequently seen in WHO grade II and III gliomas, and in 10% of GBMs. *IDH* mutations are early events that occur before codeletion of chromosomes 1p and 19q in oligodendrogliomas or *TP53* mutation in astrocytomas (adapted from Taal *et al*., 2015, 41).

1.3. Inter- and intra-tumour heterogeneity in glioblastoma

Tumour heterogeneity has long been assumed to contribute to failure of molecularly targeted cancer therapies (42). One of the first studies describing molecular subtypes in GBM revealed that there are different genomic aberrations (called copy number aberrations, CNAs), mainly involving chromosome 7 and chromosome 10 (43,44). More importantly, GBMs can be classified into various categories based on these CNA signature subclasses (43). In order to reveal the molecular basis of brain tumours, the multidimensional analyses were carried out by the TCGA consortium and deciphered differences in CNAs and pattern of somatic mutations, which classify GBMs into subtypes (15). When it comes to transcriptomic GBM subtypes, bulk-tissue RNA-seq studies identified four distinct signatures (proneural, neural, classical and mesenchymal) which are linked to specific gene aberrations in *EGFR*, *NF1*, *PDGFRA/IDH1* (21). Interestingly, recent single-cell RNA-sequencing (scRNA-seq) studies revealed that GBM cells exist in four distinct cellular states that mimic different neural cell types, are influenced by the tumour microenvironment and exhibit plasticity (22). scRNA-seq studies demonstrated that the relative frequency of cells in each state varies between GBM samples and is influenced by copy number amplifications of the *CDK4*, *EGFR*, and *PDGFRA* loci, as well as mutations in the *NF1* locus, all of which favour a specific state (22).

Intra-tumour heterogeneity is defined as the presence of multiple cell subpopulations within a single tumour from a single patient, and it allows the tumour to respond to selective pressures, contributing to tumour aggressiveness, growth, and treatment failure. Intra-tumour heterogeneity has been investigated in GBM, and its understanding is critical for assessing treatment responses and designing personalized treatment strategies. GBMs are composed of genetically distinct clones that have different tumorigenic potential (45) and evolve over time. Current GBM therapies are suboptimal since they do not target specific genomic alterations (46), and some genetic alterations, such as changes in the p53 pathway, are considered primary molecular events (47).

The other layer of intra-tumour heterogeneity in GBM is defined by spatial differences between a tumour core and a tumour periphery. The core of a GBM is characterised by high proliferation, inflammation, and necrosis, whereas the peripheral brain zone (PBZ) consists primarily of brain parenchyma with isolated tumour cell infiltrates (48). Moreover, the presence of these isolated cell groups, dispersed throughout normal brain tissues in the PBZ, explains why surgical resection is ineffective and recurrence is almost inevitable. Interestingly, the PBZ region of the tumour core is associated with the expression of proneural genes, whereas the central core expresses mesenchymal genes (49).

1.4. Glioblastoma's microenvironment

GBM is composed not only of malignant cells, but also a variety of other cells that originate either in the CNS or infiltrate from the periphery and form a dynamic complex known as a tumour microenvironment (TME). TME is critical in sustaining growth and proliferation of GBM (50,51). TME is composed of many non-malignant cells, such as reactive astrocytes, fibroblasts, endothelial cells, and numerous immune cells, all of which work together to create an immunosuppressive environment (52–54). Immune cells, particularly myeloid cells such as microglia and infiltrating bone marrow (BM)-derived macrophages, consist a significant portion of the GBM volume, accounting for up to 30 or 40% of the tumour mass (55). Glioma-associated microglia and macrophages collectively called (GAMs) accumulate within the tumour, where they play an important role in tumour progression (56). Intriguingly, microglia depletion reduces glioma growth and affects glioma patient survival (57). Communication between GAMs and glioma is driven by a number of glioma-derived factors, such as chemokine (C-X3-C motif) ligand 1 (CX3CL1), chemokine (C-C motif) ligand 2 (CCL2), the colony stimulating factor 1 (CSF-1), the stromal cell-derived factor 1 (SDF-1), granulocyte-macrophage colony-stimulating factor (GM-CSF) and protein-lysine 6-oxidase (LOX), that can act as chemoattractants for GAMs (58). Similarly, in animal experimental GBM models, tumour cells produce Ccl2, which attracts macrophages (59), so inhibiting Ccl2/Ccr2 prolongs mouse survival (60).

Toll-like receptors (TLRs), belonging to a receptor family expressed on microglia, may be useful in the establishment of these new functional states. TLRs were identified as pathogen sensors, and their expression is increased in GAMs (61). Studies of GL261 gliomas - a murine glioma cell line - demonstrated microglial TLR2-dependence of glioma growth, mediated partially by the increased production of a membrane-type metalloproteinase (MT1-MMP/MMP14), which is required for MMP2 activation and malignant glioma invasion (62). TLR2 signalling, in addition to MMP14, causes the release of MMP9 by microglia, which acts as an additional factor in degrading the extracellular matrix (ECM) to promote glioma invasion and growth (63).

During glioma progression, the ECM undergoes deposition and remodeling, changing its composition and architecture. These physical changes in the TME may facilitate glioma stiffness, specifically elevate fluid pressure (due to oedema), cell compression, and increase tumour cellular contractility, all of which promote tumour progression (64,65).

1.5. Glioblastoma's genetics: frequent somatic mutations and CNAs

Understanding the molecular pathways underlying aggressive behaviour of HGGs may lead to better treatment, effective medications and better outcomes. Somatic evolution, a process in which an accumulation of mutations leads the alteration of cancer cell genome from that of a healthy cell, promotes cancer progression (66). The development of GBM occurs through a complex sequence of different genetic aberrations (16,67,68), resulting in considerable changes in major signalling pathways. In recent years, evidence has emerged that tumours are composed of multiple populations of malignant cells harbouring specific genetic alterations (69). The interplay of cancer-predisposing constitutive genetic alterations in conjunction with known or unknown environmental risk factors, as well as somatic genetic alterations ultimately drive pre-existing glial stem cells to abnormal proliferation and malignant transformation (70,71).

The most commonly altered pathway involves receptor tyrosine kinases (RTKs) (72), which are cell-surface receptors that bind growth factors (GFs), causing dimerization of two adjacent receptors, conformational shift, kinase activation, and cross-phosphorylation of tyrosine residues initiating downstream signalling cascades (73). Another significant genetic change in GMBs involve alterations in epidermal growth factor receptor (EGFR) signalling, which can be activated through overexpression of the receptor or its ligand, amplification of the *EGFR* locus, and/or *EGFR* mutations (74,75). EGFR signalling in cancer has been the subject of considerable research for decades, mainly due to its alterations in nearly half of all human tumours (74,76). The tumour suppressor p53 pathway, which is regulated by the p53 transcription factor, plays a central role in maintaining cellular homeostasis, cell proliferation control, cellular senescence, and apoptosis to prevent damaged cells from propagating (77–79). Using next-generation sequencing on large human GBM cohorts, the

researchers defined the somatic landscape in GBM, showing that *PTEN, TP53, EGFR, PIK3CA, PIK3R1, NF1, RB1, IDH1 and PDGFRA* are the most frequently mutated genes (16).

CNAs are main somatic variations affecting the chromosome structure that result in either a gain or loss of copies in sections of DNA (80). The most prevalent CNAs identified in GBM include loss, or partial loss, of chromosomes 9 and 10; polysomy of chromosomes 7, 19, and 20; focal deletion of *CDKN2A/B* locus (9p21.3); and focal high-level amplifications of *EGFR* locus (7p11.2) (15,81).

1.6. Recurrent glioblastoma: an unsettling reality

The standard of care for most patients with newly diagnosed GBM includes surgical tumour resection followed by radio-chemotherapy (82). Because of the high invasiveness, even radical resection of the primary tumour mass is not curative and infiltrating tumour cells invariably remain within the surrounding brain, leading to disease progression or recurrence (Fig. i2) (83,84). Recurrent tumours are less responsive to therapy than primary tumours and in most cases, as tumour cells invade functioning brain areas, a second surgical resection is limited (85). In contrast to newly diagnosed GBM, there is no an additive therapy for individuals with a recurrence other than tumour excision and patients die within 12-15 months of initial diagnosis (84). Several studies found statistically significant associations between greater resection and longer overall survival (OS) (86,87). Other more advanced surgical technologies, such as the use of 5-aminolevulinic acid (5-ALA) dye for fluorescence guidance, have been shown to be more effective than traditional neuronavigation-guided surgery (88).

The biology of recurrent GBMs is largely unknown due to the fact that not all recurrent GMBs are surgically accessible and tumours had larger areas of necrotic tissues with lower vital tumour cell content than their primary counterparts (89,90). As a result, obtaining a high-quality tissue is challenging and complicates large-scale research.

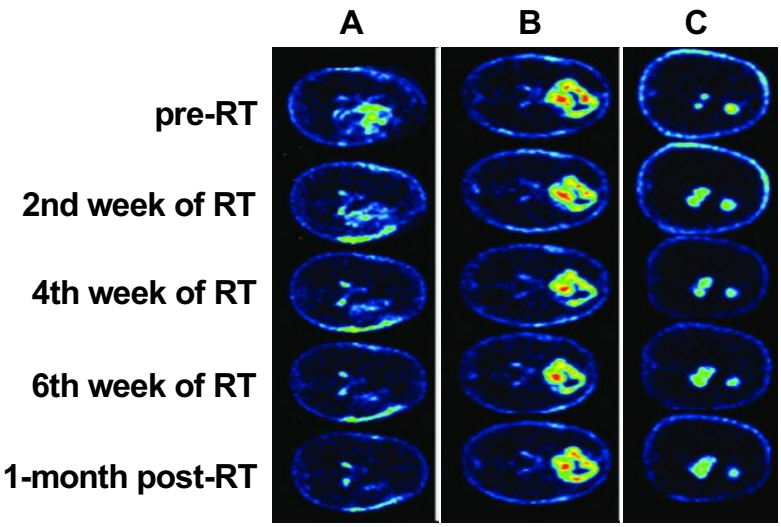


Figure i2. 18 F-Fluoromethylcholine positron emission tomography magnetic resonance imaging (MRI) in 3 glioblastoma patients. (A) Patient with GBM in the right frontal and temporal lobes; the patient is classified as a partial responder by the Response Assessment in Neuro-Oncology (RANO) criteria. (B) Patient with a bifrontal GBM; the patient was classified as stable disease using the RANO criteria. (C) Patient with multifocal GBM; a new lesion was discovered on a follow-up MRI, indicating that the patient's disease was progressing (adapted from Bolcaen *et al.*, 2017).

Recurrences are typically local, with nearly two-thirds of tumours regrowing within two centimetres of the initial tumour margin. In a more recent large-scale study, whole-exome sequencing was performed on 15 pairs of local GBM recurrences, and percentages of mutations shared between primary and recurrent tumours varied greatly between tumours (11-97% of shared mutations) (Kim *et al.*, 2015). Indeed, a direct comparison of primary and recurrent tumours, might indicate molecular changes associated with a therapy resistance of recurrent GBMs. Immunohistochemistry (IHC) or methylation-specific PCR studies revealed down-regulation of several genes coding for mismatch repair enzymes, such as *MLH1* or *MSH2* (93) and variations in DNA methylation patterns (94–96) in recurrent GMBs.

Several tumour clones may coexist within the same tumour and express diverse repertoires of oncogenic drivers in a temporo-spatial manner that can be changed by therapy. The presence of dominant subclones that can either shrink and disappear after relapse or survive and reappear as dominant subclones in the recurrent tumour, is one discernible pattern of recurrence (67,97,98). Smaller subclones can also survive therapy and resurge as dominant subclones in the recurrent tumour, or dominant clones would activate additional dormant subclones to become dominant (99). The presence of cancer stem cells is another critical element in the recurrence of GBM. Cancer stem cells (CSCs) are slow-cycling tumour cells with the enhanced self-renewal potential (100), increased resistance to radiotherapy and chemotherapy (101) through overexpression of DNA damage repair enzymes that protect CSCs from the oxygen-dependent effects of radiotherapy (102). Furthermore, GSCs undergo asymmetric cell division, resulting in fast-growing, therapy-sensitive tumour cells

(103,104). The ability of some glioma cells to dedifferentiate into CSCs shows that glioma cells may be more plastic than their neural counterparts (105).

1.7. Dysregulation of epigenomic mechanisms in glioblastoma

1.7.1. DNA methylation and the IDH phenotype

Aberrant epigenetic patterns are common features underlying the development and progression of brain tumours and are manifested by profound changes in DNA methylation and chromatin activation (106,107).

Several studies have shown that DNA methylation pattern in glioma cells differs from those in normal cells (96,108) and DNA methylation correlates negatively with a tumour grade (109). Some housekeeping genes, such as DNA repair genes and tumour suppressor genes, are frequently hypermethylated in tumour tissues (110,111) and show reduced expression, which leads to genetic instability (112). Functional silencing of tumour-associated genes is usually associated with local promoter hypermethylation and malignant phenotypes (113,114). One of the most studied promoter methylation is the epigenetic silencing of O6-methylguanine DNA methyltransferase (*MGMT*) gene. MGMT is a DNA repair protein responsible for the direct repair of TMZ-induced toxic DNA adducts, conferring drug resistance upon treatment with alkylating agents such as TMZ (115,116). In general, the Illumina human methylation 450K BeadChip array (HM450 K), which covers over 480K CpG sites and targets 96% of CpG islands in the human genome, is used to detect *MGMT* methylation (117,118). Bisulfite genomic sequencing is also regarded as a gold-standard technology for detecting DNA methylation because it allows for the identification of 5-methylcytosine at a single base-pair (bp) resolution. After treatment with sodium bisulfite, amination reactions of cytosine and 5-methylcytosine (5mC) have very different outcomes; in this case, cytosines in single-stranded DNA will be converted into uracil residues and recognized as thymine in subsequent PCR amplification and sequencing, but 5mCs will be insensitive to this conversion and will remain as cytosines, allowing 5mCs to be distinguished from unmethylated cytosines (119). After the bisulfite treatment, a PCR reaction with specific methylation primers is required to determine the methylation status at a locus of interest (119,120).

In the same line, crucial genes involved in cell cycle, DNA repair and tumour suppression could be silenced by the promoter hypermethylation in GBM (121–123). The first DNA methylation analysis identified 616 CpG sites differentially methylated between GBM and control brain tissue, and concordant CpG sites displayed an inverse correlation between the promoter methylation and expression level for several GBM genes (*B3GNT5, FABP7, ZNF217, BST2, OAS1, SLC13A5,*

GSTM5, ME1, UBXD3, TSPYL5, FAAH, C7orf13, and *C3orf14*), indicating that gene expression is tightly regulated by epigenetic mechanisms (124). Glioma CpG island methylator phenotype (G-CIMP) has been an object of many studies and is associated with the isocitrate dehydrogenase (IDH) mutational status (96,125). Patients carrying G-CIMP (G-CIMP+) tumours have shown a better prognosis than those not carrying that phenotype (G-CIMP−) (126). Specific CpG loci differentially hypermethylated in GBM patients with short- and long-term survival, includes members of the homeobox gene family (*HOXD8, HOXD13* and *HOXC4*), among other regulators (126,127).

One important way of influencing gene expression by DNA methylation is methylation of the binding site with a regulatory element, which can disrupt protein:DNA interactions. Removing site-specific DNA methylation allows binding to a regulatory element, which provides opportunities for potential therapy (128,129). A comprehensive study of the effect of CpG methylation on the binding of human TFs demonstrated that the methylated CpG sites inhibit binding of numerous TFs, including basic helix-loop-helix (bHLH), basic zipper (bZIP), and erythroblast transformation specific (ETS) members. On the contrary, TFs such as homeodomain, Pit-Oct-Unc (POU), and nuclear factor of activated T-cells (NFAT) proteins preferred to bind the methylated DNA (130).

As outlined above, DNA methylation is tightly associated with the IDH phenotype and many glioma patients harbour somatic mutations in the *IDH1* or *IDH2* genes (131,132). These mutations, which occur early in gliomagenesis, are point mutations (R132H in IDH1 and R172K in IDH2) and are found in WHO grade II, grade III gliomas (~68%) and about 5% of GBMs (~85% secondary GBMs) (133). IDH mutation contributes to oncogenesis as the altered IDH protein processes α-ketoglutarate (α-KG) to 2-hydroxyglutarate (2-HG) (134). 2-HG is an oncometabolite which inhibits the activity of many α-KG-dependent dioxygenases, including the ten-eleven translocation (TET) family of DNA hydroxylases (134,135) and histone demethylases such as lysine demethylase 4A (KDM4) and lysine demethylase 5B KDM5 (136), resulting in a hyper-methylated phenotype (Fig. i3).

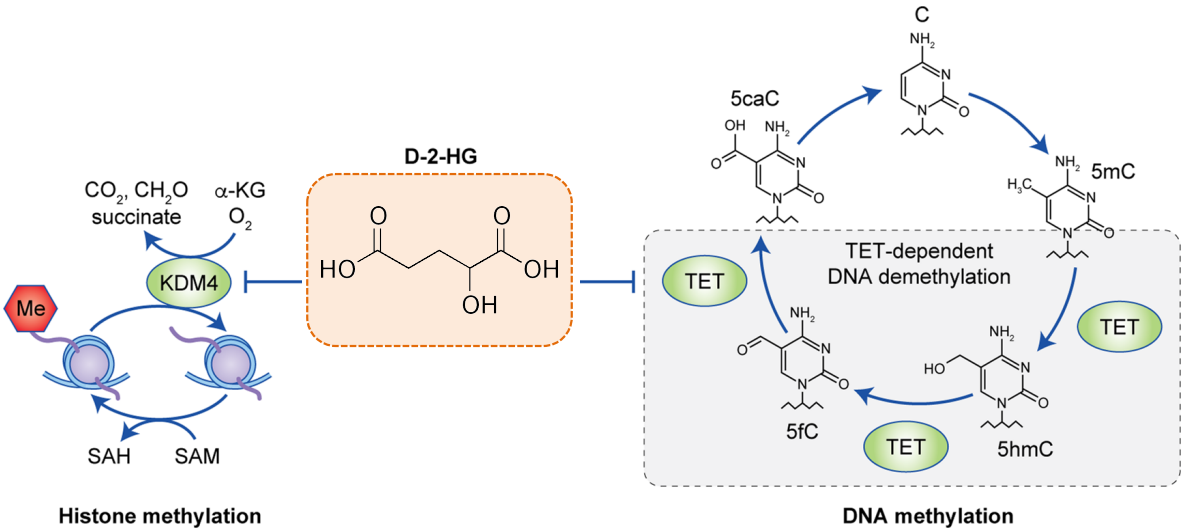


Figure i3. **The impact of *IDH1* mutations in glioma**. 2-hydroxyglutarate (D-2-HG), an oncometabolite produced from mutated *IDH1*, acts as a competitive inhibitor of lysine demethylase 4A (KDM4) or ten-eleven translocation (TET), blocking histone and DNA demethylation, respectively. SAH stands for S-Adenosyl-L-homocysteine; SAM stands for S-Adenosyl methionine; 5mc stands for 5-Methylcytosine; 5hmC stands for 5-Hydroxymethylcytosine; 5fC stands for 5-Formylcytosine; 5caC stands for 5-carboxylcytosine; C stands for Cytosine and Me stands for methyl radical (adapted from Han *et al.*, 2020).

1.7.2. Histone modification patterns and epigenetic dysregulation in gliomas

Another layer of epigenetic regulation in GBM involves histone modifications and chromatin remodeling. Histones are proteins providing a scaffold for DNA which is wrapped around core histones forming a nucleosome. N-terminal tails of histones can be post-translationally modified by methylation, acetylation, ubiquitylation and phosphorylation which alters their interactions with DNA and nuclear proteins (138,139). Chromatin relaxation through the methylation of H3K4 (Lysine 4 of Histone 3) allows transcription to be initiated, whereas chromatin closure due to methylation of H3K9 and H3K27 (Lysine 9 and Lysine 27 of Histone 3) constitutes the two main repressive mechanisms in mammalian cells (140,141). Histone acetylation is controlled by two types of histone acetyltransferases (HATs), which transfer an acetyl group (acetyl-CoA) to the α-amino groups on the N-terminal tails of histones (142,143). Histone deacetylases (HDACs) reverse this modification (144). Lysine acetylation neutralizes the lysine's positive charge, weakening histone-DNA or nucleosome-nucleosome interactions, thus inducing an open conformation of the chromatin and facilitating access of different nuclear factors to DNA (145–147). Histones (mainly H3 and H4) can also be methylated (148) and this modification is carried out by lysine methyltransferases and arginine methyltransferases, while it is reversed by lysine demethylases (149). The consequences of lysine methylation are extremely diverse. Depending on the targeted lysine, methylation can either activate or repress transcription (150). For instance, methylation of histones H3K9, H3K27, and H4K20 are mainly involved in

formation of the heterochromatin, while methylation of H3K4, H3K36, and H3K79 are associated with the euchromatin (151).

In paediatric GBMs, somatic mutations in genes coding for the H3.3-ATRX-DAXX chromatin remodeling complex proteins or histone H3F3A were detected and those alterations resulted in lengthening of telomeres, specific gene expression profiles and blockade of differentiation (152). Abnormal histone modifications can cause aberrations in gene expression, which can lead to the development and progression of gliomas. For example, HDAC1 expression is high in gliomas and knockdown of the *HDAC1* gene inhibits cell proliferation and invasion (153). HDAC inhibitors emerged as potential epigenetic drugs and are considered for treatment of gliomas (154). Clinical trials of HDAC inhibitors suberanilohydroxamic acid or valproic acid have shown that they can decrease glioma's cell cycle progression and proliferation by inhibiting G2 checkpoint kinases and proteins involved in DNA repair and mitotic spindle formation (155). Treatment with a specific inhibitor of histone deacetylase 6 (HDAC6), which is abundantly expressed in GBM, resulted in reduction of tumour growth *in vivo* (156). Other studies revealed the efficacy of HDAC inhibitors in sensitizing GBM cells to chemo- and radiotherapy (157–159).

1.7.3. Accessibility and 3D structure of the chromatin

Physical access to DNA is a highly dynamic property of chromatin that plays an essential role in establishing cellular identity. Organization of accessible chromatin across the genome reflects a network of permissible physical interactions through which enhancers, promoters, insulators and chromatin-binding factors cooperatively regulate gene expression (160). Although eukaryotic genomes are generally packed into nucleosomes, nucleosome occupancy is not uniform in the genome and can be affected by epigenetic factors or DNA binding molecules (161). In cis-regulatory elements (mainly promoters and enhancers), nucleosomes tend to be depleted, resulting in accessible chromatin where binding of transcriptional regulators is facilitated (162,163). Post-translational modifications in histones (see section 1.7.2) impact the chromatin openness, and for example, histone acetylation reduces histone-DNA binding and may contribute directly to chromatin opening and indirectly to transcriptional upregulation (164). Consequently, the analysis of TF binding sites in regulatory regions within accessible chromatin can bring insights into cell type-specific lineage factors and gene regulatory networks.

In order to investigate chromatin accessible regions, Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) has been established (165). In this very sensitive method, a genetically engineered hyperactive DNA transposase (Tn5) inserts sequencing adapters into accessible regions of chromatin and resulting sequencing reads can be used to infer regions of increased accessibility (Fig. i4), as well as to map regions of TF binding and nucleosome position

(165). Other important method to understand the intrinsic chromatin interactions in the nucleus is a Hi-C assay (166). In this method, chromatin is crosslinked with formaldehyde, then digested and re-ligated in such way that only DNA fragments that are covalently linked together form ligation products.

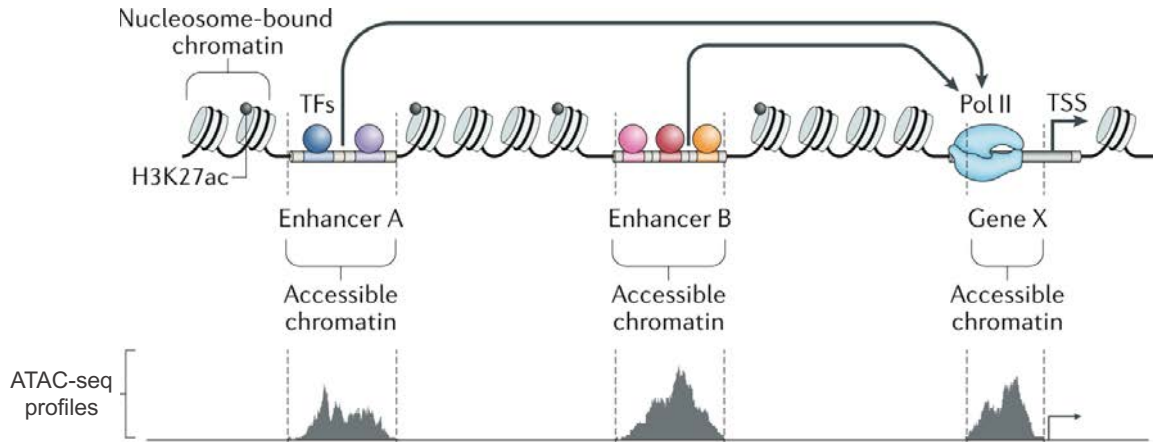


Figure i4. Chromatin landscape and chromatin accessibility. Transcription factor (TF)-bound enhancers and a gene promoter are nucleosome depleted and hence accessible. The TFs are depicted as coloured circles, and the arrows represent the 3D looping of the enhancers to the target gene promoter. TSS stands for transcription start site and Pol II stands for RNA polymerase II (adapted from Minnoye *et al.*, 2021).

Ligation products contain the information from where they originated but also where they reside, physically, in the 3D organization of the genome. Using a Hi-C chromosome conformation capture technique it is possible to delineate topologically associated domains (TADs), which are defined as regions with DNA sequences that preferentially contact each other (168). TADs have been shown in genomes of multiple species, including in human genomes (169). This technique allows to determine which specific enhancer and active promoter contacts are responsible for gene expression (170,171).

1.7.4. Chromatin remodeling

The accessibility of the transcriptional machinery to chromatin and DNA is directly related to gene expression. Cells use a variety of ATP-dependent nucleosome-remodeling complexes to perform histone sliding, ejection, or integration of histone variations (172). Chromatin remodeling complexes (CRCs) have an ATPase activity and rely on ATP hydrolysis to provide energy for chromatin structure changes (173). Altering histone-DNA contacts and thus repositioning or removing nucleosomes (174) is an important mechanism of regulating gene expression by controlling chromatin dynamics (175) (Figure i5).

One of the best studied remodeling complexes is the SWI/SNF (SWItch/Sucrose Non-Fermentable, also known as BAF) complex, an evolutionary conserved chromatin remodeling complex

composed of approximately 15 protein subunits (176,177). High prevalence of mutations in genes encoding its subunits has been reported, with nearly 25% of all cancers harbouring aberrations in one or more of these genes (178). Most SWI/SNF complexes contain many core components, including SMARCB1/BAF47, SMARCC1/BAF155, SMARCC2/BAF170 as well as one of the two mutually exclusive ATPase subunits, SMARCA4/Brg1 and SMARCA2/Brm (177). Several studies have revealed that the SWI/SNF complex plays a role in chromatin remodeling, which occurs in cis-regulatory elements such as promoters and enhancers, as well as in very large distal regions harbouring many TF-binding sites, coined super-enhancers (179–181). Polycomb group (PcG) proteins, which include Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2) proteins, are a gene-silencing system that plays a key role in multicellular development, stem cell biology and cancer (182). SWI/SNF complexes regulate Polycomb-group proteins at bivalent chromatin sites by preventing the accumulation of Polycomb complexes through ATP-dependent eviction (183). Interestingly, it has been demonstrated that the SWI/SNF complex may have a tumour-suppressive function, and that inactivating mutations or deletions fail to oppose PRC1 and PRC2, which results in an imbalance between differentiation and self-renewal and eventually leads to tumorigenesis (184).

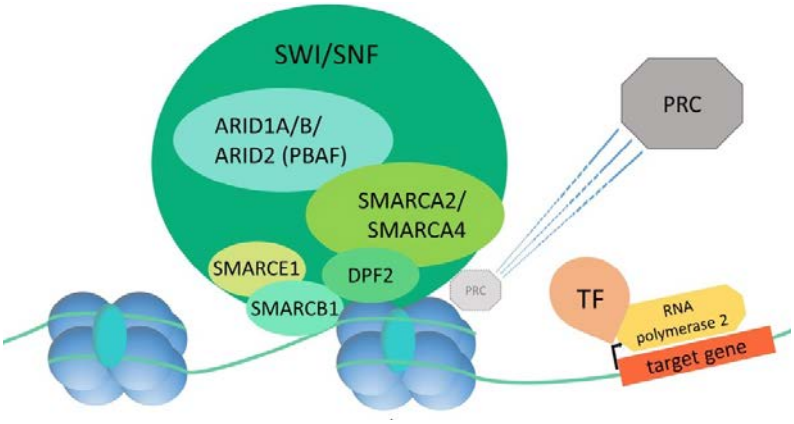


Figure i5. The SWI/SNF-dependent chromatin remodeling complex. The complex binds to DNA and histones causing nucleosome displacement to improve DNA accessibility and, as a result, the initiation of the transcription machinery. SWI/SNF has a direct PRC eviction capacity that is SMARCA4-dependent. PRC stands for polycomb repressive complex and TF stands for transcription factor (adapted from Bögershausen and Wollnik, 2018).

In gliomas, the SWI/SNF complex plays critical roles in stemness maintenance in glioma initiating cells (GICs) (186), suggesting new therapeutic routes for GBM treatment. Specific inhibitors that block the bromodomain (BRD) of the Brg1 subunit of the SWI/SNF complex, showed anti-proliferative effects and cell death of cultured GMB cells, and sensitized cells to TMZ (187).

1.8. Transcription factors in tumorigenesis

Many cancer signalling pathways are controlled by the expression of oncogenes (188). Transcription initiation is dependent on the transcriptional machinery activated in spatially and temporally coordinated manner. This machinery includes RNA polymerase II, general and specific TFs, activators binding proximal promoters and distal regulatory elements (189,190). Dysregulation of the transcriptional machinery contributes to both cancer initiation and persistence. It can occur by aberrant activation, repression and/or temporal/spatial deregulation as well as by structural changes including mutations (191). For example, most tumour cells depend on the c-Myc TF for their growth and proliferation, and its genomic locus is frequently amplified (192–194).

Many TFs have been identified as associated with cancer. The nuclear factor (NF)-κB belongs a family of five TFs (195), was found active in several cancer types and is known to influence gene transcription through a series of events starting from translocation of NFκB to the nucleus to activation of genes implicated in the regulation of cell proliferation, survival, invasion, apoptosis inhibition, metastasis and epithelial-mesenchymal transition (196,197). Another important TF is the activator protein-1 (AP-1), a dimer comprised of c-Jun, JunB, JunD, c-Fos, FRA-1 and FRA-2 proteins (198) involved mainly in cellular signalling processes (199). AP-1 is assembled via dimerization, which confers specificity and stability while also determining the composition of its leucine zipper region (200), resulting in a variety of pro-cancer effects. AP-1 complex is primarily regulated at the level of *JUN* and *FOS* gene transcription by TFs activated by mitogen-activated protein kinases (MAPKs). Post-translational modifications via phosphorylation and dephosphorylation regulate AP-1 components and upstream MAPKs (201,202). Moreover, AP-1 components regulate cancer cell proliferation through the repression of tumour suppressor genes (203), as well as through the induction of *cyclin D1* transcription (204). On the other hand, JUNB and JUND are more frequently negative regulators (205), in some cases linked to the prevention of myeloid malignancies by limiting hematopoietic stem cell proliferation (206). However, those TFs could also stimulate proliferation (207–209). Modulation of AP-1 activity may be a novel approach to reducing malignant transformation by targeting exclusively neoplastic cells (210,211).

Transcriptomic data have revealed that *HOXA3, EN1, ZIC1* and *FOXD3* genes coding for TFs are differentially expressed in GBMs when compared to lower-grade gliomas (212) (213). Other studies have described a key TF cluster with significant prognostic value in GBM patients, such as the TF cluster composed of *AHR, ATF3, BLNK, CEBPA, EGR2, FAS, FHL2, FOS, HCK, ID1, IQCG, MAFB, MYLK, MYO1B, NR2F2, PDLIM1, PDLIM4, PLK2, PRRX1, SAMSN1, SLA, SNAI2, TNFRSF11B and TWIST1* TFs (214).

1.9. Aims of the study

HGGs remain CNS tumours with the worst prognosis, despite multiple attempts to improve patient survival. Despite minor advances in surgical techniques, chemotherapy, and more focused radiation therapy, GBMs have been treated essentially in the same way for the past 30 years. Standard chemotherapy treatment with temozolomide is still used in conjunction with radiation therapy, sometimes in combination with an anti-angiogenic drug such as bevacizumab. Due to molecular and genetic complexity of HGGs, these brain tumours are impossible to treat effectively. Chemotherapy and immunotherapies do not traverse effectively the blood-brain barrier, tumours frequently mutate and become resistant to chemotherapy. Furthermore, HGGs typically recur as a result of inadequate surgery or therapy resistance. In addition, HGGs have the ability to inhibit anti-tumour responses of the immune system. Combination of those factors makes treating patients and testing novel medications extremely challenging. There is a need to learn more about tumour progression mechanisms and diagnostic markers, as well as identify new promising targets in order to provide a better therapeutic option for these patients. A search for new combinations and translation of best candidates from basic studies to clinical trials is envisioned.

Hence, the main aim of this study was to better understand the pathobiology of HGGs by the characterisation of genomic and transcriptomic profiles of HGGs after recurrence, and the analysis of transcriptomic and epigenetic deregulation in HGGs.

The specific aims were as follows:

1. **Characterisation of genomic and transcriptomic profiles of HGGs after recurrence.**
2. **Evaluation of chromatin dysregulation by studying landscapes of histone modifications and chromatin openness in HGGs and benign brain tumours.**
3. **Integration of various datasets to predict gene regulatory networks and candidate transcription factors implicated in GBM pathobiology.**
4. **Evaluation of the impacts of SMARCA2 and SMARCA4 deficiency on chromatin openness and gene regulation.**

A widely range of approaches were used in this study. We used transcriptomics (RNA-sequencing) and targeted sequencing of cancer-related genes to identify profiles of single nucleotide variants, small insertions and deletions, CNAs, gene expression alterations, and pathway dysregulations in 16 pairs (35 human resected tumours) of primary and recurrent HGGs. We searched for differences in the microenvironment of primary and recurrent HGG tumours using *in silico* cell enrichment studies and subsequent wet lab validation using immunohistochemistry and immunofluorescence. Multi-omics data, such as chromatin accessibility data (ATAC-seq), whole genome bisulfite sequencing,

histone acetylation (H3K27ac), transcriptomics profiling of GBM cells (RNA-seq), and TCGA public datasets (RNA-seq and Illumina 450k array DNA methylation) were integrated to identify glioma grade-specific TFs binding sites in human LN18 and LN229 glioma cells, as well as in GBM tissues. Finally, we used the ATAC-seq technique to investigate the impact of SMARCA2 and SMARCA4 deficiencies on chromatin accessibility in human LN18 glioma cells.

Chapter II

2. Methods

2.1. Computational analysis of recurrent high-grade gliomas

2.1.1. Description of a study cohort and sample collection

In total, 35 fresh frozen glioma samples from three different hospitals in Poland were collected (*Mazovian Bródno* Hospital in Warsaw, Medical University of Silesia in *Sosnowiec* and St. Raphael Hospital - Clinical Department of Neurosurgery in Cracow), representing grade III and IV gliomas according to World Health Organization (WHO) classification (Fig. 2.1).

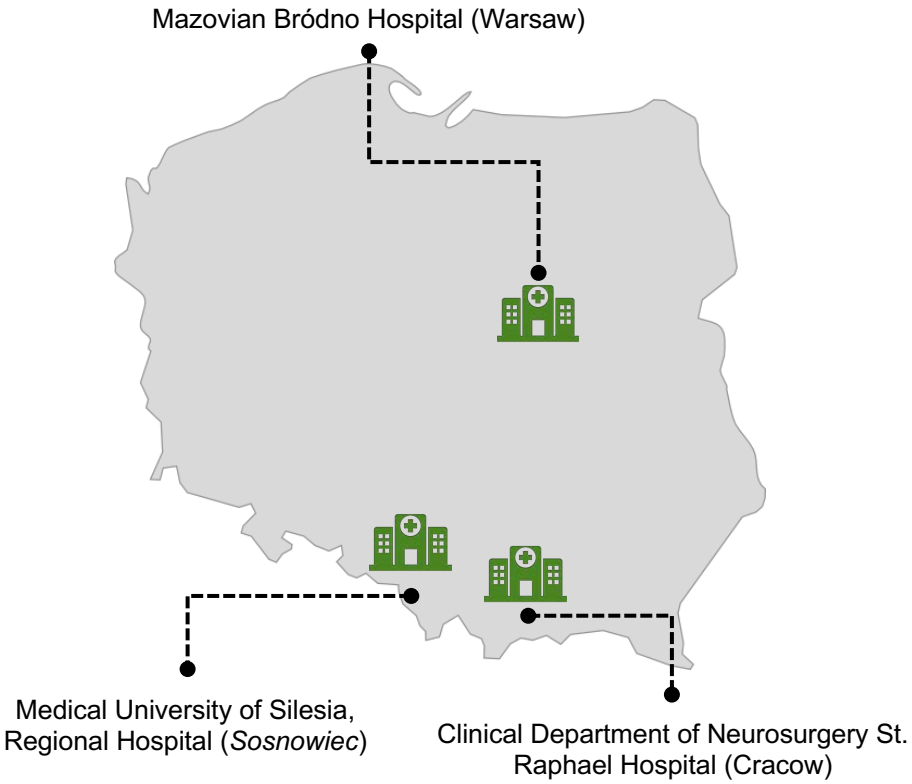


Figure 2.1. A study cohort and collaborating hospitals. The map of Poland and the hospitals which provided fresh frozen high-grade glioma samples post tumour resection of primary and recurring tumours.

DNA and RNA were extracted from each tumour and tumour-matched blood samples were collected from each patient to determine the somatic status of single nucleotide polymorphisms (SNPs), indels, and CNAs. Most of the patients underwent recurrent tumour resection once and there were two cases were patients underwent a 2nd or even a 3rd recurring tumour removal (Fig. 2.2). All patients signed an informed consent for use of their biological material for research purposes. The cohort contained 14 pairs of GBMs and 2 anaplastic astrocytomas WHO grade III. This study was approved by The Bioethics Committees of *Andrzej Frycz Modrzewski* Cracow University, St. Raphael Hospital, Cracow, Poland (Nr. 73/KBL/OIL/2015); Medical University of Silesia, *Sosnowiec*, Poland; *Mazowian Bródno* Hospital, Warsaw, Poland (Nr. KNW/0022/KB1/46/I/16).

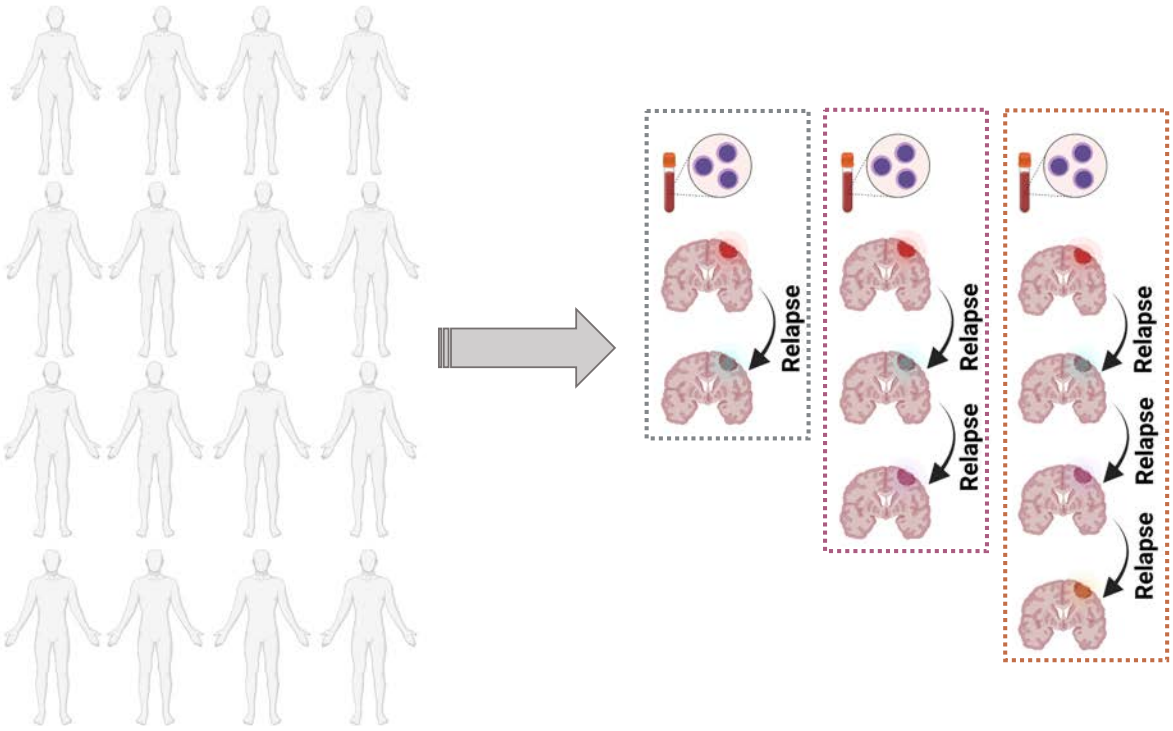


Figure 2.2. A schematic representation of participants and samples collected. Illustration of tissue collection from patients who had brain tumours resected once (14 patients), twice (1 patient), or three times (1 patient), as well as a blood sample per patients was taken. Created with BioRender.

All patients underwent standard Stupp treatment, which included surgery, radiation, and concomitant and adjuvant TMZ. Sixteen samples were original paired tumours, and 19 were recurrent paired tumours; as previously stated, two cases were individuals who underwent second and third resections.

2.1.2. Extraction of genomic DNA and RNA

Total DNA and RNA were extracted from glioma tissue samples using the TRI Reagent (Sigma Aldrich, cat no T9424-100ml), according to the manufacturer's protocol. DNA was additionally purified by phenol-chloroform extraction and precipitated by ethanol. To eliminate protein contamination, extracted DNA was treated with proteinase K (600 g/ ml) before being mixed with an equal volume of phenol-chloroform-isoamyl alcohol. The top aqueous phase was carefully transferred to a fresh tube after centrifugation, extracted with an equal volume of chloroform and centrifuged again. The top aqueous phase was treated with 5M NaCl and DNA was precipitated with 2 volumes of 100% ethanol. Samples were centrifuged, the DNA pellet was washed with 70% ethanol and dried at room temperature (RT) before being resuspended in Milli-Q (MQ) water and frozen at -20°C. The extracted DNA was used for targeted exome sequencing, while the RNA was used for RNA-seq.

2.1.3. Panel design, genomic and transcriptomic library preparation and NGS

We designed a target enrichment DNA sequencing panel with 700 cancer- and epigenetic-related genes and their adjacent regions (called 700 NGS panel for simplicity). The 7 MB ($1x10^6$ base pairs) targeted region covered cancer-related gene locations, with a focus on genes coding for epigenetic regulators (*IDH1/2*, epigenetic enzymes histone modifiers, chromatin modelers, histone chaperones). Isolated DNA from tumour samples was processed for library construction according to the user guide for the NimbleGen SeqCap EZ Library SR (v 4.2). First, 1g of genomic DNA was sheared to obtain fragments of a mean size of 300 bp in a 50 µL volume using the conventional methodology with a Covaris microTUBE screwcap and the Covaris M220 system. The sheared DNA sample was processed into a library using the KAPA Library Preparation Kit for Illumina platforms (KAPABiosystems, KR0935 v2.14), as suggested by the manufacturer. To test the size distribution and library molarity, the final libraries were run on the Agilent BioAnalyzer 2100 with the DNA High Sensitivity Kit (Agilent; 5067-4626) and Quantus Fluorometer (Promega, Madison, WI, USA). Following that, libraries were prepared for Illumina cluster formation and sequencing. The HiSeq 1500 Genome Analyzer was used to perform paired-end sequencing, which resulted in 76 bases from each end of the fragments (Illumina, San Diego, CA, USA). Tumour samples were sequenced at the enriched regions with a predetermined mean coverage of 100, while DNA extracted from whole blood samples was sequenced at a mean coverage of 30.

The quality and integrity of total RNA were determined using an RNA 6000 Nano Kit and an Agilent 2100 Bioanalyzer (Agilent Technologies, Ltd.). Strand-specific polyA enriched RNA libraries were produced according to the manufacturer's procedure using the KAPA Stranded mRNA Sample

Preparation Kit (Kapa Biosystems, MA, USA). Using poly-T oligo-attached magnetic beads, mRNA molecules were enriched from 500ng of total RNA (Kapa Biosystems, MA, USA). Using a reverse transcriptase, mRNA was fragmented and first-strand cDNA was produced. A second cDNA synthesis was carried out to produce double-stranded cDNA (dsDNA). The 3′ ends of dsDNA were adenosine-modified, and adapters were ligated (adapters from NEB, Ipswich, MA, USA). Following adapter ligation, uracil in an adapter loop structure was digested by NEB's USER enzyme (Ipswich, MA, USA). Using NEB starters, adapters containing DNA fragments were amplified by PCR (Ipswich MA, USA). The Agilent 2100 Bioanalyzer with the Agilent DNA High Sensitivity chip was used to evaluate the library (Agilent Technologies, Ltd.) The average library size was 300 bp. A Quantus fluorometer and the QuantiFluor double stranded DNA System were used to quantify the libraries (Promega, Madison, Wisconsin, USA). Libraries were run in the rapid run flow cell, as per the DNA protocol, and were paired-end sequenced (2x76bp) on HiSeq 1500 (Illumina, San Diego, CA 92122 USA).

2.1.4. Bioinformatic pipeline 1: genomic analysis and detection of CNAs

We created a bioinformatic pipeline to process targeted exome sequencing data in the first part of this study. Raw reads from the HiSeq 1500 Illumina sequencer were converted from binary base calls (BCLs) to FASTQ human-readable format in the targeted DNA-sequencing workflow (Fig. 2.3). Then, using Trimmomatic (215) (version 0.36) with default parameter values and paired-end mode, read trimming was applied to FASTQ files to remove Illumina-specific adapters, low quality 5' and 3' bases, and short reads. NextGenMap (216) (version 0.5.2) was used to align DNA sequencing reads to a human reference genome sequence (hg38), with default parameters but using the "*strata*" parameter to output only highest scoring mapped reads.

Following that, Picard Tools (217) (version 2.17.1) was used to remove PCR read duplicates, and only properly orientated and uniquely mapped reads were retained for further analysis. SAMtools (218) (version 1.5) pileup function was used on BAM files to facilitate SNP/indel calling, resulting in a pileup of reads at a single genomic locations. VarScan2 (219) was employed in these pileups to call somatic mutations. A minimum coverage of 10 reads was set for both normal and tumour samples for these somatic calls and pileup obtained from blood samples was utilized as a reference. Furthermore, because variants found in reads that only align to one strand are more likely to be false positives, strand bias variations were excluded, and only damaging coding variants with predicted Sorting Intolerant From Tolerant (SIFT) values (SIFT<0.05) were considered for downstream analysis.

Finally, the processSomatic method from VarScan2 (219) was used to extract high-confidence somatic calls based on variant allele frequency (VAF) and Fisher's exact test p-value (VAF>15%, normal VAF<5%, and a somatic p-value of <0.03). The final subset of variants was annotated with Annovar (220) (2017Jul version) employing the most recent database versions (refGene, clinvar,

cosmic, avsnp150 and dbnsfp30a). The oncodriveCLUST algorithm was used to identify genes in which mutations clustered in large spatial hot-spots, which could provide an adaptive advantage to cancer cells (221).

To infer relative changes in copy number in HGGs, we estimated somatic CNAs using data from matched tumour-normal pairings and SAMtools (version 1.5) and VarScan2 (version 2.4.3), followed by the Circular Binary Segmentation (CBS) algorithm (222). Copynumber and copycaller were employed with the default parameters but with the normal/tumour input data in consideration. At the end, somatic variations (SNPs and indels) and CNVs from Illumina reads were identified from in each patient in the cohort.

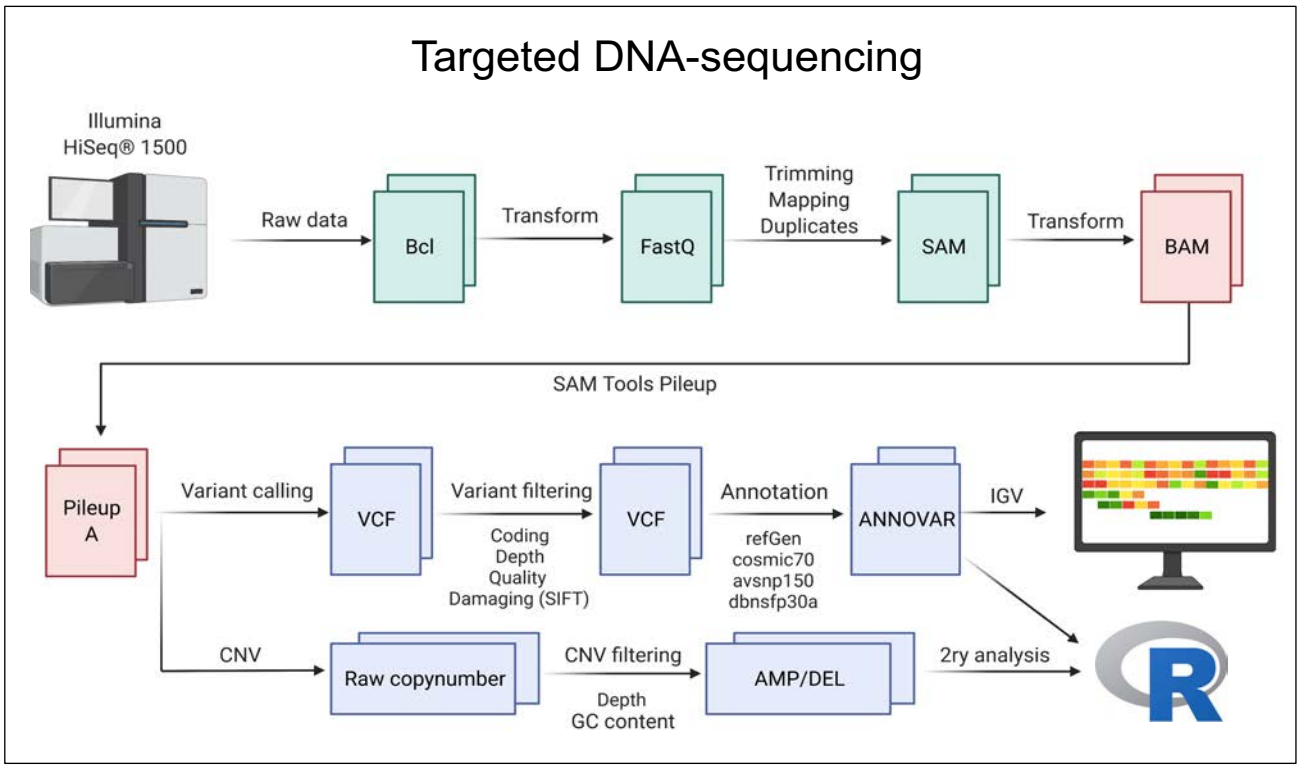


Figure 2.3. Workflow of the primary and secondary bioinformatic analyses of targeted DNA-sequencing. Illustration of data conversion, procedure followed and algorithms and tools utilized. IGV stands for Integrative Genomics Viewer, CNV stands for copy number variations and AMP/DEL stand for amplification and deletions. Created with BioRender.

2.1.5. Bioinformatic pipeline 2: detection of aberrant splicing events and transcriptomic analysis

We established a second bioinformatic pipeline to evaluate, on the one hand, the complete novel splicing events (both donor and acceptor site of splicing do not exist in transcriptome databases), and, on the other, perform classical differential gene expression analysis including *in silico*

single cell analysis (Fig. 2.4). To eliminate Illumina-specific adapters, low quality 5' and 3' bases, and short reads, we utilized the Trimmomatic tool (215) (version 0.36) with default parameter settings and paired-end mode. The resulting RNA sequencing reads were aligned to a human reference genome sequence (hg38) using the STAR aligner (223) (version 2.6) enabling the twopassMode Basic option to improve read mapping in unannotated exon junctions. Read duplicates were marked with Picard Tools MarkDuplicates (version 2.17.1). Junction annotation analysis (junction_annotation.py) from RSeQC (224) (version 2.6.5) was used and UCSC gene model annotations (hg38) were utilized to identify splice junctions in BAM files.

RNA-seq mapped reads were summarized to genes and counted in paired and reverse stranded mode (mode was determined by the method of strand-information preservation – dUTP method used in KAPA solutions) using the featureCounts method (225) (version 1.5.3). After pre-filtering low count genes (sum of the raw read counts <20), DESeq2 (226) (version 3.7) was used to analyse differentially expressed genes using a multi-factor design that included the matched sample status (recurrent/primary) to create an individual baseline for each patient. A Cook's distance cut-off of 0.5 was also employed to identify and exclude gene outliers in the cohort. Concurrently, raw counts were adjusted with a variance stabilizing transformation to better quantify gene expression differences between samples or conditions and to perform a posterior cell enrichment analysis (see section 2.1.6).

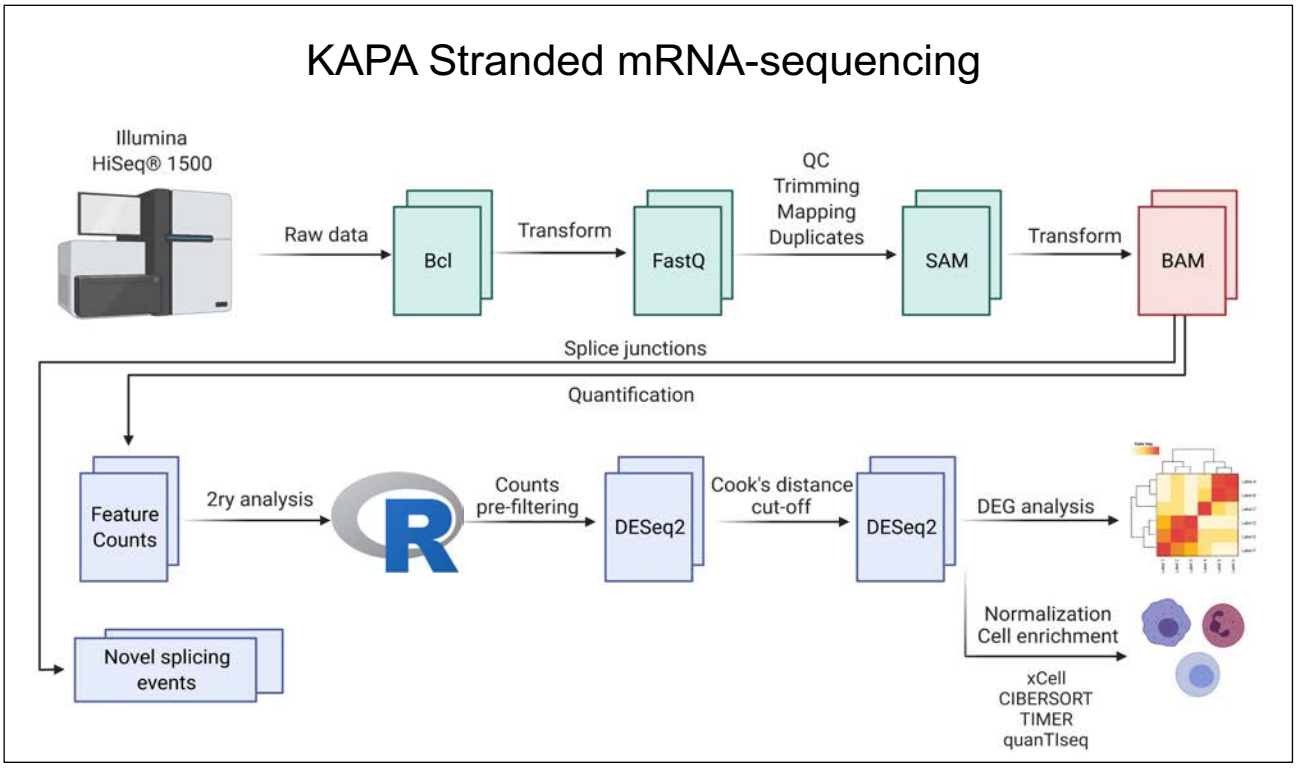


Figure 2.4. Workflow of the primary and secondary bioinformatic analysis of RNA-sequencing. Illustration of data conversion, procedure followed, and algorithms and tools utilized. QC stands for quality control, 2ry stands for secondary and DEG stands for differentially expressed gene. Created with BioRender.

