UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Intron 1-Mediated Regulation of EGFR Expression In EGFR-Dependent Malignancies

Permalink
https://escholarship.org/uc/item/7578j3r1

Author
Jameson, Nathan Meade

Publication Date
2019

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Intron 1-Mediated Regulation of *EGFR* Expression In EGFR-Dependent Malignancies

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Biomedical Sciences

by

Nathan M. Jameson

Committee in charge:

Professor Frank Furnari, Chair
Professor Xiang-Dong Fu
Professor Paul Mischel
Professor Bing Ren
Professor Robert Wechsler-Reya

2019

Copyright
Nathan M. Jameson, 2019
All rights reserved.

The dissertation of Nathan M. Jameson is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____

____

____

____

____

Chair

University of California, San Diego

2019

DEDICATION

To:

My mother, Wendy Jameson, and my father, Patric Jameson, who's support, encouragement, and love allowed me to be who and where I am today. To my wife, Rachel Jameson, and my daughter, Avery Jameson, who's love, smiles, and unwavering support helped me enjoy every day no matter the challenges.

EPIGRAPH

The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with eager feet,
Until it joins some larger way
Where many paths and errands meet.
And whither then? I cannot say
—J.R.R. Tolkien

TABLE OF CONTENTS

Signature Page . . . iii

Dedication . . . iv

Epigraph . . . v

Table of Contents . . . vi

List of Abbreviations . . . viii

List of Figures . . . xi

Acknowledgements . . . xii

Vita . . . xiv

Abstract of the Dissertation . . . xv

Chapter 1 Introduction to the oncogene EGFR and its implications in cancer . . . 1
- 1.1 EGFR Function . . . 1
- 1.2 EGFR in Cancer . . . 3
- 1.3 EGFR in Glioblastoma . . . 3
- 1.4 EGFR in Non-Small Cell Lung Cancer . . . 8
- 1.5 EGFR in Head and Neck Squamous Cell Carcinoma . . . 10
- 1.6 Summary . . . 12

Chapter 2 Introduction to epigenetics and its role in cancer . . . 13
- 2.1 Epigenetics . . . 13
- 2.2 The Nucleosome . . . 14
- 2.3 DNA Methylation . . . 14
- 2.4 Non-coding Regulatory Elements . . . 15
- 2.5 3D Genome Organization . . . 18
- 2.6 Alterations in Epigenetic Elements . . . 20
- 2.7 Summary . . . 24

Chapter 3 Identification of putative active enhancers in the first intron of *EGFR* . . . 26
- 3.1 Introduction . . . 26
- 3.2 Results . . . 29
- 3.3 Discussion . . . 41
- 3.4 Materials and Methods . . . 42
- 3.5 Acknowledgements . . . 46

Chapter 4 Genetic perturbation of *EGFR* intron 1 enhancers . . . 47
4.1 Introduction . . . 47
4.2 Results . . . 49
4.3 Summary . . . 60
4.4 Materials and Methods . . . 62
4.5 Acknowledgements . . . 65

Chapter 5 Identification and characterization of transcription factors critical for intron 1-mediated *EGFR* expression . . . 66
5.1 Introduction . . . 66
5.2 Results . . . 68
5.3 Summary . . . 86
5.4 Materials and Methods . . . 87
5.5 Acknowledgements . . . 90