2.1.6. *In silico* cell type enrichment

In order to understand changes in the tumour microenvironment during HGG progression, we performed an *in silico* cell type enrichment analysis using xCell (227) webtool. As a reference signature set, we used the experimentally curated matrix composed of immune and stroma signature cell sets. This final signature set was composed from 6,573 genes corresponding to 64 cell types and was provided by the authors of xCell (227). Weak signatures, which corresponded to signatures with xCell enrichment scores lower than 0.1, considering the score of all patients, were discarded for downstream analysis. To look for statistically significant cell type enrichments, we computed Wilcoxon test between primary and recurrent xCell scores, which were then corrected by the Benjamini-Hochberg (BH) method (padj < 0.1) and resulting cell types were represented. In an effort to validate main xCell findings, we applied an *in silico* marker gene-based approach MCP-counter (228) and three deconvolution methods, CIBERSORT (229), quanTIseq (230) and TIMER (231) to better infer the cellular composition of HGG samples using bulk gene expression data. We used the same rationale as in the xCell approach in order to identify significantly enriched cell types. However, in this case, we wanted to confirm whether specific cells identified in the original xCell analysis were highly enriched in recurrent HGGs, and we did not investigate other cell types.

2.1.7. Immunohistochemical and immunofluorescent characterisation of immune cells

Tumour slides were deparaffinated and hydrated. The pH6 Citrate Antigen Retrieval Solution (DAKO) was used for antigen retrieval, followed by 30 minutes of 10% peroxidase and blocking solution (3% NHS) for 1 hour at RT. A detection system (DAKO) was used as directed by the manufacturer, followed by haematoxylin staining (232). The following IHC antibodies were used at the indicated dilution: anti-CD83 (abcam ab87099, 1:1000), anti-CD163 (abcam 205343, 1:100). Then, immunofluorescence (IF) staining was carried out. After being transferred from -80°C storage, the slides were dried at RT for 1.5 hours. The slides were rinsed three times in PBS for 5 minutes each for rehydration before being blocked for 2 hours at RT with 10% normal serum (donkey) made in 0.1% Triton-X-100 (Tx100)/PBS. The primary antibodies were diluted in 3% serum in PBS (+Tx100) according to the manufacturer's instructions and incubated overnight at 4°C. After removing the primary antibody, the slides were washed three times in PBS for five minutes each, followed by incubation with the appropriate secondary antibody conjugated with the fluorophore, Alexa Fluor® 488 donkey anti-mouse IgG (Invitrogen A21202). The nuclei were stained with DAPI before coverslipping them and mounted with anti-fade fluorescent mounting medium (Dako, USA). Immunofluorescence

quantification was performed using ImageJ (233); all cells in 10 randomly selected fields were counted for each patient, and the percentage of those cells was calculated normalized to DAPI-stained nuclei.

2.2. Multi-omics analysis for TFBS detection in accessible chromatin regions

2.2.1. Human glioma cell lines and surgically resected tumours

Human established GBM LN18 and LN229 cells were obtained from the American Type Culture Collection (ATCC). These cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% foetal bovine serum (ThermoFisher Scientific) and 100 units/mL penicillin and 100 µg/mL streptomycin. Freshly resected glioma specimens were acquired from two neurosurgical hospitals: The Medical University of Warsaw and the Mazovian *Bródno* Hospital. The tissue collection protocol was approved as described in (234). Tumour samples were transported in DMEM/F-12 medium on ice and processed immediately after surgical resection. Tumour samples were transferred to cold PBS, minced with sterile scissors and a scalpel on a Petri dish kept on ice, and then homogenized with a chilled manual glass mincer.

2.2.2. DNA and RNA extraction from glioma samples

Total DNA was isolated from of 5-100 mg of tissues (depending on the starting specimen size) using Tri-Reagent extraction (Sigma-Aldrich, Munich, Germany). NanoDrop 2000 was used to determine DNA purity and screen for potential contaminants in DNA samples (Thermo Scientific, NanoDrop products, Wilmington, USA).

Total RNA from glioma cells was isolated using Qiagen RNeasy kit. Briefly, $1x10^6$ cells were lysed with 350 µl RLT buffer provided by the manufacturer supplemented with 1% B-mercaptoethanol. Extraction procedure has been carried out according to manufacturer instructions. Total RNA was eluted with 25 µl sterile H_20. The concentration of isolated RNA was evaluated by Nanodrop 2000 measurement (Thermo Scientific, NanoDrop products, Wilmington, USA).

2.2.3. ATAC-sequencing

Tumour sample aliquots corresponding to 50-100 mg of tissue were drawn through a syringe needle between 45 and 55 times. Mechanical homogenization was followed by 5 minutes of centrifugation at 2400g at 4°C. Each pellet was resuspended in 10 ml of cold lysis buffer L1 (50 mM HEPES KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 10% glycerol, 5% NP-40, 0.25% Triton X-100, containing proteinase inhibitor cocktail) and shaken for 20 minutes at 4°C. The tissue was then mechanically disrupted again, residual debris was pre-cleared by filtration through an 80-g/mL streptomycin nylon mesh filter, and the lysis buffer was replaced with PBS. Cells were automatically counted using the NucleoCounter NC-100, and 50,000 cells were lysed as previously stated (165). The reactions were filtered using Zymo DNA Clean and Concentrator 5 columns. The steps of the ATAC-seq library preparation were carried out as previously described (165). Finally, ATAC-seq libraries were visualized on a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) and the electropherograms produced were used to estimate DNA concentration. Libraries were run in the Rapid Run flow cell and paired-end sequenced (2x76bp) on HiSeq 1500 (Illumina, San Diego, CA 92122 USA).

2.2.4. ATAC-sequencing data processing

The FastQC tool was used to evaluate the quality of raw FASTQ data (235). After trimming ATAC-seq reads with the FASTQ trimmer (236), only reads with a quality of 10 or higher were considered. Reads with incorrect pairing and a length of less than 20 bp were discarded. Using the default parameters, the Bowtie2 aligner (237) was used to map the reads to the human genome (hg38). Only high-quality reads (MAPQ > 30), correctly paired read mates, and uniquely mapped reads were considered for downstream analysis. PicardTools (217) was also used to find and eliminate PCR duplicates. The following parameters in MACS2 were used to center a 200 bp window on the Tn5 binding site (5' ends of reads represent the cut sites), which is more accurate for ATAC-seq peaks: *--broad --nomodel --shift -100 --extsize 200*. Resulting peaks were then intersected with human ENCODE blacklisted genomic regions to eliminate anomalous and unstructured signals from our NGS experiment (hg38).

2.2.5. Selection of differentially expressed genes between glioma grades

We used TCGA data to find overexpressed genes in GBM (WHO grade IV) compared to gliomas of WHO grade II (GII) using normalized RNA-seq expression values (Fragments Per Kilobase of transcript per Million mapped reads, FPKM). RNA-seq data from 408 glioma patients (248 GII gliomas and 160 GIV gliomas) were analysed. The biotype of genes from RNA-seq data was restricted to protein-coding genes (ensembl 98 annotation). We randomly sampled 20 GII and 20 GIV patients from the normalized count matrix (n=200 times) to maximize statistical power and robustness of the gene selection. The sample function from the base R library (version 3.6.2) was used as a sampling technique, with each sample having an equal chance of being chosen.

We then calculated Student's t-test p-values for all the genes for each of 200 random 20 GII vs 20 GIV comparisons. The p-values obtained from each of the 200 comparisons were then corrected using multiple testing (false discovery rate, FDR), and the means of these adjusted p-values from all of these comparisons were calculated. Only genes that differed significantly between glioma grades (GIV vs. GII) were kept (adjusted p-value means < 0.01). On the same TCGA dataset, we used DESeq2 methods (226) to detect gene expression direction using log fold change (logFC). Overexpressed genes in GBMs were genes that changed significantly based on FDR and had a positive logFC. To investigate the biological significance of our gene selection, we performed pathway enrichment analysis with the ClusterProfiler (238) R library using the Gene Ontology (Biological Processes) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) databases. To confirm the pathway enrichment, we performed a Gene Set Enrichment Analysis (GSEA) on selected DEGs (using the KEGG and Reactome databases) to gain insight into their over-representation in various biological pathways.

2.2.6. Extraction of RNA from human glioma cells and RNA-seq processing

The Qiagen RNeasy kit was used to isolate total RNA from glioma cells. In brief, $1x10^6$ cells were lysed with 350 µl RLT buffer provided by the manufacturer and supplemented with 1% B-mercaptoethanol. The extraction procedure was performed according to the manufacturer's instructions, as detailed in section 2.2.2. The Nanodrop 2000 was used to evaluate potential contaminants and the concentration of isolated RNA (Thermo Scientific, NanoDrop products, Wilmington, USA). We performed RNA-seq to gain insights into TF expression at the mRNA level in both LN18 and LN229 GBM cell lines. The KAPA Stranded mRNA Sample Preparation Kit was used to build polyA enriched RNA libraries (Kapa Biosystems, Wilmington, MA, USA). Trimmomatic (215) (version 0.36) was used with default parameters for the transcriptomic analysis to remove Illumina

adapters and low-quality reads. Then, RNA sequencing reads were aligned to a reference human genome sequence (hg38) with the twopassMode Basic choice enabled in STAR aligner (223) (version 2.6) and all other parameters were set to default. Only properly oriented pairs of reads were considered for downstream analysis. MarkDuplicates from Picard Tools (217) was used to flag read duplicates and obtain optical duplication estimation (version 2.17.1). The RNA-seq mapped reads, in paired and reverse stranded mode, were summarized and counted by genes using featureCounts software (225) (version 1.5.3). At this stage, only genes that were uniquely mapped and had MAPQ mapping quality values of 255 were considered. Finally, raw counts from featureCounts were converted to FPKM values, and genes encoding various TFs were selected for further investigation.

2.2.7. Transcription factor binding prediction using ATAC-seq data

The human HOCOMOCO v11 (version 11) motif database (239) in the MEME motif format was used to find TFs that could potentially bind to promoters with open chromatin regions. Using the FIMO tool (240), position weight matrices (PWMs) were used to scan the human genome (hg38) FASTA file. The background nucleotide frequency from hg38 was used and all motif occurrences with a p-value less than $1e^{-4}$ on both DNA strands were considered. Motifs found on the mitochondrial genome were discarded for the subsequent analysis. Overall, motif occurrences were computed independently for each of the 735 motifs models. The BMO algorithm (241) is an unsupervised method that estimates the likelihood that a given motif instance is bound using a negative binomial model of ATAC-seq fragments and the number of co-occurring motifs. BMO was used to classify TF binding in human GBM cell lines and human GBM samples. For further analysis, only motif instances expected to be bound with adjusted p-values below 0.05 (Benjamini-Yekutieli correction procedure) were used. Only motif instances predicted to be bound at the same chromosomal localization (for a particular TF model) in LN18 and LN229 human GBM cell lines were considered when determining TFs prediction. We intersected the resulting TFBS with the promoters of protein coding genes after selecting motif instances that were common in both glioma cell lines. Transcription start sites (TSSs) and their flanking DNA regions upstream (1.5 kb) and downstream (1.5 kb) were used to identify gene promoters. We chose to keep things simple by only considering one TFBS per promoter, so if a specific TF was predicted to be bound twice in a promoter, we count it as one. By focusing on the top TF regulators, we could determine the relevance of a specific TF and its relationship to gene dysregulation. Finally, the TFBS found in the LN18 and LN229 GBM cell lines was cross-referenced with BMO results from two human GBM samples.

2.2.8. ChIP-sequencing

The QIAseq Ultra Low Input Library Kit was used to create DNA libraries from chromatin immunoprecipitation with the appropriate antibodies (QIAGEN, Hilden, Germany). End-repair DNA was used, adenosines were added to the 3′ ends of dsDNA to create “sticky-ends”, and adapters were ligated (adapters from NEB, Ipswich, MA, USA). Following adapter ligation, uracil was digested in an adapter loop structure by USER enzyme from NEB (Ipswich, MA, USA). Using NEB starters, adapters containing DNA fragments were amplified by PCR (Ipswich MA, USA). The Agilent 2100 Bioanalyzer with the Agilent DNA High Sensitivity chip was used to evaluate the library's quality (Agilent Technologies, Ltd.) To quantify and evaluate the obtained samples, the Nanodrop spectrophotometer (Thermo Scientific, NanoDrop products, Wilmington, USA), Quantus fluorometer (Promega Corporation, Madison, USA), and 2100 Bioanalyzer were used to quantify and evaluate the obtained samples (Agilent Technologies, Santa Clara, USA). The average library size was 300 bp. Libraries were run in the rapid run flow cell and were single-end sequenced (65 bp) in the rapid run flow cell on HiSeq 1500 (Illumina, San Diego, CA 92122 USA).

2.2.9. Comparison of H3K27ac histone modification across glioma grades

We had acquired histone ChIP-seq data from gliomas of different grades from a previous study conducted in our laboratory (234). We focused on activated enhancers from eight diffuse astrocytoma (DA) patients and ten GBMs. We then used DESeq2 method to identify H3K27ac ChIP-seq signal differences within enhancer peaks to better capture the differences in active enhancer marks between glioma grades. First, we filtered out peaks found in only one tumour sample (DA=7 patients, GBM=10 patients), and the resulting peakset was used to count single-end reads from BAM files using the featureCounts (225) tool. Then, we created a DESeq2 object after obtaining counts in each of the glioma enhancers to test for H3K27ac signal differences between GBM and DA tumours, and regions with adjusted p-values 0.05 were considered.

2.2.10. Annotation of glioma enhancers and their association with TFBS

The presence of H3K27ac peaks in non-promoter regions determined the genomic ranges of active enhancers. We started with a set of active enhancers identified in the study conducted in our group (234), which were discovered after analysing H3K27ac ChIP-seq experiments from GBMs, DAs, and PAs tumours. These H3K27ac peaks were found in at least 5 GBM patients and at least 5 DAs

patients across all analysed samples (234). First, we used the ChIPseeker (version 1.28.3) library's peakAnnotation function (242) to additionally pre-filter potential H3K27ac peaks near TSS regions. The resulting set of glioma enhancers was then intersected by chromosomal coordinates with predicted TFBS in glioma cell lines using the tidygenomics R library's genome intersection function (version 0.1.2). At this point, we only considered the TFBS motifs shared between LN18 and LN229 glioma cell lines. Furthermore, we performed an integrative analysis of TFBS motifs in enhancers to model the relationship between each TF model and all distal-regulatory regions. Using the phyper function in R, we calculated probabilities based on the cumulative distribution function of the hypergeometric distribution. The p-values obtained for each of the TF models represent the probability of obtaining the observed number of motif instances or higher within glioma enhancers.

2.2.11. DNA methylation sequencing

EZ DNA Methylation-Lightning Kit was used to bisulfite-convert DNA samples (Zymo Research, Irvine, CA, USA). SeqCap Epi CpGiant Enrichment Kit (Hoffmann-La Roche, Basel, Switzerland) probes were used to enrich each Bisulfite-Converted Sample Library in the predetermined distinct genomic regions of 80.4 Mb capture size, which included 5.6 million CpG sites on both DNA strands. The libraries were created using the "NimbleGen SeqCap Epi Library Workshop Protocol, v1.0" and "SeqCap Epi Enrichment System User's Guide, v1.2" from Hoffmann-La Roche. In brief, genomic DNA concentration was determined using a Quantus Fluorometer with QuantiFluor dsDNA System (Promega, Madison, WI, USA). One µg input DNA, as well as 165 pg Bisulfite-Conversion Control (Lambda phage unmethylated gDNA; SeqCap Epi Accessory Kit; Hoffmann-La Roche) were fragmented using Focused-ultrasonicator Covaris M220 (Covaris, Inc., Woburn, MA, USA) to an average size of 200 ± 20 bp. The DNA fragments were measured on a 2100 Bioanalyzer using the High Sensitivity DNA Kit (Agilent Technologies, Inc., Santa Clara, CA, USA). With the KAPA LTP Library Preparation Kit (KAPA Biosystems, Wilmington, USA), SeqCap Adapter Kit A and B (Hoffmann-La Roche), and DNA purification beads, the DNA fragments were "End-Repaired," "A-tailed," and the index adapters were ligated (Agencourt AMPure XP Beads; SeqCap EZ Pure Capture Bead Kit; Hoffmann-La Roche). Following that, adapter-enhanced DNA fragments were size-selected using Agencourt AMPure XP beads (SeqCap EZ Pure Capture Bead Kit) and Solid Phase Reversible Immobilization technology to exclude DNA fragments larger than 450 and smaller than 250 bp.

Next, the libraries were bisulfite transformed using the Zymo Research EZ DNA Methylation-Lightning Kit and amplified using Pre-Capture Ligation Mediated PCR (LM -PCR). The content and concentrations of Amplified Bisulfite-Converted Sample Libraries were calculated using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and a Quantus using QuantiFluor dsDNA method (Promega) after purification on Agencourt AMPure XP beads (SeqCap EZ Pure

Capture Bead Kit). In addition, the size of DNA fragments was measured on a 2100 Bioanalyzer using the High Sensitivity DNA Kit (Agilent Technologies, Inc.).

Next, 1 µg of each Amplified Bisulfite-Converted Sample Library was hybridized (47 °C, 67 ± 2 h) with probes from SeqCap Epi CpGiant Enrichment Kit (Hoffmann-La Roche), bound to the Capture Beads (SeqCap EZ Pure Capture Bead Kit; Hoffmann-La Roche) and sequentially washed out of contamination and unspecific DNA in buffers of SeqCap Hybridization and Wash Kit (Hoffmann-La Roche). Finally, the Captured Bisulfite-Converted Sample Libraries were amplified in Post-Capture LM-PCR, cleaned up using Agencourt AMPure XP Beads (SeqCap EZ Pure Capture Bead Kit) and the Amplified Captured Bisulfite-Converted Sample Libraries were submitted to the last quality check where the quality and the concentrations of the final libraries were determined using NanoDrop (Thermo Fisher Scientific) and Quantus with QuantiFluor dsDNA System (Promega), respectively. A size of the obtained DNA fragments was also analysed using the High Sensitivity DNA Kit on a 2100 Bioanalyzer (Agilent Technologies, Inc.). Libraries were run in the rapid run flow cell and were paired-end sequenced (2x76 bp) on HiSeq 1500 (Illumina, San Diego, CA 92122 USA) (234).

2.2.12. Analysis of DNA methylation in published glioma datasets

The methylation analysis workflow was carried out using the CytoMeth tool (https://github.com/mdraminski/CytoMeth), which takes FASTQ files as inputs and returns the calculated DNA methylation levels (ß-values) at the bp level. Generally, this automated workflow includes: FastQC (235) to assess read quality, BSMAP (243) to map reads to the hg38 reference genome, Picard Tools (217) to remove PCR duplicates and methratio.py to assess coverage statistics and assign methylation levels returned as beta-values. The minimal bisulphite conversion was set to ~99%. The cytosines in CpG and non-CpG contexts with at least ten reads of coverage were further examined. The analysis covered various glioma samples: GII/GIII (n=4), GIV (n=10) and IDH mutant samples (n=4; three from GII/GIII and one from GIV). In the end, each sample yielded ~3.5 x 10^6 of well-covered cytosines. However, due to DNA degeneration, the total number of cytosines shared by all samples was only 350k. As a result, the following analysis focused on differentially methylated regions rather than individual cytosines.

The DiffMeth module was used to investigate the variability of DNA methylation within promoter regions (+2000/-500bp from TSS). DiffMeth identifies statistically significant differences in the methylation levels of specific DNA regions between defined groups of samples. The first significance criterion of the DiffMeth module was used in this case. It was analysed with a standard $\chi 2$ statistical test, in which all groups were compared to one another (pair by pair). The $\chi 2$ test compares the distribution of beta values assigned to predefined ranges reflecting hypo-, medium-, and hyper-

methylated cytosines: [0.0-0.2], [0.2-0.6], and [0.6-1.0], and the p-values obtained are corrected with FDR. In this case, the null hypothesis was rejected when FDR<0.05.

In addition, to avoid the issue of too many beta values aggregations from long DNA sequences such as promoters, the DiffMeth module detects a short C-rich region and applies the previously described statistical pipeline to short-regions rather than whole promoters. DiffMeth was set to detect short regions of similar length to TFBS with a median length of 22 bp. Because the same set of c-Jun target promoters differed in DNA methylation levels when computed for all cytosines (CpG and non-CpG) as well as just cytosines in the CpG context, only the results for cytosines in the CpG context were reported.

2.2.13. Analysis of DNA methylation in TCGA data

The p sites, selected from c-Jun targeted genes regulatory regions (as described in 2.2.7 and 2.2.14 sections), were intersected with promoter regions defined as TSS ± 1 kb with the annotation for the hg19 human genome obtained from feb2014.archive.ensembl.org. TCGA GBM and LGG 450k DNA methylation datasets were downloaded from gdac.broadinstitute.org. For each defined promoter region, the median beta-values of DNA methylation were calculated per each sample. FPKM-normalized TCGA data was also uploaded, and Pearson correlation was calculated for selected genes for samples with matching DNA methylation and RNA-seq data.

Furthermore, in the context of DNA methylation in glioma enhancers, we searched the TCGA for information on DNA methylation in these enhancers and used the available CpG (cg02258482, cg12155676 and cg08003402). Specifically, we focused on the c-Jun binding on these enhancers ±20 bp flanking regions.

2.2.14. Distal-range intra-chromosomal contacts between glioma enhancers and gene promoters

In order to better describe enhancers, we used Hi-C data from high-resolution 3D maps of chromatin contacts from developing human brains (171) to identify enhancer–promoter contacts that might function in gliomas. We searched for contacts of common active enhancers using Hi-C data and found that 5,530 of the 10,673 common enhancers had significant contacts in our samples with any region within the 2 Mb range (234). We, then, found enhancer-promoter contacts for those common active glioma enhancers that contained at least one c-Jun TFBS. Each promoter was then assigned to the nearest gene.

2.2.15. Human survival analyses of c-Jun and c-Jun's targets

The TCGA data was used for survival analyses. Depending on the type of analysis, only patients with GBM or lower-grade gliomas were chosen from the cohort. Patients were divided into two subgroups based on their level of c-Jun target expression (high mRNA and low mRNA levels). Log-rank tests were used to confirm the association between expression levels of c-Jun targets and patient survival. Finally, Kaplan-Maier plots were computed for each of the genes that could be controlled by c-Jun, and censored patient's data were included in the analyses.

2.2.16. Gene expression profiling in Pan-cancer and paired normal tissues

In order to determine *JUN*'s expression in different human cancers, including in brain tumours, we used the Gene Expression Profiling Interactive Analysis (GEPIA2). Transcripts per million (TPM) values were taken from the different TCGA and Genotype-Tissue Expression (GTEx) datasets and brought together in the form of median expression in each cancer and in the paired normal tissues that were extracted, and used as an input to R. In the presented analysis, the following cancers and corresponding healthy tissues were examined: Adrenocortical carcinoma (ACC) and Adrenal Gland; Bladder Urothelial Carcinoma (BCLA) and bladder; Breast invasive carcinoma (BRCA) and breast; Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) and cervix uteri; Colon adenocarcinoma (COAD) and colon; Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC) and blood, Oesophageal carcinoma (ESCA) and oesophagus; Glioblastoma multiforme (GBM) and brain; Kidney Chromophobe (KICH) and kidney; Kidney renal clear cell carcinoma (KIRC) and kidney; Kidney renal papillary cell carcinoma (KIRP) and kidney; Acute Myeloid Leukaemia (LAML) and bone marrow; Brain Lower Grade Glioma (LGG) and brain; Liver hepatocellular carcinoma (LIHC) and liver; Lung adenocarcinoma (LUAD) and lung; Lung squamous cell carcinoma (LUSC) and lung; Ovarian serous cystadenocarcinoma (OV) and ovary; Pancreatic adenocarcinoma (PAAD) and pancreas; Prostate adenocarcinoma (PRAD) Prostate, Rectum adenocarcinoma (READ) and colon; Skin Cutaneous Melanoma (SKCM) and skin; Stomach adenocarcinoma (STAD) and stomach; Testicular Germ Cell Tumours (TGCT) and testis; Thyroid carcinoma (THCA) and thyroid; Thymoma (THYM) and blood; Uterine Corpus Endometrial Carcinoma (UCEC) and uterus; Uterine Carcinosarcoma (UCS) and uterus.

The JUN mRNA expression profile was compared between tumour samples (TCGA) and paired normal tissues (TCGA normal + GTEx normal), and statistical significance was determined using one-way ANOVA and disease state (Tumour versus healthy tissue of tumour origin).

2.2.17. Cell culture, nuclear extracts and Electrophoretic Mobility Shift Analysis (EMSA)

Glioma LN18 and LN229 cells were cultured in a DMEM medium. Patient-derived GBM cell cultures (WG12) were set up as described (29) and cultured in a DMEM/F12 GlutaMAX medium. Normal human astrocytes (NHA) (Lonza Walkersville, USA) were cultured in a commercial medium as described in (244). All media were supplemented with 10% FBS (Gibco, USA) and antibiotics (100 U/mL penicillin, 100 µg/mL streptomycin), then cells were cultured in a humidified atmosphere of CO_2/air (5%/95%) at 37 °C.

For the EMSA probe we used the oligonucleotide containing the c-Jun motif: G-T-G-A-G-T-C-A-C-C from the human *VIMENTIN* promoter. Oligonucleotides: biotin-labelled c-Junbiof: C-A-G-G-G-C-G-C-G-G-T-G-A-G-T-C-A-C-C-G-C-C-G-G-T-G-A-C-T-A-A-G-3' and unlabelled c-Junr: 5'-C-T-T-A-G-T-C-A-C-C-G-G-C-G-G-T-G-A-C-T-C-A-C-C-G-C-G-C-C-C-3', and c-Junf: 5'-C-A-G-G-G-C-G-C-G-G-T-G-A-G-T-C-A-C-C-G-C-C-G-G-T-G-A-C-T-A-3' were purchased from Metabion. Oligos were dissolved in water, heated to 90^0C and let to anneal for 30 min.

Nuclear extracts were prepared using a nuclear extraction kit: NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Scientific cat. no. 78833) according to the manufacturer's instructions. Protein concentration was measured using THERMO Labsystems Multiscan EX at wave length 570 nm using Bradford Reagent (Sigma Life Science cat no. B6916) and a bovine serum albumin standard (Thermo Scientific cat no. 23209) for calibration.

EMSA was performed using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific cat. no. 20148) according to manufacturer's instructions. Binding reactions contained: 40 fmol dsDNA, 5µg of protein nuclear extracts and 10 mM Tris pH 7.5 buffer with 50 mM KCl, 1 mM DTT, 1.5 mM $MgCl_2$, 1.5 µg Poly (dI-dC). DNA binding reactions were performed in 30 µl. The reaction mixtures were incubated for 30 min at RT and subjected to electrophoresis (70 V, 8°C) in 6% polyacrylamide gels with 10% glycerol and Tris–borate–EDTA buffer. Then, electrophoretically separated material was transferred onto a 0.45 µm Biodyne nylon membrane (Thermo Scientific cat. no. 77016) in Tris–borate–EDTA buffer and detected by chemiluminescence using Chemidoc camera (Bio-Rad).

2.3. Identification of changes in chromatin accessibility in SMARCA2 and SMARCA4 knock down in GBM cell lines

2.3.1. Cell culture and maintenance

Human LN18 GBM cells were obtained from American Type Culture Collection provided (ATCC). LN18 cells were grown in DMEM (Gibco, Invitrogen). Both basal media were supplemented with 10% foetal bovine serum (Gibco, Invitrogen), 100 units/mL penicillin, and 100 g/mL streptomycin. Cells were seeded at a density of 10^5 cells/cm2 in 96-well plates (for MTT assay), 24-well plates (for BrdU incorporation assays), or 12-well plates (for silencing assay or Western blotting). Cells were allowed to grow as adherent cell cultures for 24 hours before each experiment.

2.3.2. SMARCA2 and SMARCA4 silencing

LN18 cells were plated in 12-well plates for 24 hours to achieve 80-90 % confluence before the transfection assays. Two siRNAs (20 ng of ON-TARGETplus siRNA, DharmaCon, GE) against *SMARCA2* (a, #J-017253-05-0002 and b, #J-017253-06-0002) and *SMARCA4* (a, #J-010431-05-0002 and b, #J-010431-06-0002) were used individually or in combination. As an internal control, 20 ng of ON-TARGETplus Non-targeting Pool (#D-001810-10-05) was used. The LN18 cells were transfected with Lipofectamine 2000 (Thermo Fisher Scien tific) and DharmaFECT 1 (DharmaCon, GE), respectively, according to the manufacturer's protocol. Briefly, the medium was removed and the wells were filled with an Opti-MEM I reduced serum media (Thermo Fisher Scientific). The transfection mix included Opti-MEM, Lipofectamine 2000, or DharmaFECT 1, as well as the siRNA of interest. After adding this mixture to the wells, the plates were incubated for 6h before adding fresh, full media to the cells. Forty-eight hours after transfection, the cells were harvested and used for RNA or protein isolation.

2.3.3. Quantitative real-time PCR

Total RNA was isolated from siRNA treated and untreated cells using the Roche High Pure RNA Isolation Kit and used as a template to synthesize cDNA by combining oligo(dT)15 primers (2.5 mmol/L) with 200 units of SuperScript III Reverse transcriptase (Invitrogen). Real-time PCR amplifications were performed in duplicates in a 20 µL reaction volume containing cDNA, TaqMan Universal Master Mix II, no UNG (Uracil-DNA glycosylases, Thermo Fisher Scientific), and TaqMan

probes complementary to *SMARCA4* (#Hs00231324 m1), *SMARCA2* (#Hs01030846 m1) or *GAPDH* (#Hs02758991 g1, an internal control), using the QuantStudioTM 12K Flex Real-Time PCR System and software (Thermo Fisher Scientific). Data were analysed using the relative quantification ($^{\Delta\Delta}$Ct) method. The amount of target mRNA was first normalized to the level of expression of the *GAPDH* gene, which was amplified from the same sample, and then to untreated controls.

2.3.4. Assaying cell viability using MTT assay

The conversion of MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, final concentration of 0.5 mg/mL) to formazan in living cells was used to determine cell viability. LN18 cells were seeded at a density of 4×10^3 cells per well in 96-well plates. The cells were transfected with the appropriate siRNA for 24 hours, the media was changed after 6 hours, and MTT (Sigma-Aldrich) was added to the wells 48 hours later. The resulting formazan crystals were dissolved by adding 200 µl of DMSO (Sigma-Aldrich) to the wells after 2 hours of incubation with MTT at 37°C. A spectrophotometer was used to measure optical density at 570nm. The experiments were repeated three times in duplicates.

2.3.5. Cell proliferation assay using BrdU assay

Cell Proliferation BrdU Kit (Roche) was used to measure glioma cell proliferation in 24-well plates seeded at a density of 5×10^5 cells per well. After adding BrdU for 2 hours, the rate of incorporation was measured according to the manufacturer's protocol. The experiments were repeated three times in duplicates.

2.3.6. Isolation of proteins and Western Blot

After 48 hours of gene silencing, the cells were harvested. The wells were rinsed with 1x PBS before being scraped into a buffer containing protease and phosphatase inhibitors [Tris HCl pH 6.8 (20 mM), Sodium Chloride (137 mM), β-glycerophosphate (25 mM), Sodium Pyrophosphate (2 mM), EDTA (2 mM), Sodium Orthovanadate (1 mM), Triton X-100 (1%), Glycerol (10%), Leupeptin (5 µg/ml), Aprotinin (5 µg/ml), Benzamidine (2 mM), DTT (0.5 mM), and PMSF (1 mM)]. Following this, the protein concentration was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Protein extracts were resolved on SDS-PAGE before electrophoretic transfer onto a

nitrocellulose membrane. Antibodies recognizing SMARCA2 (#ab15597, diluted 1:4000), and SMARCA4 (#ab4081, diluted 1:4000) were both purchased from Abcam. Horseradish peroxidase-conjugated anti-rabbit IgG (#PI-1000 diluted 1:10000), and horseradish peroxidase-conjugated anti-mouse IgG (#PI-2000 diluted 1:10000) were obtained from Vector Laboratories. The immunocomplexes were visualized by using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific). GAPDH detection using anti-GAPDH antibody (#MAB374, diluted 1:1000) was used as a loading control. The molecular weight of proteins was estimated with pre-stained protein markers (Cozy™ Prestained Protein Ladder, HighQu, Germany). Densitometry analysis was performed using ImageJ software.

2.3.7. ATAC-sequencing

LN18 cells from treated and control groups were trypsinised and counted. Following that, 20,000 cells were lysed as previously described (165). The transposition reaction was then carried out using Illumina's Nextera DNA Library Preparation kit (165). Zymo DNA Clean and Concentrator 5 columns were used to clean up the reactions. ATAC-seq library preparation was conducted as previously described (165). The ATAC-seq libraries were then visualized on a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA), and the chromatograms generated were used to estimate DNA concentration.

2.3.8. *SMARCA2* and *SMARCA4* gene expression in TCGA

TCGA GBM and LGG level 3 RNA-seq data (aligned by Tophat2 and gene expression counted by HTseq) were uploaded to R. Gene expression levels as FPKM were extracted for *SMARCA2* and *SMARCA4* genes. Visualization of gene expression differences between grades was done in R.

2.3.9. Experimental statistical analyses

Using all GII, GIII, and GIV samples, gene expression of *SMARCA2* and *SMARCA4* was analysed using ANOVA with no assumption of equal variance (Welch one-way test). FDR correction was applied and any significant differences in mean values between GII, GIII, and GIV were determined and denoted as significant using the following annotations: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001. All experiments were carried out in duplicates or triplicates and were

repeated three times. The standard error of the mean is used to express numerical results. The one-way ANOVA test was used to perform statistical analysis of western blot densitometry, and significant differences were denoted as *p value < 0.05, **p value < 0.01, ***p value < 0.001.

2.3.10. ATAC-seq data processing in SMARCA2/4 deficient cells

FASTQ file quality was checked using the FastQC (235) tool, and adapter sequences, as well as low-quality bases, were removed using Trimmomatic (215). Bowtie2 (237) was used to map the resulting reads to the human genome (hg38) using the following settings in order to achieve a sensitive and accurate mapping while minimizing time: *bowtie2 --local --very-sensitive-local*. To improve biological reproducibility, PCR duplicates were identified and removed using Picard Tools' MarkDuplicates method (217), and only mapping quality MAPQ > 10 and properly and uniquely paired reads were selected using SAMtools (218). Then, mitochondrial DNA reads were discarded, as were ENCODE blacklisted regions.

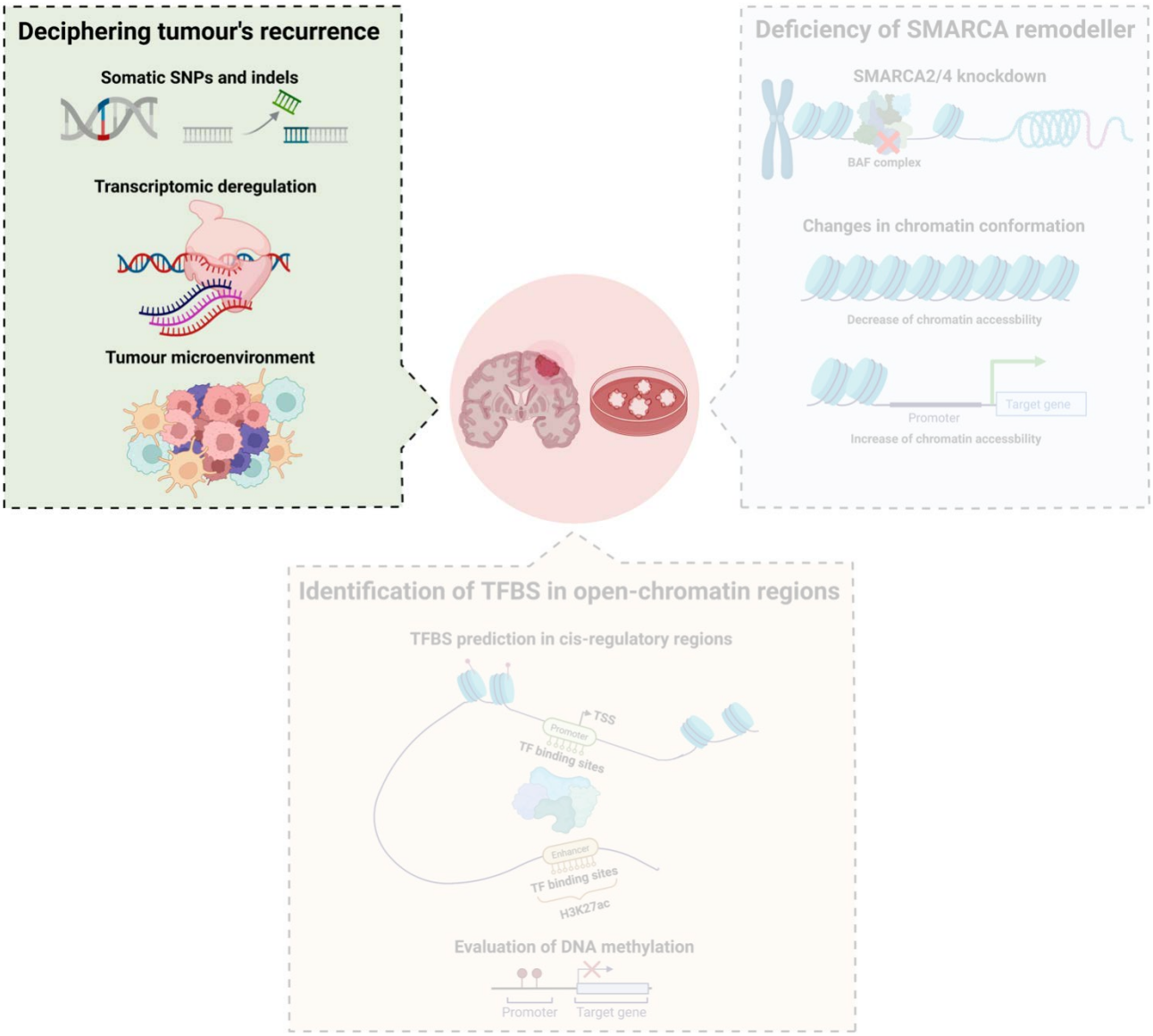
ATAC-seq fragment size distribution was calculated using the ATACseqQC tool (245) as part of the ATAC-seq post-alignment quality evaluation to distinguish nucleosome-free regions (NFR) peaks, which are expected to be enriched around transcription start sites of genes, and mono- and di-nucleosomes (200 bp, 400 bp), which are enriched in flanking or distal regions around TSS. TSS profiles were calculated and plotted using ChIPseeker (242), and the total pool of ATAC-seq fragments was considered for downstream analysis. ATAC-seq peaks were called using MACS2 (246) separately for each replicate and with the following settings to visualize the chromatin accessibility landscape using the information from whole read fragments without modelling or artificial extension: *macs2 callpeak -t -f BAMPE -g hs -B -q 0.05*. Only peaks found in both replicates and uniquely identifying in mockCTRL/siCTRL or SMARCA4/2 knockdown groups were considered.

To better capture the differences in chromatin openness between groups, we used DESeq2 to identify ATAC-seq signal differences within peaks. First, we filtered out not consistent peaks (peaks identified only in 1 sample); the resulting peakset was used to count read fragments in each of the samples using BAM files with the *isPairedEnd = T* and *maxFragLength = 100* parameters of FeatureCounts (225), which defines the paired-end experiment and only quantifies nucleosome-free regions. After obtaining fragment counts in nucleosome free regions, we created a DESeq2 object to test for ATAC-seq signal differences between groups, and we then annotated identified differential ATAC-seq regions to genes using the ChIPseeker (242) and rGREAT (247) tools. Lastly, enrichGO method from clusterProfiler (238) was used to obtain Gene-Ontology - Biological Processes - terms.

Chapter III

3. Results

3.1. Somatic mutations and copy number aberrations in primary and recurrent HGGs



3.1.1. Somatic mutational landscape in progression

Mutational landscape covering the 700 NGS panel was analysed in 35 high grade glioma samples obtained from 16 patients after resection of a primary tumour lesion and after tumour recurrence, which in two patients occurred more than once. In both primary and recurrent tumour samples, an average of 20 high-confidence somatic mutations were discovered per sample (Table 3.1). Using data form 16 matched pairs of primary and recurrent HGGs we wanted to assess how many mutations detected in an initial lesion remained after the tumour relapsed. A comparison of somatic mutations found in primary and recurrent tumours revealed that in the majority of patients (14/16, 87.5 percent) there were more somatic variants unique either to a primary or recurrent tumour sample than those shared between the two disease stages (Fig. 3.1A).

Table 3.1: Clinical information from the studied cohort. [a]High-grade glioma type where GBM indicates glioblastoma, AA indicates anaplastic astrocytoma and AOD indicates anaplastic oligodendroglioma. [b]Tumour localization where *sin* denotes *sinister* (left) and *dex* denotes *dexter* (right). [c]Survival time after initial diagnosis. [d]High-confidence somatic variants. [e]Adjuvant therapy where R indicates radiotherapy, CTX indicates chemotherapy and DEX indicates dexamethasone.

Patient	Diagnosis[a]	Age	Sex	Localization[b]	Type	Surv.[c]	Vars.[d]	Therapy[e]
PG1	AA	26	F	frontalis dex	primary	46	16	none
R1G1	AA	26	F	frontalis dex	recurrent	46	19	none
R2G1	AA	26	F	frontalis dex	recurrent	46	23	none
PG2	GBM	60	F	frontal dex	primary	14	78	R, CTX, DEX
R1G2	GBM	60	F	frontal dex	recurrent	14	15	R, CTX, DEX
PG3	GBM	55	F	frontal lobe	primary	16	39	R, DEX
R1G3	GBM	56	F	frontal lobe	recurrent	16	42	R, DEX
PG4	GBM	64	M	parietal dex	primary	12	35	R, CTX, DEX
R1G4	GBM	64	M	temporal dex	recurrent	12	13	R, CTX, DEX
PG5	GBM	59	M	parietooccipital sin	primary	22	11	none
R1G5	GBM	59	M	parietooccipital sin	recurrent	22	69	none
R2G5	GBM	59	M	parietooccipital sin	recurrent	22	7	none
R3G5	GBM	59	M	parietooccipital sin	recurrent	22	35	none
PG6	GBM	62	F	temporal sin	primary	10	33	R, CTX, DEX
R1G6	GBM	62	F	temporal sin	recurrent	10	11	R, CTX, DEX
PG7	GBM	34	F	parietal dex	primary	23	12	R, CTX, DEX
R1G7	GBM	34	F	parietal dex	recurrent	23	19	R, CTX, DEX
PG8	GBM	50	M	temporal sin	primary	13	16	R, CTX, DEX
R1G8	GBM	50	M	temporal sin	recurrent	13	38	R, CTX, DEX
PG9	AOD	43	F	frontal sin	primary	23	5	R, CTX, DEX
R1G9	AOD	43	F	frontal sin	recurrent	23	5	R, CTX, DEX
PG10	GBM	60	M	temporal dex	primary	16	16	R, CTX, DEX
R1G10	GBM	60	M	temporal dex	recurrent	16	13	R, CTX, DEX
PG11	GBM	44	M	parietotemporal sin	primary	19	11	R, CTX, DEX
R1G11	GBM	44	M	parietotemporal sin	recurrent	19	6	R, CTX, DEX
PG12	GBM	47	F	temporal sin	primary	26	11	R, CTX, DEX
R1G12	GBM	47	F	temporal sin	recurrent	26	11	R, CTX, DEX
PG13	GBM	71	F	temporalis dex	primary	16	15	R, CTX, DEX
R1G13	GBM	71	F	temporalis dex	recurrent	16	19	R, CTX, DEX
PG14	GBM	70	F	frontal dex	primary	8	28	R, CTX, DEX

R1G14	GBM	70	F	frontal dex	recurrent	8	22	R, CTX, DEX
PG15	GBM	44	F	temporal sin	primary	14	6	R, CTX, DEX
R1G15	GBM	44	F	temporal sin	recurrent	14	11	R, CTX, DEX
PG16	GBM	77	F	parietal dex	primary	5	19	R, DEX
R1G16	GBM	77	F	parietal dex	recurrent	5	17	R, DEX

In the genomic somatic analysis, which included non-synonymous mutations that could directly affect protein structure, we discovered that *TP53* (26%) was the most frequently altered gene in both primary and recurrent HGGs, followed by *PTEN* (23%), *PIK3R1* (20%), and *IDH1* (17%) (Fig. 3.1B). Other less frequently mutated genes in the cohort included ATRX (11%), *EGFR* (11%), and *PIK3CA* (11%), which is consistent with previous research (16,67). Interestingly, we discovered a specific frame-shift mutation in the *ZNF384* gene in four HGG samples (Fig. 3.2A).

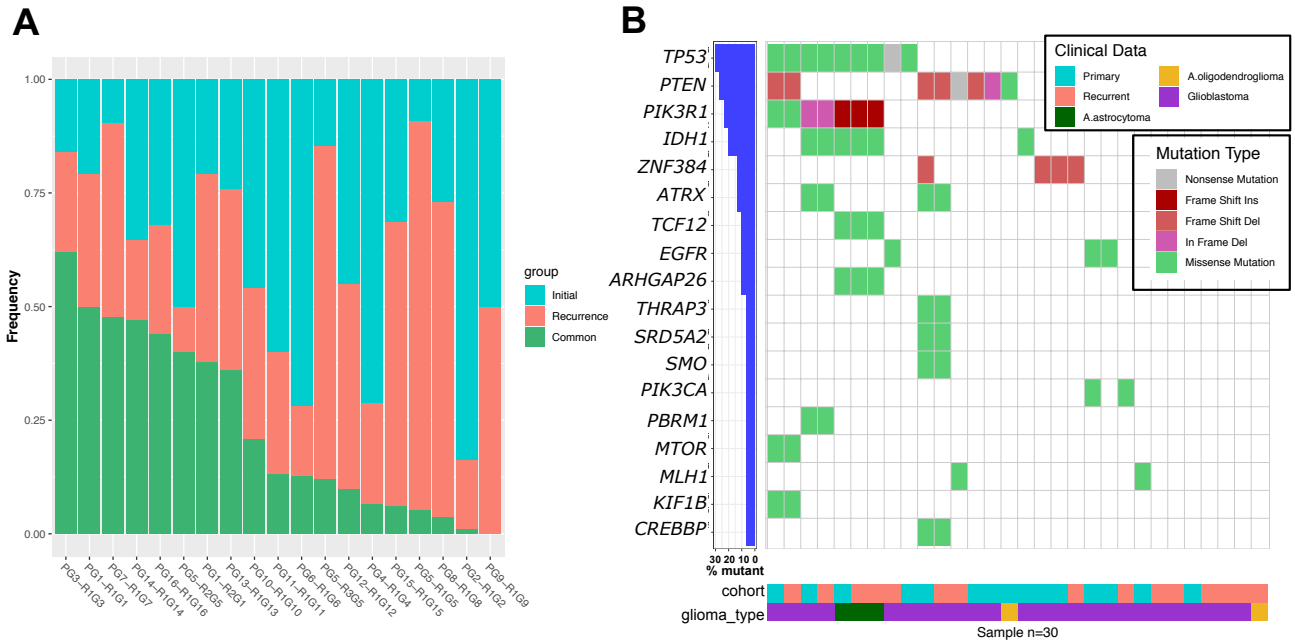


Figure 3.1. Mutational landscape of high-grade gliomas. (A) Frequency of high-confidence nonsynonymous somatic mutations using empirically-derived criteria from VarScan2 (tumour VAF>15%, normal VAF<5%, somatic p-value of <0.03) occurring exclusively in primary HGGs (cyan), exclusively in recurrent HGGs (salmon) and in both (green). (B) Mutation waterfall plot showing somatic mutation types and status found in at least 2 samples and with SIFT values < 0.05 in the case of SNP. Types of somatic alterations are ordered by the frequency of the occurrence in the studied cohort.

The *ZNF384* gene encodes a C2H2-type zinc finger protein that acts as a TF for extracellular matrix genes (248). Although the detected somatic mutation is outside of any domain of the protein (Fig. 3.2A), it could have the potential to affect protein stability. Furthermore, we discovered that high *ZNF384* expression was inversely related to patient survival in the TCGA-GBM/LGG dataset (Fig. 3.2B), indicating the importance of an in-frame mutations in this gene that could lead to gene expression dysregulation.

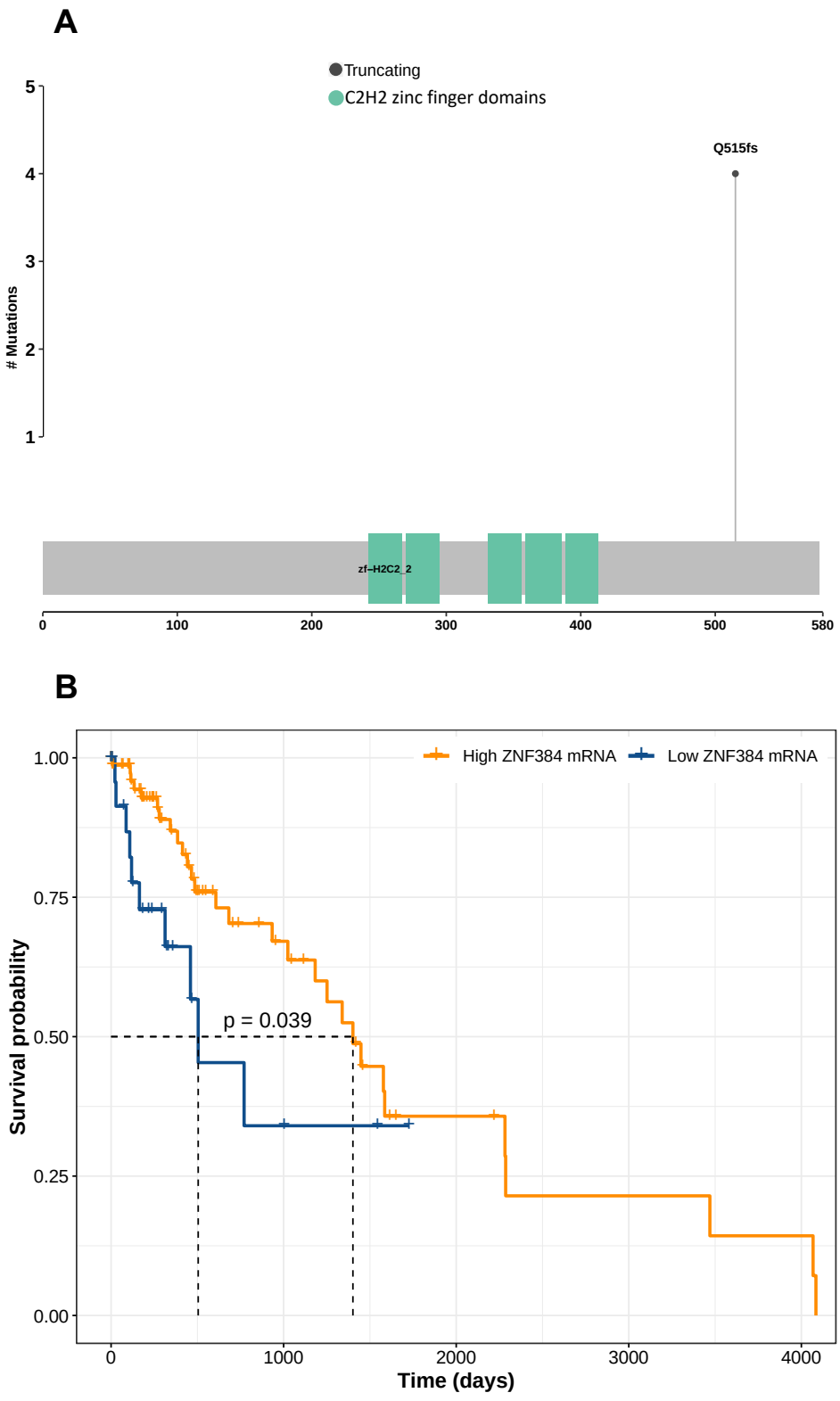


Figure 3.2. Genetic alterations in the *ZNF384* gene and patient's survival depending on its expression. (A) Mutational hot-spot and its effect on the ZNF384 protein structure, where domains and aminoacidic changes are indicated. (B) Kaplan-Meier survival analysis between High-grade gliomas (HGGs) with low- and high-expression of the *ZNF384* (TCGA-GBM/LGG data). Dashed lines represent the median survival for each of the groups. Log Rank Test was used for a statistical analysis.

The most frequent variant type in the cohort was a missense mutation where a cytosine (C) was substituted by a thymine (T) (Fig. 3.3A). Concurrently, we detected a number of transitions (Ti) and transversions (Tv) in both primary (Fig. 3.3B) and recurrent HGG groups (Fig. 3.3C). Tv are more likely to alter the amino acid sequence of proteins due to larger changes in the shape of the DNA backbone with a bigger impact on regulatory DNA (249). We noticed that the ratio of Ti/Tv slightly decreases in recurrent HGG; however, this change was negligible.

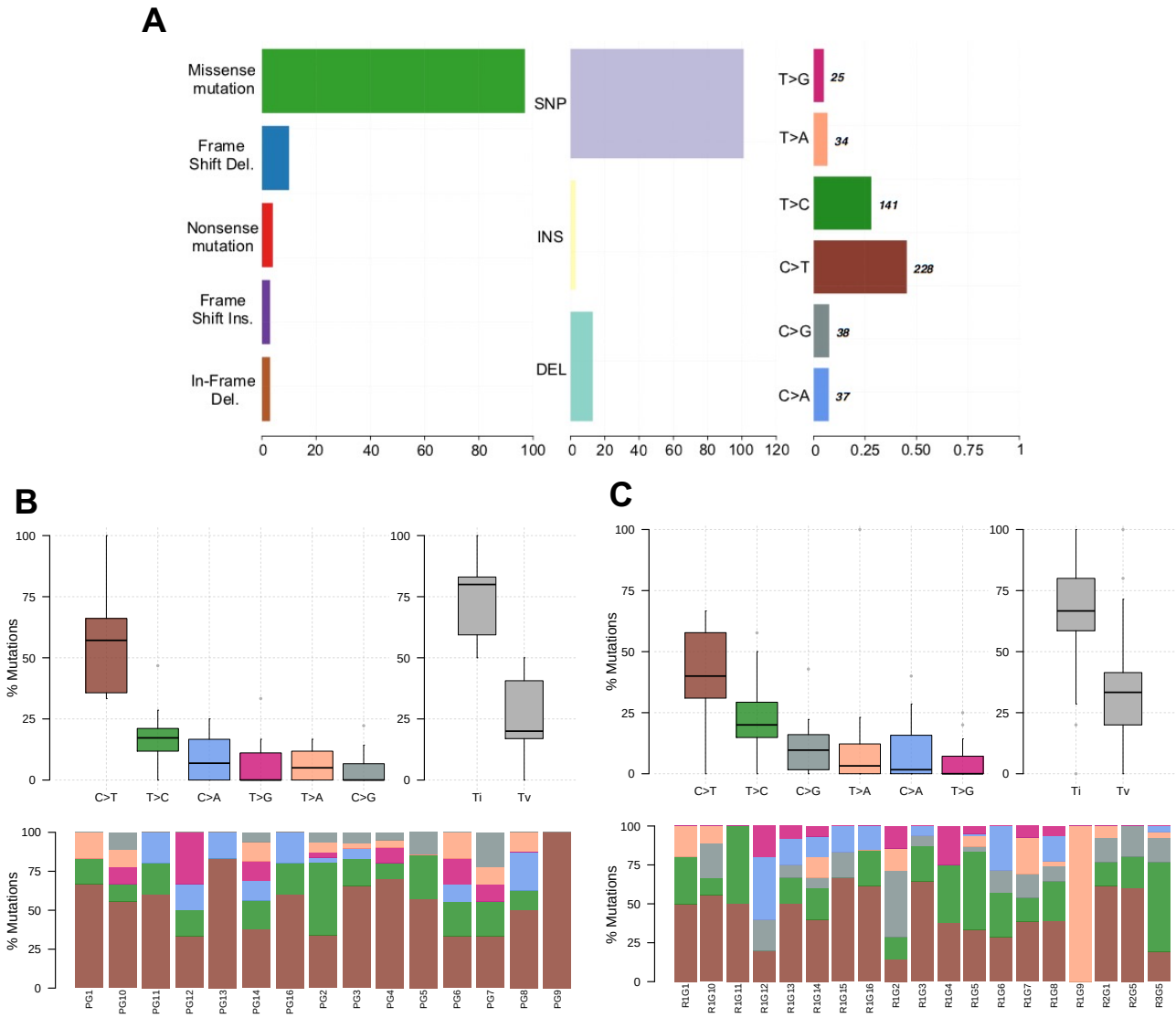


Figure 3.3. Somatic alterations of various types. (A) A summary plot depicting the most common type of somatic mutations and nucleotide substitution across all cohorts. Boxplots depicting the distribution of nucleotide conversions across primary (B) and recurrent (C) HGG samples, as well as the overall transition and transversion frequencies. Data for each individual patient are depicted individually in the bottom stacked bar plots.

We used the oncodriveCLUST algorithm to identify spatial clustering hot-spots that could provide an adaptive advantage to tumour cells and consequently, a positive selection during the clonal tumour evolution (221). The results demonstrated that *TP53*, *IDH1* and *PIK3R1* are the most

frequently mutated genes, harbouring significantly clustered mutations corresponding to specific protein regions (Fig. 3.4A). The same analysis performed on primary and recurrent HGG samples as independent cohorts, revealed the *ZNF384* gene as a potential novel candidate driver gene in primary HGGs cohort, as the specific mutation clustering was detected only in primary tumours (Fig. 3.4B). Gene hot-spots in the recurrent cohort were similar to the entire cohort (Fig. 3.4C).

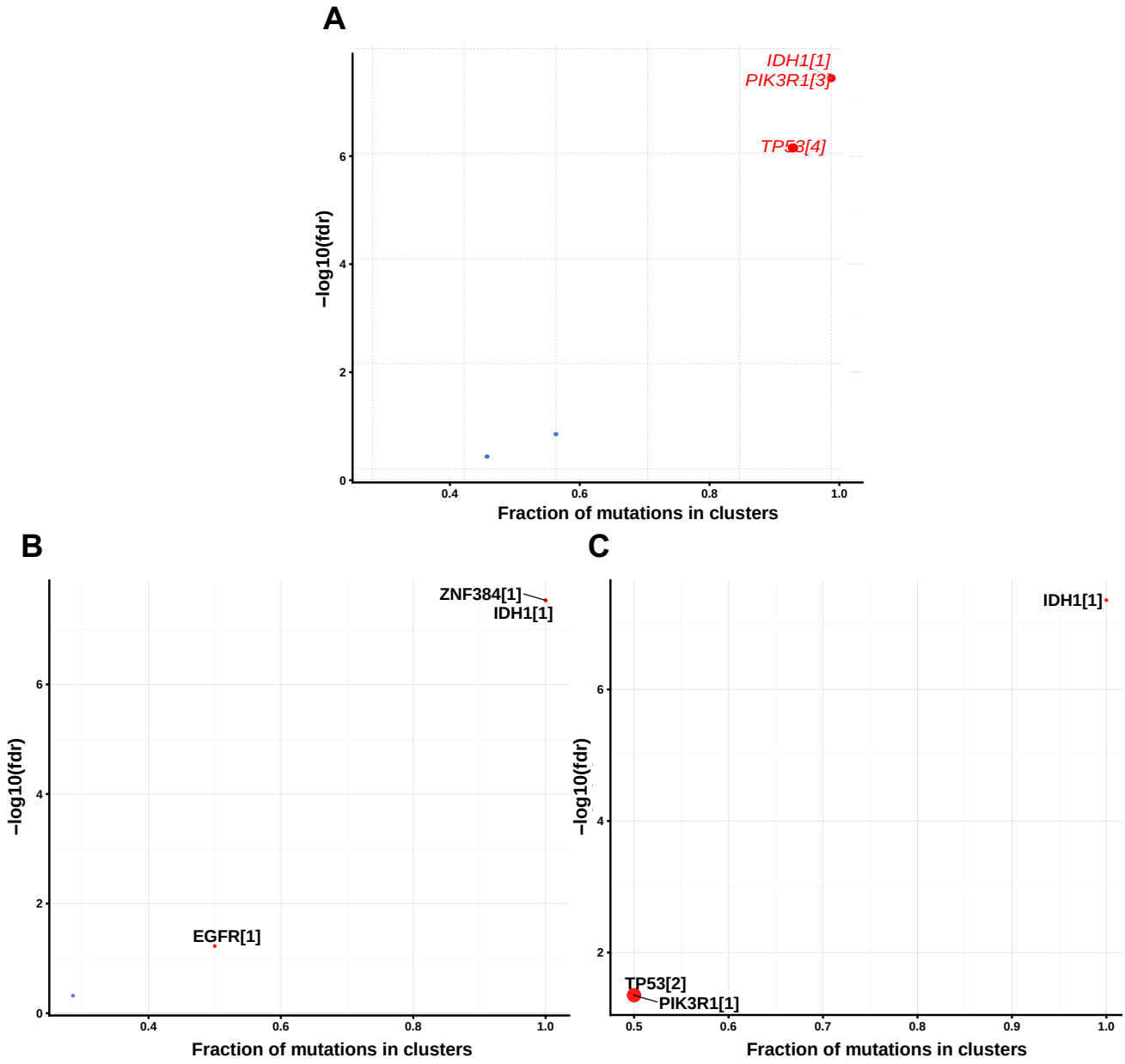


Figure 3.4. Onco-driver gene identification using spatial clustering. (A) Positional clustering of onco-drivers in the cohort, using oncodriveCLUST algorithm, corrected by FDR < 0.1. Numbers enclosed in square brackets represent a number of clusters found per gene and dots in red correspond to statistically significant clusters. Potential onco-driver genes in primary. (B) and recurrent cohorts (C), respectively.

3.1.2. Copy number aberrations in primary and recurrent tumours

As stated previously, we were able to collect blood samples from the same patient as well as tumour tissue, which served as a source of reference DNA. Using data from matched tumour-normal pairs, we computed somatic CNAs and discovered the presence of repeated and consistent CNAs, primarily on chromosomes 7 and 10 (Fig. 3.5), indicating frequent DNA duplications or deletions in these areas.

Figure 3.5. Copy number aberrations in progression of HGGs. CNA segments at the cohort level, with rows representing individual patients, where "P" denotes primary tumour and "R" denotes recurrent tumour. The UCSC hg38 genome was used to calculate the chromosomal coordinates, which are denoted in columns by the corresponding chromosome number. CNA calling was done with the circular binary segmentation (CBS) algorithm, and single point outliers were smoothed before the analysis.

Amplification of the *EGFR* was found in the majority of the HGGs (75%, 12/16), both primary and recurrent (Fig. 3.6A). This amplification persisted after recurrence, with varying degrees of the intensity among individuals. Furthermore, after recurrence, *PTEN* deletion was not detected in several HGGs (Fig. 3.6B). Surprisingly, *EGFR* and *PTEN* copy numbers were inversely correlated; in tumours with higher levels of *EGFR* amplification, a higher level of *PTEN* deletion was found in both primary and recurrent cohorts (Fig. 3.6C, Fig. 3.6D).

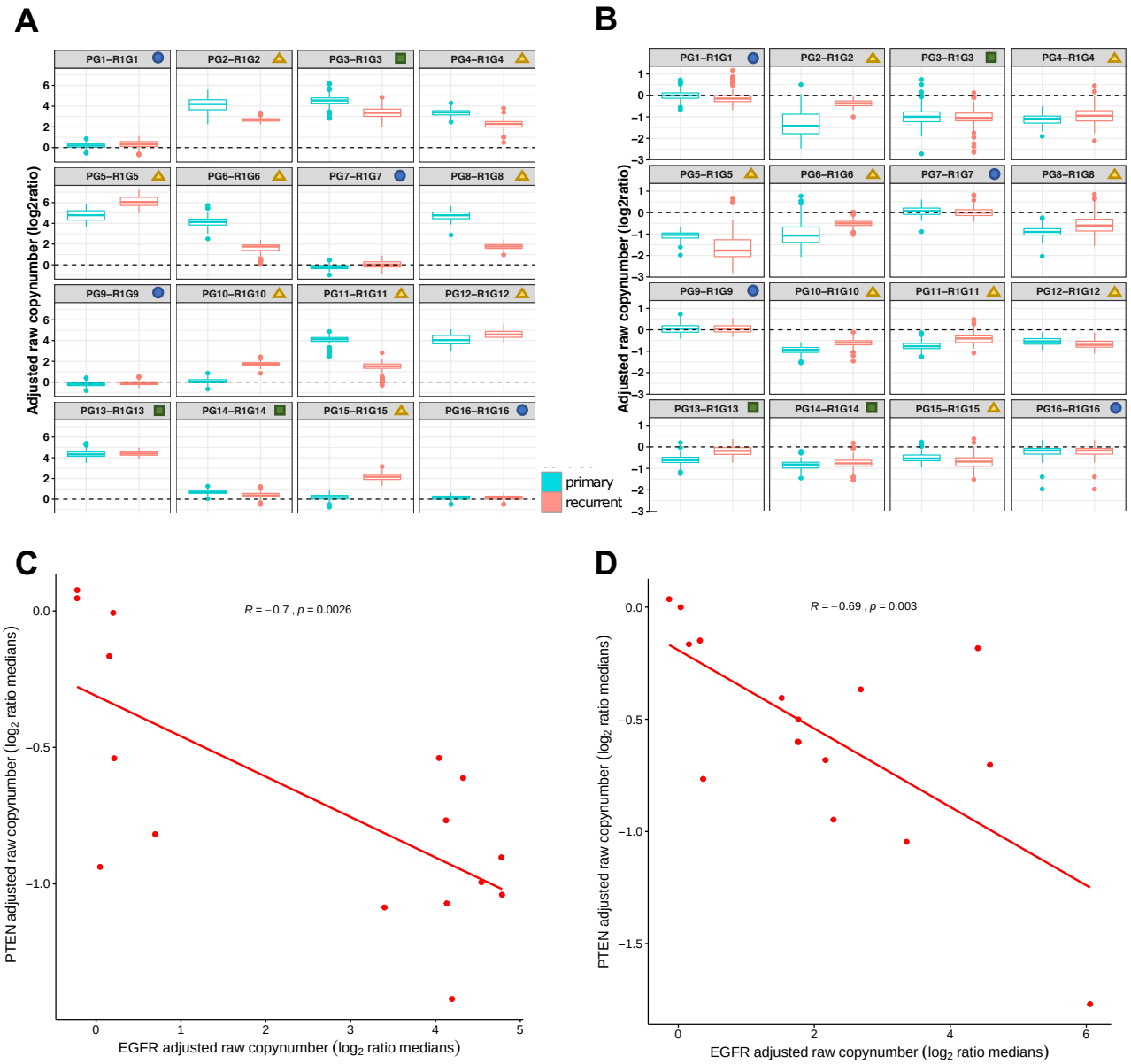


Figure 3.6. Focal CNAs in progression of HGGs. The adjusted log ratio between blood DNA (reference) and tumour DNA is used to represent focal copy number aberrations of (A) *EGFR* and (B) *PTEN* in primary (teal)-recurrent (red) paired boxplots. The top right corner blue icons indicate that there are no focal CNAs; the top right corner yellow triangle indicates that there are inversely correlated *EGFR* and *PTEN* focal CNAs; and the top right corner green squares indicate that there is no correlation. Correlation plots demonstrating somatic copy number changes of *EGFR* and *PTEN* in tumour samples. Each dot represents the adjusted copy number median of *EGFR* and *PTEN* for each of the primary (C) and recurrent (D) HGGs relative to the matched reference blood DNA.

3.1.3. Gene splicing deregulation

We employed the RSeQC package to comprehensively evaluate the RNA-seq data from HGG samples, test sequence quality, GC, PCR and nucleotide composition bias, sequencing depth, strand specificity, coverage uniformity and read genomic distribution. We found no significant differences in these parameters between primary and recurrent cohorts (data not shown). We observed, however, an increase in a proportion of complete novel splicing events in recurrent HGGs in this analysis (Fig. 3.7), suggesting higher transcriptomic variability at recurrence. This increase in the alternative splicing could produce transcriptomic instability, novel transcripts and potentially non-functional proteins.

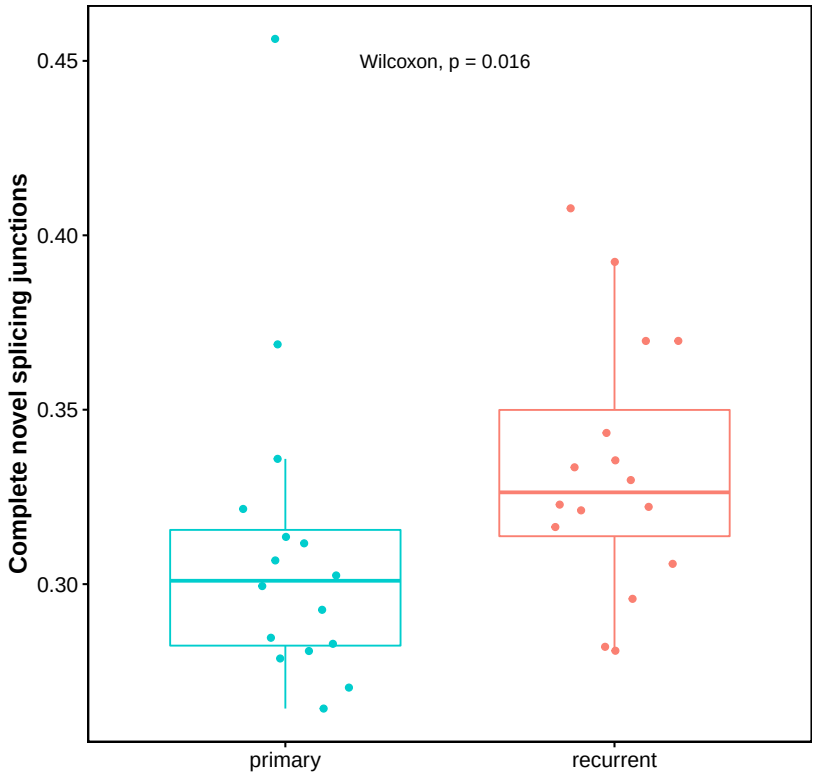


Figure 3.7. Increased novel splicing events in HGGs upon recurrence. A reference gene model from RefSeq was used to detect complete novel splicing junction alterations in primary and recurrent HGGs (hg38). The data represents the arcsin transformation, and the p-value was calculated using the Wilcoxon signed-rank test.

3.2. Transcriptomic dysregulation and changes in the tumour microenvironment

3.2.1. Gross changes in transcription characterise recurrent HGGs

Transcriptomic profiles of primary and recurrent HGGs were generated using RNA-seq analysis. The differentially expressed genes between the two cohorts were subjected to functional enrichment analyses using the Kyoto Encyclopaedia of Genes and Genomes (KEGG) and the Reactome database. KEGG analysis revealed that in recurrent HGGs, genes related to spliceosome, DNA replication, and cell cycle activity were down-regulated when compared with primary tumours (Fig. 3.8A). In contrast, fatty acid-related pathways such as the phosphatidylinositol signalling system and sphingolipid metabolism were found to be up-regulated in recurrent HGGs. The Reactome pathway analysis (Fig. 3.8B) showed a large group of up-regulated genes involved in interferon (IFN) signalling, IFN stimulation, phosphatidylinositol (PI) signalling, and sphingolipid metabolism in recurrent HGGs, which was consistent with the KEGG analysis. In recurrent HGGs, cell cycle, DNA replication, DNA repair, and spliceosome genes were all down-regulated (Fig. 3.8B).

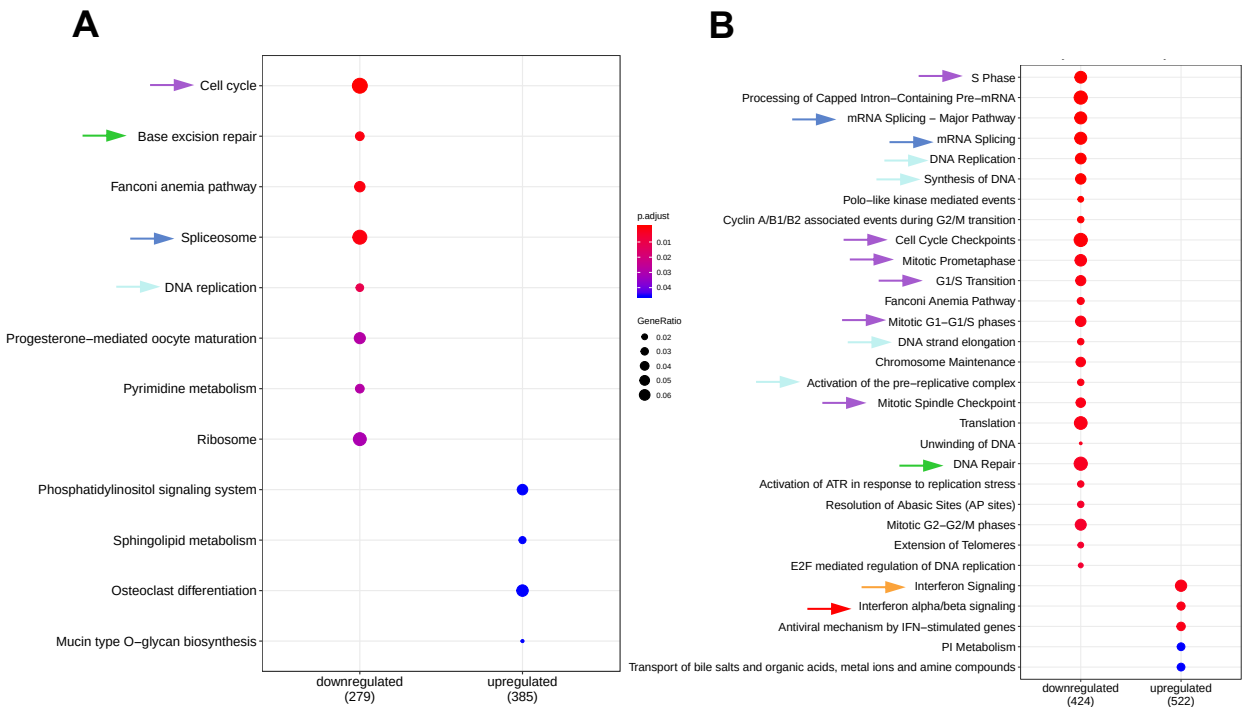


Figure 3.8. Transcriptomic pathway alterations in recurrent HGGs. KEGG (A) and Reactome (B) analyses of differentially expressed genes in recurrent HGGs samples reveal several altered regulatory pathways. To correct for multiple testing, raw counts were pre-filtered (>10 reads within the cohort), a Cook's distance cut-off of 0.5 was imposed, and the Benjamini-Hochberg (BH) procedure was used. From the two independent pathway enrichment analyses, arrows indicate similar pathways: purple arrows indicate cell cycle related pathways, green arrows indicate DNA repair related pathways, dark blue arrows indicate spliceosome, and cyan arrows indicate DNA replication related pathways.

We found significant transcriptomic profile differences when primary and recurrent tumours were compared. We discovered 1,696 genes that were differentially expressed between the two tumour stages (Fig. 3.9). In particular, 735 genes were down-regulated in recurrent tumours, while 961 genes were up-regulated after relapse (Fig. 3.9).

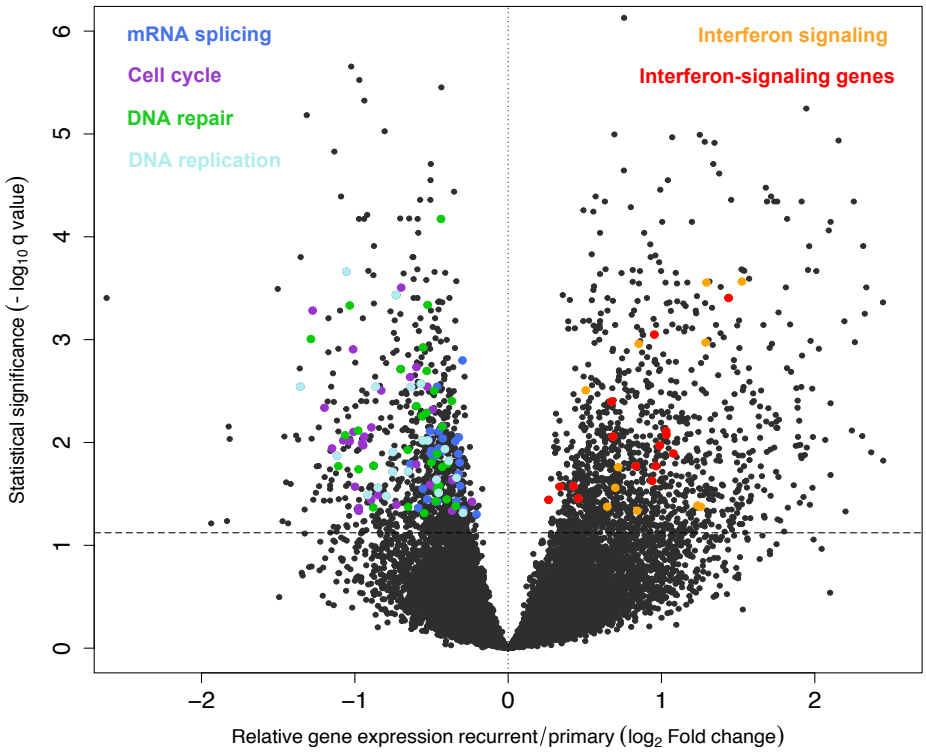


Figure 3.9. Differentially expressed genes and pathways in recurrent vs primary HGGs. The volcano plot shows down- (735) and up-regulated (961) genes in recurrent HGGs relative to primary HGGs (log2 fold change < 0 and log2 fold change > 0, respectively, and BH FDR-correction P < 0.05). Each dot above the dashed line (corresponding to a q-value of 0.05) represents a significantly changed gene. Genes from selected functional KEGG and Reactome categories are coloured differently, as indicated.

3.2.2. Cell type enrichment differences after relapse

Among differentially overexpressed genes in recurrent versus primary HGGs, we found many genes that were related to the immune response. Therefore, to further investigate potential differences in the composition of the tumour microenvironment upon HGG relapse, we performed a computational deconvolution of bulk RNA-seq profiles using the xCell tool (227). The procedure resembles cell sorting based on transcriptomic signatures of hematopoietic cells. We found that some cell enrichment scores differ remarkably between primary and recurrent HGG. The most striking changes were observed primarily in the abundance of M2 (pro-tumorigenic) macrophages, immature dendritic cells

(iDC), T helper cells (Th1), but also in megakaryocyte–erythroid progenitor (MEP) cells and pro B-cells, among others (Fig. 3.10). Noteworthy, M2 and iDC signatures were enriched, while Th1 scores were lower in the HGG samples collected after relapse.

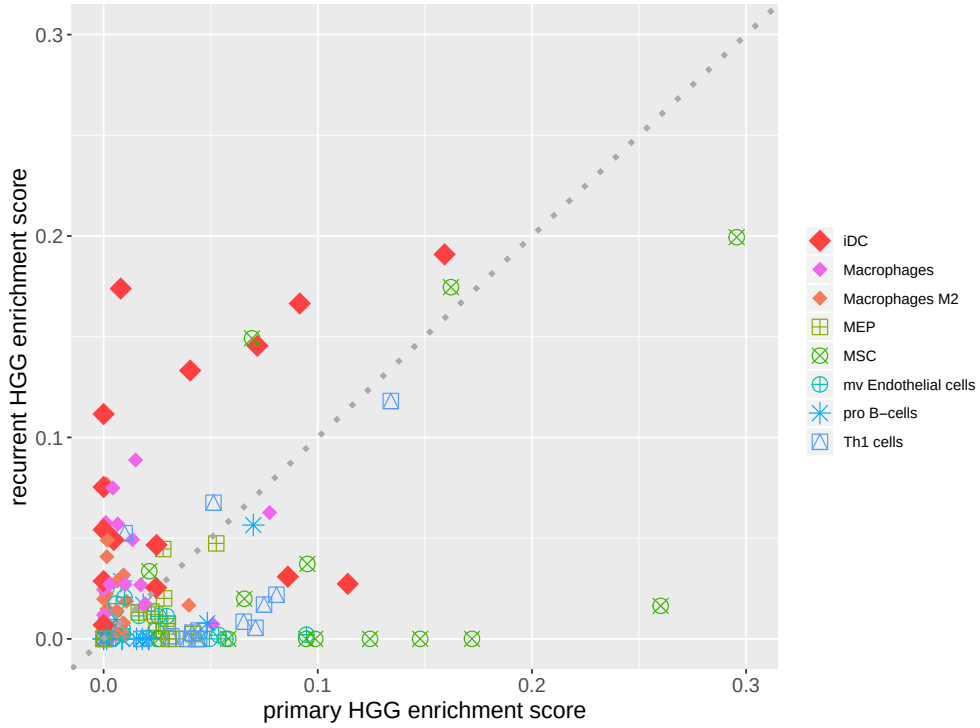


Figure 3.10. Immune cell heterogeneity in primary and recurrent HGGs. The analysis of cell type enrichment from normalized gene expression data using a 64-immune and stroma signature set reveals significant differences in enrichment scores between primary and recurrent cohorts. iDC stands for immature dendritic cells; MEP stands for Megakaryocyte–erythroid progenitor cells and MSC stands for mesenchymal stem cells. The Wilcoxon signed-rank test was used to calculate p-values, and the most significant signatures are shown.

We cross-validated these findings using the CIBERSORT and QuantiSEQ approaches (Fig. 3.11A, Fig. 3.11B). Both algorithms revealed an increase in M2 macrophages signature in the recurrent state. Macrophages may acquire distinct phenotypes: M2 macrophages play immunosuppressive functions and their presence in the tissue is associated with a reduction of inflammation, while M1-polarized macrophages orchestrate immune response (250,251). Interestingly, we discovered that the M1 macrophage enrichment was generally low in both tumour stages, whereas the scores for the unpolarized M0 cell population were higher in primary HGGs. In order to estimate the DC content in analysed tumour samples with an independent method we carried out bioinformatic predictions using Mcp-COUNTER and TIMER. These approaches were only able to evaluate myeloid dendritic cells (Fig. 3.11C, Fig. 3.11D), and we could not find any available computational tools for the detection of iDC apart from xCell. Using those algorithms, we found a higher dendritic cell enrichment scores in the recurrent HGGs, but the differences were not significant.

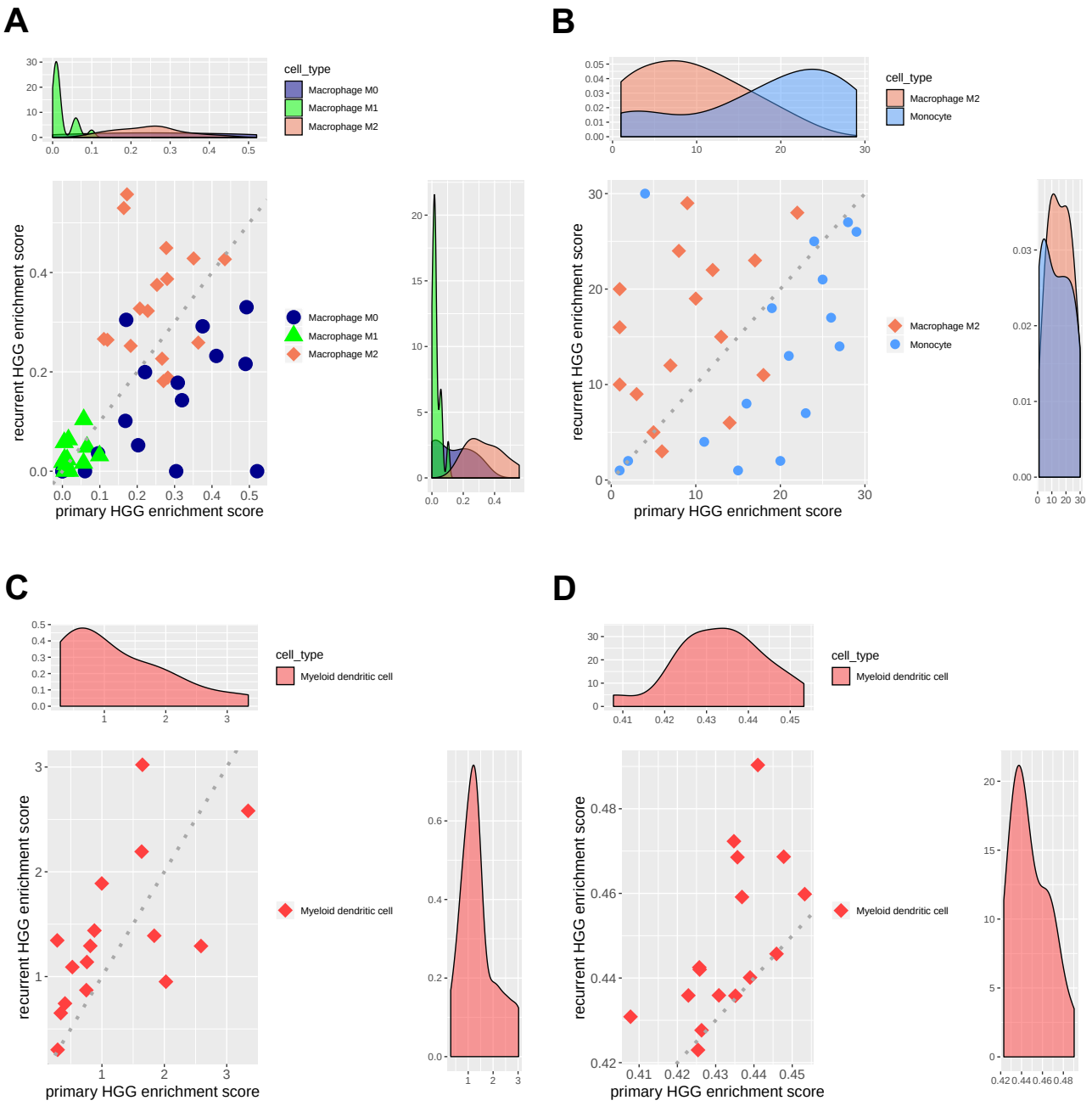


Figure 3.11. Heterogeneity of immune signatures of macrophages and DC in primary and recurrent HGGs. Analysis of a cell type enrichment from normalized gene expression using (A) CIBERSORT, (B) QuantiSEQ, (C) mcp-COUNTER, and (D) TIMER approaches to depict macrophages, monocytes, and myeloid dendritic cells. For each of the analyses, scatter- and density plots are shown. Top density plots depict the x-axis distribution of cell enrichment scores, while right density plots depict the y-axis distribution.

3.2.3. Expression of immunoglobulins, M2 macrophages and dendritic cell markers

Numerous mRNA coding for immunoglobulins and immunoglobulin-related molecules were found to be expressed at higher levels in recurrent HGGs, but the level of expression in recurrent tumours was highly variable (Fig. 3.12A). Furthermore, the majority of these were low expressed and did not meet the pre-filtering criteria for differential gene expression analysis. The increased expression of tumour-associated M2 macrophage markers (252) in HGGs (Fig. 3.12B) emphasizes the important role of pro-tumorigenic macrophages in HGG progression. In the data set, we focused on mRNA levels of specific differentially expressed markers for M2 macrophages, iDC, and DC (Fig. 3.12C); upon recurrence, we observed an increase in mRNA markers for iDC (CD209) and M2 macrophages (CD163), but not in DC markers (CD83).

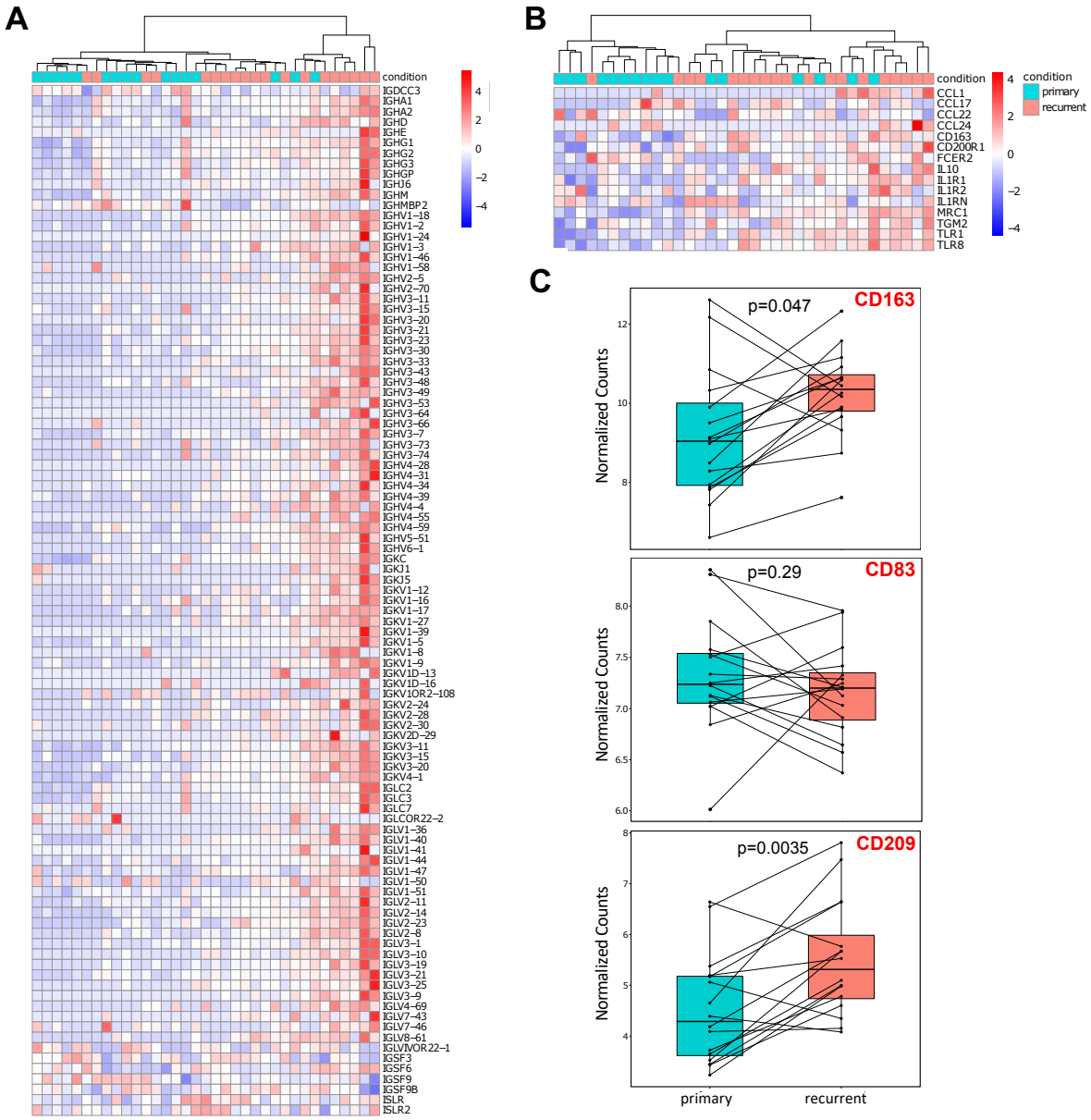


Figure 3.12. Dysregulation of immunoglobulins expression in HGGs and changes in glioma markers for infiltrating macrophages and DC markers upon tumour recurrence. Heatmaps depict mRNA expression levels of (A) immunoglobulins and immunoglobulin-related molecules and (B) M2 macrophage markers. DESeq2's variance stabilizing transformation method was used to normalize raw count values, and Ward's method was used to cluster samples and to identify the strongest clustering structure. (C) Box plots represent mRNA levels of selected immune cell markers in matched pairs of primary and recurrent HGGs samples: CD163 (M2 macrophages), CD83 (DC), and CD209 (iDC).

3.2.4. Accumulation of pro-tumorigenic macrophages and immunosuppressive dendritic cells

To validate our findings on the enrichment of transcriptomic signatures of certain immune cells upon tumour recurrence, we decided to employ immunohistochemistry staining and check the abundance of selected cell types in the sections of matched primary and recurrent human HGGs from the same patients. We used an antibody that recognizes the scavenger receptor cysteine-rich (SRCR/CD163) for detection of M2 macrophages, an antibody against dendritic cell-specific ICAM-3-Grabbing non-integrin 1 (DC-SIGN/CD209) for detecting mDC and iDC and an anti-CD83 antibody for detecting mDC. In the recurrent GBMs, we found higher numbers of CD209+ cells as compared to primary tumours (Fig. 3.13A). The observed changes were validated by counting CD209+ cells (Fig. 3.13B). There were no differences between primary and recurrent HGGs when an anti-mDC antibody was used, confirming that only immature cells that stained with CD209 were enriched. We also found more CD163+ cells in the brain parenchyma and perivascular spaces in the recurrent GBMs than in primary GBMs, indicating an accumulation of anti-inflammatory M2 macrophages with the disease relapse.

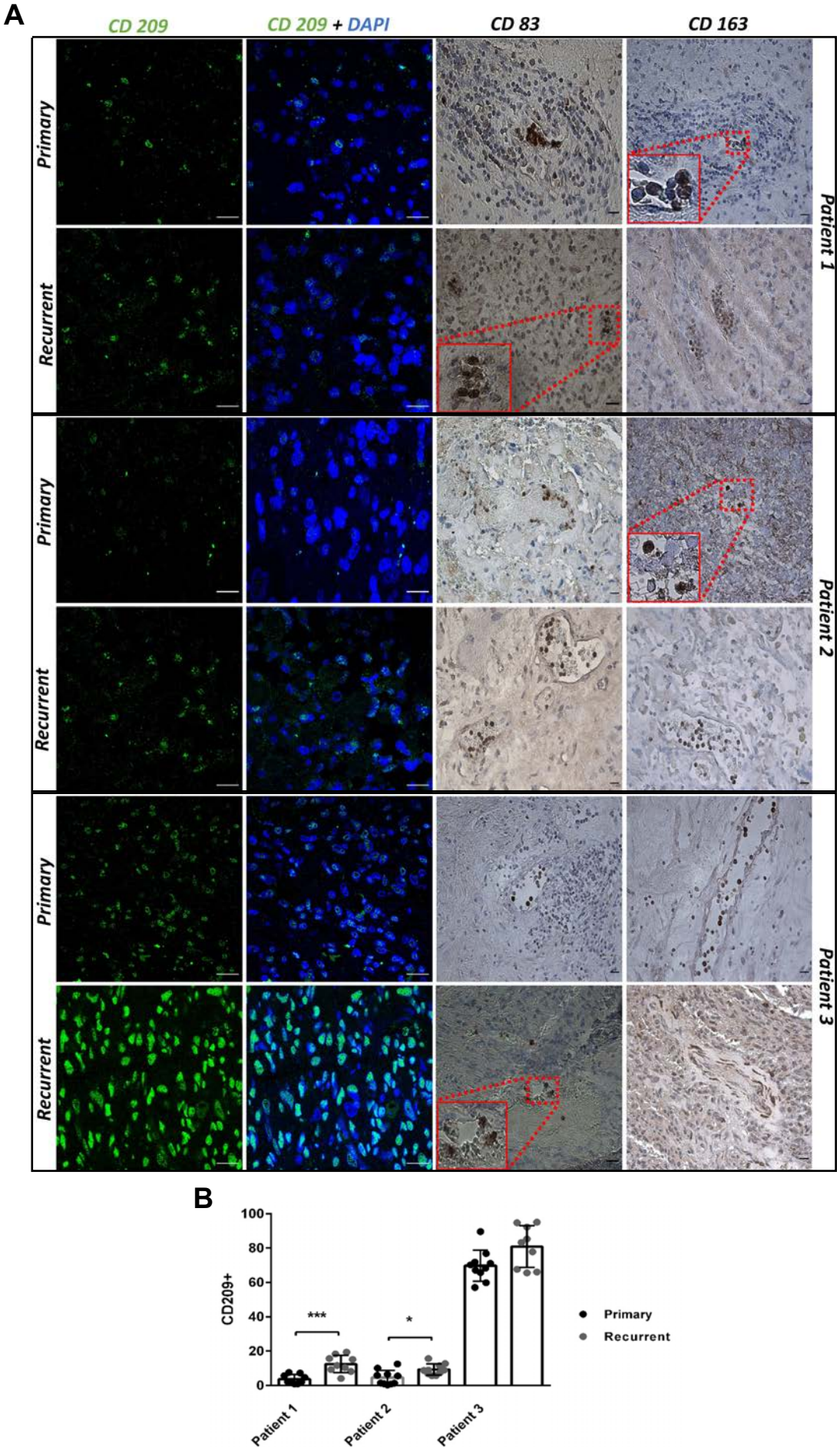
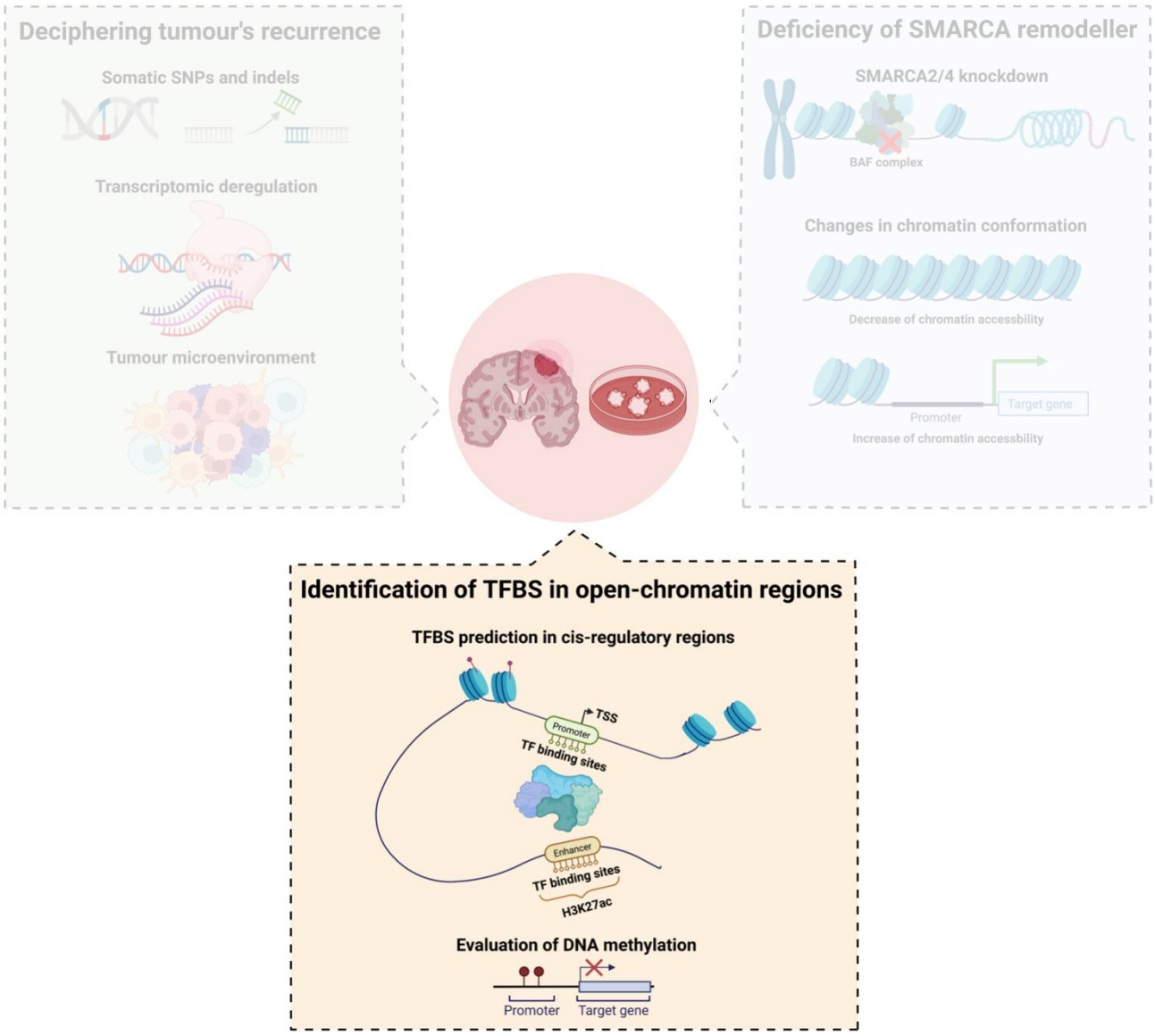

A
CD 209
CD 209 + DAPI
CD 83
CD 163
Primary
Recurrent
Patient 1
Primary
Recurrent
Patient 2
Primary
Recurrent
Patient 3
B
CD209+
100
80
60
40
20
0
***
*
Patient 1
Patient 2
Patient 3
Primary
Recurrent

Figure 3.13. Accumulation of pro-tumorigenic macrophages and immature dendritic cells in recurrent GBMs. (A) Representative confocal microscopy images show that recurrent GBMs have a higher number of immature dendritic cells (CD209+) than primary GBMs; cell nuclei are counterstained with DAPI (blue), scale bar: 20 μm. The images show mature dendritic cells (CD83+) and phagocytic microglia/macrophages (CD163+). (B) CD209 positive cells were quantified in relation to total cells. The cells were counted from 10 randomly selected fields using ImageJ software, and p-values were calculated using the Mann-Whitney nonparametric test.

3.3. Multi-omics integration and identification of TFBS in open-chromatin regions



3.3.1. Identification of TF binding sites in open-chromatin regions related to glioma malignancy

We obtained TFBS predictions in both GBM cell lines (LN18 and LN229) and GBM specimens (GBM1 and GBM2) using chromatin accessibility data (ATAC-seq), as described in section 2.2.7. We found a strong overlap of TFBS calls in the cell lines (Fig. 3.14A), as well as in GBM samples (Fig. 3.14B). Only the TFBS predictions found in both cell lines (145123) were taken into account for downstream analysis. We created heatmaps of all the peaks to learn more about the ATAC-seq signal enrichment near the Transcription Start Site (TSS), and the precise distribution of these promoter

peaks revealed a clear higher accumulation in the vicinity of the TSS for samples analysed (Fig. 3.14C), confirming the quality of both the annotation and our data.

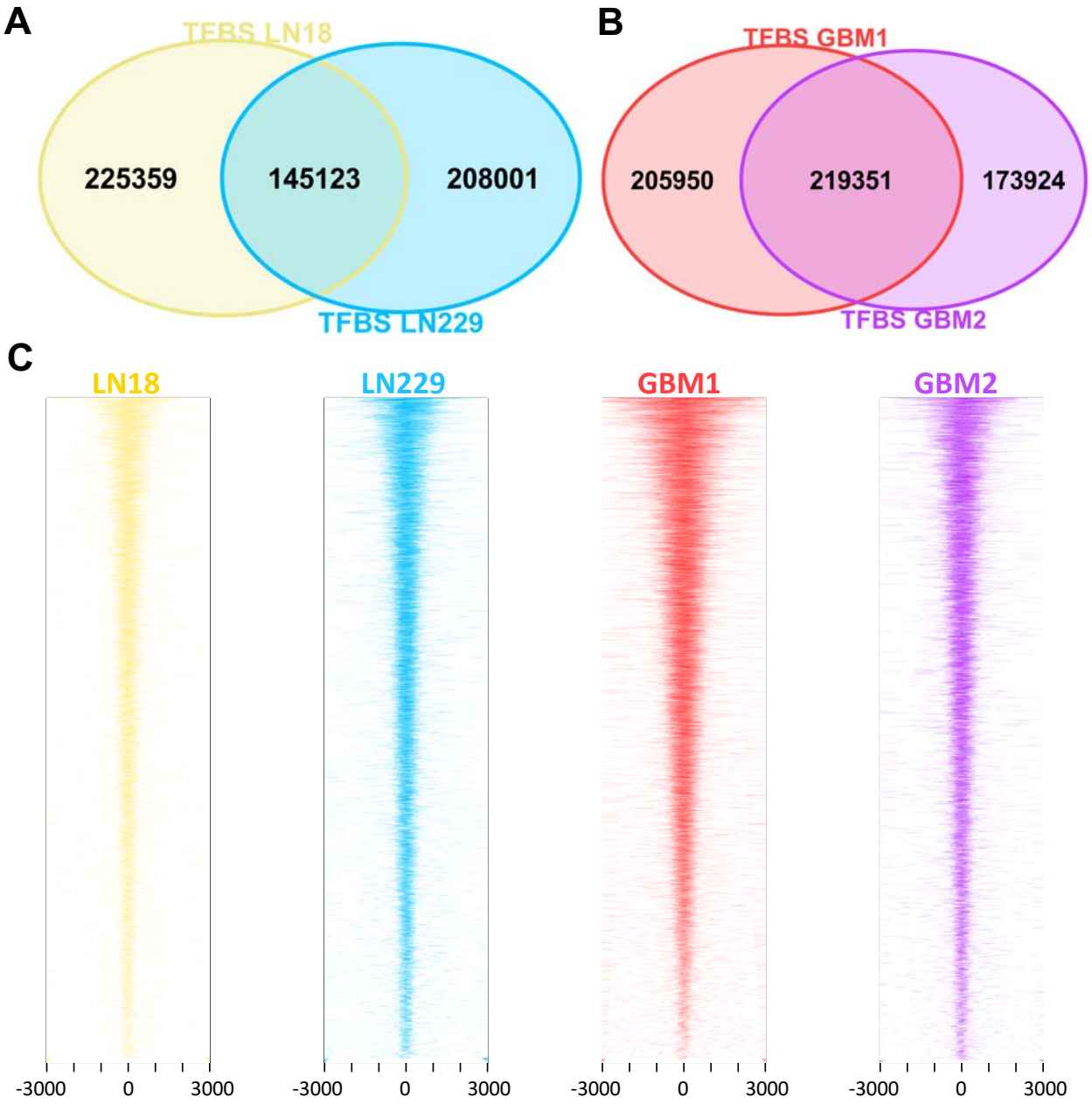


Figure 3.14. Characterisation of TFBS in open chromatin regions in glioblastoma specimens and glioblastoma cell lines. Total number of transcription factor binding sites (TFBS) predicted in open-chromatin regions using ATAC-seq fragments and position-weight matrices (PWMs) motifs in (A) established human glioma LN18 and LN229 cells and in (B) glioblastoma samples. (C) Profile heatmap of total ATAC-seq peaks identified around transcription start sites (TSS) in the cell lines and GBM specimens, color-coded as in the previous Venn diagrams.

Next, we verified how many different transcription factor motifs were found in open-chromatin gene promoters, and we discovered that AP2D, PAX5, ZFX, KLF4 and SP2 motifs were the most

abundant, among others (Fig. 3.15), suggesting that they are potentially involved in the regulation of many genes in gliomas.

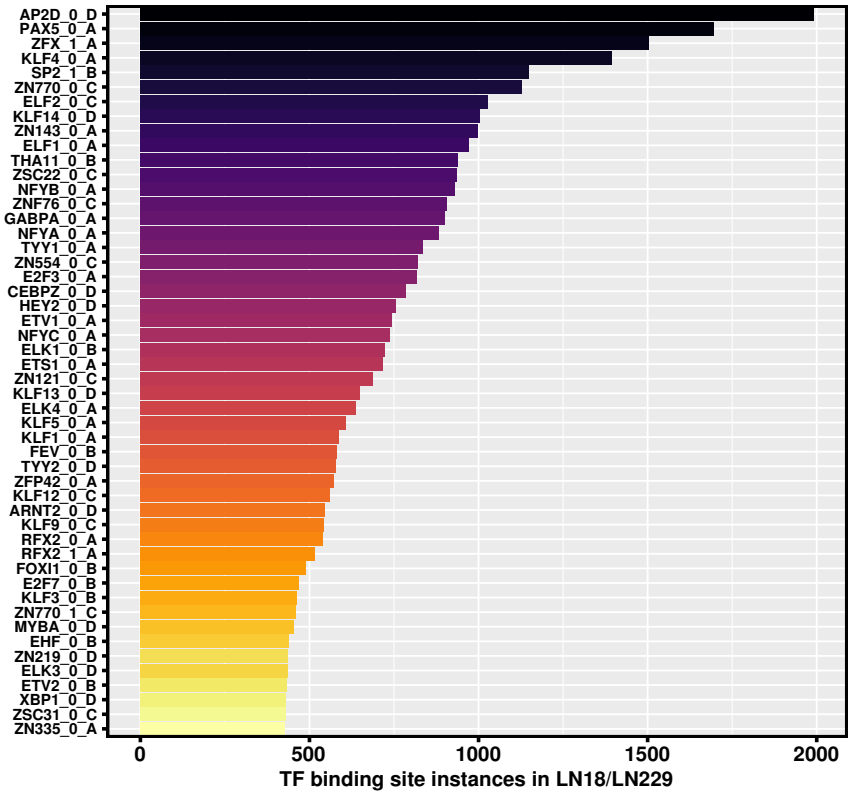


Figure 3.15. Most abundant TFBS in open-chromatin regions. Total number of predicted transcription factor binding sites (TFBS) in open-chromatin regions using ATAC-seq fragments and position-weight matrices (PWMs) in established human LN18 and LN229 glioma cells.

Subsequently, as described in section 2.2.5, we examined the landscape of open chromatin regions, focusing on genes highly expressed in low-grade (WHO GII) specimens and genes highly expressed in high-grade (WHO GIV) glioma tumours. Because we lacked chromatin accessibility from low-grade (WHO GII) patients, we used TCGA data to identify genes that were significantly over-expressed in GBMs when compared to LGG and inferred TFBS calls from glioblastoma cell lines. Here, we focused exclusively on TF motifs found in the promoter regions of genes differentially expressed upon grade GIV or upon grade GII gliomas. TFBS found in the promoter region of genes overexpressed in either grade IV or grade II glioma were referred to as "generic TFBS," whereas TFBS found only in the promoter region of genes overexpressed in a specific glioma grade were referred to as "grade-specific TFBS" (Fig. 3.16).

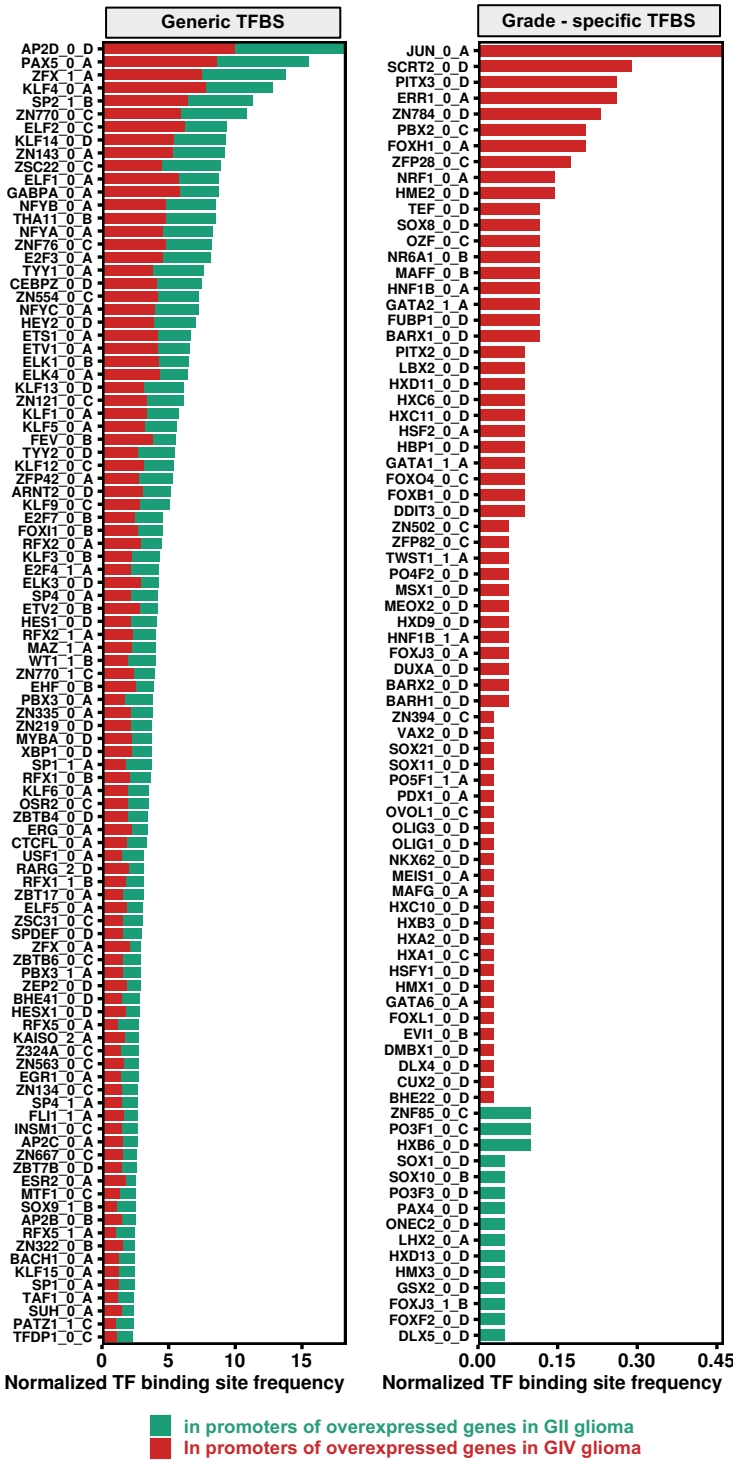


Figure 3.16. Most abundant TFBS in genes within open-chromatin regions. Prediction of "generic" (left panel) and "grade-specific" (right panel) transcription factor binding sites (TFBS) in the promoters (TSS ± 1.5 kb) of differentially expressed genes (DEGs) between grade IV and II gliomas. The abscissa is a normalization factor for TFBS occurrences that takes into account the total number of DEGs within a given glioma grade.