Chapter 6 Conclusion . . . 91

Bibliography . . . 98

LIST OF ABBREVIATIONS

3C	Chromosome Conformation Capture
4C	Circular Chromosome Conformation Capture
5-aza	5-Azacytidine
5hmC	5-Hydroxymethylcytosine
5mC	5-Methylcytosine
a-KG	Alpha-Ketoglutarate
ACH	Active Chromatin Hub
aKD	Asymmetric Kinase Domain
AML	Acute Myeloid Leukemia
AP-1	Activating Protein 1
ATAC-Seq	Transposase-Accessible Chromatin with Sequencing
ATP	Adenosine Triphosphate
BAC	Bacterial Artificial Chromosome
BET	Bromodomain and Extraterminal Domain
BRAF	v-Raf Murine Viral Oncogenes Homolog B1
BRD4	Bromodomain-Containing Protein 4
Cas9	*Streptococcus pyogenes* Cas9 (spCas9)
Cas9	CRISPR-Associated
CE	Constituent Enhancer
CE1	Constituent Enhancer 1
CE2	Constituent Enhancer 2
ChIP-seq	Chromatin Immunoprecipitation with Sequencing
CHORI	Childrens Hospital Oakland Research Institute
CNS	Central Nervous System
CRISPR	Clustered Regulatory Interspaced Short Palindromic Repeats
CRISPRa	CRISPR Activation
CRISPRi	CRISPR Interference
CTCF	CCCTC-Binding Factor
DBD	DNA Binding Domain
DM	Double Minute
DNase I	Deoxyribonuclease I
DNMT	DNA Methyltransferase
DSB	Double Strand Break
ECD	Extracellular Domain
ecDNA	Extrachromosomal DNA
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
ENCODE	Encyclopedia of DNA Elements
eRNA	Enhancer RNA

ESC	Embryonic Stem Cell
EZH2	Enhancer of Zeste 2
FDA	Food and Drug Administration
FPKM	Fragments Per Kilobase of Transcript Per Million
GBM	Glioblastoma
GFP	Green Fluorescent Protein
GOF	Gain of Function
GR	Glucocorticoid Receptor
gRNA	Guide RNA
GSC	Glioblastoma Stem Cell
GTF	General Transcription Factor
GWAS	Genome Wide Association Study
H3K27Ac	Histone 3 Lysine 27 Acetylation
H3K27me3	Histone 3 Lysine 27 Trimethylation
H3K4me1	Histone 3 Lysine 4 Monomethylation
H3K4me3	Histone 3 Lysine 4 Trimethylation
HAT	Histone Acetyltransferase
HDAC	Histone Deacetylase
HDR	Homology Driven Repair
HIV	Human Immunodeficiency Virus
HMT	Histone Methyltransferase
HNSCC	Head and Neck Squamous Cell Carcinoma
HOMER	Hypergeometric Optimization of Motif EnRichment
HPV	Human Papilloma Virus
HRP	Horesradish Peroxidase
HSR	Homogeneously Staining Region
IDH	Isocitrate Dehydrogenase
IGV	Integrative Genomics Viewer
IHC	Immunohistochemistry
Indel	Insertion or Deletion
IR	Insulin Receptor
JunDN	c-Jun Dominant Negative
KRAB	Krüppel-Associated Box
LCR	Locus Control Region
LGG	Low Grade Glioma
LOF	Loss of Function
MACS2	Model-based Analysis of ChIP-Seq
MAPK	Mitogen-activated Protein Kinase
MBD	Methylated-CpG-Binding Domain
MED1	Mediator Complex
MGMT	O6-Methylguanine-DNA Methyltransferase