We found "generic TFBS" motifs in similar proportions in the promoters of GIV and GII genes, indicating that they could be involved in a variety of glioma processes regardless of the tumour grade (Fig. 3.16). However, we did find some TFBS that were present only in actively transcribed GIV-genes

and they represented sites for JUN, SCRT2, PITX3, ERR1, or ZN784. Several TFBS such as ZNF85, PO3F1, HX36, SOX1, or SOX10 were present in within actively transcribed GII-genes (Fig. 3.16). Subsequently, we used the HOCOMOCO database to annotate those grade-specific TFs into TF families, and we found that some TF families may be more relevant (unique) for transcription regulation in GIV and some in GII grade gliomas (Fig. 3.17A, Fig. 3.17B). There were TF families, such as NK-related factors, HOX-related factors, and POU-related factors, for which the TFBS were enriched in the actively transcribed genes in both GII or GIV gliomas. Certain TF families, on the other hand, such as Tal-related factors, GATA-type zinc fingers, or Jun-related factors, were present only in the GIV glioma genes (Fig. 3.17A). This finding implies that certain TF families may play more important roles in GBMs.

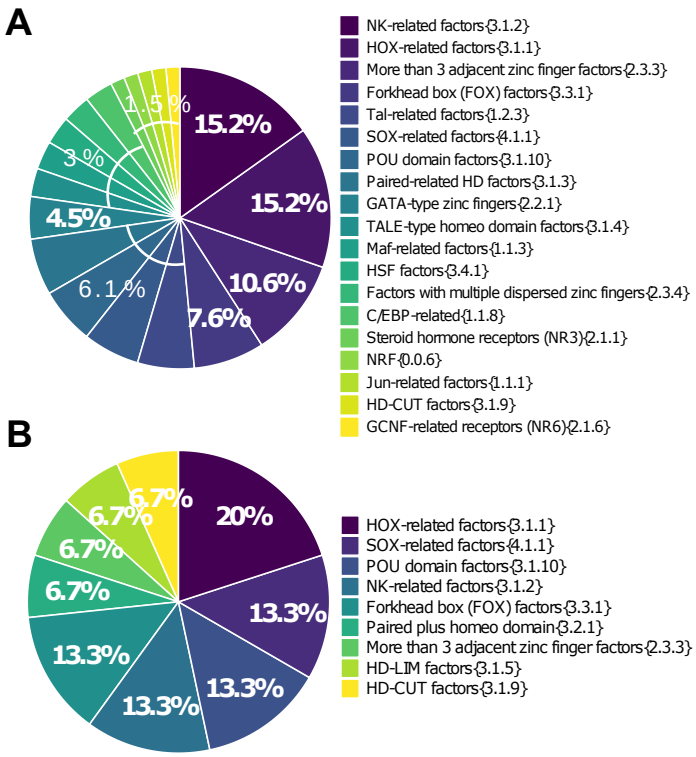


Figure 3.17. Transcription factor families from identified TFBS. HOCOMOCO v11 (version 11) transcription factor (TF) families from overexpressed genes in GIV-specific TFBS (A) and GII-specific TFBS (B).

3.3.2. Transcriptomic profiles of GBM and LGG gliomas in TCGA

Because gene selection was an important step in defining grade-specific TFBS, we looked at transcriptomic differences between GBMs (grade IV glioma) and LGGs (focusing only on WHO grade II glioma) using TCGA data to see if our selection of differentially expressed genes (DEGs) (as described in section 2.2.5) between gliomas of different grades had any biological significance. As

many previous studies have shown (253–255), patient samples clustered based on the glioma grade (Fig. 3.18A), and the number of transcriptomic deregulations was remarkable (Fig. 3.18B).

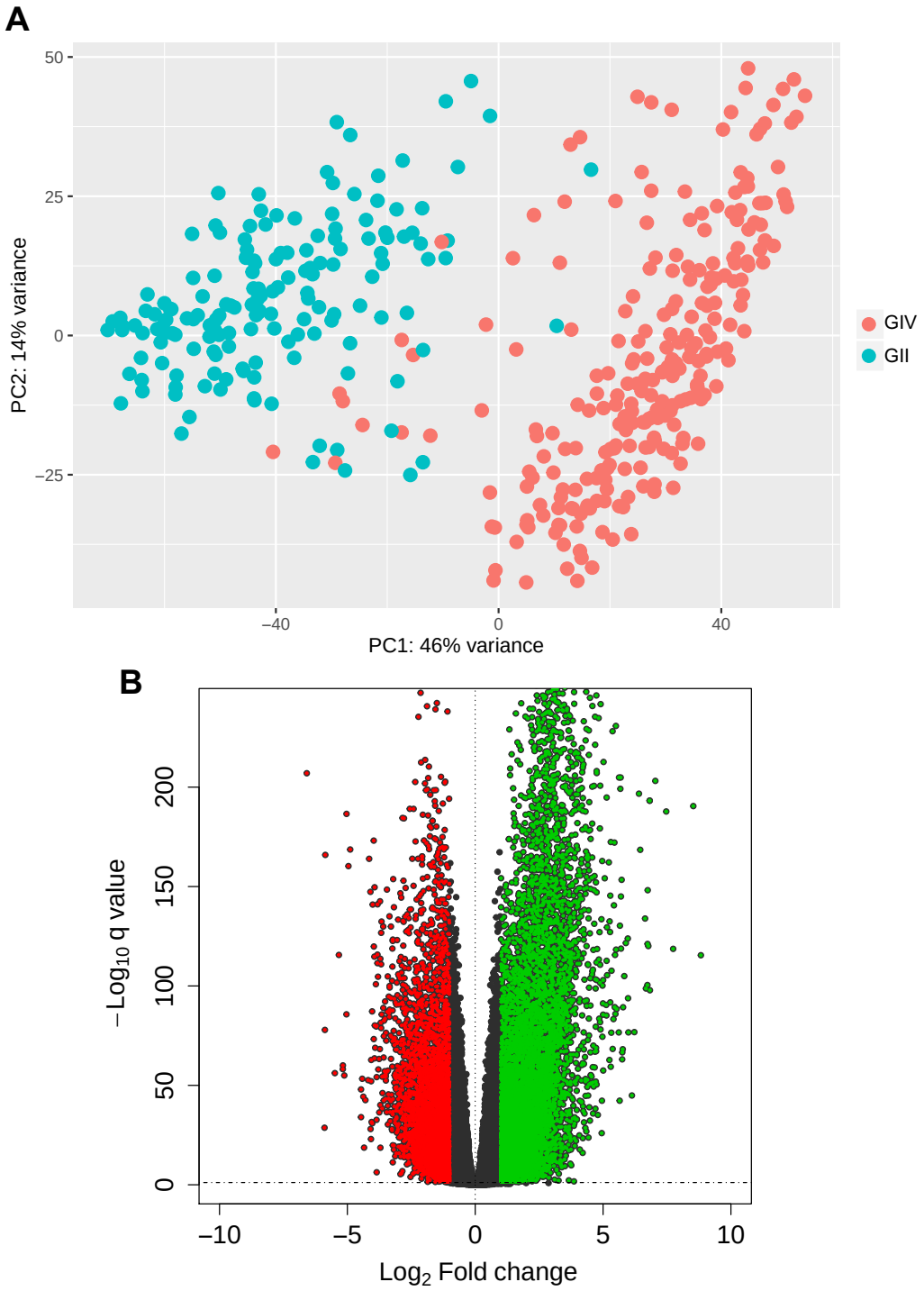


Figure 3.18. Transcriptomic differences between high- and low-grade gliomas. (A) Principal component analysis (PCA) was used to plot TCGA samples (248 GII gliomas and 160 GIV gliomas). The first two principal components (PCs) are plotted and coloured according to the patient's glioma grade. The TCGA's normalized RNA-seq expression data were used to perform PCA. The axis label displays the percentage of variation accounted for by each principal component. (B) Volcano plot depicting the relevant gene expression differences between glioma grades (GIV vs GII). Green and red dots represent statistically significant up-regulated genes (DESeq2 methods, padj < 0.01) in GIV gliomas or GII gliomas, respectively, with logFC >1. The q-value threshold is indicated by a dotted horizontal line.

The overexpressed genes in GIV gliomas are immune- and cell cycle-related (Fig. 3.19, Fig. 3.20), whereas in GII gliomas overexpressed genes are synaptic- and neuron-related, as shown using the GSEA-based analysis (Fig. 3.19). Furthermore, using the same rationale, we performed an additional pathway enrichment analysis with the ClusterProfiler R library and found that genes up-regulated in GIV gliomas were associated with the P53 signalling pathway, cell cycle, IL-17, nucleosome assembly, and ECM organization, among other elements, whereas genes up-regulated in GII gliomas were associated with neuroactive interactions, GABAergic and synapses and synaptic plasticity (Fig. 3.20), which is in line with previous analyses on glioma of different grades (256–259).

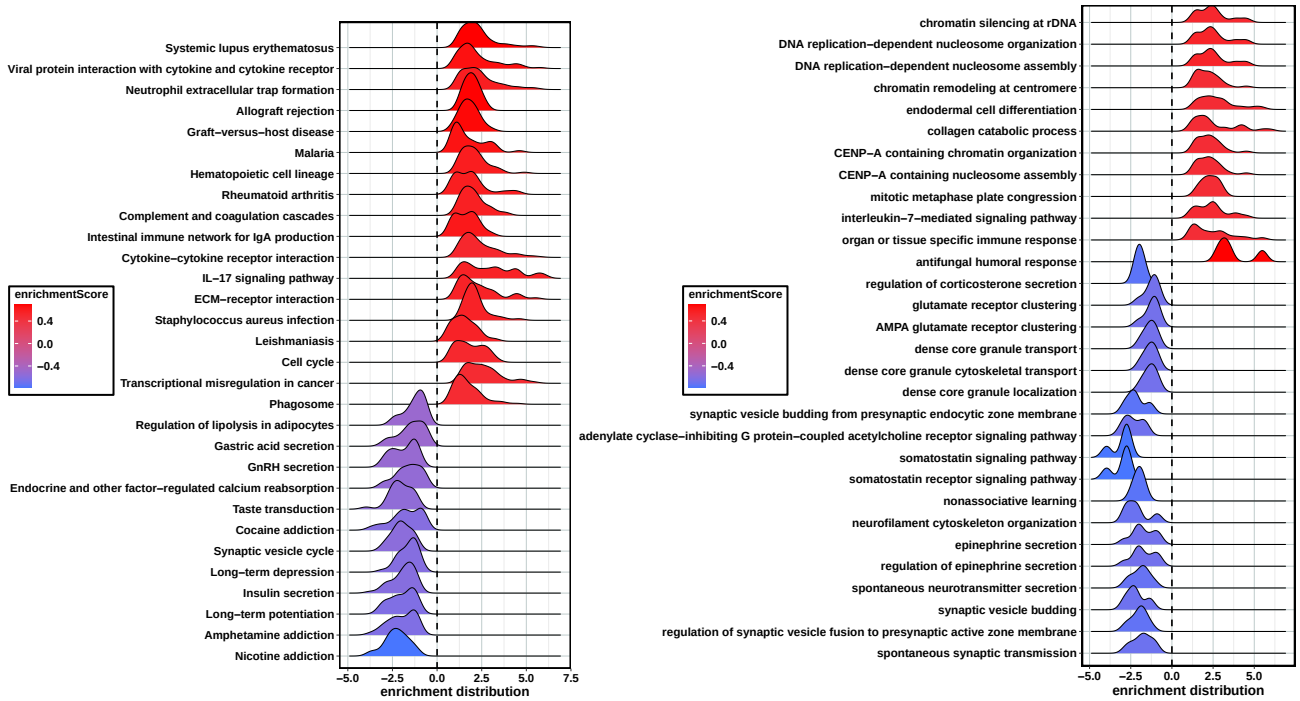


Figure 3.19. Gene set enrichment analysis. Over-represented pathways and their corresponding enrichment distribution and scores in GIV gliomas (red, 10% positive enriched categories) versus GII gliomas (blue, 90% negative enriched categories) determined by the KEGG (left) and Gene Ontology: Biological Processes (right) databases.

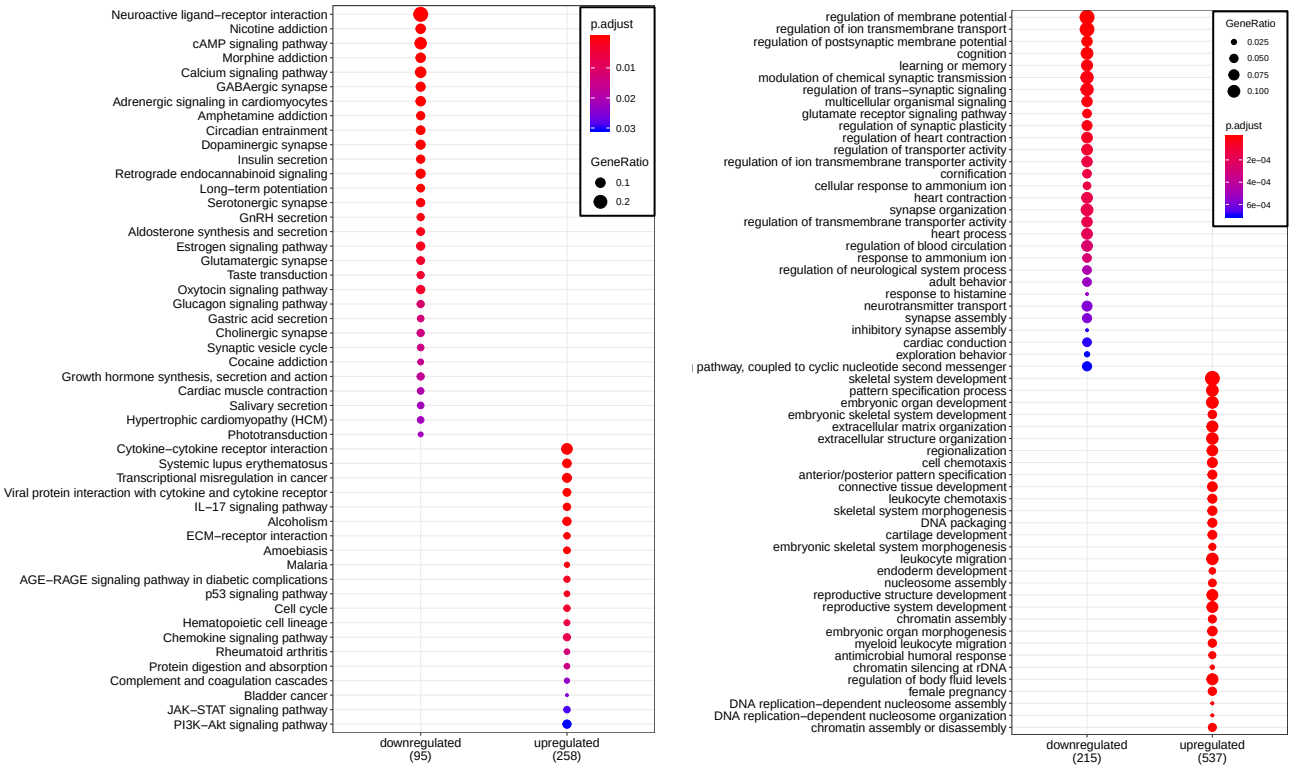


Figure 3.20. Pathway enrichment analysis. Pathway enrichment of differentially expressed genes (up logFC>2.5 and padj<0.01; down logFC<-2.5 and padj<0.01) between GIV gliomas and GII gliomas using (A) KEGG and (B) Gene Ontology: Biological Processes databases genes reveals several regulatory pathways. Raw counts were pre-filtered (> 5 reads within the cohort), and multiple testing was corrected using the Benjamini–Hochberg (BH) procedure.

3.3.3. c-Jun expression is deregulated in many cancer types and in malignant gliomas

According to our results, the c-Jun TF is predicted to bind exclusively in the promoter regions of several genes highly overexpressed in GBM. We investigated if this observation holds in the Pan-Cancer and glioma datasets. We found that *JUN* expression varies significantly between cancer types (Fig. 3.21A). Several cancers (BLCA, BRCA, SKCM, CESC, OV, LUSC, UCEC, LUAD, and UCS, please see section 2.2.16) had significantly lower *JUN* expression when compared to non-tumorous tissues. Only two cancer types, thymoma (THYM) and GBM, had significantly higher *JUN* expression in cancer tissue compared to non-tumorous tissue. Then, we determined *JUN* expression in gliomas using the LGG/GBM TCGA datasets and found the increasing *JUN* expression in high grade gliomas (through GIII to GIV) (Fig. 3.21B).

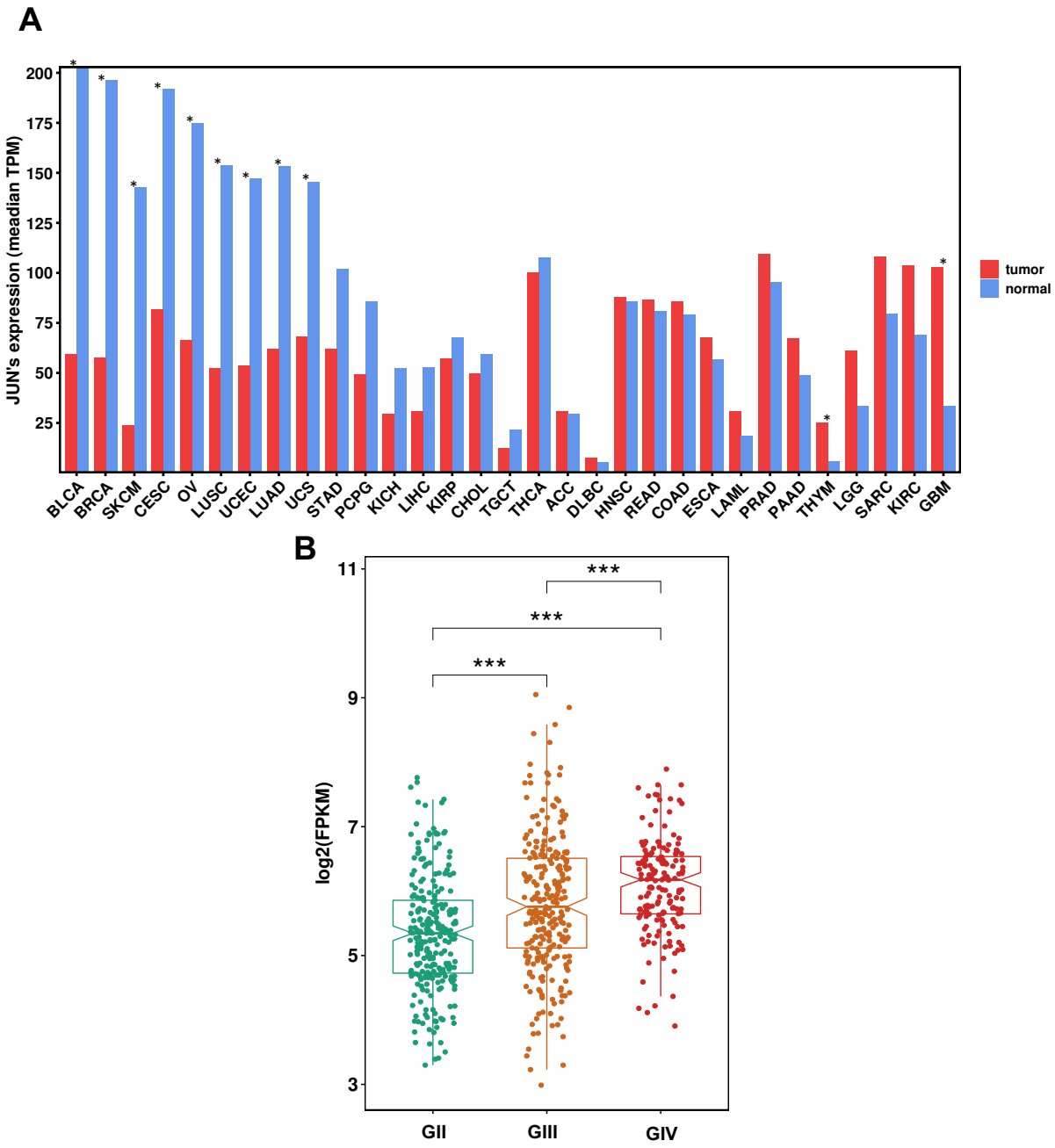


Figure 3.21. *JUN* expression in pan-cancer and across glioma grades. (A) *JUN* mRNA expression profile ordered by expression differences between tumour samples (TCGA) and paired normal tissues (TCGA normal + GTEx normal). The differential expression was calculated using one-way ANOVA between the disease states (tumour or normal, *p-value < 0.05). (B) *JUN* mRNA expression across glioma grades using the TCGA data. The differential expression was calculated using Wilcoxon rank-sum statistical test (*p-value < 0.05, **p-value < 0.01 and ***p-value < 0.001).

3.3.4. Integration of ATAC-seq peaks and cis-regulatory regions in the context of c-Jun

Because of the high intratumoural heterogeneity and highly complex tumour microenvironment in glioblastoma, we decided to validate our predictions using established glioblastoma cell lines. ATAC-seq peaks consistent between LN18 and LN229 cells (Fig. 3.22, 1st outer track) showed considerable similarity to open-chromatin regions in glioblastoma specimens (Fig. 3.22, 2nd outer track). Focusing on cell lines calls, we identified 101,962 TFBS in promoter regions (TSS ± 1.5 kb), accounting for 81.29% of all TFBS predictions; whereas, only 18.706% of TFBS were found outside of promoter regions (Fig. 3.22, 3rd outer track). We found 24 TF binding sites in the promoters of GII glioma genes and 240 TF binding sites in the promoters of GIV glioma genes when we looked for TFBS within the promoters of overexpressed genes in glioma GIV and glioma GII (Fig. 3.22, 4th outer track).

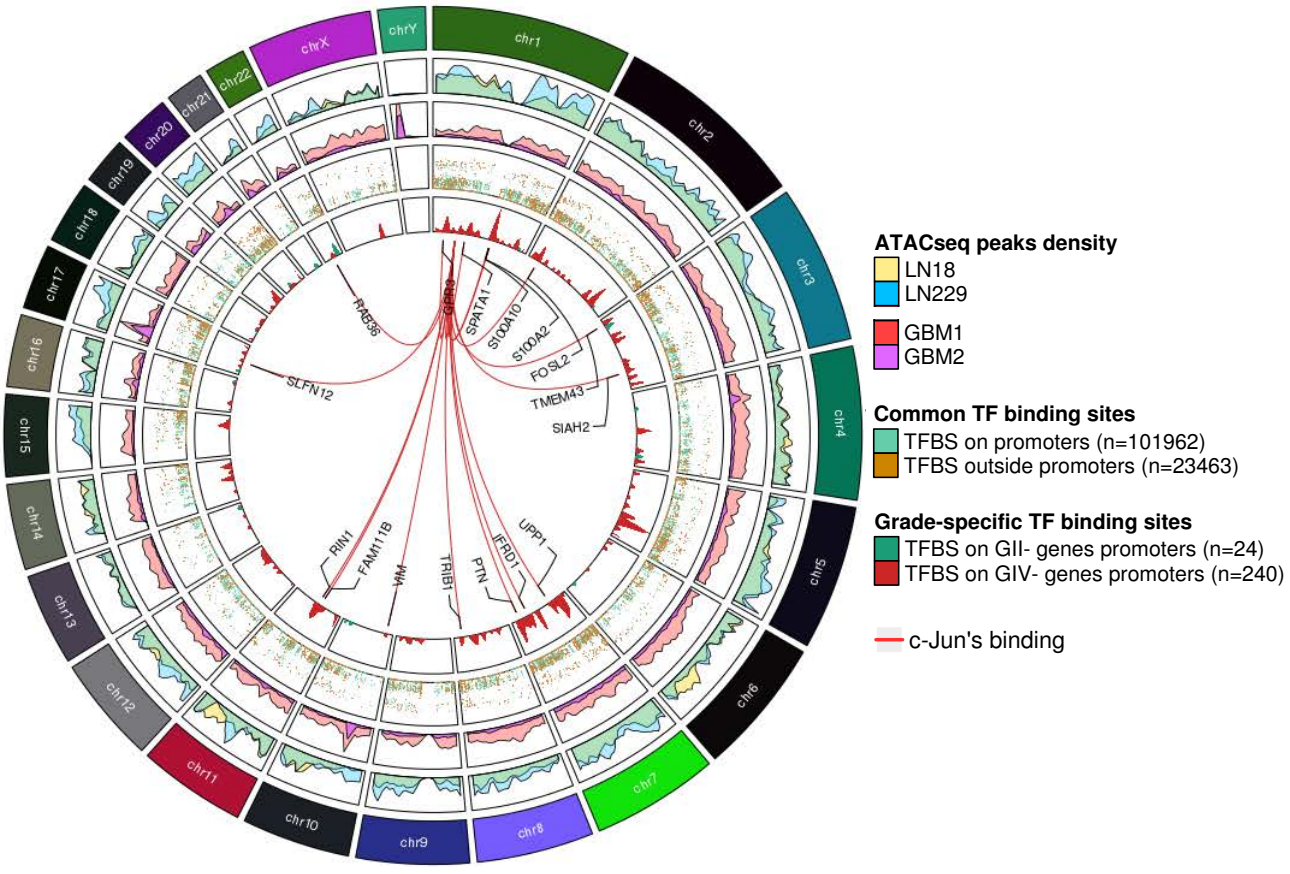


Figure 3.22. Chromatin accessibility profiling and TFBS prediction in cis-regulatory elements of over-expressed genes in gliomas of a given grade. ATAC-seq peaks identified in glioma cell lines and glioblastoma specimens (1st and 2nd track). TFBS prediction in the promoters and outside (3rd track). Grade specific TFBS predictions on overexpressed genes (4th track). The *JUN* locus (chr1:58,776,845-58,784,048) is linked to each of the c-Jun-controlled genes in glioblastoma by red lines.

Interestingly, many c-Jun binding sites were found in the promoters of genes involved in immune-related signalling, such as *IFRD1*, *UPP1*, and *SLNF12*; cell proliferation, migration, and invasiveness, such as *VIM*, *FOSL2*, *PTN*, *SIAH2*, *S100A2*, *S100A10*, and *FAM111B*; and radioresistance, such as *TRIB1*. The expression levels of all of these genes were significantly up-regulated in glioblastoma when compared with grade II gliomas (Fig. 3.23).

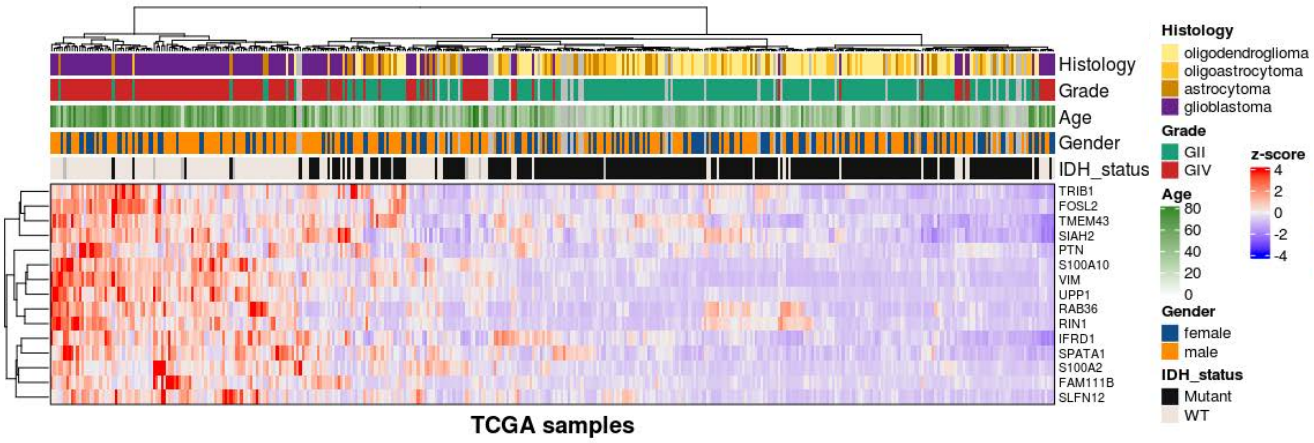


Figure 3.23. Supervised hierarchical clustering of WHO grade II and IV glioma samples based on selected genes. Heatmaps depict expression of genes with predicted c-Jun binding sites in their promoters. Transcriptomic data from grade II and grade IV gliomas linked with patient-related information are from TCGA and c-Jun binding sites predictions are based on our open chromatin data. Patients with missing clinical information on histology, grade, age or gender are illustrated in grey.

Several Jun-related factors (Jun-B, JunD) and Fos-related factors (c-Fos, FRA1, FRA2, and FosB) have overlapping binding sites with the c-Jun binding sites within gene regulatory regions of potential c-Jun targets (Fig. 3.24). This indicates that c-Jun may interact with other TFs in specific promoters to regulate gene expression, e.g., by forming the AP-1 complex.

In this line, a search for these TFBS in open chromatin sites from tumour samples was performed to see if the predictions for c-Jun and Jun-related factors generated in our studies on cultured glioma cells held true in tumour samples. The same putative TFBS were identified in a half of the gene promoters in glioblastoma samples (Fig. 3.24, GBM1 and GBM2 TFBS predictions). The results validate the presented approach and the usefulness of generating predictions on cultured glioma cells.

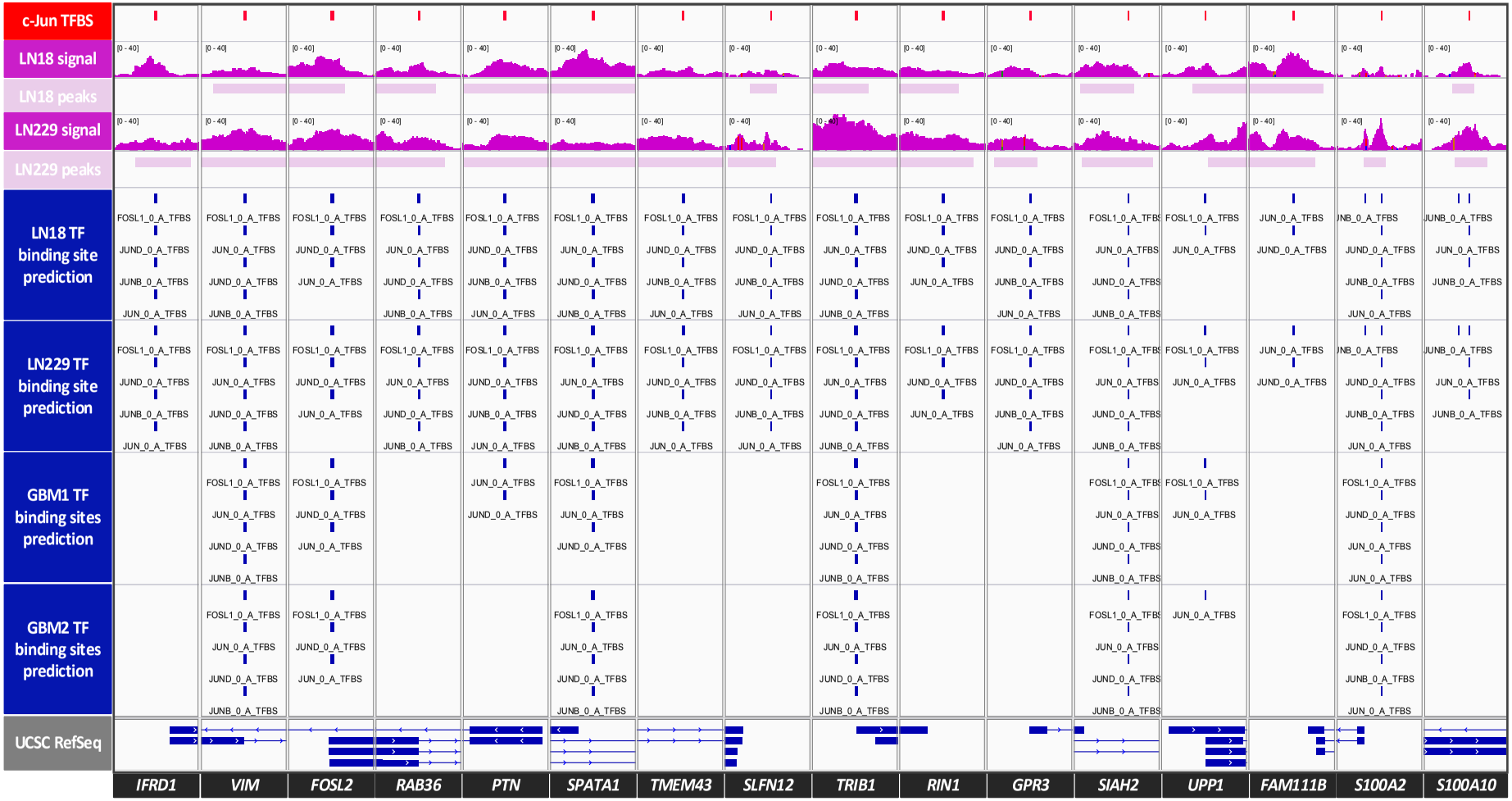


Figure 3.24. c-Jun binding prediction in cis-regulatory elements of targeted genes. Prediction of c-Jun binding (1st track) on cis-regulatory regions of over-expressed GIV genes, ATAC-seq signal (2nd and 4th tracks), and MACS2 peaks (3rd and 5th tracks). For glioma cell lines (6th and 7th tracks) and glioblastoma specimens, TFBS clusters of basic leucine zipper factors in promoters are shown in blue for glioma cell lines (6th and 7th tracks) and glioblastoma specimens (8th and 9th tracks).

3.3.5. Transcriptomic analysis of grade-specific transcription factors

The presented results suggest that TFs specific for gliomas of high malignancy grades (Fig. 3.16) may regulate genes crucial for glioma progression via cis-regulatory mechanisms in open-chromatin regions. Therefore, we hypothesized that genes coding for these TFs might be upregulated in glioblastomas. We examined the gene expression of WHO grade IV glioma specific TFs (64 genes) using hierarchical clustering of TCGA glioma WHO grade II and grade IV samples (Fig. 3.25). We found that while most of them are highly expressed in GBMs, some are more prominent in WHO grade II gliomas, suggesting their potential roles as transcriptional repressors in WHO GII gliomas. Many HOX-related genes *(HOXD11, HOXD9, HOXC10, HOXC11, HOXC6, HOXB3, HOXA2, HOXA1*) were found to be associated with the grade and histological type of gliomas. Further, we also found that the *c-JUN* gene is significantly up-regulated in glioblastomas when compared to WHO GII tumours (Fig. 3.25, arrowed). Some TF coding genes, on the other hand, were not highly expressed in glioblastomas and even showed higher expression levels in WHO grade II gliomas (Fig. 3.25, bottom cluster).

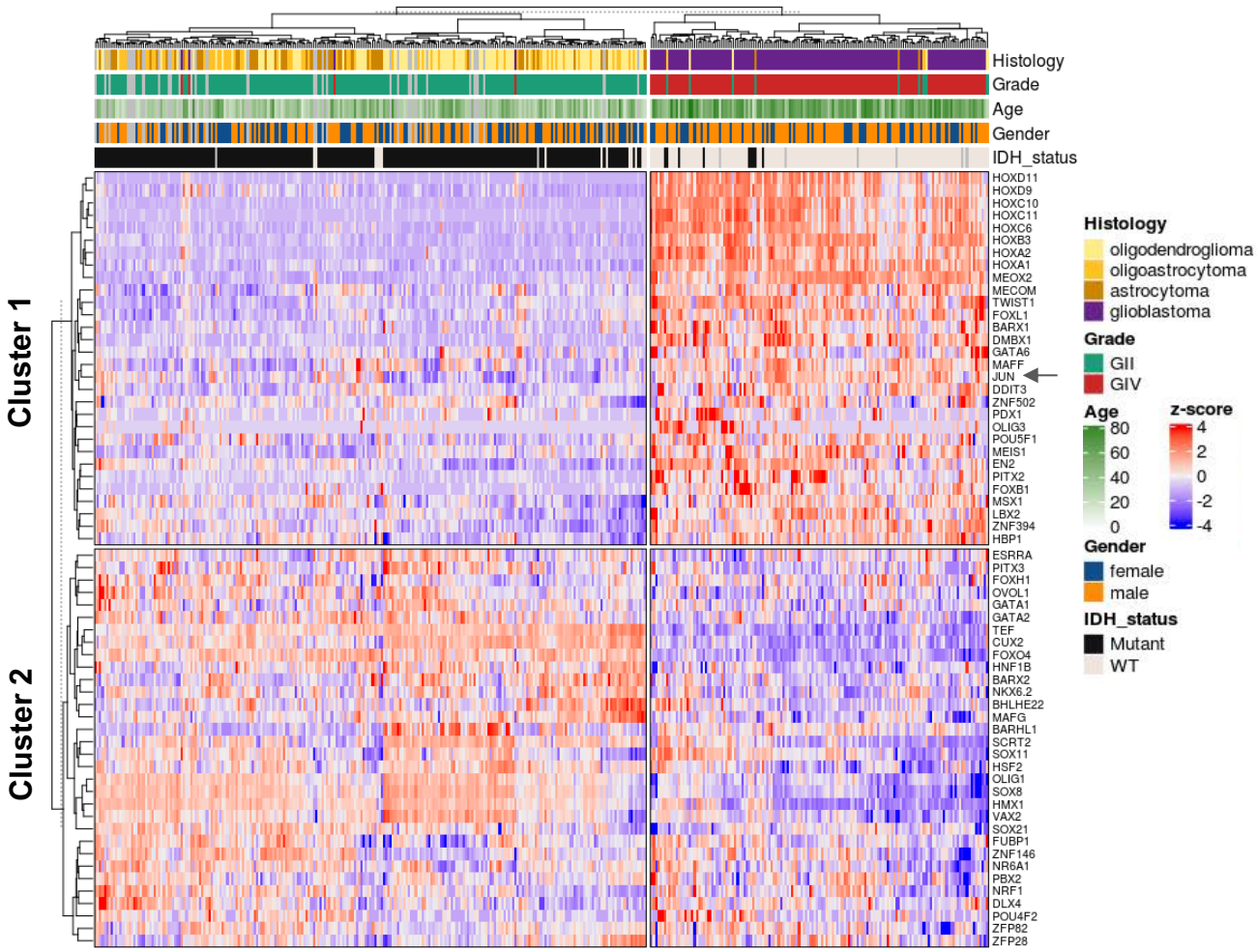


Figure 3.25. Deregulation of expression of selected genes coding for transcription factor across gliomas of various grades. Hierarchical clustering of genes coding for transcription factors (TFs) predicted to be bound in the promoters of genes overexpressed in WHO grade IV gliomas. A Ward's minimum variance method was

used to cluster the TCGA patients and genes, which identifies the strongest clustering structure (Cluster 1 and Cluster 2). Patients with missing clinical information on histology, grade, age or gender are illustrated in grey.

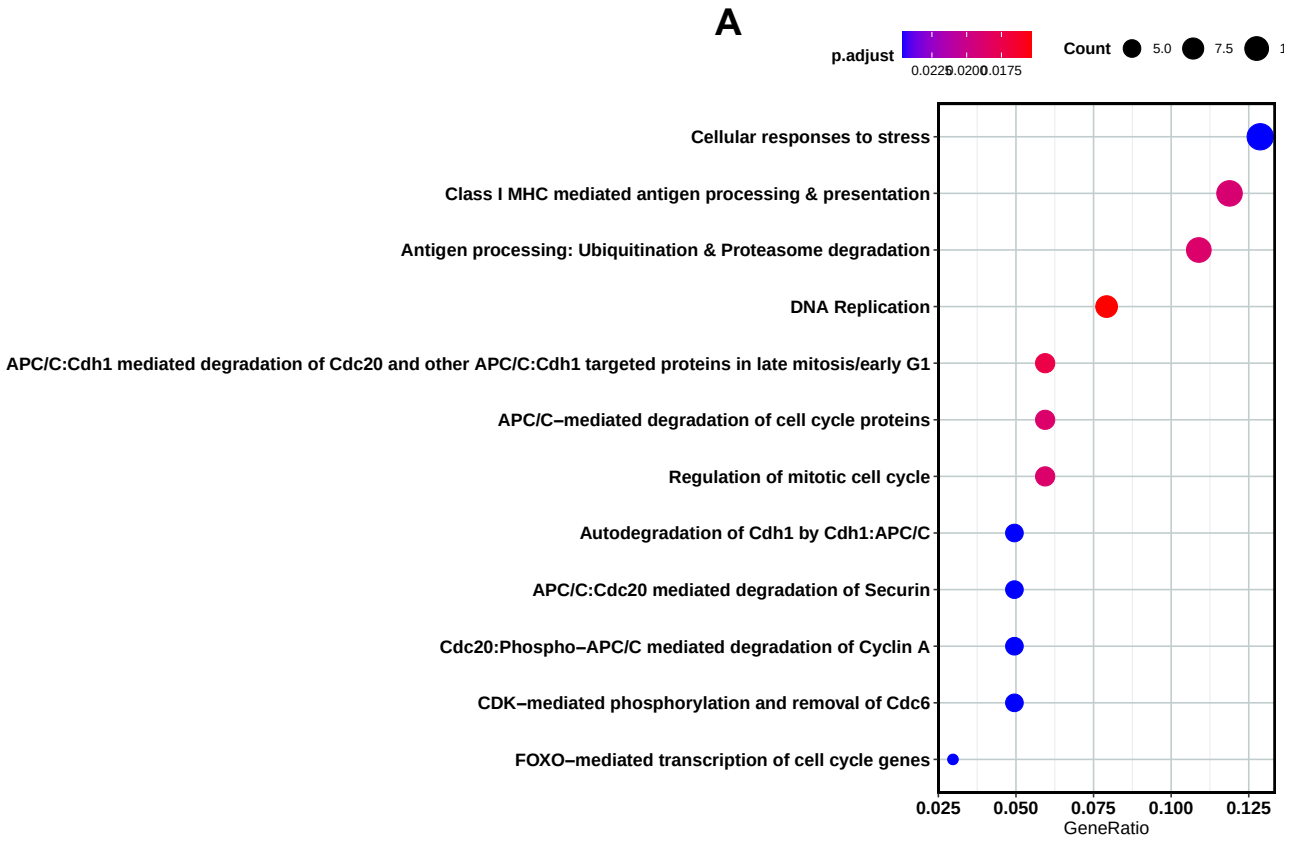
Additionally, to evaluate the levels of these TFs in cultured glioma cells and tumour specimens, expression of each TF coding gene was depicted (Fig. 3.26). HOX-related genes were significantly overexpressed in WHO grade IV versus grade II gliomas (Wilcoxon rank-sum and BH padj <0.05). Moreover, crucial glioblastoma TFs such as MEOX2, TWIST1, MAFF, DDIT3, MEIS1 were overexpressed as well in this type of tumour. The vast majority of the genes coding for TFs associated with higher glioma grades (Fig. 3.26, cluster 1) were statistically significantly overexpressed in high grade gliomas (90% of the cases, 27/30 TFs), whereas the genes coding for TFs that were associated with lower grade gliomas (Fig. 3.26, cluster 2), showed significantly different expression across glioma grades in only 53.15% of the cases (17/32 TFs). Expression of TF encoding genes in LN18 and LN229, and the gene expression medians were consistent with the patterns detected in the tumour samples. However, due to the low number of samples with RNA-seq data from the cultured cells, statistical significance was not calculated (Fig 3.26). Some genes coding for TFs (PDX1, OLIG3, POU5F1, PITX3, FOXH1, OVOL1, GATA1, HNF1B, BARX2, POU42F) were expressed at a very low level (Fig. 3.26). This implies that even though their motifs were found in the promoter region of key genes, they might play specific roles in processes that are not recapitulated in cultured cells. On the other hand, increased expression of certain TFs and enrichment of their binding sites in the open chromatin regions in both cultured cells and glioblastoma specimens point to the vital role of these TFs for the tumour pathology.

Figure 3.26. Deregulation of TF expression in gliomas of different grades and in cultured glioma cells. TCGA RNA-seq data (grade II: 248 patients, grade IV: 160 patients) and RNA-seq data from cultured LN18 and LN229 glioma cells were used to visualize transcription factor (TF) expression from the two main dendrogram clusters. The Wilcoxon rank-sum statistical test and the Benjamini-Hochberg (BH) correction were used to calculate the adjusted p-values for statistical differences between glioma grades. In glioma patients, TFs with a logarithmic expression of zero or nearly zero had no statistical validity.

3.3.6. Specific TF-targeted genes are associated with glioma progression

Starting with WHO grade IV-specific TFs (Fig. 3.16, right panel), we selected all the genes (166 genes) that had at least one TF motif (WHO grade IV-specific) instance prediction in the promoter and were significantly upregulated in GBMs compared to LGGs (student's t-test and FDR < 0.05). Then, we performed a pathway enrichment analysis on all of these genes to learn more about their biological functions (Fig. 3.27A, Fig. 3.27B). We found that many of those genes were involved in cellular stress responses, DNA replication, cell cycle, and antigen processing and presentation. This

suggests that identified WHO grade IV-specific TFs may influence expression of critical genes involved in glioma progression. Furthermore, expression of some genes showed association with binding of the JUN kinase (Fig. 3.27B, highlighted), which is required for the phosphorylation and activation of many TFs like c-Jun and c-Jun related factors (260,261).



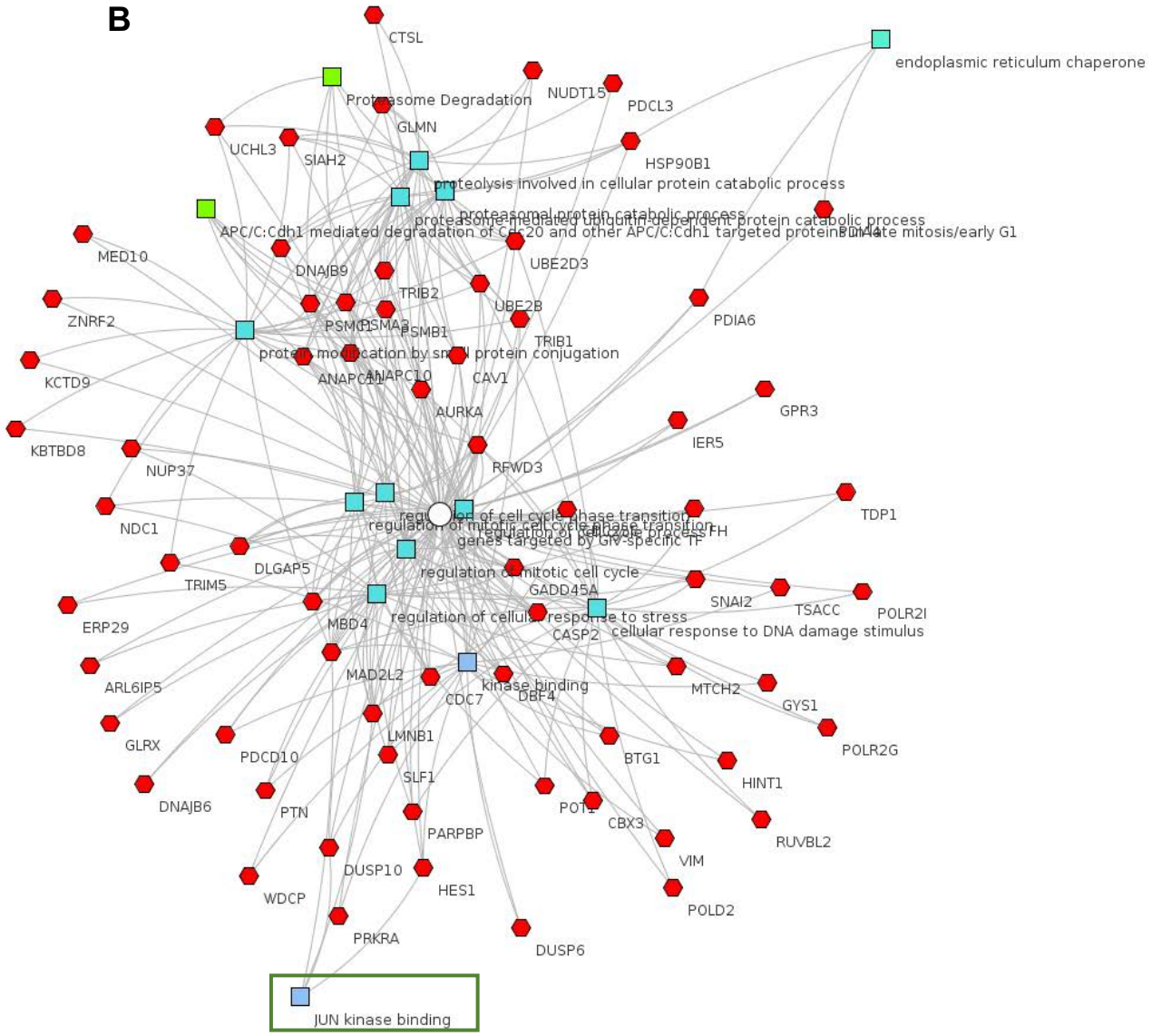


Figure 3.27. Putative targets of GIV glioma-specific transcription factors. (A) Reactome analysis of genes containing grade IV-specific transcription factor (TF) motifs in their promoters (TSS ± 1.5 kb) displaying regulatory pathways. To account for multiple testing, the BH procedure was used. (B) Top enriched categories of canonical pathways (toppcluster) containing significantly up-regulated genes in GBMs and containing a glioblastoma-specific TF motif in cis-regulatory regions. GO: Molecular Function (blue), GO: Biological Processes (cyan) and WikiPathways (green) databases were used to link genes (red) and their biological implications.

3.3.7. Expression of JUN positively correlates with expression of its targets

We hypothesized that high levels of c-Jun will increase the mRNA levels of its target genes. We calculated the correlation between *JUN* mRNA and the expression of c-Jun targets in WHO grade II and IV gliomas in the TCGA dataset (Fig. 3.28). We found a positive and significant Pearson's correlation (p.adjusted < 0.05) in all of the cases, with the highest positive correlation for genes

encoding interferon related developmental regulator 1 (*IFRD1*), vimentin (*VIM*), and FOS Like 2 (*FOSL2*). Using publicly available reverse protein phase assay (RPPA) data, we checked the level of the phosphorylated c-Jun (serine 73, S73) and expression of sixteen genes (Fig. 3.29). The higher levels of phosphorylated c-Jun significantly correlated with mRNA levels of c-Jun predicted targets. *FOSL2* and *VIM* were the most positively correlated targets, judging by the mRNA-to-mRNA correlation (Fig. 3.28) but also from the phosphorylated c-Jun-to-mRNA (Fig. 3.29).

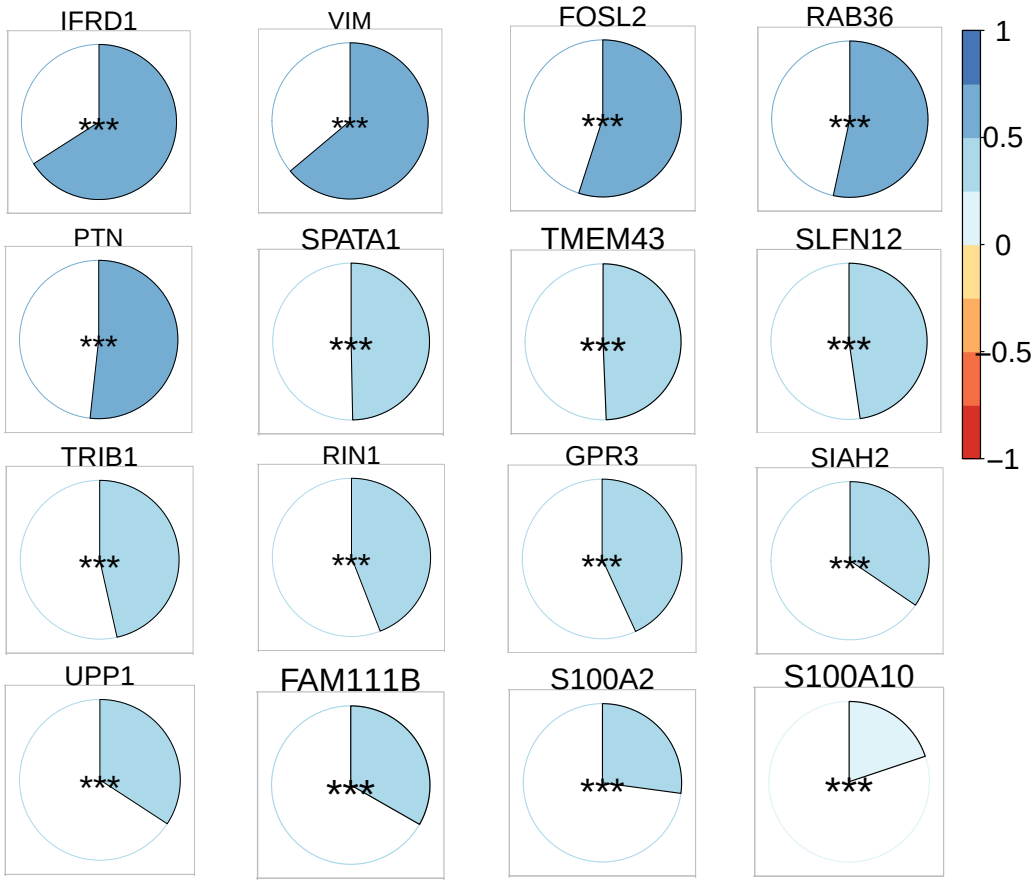


Figure 3.28. Correlation between *JUN* expression and target gene expression at the mRNA level. Genes are ordered based on the obtained Pearson correlation (blue to red scale) and associated p-values were corrected by multiple testing (*padj < 0.05, **padj < 0.01 and ***padj < 0.001).