MMAF	Monomethyl Auristatin F
mTOR	Mechanistic Target of Rapamycin
ncRNA	Non-Coding RNA
NGS	Next Generation Sequencing
NHEJ	Non-homologous End Joining
NMD	Nonsense-Mediated Decay
NSCLC	Non-Small Cell Lung Cancer
O-T	Off-Target
OS	Overall Survival
PAM	Protospacer-Adjacent Motif
PI3K	Phosphatidylinositol 3-Kinase
PIC	Pre-Initiation Complex
PIP3	Phosphatidylinositol 3,4,5-Triphosphate
PKB/Akt	Protein Kinase B
Pol II	RNA Polymerase II
PRC3	Polycomb Repressive Complex 2
Raf	Rapidly Accelerated Fibrosarcoma
Ras	Rat Sarcoma
RNA-Seq	RNA Sequencing
ROSE	Rank Ordering of Super Enhancers
RTK	Receptor Tyrosine Kinase
SAM	S-Adenosylmethionine
SCLC	Small Cell Lung Cancer
SE	Super Enhancer
SH2	Src Homology 2
sKD	Symmetric Kinase Domain
SNP	Single Nucleotide Polymorphism
STARR-Seq	Self-Transcribing Active Regulatory Region Sequencing
T-ALL	T-Cell Acute Lymphoblastic Leukemia
TAD	Topologically Associating Domain
TCGA	The Cancer Genome Atlas
TDG	Thymine DNA Glycosylase
TET1	Ten-eleven Translocation Methylcytosine Dioxygenase 1
TF	Transcription Factor
TKI	Tyrosine Kinase Inhibitor
TRE	TPA-Response Element
TSS	Transcription Start Site
WHO	World Health Organization
wtEGFR	Wild-Type EGFR

LIST OF FIGURES

Figure 3.1: Chromatin landscape of wild-type *EGFR* 30
Figure 3.2: Identification of Super enhancers by ROSE 33
Figure 3.3: GSCs and GBM primary tumors confirm putative *EGFR* intron 1 enhancer presence. 35
Figure 3.4: Identification of critical constituent enhancers in *EGFR* intron 1 of GBM. . . 37
Figure 3.5: Identification of critical constituent enhancers in *EGFR* intron 1 of HNSCC. 38
Figure 3.6: Enhancer-promoter interaction by chromatin conformation capture (3C) . . . 40

Figure 4.1: CRISPR/Cas9-mediated deletion of CE1 and CE2. 51
Figure 4.2: CRISPR/Cas9-mediated deletion of CE1 and CE2 in GBM and HNSCC cell lines reduces *EGFR* expression 53
Figure 4.3: CRISPR/Cas9-mediated double deletion of CE1 and CE2 in GBM and HNSCC cell lines further reduces *EGFR* expression. 54
Figure 4.4: Targeting dCas9-KRAB to the CEs decreases *EGFR* gene transcription and reduces proliferation in HN12 cells. 58
Figure 4.5: Targeting dCas9-KRAB to the CEs decreases *EGFR* gene transcription SF767 cells. 59

Figure 5.1: AP-1 family transcription factors are identified as possible modulators of *EGFR* intron 1 enhancers. 70
Figure 5.2: AP-1 family members bind to *EGFR* Intron 1. 73
Figure 5.3: AP-1 family members modulate *EGFR* expression. 75
Figure 5.4: BET bromodomain proteins bind in *EGFR* intron 1 acetylated regions. . . . 78
Figure 5.5: JQ1 treatment reduces *EGFR* transcription through inhibition of BRD2 and BRD4 activity in HN12 cells. 82
Figure 5.6: JQ1 treatment reduces *EGFR* transcription through inhibition of BRD2 and BRD4 activity in SF767 cells. 84

Figure 6.1: Summary of current anti-EGFR strategies and findings from this study. . . . 97

ACKNOWLEDGEMENTS

There are so many people who deserve credit for helping me get to where I am today, my hope is that I can do them all justice by acknowledging their contributions.

First, I want to thank my advisor Dr. Frank Furnari for the exceptional mentorship he has provided me over the last 5 years. I was incredibly fortunate to identify an advisor who would be available, invested, and a friend. I'll always remember coming into his office, fresh from a difficult meeting where I was told to do a full rewrite of my first big assignment as a PhD student. He was willing to drop everything he was doing and help me work through my ideas and provide suggestions of his own to make the proposal satisfactory to all my committee. I was dangerously close to quitting that day, and he pulled me from the ledge and helped me realize I could do it.

Next, I have to thank my parents. Even though they were hours away and dealing with their own struggles, they were always there to lend a helpful word when things got difficult. Even more important, I consider them to be among my best friends. They would come visit and we would always have a great time, helping me to escape from whatever stress was in my life and just enjoy the moment. Thank you for never pressuring me to be more than I could be, and helping me realize the potential I had in myself.