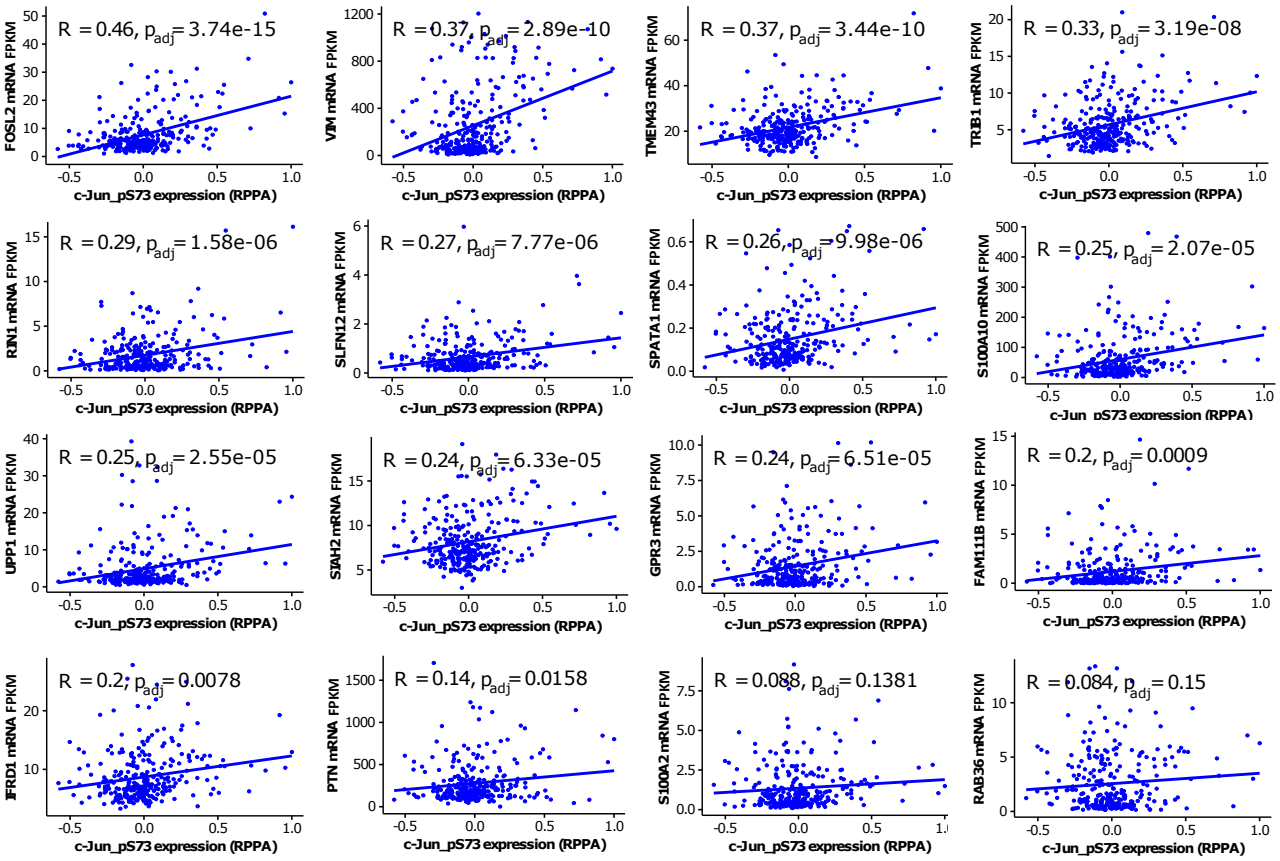


Figure 3.29. Correlation between c-Jun phosphorylation and target gene expression at the mRNA level. Genes are ordered based on the obtained Pearson correlation and associated p-values were corrected by multiple testing. RPPA stands for reverse-phase protein array and FPKM stands for Fragments Per Kilobase of transcript per Million mapped reads.

3.3.8. Survival analyses based on expression of c-Jun target genes

We also analysed the patient survival considering the expression of the *JUN* and its target genes in GBM (Fig. 3.30) and LGG samples (Fig. 3.31). High expression levels of *FOSL2*, *GPR3*, *RIN1* and *UPP1* (Log-rank p-values<0.05) were associated with a worse prognosis in GBM patients. The same analysis limited to patient survival in the LGG group showed that high expression of *JUN* and c-Jun targets was associated with a worse prognosis in all cases.

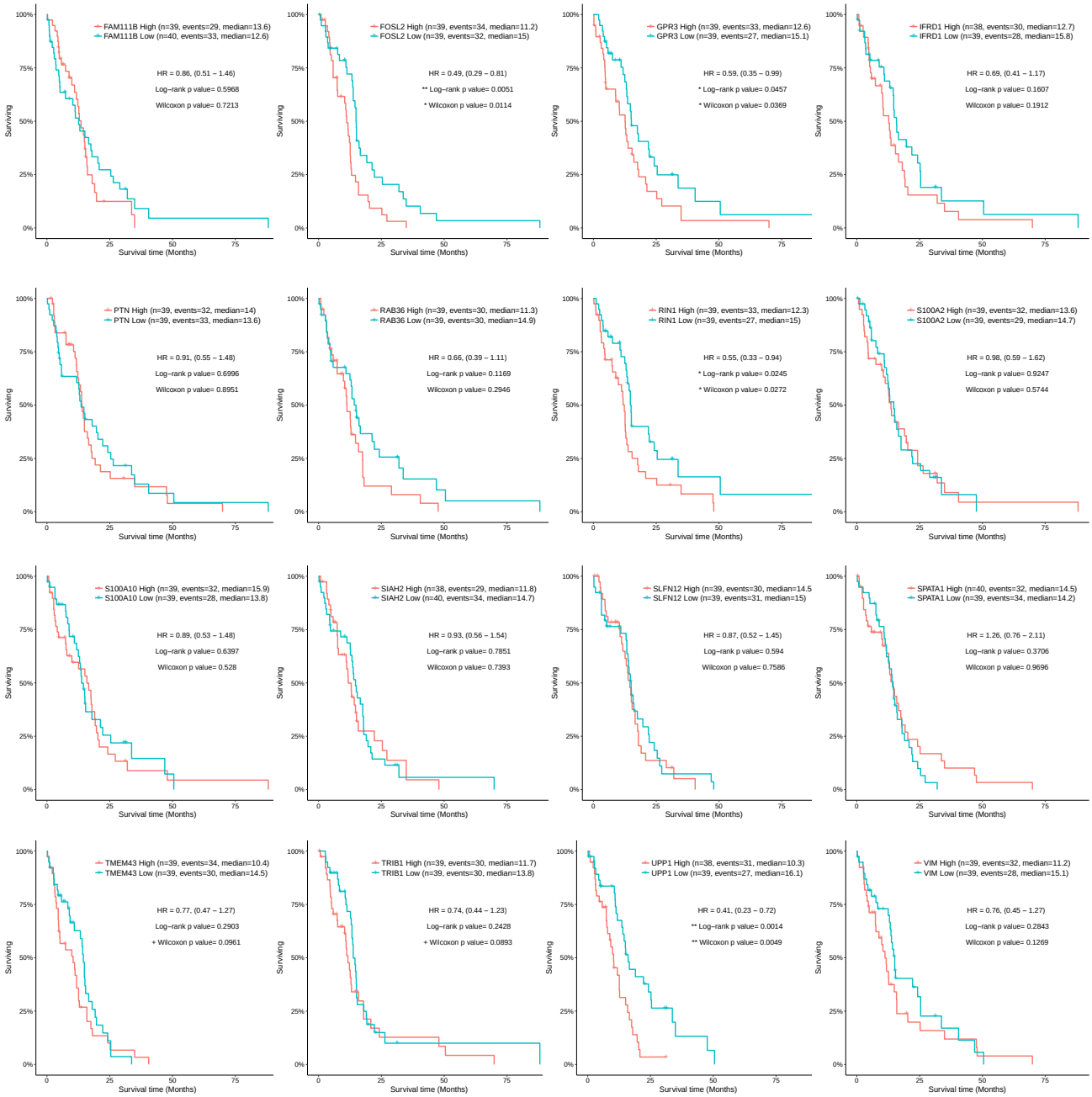


Figure 3.30. Expression of c-Jun targets and GBM patient survival (TCGA). Median overall survival (OS) was estimated based on gene expression of c-Jun target genes using the Kaplan-Meier method (cut-off defined by splitting datasets in 25% lower expressing patients vs 25% higher expressing patients). Vertical marks indicate censorship, the number of patients and median survival in months per group are shown, and the hazard ratio (HR) is defined. The log-rank test and Wilcoxon p-values are displayed.

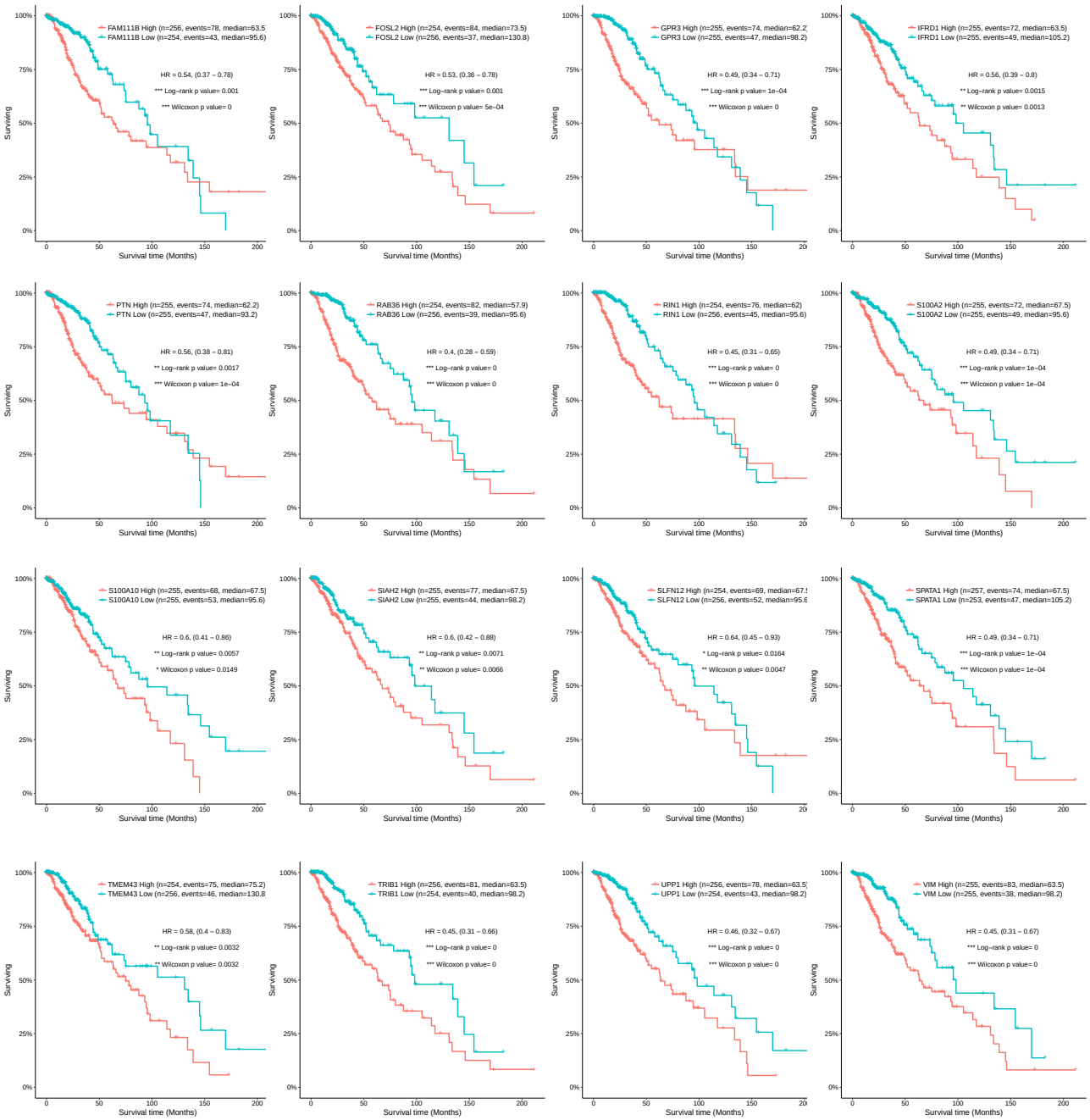


Figure 3.31. Expression of c-Jun targets and LGG patient survival. Median overall survival (OS) was estimated based on gene expression of c-Jun target genes using the Kaplan-Meier method (cut-off defined by splitting datasets in 25% lower expressing patients vs 25% higher expressing patients). Vertical marks indicate censorship, the number of patients and median survival in months per group are shown, and the hazard ratio (HR) is defined. The log-rank test and Wilcoxon p-values are displayed (*p-value < 0.05, **p-value < 0.01 and ***p-value < 0.001).

3.3.9. Distal regulatory regions in gliomas are highly abundant in motifs for the c-Jun and other bZIP TFs

Gene regulation is a multi-step process involving binding of TFs to promoters and enhancers, which are gene-distal cis-regulatory sequences that control the spatiotemporal and quantitative expression dynamics of target genes (262,263). We had identified and described the enhancers landscape in gliomas based on the analysis of active histone H3K27ac marks by ChIP-seq in PAs, DAs and GBMs (234). In the present analysis, we integrated H3K27ac peaks (potential enhancers) that were recurrent in GBMs (found in at least 5 patients) and in DAs (found in at least 5 patients) and we found that the number of consistent enhancers was lower in DAs than in a GBMs cohort (Fig. 3.32 first outer track). Furthermore, from the previous step, we selected active enhancers that were common in all glioma grades and defined a set of putative glioma enhancers (Fig. 3.32, second outer track). Subsequently, we searched for all predicted TF motifs within these regulatory regions, which yielded 7,571 TF motif instances in total (Fig. 3.32, third outer track). In glioma enhancers, a total of 94 binding sites for the c-Jun were identified (Fig. 3.32, third outer track, brown marks).

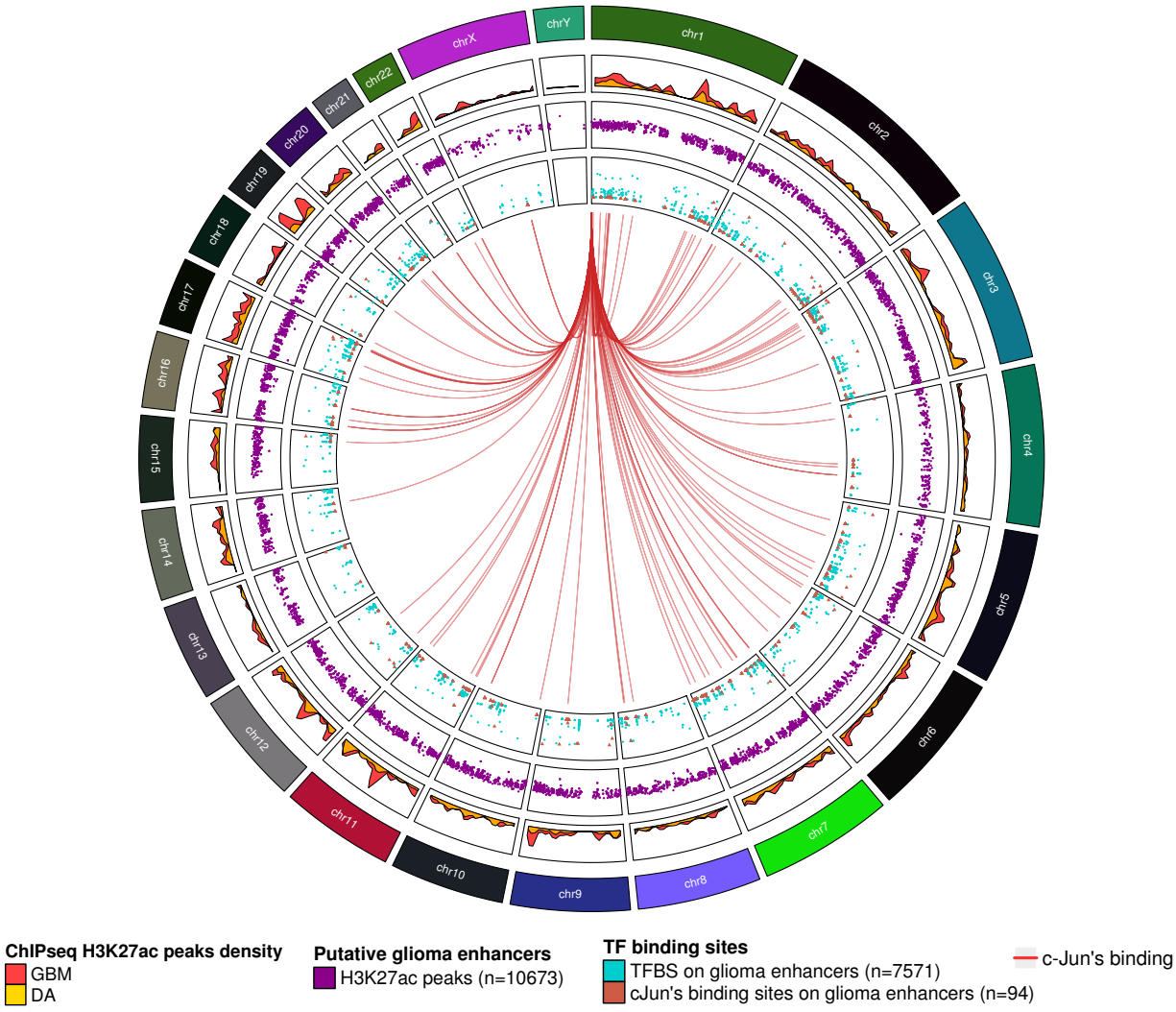
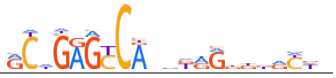


Figure 3.32. Glioma enhancers and their intersection with TFBS. The glioma enhancer atlas depicts: 1) The ChIP-seq H3K27ac peak density in glioblastomas (GBMs) and diffuse astrocytomas (DAs) patient samples (1st track); 2) Identified putative glioma enhancers (234) (shown in the second track); 3) A total number of TFBS motif predictions within enhancers (in turquoise) and the total number of c-Jun binding sites (in brown) are shown separately (3rd track). Each JUN motif found in glioma enhancers is linked to the *JUN* gene's chromosomal position.

We calculated hypergeometric probabilities to quantify the enrichment of different TFs within glioma enhancers and demonstrated that several bZIP TFs, including c-Jun, were significantly abundant (Table 3.2, Suppl. Table 1). This finding suggests that the bZIP TF class, which includes the Fos-, Jun- and Maf-related families, could be important in the regulation of gene expression not only at the gene promoters but also distal-regulatory elements in GBMs. In fact, glioma consensus H3K27ac peaks are mostly found in distal intergenic regions, followed by intronic regions (Fig. 3.33A). This suggest that many DNA sequences in introns harbour important elements for tumour-related transcriptional regulation.

Finding so many enhancers enriched in c-Jun motifs and other bZIP TF motifs (Fig. 3.33B) was unexpected and required more attention. In particular, all 94 c-Jun motifs identified in glioma enhancers from H3K27ac ChIP-seq analysis (Table 3.2, Suppl. Table 1) were found in ATAC-seq LN18 and LN229 cells. Some of these instances were not found in the glioblastoma ATAC-seq dataset (Fig. 3.33, GBM1 and GBM2 TFBS predictions). Comparison of enhancer H3K27ac peaks detected in GBM and DA specimens revealed only few significantly different regions between these two sample cohorts (Suppl. Table 2). This shows that most of the regions have activating histone marks both in DAs and GBMs.

Table 3.2. Top 15 TF binding probabilities in glioma enhancers. Obtained p-values were calculated by the hypergeometric test and corrected using the Benjamini-Hochberg method.

TF model	Occurrences in enhancers	Occurrences in the genome	Hypergeometric test (p-value)	Hypergeometric test (adj. p-value)	Consensus sequence
FOSL1_0_A	128	597	1,20723E-42	7,58142E-40	
FOSL2_0_A	137	726	4,86818E-39	1,52861E-36	
JUNB_0_A	115	544	7,50317E-38	1,57066E-35	
JUND_0_A	123	623	3,12176E-37	4,90116E-35	
FOS_0_A	119	626	1,9905E-34	2,50007E-32	
BACH2_0_A	67	273	1,21946E-26	1,27636E-24	
JUN_0_A	94	524	1,38731E-25	1,24462E-23	
NFE2_0_A	71	344	2,45514E-23	1,92728E-21	
BACH1_0_A	87	494	3,08603E-23	2,15336E-21	
FOSB_0_A	96	585	3,7038E-23	2,32598E-21	
NF2L2_0_A	63	279	4,66643E-23	2,66411E-21	
MAFB_0_B	33	147	9,6426E-13	5,04629E-11	
ZN554_1_D	48	302	7,42587E-12	3,58726E-10	
MAFF_1_B	30	134	1,0909E-11	4,89348E-10	
MAFK_0_A	16	41	8,72968E-11	3,65483E-09	

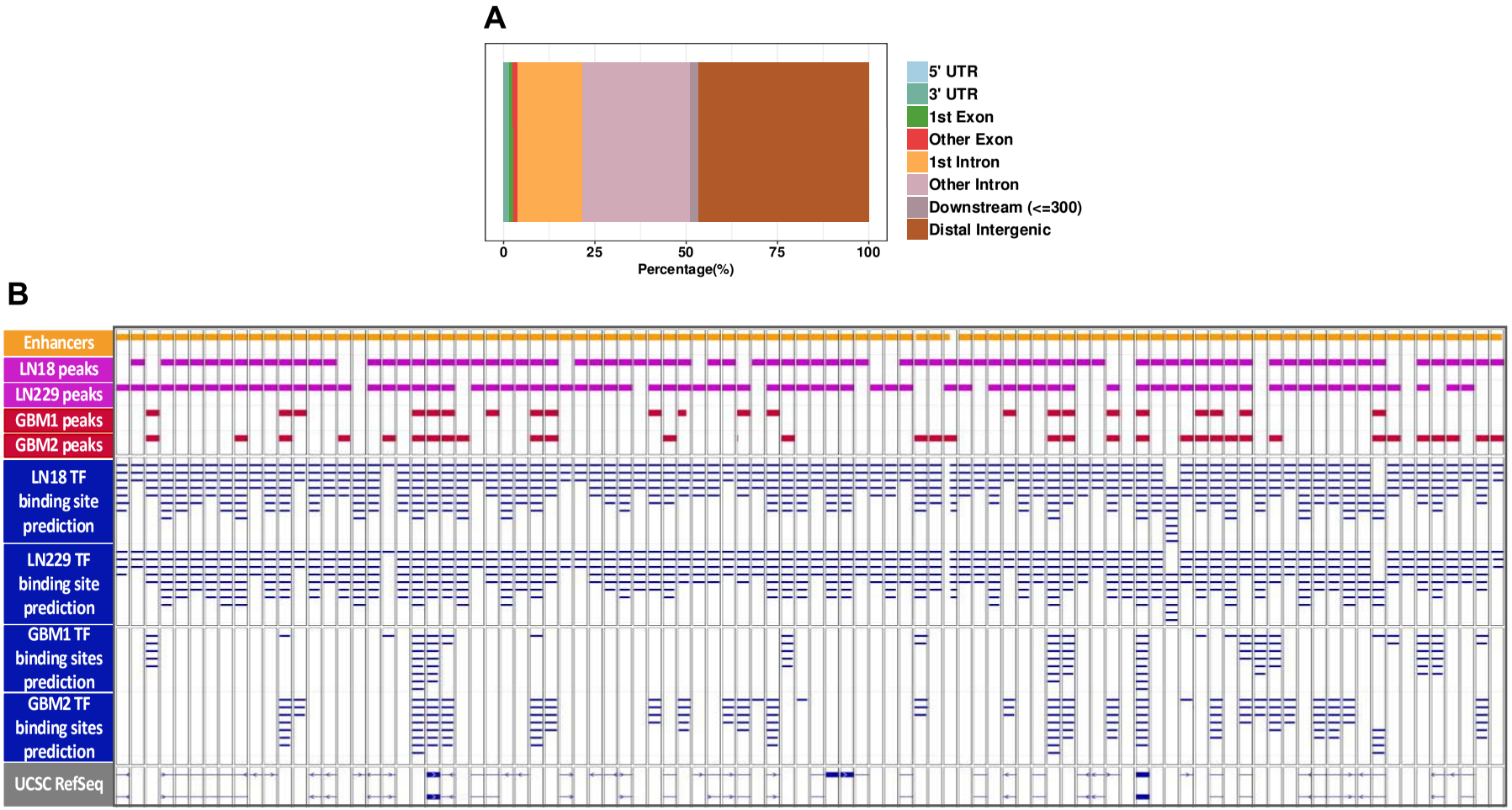


Figure 3.33. TFBS prediction in glioma enhancers in the context of c-Jun binding sites. (A) Distribution of glioma enhancer feature (GBM and DA H3K27ac peaks). (B) Glioma enhancers (n=94) harbouring bZIP binding sites and chromatin openness in glioma cell lines and glioblastoma specimens, as well as leucine zipper-related transcription factors (top candidates from Table2). The empty UCSC RefSeq track represents an intergenic region, whereas the regular blue lines or rectangle boxes represent intronic or exonic regions).

3.3.10. Genes affected by long-range contacts with enhancers harbouring c-Jun TFBS

We employed Hi-C chromatin data from a human brain development conformation analysis (171) to identify all enhancer-promoter contacts. This approach revealed several connections between enhancers with a c-Jun binding site and the promoters in the nearby regions. Enhancers with c-Jun binding sites were interacting with promoters of several protein coding genes, including *CDK18, ZBTB18, PRDM16, IRF2BPL, SCNN1G* and *TRIB2* (Suppl. Table 3). In the next step, we examined whether the expression of these genes correlates with *JUN* mRNA levels in glioma samples (WHO grade II to grade IV) and we found a significant, positive correlation for the majority of these gene targets, which was even more evident in WHO GIV gliomas (Fig. 3.34A). This result suggested that c-Jun by binding to adjacent enhancers may regulate expression of these genes.

Furthermore, we investigated whether DNA methylation in these enhancers could influence gene expression. TCGA data on DNA methylation at enhancers with c-Jun binding was downloaded, and we obtained data for three cytosines (cg02258482, cg12155676 and cg08003402) which were within c-Jun's motif ± 20 bp flanking regions. We analysed if the presence of CpG methylation at the enhancers harbouring a predicted c-Jun binding site correlated with the target gene expression. We only found four different promoters and their corresponding genes that were linked to identified enhancers. We found both positive and negative correlation in various enhancers (Fig. 3.34B, 3.34C, 3.34D, 3.34E), with the level and significance varying according to glioma grade.

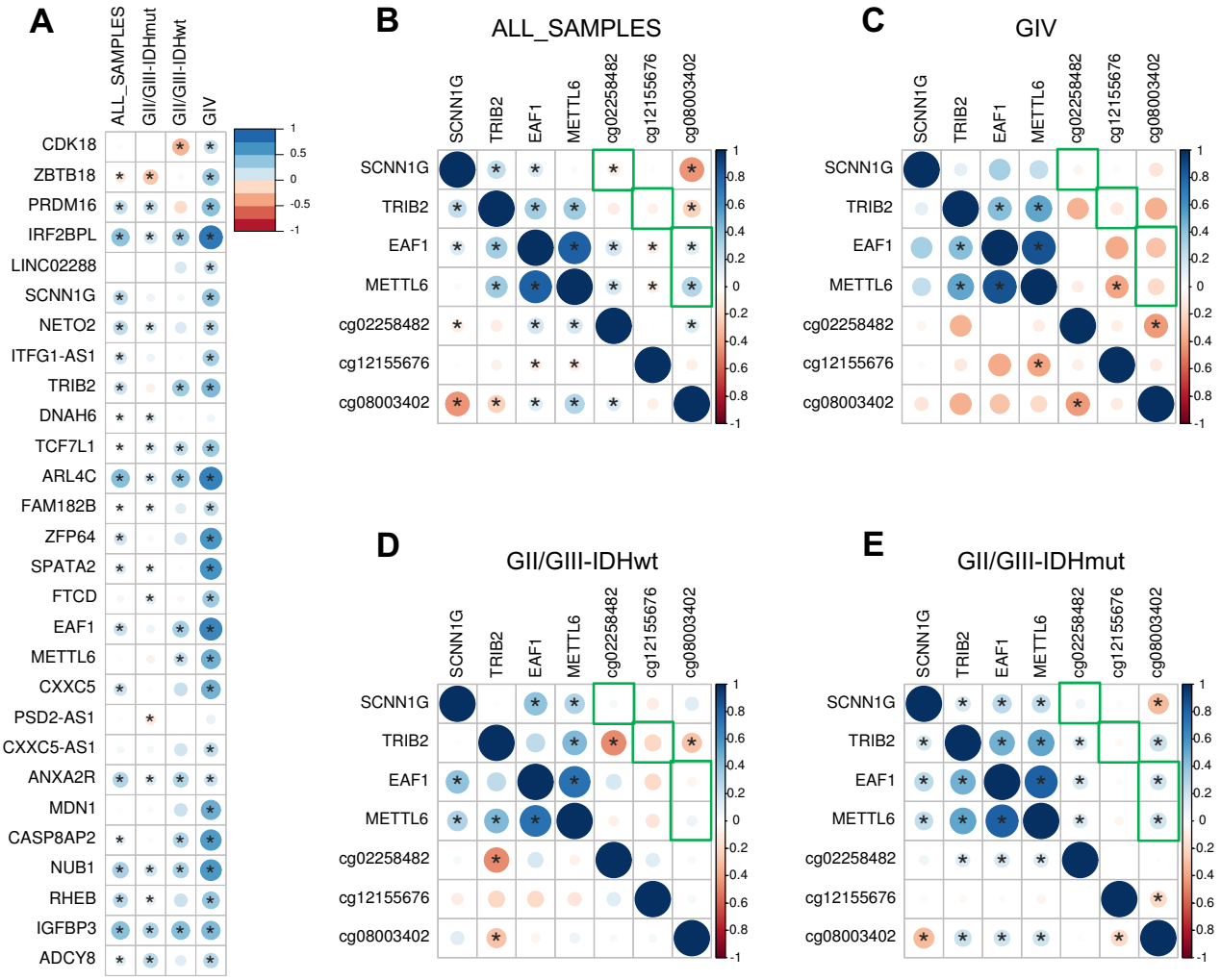


Figure 3.34. Long-range distance c-Jun targets in TCGA. (A) Expression of *JUN* and its putative enhancer-targeted genes was acquired from TCGA (LGG/GBM RNA-seq datasets) and correlated in ALL gliomas (GII, III, IV) and separately in GII/III IDH mutant, GII/III IDH wildtype and in GIV tumours (all were IDH-wildtype) and represented as a correlation plot. Data from TCGA LGG/GBM RNA-seq and 450k DNA methylation array was correlated in ALL gliomas (GII, III, IV) (B) and separately (C) in GIV tumours (all were IDH-wildtype), (D) in GII/III IDH mutant, (E) and in GII/III IDHT wildtype and represented as correlation plots. Biologically-relevant pairs (predicted from Hi-C data) of DNA methylation and gene expression are highlighted in green squares; these pairs were obtained by associating DNA methylation sites that may affect the putative enhancer c-JUN TFBS, which targets these specific genes. Pearson correlation significance is highlighted (*adjusted p-value < 0.05).

3.3.11. DNA methylation in the gene promoter of specific c-Jun targets differs in gliomas of various grades

DNA methylation regulates gene expression and, more importantly, it can affect a TF binding affinity to the DNA (128,129). As a result, in addition to examining DNA methylation patterns at enhancers (as described in section 3.3.10), we examined methylation patterns at promoters because

methylation regulation of gene expression in these regions is better understood than regulation by methylation in enhancers. Therefore, we examined methylation patterns in the promoters (+2 kb/-500 bp from TSS) of c-Jun target genes and in the *JUN* promoter in the tumour samples from low- and high-grade glioma patients (Fig. 3.35). We found primarily two clusters of genes that may be regulated by c-Jun TF: one cluster of genes with high DNA methylation in IDHmut and IDHwt GII/GIII but low in GIV tumours (*S100A10, S100A2, IFRD1, RUN1, RAB36, UPP1, SLFN12* and *VIM*), and the second cluster of genes with constant low DNA methylation (beta values ~0) regardless of a tumour grade (*PTN, FOSL2, FAM111B, SIAH2, SPATA1, TMEM43, TRIB1* and *GPR3*).

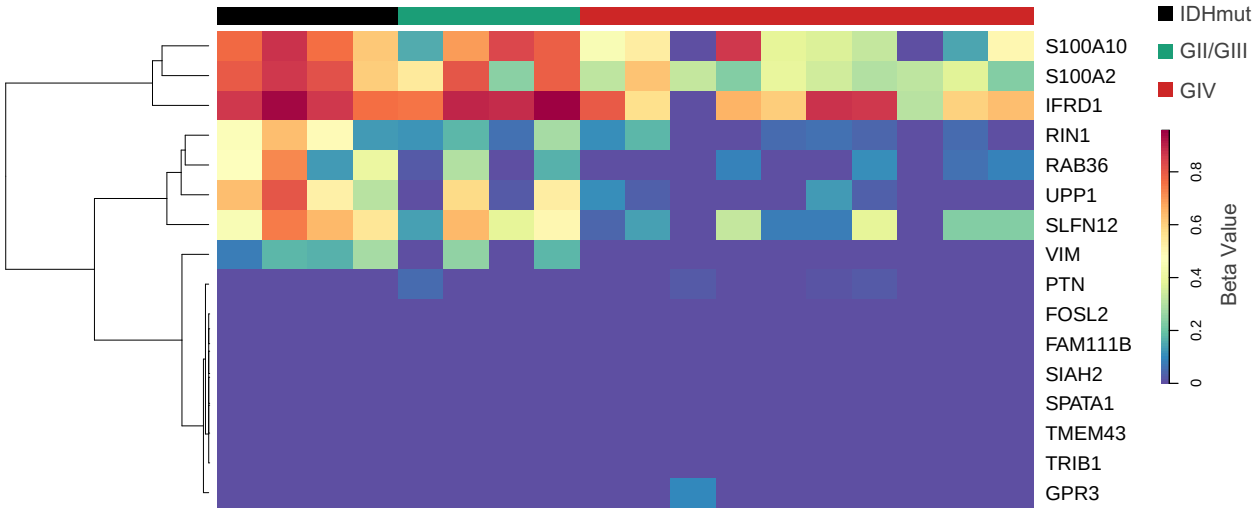


Figure 3.35. Median DNA methylation in the promoters of c-Jun targets. Heatmap of the supervised hierarchical clustering analysis showing median DNA methylation in the promoters of c-Jun targets (+2 kb/-500 bp from TSS) in glioma samples (4 IDHmut grade II/grade III, 4 IDHwt grade II/grade III and 10 grade IV gliomas). DNA methylation is expressed in beta values, 0 being hypomethylated and 1 hypermethylated cytosines.

Many c-JUN target genes did not differ in the promoter methylation between low- and high-grade glioma samples, including *IFRD1, S100A2*, which were at the verge of statistical significance (Fig. 3.36A). In this analysis all samples of WHO grade II/III gliomas were merged independently of IDH mutation status. The promoter methylation pattern of a second group of c-Jun target genes clearly differed (Fig. 3.36B) and DNA methylation at these regions was lower in GIV tumours (Chi-square test for two independent groups, FDR<0.05). This suggests that epigenetic mechanisms could control expression of these genes by affecting TF binding to the regulatory regions in DNA.

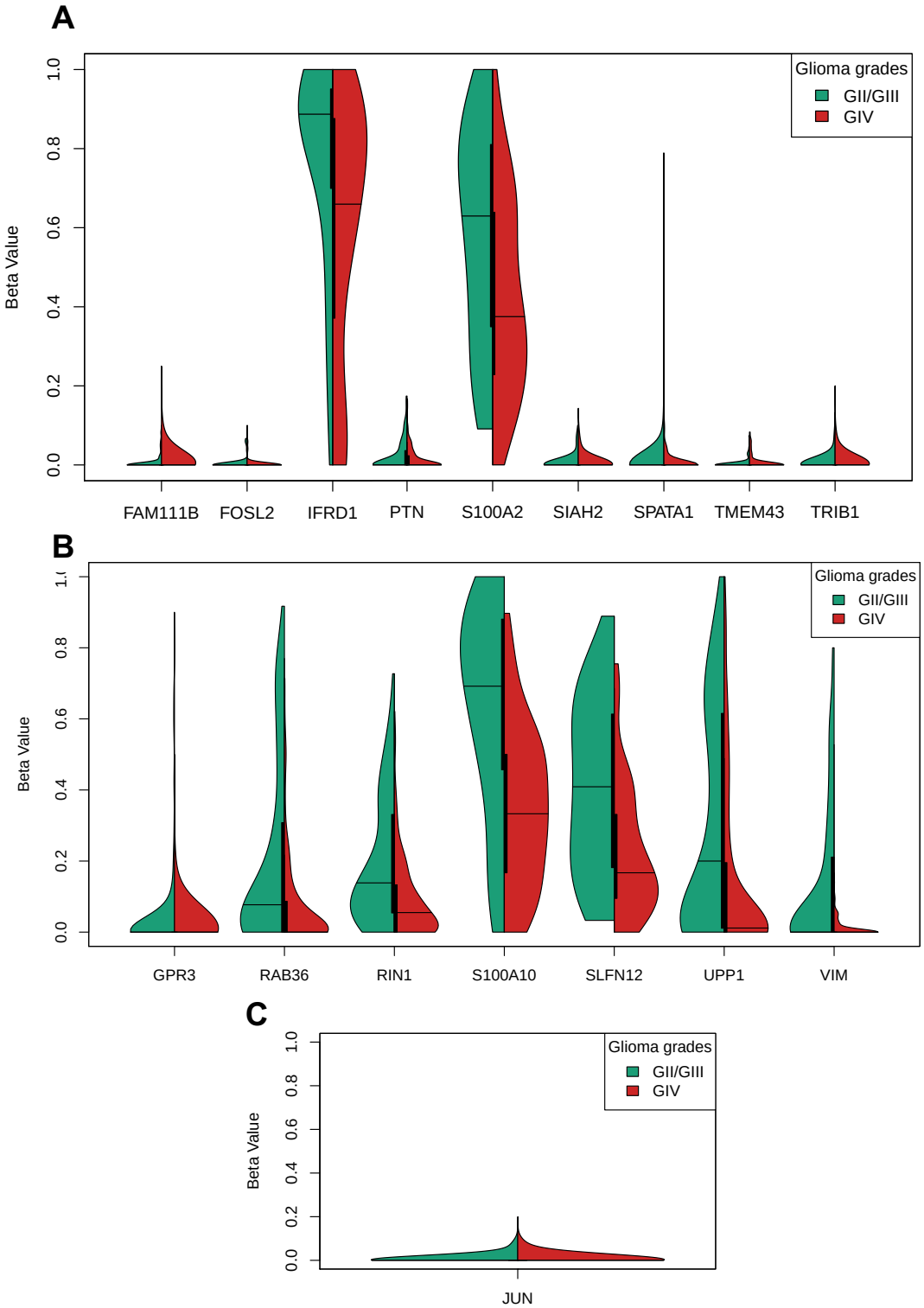


Figure 3.36. DNA methylation in the promoter of c-Jun targets and *JUN* coding gene. Statistically non-significant (A) and statistically significant (B) differences in beta values distribution of gene promoters with the predicted c-JUN TFBS in GII/GIII and GIV glioma samples. (C) Beta values distribution in the JUN promoter.

Methylation of the c-Jun promoter was low and similar in GII/GIII vs GIV glioma samples (Fig. 3.36C) suggesting that its differential expression is not regulated by DNA methylation. These findings

suggest that DNA methylation patterns at the promoters of c-Jun-regulated genes differ depending on the gliomas WHO malignancy grades.

We further analysed methylation patterns on cytosine-rich regions near or overlapping with c-Jun binding sites in gliomas. The c-Jun motif ("dvTGAGT**C**AYh", HOCOMOCOv11 model) contains 1 cytosine. Methylation of the cytosine within the c-Jun motif in c-Jun-targeted gene promoters was determined. There were no differences in the methylation of this specific residue in the c-Jun motif (Fig. 3.37). However, CpG methylation in the flanking regions of the predicted c-Jun binding differed between LGGs and HGGs (Fig. 3.37, brown boxes).

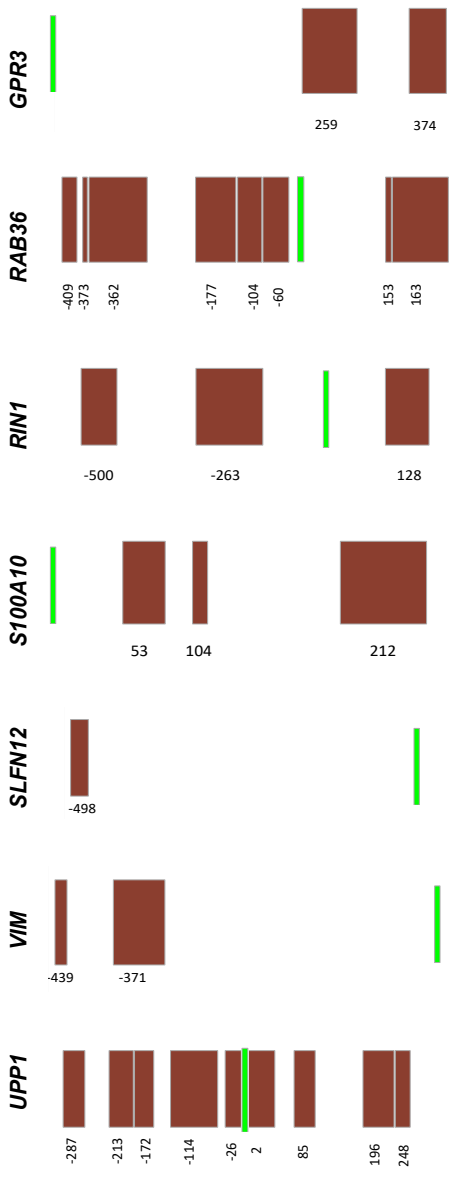


Figure 3.37. Methylation of DNA in C-rich regions of c-Jun targeted promoters. Distance of c-Jun motif to the beginning of differentially methylated C-rich regions between high- and low-grade gliomas. Green boxes represent a c-Jun predicted binding site while brown boxes show each C-rich region that was found significantly differently methylated between low- and high-grade glioma samples (Chi-squared test at significance level $p<0.05$).

To cross-validate DNA methylation levels in the promoters of c-Jun targets, we used the TCGA data (see section 2.2.13). Despite the fact that our dataset and the TCGA dataset had different cytosine coverage, the TCGA dataset contains more samples, and offers an independent dataset. In this analysis, we found that the IDH phenotype causes a clear hypermethylated pattern in the majority of the genes (Fig. 3.38A). Moreover, many c-Jun targets (*PTN, SIAH2, FAM111B, TMEM43, FOSL2, TRIB1,* and *SPATA1*) are hypomethylated regardless of tumour grade and IDH status (Fig 3.35A). When comparing GII/GIII-IDHmut gliomas with GIV gliomas, remarkable methylation differences in the promoters of *S100A2, RAB36, RIN1, UPP1* and *VIM* were detected. IDHwt LGGs have similar patterns of DNA methylation in those genes as GBMs. In most cases, DNA methylation was found to be negatively correlated with gene expression, which is known as the classic DNA methylation effect on gene expression (Fig. 3.38B).

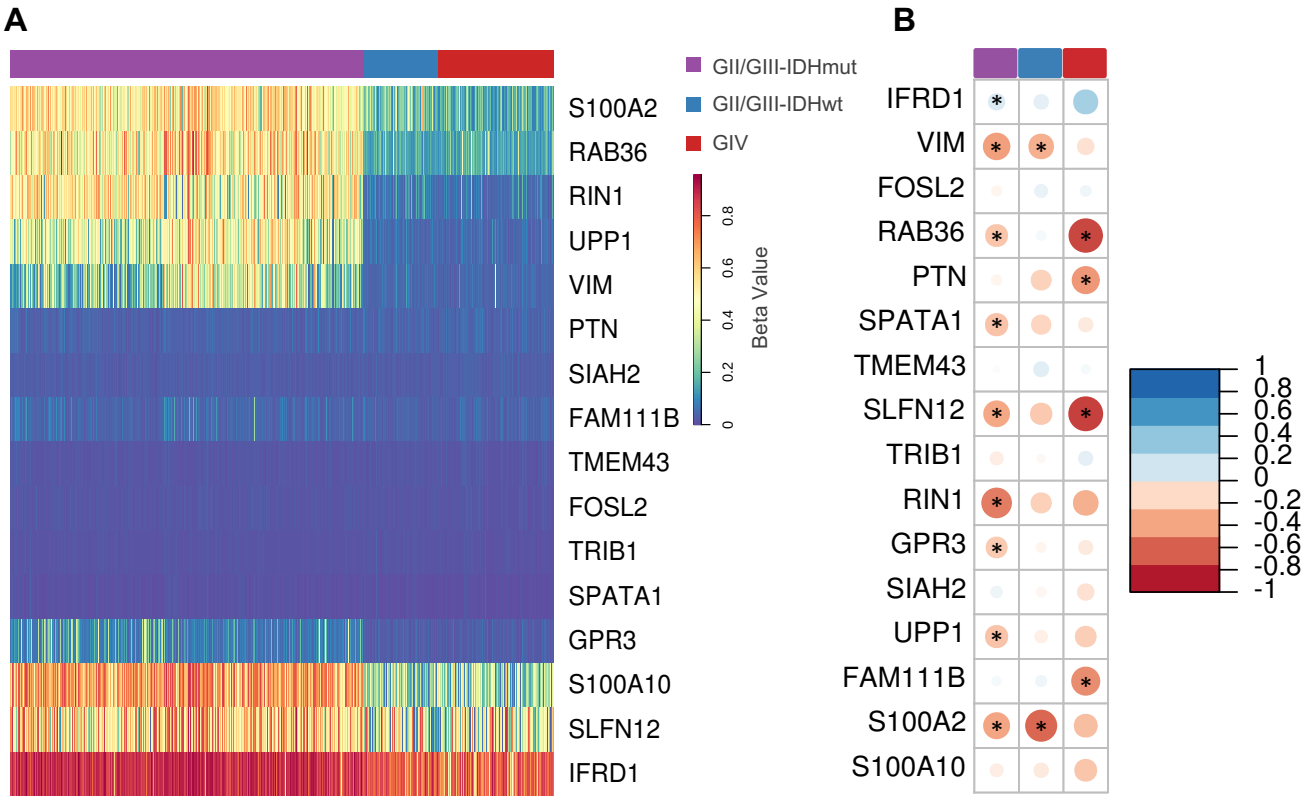


Figure 3.38. Cross-validation of DNA methylation and gene expression interrelationship for the predicted c-Jun targets in TCGA. (A) Heatmap of the supervised hierarchical clustering analysis showing median DNA methylation within the gene promoters of c-Jun targets (TSS ± 1 kb) in gliomas (369 IDHmut and 69 IDHwt grade II/grade III, and 121 grade IV gliomas). DNA methylation is expressed in beta values, with 0 indicating hypomethylated cytosines and 1 indicating hypermethylated cytosines. (B) Correlation of DNA methylation with expression of c-Jun target genes in 356 GII/GIII-IDHmut (violet), 78 GII/GIII-IDHwt (blue), and 44 grade IV gliomas (red). The graph depicts Pearson's correlation (blue to red scale) and its significance is highlighted (*adjusted p-value < 0.05).

3.3.12. c-Jun binds to the *VIMENTIN* gene promoter

The presented data showed an interesting link between c-Jun and expression of the gene coding for Vimentin, an intermediate filament that provides support and anchoring to the cell, and is involved in cell migration, adhesion, and cell division (264,265). Vimentin is overexpressed in several cancers and has been linked to accelerated tumour growth, invasion, and a poor prognosis (266,267).

To verify if c-Jun binds to the *VIMENTIN* promoter, we performed an electrophoretic mobility shift assay (EMSA) using nuclear extracts from LN229 and LN18 glioblastoma cells, LGG-derived cell cultures (WG12) and normal human astrocytes (NHA) (Fig. 3.39A). We found nuclear extracts from glioma cells bound a fragment of DNA from the *VIMENTIN* promoter producing a clear shift of the labelled probe. The probe was further shifted after adding anti-c-Jun antibody (Fig. 3.39B), that indicates the presence of c-Jun in the DNA-protein complex. Reduction of DNA-protein complexes after adding the unlabelled probe competing for c-Jun binding confirms a specificity of the binding. The reduced binding of nuclear extracts from NHA or WG12 cells to the *VIMENTIN* promoter probe (Fig. 3.39A) indicates the lower amount of c-Jun protein able to bind to the *VIM* promoter.

The EMSA results from three experiments were evaluated by densitometry and quantified. Significant differences in the c-Jun binding to the *VIMENTIN* promoter were found between glioblastoma cells and LGG-derived cells (Fig. 3.39C). Surprisingly, the c-Jun binding to the *VIMENTIN* promoter was relatively high in NHA, nevertheless, the shifted bands were less intense than those detected with extracts from LN18 and LN229 cells (Fig. 3.39). The findings suggest that c-Jun binds to the *VIMENTIN* promoter in glioma cells and there are differences in c-Jun levels or activation between glioma cells, with the higher binding in glioblastoma cells.

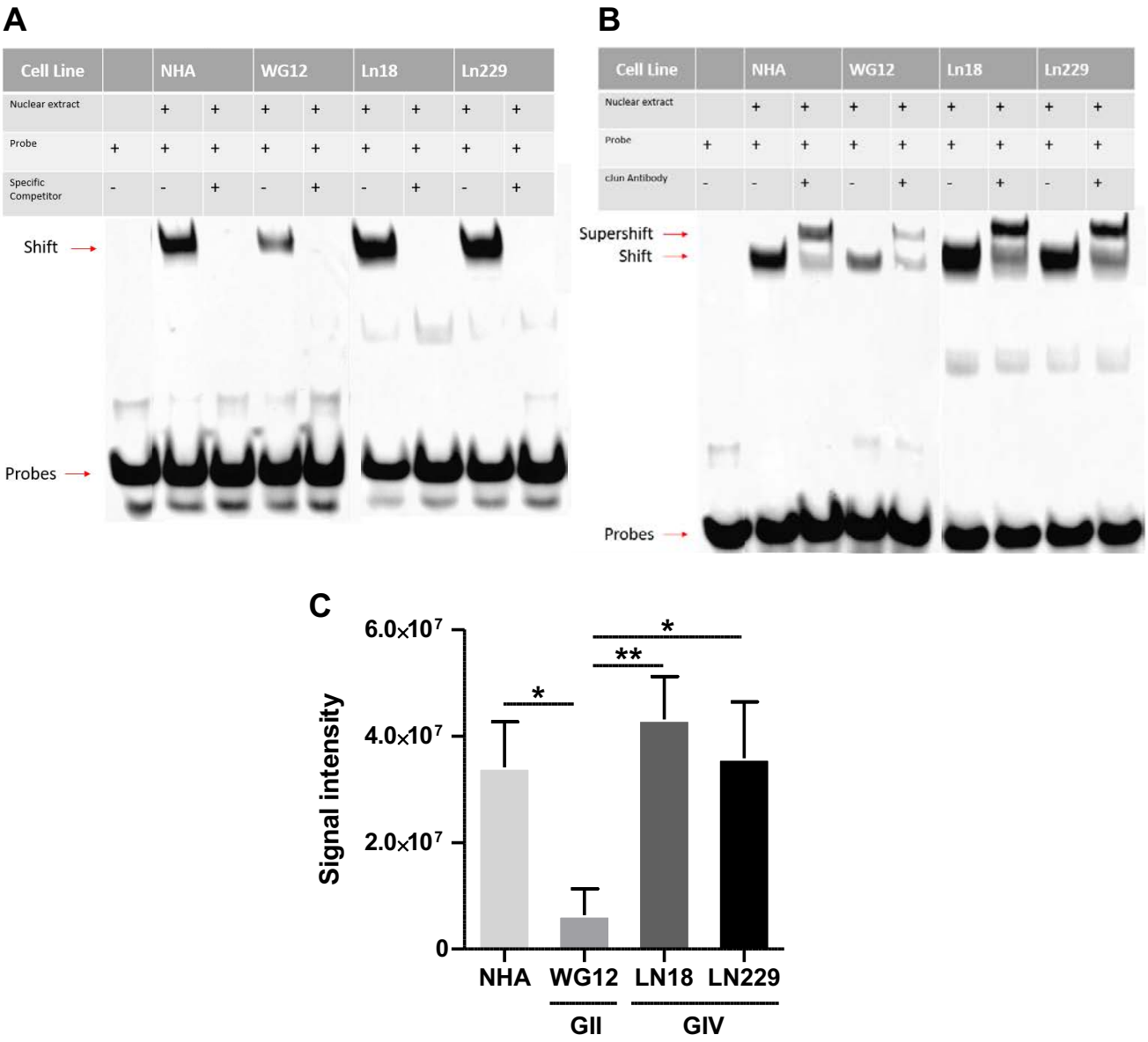
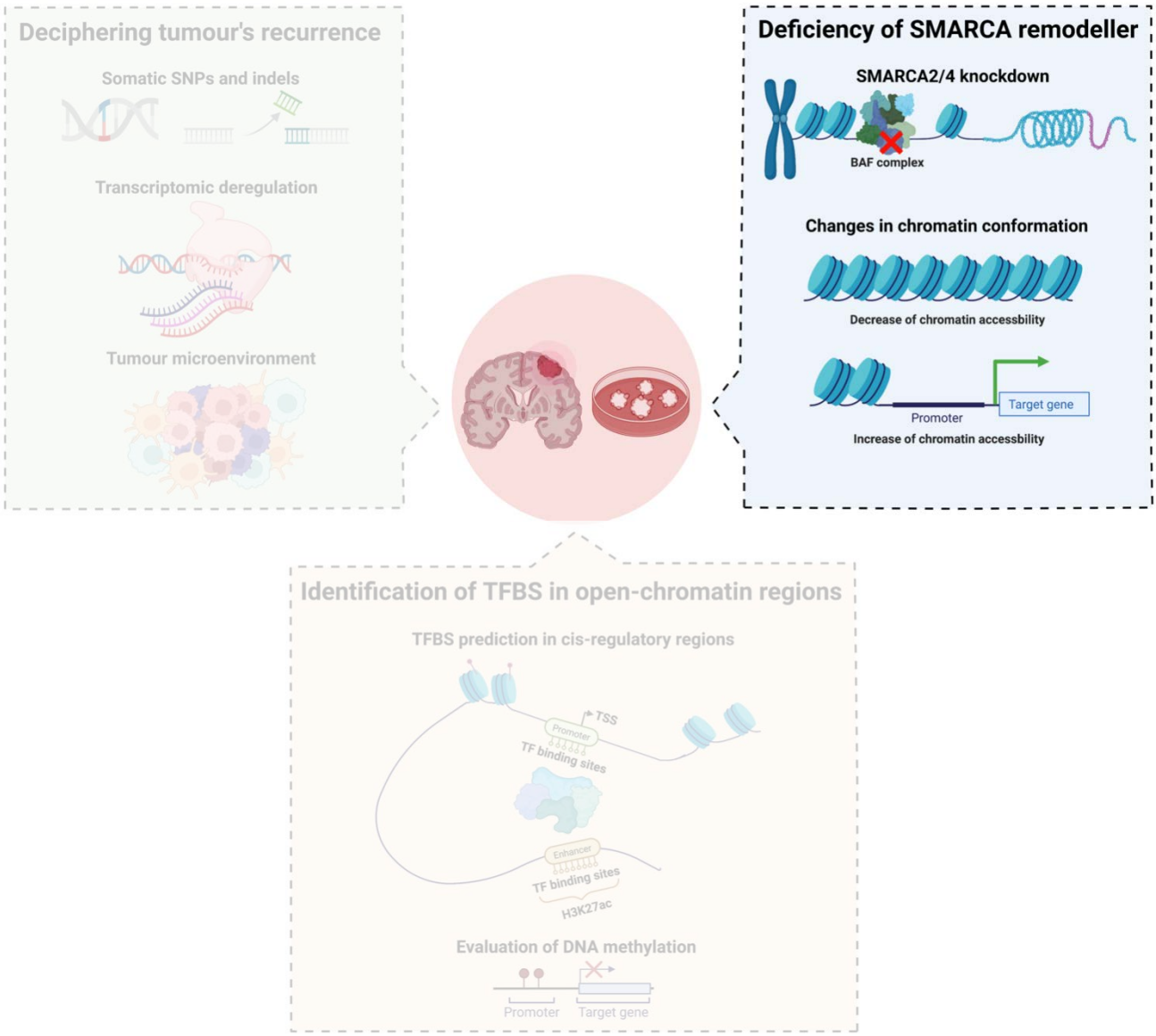


Figure 3.39. EMSA and EMSA supershift experiments confirms c-Jun binding to the *VIM* promoter. (A) Electrophoretic mobility shift assay (EMSA) using DNA probes alone, with nuclear proteins or in combination with the competitor (50-fold excess of the unlabelled probe). Nuclear extracts were isolated from normal astrocytes (NHA), patient-derived LGG cells (WG12), and LN18 and LN229 glioblastoma cells. (B) Anti-c-Jun antibody added to the reaction prior to gel loading produced a supershift, which confirms the presence of c-Jun in the DNA-protein complex. (C) Intensities of the shifted bands from individual experiments were determined by densitometry of blots and are presented as means ±SEM of three replicates; *P<0.05, **P<0.01 one-way ANOVA and Tukey multiple comparison post-test.