Third, I want to thank my wonderful wife Rachel. Thank you for sticking through the tough times, traveling to and from San Diego and Phoenix nearly every month so that we could continue to build our relationship, love each other, and have fun together. You've given me a reason to look forward to coming home every day, and now that you've given me the greatest gift in Avery there's even more for me to look forward to every day.

To all my friends, teachers, mentors, collaborators, labmates, acquaintances, and everyone who I've interacted with directly or indirectly on this journey, thank you. No matter the contribution, you've been a part of my success so you deserve your part of the credit.

Chapter 3, chapter 4, and chapter 5 in part, have been submitted for publication of the material as it may appear in Cancer Research; **Jameson, N.M.**, Ma, J., Benitez, J., Izurieta, A., Han, J.Y., Mendez, R., Parisian, A., Furnari, F., AACR Publications, 2019. The dissertation

author was the primary investigator and author of this paper.

VITA

2013	Bachelor of Science in Biological Sciences, Arizona State University
2019	Doctor of Philosophy in Biomedical Sciences, University of California, San Diego

PUBLICATIONS

Whitsett TG, Cheng E, Inge L, Asrani K, **Jameson NM**, Hostetter G, Weiss GJ, Kingsley CB, Loftus JC, Bremner R, Tran NL, Winkles JA. Elevated expression of Fn14 in non-small cell lung cancer correlates with activated EGFR and promotes tumor cell migration and invasion. *American Journal of Pathology*. 2012. 181(1):111-20.

Dhruv HD, Whitsett TG, **Jameson NM**, Patel F, Winkles JA, Berens ME, Tran NL. (2013). Tumor necrosis factor-like weak inducer of apoptosis (TWEAK) promotes glioblastoma cell chemotaxis via Lyn activation. *Carcinogenesis*. 2014. 35(1):218-26.

Benitez JA, Ma J, D'Antonio M, Boyer A, Camargo MF, Zanca C, Kelly S, Khodadadi-Jamayran A, **Jameson NM**, Andersen M, Miletic H, Saberi S, Frazer KA, Cavenee WK, Furnari FB. PTEN regulates glioblastoma oncogenesis through chromatin-associated complexes of DAXX and histone H3.3. *Nature Communications*. 2017. 8: p. 15223.

Zanca C, Villa GR, Benitez JA, Thorne AH, Koga T, D'Antonio M, Ikegami S, Ma J, Boyer AD, Banisadr A, **Jameson NM**, Parisian AD, Eliseeva OV, Barnabe GF, Liu F, Wu S, Yang H, Wykosky J, Frazer KA, Verkhusha VV, Isaguliants MG, Weiss WA, Gahman TC, Shiau AK, Chen CC, Mischel PS, Cavenee WK, Furnari FB. Glioblastoma cellular crosstalk converges on NF-B to attenuate EGFR inhibitor sensitivity. *Genes and Development*. 2017. 31: 1212-1227

ABSTRACT OF THE DISSERTATION

Intron 1-Mediated Regulation of *EGFR* Expression In EGFR-Dependent Malignancies

by

Nathan M. Jameson

Doctor of Philosophy in Biomedical Sciences

University of California, San Diego, 2019

Professor Frank Furnari, Chair

The epidermal growth factor receptor is known to be overexpressed in numerous solid tumor types and has been the subject of extensive therapeutic development efforts. Much of the research on EGFR is focused on protein dynamics and downstream signaling, however few studies have explored how the gene is regulated transcriptionally. Here, we identified two novel enhancers (CE1 and CE2) present within the first intron of the *EGFR* gene in models of glioblastoma (GBM) and head and neck squamous cell carcinoma (HNSCC). CE1 and CE2 contain open chromatin and H3K27Ac histone marks, functionally enhance transcription in reporter assays, and interact with the *EGFR* promoter. Genetic deletion of CE1 and CE2 by CRISPR/Cas9 editing significantly reduces *EGFR* transcript levels, with double deletion exercising an additive