3.4. Altering chromatin accessibility by SMARCA2 and SMARCA4 knockdown in human glioblastoma cells



3.4.1. *SMARCA2* and *SMARCA4* expression in gliomas and in SMARCA-deficient cells

Chromatin remodeling SWI/SNF complexes contain many core components, including SMARCB1/BAF47, SMARCC1/BAF155, SMARCC2/BAF170 as well as one of the two mutually exclusive ATPase subunits, SMARCA4/Brg1 and SMARCA2/Brm (268). Deleterious mutations in genes encoding its subunits have been reported in 25% of cancers, including glioblastomas (178).

We sought to study how modulation of SMARCA2 and/or SMARCA4 expression would affect chromatin openness, gene expression and functions of glioma cells.

Using TCGA datasets, we first determined the mRNA expression levels of *SMARCA2* and *SMARCA4* in gliomas. We found significant down-regulation of their mRNA levels in WHO grade III and grade IV gliomas as compared to lower-grade tumours and control brain samples (Fig. 3.40A, Fig. 3.40B). To reduce SMARCA2 and SMARCA4 expression in human glioma cells, LN18 glioma cells were transfected with siRNAs targeting SMARCA2 or SMARCA4 or combination of both. The silencing efficacy was confirmed by significant changes in expression of *SMARCA2* (Fig. 3.40C) and *SMARCA4* mRNA levels in glioma cells 48 hours post-transfection, determined with quantitative PCR (Fig. 3.40D). The mock transfected cells and cells transfected with the non-targeting siRNA were used as controls. The siRNAs used in this experiment were two commercial siRNAs (a and b) with a confirmed efficacy of silencing. Knockdown of SMARCA2 was less effective and the gene expression was reduced by 50%.

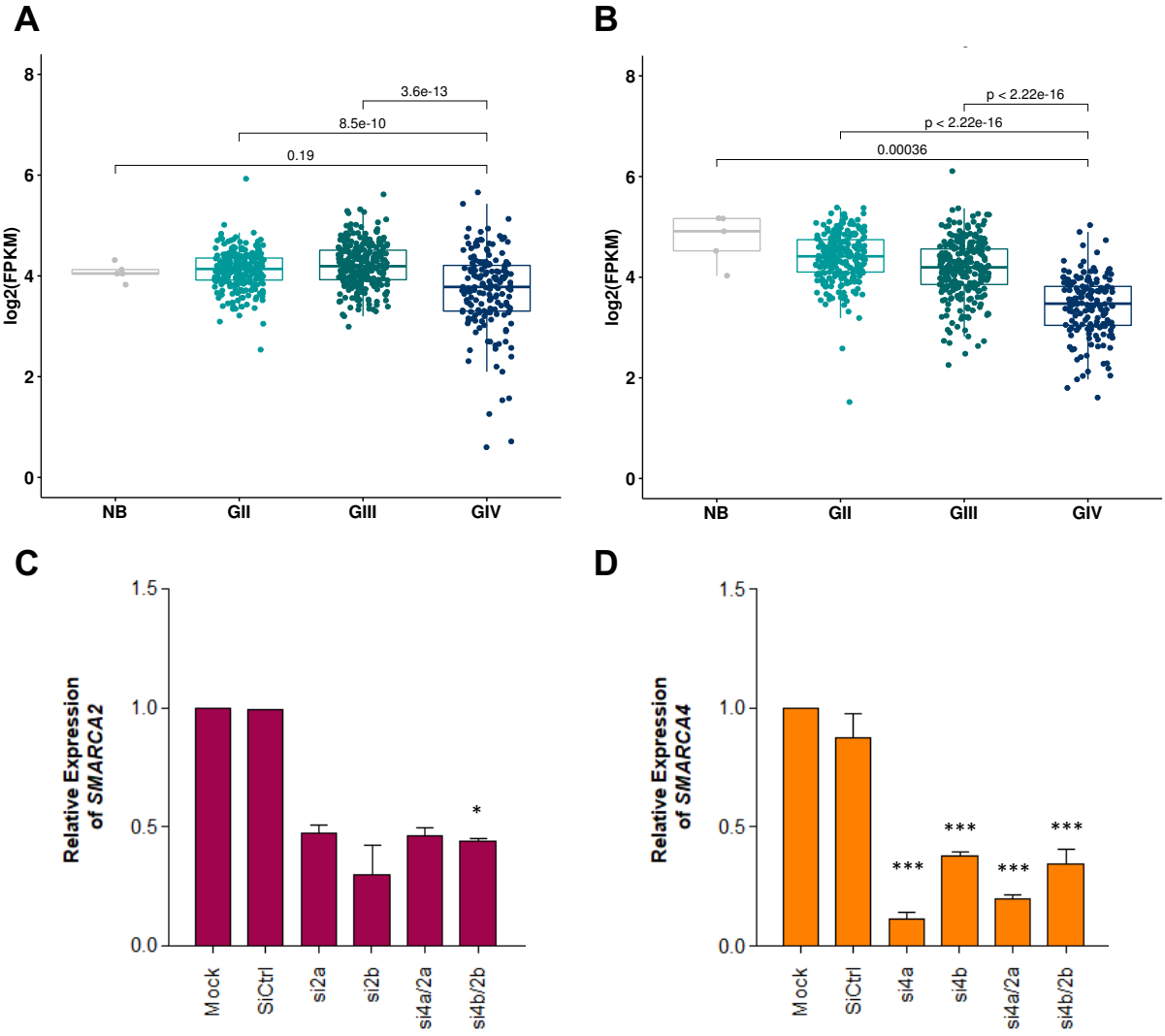


Figure 3.40. *SMARCA2* and *SMARCA4* expression in glioma samples and gene silencing in LN18 glioma cells. mRNA expression of *SMARCA2* (A) and *SMARCA4* (B) across glioma grades using TCGA data. Wilcoxon rank-sum statistical test was used to calculate differential expression (*p-value 0.05, **p-value 0.01 and ***p-

value 0.001). *SMARCA2* (C) and *SMARCA4* (D) relative expression in LN18 glioma cells post-transfection with siRNA. The ANOVA test was used to analyse *SMARCA2* and *SMARCA4* gene expression with no assumption of equal variance (Welch one-way test). Data are shown as means from 3 independent experiments.

3.4.2. Reduction of SMARCA2 and SMARCA4 protein levels in deficient cells

Western blot analysis revealed a significant down-regulation of SMARCA2 and SMARCA4 protein levels in both controls and SMARCA2, SMARCA4, or *SMARCA2/4* siRNA transfected cells (Fig. 3.41A). Densitometry of immunoblots from 3 experiments revealed statistically significant down-regulation of SMARCA2 and SMARCA4 proteins in cells transfected with specific siRNAs (Fig. 3.41B, Fig. 3.41C).

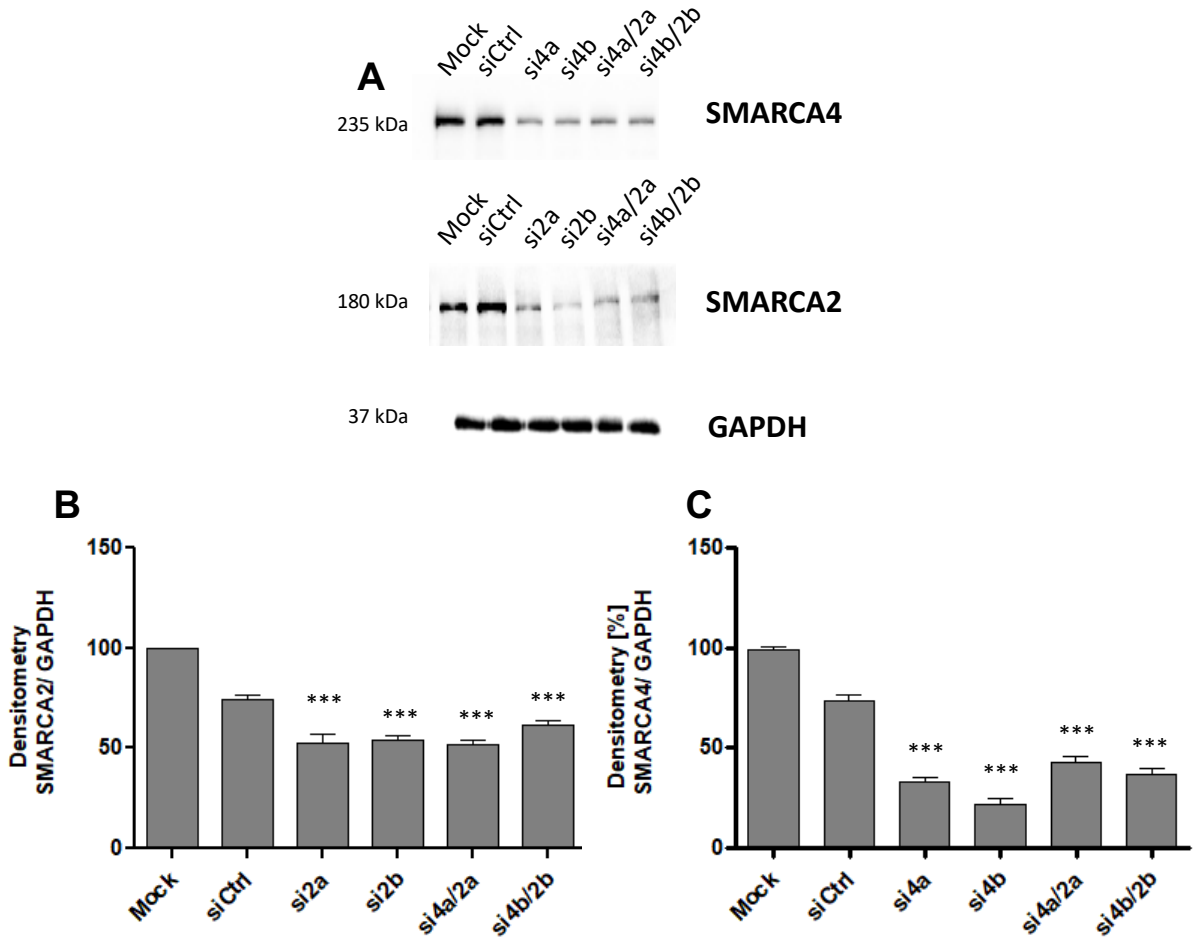


Figure 3.41. Western blots detecting SMARCA2 and SMARCA4 protein levels in LN18 glioma cells. (A) Western blotting was used to evaluate changes in protein levels in mock control (Mock), control (siCtrl), and *SMARCA2/SMARCA4* siRNA transfected cells. GAPDH detection was used to normalize the levels of SMARCA2 and SMARCA4. Densitometric analysis of SMARCA2 (B) and SMARCA4 (C) levels compared to GAPDH levels was performed. Non-parametric tests (Kruskal-Wallis) were used to analyse western blot densitometry, and significant differences were denoted as *p value < 0.05, **p value < 0.01, ***p value < 0.001.

3.4.3. Cell viability and proliferation of glioma cells after gene silencing

In order to assess the cell viability and proliferation after SMARCA2 and SMARCA4 silencing, we performed MTT metabolism and BrdU incorporation assays. The results indicated reduced cell viability and cellular proliferation of cells transfected with specific siRNAs compared to siCtrl transfected cells (Fig. 3.42).

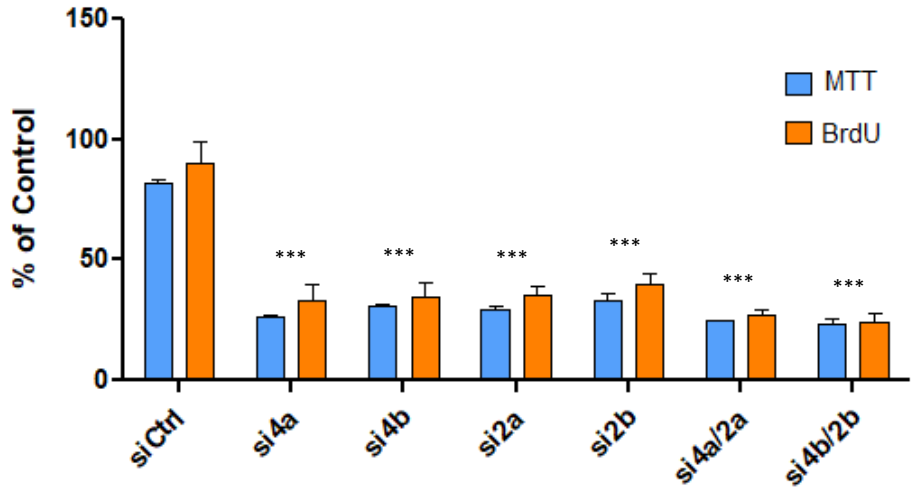


Figure 3.42. The effect of SMARCA2 and SMARCA4 knockdown on cell proliferation and viability. MTT metabolism and BrdU assays were used to evaluate cell viability and proliferative activity, respectively. Results are expressed as percentages of the control (n=3).

3.4.4. Chromatin accessibility in SMARCA2/4 depleted cells

To better understand the role of specific SMARCA proteins in glioma cells, we performed ATAC-seq to analyse global changes in chromatin accessibility in LN18 glioblastoma cells depleted of a specific SMARCA protein or both proteins. We examined signal distribution in promoter peaks, which revealed the higher accumulation in the vicinity of the TSS in cells depleted of a specific SMARCA protein or both proteins (Fig. 3.43A). In this approach, we examined the peak distribution around TSSs presenting signal enrichments as heatmaps in all tested groups (Fig. 3.43B, upper panel) and in group-specific open-chromatin regions (Fig. 3.43B, bottom panel). While many similarities in terms of accessible promoters were detected, we also found specific regions that are only open in controls or only open in cells depleted of *SMARCA2/4* (Fig. 3.43B, bottom panel). We found that the unique peaks identified in the controls are localised within distal open-chromatin regions, whereas the unique peaks identified in *SMARCA4* and *SMARCA4/SMARCA2* depleted cells are more concentrated around TSSs.

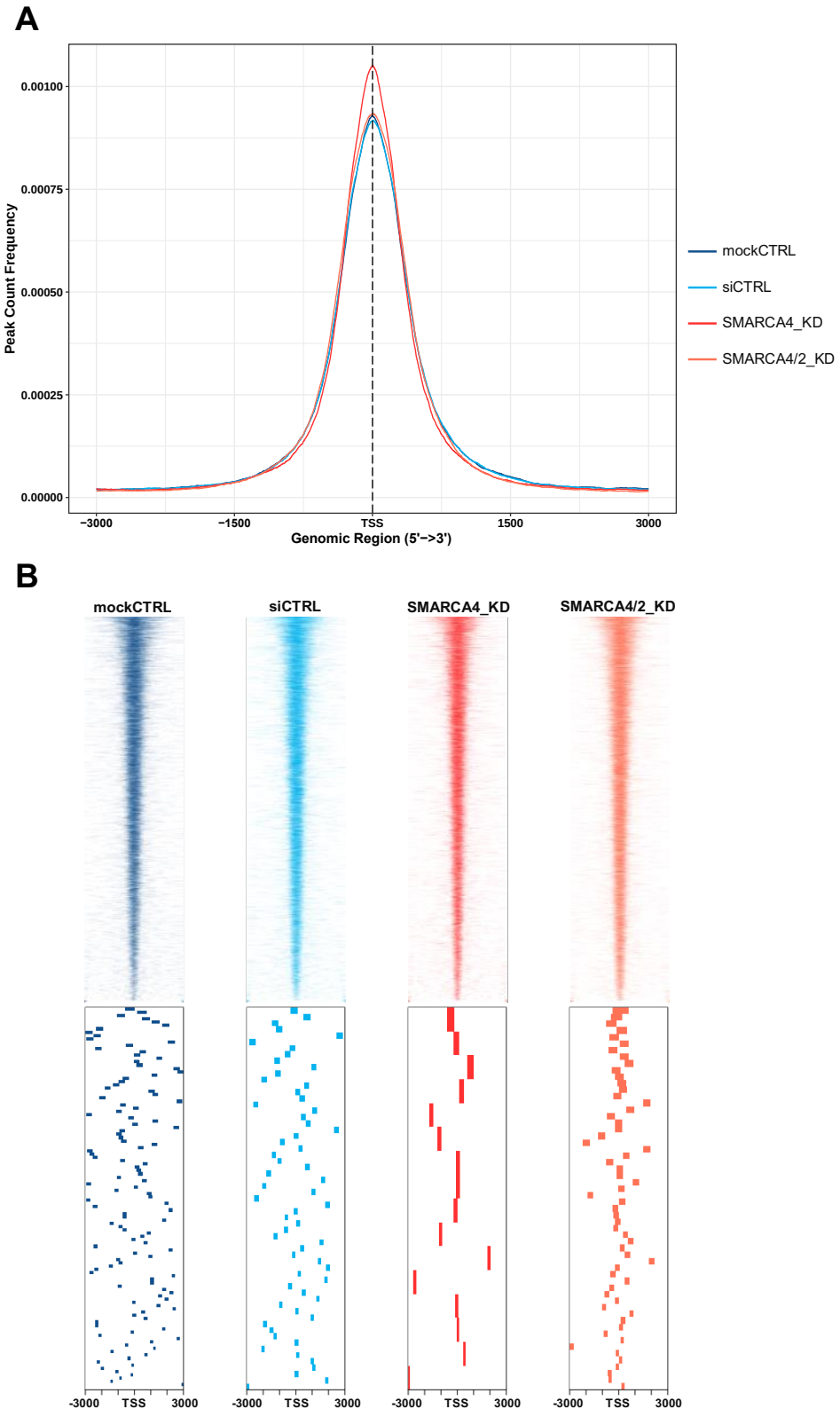


Figure 3.43. Landscape of chromatin accessibility in control and SMARCA2/4 depleted cells. (A) Transcription start site (TSS) enrichment profiles showing nucleosome-free regions enriched around TSSs and in the near flanking regions (TSS ± 3 kb) per group and replicate. (B) Profile heatmap of total (top panels) and unique (bottom panels) ATAC-seq peaks identified in a TSS ± 3 kb window and color-coded based on the previous plot. Unique peaks were identified after using *"bedtools substract -a"* command when comparing controls and SMARCA knockdown (KD) groups.

3.4.5. Identification of open chromatin regions and changes in the chromatin openness in SMARCA-depleted cells

In order to evaluate the effect of SMARCA silencing on the distribution of open chromatin spots throughout genomic regions, we compared the genomic location of ATAC-seq peaks identified in different experimental groups. The peaks present only in the controls or SMARCA4-depleted cells had a similar pattern of peak distribution, with the majority of peaks found in distal intergenic, intronic, and promoter regions. However, distribution of open chromatin peaks differed in the SMARCA4/2-depleted cells versus the other groups, with more abundant peaks in the gene promoters (Fig. 3.44A, 3.44B). This comparison revealed the increased chromatin accessibility in the promoters or proximal regulatory regions of some protein-coding genes (Fig. 3.44C). Moreover, regions with the decreased chromatin accessibility were also identified in SMARCA4/2 depleted cells (Fig. 3.44C).

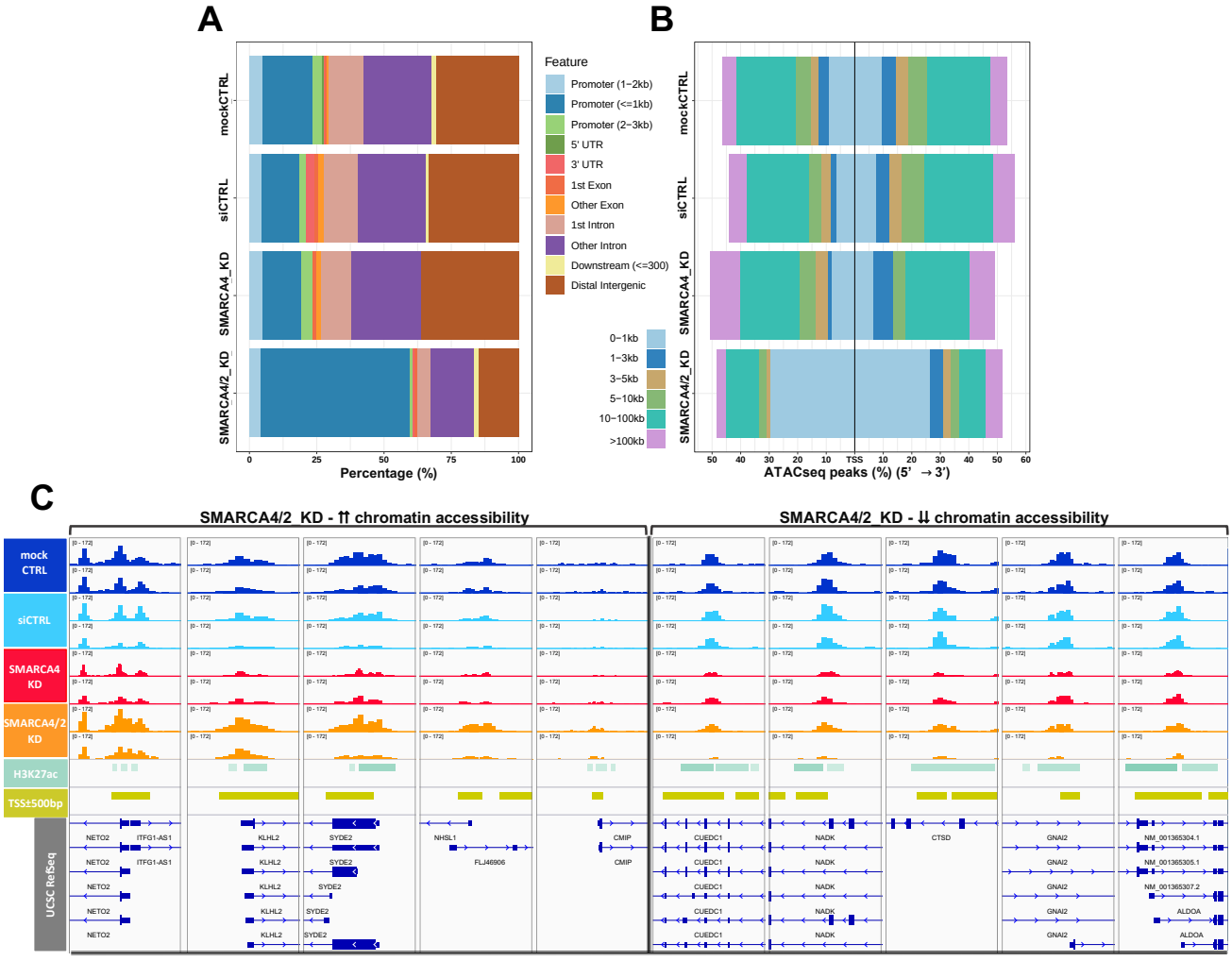


Figure 3.44. Peak distribution and differences in ATAC-seq peaks in SMARCA depleted cells. (A) A bar plot depicts the genomic annotation of the unique ATAC-seq peaks in relation to the closest annotated gene. (B) The location of these peaks in genomic regions in the proximity of TSSs. (C) Significant regional differences in chromatin accessibility in *SMARCA4/2* depleted cells (KD) (orange) compared to siCTRL transfected cells. The top 5 significant differences in open chromatin peaks are shown.

3.4.6. Open chromatin changes are associated with specific transcriptomic programs

We assessed which genes and biological pathways are affected by *SMARCA4/2* silencing in glioma LN18 cells. The identified changes in regulatory regions of certain genes indicated that genes involved in specific pathways related to cell projection organization, neuron projection development, and system development, among others, are associated with the increased chromatin accessibility (Fig. 3.45A). In *SMARCA4/2* depleted cells, regions controlling genes that belong to pathways related to protein localization, signal transduction, and cell communication show a decrease in chromatin accessibility (Fig. 3.45B).

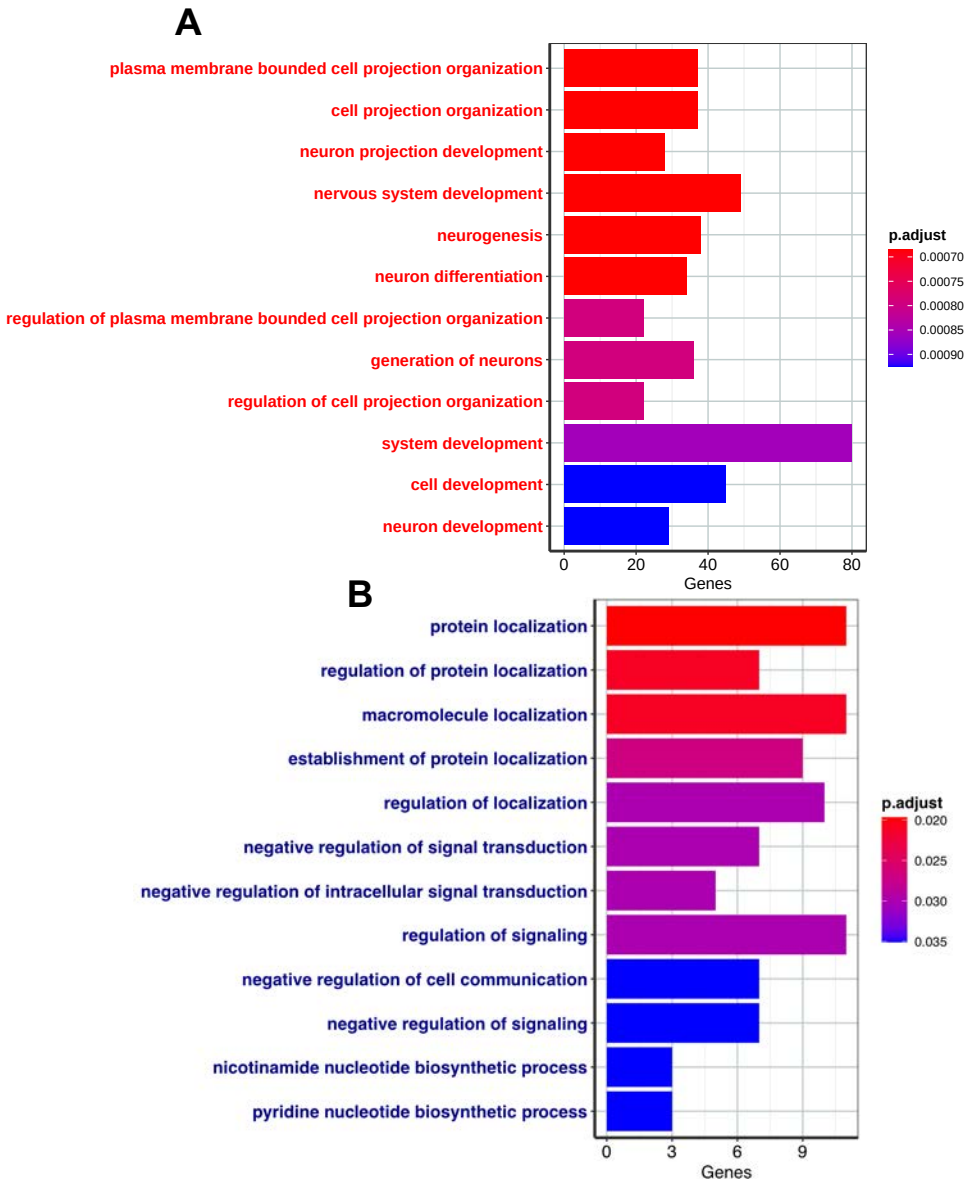


Figure 3.45. Biological implications of SMARCA4/2 silencing in glioma cells. Analysis of Gene Ontology (GO: Biological Processes) of increased (A) and decreased (B) chromatin-accessible regions between *SMARCA4/2* depleted cells and the controls. Only open chromatin regions within the promoters (TSS ± 500 bp) were chosen and annotated.

While analysing genes associated within the altered pathways, we noticed that some crucial genes involved in the TGFB pathway such as the receptor-activated *SMAD1* or the *BRPR1A* (encoding Bone morphogenetic protein receptor type-1A, which binds and activates SMAD transcriptional regulators) are associated with increased of chromatin accessibility in their promoters (Fig. 3.46).

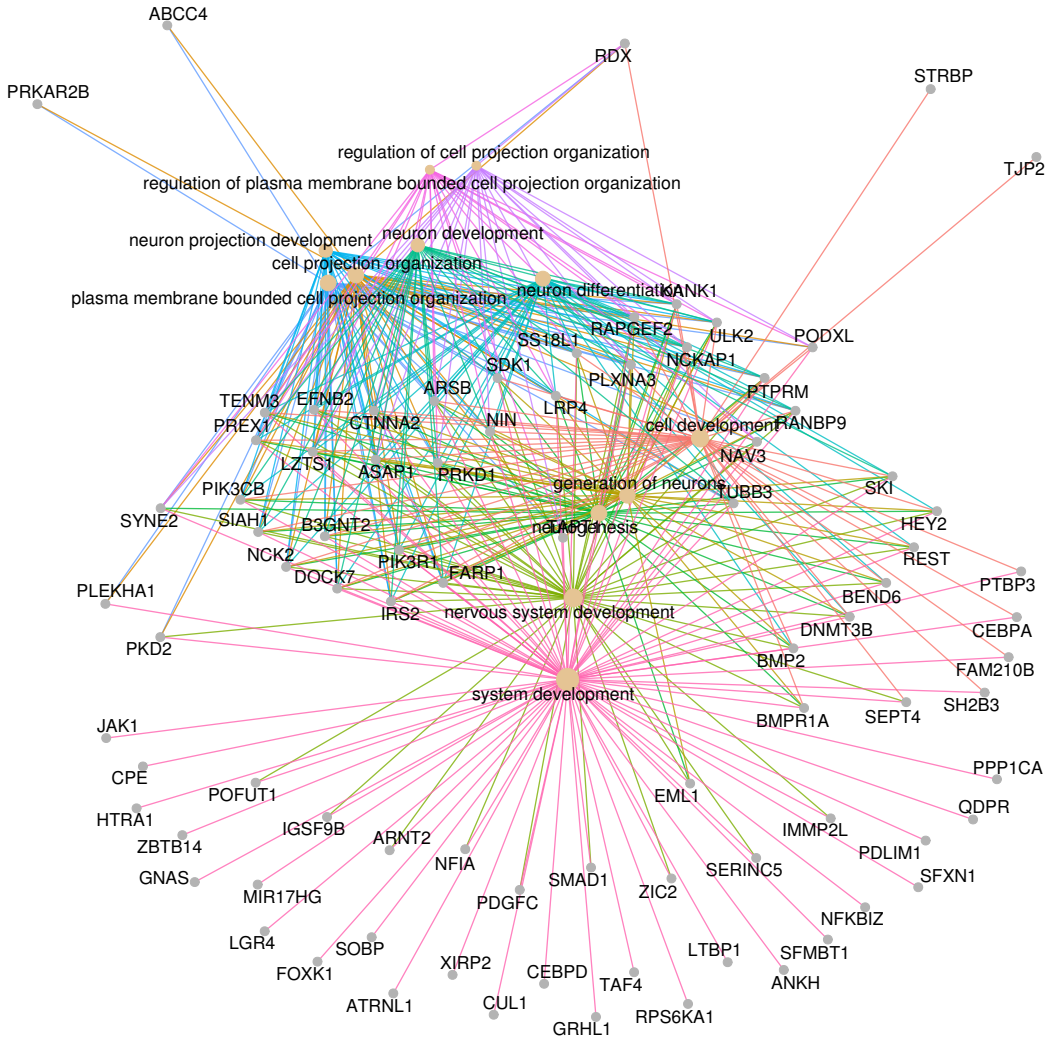


Figure 3.46. Gene Ontology analysis and genes involved in specific pathways in SMARCA4/2 depleted cells. (A) Gene ontology (GO: Biological Processes) analysis of increased and decreased chromatin-accessible regions between *SMARCA4/2* depleted cells and the control cells. Only open chromatin regions within the promoters (TSS ± 500 bp) are indicated.

3.4.7. Components of the TGF-β signalling pathway are up-regulated in SMARCA2/4-depleted glioma cells

To verify the main findings, we performed Western blotting analysis of proteins focusing on the TGF-β pathway-related factors. We determined the levels of SMAD1, SMAD3 and SMAD6 TFs, as well as JAK1, BMPR1A, and TGFBR2 in SMARCA-depleted cells and controls (Fig. 3.47A). The levels of TGFBR2 (transforming growth factors receptor 2) and SMAD3 increased in both SMARCA4- and SMARCA2/4-depleted cells (Fig. 3.47C). Because TGFBR2 is required to phosphorylate SMAD3 (269), this increase in expression has an effect on SMAD3 phosphorylation levels. In the SMARCA2/4-depleted cells, the SMAD3 level was increased. In addition, we observed a decrease in the level of SMAD6, which is a negative regulator of TGF-β signalling. Furthermore, we noticed an increase in BMPR1A levels, but the levels of this protein were high in the siCTRL cells. The levels of SMAD1 (Fig. 3.47B) and SMAD3 (Fig. 3.47D) were higher in SMARCA4- and SMARCA2/4-depleted cells than in controls, which confirms our predictions from the computational analysis of chromatin openness. The levels of JAK1 and TGFBR2 were likewise upregulated (Fig. 3.47A, Fig. 3.47C).

To further verify our findings on the activation of TGF-β signalling, we performed a gene reporter assay by transfecting cells with the vector carrying a *luciferase* gene under the control of SMAD-dependent promoter. SMAD proteins after activation of TGF-β signalling bind and induce expression of target genes via the CAGA motifs. Therefore, 48 h after transfection with specific or control siRNAs, the cells were additionally transfected with a (CAGA)-luciferase reporter vector and 24 h later the cells were lysed and luciferase activity was determined. We found the increased activity of the CAGA reporter in *SMARCA2/4*-depleted cells, indicating that the expression and activity of components of TGF-β signalling pathway was augmented in SMARCA4- and SMARCA2/4-depleted LN18 glioma cells (Fig. 3.47E).

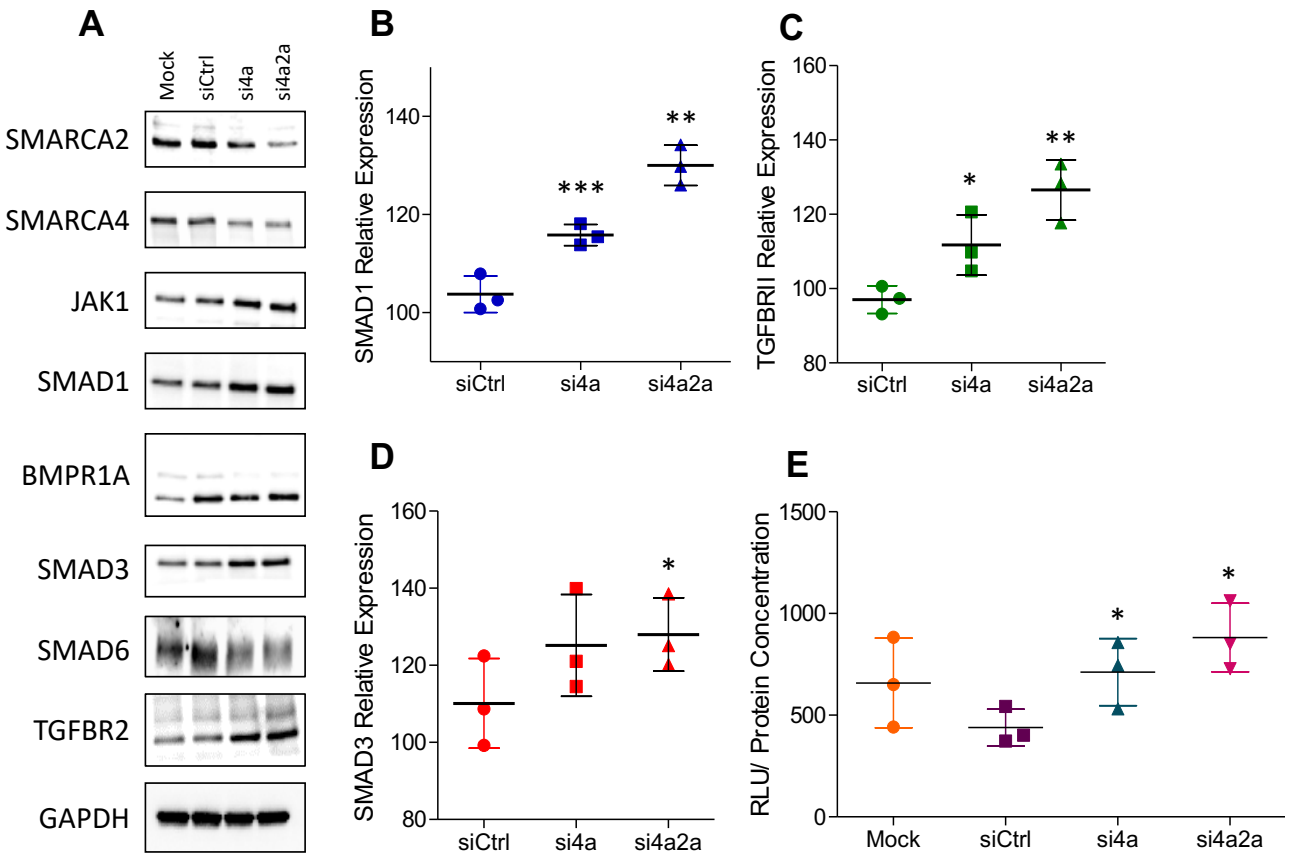


Figure 3.47. The expression of mediators of the TGF-β signalling pathway and its activation in SMARCA4- and SMARCA2/4-depleted cells. (A) Western blotting analysis of proteins involved the TGF-β signalling pathway in extracts from mock, siCtrl-transfected and *SMARCA4*- (si4a) or *SMARCA2/4*-depleted cells (si4a2a). Levels of SMAD1, TGFBRII and SMAD3 proteins in control and depleted cells are shown on representative Western blots. (B-D) Densitometry analysis of immunoblots from 3 independent experiments. (E) A (CAGA)-Luciferase reporter assay showing the increased activity of the TGFβ-dependent pathway in *SMARCA4*- (si4a) or *SMARCA2/4*-depleted cells (si4a2a). All values shown in the graph are final readouts of RLU (Relative Light Units) after normalization to a protein concentration; the results are presented as a fold change versus values determined in siCtrl cells. Statistical significance was estimated with One-way ANOVA (*p<0.05, **p<0.01, ***p<0.001).

Chapter IV

4. Discussion and perspectives

HGGs remain incurable brain tumours with the worst prognosis and the most debilitating symptoms. GBM has a current life expectancy of about 12 months after diagnosis. These brain tumours are also difficult to treat due to their molecular and genetic complexity. HGG is a multifaceted disease, in which several interconnected biological components - somatic mutations, deregulation of TFs, aberrant histone modification and DNA methylation patterns resulting in long-term chromatin remodeling changes occur at the same time, contributing to the disease aggressiveness. Understanding disease pathobiology and identification of progression markers and new promising targets may pave a way to designing new therapeutic combinations and better disease control.

The current thesis contributes to a better understanding of glioma pathobiology in several ways: 1) by characterisation of genomic and transcriptomic profiles in HGGs after recurrence; 2) by identification of mechanisms of transcription deregulation and chromatin remodeling in HGGs when compared to benign gliomas; 3) by defining the role of SMARCA2/4 remodelers in chromatin regulation.

4.1. Emerging evidence of clonal evolution and transcription deregulation during malignant glioma progression

Understanding the genomic and transcriptomic changes that underpin HGG progression and recurrence, as well as the associated changes in the tumour microenvironment, provides crucial insights into the evolution of malignant gliomas and may help in designing better treatments. We performed extensive genetic and transcriptomic analyses on 16 pairs of primary-recurrent HGGs. The vast majority of those were GBMs and the cohort represented mostly the cases of first and second resection. The results presented in this study show the presence of some well-known somatic mutations in genes such as *TP53, PTEN, PIK3R1, IDH1, ATRX,* and *PIK3CA. PIK3R1* appears to be a potential cancer driver gene. Somatic variants shared by primary and recurrent HGGs account for a small proportion of all detected variants. Most of the pathogenic variants detected at recurrence were not found in the primary tumours. This finding suggests a sub-clone substitution, with new sub-clones

harbouring different mutations as proposed by Brennan and colleagues (16) or by a neutral evolution and polyclonal re-emergence as suggested by other researchers (270,271).

Loss of chr10 and amplification of chr7 are common alterations in GBMs (106) and they were detected in the current CNA analysis. Profiles of CNAs were similar in primary and recurrent HGGs (Fig. 3.5). The computational methods employed in this study to detect CNAs (see section 2.1.4) used the data obtained with the 700 NGS gene panel which provides only an estimation of the true CNA landscape. Higher-resolution copy number analysis using whole genome sequencing or microarray-based CNA analysis, could provide a result with a greater accuracy regarding focal CNAs. Nevertheless, focusing on *EGFR* and *PTEN* CNAs, we found an inverse correlation between *EGFR* and *PTEN* focal aberrations (Fig. 3.6). This is consistent with the fact that *EGFR* amplification and deletion of *PTEN* produce similar consequences in deregulation of intracellular signalling that activate pro-proliferative and pro-survival pathways (272–275).

The GLASS consortium, which attempts to collect a comprehensive and longitudinal information about glioma progression, presented similar findings (19). The number of somatic variants specific to a primary or recurrent tumour was higher than those shared between the two disease stages for the majority of the tumours, which is consistent with our findings. Furthermore, their CNA analysis of samples collected during tumour progression indicates that chromosome 7 amplification and chromosome 10 deletion are found in both early and late stages of tumour evolution, particularly in IDH-WT specimens (18). Most of the genes established as oncogenic drivers in the GLASS study occurred in our analysis, except from *PIK3R1,* which we found as a new potential oncogenic driver gene.

We discovered a novel frame-shift insertion in the *ZNF384* gene that may affect protein stability. ZNF384 TF is involved in the pathobiology of acute lymphoblastic leukaemia through a fusion with *TET* family genes (276). ZNF384 (also known as NMP4, nuclear matrix protein 4) is a nucleocytoplasmic shuttling protein that suppresses bone anabolism by repressing genes involved in the osteogenic lineage commitment and mineralization (277). NMP4 may act as a transcriptional repressor of c-Myc and Gadd34 TFs, inhibiting ribosome biogenesis and global protein synthesis (277). In the survival analysis, we discovered that high *ZNF384* expression was inversely related to survival (Fig. 3.2), which is another argument supporting its role in gliomagenesis. The precise role of ZNF384 in HGGs requires more advanced studies using biochemical and gene editing techniques.

Transcriptomics analyses of HGGs at diagnosis and relapse show that genes involved in mRNA splicing, cell cycle, and DNA repair are down-regulated, while genes involved in interferon signalling are up-regulated at recurrence (Fig. 3.8, Fig. 3.9). The pathway enrichment analysis shows a strong evidence of deregulation of mRNA splicing, which is consistent with detection of differences in splicing isoforms between primary and recurrent tumours (Fig. 3.7). The abundance of splice variants and down-regulation of genes coding for the splicing machinery suggests the disruption of spliceosome functions and mRNA processing in recurrent HGGs. Tumour-specific splicing isoforms

may function differently than canonical isoforms found in healthy tissues. Because the splicing machinery components are druggable, this deregulation could be used to improve clinical outcomes (278,279).

Functional analysis of differentially expressed genes revealed the upregulation of IFN signalling-related genes in the recurrent HGGs that suggest a remodeling of TME upon tumour relapse. Interferons (IFNs) have anti-tumour (280) and anti-proliferative activity against glioblastoma cells (281,282). However, autocrine IFN signalling contributes to glioma cell immune evasion (283) and immunoediting (284). Many genes involved in the cell cycle regulation and cell cycle checkpoints may be down-regulated as a result of IFN up-regulation (285). The up-regulation of phosphatidylinositol (PI) signalling (Fig. 3.8, Fig. 3.9) and PI metabolism in recurrent HGGs suggests deregulation of intracellular lipid signalling and membrane trafficking, which could reflect the enhanced migration and invasion of tumour cells (286–288).

A growing evidence demonstrates that TME plays an important role in shaping tumour plasticity and aggressiveness. Our computational analysis of the immune microenvironment indicates the abundance of immune cells, primarily pro-tumorigenic (M2) macrophages, immature dendritic cells, and T helper cells, thus TME was found to vary remarkably during the progression of HGGs. The *in silico* predictions generated with the xCell method were verified with several computational methods and validated by immunocytochemistry on tumour sections. Dendritic cells are central regulators of the adaptive immune response, responsible for cancer recognition and eradication (289,290). Activated, mature DCs are the main antigen-presenting cells for initiating adaptive immune responses, whereas immature DCs are implicated in tolerance and induction of regulatory T cells (291–293). Antigen-presenting function of DCs is lost or inefficient in malignant gliomas (294,295). Furthermore, tumour-infiltrating DCs (TIDCs) may influence tumour progression, as evidenced in patients with relapsed prostate cancers who had higher densities of immature TIDCs than their primary tumours (294). Our findings revealed an accumulation of immature dendritic cells (CD209+ cells) in recurrent HGGs, which could impair anti-tumour responses and augment the immunosuppression. The presence of a strong pro-tumorigenic macrophage signature and an increase of phagocytic CD163+ cells (296,297) in recurrent HGGs, which is partially in line with GLASS's *in silico* cell enrichment results (18), indicates a tumour-supportive microenvironment and post-recurrence immunosuppression potentiation. These processes may be the primary impediments to effective glioma immunotherapy and must be overcome prior to the introduction of frontline immunotherapies in GBM.

4.2. Identification of HGG-specific TFs, their regulatory networks and translational potential as new therapy targets

Dysregulation of the transcriptional machinery may contribute to both cancer initiation and persistence. It can occur by aberrant activation, repression and/or temporal/spatial dyscoordination of crucial gene expression, as well as by protein structural changes caused by mutations. TFs bind to DNA mainly in nucleosome-depleted regions and their aberrant activity in cancer may be influenced by a number of direct and indirect processes, including gene amplification, point mutations and changes in expression. Moreover, DNA methylation and histone modifications may affect their action by opening or closing the chromatin.

Using chromatin accessibility data generated by ATAC-seq, we identified HGG-specific TFs binding sites in human LN18 and LN229 glioma cells, as well as in glioblastoma tumour samples. We found that many TFBS are present in the same open chromatin regions in both cultured cells and glioblastoma specimens, but some of them were found only in one sample, indicating that the patterns of chromatin openness can be variable (Fig. 3.14). We focused our efforts exclusively on TFBS predictions that were detected in both GBM cell lines.

Interestingly, the majority of open-chromatin regions identified by ATAC-seq are enriched in promoter regions, confirming that nucleosome-free fragments are enriched in the proximity of TSSs, while mono-nucleosome fragments are depleted at the proximity of TSSs but enriched at flanking regions (298). Some TFBS predictions, such as binding sites for AP2D, PAX5, and ZFX, were highly prevalent within open-chromatin areas (Fig. 3.15). AP2D, which codes for the TF activator protein 2 (AP-2) is involved in a variety of pathological and carcinogenic processes (299), while PAX5 TF promotes tumour malignancy and is linked to other well-documented and highly expressed TFs in astrocytomas, such as MYC, FOS, or JUN. The relation between the expression of human epidermal growth factor receptor 2 (HER2) and AP-2 was demonstrated (300). The zinc finger and X-linked transcription factor (ZFX) TF is associated with proliferation, tumorigenesis, and patient survival in a variety of human cancers (301). It maintains GSC self-renewal and tumorigenic potential by upregulating c-Myc expression (302). A close inspection of overexpressed genes in glioblastomas when compared with benign gliomas revealed specific TFBS for in malignant tumours. In other words, those TFBS were only present in the promoters of genes highly expressed in GBMs (Fig. 3.16). At this point, we hypothesised that this observation may indicate the contribution of a specific TF to glioma malignancy. A transcriptional activity of a TF is regulated at various levels. While the presence of TFBS in cis-regulatory regions indicates a potential TF-gene regulation, a direct measurement of TF activity is a more accurate estimate to understand transcriptional responses (303,304).

One of most interesting findings in that part of the analysis is that we identified WHO grade IV-specific TFBS (Fig. 3.16). The most striking finding concerns c-Jun TF, a well-documented proto-oncogene which is involved in proliferation, angiogenesis, migration and apoptosis in several cancers (208,305). In the pan-cancer analysis we found that JUN gene expression is higher in normal tissues when compared with their corresponding malignant lesions in several cancer types: BLCA, BRCA, SKCM, CESC, OV, LUSC, UCEC, LUAD, and UCS (Fig. 3.21). This suggests that deactivating the tumour-suppressive c-Jun protein may cause normal cells to transform (306). In contrast, in some cancers c-Jun may act as a pro-tumorigenic factor, which is consistent with previous findings (306–308). We found significantly higher expression of the *JUN* gene in thymomas and GBMs (Fig. 3.21), therefore we hypothesized that this TF might be a key player in malignant processes in these tumours. Moreover, the analysis of *c-JUN* expression across gliomas of different WHO grades, revealed a positive association with increasing tumour malignancy, indicating its importance in HGGs (Fig. 3.21).

Based on ATAC-seq data c-Jun binding motifs were detected in the promoter regions of 16 genes overexpressed in GBMs (Fig. 3.22), which makes this specific TFBS the most abundant TFBS across GBM-specific genes. Many of these genes, such as *VIM* (309)*, FOSL2* (310)*, PTN* (311)*, GPR3* (312)*, SIAH2* (313,314)*, UPP1* (315)*, S100A2* (316) are associated with cell migration and mesenchymal transition in several cancers. Intriguingly, when checking the influence of expression of these genes on patient's survival, all of them were predictive in LGGs but not in GBMs (Fig. 3.30, Fig. 3:31). Increased gene expression was correlated with poor clinical outcome in LGG. One explanation for such a lack of association in GBMs could be that the expression of these genes in GBM patients is very high, making the distinction between low-expressing and high-expressing patients within this grade more difficult, whereas in LGG the distinction is clearer (Fig. 3.30, Fig. 3:31).

Moreover, the observation that c-Jun target genes show a greater prognostic value in WHO GII than in GIV tumours just reinforces the notion that c-Jun is important in transformation of lower-grade, more benign tumours into highly aggressive glioblastomas. Furthermore, the putative TFBS for several Jun-related factors (Jun-B, JunD) and fos-related factors (c-Fos, FRA1, FRA2, and FosB) overlap with c-Jun binding sites (Fig. 3.14). This suggest that c-Jun interacts with other oncogenic TFs at the specific promoters to regulate gene expression by forming the AP-1 complex (317). In fact, c-Jun needs to form heterodimers with different bZIP proteins to regulate gene expression (318,319). c-Jun/c-Jun homodimers occur in cells and they are more stable and transcriptionally active than c-Jun/c-Fos heterodimers (320). Notably, a choice of a dimerizing partner influences not only the DNA recognition properties but also the regulatory function of a given bZIP (321). Post-translational modifications regulate the c-Jun/c-Fos heterodimer formation via mitogen-activated protein kinase (MAPK) cascades (322), causing their activation and subsequent transcriptional regulation of targeted genes via AP-1 binding sites (323).

Our TFBS analysis was also extended to distal regulatory regions. In glioma enhancers, we identified the enrichment of c-Jun binding motifs, together with motifs for other bZIP TFs (Fig. 3.33B,

Table 2). This part of the analysis suggested that besides gene promoters, c-Jun and its protein partners may be important for gene regulation at the distal regulatory elements. Thus, changes in AP-1 complex components expression or activity may have a significant impact on transcriptional networks in GBM. Finally, the comparison of the H3K27ac signal between GBMs and DAs enhancers demonstrated only a few regions with significantly higher signal towards GBMs (Suppl. Table 2). This shows that most of the studied regions have activating histone marks both in DAs and GBMs and suggests common mechanisms for gene regulation at the level of enhancers in gliomas.

We noticed that the expression of genes coding GBM specific TFs is high in HGGs and low in LGGs (Fig. 3.25, Fig. 3.26), as is exemplified by the expression of *HOX genes, JUN,* and *TWIST*, among others. These genes have been previously reported as more abundantly expressed in HGGs (324,325). These observations, along with a positive correlation between *c-Jun* mRNA or protein activity and target gene expression, support the conclusion that c-Jun most likely regulates the predicted targets in GBMs (Fig. 3.28, Fig. 3.29).

Among computationally unravelled putative c-Jun target genes in GBMs, a case of an intermediate filament Vimentin piqued our interest as a critical protein linked to increased metastatic potential and cell migration in various cancers (265,326,327). *VIMENTIN* expression is a poor prognostic factor in glioblastomas (328). The intrinsic mechanisms governing *VIMENTIN* overexpression in HGGs were unknown. Therefore, we biochemically validated c-Jun binding to the *VIMENTIN* promoter in glioblastoma cells using EMSA and supershift assay. We demonstrated that 1) c-Jun binds to the *VIMENTIN* promoter, as predicted by our computational analyses; 2) more DNA-protein complexes (c-Jun-*VIMENTIN* promoter) were detected in GBM cells than in the LGG patient-derived cells (Fig. 3.37). The quantification of EMSA results suggests that c-Jun binding to the *VIMENTIN* promoter is more stable and potent, or that the amount of activated c-Jun in GBM cells is higher.

DNA methylation regulates gene expression by affecting TF binding affinity to the DNA (128,129). To study a potential impact of DNA methylation, we examined methylation levels in the promoters (+2 kb/-500 bp from TSS) of c-Jun targets and in the promoter of *JUN* in gliomas of various grades. We noticed differences in DNA methylation in some genes with c-Jun motifs (*S100A10, S100A2, IFRD1, RUN1, RAB36, UPP1, SLFN12* and *VIM*) (Fig. 3.34, Fig. 3.35). The majority of these differences was detected in flanking regions rather than in c-Jun binding sites when C-rich regions were examined (Fig. 3.36). According to recent research, the JUN binding site motif, along with FOSL2 or CREB1, is highly affected by DNA methylation disorder and is associated with increased cell stress, suggesting the importance of these TFs in epigenetic intratumoural heterogeneity (329).

We have experimentally tested whether differences in DNA methylation in these regions in cultured glioma cells and in normal astrocytes might have a biological significance, and if c-Jun binding to the DNA is affected. These results are currently collected.

4.3. Knockdown of SMARCA chromatin remodelers affects crucial signalling proteins in gliomas

Transcription initiation is regulated by the accessibility of the transcriptional machinery to chromatin and DNA. To perform histone sliding, ejection, or integration of histone variations, cells use a variety of ATP-dependent nucleosome-remodeling complexes (172). These complexes have an ATPase activity and may directly change the three-dimensional structure of chromatin (173) by altering histone-DNA contacts (174) and controlling chromatin dynamics (175). The chromatin remodeling carried out by the SWI/SNF complex typically occurs in cis-regulatory elements and in super-enhancers (179–181). Deleterious mutations in *SMARCA2/4* genes occur in cancers, including GBMs (178,330) and may affect functions of the SWI/SNF complex.

The level of genes encoding SMARCA2 and SMARCA4 are significantly down-regulated with increasing glioma malignancy grades in the TCGA dataset (Fig. 3.40A, Fig. 3.40B). To evaluate the impact of SMARCA2/4 on gene expression in glioma cells, we efficiently silenced *SMARCA2* and/or *SMARCA4* in human LN18 GBM cells (Fig. 3.40D). Functional assays such as MTT metabolism and cell proliferation (BrdU) assays were performed on SMARCA2/4 depleted cells. The MTT metabolism analysis showed that LN18 cells transfected with siSMARCA2 and siSMARCA4 have lower viability and proliferation than cells transfected with siCtrl (Fig. 3.42). This observation was unexpected as we observed a decrease of SMARCA expression upon GBMs in TCGA dataset. These findings suggest that the chromatin remodelers SMARCA2 and SMARCA4 might have a tumour suppressive function in GBM cells.

To better comprehend the role of SMARCA proteins in chromatin reorganization, we performed chromatin accessibility assays (ATAC-seq) on the LN18 cells and studied the ATAC-seq signal enrichment in gene promoter regions, which marked nucleosome-free regions (Fig. 3.43A). We found that the majority of changes in chromatin openness occurred in SMARCA2+SMARCA4 depleted cells rather than the in single knock-out group. Therefore, the absence of just one SMARCA protein may not be sufficient to induce gross impact on chromatin openness (Fig. 3.43, Fig. 3.44). The detailed analysis of ATAC-seq signals showed the increased chromatin accessibility in the promoters or close regions of some protein-coding genes (Fig. 3.44C), although certain regions with the decreased chromatin accessibility occurred in SMARCA4/2-depleted cells (Fig. 3.44C). This finding suggests that the absence of both SMARCA proteins causes bidirectional changes in a chromatin structure: on the one hand, some regions closer to TSS regions become more open/accessible, while on the other hand, more distal regions become less accessible (Fig. 3.44, Fig. 3.45). The functional analysis of genes assigned to open chromatin regions in the SMARCA4/2–depleted cells revealed an enrichment of pathways related to cell projection organization, neuron projection development and system development, among others (Fig. 3.45A). In contrast, pathways related to protein localization, signal

transduction and cell communication were associated with a decrease in the chromatin accessibility in SMARCA4/2-depleted cells (Fig. 3.45B). These differences show that changes in the chromatin accessibility may influence expression of specific genes, thereby affecting various biological pathways.

We discovered that some critical genes involved in the TGF-β pathway, such as receptor-activated *SMAD1* or Bone morphogenetic protein receptor type-1A (*BMPR1A*), are localized in the regions with increased chromatin openness (Fig. 3.46). BMPR1A is a crucial protein receptor which activates SMAD transcriptional regulators (331). After BMP receptors phosphorylate and activate SMAD1, the protein mediates transcriptional regulation of DNA damage and oncogenesis pathways. (332). Our findings suggest that expression of these two genes might be dependent on chromatin remodeling carried out by SMARCA2 and SMARCA4 proteins. We validated changes in TGF-β signalling related factors by Western blotting and (GAGA)-dependent luciferase reporter assay to confirm computational findings. Levels of SMAD1, SMAD3, and TGFBR2 were increased in *SMARCA4/2*-depleted cells (Fig. 3.47). Bioactive TGF-β binds to TGFBR2 on the cell surface triggering formation of a receptor complex with TGFBR1, which is required for effector proteins SMAD2 and SMAD3 phosphorylation and their nuclear translocation (333,334). Therefore, increased expression of TGFBR2 and SMAD3 may lead to augmented activation of SMAD-dependent transcription (335,336). We also found a decrease in expression of SMAD6 (one of the negative regulators of SMAD signalling). Thus, knockdown of SMARCA2 and SMARCA4 affected expression of several TGF-β signalling pathway components, leading to upregulation of this signalling pathway in LN18 glioma cells. The results underline the role of the SWI/SNF remodeling complex in controlling expression of *TGFR2* and *SMAD1/SMAD3* (Fig. 3.47E). As a result, deleterious mutations reported in *SMARCA2* or *SMARCA4* (178,330) may eventually affect the functionality of TGF-β signalling in glioma cells.

4.4. Summary and conclusions

In this study we characterised genomic and transcriptomic profiles of HGGs after recurrence, and identified some mechanisms of transcriptomic and epigenetic deregulation in HGGs all of which provide a valid information on the pathobiology of HGGs. Intersection of various high-throughput data allowed us to precisely pinpoint the pathogenic factors driving the disease, decipher potential mechanisms behind treatment failure and indicate new potential therapy targets. The presented results pointed to several gene regulatory networks that are overactivated in malignant glioma cells and HGGs. Chromatin accessibility analysis of *SMARCA* depleted cells revealed a bidirectional change in the chromatin structure.

The specific findings can be listed as follows:

1. **We identified distinct genomic alterations in HGGs after recurrence that suggest a sub-clone substitution during tumour progression. A novel frame-shift insertion in the *ZNF384* gene, which could affect the protein stability, has been identified. The presence of focal CNAs in the *EGFR* and *PTEN* genes was found to be inversely correlated in both primary and recurrent tumours, suggesting a co-dependency.**

2. **Several genes coding for components of the spliceosome machinery were down-regulated upon recurrence. Moreover, transcriptomic profiles after recurrence revealed changes in the tumour microenvironment such as the enrichment of M2 macrophages and immature dendritic cells, which are indicative of immunosuppression.**

3. **c-Jun binding motifs were found in crucial regulatory sites of GBM-specific genes revealing c-Jun as a master regulator in GBMs. We experimentally validated c-Jun binding to the *VIMENTIN* gene promoter.**

4. **Knockdown of SMARCA2 and SMARCA4 chromatin remodelers affected chromatin accessibility in human glioblastoma cells, particularly in the promoters of TGF-β signalling related genes. Levels of TGFR2, SMAD1 and SMAD3 proteins increased as a consequence of SMARCA2/4 silencing.**

A few mechanisms still need to be experimentally verified. For example, the EMSA experiment on c-Jun binding is performed with the methylated and unmethylated probes from the *UPP1* promoter to estimate if DNA methylation affects the c-Jun binding to the target gene. The experiments are in progress. Furthermore, a transcriptional activity of AP-1 complex can be measured with a TRE (TPA responsive element) reporter luciferase assay. Moreover, we envision that further studies of glioma cell migration, invasion or matrix reorganization in SMARCA2/4 depleted glioma cells will provide more information.

References

1. Adamson DC, Rasheed BAK, McLendon RE, Bigner DD. Central nervous system. Cancer Biomarkers. 2011;
2. Germano I, Swiss V, Casaccia P. Primary brain tumors, neural stem cell, and brain tumor cancer cells: Where is the link? Neuropharmacology. 2010.
3. Leonard A, Wolff JE. Etoposide improves survival in high-grade glioma: A meta-analysis. Anticancer Res. 2013;
4. Fisher JL, Schwartzbaum JA, Wrensch M, Wiemels JL. Epidemiology of Brain Tumors. Neurol. Clin. 2007.
5. Claus EB, Walsh KM, Wiencke JK, Molinaro AM, Wiemels JL, Schildkraut JM, et al. Survival and low-grade glioma: The emergence of genetic information. Neurosurg Focus. 2015;
6. Ostrom QT, Patil N, Cioffi G, Waite K, Kruchko C, Barnholtz-Sloan JS. CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2013-2017. Neuro Oncol. 2020;
7. Grossman SA, Ye X, Piantadosi S, Desideri S, Nabors LB, Rosenfeld M, et al. Survival of patients with newly diagnosed glioblastoma treated with radiation and temozolomide in research studies in the United States. Clin Cancer Res. 2010;
8. Wang D, Wang C, Wang L, Chen Y. A comprehensive review in improving delivery of small-molecule chemotherapeutic agents overcoming the blood-brain/brain tumor barriers for glioblastoma treatment. Drug Deliv. 2019;
9. Tran B, Rosenthal MA. Survival comparison between glioblastoma multiforme and other incurable cancers. J. Clin. Neurosci. 2010.
10. Strobel H, Baisch T, Fitzel R, Schilberg K, Siegelin MD, Karpel-Massler G, et al. Temozolomide and Other Alkylating Agents in Glioblastoma Therapy. Biomedicines. 2019;
11. Jiang H, Yu K, Li M, Cui Y, Ren X, Yang C, et al. Classification of Progression Patterns in Glioblastoma: Analysis of Predictive Factors and Clinical Implications. Front Oncol. 2020;
12. Eijgelaar RS, Bruynzeel AME, Lagerwaard FJ, Müller DMJ, Teunissen FR, Barkhof F, et al. Earliest radiological progression in glioblastoma by multidisciplinary consensus review. J Neurooncol. 2018;
13. Poon MTC, Sudlow CLM, Figueroa JD, Brennan PM. Longer-term (≥ 2 years) survival in patients with glioblastoma in population-based studies pre- and post-2005: a systematic review and meta-analysis. Sci Rep. 2020;
14. Perry A, Wesseling P. Histologic classification of gliomas. Handb Clin Neurol. 2016.
15. McLendon R, Friedman A, Bigner D, Van Meir EG, Brat DJ, Mastrogianakis GM, et al. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;
16. Brennan CW, Verhaak RGW, McKenna A, Campos B, Noushmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. Cell. 2013;
17. Brat DJ et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med. 2015;
18. Barthel FP, Johnson KC, Varn FS, Moskalik AD, Tanner G, Kocakavuk E, et al. Longitudinal molecular trajectories of diffuse glioma in adults. Nature. 2019;
19. Aldape K, Amin SB, Ashley DM, Barnholtz-Sloan JS, Bates AJ, Beroukhim R, et al. Glioma through the looking GLASS: Molecular evolution of diffuse gliomas and the Glioma Longitudinal Analysis Consortium. Neuro Oncol. 2018;
20. Wang Q, Hu B, Hu X, Kim H, Squatrito M, Scarpace L, et al. Tumor Evolution of Glioma-

Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. Cancer Cell. 2017;
21. Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell. 2010;
22. Neftel C, Laffy J, Filbin MG, Hara T, Shore ME, Rahme GJ, et al. An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. Cell. 2019;
23. Perng P, Lim M. Immunosuppressive mechanisms of malignant gliomas: Parallels at non-CNS sites. Front. Oncol. 2015.
24. Ellert-Miklaszewska A, Dabrowski M, Lipko M, Sliwa M, Maleszewska M, Kaminska B. Molecular definition of the pro-tumorigenic phenotype of glioma-activated microglia. Glia. 2013;
25. Hussain SF, Yang D, Suki D, Aldape K, Grimm E, Heimberger AB. The role of human glioma-infiltrating microglia/macrophages in mediating antitumor immune responses. Neuro Oncol. 2006;
26. Maas SLN, Abels ER, Van De Haar LL, Zhang X, Morsett L, Sil S, et al. Glioblastoma hijacks microglial gene expression to support tumor growth. J Neuroinflammation. 2020;
27. Kaminska B, Mota M, Pizzi M. Signal transduction and epigenetic mechanisms in the control of microglia activation during neuroinflammation. Biochim Biophys Acta - Mol Basis Dis. 2016;
28. Gabrusiewicz K, Ellert-Miklaszewska A, Lipko M, Sielska M, Frankowska M, Kaminska B. Characteristics of the alternative phenotype of microglia/macrophages and its modulation in experimental gliomas. PLoS One. 2011;
29. Ciechomska IA, Przanowski P, Jackl J, Wojtas B, Kaminska B. BIX01294, an inhibitor of histone methyltransferase, induces autophagy-dependent differentiation of glioma stem-like cells. Sci Rep. 2016;
30. Collins VP. Brain tumours: Classification and genes. Neurol. Pract. 2004.
31. Bai J, Varghese J, Jain R. Adult Glioma WHO Classification Update, Genomics, and Imaging. Top Magn Reson Imaging. 2020;
32. Chen R, Smith-Cohn M, Cohen AL, Colman H. Glioma Subclassifications and Their Clinical Significance. Neurotherapeutics. 2017.
33. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO classification of tumors of the central nervous system: A summary. Neuro Oncol. 2021;
34. Marquet G, Dameron O, Saikali S, Mosser J, Burgun A. Grading glioma tumors using OWL-DL and NCI Thesaurus. AMIA Annu Symp Proc. 2007;
35. Collins VP, Jones DTW, Giannini C. Pilocytic astrocytoma: pathology, molecular mechanisms and markers. Acta Neuropathol. 2015.
36. Marko NF, Weil RJ. The molecular biology of WHO grade II gliomas. Neurosurg Focus. 2013;
37. Mair MJ, Geurts M, van den Bent MJ, Berghoff AS. A basic review on systemic treatment options in WHO grade II-III gliomas. Cancer Treat. Rev. 2021.
38. Ohgaki H, Kleihues P. The definition of primary and secondary glioblastoma. Clin. Cancer Res. 2013.
39. Ahir BK, Engelhard HH, Lakka SS. Tumor Development and Angiogenesis in Adult Brain Tumor: Glioblastoma. Mol. Neurobiol. 2020.
40. Mei X, Chen YS, Chen FR, Xi SY, Chen ZP. Glioblastoma stem cell differentiation into endothelial cells evidenced through live-cell imaging. Neuro Oncol. 2017;
41. Taal W, Bromberg JEC, van den Bent MJ. Chemotherapy in glioma. CNS Oncol. 2015.
42. Eskilsson E, Røsland G V., Solecki G, Wang Q, Harter PN, Graziani G, et al. EGFR heterogeneity and implications for therapeutic intervention in glioblastoma. Neuro Oncol.

2018;
43. Nigro JM, Misra A, Zhang L, Smirnov I, Colman H, Griffin C, et al. Integrated array-comparative genomic hybridization and expression array profiles identify clinically relevant molecular subtypes of glioblastoma. Cancer Res. 2005;
44. Sturm D, Bender S, Jones DTW, Lichter P, Grill J, Becher O, et al. Paediatric and adult glioblastoma: Multiform (epi)genomic culprits emerge. Nat. Rev. Cancer. 2014.
45. Stieber D, Golebiewska A, Evers L, Lenkiewicz E, Brons NHC, Nicot N, et al. Glioblastomas are composed of genetically divergent clones with distinct tumourigenic potential and variable stem cell-associated phenotypes. Acta Neuropathol. 2014;
46. Neilsen BK, Sleightholm R, McComb R, Ramkissoon SH, Ross JS, Corona RJ, et al. Comprehensive genetic alteration profiling in primary and recurrent glioblastoma. J Neurooncol. 2019;
47. Kim H, Zheng S, Amini SS, Virk SM, Mikkelsen T, Brat DJ, et al. Whole-genome and multisector exome sequencing of primary and post-treatment glioblastoma reveals patterns of tumor evolution. Genome Res. 2015;
48. Klopp AH, Spaeth EL, Dembinski JL, Woodward WA, Munshi A, Meyn RE, et al. Tumor irradiation increases the recruitment of circulating mesenchymal stem cells into the tumor microenvironment. Cancer Res. 2007;
49. Jin X, Kim LJY, Wu Q, Wallace LC, Prager BC, Sanvoranart T, et al. Targeting glioma stem cells through combined BMI1 and EZH2 inhibition. Nat Med. 2017;
50. Yekula A, Yekula A, Muralidharan K, Kang K, Carter BS, Balaj L. Extracellular Vesicles in Glioblastoma Tumor Microenvironment. Front. Immunol. 2020.
51. Simon T, Jackson E, Giamas G. Breaking through the glioblastoma micro-environment via extracellular vesicles. Oncogene. 2020.
52. Antunes ARP, Scheyltjens I, Duerinck J, Neyns B, Movahedi K, Van Ginderachter JA. Understanding the glioblastoma immune microenvironment as basis for the development of new immunotherapeutic strategies. Elife. 2020;
53. Gargini R, Segura-Collar B, Sánchez-Gómez P. Cellular plasticity and tumor microenvironment in gliomas: The struggle to hit a moving target. Cancers (Basel). 2020.
54. DeCordova S, Shastri A, Tsolaki AG, Yasmin H, Klein L, Singh SK, et al. Molecular Heterogeneity and Immunosuppressive Microenvironment in Glioblastoma. Front. Immunol. 2020.
55. Glass R, Synowitz M. CNS macrophages and peripheral myeloid cells in brain tumours. Acta Neuropathol. 2014.
56. Russo C, Lisi L, Tentori L, Navarra P, Graziani G, Combs C. Exploiting Microglial Functions for the Treatment of Glioblastoma. Curr Cancer Drug Targets. 2016;
57. Hambardzumyan D, Gutmann DH, Kettenmann H. The role of microglia and macrophages in glioma maintenance and progression. Nat. Neurosci. 2015.
58. Roesch S, Rapp C, Dettling S, Herold-Mende C. When immune cells turn bad—tumor-associated microglia/macrophages in glioma. Int. J. Mol. Sci. 2018.
59. Platten M, Kretz A, Naumann U, Aulwurm S, Egashira K, Isenmann S, et al. Monocyte chemoattractant protein-1 increases microglial infiltration and aggressiveness of gliomas. Ann Neurol. 2003;
60. Chen X, Wang Y, Nelson D, Tian S, Mulvey E, Patel B, et al. CCL2/CCR2 regulates the tumor microenvironment in HER-2/neu-driven mammary carcinomas in mice. PLoS One. 2016;
61. Hu F, Dzaye ODA, Hahn A, Yu Y, Scavetta RJ, Dittmar G, et al. Glioma-derived versican promotes tumor expansion via glioma-associated microglial/macrophages Toll-like receptor 2 signaling. Neuro Oncol. 2015;
62. Vinnakota K, Hu F, Ku MC, Georgieva PB, Szulzewsky F, Pohlmann A, et al. Toll-like

receptor 2 mediates microglia/brain macrophage MT1-MMP expression and glioma expansion. Neuro Oncol. 2013;
63. Hu F, Ku MC, Markovic D, Dzaye OD, Lehnardt S, Synowitz M, et al. Glioma-associated microglial MMP9 expression is upregulated by TLR2 signaling and sensitive to minocycline. Int J Cancer. 2014;
64. Pickup MW, Mouw JK, Weaver VM. The extracellular matrix modulates the hallmarks of cancer. EMBO Rep. 2014;
65. Netti PA, Berk DA, Swartz MA, Grodzinsky AJ, Jain RK. Role of extracellular matrix assembly in interstitial transport in solid tumors. Cancer Res. 2000;
66. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009.
67. Wang J, Cazzato E, Ladewig E, Frattini V, Rosenbloom DIS, Zairis S, et al. Clonal evolution of glioblastoma under therapy. Nat Genet. 2016;
68. Sakthikumar S, Sakthikumar S, Roy A, Haseeb L, Pettersson ME, Sundström E, et al. Whole-genome sequencing of glioblastoma reveals enrichment of non-coding constraint mutations in known and novel genes. Genome Biol. 2020;
69. Loeb LA. Human cancers express mutator phenotypes: Origin, consequences and targeting. Nat. Rev. Cancer. 2011.
70. Kyritsis AP, Bondy ML, Rao JS, Sioka C. Inherited predisposition to glioma. Neuro. Oncol. 2010.
71. Wang X, Zhou R, Xiong Y, Zhou L, Yan X, Wang M, et al. Sequential fate-switches in stem-like cells drive the tumorigenic trajectory from human neural stem cells to malignant glioma. Cell Res. 2021;
72. Carrasco-García E, Saceda M, Martínez-Lacaci I. Role of Receptor Tyrosine Kinases and Their Ligands in Glioblastoma. Cells. 2014;
73. Du Z, Lovly CM. Mechanisms of receptor tyrosine kinase activation in cancer. Mol. Cancer. 2018.
74. Hatanpaa KJ, Burma S, Zhao D, Habib AA. Epidermal growth factor receptor in glioma: Signal transduction, neuropathology, imaging, and radioresistance1. Neoplasia. 2010.
75. An Z, Aksoy O, Zheng T, Fan QW, Weiss WA. Epidermal growth factor receptor and EGFRvIII in glioblastoma: Signaling pathways and targeted therapies. Oncogene. 2018.
76. Xu H, Zong H, Ma C, Ming X, Shang M, Li K, et al. Epidermal growth factor receptor in glioblastoma (Review). Oncol Lett. 2017;
77. Zhang Y, Dube C, Gibert M, Cruickshanks N, Wang B, Coughlan M, et al. The p53 pathway in glioblastoma. Cancers (Basel). 2018.
78. Lane DP. p53, guardian of the genome. Nature. 1992.
79. Kastenhuber ER, Lowe SW. Putting p53 in Context. Cell. 2017.
80. Freire P, Vilela M, Deus H, Kim YW, Koul D, Colman H, et al. Exploratory analysis of the copy number alterations in glioblastoma multiforme. PLoS One. 2008;
81. Lopez-Gines C, Cerda-Nicolas M, Gil-Benso R, Pellin A, Lopez-Guerrero JA, Callaghan R, et al. Association of chromosome 7, chromosome 10 and EGFR gene amplification in glioblastoma multiforme. Clin Neuropathol. 2005;
82. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB, et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. N Engl J Med. 2005;
83. Wilson TA, Karajannis MA, Harter DH. Glioblastoma multiforme: State of the art and future therapeutics. Surg Neurol Int. 2014;
84. Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol. 2009;

85. Weller M, Cloughesy T, Perry JR, Wick W. Standards of care for treatment of recurrent glioblastoma-are we there yet? Neuro. Oncol. 2013.
86. Stummer W, Pichlmeier U, Meinel T, Wiestler OD, Zanella F, Reulen HJ. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. Lancet Oncol. 2006;
87. Keles GE, Anderson B, Berger MS. The effect of extent of resection on time to tumor progression and survival in patients with glioblastoma multiforme of the cerebral hemisphere. Surg Neurol. 1999;
88. Zhao S, Wu J, Wang C, Liu H, Dong X, Shi C, et al. Intraoperative Fluorescence-Guided Resection of High-Grade Malignant Gliomas Using 5-Aminolevulinic Acid-Induced Porphyrins: A Systematic Review and Meta-Analysis of Prospective Studies. PLoS One. 2013.
89. Barajas RF, Chang JS, Segal MR, Parsa AT, McDermott MW, Berger MS, et al. Differentiation of recurrent glioblastoma multiforme from radiation necrosis after external beam radiation therapy with dynamic susceptibility- weighted contrast-enhanced perfusion MR imaging. Radiology. 2009;
90. Liu S, Wang Y, Xu K, Wang Z, Fan X, Zhang C, et al. Relationship between necrotic patterns in glioblastoma and patient survival: Fractal dimension and lacunarity analyses using magnetic resonance imaging. Sci Rep. 2017;
91. BOLCAEN J, ACOU M, DESCAMPS B, KERSEMANS K, DEBLAERE K, VANHOVE C, et al. PET for Therapy Response Assessment in Glioblastoma. Glioblastoma. 2017.
92. Kim J, Lee IH, Cho HJ, Park CK, Jung YS, Kim Y, et al. Spatiotemporal Evolution of the Primary Glioblastoma Genome. Cancer Cell. 2015;
93. Stark AM, Doukas A, Hugo HH, Hedderich J, Hattermann K, Mehdorn HM, et al. Expression of DNA mismatch repair proteins MLH1, MSH2, and MSH6 in recurrent glioblastoma. Neurol Res. 2015;
94. Martinez R, Setien F, Voelter C, Casado S, Quesada MP, Schackert G, et al. CpG island promoter hypermethylation of the pro-apoptotic gene caspase-8 is a common hallmark of relapsed glioblastoma multiforme. Carcinogenesis. 2007;
95. Klughammer J, Kiesel B, Roetzer T, Fortelny N, Nemc A, Nenning KH, et al. The DNA methylation landscape of glioblastoma disease progression shows extensive heterogeneity in time and space. Nat Med. 2018;
96. de Souza CF, Sabedot TS, Malta TM, Stetson L, Morozova O, Sokolov A, et al. A Distinct DNA Methylation Shift in a Subset of Glioma CpG Island Methylator Phenotypes during Tumor Recurrence. Cell Rep. 2018;
97. Andor N, Harness J V., Müller S, Mewes HW, Petritsch C. Expands: Expanding ploidy and allele frequency on nested subpopulations. Bioinformatics. 2014;
98. Soeda A, Hara A, Kunisada T, Yoshimura SI, Iwama T, Park DM. The evidence of glioblastoma heterogeneity. Sci Rep. 2015;
99. Li SC, Lee KL, Luo J. Control dominating subclones for managing cancer progression and posttreatment recurrence by subclonal switchboard signal: Implication for new therapies. Stem Cells Dev. 2012;
100. Lyle S, Moore N. Quiescent, slow-cycling stem cell populations in cancer: A review of the evidence and discussion of significance. J. Oncol. 2011.
101. Chen K, Zhang C, Ling S, Wei R, Wang J, Xu X. The metabolic flexibility of quiescent CSC: implications for chemotherapy resistance. Cell Death Dis. 2021.
102. Bao S, Wu Q, McLendon RE, Hao Y, Shi Q, Hjelmeland AB, et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature. 2006;
103. Lathia JD, Hitomi M, Gallagher J, Gadani SP, Adkins J, Vasanji A, et al. Distribution of CD133 reveals glioma stem cells self-renew through symmetric and asymmetric cell divisions. Cell Death Dis. 2011;

104. Richichi C, Brescia P, Alberizzi V, Fornasari L, Pelicci G. Marker-independent method for isolating slow-dividing cancer stem cells in human glioblastoma. Neoplasia (United States). 2013;
105. Friedmann-Morvinski D, Verma IM. Dedifferentiation and reprogramming: Origins of cancer stem cells. EMBO Rep. 2014.
106. Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, et al. Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. Cell. 2016;
107. Hall AW, Battenhouse AM, Shivram H, Morris AR, Cowperthwaite MC, Shpak M, et al. Bivalent chromatin domains in glioblastoma reveal a subtype-specific signature of glioma stem cells. Cancer Res. 2018;
108. Liao P, Ostrom QT, Stetson L, Barnholtz-Sloan JS. Models of epigenetic age capture patterns of DNA methylation in glioma associated with molecular subtype, survival, and recurrence. Neuro Oncol. 2018;
109. Kraus TFJ, Globisch D, Wagner M, Eigenbrod S, Widmann D, Münzel M, et al. Low values of 5-hydroxymethylcytosine (5hmC), the "sixth base," are associated with anaplasia in human brain tumors. Int J Cancer. 2012;
110. Esteller M. CpG island hypermethylation and tumor suppressor genes: A booming present, a brighter future. Oncogene. 2002.
111. Lahtz C, Pfeifer GP. Epigenetic changes of DNA repair genes in cancer. J Mol Cell Biol. 2011;
112. Ehrlich M. DNA methylation in cancer: Too much, but also too little. Oncogene. 2002.
113. Moarii M, Boeva V, Vert JP, Reyal F. Changes in correlation between promoter methylation and gene expression in cancer. BMC Genomics. 2015;
114. Cheung HH, Lee TL, Rennert OM, Chan WY. DNA methylation of cancer genome. Birth Defects Res. Part C - Embryo Today Rev. 2009.
115. Yu W, Zhang L, Wei Q, Shao A. O6-Methylguanine-DNA Methyltransferase (MGMT): Challenges and New Opportunities in Glioma Chemotherapy. Front. Oncol. 2020.
116. Oldrini B, Vaquero-Siguero N, Mu Q, Kroon P, Zhang Y, Galán-Ganga M, et al. MGMT genomic rearrangements contribute to chemotherapy resistance in gliomas. Nat Commun. 2020;
117. Trabelsi S, Mama N, Ladib M, Karmeni N, Haddaji Mastouri M, Chourabi M, et al. MGMT methylation assessment in glioblastoma: MS-MLPA versus human methylation 450K beadchip array and immunohistochemistry. Clin Transl Oncol. 2016;
118. Wang Z, Wu XL, Wang Y. A framework for analyzing DNA methylation data from Illumina Infinium HumanMethylation450 BeadChip. BMC Bioinformatics. 2018;
119. Li Y, Tollefsbol TO. DNA methylation detection: Bisulfite genomic sequencing analysis. Methods Mol Biol. 2011;
120. Pharo HD, Honne H, Vedeld HM, Dahl C, Andresen K, Liestøl K, et al. Experimental factors affecting the robustness of DNA methylation analysis. Sci Rep. 2016;
121. Ohta T, Watanabe T, Katayama Y, Yoshino A, Yachi K, Ogino A, et al. Aberrant promoter hypermethylation profile of cell cycle regulatory genes in malignant astrocytomas. Oncol Rep. 2006;
122. Zheng S, Houseman EA, Morrison Z, Wrensch MR, Patoka JS, Ramos C, et al. DNA hypermethylation profiles associated with glioma subtypes and EZH2 and IGFBP2 mRNA expression. Neuro Oncol. 2011;
123. Mazor T, Pankov A, Johnson BE, Hong C, Hamilton EG, Bell RJA, et al. DNA Methylation and Somatic Mutations Converge on the Cell Cycle and Define Similar Evolutionary Histories in Brain Tumors. Cancer Cell. 2015;
124. Etcheverry A, Aubry M, de Tayrac M, Vauleon E, Boniface R, Guenot F, et al. DNA

methylation in glioblastoma: Impact on gene expression and clinical outcome. BMC Genomics. 2010;
125. Malta TM, De Souza CF, Sabedot TS, Silva TC, Mosella MS, Kalkanis SN, et al. Glioma CpG island methylator phenotype (G-CIMP): Biological and clinical implications. Neuro Oncol. 2018;
126. Shinawi T, Hill VK, Krex D, Schackert G, Gentle D, Morris MR, et al. DNA methylation profiles of long- and short-term glioblastoma survivors. Epigenetics. 2013;
127. Paço A, de Bessa Garcia SA, Freitas R. Methylation in HOX Clusters and Its Applications in Cancer Therapy. Cells. 2020.
128. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature. 2000;
129. Héberlé É, Bardet AF. Sensitivity of transcription factors to DNA methylation. Essays Biochem. 2019.
130. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science (80-). 2017;
131. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, et al. IDH1 and IDH2 Mutations in Gliomas . N Engl J Med. 2009;
132. Yang H, Ye D, Guan KL, Xiong Y. IDH1 and IDH2 mutations in tumorigenesis: Mechanistic insights and clinical perspectives. Clin. Cancer Res. 2012.
133. Huang LE. Friend or foe-IDH1 mutations in glioma 10 years on. Carcinogenesis. 2019.
134. Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature. 2009;
135. Jin G, Reitman ZJ, Spasojevic I, Batinic-Haberle I, Yang J, Schmidt-Kittler O, et al. 2-hydroxyglutarate production, but not dominant negative function, is conferred by glioma-derived NADP+-dependent isocitrate dehydrogenase mutations. PLoS One. 2011;
136. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim SH, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer Cell. 2011;
137. Han S, Liu Y, Cai SJ, Qian M, Ding J, Larion M, et al. IDH mutation in glioma: molecular mechanisms and potential therapeutic targets. Br. J. Cancer. 2020.
138. Kim YZ. Altered Histone Modifications in Gliomas. Brain Tumor Res Treat. 2014;
139. Turner BM. Reading signals on the nucleosome with a new nomenclature for modified histones. Nat. Struct. Mol. Biol. 2005.
140. Wong CM, Wong CCL, Ng YL, Au SLK, Ko FCF, Ng IOL. Transcriptional repressive H3K9 and H3K27 methylations contribute to DNMT1-mediated DNA methylation recovery. PLoS One. 2011;
141. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. Exp. Mol. Med. 2017.
142. Gräff J, Tsai LH. Histone acetylation: Molecular mnemonics on the chromatin. Nat. Rev. Neurosci. 2013.
143. Hu Y, Lu Y, Zhao Y, Zhou DX. Histone Acetylation Dynamics Integrates Metabolic Activity to Regulate Plant Response to Stress. Front. Plant Sci. 2019.
144. Gallinari P, Di Marco S, Jones P, Pallaoro M, Steinkühler C. HDACs, histone deacetylation and gene transcription: From molecular biology to cancer therapeutics. Cell Res. 2007;
145. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011.
146. Dueva R, Akopyan K, Pederiva C, Trevisan D, Dhanjal S, Lindqvist A, et al. Neutralization of the Positive Charges on Histone Tails by RNA Promotes an Open Chromatin Structure. Cell Chem Biol. 2019;
147. Shoaib M, Chen Q, Shi X, Nair N, Prasanna C, Yang R, et al. Histone H4 lysine 20 mono-

methylation directly facilitates chromatin openness and promotes transcription of housekeeping genes. Nat Commun. 2021;
148. Duan J. Chromatin modification and remodeling in schizophrenia. Chromatin Signal Neurol Disord. 2019.
149. Chen Y, Ren B, Yang J, Wang H, Yang G, Xu R, et al. The role of histone methylation in the development of digestive cancers: a potential direction for cancer management. Signal Transduct. Target. Ther. 2020.
150. Jambhekar A, Dhall A, Shi Y. Roles and regulation of histone methylation in animal development. Nat. Rev. Mol. Cell Biol. 2019.
151. Vakoc CR, Sachdeva MM, Wang H, Blobel GA. Profile of Histone Lysine Methylation across Transcribed Mammalian Chromatin. Mol Cell Biol. 2006;
152. Schwartzentruber J, Korshunov A, Liu XY, Jones DTW, Pfaff E, Jacob K, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature. 2012;
153. Wang XQ, Bai HM, Li ST, Sun H, Min LZ, Tao BB, et al. Knockdown of HDAC1 expression suppresses invasion and induces apoptosis in glioma cells. Oncotarget. 2017;
154. Zang L, Kondengaden SM, Che F, Wang L, Heng X. Potential epigenetic-based therapeutic targets for glioma. Front. Mol. Neurosci. 2018.
155. Cornago M, Garcia-Alberich C, Blasco-Angulo N, Vall-llaura N, Nager M, Herreros J, et al. Histone deacetylase inhibitors promote glioma cell death by G2 checkpoint abrogation leading to mitotic catastrophe. Cell Death Dis. 2014;
156. Auzmendi-Iriarte J, Saenz-Antoñanzas A, Mikelez-Alonso I, Carrasco-Garcia E, Tellaetxe-Abete M, Lawrie CH, et al. Characterization of a new small-molecule inhibitor of HDAC6 in glioblastoma. Cell Death Dis. 2020;
157. Karagiannis TC, El-Osta A. Modulation of cellular radiation responses by histone deacetylase inhibitors. Oncogene. 2006.
158. Shi W, Palmer JD, Werner-Wasik M, Andrews DW, Evans JJ, Glass J, et al. Phase I trial of panobinostat and fractionated stereotactic re-irradiation therapy for recurrent high grade gliomas. J Neurooncol. 2016;
159. Drappatz J, Lee EQ, Hammond S, Grimm SA, Norden AD, Beroukhim R, et al. Phase i study of panobinostat in combination with bevacizumab for recurrent high-grade glioma. J Neurooncol. 2012;
160. Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. Nat. Rev. Genet. 2019.
161. Daenen F, van Roy F, De Bleser PJ. Low nucleosome occupancy is encoded around functional human transcription factor binding sites. BMC Genomics. 2008;
162. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature. 2012;
163. Sheffield NC, Furey TS. Identifying and characterizing regulatory sequences in the human genome with chromatin accessibility assays. Genes (Basel). 2012.
164. Turner BM. Defining an epigenetic code. Nat Cell Biol. 2007;
165. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;
166. Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. Hi-C: A comprehensive technique to capture the conformation of genomes. Methods. 2012;
167. Minnoye L, Marinov GK, Krausgruber T, Pan L, Marand AP, Secchia S, et al. Chromatin accessibility profiling methods. Nat Rev Methods Prim. 2021;
168. Szabo Q, Bantignies F, Cavalli G. Principles of genome folding into topologically associating domains. Sci. Adv. 2019.

169. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;
170. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell. 2017;
171. Won H, De La Torre-Ubieta L, Stein JL, Parikshak NN, Huang J, Opland CK, et al. Chromosome conformation elucidates regulatory relationships in developing human brain. Nature. 2016;
172. Clapier CR, Iwasa J, Cairns BR, Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat. Rev. Mol. Cell Biol. 2017.
173. Stanton BZ, Hodges C, Crabtree GR, Zhao K. A General Non-Radioactive ATPase Assay for Chromatin Remodeling Complexes. Curr Protoc Chem Biol. 2017;
174. Kwon H, Imbalzano AN, Khavari PA, Kingston RE, Green MR. Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. Nature. 1994;
175. Sokpor G, Xie Y, Rosenbusch J, Tuoc T. Chromatin remodeling BAF (SWI/SNF) complexes in neural development and disorders. Front. Mol. Neurosci. 2017.
176. Kadam S, McAlpine GS, Phelan ML, Kingston RE, Jones KA, Emerson BM. Functional selectivity of recombinant mammalian SWI/SNF subunits. Genes Dev. 2000;
177. Chen G, Zhou H, Liu B, Wang Y, Zhao J, Giancotti FG, et al. A heterotrimeric SMARCB1–SMARCC2 subcomplex is required for the assembly and tumor suppression function of the BAF chromatin-remodeling complex. Cell Discov. 2020.
178. Mittal P, Roberts CWM. The SWI/SNF complex in cancer — biology, biomarkers and therapy. Nat. Rev. Clin. Oncol. 2020.
179. Hu G, Schones DE, Cui K, Ybarra R, Northrup D, Tang Q, et al. Regulation of nucleosome landscape and transcription factor targeting at tissue-specific enhancers by BRG1. Genome Res. 2011;
180. Tolstorukov MY, Sansam CG, Lu P, Koellhoffer EC, Helming KC, Alver BH, et al. Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. Proc Natl Acad Sci U S A. 2013;
181. Wang X, Lee RS, Alver BH, Haswell JR, Wang S, Mieczkowski J, et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. Nat Genet. 2017;
182. Simon JA, Kingston RE. Mechanisms of Polycomb gene silencing: Knowns and unknowns. Nat. Rev. Mol. Cell Biol. 2009.
183. Wilson BG, Wang X, Shen X, McKenna ES, Lemieux ME, Cho YJ, et al. Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. Cancer Cell. 2010;
184. Kadoch C, Williams RT, Calarco JP, Miller EL, Weber CM, Braun SMG, et al. Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states. Nat Genet. 2017;
185. Bögershausen N, Wollnik B. Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. Front. Mol. Neurosci. 2018.
186. Hiramatsu H, Kobayashi K, Kobayashi K, Haraguchi T, Ino Y, Todo T, et al. The role of the SWI/SNF chromatin remodeling complex in maintaining the stemness of glioma initiating cells. Sci Rep. 2017;
187. Yang C, Wang Y, Sims MM, He Y, Miller DD, Pfeffer LM. Targeting the bromodomain of brg-1/brm subunit of the swi/snf complex increases the anticancer activity of temozolomide in glioblastoma. Pharmaceuticals. 2021;
188. Heppler LN, Frank DA. Targeting Oncogenic Transcription Factors: Therapeutic Implications of Endogenous STAT Inhibitors. Trends in Cancer. 2017.
189. Thomas MC, Chiang CM. The general transcription machinery and general cofactors. Crit. Rev. Biochem. Mol. Biol. 2006.

190. van Steensel B, Furlong EEM. The role of transcription in shaping the spatial organization of the genome. Nat. Rev. Mol. Cell Biol. 2019.
191. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. Cell. 2013.
192. Littlewood TD, Kreuzaler P, Evan GI. All Things to All People. Cell. 2012.
193. Miller DM, Thomas SD, Islam A, Muench D, Sedoris K. c-Myc and cancer metabolism. Clin. Cancer Res. 2012.
194. Hoffman B, Liebermann DA. Apoptotic signaling by c-MYC. Oncogene. 2008.
195. Karin M, Cao Y, Greten FR, Li ZW. NF-κB in cancer: From innocent bystander to major culprit. Nat. Rev. Cancer. 2002.
196. Xia Y, Shen S, Verma IM. NF-κB, an active player in human cancers. Cancer Immunol. Res. 2014.
197. Taniguchi K, Karin M. NF-B, inflammation, immunity and cancer: Coming of age. Nat. Rev. Immunol. 2018.
198. Suzuki T, Okuno H, Yoshida T, Endo T, Nishina H, Iba H. Difference in transcriptional regulatory function between c-Fos and Fra-2. Nucleic Acids Res. 1991;
199. Leppä S, Eriksson M, Saffrich R, Ansorge W, Bohmann D. Complex Functions of AP-1 Transcription Factors in Differentiation and Survival of PC12 Cells. Mol Cell Biol. 2001;
200. Mason JM, Schmitz MA, Müller KM, Arndt KM. Semirational design of Jun-Fos coiled coils with increased affinity: Universal implications for leucine zipper prediction and design. Proc Natl Acad Sci U S A. 2006;
201. Cargnello M, Roux PP. Activation and Function of the MAPKs and Their Substrates, the MAPK-Activated Protein Kinases. Microbiol Mol Biol Rev. 2011;
202. Kaminska B. MAPK signalling pathways as molecular targets for anti-inflammatory therapy - From molecular mechanisms to therapeutic benefits. Biochim Biophys Acta - Proteins Proteomics. 2005.
203. Hettinger K, Vikhanskaya F, Poh MK, Lee MK, de Belle I, Zhang JT, et al. c-Jun promotes cellular survival by suppression of PTEN. Cell Death Differ. 2007;
204. Bakiri L, Lallemand D, Bossy-Wetzel E, Yaniv M. Cell cycle-dependent variations in c-Jun and JunB phosphorylation: A role in the control of cyclin D1 expression. EMBO J. 2000;
205. Zenz R, Eferl R, Scheinecker C, Redlich K, Smolen J, Schonthaler HB, et al. Activator protein 1 (Fos/Jun) functions in inflammatory bone and skin disease. Arthritis Res. Ther. 2008.
206. Santaguida M, Schepers K, King B, Sabnis AJ, Forsberg EC, Attema JL, et al. JunB Protects against Myeloid Malignancies by Limiting Hematopoietic Stem Cell Proliferation and Differentiation without Affecting Self-Renewal. Cancer Cell. 2009;
207. Piechaczyk M, Farràs R. Regulation and function of JunB in cell proliferation. Biochem Soc Trans. 2008.
208. Vleugel MM, Greijer AE, Bos R, van der Wall E, van Diest PJ. c-Jun activation is associated with proliferation and angiogenesis in invasive breast cancer. Hum Pathol. 2006;
209. Watanabe T, Hiasa Y, Tokumoto Y, Hirooka M, Abe M, Ikeda Y, et al. Protein Kinase R Modulates c-Fos and c-Jun Signaling to Promote Proliferation of Hepatocellular Carcinoma with Hepatitis C Virus Infection. PLoS One. 2013;
210. Matthews C, Colburn N, Young M. AP-1 a Target for Cancer Prevention. Curr Cancer Drug Targets. 2007;
211. Mees C, Nemunaitis J, Senzer N. Transcription factors: Their potential as targets for an individualized therapeutic approach to cancer. Cancer Gene Ther. 2009.
212. Cheng Q, Huang C, Cao H, Lin J, Gong X, Li J, et al. A Novel Prognostic Signature of Transcription Factors for the Prediction in Patients With GBM. Front Genet. 2019;
213. Qin G, Hu B, Li X, Li R, Meng Y, Wang Y, et al. Identification of Key Differentially Expressed Transcription Factors in Glioblastoma. J Oncol. 2020;
214. Li ZH, Guan YL, Zhang G Bin. Genomic Analysis of Glioblastoma Multiforme Reveals a Key

Transcription Factor Signature Relevant to Prognosis and the Immune Processes. Front Oncol. 2021;
215. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;
216. Sedlazeck FJ, Rescheneder P, Von Haeseler A. NextGenMap: Fast and accurate read mapping in highly polymorphic genomes. Bioinformatics. 2013;
217. Broad Institute. Picard Tools - By Broad Institute. Github. 2009.
218. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;
219. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012;
220. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;
221. Tamborero D, Gonzalez-Perez A, Lopez-Bigas N. OncodriveCLUST: Exploiting the positional clustering of somatic mutations to identify cancer genes. Bioinformatics. 2013;
222. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics. 2007;
223. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;
224. Wang L, Wang S, Li W. RSeQC: Quality control of RNA-seq experiments. Bioinformatics. 2012;
225. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;
226. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;
227. Aran D, Hu Z, Butte AJ. xCell: Digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;
228. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;
229. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;
230. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. quanTIseq: quantifying immune contexture of human tumors. bioRxiv. 2017;
231. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res. 2017;
232. Fischer AH, Jacobson K a, Rose J, Zeller R. Hematoxylin and Eosin (H & E) staining. CSH Protoc. 2005;
233. Rasband W. ImageJ [Software]. U S Natl Institutes Heal Bethesda, Maryland, USA. 2015;
234. Stępniak K, Machnicka MA, Mieczkowski J, Macioszek A, Wojtaś B, Gielniewski B, et al. Mapping chromatin accessibility and active regulatory elements reveals pathological mechanisms in human gliomas. Nat Commun. 2021;12.
235. Andrews S. FastQC - A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Babraham Bioinforma. 2010;
236. Gordon A, Hannon GJ. Fastx-toolkit. FASTQ/A short-reads pre-processing tools. Unpubl http//hannonlab cshl edu/fastx_ toolkit. 2010;
237. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;
238. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: An R package for comparing biological themes among gene clusters. Omi A J Integr Biol. 2012;

239. Kulakovskiy I V., Vorontsov IE, Yevshin IS, Sharipov RN, Fedorova AD, Rumynskiy EI, et al. HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucleic Acids Res. 2018;
240. Grant CE, Bailey TL, Noble WS. FIMO: Scanning for occurrences of a given motif. Bioinformatics. 2011;
241. D'Oliveira Albanus R, Kyono Y, Hensley J, Varshney A, Orchard P, Kitzman JO, et al. Chromatin information content landscapes inform transcription factor and DNA interactions. Nat Commun. 2021;
242. Yu G, Wang LG, He QY. ChIP seeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;
243. Xi Y, Li W. BSMAP: Whole genome bisulfite sequence MAPping program. BMC Bioinformatics. 2009;
244. Ahmed SU, Carruthers R, Gilmour L, Yildirim S, Watts C, Chalmers AJ. Selective inhibition of parallel DNA damage response pathways optimizes radiosensitization of glioblastoma stem-like cells. Cancer Res. 2015;
245. Ou J, Liu H, Yu J, Kelliher MA, Castilla LH, Lawson ND, et al. ATACseqQC: A Bioconductor package for post-alignment quality assessment of ATAC-seq data. BMC Genomics. 2018;
246. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;
247. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010;
248. Gocho Y, Kiyokawa N, Ichikawa H, Nakabayashi K, Osumi T, Ishibashi T, et al. A novel recurrent EP300-ZNF384 gene fusion in B-cell precursor acute lymphoblastic leukemia. Leukemia. 2015.
249. Guo C, McDowell IC, Nodzenski M, Scholtens DM, Allen AS, Lowe WL, et al. Transversions have larger regulatory effects than transitions. BMC Genomics. 2017;
250. Chen Y, Song Y, Du W, Gong L, Chang H, Zou Z. Tumor-associated macrophages: An accomplice in solid tumor progression. J. Biomed. Sci. 2019.
251. Vidyarthi A, Agnihotri T, Khan N, Singh S, Tewari MK, Radotra BD, et al. Predominance of M2 macrophages in gliomas leads to the suppression of local and systemic immunity. Cancer Immunol Immunother. 2019;
252. Szulzewsky F, Pelz A, Feng X, Synowitz M, Markovic D, Langmann T, et al. Glioma-associated microglia/macrophages display an expression profile different from M1 and M2 polarization and highly express Gpnmb and Spp1. PLoS One. 2015;
253. Seifert M, Garbe M, Friedrich B, Mittelbronn M, Klink B. Comparative transcriptomics reveals similarities and differences between astrocytoma grades. BMC Cancer. 2015;
254. Xu B. Prediction and analysis of hub genes between glioblastoma and low-grade glioma using bioinformatics analysis. Medicine (Baltimore). 2021;
255. Karunasena E, McIver LJ, Rood BR, Wu X, Zhu H, Bavarva JH, et al. Somatic intronic microsatellite loci differentiate glioblastoma from lower-grade gliomas. Oncotarget. 2014;
256. Shen Y, Grisdale CJ, Islam SA, Bose P, Lever J, Zhao EY, et al. Comprehensive genomic profiling of glioblastoma tumors, BTICs, and xenografts reveals stability and adaptation to growth environments. Proc Natl Acad Sci U S A. 2019;
257. Zheng Q, Diao S, Wang Q, Zhu C, Sun X, Yin B, et al. IL-17A promotes cell migration and invasion of glioblastoma cells via activation of PI3K/AKT signalling pathway. J Cell Mol Med. 2019;
258. Jung E, Alfonso J, Osswald M, Monyer H, Wick W, Winkler F. Emerging intersections between neuroscience and glioma biology. Nat. Neurosci. 2019.
259. Geng RX, Li N, Xu Y, Liu J hui, Yuan F en, Sun Q, et al. Identification of core biomarkers

associated with outcome in glioma: Evidence from bioinformatics analysis. Dis Markers. 2018;
260. Johnson GL, Nakamura K. The c-jun kinase/stress-activated pathway: Regulation, function and role in human disease. Biochim. Biophys. Acta - Mol. Cell Res. 2007.
261. Pulverer BJ, Kyriakis JM, Avruch J, Nikolakaki E, Woodgett JR. Phosphorylation of c-jun mediated by MAP kinases. Nature. 1991;
262. Azofeifa JG, Allen MA, Hendrix JR, Read T, Rubin JD, Dowell RD. Enhancer RNA profiling predicts transcription factor activity. Genome Res. 2018;
263. Panigrahi A, O'Malley BW. Mechanisms of enhancer action: the known and the unknown. Genome Biol. 2021.
264. Gan Z, Ding L, Burckhardt CJ, Lowery J, Zaritsky A, Sitterley K, et al. Vimentin Intermediate Filaments Template Microtubule Networks to Enhance Persistence in Cell Polarity and Directed Migration. Cell Syst. 2016;
265. Duarte S, Viedma-Poyatos Á, Navarro-Carrasco E, Martínez AE, Pajares MA, Pérez-Sala D. Vimentin filaments interact with the actin cortex in mitosis allowing normal cell division. Nat Commun. 2019;
266. Satelli A, Li S. Vimentin in cancer and its potential as a molecular target for cancer therapy. Cell. Mol. Life Sci. 2011.
267. Chakraborty S, Kumar A, Faheem MM, Katoch A, Kumar A, Jamwal VL, et al. Vimentin activation in early apoptotic cancer cells errands survival pathways during DNA damage inducer CPT treatment in colon carcinoma model. Cell Death Dis. 2019;
268. Alver BH, Kim KH, Lu P, Wang X, Manchester HE, Wang W, et al. The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. Nat Commun. 2017;
269. Vander Ark A, Cao J, Li X. TGF-β receptors: In and beyond TGF-β signaling. Cell. Signal. 2018.
270. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin A V., et al. Signatures of mutational processes in human cancer. Nature. 2013;
271. Walens A, Lin J, Damrauer JS, McKinney B, Lupo R, Newcomb R, et al. Adaptation and selection shape clonal evolution of tumors during residual disease and recurrence. Nat Commun. 2020;
272. Ettl T, Baader K, Stiegler C, Müller M, Agaimy A, Zenk J, et al. Loss of PTEN is associated with elevated EGFR and HER2 expression and worse prognosis in salivary gland cancer. Br J Cancer. 2012;
273. Li J, Liang R, Song C, Xiang Y, Liu Y. Prognostic significance of epidermal growth factor receptor expression in glioma patients. Onco Targets Ther. 2018;
274. Han F, Hu R, Yang H, Liu J, Sui J, Xiang X, et al. PTEN gene mutations correlate to poor prognosis in glioma patients: A meta-analysis. Onco Targets Ther. 2016;
275. Benitez JA, Ma J, D'Antonio M, Boyer A, Camargo MF, Zanca C, et al. PTEN regulates glioblastoma oncogenesis through chromatin-associated complexes of DAXX and histone H3.3. Nat Commun. 2017;
276. Hirabayashi S, Ohki K, Nakabayashi K, Ichikawa H, Momozawa Y, Okamura K, et al. ZNF384-related fusion genes define a subgroup of childhood B-cell precursor acute lymphoblastic leukemia with a characteristic immunotype. Haematologica. 2017;
277. Young SK, Shao Y, Bidwell JP, Wek RC. Nuclear matrix protein 4 is a novel regulator of ribosome biogenesis and controls the unfolded protein response via repression of Gadd34 expression. J Biol Chem. 2016;
278. Blijlevens M, Li J, van Beusechem VW. Biology of the mrna splicing machinery and its dysregulation in cancer providing therapeutic opportunities. Int. J. Mol. Sci. 2021.
279. Fong JY, Pignata L, Goy PA, Kawabata KC, Lee SCW, Koh CM, et al. Therapeutic Targeting

of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation. Cancer Cell. 2019;
280. Yang CH, Wang Y, Sims M, Cai C, He P, Häcker H, et al. MicroRNA203a suppresses glioma tumorigenesis through an ATM-dependent interferon response pathway. Oncotarget. 2017;
281. Kang DC, Gopalkrishnan R V., Lin L, Randolph A, Valerie K, Pestka S, et al. Expression analysis and genomic characterization of human melanoma differentiation associated gene-5, mda-5: A novel type I interferon-responsive apoptosis-inducing gene. Oncogene. 2004;
282. Du Z, Cai C, Sims M, Boop FA, Davidoff AM, Pfeffer LM. The effects of type I interferon on glioblastoma cancer stem cells. Biochem Biophys Res Commun. 2017;
283. Silginer M, Nagy S, Happold C, Schneider H, Weller M, Roth P. Autocrine activation of the IFN signaling pathway may promote immune escape in glioblastoma. Neuro Oncol. 2017;
284. Gangoso E, Southgate B, Bradley L, Rus S, Galvez-Cancino F, McGivern N, et al. Glioblastomas acquire myeloid-affiliated transcriptional programs via epigenetic immunoediting to elicit immune evasion. Cell. 2021;
285. Tanabe T, Kominsky SL, Subramaniam PS, Johnson HM, Torres BA. Inhibition of the glioblastoma cell cycle by type I IFNs occurs at both the G1 and S phases and correlates with the upregulation of p21(WAF1/CIP1). J Neurooncol. 2000;
286. Gozzelino L, De Santis MC, Gulluni F, Hirsch E, Martini M. PI(3,4)P2 Signaling in Cancer and Metabolism. Front. Oncol. 2020.
287. Chin YR, Toker A. Function of Akt/PKB signaling to cell motility, invasion and the tumor stroma in cancer. Cell. Signal. 2009.
288. Koundouros N, Poulogiannis G. Reprogramming of fatty acid metabolism in cancer. Br. J. Cancer. 2020.
289. Granucci F, Zanoni I, Feau S, Ricciardi-Castagnoli P. Dendritic cell regulation of immune responses: A new role for interleukin 2 at the intersection of innate and adaptive immunity. EMBO J. 2003.
290. Granucci F, Zanoni I, Ricciardi-Castagnoli P. Central role of dendritic cells in the regulation and deregulation of immune responses. Cell. Mol. Life Sci. 2008.
291. Kim R, Emi M, Tanabe K. Functional roles of immature dendritic cells in impaired immunity of solid tumour and their targeted strategies for provoking tumour immunity. Clin. Exp. Immunol. 2006.
292. Fucikova J, Palova-Jelinkova L, Bartunkova J, Spisek R. Induction of tolerance and immunity by dendritic cells: Mechanisms and clinical applications. Front. Immunol. 2019.
293. Moser M. Dendritic cells in immunity and tolerance - Do they display opposite functions? Immunity. 2003.
294. Ma Y, Shurin G V., Peiyuan Z, Shurin MR. Dendritic cells in the cancer microenvironment. J. Cancer. 2013.
295. Ludewig P, Gallizioli M, Urra X, Behr S, Brait VH, Gelderblom M, et al. Dendritic cells in brain diseases. Biochim Biophys Acta - Mol Basis Dis. 2016;
296. Skytthe MK, Graversen JH, Moestrup SK. Targeting of cd163+ macrophages in inflammatory and malignant diseases. Int J Mol Sci. 2020;
297. Aras S, Raza Zaidi M. TAMeless traitors: Macrophages in cancer progression and metastasis. Br. J. Cancer. 2017.
298. Yan F, Powell DR, Curtis DJ, Wong NC. From reads to insight: A hitchhiker's guide to ATAC-seq data analysis. Genome Biol. 2020.
299. Kolat D, Kaluzinska Z, Bednarek AK, Pluciennik E zbieta. The biological characteristics of transcription factors AP-2α and AP-2γ and their importance in various types of cancers. Biosci Rep. 2019;
300. Stuart ET, Kioussi C, Gruss P, Aguzzi A. PAX5 Expression Correlates with Increasing Malignancy in Human Astrocytomas. Clin Cancer Res. 1995;

301. Rhie SK, Yao L, Luo Z, Witt H, Schreiner S, Guo Y, et al. ZFX acts as a transcriptional activator in multiple types of human tumors by binding downstream from transcription start sites at the majority of CpG island promoters. Genome Res. 2018;
302. Fang X, Huang Z, Zhou W, Wu Q, Sloan AE, Ouyang G, et al. The zinc finger transcription factor ZFX Is required for maintaining the tumorigenic potential of glioblastoma stem cells. Stem Cells. 2014;
303. Kim HD, O'Shea EK. A quantitative model of transcription factor-activated gene expression. Nat Struct Mol Biol. 2008;
304. Chua G, Morris QD, Sopko R, Robinson MD, Ryan O, Chan ET, et al. Identifying transcription factor functions and targets by phenotypic activation. Proc Natl Acad Sci U S A. 2006;
305. Brennan A, Leech JT, Kad NM, Mason JM. Selective antagonism of cJun for cancer therapy. J. Exp. Clin. Cancer Res. 2020.
306. Zhang J, Chen YH, Lu Q. Pro-oncogenic and anti-oncogenic pathways: Opportunities and challenges of cancer therapy. Futur. Oncol. 2010.
307. Johnson R, Spiegelman B, Hanahan D, Wisdom R. Cellular transformation and malignancy induced by ras require c-jun. Mol Cell Biol. 1996;
308. Mariani O, Brennetot C, Coindre JM, Gruel N, Ganem C, Delattre O, et al. JUN Oncogene Amplification and Overexpression Block Adipocytic Differentiation in Highly Aggressive Sarcomas. Cancer Cell. 2007;
309. Wu S, Du Y, Beckford J, Alachkar H. Upregulation of the EMT marker vimentin is associated with poor clinical outcome in acute myeloid leukemia. J Transl Med. 2018;
310. Meyer-Schaller N, Cardner M, Diepenbruck M, Saxena M, Tiede S, Lüönd F, et al. A Hierarchical Regulatory Landscape during the Multiple Stages of EMT. Dev Cell. 2019;
311. Shi Y, Ping YF, Zhou W, He ZC, Chen C, Bian BSJ, et al. Tumour-associated macrophages secrete pleiotrophin to promote PTPRZ1 signalling in glioblastoma stem cells for tumour growth. Nat Commun. 2017;
312. Meng F, Li W, Li C, Gao Z, Guo K, Song S. CCL18 promotes epithelial-mesenchymal transition, invasion and migration of pancreatic cancer cells in pancreatic ductal adenocarcinoma. Int J Oncol. 2015;
313. Kim H, Claps G, Möller A, Bowtell D, Lu X, Ronai ZA. Siah2 regulates tight junction integrity and cell polarity through control of ASPP2 stability. Oncogene. 2014;
314. Chen A, Wong CSF, Liu MCP, House CM, Sceneay J, Bowtell DD, et al. The ubiquitin ligase Siah is a novel regulator of Zeb1 in breast cancer. Oncotarget. 2015;
315. Guan Y, Bhandari A, Zhang X, Wang O. Uridine phosphorylase 1 associates to biological and clinical significance in thyroid carcinoma cell lines. J Cell Mol Med. 2019;
316. Naz S, Bashir M, Ranganathan P, Bodapati P, Santosh V, Kondaiah P. Protumorigenic actions of S100A2 involve regulation of PI3/Akt signaling and functional interaction with Smad3. Carcinogenesis. 2014;
317. Shaulian E, Karin M. AP-1 in cell proliferation and survival. Oncogene. 2001.
318. Malnou CE, Brockly F, Favard C, Moquet-Torcy G, Piechaczyk M, Jariel-Encontre I. Heterodimerization with different jun proteins controls c-Fos intranuclear dynamics and distribution. J Biol Chem. 2010;
319. Gustems M, Woellmer A, Rothbauer U, Eck SH, Wieland T, Lutter D, et al. C-Jun/c-Fos heterodimers regulate cellular genes via a newly identified class of methylated DNA sequence motifs. Nucleic Acids Res. 2014;
320. Kullmann MK, Pegka F, Ploner C, Hengst L. Stimulation of c-Jun/AP-1-Activity by the Cell Cycle Inhibitor p57Kip2. Front Cell Dev Biol. 2021;
321. Rodríguez-Martínez JA, Reinke AW, Bhimsaria D, Keating AE, Ansari AZ. Combinatorial bZIP dimers display complex DNA-binding specificity landscapes. Elife. 2017;

322. Hess J, Angel P, Schorpp-Kistner M. AP-1 subunits: Quarrel and harmony among siblings. J Cell Sci. 2004;
323. Eferl R, Wagner EF. AP-1: A double-edged sword in tumorigenesis. Nat. Rev. Cancer. 2003.
324. Elias MC, Tozer KR, Silber JR, Mikheeva S, Deng M, Morrison RS, et al. TWIST is expressed in human gliomas and promotes invasion. Neoplasia. 2005;
325. Kurscheid S, Bady P, Sciuscio D, Samarzija I, Shay T, Vassallo I, et al. Chromosome 7 gain and DNA hypermethylation at the HOXA10 locus are associated with expression of a stem cell related HOX-signature in glioblastoma. Genome Biol. 2015;
326. Velez-Delvalle C, Marsch-Moreno M, Castro-Muñozledo F, Galván-Mendoza IJ, Kuri-Harcuch W. Epithelial cell migration requires the interaction between the vimentin and keratin intermediate filaments. Sci Rep. 2016;
327. Zhu QS, Rosenblatt K, Huang KL, Lahat G, Brobey R, Bolshakov S, et al. Vimentin is a novel AKT1 target mediating motility and invasion. Oncogene. 2011;
328. Zhao J, Zhang L, Dong X, Liu L, Huo L, Chen H. High Expression of Vimentin is Associated with Progression and a Poor Outcome in Glioblastoma. Appl Immunohistochem Mol Morphol. 2018;
329. Johnson KC, Anderson KJ, Courtois ET, Gujar AD, Barthel FP, Varn FS, et al. Single-cell multimodal glioma analyses identify epigenetic regulators of cellular plasticity and environmental stress response. Nat Genet. 2021;
330. Shain AH, Pollack JR. The Spectrum of SWI/SNF Mutations, Ubiquitous in Human Cancers. PLoS One. 2013;
331. Kretzschmar M, Liu F, Hata A, Doody J, Massagué J. The TGF-ß mediator Smad1 is directly phosphorylated and functionally activated by the BMP receptor kinase. Genes Dev. 1997;
332. Chau JFL, Jia D, Wang Z, Liu Z, Hu Y, Zhang X, et al. A crucial role for bone morphogenetic protein-Smad1 signalling in the DNA damage response. Nat Commun. 2012;
333. Massagué J. TGFβ signalling in context. Nat. Rev. Mol. Cell Biol. 2012.
334. Batlle E, Massagué J. Transforming Growth Factor-β Signaling in Immunity and Cancer. Immunity. 2019.
335. Ramachandran A, Vizán P, Das D, Chakravarty P, Vogt J, Rogers KW, et al. TGF-β uses a novel mode of receptor activation to phosphorylate SMAD1/5 and induce epithelial-to-mesenchymal transition. Elife. 2018;
336. van Caam A, Madej W, Garcia de Vinuesa A, Goumans MJ, ten Dijke P, Blaney Davidson E, et al. TGFβ1-induced SMAD2/3 and SMAD1/5 phosphorylation are both ALK5-kinase-dependent in primary chondrocytes and mediated by TAK1 kinase activity. Arthritis Res Ther. 2017;

Chapter V

5. Appendix

5.1. Supplementary tables

All supplementary tables can be found annexed with the presented thesis.

Supplementary Table 1. TF binding probabilities in glioma enhancers

Supplementary Table 2. Significant differences in histone H3K27ac ChiP-seq signal between GBMs and DAs

Supplementary Table 3. Prediction of c-Jun binding instances in glioma enhancers

Supplementary Table 1. **TF binding probabilities in glioma enhancers.** Hypergeometric test was used and obtained p-values were corrected using the Benjamini-Hochberg (BH) method.

TF model	Occurrences in enhancers	Occurrences in the genome	Hypergeometric test (p-value)	Hypergeometric test (adj. p-value)
FOSL1_0_A	128	597	1,20723E-42	7,58142E-40
FOSL2_0_A	137	726	4,86818E-39	1,52861E-36
JUNB_0_A	115	544	7,50317E-38	1,57066E-35
JUND_0_A	123	623	3,12176E-37	4,90116E-35
FOS_0_A	119	626	1,9905E-34	2,50007E-32
BACH2_0_A	67	273	1,21946E-26	1,27636E-24
JUN_0_A	94	524	1,38731E-25	1,24462E-23
NFE2_0_A	71	344	2,45514E-23	1,92728E-21
BACH1_0_A	87	494	3,08603E-23	2,15336E-21
FOSB_0_A	96	585	3,7038E-23	2,32598E-21
NF2L2_0_A	63	279	4,66643E-23	2,66411E-21
MAFB_0_B	33	147	9,6426E-13	5,04629E-11
ZN554_1_D	48	302	7,42587E-12	3,58726E-10
MAFF_1_B	30	134	1,0909E-11	4,89348E-10
MAFK_0_A	16	41	8,72968E-11	3,65483E-09
MAFK_1_A	39	239	2,919E-10	1,14571E-08
MAFF_0_B	17	56	2,13161E-09	7,8744E-08
NFIA_0_C	36	234	6,92026E-09	2,4144E-07
SOX10_0_B	15	47	8,67493E-09	2,86729E-07
NRL_0_D	36	239	1,22002E-08	3,83085E-07
MAF_0_A	19	81	2,76898E-08	8,28056E-07
SOX9_0_B	20	90	3,21318E-08	9,17217E-07
GCR_0_A	16	64	1,31334E-07	3,58599E-06
NFIC_1_A	25	146	1,61645E-07	4,2297E-06
NFIC_0_A	45	380	3,32432E-07	8,3507E-06
NFIB_0_D	27	189	2,16429E-06	5,2276E-05
BCL6_0_A	21	128	3,10189E-06	7,21477E-05
SOX4_0_B	24	164	5,06744E-06	0,000110867
FEZF1_0_C	21	132	5,11964E-06	0,000110867
NDF1_0_A	16	87	1,017E-05	0,000212892
SOX3_0_B	22	150	1,18536E-05	0,000240131
GCR_1_A	16	97	4,15416E-05	0,000790252
TEAD1_0_A	20	140	4,25137E-05	0,000790252
TEAD2_0_D	22	163	4,40171E-05	0,000790252
ANDR_1_A	13	68	4,40427E-05	0,000790252
NANOG_1_B	13	69	5,16858E-05	0,000898085
FOXO1_0_A	19	131	5,29127E-05	0,000898085
ATF4_0_A	14	81	7,2825E-05	0,00120353
PRDM4_0_D	18	124	8,13106E-05	0,001309309
TWST1_1_A	10	45	8,78213E-05	0,001378794
BARX1_0_D	10	46	0,000106986	0,00163872
ATF2_1_B	17	116	0,000112341	0,001679761
RARA_1_A	15	96	0,000134278	0,001961082
CEBPG_0_B	12	66	0,000142007	0,002026824
ZN708_1_D	15	97	0,000151242	0,002110664
NFAC2_0_B	11	58	0,000180341	0,002462044
NANOG_0_A	10	49	0,000186941	0,00247284
MYB_0_A	23	192	0,000189007	0,00247284
COT1_1_C	26	230	0,000193813	0,002483973
RUNX2_0_A	12	69	0,000219981	0,002762964
FOXO4_0_C	8	33	0,000234081	0,002882412
ZN317_0_C	9	42	0,000263988	0,003188165
IRF5_0_D	13	81	0,000280129	0,003319267
RUNX1_0_A	22	185	0,00028546	0,003319797
PRGR_0_A	8	34	0,00029216	0,003335933
MEF2B_0_A	12	72	0,000331739	0,003720213
NGN2_0_D	14	94	0,000369125	0,004066847
ZIC3_0_B	19	152	0,000383498	0,004152356
SOX2_0_A	20	165	0,000410434	0,00436869

BATF_0_A	15	107	0,000451804	0,004675901
IRF7_0_C	13	85	0,000454188	0,004675901
RUNX3_0_A	21	180	0,000500504	0,005069624
SMAD4_0_B	21	183	0,000622913	0,00620935
FOXC2_0_D	11	68	0,000752297	0,007356144
TEAD4_0_A	14	101	0,000773098	0,007356144
ZN350_0_C	14	101	0,000773098	0,007356144
PPARA_1_B	12	79	0,000789506	0,007400141
FOXK1_0_A	11	70	0,000965168	0,008779099
ETV6_0_D	17	139	0,000969707	0,008779099
TAL1_1_A	13	92	0,000978562	0,008779099
TLX1_0_D	21	190	0,001013377	0,00896339
NR4A1_0_A	13	94	0,001198712	0,010318537
PPARD_0_D	11	72	0,001225472	0,010318537
STA5A_0_A	17	142	0,001230693	0,010318537
COE1_0_A	12	83	0,001232309	0,010318537
RXRB_0_C	16	130	0,001259109	0,010404216
RFX2_0_A	46	551	0,001335851	0,010894989
HIC2_0_D	20	182	0,001412651	0,011373652
PEBB_0_C	16	132	0,001480275	0,011767253
ANDR_2_A	14	110	0,00179	0,013898333
HLF_0_C	8	44	0,001792619	0,013898333
ZN121_0_C	57	732	0,001975985	0,015133154
MEF2A_0_A	10	66	0,002144161	0,016223287
HXA9_0_B	6	27	0,002303846	0,017078211
OLIG2_1_B	13	101	0,002318392	0,017078211
SUH_0_A	25	258	0,002338736	0,017078211
IRF2_0_A	12	90	0,0024927	0,017993279
REST_0_A	17	152	0,002568478	0,01832959
HAND1_0_D	18	168	0,003129407	0,022081656
PPARG_1_A	11	81	0,003205135	0,022364717
SOX2_1_A	10	70	0,003339456	0,023045918
FOXQ1_0_C	8	49	0,003621519	0,024720804
PO5F1_0_A	7	39	0,00369344	0,024940652
ZN410_0_D	5	21	0,003887004	0,025968495
SMCA1_0_C	12	96	0,004269769	0,028225423
ZN449_0_C	21	215	0,004534445	0,029662831
RFX4_0_D	21	216	0,004781415	0,030955963
BATF_1_A	15	136	0,005061398	0,032378869
HNF4G_0_B	11	86	0,005104312	0,032378869
NFAC3_0_B	11	87	0,005573419	0,035001068
HBP1_0_D	6	32	0,005644094	0,03509397
FOXP2_0_C	9	64	0,005824031	0,035509628
GRHL1_0_D	9	64	0,005824031	0,035509628
ZN586_0_C	17	166	0,006296596	0,038021754
MEF2C_0_A	9	65	0,00645353	0,038598253
ATOH1_0_B	13	114	0,006602376	0,039115962
ZBT7B_0_D	40	508	0,006980535	0,040969868
TBX20_0_D	14	128	0,007117786	0,04130754
RFX3_0_B	18	182	0,007191602	0,04130754
SMAD3_0_B	20	210	0,007235397	0,04130754
COT2_0_A	12	103	0,007493118	0,042393497
SRF_0_A	6	34	0,007665225	0,04298001
HXD10_0_D	5	25	0,008534313	0,047409148
STAT6_0_B	12	105	0,008698034	0,047409148
RFX1_1_B	34	422	0,008724606	0,047409148
ASCL2_0_D	13	118	0,008757104	0,047409148
NFAC1_1_B	7	46	0,009364837	0,049421156
ZN490_0_C	7	46	0,009364837	0,049421156
ZN713_0_D	7	46	0,009364837	0,049421156
STF1_0_B	13	120	0,010023334	0,05160316
NR4A3_0_D	12	107	0,010048778	0,05160316
MYNN_0_D	6	36	0,010160959	0,05160316
NR1H3_1_B	8	58	0,010186785	0,05160316

RFX2_1_A	42	552	0,010189159	0,05160316
ETV4_0_B	20	219	0,011217688	0,056357664
FOXA1_0_A	9	71	0,011390117	0,056769791
FOXD3_0_D	11	96	0,011486229	0,056798045
ZIM3_0_C	10	84	0,012069486	0,058974492
P73_0_A	15	150	0,012114187	0,058974492
NR6A1_0_B	4	18	0,012566557	0,060706137
SOX18_0_D	10	85	0,01305649	0,06259142
MEF2D_0_A	8	61	0,013648712	0,064934782
HNF1A_0_C	5	28	0,013873201	0,065506542
RARB_0_D	17	181	0,014241294	0,066742782
PRGR_1_A	9	74	0,014720258	0,068139385
NKX28_0_C	6	39	0,014910282	0,068139385
FOXD1_0_D	8	62	0,014973304	0,068139385
IRF1_0_A	8	62	0,014973304	0,068139385
RORA_0_C	10	87	0,015210546	0,068503473
DBP_0_B	4	19	0,015271475	0,068503473
HXC9_0_C	5	29	0,016059022	0,07102159
PITX2_0_D	5	29	0,016059022	0,07102159
MCR_0_D	7	51	0,016180053	0,071056455
REL_0_B	11	101	0,016374744	0,07141208
PO4F2_0_D	3	11	0,01707325	0,073944834
COT2_1_A	15	157	0,017748611	0,076343339
FOXF1_0_D	6	41	0,018818968	0,080291139
SOX17_0_C	10	90	0,018922116	0,080291139
GCM1_0_D	10	91	0,020295388	0,085540294
ERR2_0_A	5	31	0,021090831	0,08771551
NDF2_0_B	5	31	0,021090831	0,08771551
TBX2_0_D	12	119	0,021838881	0,09022906
NR4A2_0_C	8	67	0,023019451	0,093871527
TEAD3_0_D	8	67	0,023019451	0,093871527
ERR1_0_A	7	55	0,023715824	0,095422536
BRAC_1_B	5	32	0,023951338	0,095422536
OZF_0_C	5	32	0,023951338	0,095422536
OLIG2_0_B	15	163	0,024007581	0,095422536
BATF3_0_B	8	68	0,024934331	0,098482766
STAT3_0_A	18	208	0,025217746	0,098718059
HXC6_0_D	4	22	0,025465487	0,098718059
HXD11_0_D	4	22	0,025465487	0,098718059
FOXA2_0_A	8	69	0,026957654	0,102333787
FOXA3_0_B	8	69	0,026957654	0,102333787
MAFG_0_A	5	33	0,027050014	0,102333787
NR2E3_0_C	5	33	0,027050014	0,102333787
ZNF41_0_C	12	124	0,028987345	0,109006304
RORG_0_C	11	111	0,030578883	0,112961992
TF7L1_0_B	11	111	0,030578883	0,112961992
ZN680_0_C	11	111	0,030578883	0,112961992
AP2A_0_A	19	228	0,030822386	0,113195664
LYL1_0_A	15	169	0,03178643	0,116057431
TBX19_0_D	3	14	0,03351888	0,121675472
TF7L2_0_A	7	60	0,036136582	0,129853542
NFAC1_0_B	8	73	0,036185302	0,129853542
ISL2_0_D	5	36	0,037824714	0,134202941
TEF_0_D	5	36	0,037824714	0,134202941
THB_1_D	8	74	0,038787126	0,135866646
DDIT3_0_D	4	25	0,038942669	0,135866646
ERR3_0_B	4	25	0,038942669	0,135866646
ZIC2_0_D	15	174	0,039558843	0,137253888
NFKB2_0_B	9	88	0,040097606	0,138358771
MEIS3_0_D	6	49	0,041273849	0,140869441
TWST1_0_A	6	49	0,041273849	0,140869441
ZBT18_0_C	12	131	0,041636392	0,141338672
ESR1_0_A	13	146	0,042475072	0,143410459
P63_0_A	13	147	0,044461343	0,149314028

SIX2_0_A	11	119	0,047019679	0,157065737
FOXO3_0_B	7	64	0,048701264	0,160970495
ZNF85_1_C	7	64	0,048701264	0,160970495
RFX5_1_A	23	305	0,050549813	0,166205669
CEBPE_0_A	5	39	0,050888471	0,166447706
STAT1_0_A	19	244	0,054518816	0,177398014
NKX61_0_B	3	17	0,055785494	0,178741876
SOX1_0_D	3	17	0,055785494	0,178741876
HMX3_0_D	4	28	0,055785681	0,178741876
PTF1A_1_B	12	138	0,057696994	0,183927474
SNAI2_0_A	8	81	0,060502672	0,191897365
FOXJ2_0_C	4	29	0,062141946	0,196106243
SNAI1_0_C	4	30	0,068861883	0,216226314
NR1H4_0_B	14	173	0,069620421	0,21752052
ZN549_0_C	11	128	0,071712594	0,222948064
IKZF1_0_C	16	206	0,074082897	0,229182559
MEIS1_0_A	4	31	0,075939076	0,233773235
NR5A2_0_B	7	71	0,076667607	0,23486467
FOXF2_0_D	5	44	0,077825307	0,237253849
P73_1_A	13	161	0,079270387	0,240491801
HSF1_1_A	6	58	0,081047634	0,24353069
ZN384_0_C	6	58	0,081047634	0,24353069
NFIL3_0_D	3	20	0,083410543	0,249437244
FOXC1_0_C	7	73	0,086078813	0,256196656
HIC1_0_C	9	103	0,089496888	0,265113423
HNF4A_0_A	7	74	0,091019735	0,266194899
THA_1_D	7	74	0,091019735	0,266194899
ISL1_0_A	4	33	0,091133604	0,266194899
P53_0_A	8	89	0,093134824	0,26921403
SIX1_0_A	10	119	0,093213324	0,26921403
TFE2_0_A	9	104	0,093686502	0,26921403
IRF4_0_A	12	150	0,093881963	0,26921403
ZN341_1_C	13	166	0,095020325	0,271239836
RARG_1_B	10	120	0,097193775	0,276188645
STAT4_0_A	16	216	0,101612408	0,287444109
GLIS2_0_D	8	91	0,102605993	0,287663231
NKX21_0_A	8	91	0,102605993	0,287663231
ZNF8_0_C	3	22	0,104466667	0,291578076
ZN418_0_C	13	169	0,105327884	0,29172777
SOX10_1_A	10	122	0,105449369	0,29172777
DLX5_0_D	2	11	0,109493302	0,301586813
ZNF18_0_C	7	78	0,112325162	0,308035816
ATF2_0_B	21	303	0,11509805	0,313426743
TF2LX_0_D	3	23	0,115698777	0,313426743
AP2C_0_A	22	320	0,115788224	0,313426743
PPARA_0_B	10	125	0,118562555	0,319559162
MEIS1_1_B	11	141	0,119766168	0,320055972
ZN708_0_C	11	141	0,119766168	0,320055972
TFCP2_0_D	18	256	0,123964067	0,329870484
ZN250_0_C	5	51	0,125932562	0,333694722
CDC5L_0_D	2	12	0,127013191	0,333741775
NKX62_0_D	2	12	0,127013191	0,333741775
CEBPB_0_A	8	96	0,128516326	0,336284388
ZN418_1_D	9	112	0,131196032	0,341871818
BC11A_0_A	10	128	0,132535463	0,343935004
PBX3_1_A	21	311	0,138505638	0,357948727
ASCL1_0_A	9	114	0,141648331	0,363934538
PAX8_0_D	16	228	0,141980831	0,363934538
SMAD1_0_D	6	68	0,143354932	0,365962996
MIXL1_0_D	2	13	0,145128883	0,367503785
TBX4_0_D	2	13	0,145128883	0,367503785
PBX1_0_A	14	196	0,146528726	0,369558394
GSX1_0_D	1	3	0,148486591	0,371512268
LHX4_0_D	1	3	0,148486591	0,371512268

GLIS3_0_D	6	69	0,150542258	0,373678015
ZKSC3_0_D	6	69	0,150542258	0,373678015
ZN322_0_B	23	350	0,153600637	0,379768503
ZN528_0_C	9	117	0,158089954	0,389335259
ESR1_1_A	11	150	0,161476267	0,396121467
GLI1_0_D	10	134	0,162955315	0,397120409
HNF1B_1_A	2	14	0,163730551	0,397120409
MNX1_0_D	4	41	0,164212298	0,397120409
BHE40_0_A	15	217	0,164412908	0,397120409
E2F8_0_D	13	184	0,166287412	0,399964481
FOXP1_0_A	5	56	0,166864162	0,399964481
STAT2_0_A	6	72	0,173028653	0,413163475
NKX31_0_C	4	42	0,174520981	0,415148394
RARG_0_B	8	104	0,176177946	0,415938909
TBX21_0_A	8	104	0,176177946	0,415938909
HNF1B_0_A	3	28	0,177649097	0,417841323
ITF2_0_C	8	105	0,182626553	0,426568885
OVOL1_0_C	2	15	0,182718201	0,426568885
MYBB_0_D	4	43	0,18504361	0,43039773
RXRA_1_A	9	122	0,187409088	0,434291171
COT1_0_C	13	189	0,189888058	0,438418016
NKX22_0_D	11	156	0,19279242	0,441979884
VSX2_0_D	1	4	0,192910571	0,441979884
HTF4_0_A	9	123	0,193542147	0,441979884
ZN554_0_C	53	898	0,195652751	0,445180896
STAT1_1_A	6	75	0,196792431	0,44615757
RX_0_D	2	16	0,20200095	0,455325187
MYOG_0_B	10	141	0,20228619	0,455325187
FIGLA_0_D	3	30	0,204545175	0,457132988
FOXH1_0_A	3	30	0,204545175	0,457132988
NR1H4_1_B	4	45	0,206671649	0,458621185
OTX2_0_A	4	45	0,206671649	0,458621185
ZBT48_0_C	12	176	0,209880914	0,464102865
NR1D1_0_B	10	144	0,220273757	0,485375156
ETV3_0_D	5	62	0,221747143	0,486913307
ETV2_0_B	27	442	0,226025886	0,49457929
BCL6B_0_D	4	47	0,228977864	0,496938758
SRBP1_0_A	8	112	0,23040534	0,496938758
BRAC_0_A	5	63	0,231386072	0,496938758
LEF1_0_A	5	63	0,231386072	0,496938758
NR1H2_0_D	3	32	0,232285624	0,496938758
ZEP1_0_D	11	163	0,232435487	0,496938758
GSX2_0_D	1	5	0,235017215	0,496938758
HOMEZ_0_D	1	5	0,235017215	0,496938758
NOTO_0_D	1	5	0,235017215	0,496938758
PO6F2_0_D	1	5	0,235017215	0,496938758
HXC8_0_D	4	48	0,240348342	0,501422082
MAX_0_A	4	48	0,240348342	0,501422082
NR2C1_0_C	4	48	0,240348342	0,501422082
HXD3_0_D	2	18	0,241129727	0,501422082
HXD8_0_D	2	18	0,241129727	0,501422082
STA5B_0_A	7	97	0,242847284	0,502401684
NR2F6_0_D	8	114	0,244809283	0,502401684
BMAL1_0_A	9	131	0,245443477	0,502401684
CUX1_0_C	3	33	0,246400826	0,502401684
NR2E1_0_D	3	33	0,246400826	0,502401684
SOX8_0_D	3	33	0,246400826	0,502401684
IRX2_0_D	16	253	0,249636111	0,50735106
RFX5_0_A	17	271	0,251505876	0,509502226
PRDM6_0_C	3	34	0,260642927	0,525011324
NKX32_0_C	2	19	0,260833651	0,525011324
PRD14_0_A	3	35	0,274985845	0,551728787
ZN768_0_C	10	153	0,27761481	0,554036866
BARH2_0_D	2	20	0,28054733	0,554036866

HXC12_0_D	2	20	0,28054733	0,554036866
LHX2_0_A	2	20	0,28054733	0,554036866
PITX3_0_D	2	20	0,28054733	0,554036866
HLTF_0_D	9	137	0,287127056	0,565253264
VENTX_0_D	2	21	0,300216125	0,587090088
ZN547_0_C	9	139	0,301427563	0,587090088
RHXF1_0_D	5	70	0,301620416	0,587090088
ZN331_0_C	12	192	0,301958755	0,587090088
CR3L1_0_D	5	71	0,311941346	0,595183493
EOMES_0_D	5	71	0,311941346	0,595183493
VDR_0_A	5	71	0,311941346	0,595183493
FOXD2_0_D	1	7	0,312755657	0,595183493
HMX1_0_D	1	7	0,312755657	0,595183493
ONEC3_0_D	1	7	0,312755657	0,595183493
PO3F3_0_D	1	7	0,312755657	0,595183493
DPRX_0_D	3	38	0,318374498	0,60128381
GSC_0_D	2	22	0,319791071	0,60128381
HMX2_0_D	2	22	0,319791071	0,60128381
PIT1_0_C	2	22	0,319791071	0,60128381
IRF9_0_C	4	55	0,322663162	0,604873033
HSF1_0_A	3	39	0,332881126	0,622170675
GATA1_0_A	4	56	0,334647701	0,623616487
P63_1_A	6	91	0,338561176	0,626574868
HES7_0_D	12	198	0,339041659	0,626574868
PO3F1_0_C	2	23	0,339228432	0,626574868
IRF8_0_B	7	109	0,342060974	0,629953934
ZFP28_0_C	3	40	0,347374285	0,63786857
MSX1_0_D	1	8	0,348610607	0,638272482
SOX9_1_B	15	255	0,354083379	0,64640803
NR1H3_0_B	6	93	0,357289198	0,650369902
RARA_2_A	5	76	0,364065109	0,658884404
TF65_0_A	5	76	0,364065109	0,658884404
CEBPA_0_A	6	94	0,366679892	0,661709691
TCF7_0_A	4	59	0,370661186	0,666977722
ZN524_0_D	11	185	0,372191819	0,667818464
TBX3_0_C	10	168	0,38069857	0,680650888
PO4F3_0_D	1	9	0,382595165	0,680650888
PO5F1_1_A	1	9	0,382595165	0,680650888
ZN589_0_D	8	132	0,383854556	0,68096232
LHX6_0_D	13	224	0,387138184	0,6848529
ARNT_0_B	8	133	0,391838077	0,689832818
ELF3_0_A	13	225	0,393248645	0,689832818
THA_0_C	13	225	0,393248645	0,689832818
ELK3_0_D	23	411	0,395058819	0,691077823
KAISO_2_A	16	283	0,407038875	0,710056704
SPI1_0_A	9	154	0,412308328	0,715795903
BARX2_0_D	1	10	0,414806888	0,715795903
PAX4_0_D	1	10	0,414806888	0,715795903
HXC13_0_D	2	27	0,414888071	0,715795903
PITX1_0_D	4	63	0,418401462	0,719879775
KLF8_0_C	6	100	0,423048072	0,723907872
ZN816_1_C	6	100	0,423048072	0,723907872
TGIF2_0_D	2	28	0,433137223	0,739158087
MGAP_0_D	1	11	0,445338247	0,757919835
ESR2_0_A	14	253	0,449392053	0,760695981
PKNX1_0_B	14	253	0,449392053	0,760695981
AIRE_0_C	2	29	0,451075029	0,761492253
SRBP2_0_B	10	179	0,458274472	0,771110088
ATF3_0_A	11	198	0,459227983	0,771110088
ZN382_0_C	4	67	0,465243749	0,779128198
HXA7_0_D	2	30	0,468684304	0,782802508
RELB_0_C	8	143	0,471361793	0,785186223
PAX2_0_D	3	49	0,474341547	0,788059502
ZN667_0_C	16	296	0,478348735	0,792620067

GATA5_0_D	2	31	0,485950504	0,80001178
ZBT49_0_D	2	31	0,485950504	0,80001178
GSC2_0_D	9	164	0,486631369	0,80001178
PRDM1_0_A	9	166	0,501256636	0,816246886
P5F1B_0_D	1	13	0,501705889	0,816246886
PDX1_0_A	1	13	0,501705889	0,816246886
VAX2_0_D	1	13	0,501705889	0,816246886
HES5_0_D	6	109	0,505877939	0,82090787
EPAS1_0_B	5	90	0,508000681	0,822227906
SMAD2_0_A	8	148	0,510311574	0,823844906
MYOD1_0_A	10	187	0,513810036	0,827365904
SPZ1_0_D	12	226	0,516757928	0,829984601
PTF1A_0_B	17	323	0,520139304	0,831164078
Z324A_0_C	17	323	0,520139304	0,831164078
ZN282_0_D	4	72	0,521738796	0,831603969
ZBTB6_0_C	16	305	0,527008517	0,83265856
HXD13_0_D	1	14	0,527703992	0,83265856
LBX2_0_D	1	14	0,527703992	0,83265856
ONEC2_0_D	1	14	0,527703992	0,83265856
ID4_0_D	4	73	0,53269688	0,838430177
FOXO6_0_D	2	34	0,535580122	0,840860792
BHA15_0_B	8	152	0,540777463	0,846903359
CR3L2_0_D	7	133	0,544945452	0,849195393
SPIB_0_A	7	133	0,544945452	0,849195393
ZIC1_0_B	8	153	0,54827832	0,852268265
GMEB2_0_D	2	35	0,551373715	0,852268265
FOXB1_0_D	1	15	0,552345834	0,852268265
HXA10_0_C	1	15	0,552345834	0,852268265
CRX_0_B	2	36	0,566783666	0,870829693
MLXPL_0_D	6	116	0,567148638	0,870829693
PLAL1_0_D	4	77	0,575218367	0,875401831
ZN329_0_C	6	117	0,575600518	0,875401831
GATA4_0_A	1	16	0,575702159	0,875401831
GATA6_0_A	1	16	0,575702159	0,875401831
BPTF_0_D	2	37	0,581807016	0,880729532
CDX2_0_A	20	394	0,582010758	0,880729532
NR1I3_0_C	4	78	0,585500093	0,883879948
THB_0_C	8	159	0,592157922	0,891786991
GFI1B_0_A	4	79	0,595635234	0,893913177
HSF4_0_D	1	17	0,597840023	0,893913177
PO4F1_0_D	1	17	0,597840023	0,893913177
RXRG_0_B	3	59	0,600412321	0,89562693
RARG_2_D	15	300	0,603787524	0,898527405
CEBPD_0_C	5	101	0,611105395	0,904810379
BHE23_0_D	3	60	0,611873637	0,904810379
ZN436_0_C	11	222	0,612331865	0,904810379
ZFX_0_A	22	442	0,619905879	0,91385186
EHF_0_B	20	403	0,622471931	0,91428526
BRCA1_0_D	3	61	0,62311161	0,91428526
MYOD1_1_A	6	124	0,632311229	0,9256211
HXD9_0_D	1	19	0,638711278	0,930651235
TBR1_0_D	1	19	0,638711278	0,930651235
FLI1_1_A	12	248	0,646032049	0,939139183
FOXM1_0_A	3	64	0,655469117	0,948466834
P53_1_A	3	64	0,655469117	0,948466834
GATA3_0_A	1	20	0,657562012	0,949307916
PAX1_0_D	3	65	0,665799695	0,957581381
ATF6A_0_B	10	211	0,666342458	0,957581381
WT1_1_B	24	494	0,66869879	0,958773607
PATZ1_1_C	14	294	0,674279812	0,964021828
ZSC16_0_D	1	21	0,675429306	0,964021828
ZN134_0_C	14	295	0,67909757	0,967059578
MYC_0_A	3	67	0,685776761	0,973273151
GATA2_0_A	2	45	0,688110317	0,973273151

ZEB1_0_A	2	45	0,688110317	0,973273151
PPARG_0_A	7	153	0,691071289	0,973529433
HEY1_0_D	8	174	0,691665879	0,973529433
TFAP4_0_A	5	111	0,692942128	0,973529433
TAL1_0_A	3	68	0,695424168	0,974835663
NKX23_0_D	2	46	0,699704731	0,978651605
ZBT7A_0_A	4	91	0,705040644	0,983923388
HXD12_0_D	1	23	0,708416098	0,986441928
CLOCK_0_C	6	138	0,731250174	1
ZNF41_1_C	3	72	0,731765364	1
HMBX1_0_D	2	49	0,732348437	1
ZN350_1_D	9	206	0,751939448	1
ZN320_0_C	7	164	0,757044042	1
ERG_0_A	17	377	0,763876137	1
ARI5B_0_C	1	27	0,764673928	1
NR1D1_1_D	1	27	0,764673928	1
ZSCA4_0_D	1	27	0,764673928	1
SPDEF_0_D	12	274	0,771231594	1
XBP1_0_D	18	402	0,778457908	1
MZF1_0_B	3	78	0,779745803	1
ZN563_0_C	17	382	0,781525288	1
ZN423_0_D	1	29	0,788591566	1
SCRT2_0_D	2	55	0,788612613	1
GATA1_1_A	1	30	0,799622851	1
RFX1_0_B	31	679	0,800934328	1
CREM_0_C	11	260	0,800954028	1
NR1I2_0_C	3	82	0,807627435	1
ARNT2_0_D	24	537	0,808335533	1
HXA13_0_C	1	31	0,810078596	1
ZN652_0_D	1	31	0,810078596	1
HINFP_0_C	7	175	0,81221623	1
MAFA_0_D	12	285	0,813714271	1
RXRA_0_A	8	198	0,81541443	1
BHE41_0_D	12	286	0,817267095	1
GCM2_0_D	1	32	0,819988825	1
PO2F2_0_A	1	32	0,819988825	1
ZEP2_0_D	13	310	0,82537132	1
MYBA_0_D	17	396	0,826019434	1
GLI3_0_B	3	85	0,826530414	1
ZN136_0_C	1	33	0,829381996	1
CTCF_0_A	11	268	0,830400641	1
GLIS1_0_D	2	61	0,834174097	1
GLI2_0_D	2	62	0,840842422	1
ZN816_0_C	4	113	0,846517137	1
NFAT5_0_D	8	206	0,846927629	1
ZN140_0_C	3	89	0,84923924	1
EGR3_0_D	5	140	0,860073476	1
GRHL2_0_A	3	94	0,873946176	1
ZBTB4_0_D	14	349	0,874587782	1
HXB13_0_A	1	40	0,882754018	1
HEN1_0_C	12	309	0,885598788	1
SCRT1_0_D	1	41	0,888872373	1
USF1_0_A	13	334	0,891681436	1
TFEB_0_C	6	173	0,892364171	1
TBP_0_A	1	42	0,894671488	1
USF2_0_A	13	340	0,904589032	1
GFI1_0_C	2	75	0,907852212	1
VDR_1_A	2	76	0,911723967	1
RARA_0_A	8	230	0,916369412	1
ZN134_1_C	4	132	0,917887063	1
ZKSC1_0_B	1	47	0,919431921	1
PAX6_0_C	2	80	0,92573348	1
FOXI1_0_B	18	468	0,930566776	1
SALL4_0_B	4	137	0,93088629	1

MXI1_1_A	2	85	0,940300723	1
ZN563_1_C	9	268	0,942339732	1
INSM1_0_C	8	244	0,942780187	1
NR0B1_0_D	5	169	0,943322189	1
KAISO_1_A	2	87	0,945330272	1
TFE3_0_B	8	247	0,94737352	1
MTF1_0_C	8	249	0,950251104	1
ETV5_0_C	4	148	0,953160413	1
CEBPZ_0_D	31	780	0,954763636	1
NOBOX_0_C	1	58	0,955320733	1
MBD2_0_B	1	59	0,957652562	1
ZIC4_0_D	6	205	0,959051143	1
MXI1_0_A	5	180	0,960738066	1
NKX25_0_B	3	127	0,964324819	1
ZN335_0_A	15	433	0,966604755	1
HESX1_0_D	9	289	0,967584973	1
HES1_0_D	13	387	0,967836707	1
NR2C2_0_B	3	130	0,968373189	1
ETS1_0_A	27	712	0,968855938	1
ZSC31_0_C	15	438	0,970313677	1
ZFP42_0_A	17	486	0,970876284	1
CREB5_0_D	2	106	0,976691562	1
MLX_0_D	3	141	0,979797384	1
ATF7_0_D	2	111	0,981454518	1
MYCN_0_A	5	202	0,981772801	1
CREB3_0_D	2	113	0,98308228	1
ZN257_0_C	5	205	0,983634069	1
PAX5_0_A	104	2432	0,986189681	1
RREB1_0_D	1	81	0,986978839	1
ATF1_0_B	5	213	0,987760368	1
KAISO_0_A	4	184	0,987883116	1
ZN264_0_C	1	84	0,988913329	1
CTCFL_0_A	14	455	0,989522613	1
JDP2_0_D	3	157	0,989643398	1
E2F5_0_B	1	86	0,990040566	1
ETS2_0_B	1	87	0,990560454	1
CREB1_0_A	2	126	0,990738814	1
PBX3_0_A	9	334	0,991541541	1
HIF1A_0_C	1	90	0,991962896	1
ATF2_2_C	2	131	0,992671135	1
OSR2_0_C	18	582	0,994907756	1
ZNF85_0_C	1	100	0,995298348	1
SMCA5_0_C	2	142	0,995637237	1
E4F1_0_D	1	102	0,99577644	1
ZN770_1_C	31	912	0,996513134	1
ZF64A_0_D	12	446	0,996867862	1
MITF_0_A	6	286	0,99756851	1
E2F2_0_B	2	156	0,997760844	1
ZN214_0_C	1	120	0,998391198	1
E2F6_0_A	3	201	0,998473443	1
Z354A_0_C	1	121	0,998475198	1
ZNF76_0_C	37	1100	0,998712173	1
SPIC_0_D	2	176	0,999146068	1
ZN260_0_C	2	177	0,999186497	1
ELF5_0_A	5	284	0,999217381	1
ZBTB4_1_D	2	179	0,999261762	1
E2F4_1_A	8	380	0,999314514	1
ZN143_0_A	41	1234	0,999447915	1
FEV_0_B	14	558	0,999478884	1
NFYC_0_A	23	796	0,999500433	1
KLF13_0_D	18	685	0,999709813	1
ETV1_0_A	19	712	0,999713569	1
E2F3_0_A	21	779	0,999810367	1
AP2B_0_B	5	323	0,999843663	1

KLF5_0_A	37	1206	0,999888858	1
NFYA_0_A	25	908	0,999904797	1
THA11_0_B	33	1136	0,99994817	1
ELK1_0_B	16	692	0,999957881	1
NFYB_0_A	26	969	0,99996437	1
AP2D_0_D	129	3416	0,999974205	1
GABPA_0_A	23	901	0,99997479	1
ZN263_1_A	1	198	0,999975484	1
ELK4_0_A	13	627	0,99997977	1
TFDP1_0_C	2	270	0,999991826	1
ZFX_1_A	131	3530	0,99999189	1
ELF1_0_A	23	940	0,999992291	1
HEY2_0_D	24	970	0,999992893	1
PURA_0_D	3	319	0,999993593	1
E2F4_0_A	5	407	0,999995867	1
TYY1_0_A	13	696	0,999998328	1
ZSC22_0_C	53	1791	0,99999889	1
KLF4_0_A	70	2233	0,99999934	1
ZN219_0_D	15	806	0,999999718	1
ZBT17_0_A	16	842	0,999999767	1
EGR2_0_A	5	489	0,999999898	1
ELF2_0_C	20	994	0,999999924	1
TYY2_0_D	4	456	0,99999993	1
TAF1_0_A	7	603	0,99999998	1
ZN263_0_A	8	654	0,999999989	1
KLF14_0_D	32	1401	0,99999999	1
KLF6_0_A	23	1150	0,999999993	1
KLF12_0_C	19	1030	0,999999993	1
ZN281_0_A	13	840	0,999999993	1
EGR1_0_A	14	881	0,999999995	1
KLF9_0_C	16	999	0,999999999	1
EGR4_0_D	4	578	1	1
E2F7_0_B	7	721	1	1
MAZ_1_A	7	728	1	1
PLAG1_0_D	1	450	1	1
SP4_0_A	21	1238	1	1
ZN467_0_C	3	583	1	1
VEZF1_0_C	4	642	1	1
EGR2_1_A	6	733	1	1
PROX1_0_D	4	650	1	1
TBX1_0_D	5	697	1	1
KLF15_0_A	9	873	1	1
KLF1_0_A	20	1264	1	1
KLF3_0_B	24	1408	1	1
SP4_1_A	3	637	1	1
SP1_1_A	7	833	1	1
ZN148_0_D	3	657	1	1
ZN770_0_C	90	3318	1	1
KLF16_0_D	4	714	1	1
TBX15_0_D	9	952	1	1
ZN341_0_C	1	566	1	1
PATZ1_0_C	2	639	1	1
MAZ_0_A	4	804	1	1
SP2_1_B	44	2207	1	1
WT1_0_C	2	748	1	1
SP1_0_A	5	981	1	1
SP3_0_B	2	784	1	1

Supplementary Table 2. **Significant differences in histone H3K27ac ChIP-seq signal between GBMs and DAs.** c-Jun transcription factor (TF) binding prediction in glioblastoma LN18 and LN229 cell lines (DESeq2 methods, padj < 0.05) in the context of glioma enhancers. Width represents the size of the enhancers; baseMean represents the mean of normalized counts of all samples normalized by the sequencing depth; log2FC stand for log2 Fold Change.

Chromosome	Start	End	Width	BaseMean	Log2FC	P-value	Adj. p-value
chr5	139696965	139696975	2523	145,7748534	0,97444	0,001004323	0,045145502
chr7	2758046	2758056	2280	135,6851289	1,06061	5,0239E-05	0,010542645
chr7	2758077	2758087	2280	135,6851289	1,06061	5,0239E-05	0,010542645

Supplementary Table 3. **Prediction of c-Jun binding instances in glioma enhancers**. BMO scores correspond to the significance of a motif prediction (-log10 adjusted p-value of a motif to be bound in a specific open-chromatin region). Targeted chromosome conformation capture (Hi-C), if any, are shown as well as DNA methylation information from the TCGA data.

Chromosome	Start	End	BMO score LN18	BMO score LN229	Targeted gene (Hi-C)	450k information
chr1	19011183	19011193	1,904704368	1,40475917	-	no
chr1	205284858	205284868	1,904704368	2,829492378	CDK18	no
chr1	209747897	209747907	1,904704368	1,531739163	-	yes
chr1	223726881	223726891	1,646572432	1,777844606	-	no
chr1	230104846	230104856	2,787223126	3,181933755	-	no
chr1	244274002	244274012	2,03330937	2,026325852	ZBTB18	no
chr1	31701225	31701235	1,387345357	2,611794724	-	no
chr1	3313298	3313308	1,387345357	4,250343025	PRDM16	no
chr1	36373958	36373968	1,904704368	3,713547308	-	no
chr1	37471873	37471883	2,03330937	2,15326143	-	no
chr1	39192312	39192322	2,787223126	3,713547308	-	no
chr1	68384507	68384517	4,944140778	1,531739163	-	no
chr1	88327824	88327834	1,769092551	3,713547308	-	no
chr10	116947298	116947308	4,158371778	2,651162477	-	no
chr10	116947319	116947329	4,027636422	2,899381334	-	no
chr10	6275483	6275493	1,387345357	2,026325852	-	no
chr11	120548105	120548115	1,769092551	2,15326143	-	no
chr11	129143526	129143536	4,691052386	3,394364815	-	no
chr11	20112392	20112402	3,527439214	3,394364815	-	no
chr11	6357388	6357398	1,646572432	2,15326143	-	no
chr11	9364528	9364538	2,03330937	2,501984762	-	no
chr12	20212883	20212893	2,168977805	5,26176774	-	no
chr12	75978050	75978060	3,527439214	2,269732217	-	no
chr14	76918968	76918978	2,667700422	1,898783406	IRF2BPL, LINC02288	no
chr15	87477930	87477940	1,516284635	1,654098339	-	no
chr16	23423619	23423629	1,516284635	1,898783406	SCNN1G	yes
chr16	48123282	48123292	5,196429652	3,624482619	NETO2, ITFG1-AS1	no
chr16	48123305	48123315	5,196429652	3,931520317	-	no
chr16	9047972	9047982	3,101542182	4,188570246	-	yes

chr16	9048017	9048027	2,80066288	3,504784361	-	yes
chr17	16286484	16286494	1,387345357	1,40475917	-	no
chr17	44138372	44138382	2,03330937	2,026325852	-	no
chr17	78862177	78862187	2,168977805	4,633093625	-	yes
chr17	78862427	78862437	1,646572432	4,940725081	-	no
chr17	80833259	80833269	2,03330937	4,36349031	-	no
chr17	82217823	82217833	1,387345357	1,654098339	-	no
chr18	9017284	9017294	3,181847607	3,059958736	-	no
chr19	2169036	2169046	1,516284635	1,777844606	-	no
chr2	11830409	11830419	2,297108203	3,713547308	TRIB2	yes
chr2	12177165	12177175	1,904704368	2,393142332	-	no
chr2	200406052	200406062	2,297108203	2,393142332	-	no
chr2	205025261	205025271	1,387345357	2,15326143	-	no
chr2	20579473	20579483	2,787223126	7,436162975	-	no
chr2	223613229	223613239	2,297108203	2,15326143	-	no
chr2	233974672	233974682	1,904704368	2,93704733	-	yes
chr2	64831516	64831526	3,662392041	2,93704733	-	no
chr2	84920038	84920048	1,904704368	3,181933755	DNAH6, TCF7L1	no
chr20	19993346	19993356	2,168977805	2,501984762	ARL4C	no
chr20	25237968	25237978	2,419224739	6,263158077	FAM182B	no
chr20	51366407	51366417	2,667700422	3,282108363	-	no
chr20	51750327	51750337	3,527439214	1,777844606	ZFP64, SPATA2	no
chr21	45364146	45364156	1,904704368	1,898783406	FTCD	no
chr21	45364276	45364286	1,387345357	2,611794724	-	no
chr22	32532670	32532680	1,387345357	2,715741145	-	no
chr22	49970917	49970924	2,156797248	2,529206338	-	no
chr22	49970968	49970978	2,667700422	3,624482619	-	no
chr3	109685488	109685488	1,646572432	2,269732217	-	no
chr3	11096707	11096717	3,415660408	3,282108363	-	no
chr3	15269525	15269535	1,387345357	1,654098339	EAF1, METTL6	yes
chr3	197394709	197394719	1,387345357	1,777844606	-	no
chr3	19889495	19889505	6,191413093	5,737966727	-	no
chr3	27534176	27534186	4,876221047	2,501984762	-	no
chr3	42080753	42080763	1,904704368	2,026325852	-	no
chr4	121711767	121711777	3,527439214	2,269732217	-	no
chr4	128387189	128387199	2,03330937	2,611794724	-	yes
chr4	145808464	145808474	3,499615518	1,625315008	-	no
chr4	145808514	145808524	3,499615518	1,625315008	-	no
chr4	25173682	25173692	1,646572432	2,393142332	-	no
chr5	139696965	139696975	1,387345357	1,777844606	CXXC5, PSD2-AS1, CXXC5-AS1	no
chr5	154682983	154682993	1,516284635	2,715741145	-	no
chr5	173764632	173764642	1,904704368	3,282108363	-	yes
chr5	42950825	42950835	3,288109134	1,654098339	ANXA2R	no
chr5	77893700	77893710	1,769092551	1,898783406	-	no
chr6	10316456	10316466	3,916909957	3,624482619	-	no
chr6	129501546	129501556	1,769092551	1,40475917	-	no
chr6	4358755	4358765	3,181847607	3,624482619	-	no

chr6	52517739	52517749	2,03330937	2,15326143	-	no
chr6	90479417	90479427	1,646572432	1,898783406	MDN1, CASP8AP2	no
chr7	106065101	106065111	2,667700422	3,624482619	-	no
chr7	116272213	116272223	3,527439214	2,269732217	-	no
chr7	139681013	139681023	2,03330937	2,829492378	-	no
chr7	151754387	151754397	4,790033657	6,181280391	NUB1, RHEB	no
chr7	2758046	2758056	2,156797248	4,382462427	-	no
chr7	2758077	2758087	2,003491971	4,61049481	-	no
chr7	35003857	35003867	2,787223126	5,462107257	-	no
chr7	37032162	37032172	2,297108203	4,633093625	-	no
chr7	47253768	47253778	1,769092551	2,715741145	IGFBP3	no
chr7	76291698	76291708	2,297108203	1,654098339	-	no
chr8	110894561	110894571	5,170802091	3,394364815	-	no
chr8	129984484	129984494	1,904704368	1,531739163	-	no
chr8	131856985	131856995	4,691052386	4,737447223	ADCY8	no
chr9	96339884	96339894	1,395563439	2,278427627	-	no
chrX	103152371	103152381	2,54482235	3,167943304	-	no
chrX	103152434	103152442	1,851718478	2,278427627	-	no

5.2. Published articles and manuscripts in preparation

Parts of the work presented here have been published or are in the process of being published:

- **Adria-Jaume Roura**, Bartlomiej Gielniewski, Paulina Pilanc, Paulina Szadkowska, Marta Maleszewska, Sylwia K. Krol, Ryszard Czepko, Wojciech Kaspera, Bartosz Wojtas, and Bozena Kaminska. "Identification of the immune gene expression signature associated with recurrence of High-Grade Gliomas." *Journal of Molecular Medicine* 99, no. 2 (2020): 241-55. doi:10.1007/s00109-020-02005-7.

- **Adria-Jaume Roura**, Paulina Szadkowska, Michal J. Dabrowski, Karolina Stepniak, Bartosz Wojtas, and Bozena Kaminska. "The oncogenic transcription factor c-Jun regulates critical over-expressed genes in Glioblastoma and is widely involved in distal-regulatory glioma elements."

- Chinchu Jayaprakash, **Adria-Jaume Roura**. Bartosz Wojtas, Bartek Gielniewski, Paulina Szadkowska, Sylwia K. Krol. "Knockdown of *SMARCA4* and *SMARCA2*, subunits of the SWI/SNF chromatin remodeling complex, deregulates open chromatin and transcription profiles in human gliomas".

Other published works:

- Aleksandra Ellert-Miklaszewska, Natalia Ochocka, Marta Maleszewska, Ling Ding, Erik Laurini, Yifan Jiang, **Adria-Jaume Roura**, Suzanne Giorgio, Bartlomiej Gielniewski, Sabrina Pricl, Ling Peng, and Bozena Kaminska. "Efficient and Innocuous Delivery of Small Interfering RNA to Microglia Using an Amphiphilic Dendrimer Nanovector." *Nanomedicine* 14, no. 18 (2019): 2441-459. doi:10.2217/nnm-2019-0176.

- Ilona E. Grabowicz, Bartek Wilczyński, Bożena Kamińska, **Adria-Jaume Roura**, Bartosz Wojtaś, and Michał J. Dąbrowski. "The Role of Epigenetic Modifications, Long-range Contacts, Enhancers and Topologically Associating Domains in the Regulation of Glioma Grade-specific Genes." *Scientific Reports*11, no. 1 (2021). doi:10.1038/s41598-021-95009-3.

- Paulina Pilanc, Kamil Wojnicki, **Adria-Jaume Roura**, Salwador Cyranowski, Aleksandra Ellert-Miklaszewska, Natalia Ochocka, Bartłomiej Gielniewski, Marcin M. Grzybowski, Roman Błaszczyk, Paulina S. Stańczak, Paweł Dobrzański, and Bozena Kaminska. "A Novel Oral

Arginase 1/2 Inhibitor Enhances the Antitumor Effect of PD-1 Inhibition in Murine Experimental Gliomas by Altering the Immunosuppressive Environment." *Frontiers in Oncology*11 (2021). doi:10.3389/fonc.2021.703465.

Other works in preparation:

- Bartlomiej Gielniewski, Katarzyna Poleszak, **Adria-Jaume Roura**, Paulina Szadkowska, Sylwia K. Krol, Rafal Guzik, Paulina Wiechecka, Marta Maleszewska, Beata Kaza, Andrzej Marchel, Tomasz Czernicki, Andrzej Koziarski, Grzegorz Zielinski, Andrzej Styk, Maciej Kawecki, Cezary Szczylik, Ryszard Czepko, Mariusz Banach, Wojciech Kaspera, Wojciech Szopa, Mateusz Bujko, Bartosz Czapski, Miroslaw Zabek, Ewa Izycka-Swieszewska, Wojciech Kloc, Pawel Nauman, Joanna Cieslewicz, Bartosz Wojtas, and Bozena Kaminska. "The Novel, Recurrent Mutation in the TOP2A Gene Results in the Enhanced Topoisomerase Activity and Transcription Deregulation in Glioblastoma." 2020. doi:10.1101/2020.06.17.158477 (bioRxiv, under revision in PLOS Genetics).

- Maria Banqueri, **Adria-Jaume Roura**, Anna Kiryk, Marie-Eve Tremblay, Bozena Kaminska. “Transcriptomic responses of microglia to a chronic, unpredictable, mild stress in the prefrontal cortex and hippocampus in a murine model of depression.”

- Malgorzata Perycz, Marta Jardanowska, **Adria-Jaume Roura**, Bartlomiej Gielniewski, Karolina Stepniak, Michal J Dabrowski, Michal Draminski, Bozena Kaminska, Bartosz Wojtas. “REST transcription factor holds the balance between the invasion and cell differentiation in IDH-mutant and IDH-wild type gliomas”.