effect. Similarly, targeted repression of CE1 and CE2 by dCas9-KRAB targeting demonstrates repression of transcription similar to that of genomic deletion. We identify AP-1 transcription factor family members in concert with BET bromodomain proteins as candidate modulators of CE1 and CE2 activity in HNSCC and GBM through *de novo* motif identification and validate their presence in these enhancers. Genetic inhibition of AP-1 or pharmacologic disruption of BET/AP-1 binding results in downregulated EGFR protein and transcript levels, further confirming a role for these factors in CE1 and CE2. Our results identify and characterize these novel enhancers, shedding light on the role that epigenetic mechanisms play in regulating *EGFR* transcription in EGFR-dependent cancer types.

Chapter 1

Introduction to the oncogene EGFR and its implications in cancer

1.1 EGFR Function

The epidermal growth factor receptor (EGFR) was the first receptor tyrosine kinase (RTK) to be discovered [1] and has served as the basis for discoveries of the principles and paradigms that underlie the action of other RTKs [2]. EGFR is a member of the ErbB family of RTKs, and is also known as ErbB1 or HER1. This family consists of four members, including ErbB2 (HER2), ErbB3 (HER3) and ErbB4 (HER4). Like all RTKs, EGFR is composed of an extracellular ligand binding domain, a single transmembrane domain, and a cytoplasmic domain which includes the protein tyrosine kinase core. With the exception of the insulin receptor (IR) family of RTKs, all known RTKs exist as monomers in the cell membrane. Canonical EGFR activation is mediated by binding of the epidermal growth factor (EGF) [3], however there are at least six growth factors which can bind to and activate EGFR [4–8]. For EGFR, as with all RTKs, activation is stimulated by ligand binding, inducing dimerization of two monomers [9] and autophosphorylation of multiple tyrosine residues in the cytoplasmic tail which serve as binding sites for signaling proteins containing Src homology 2 (SH2) domains [10]. Unlike some other RTKs which form homodimers exclusively (PDGFR, VEGFR), EGFR can form both homodimers

and heterodimers with other members of the ErbB family and other RTKs [11, 12]. ErbB2 is the preferred dimerization partner of all the ErbB receptors [13], and the heterodimer of EGFR and ErbB2 demonstrates robust signaling activity [14]. EGFR is involved in a number of downstream signaling pathways, and heterodimerization with different ErbB partners and other RTKs, including c-Met, modulates the phosphorylation of the cytoplasmic tail, expanding its downstream signaling potential [11,12].

Many critical signaling pathways are induced by the activation of EGFR. One such pathway is the Ras (Rat Sarcoma)/Raf (Rapidly Accelerated Fibrosarcoma)/MAPK (Mitogen-activated protein kinase) pathway. Following EGFR phosphorylation, the adaptor protein Grb2 binds directly to specific docking sites on the receptor [15]. This interaction recruits Sos and activates Ras, which in turn activates Raf1, which is then able to phosphorylate MAPK1/2 (also known as ERK1/2) [16, 17]. These activated MAPKs are translocated into the nucleus where they phosphorylate transcription factors which promote cell proliferation [18]. In addition to growth, Ras downstream effectors include proteins involved in migration, adhesion, cytoskeletal integrity, survival and differentiation [19]. The importance of this pathway is underscored by the prevalence of alterations in various members of the pathway in many solid tumors. Gain-of-function mutations in the three Ras family members in humans (KRas, HRas and NRas) together are found in up to 30% of all human tumors [20]. The frequency of alterations in this pathway decreases as one moves further downstream, however mutations in v-raf murine sarcoma viral oncogenes homolog B1 (BRAF), including $\text{BRAF}^{\text{V600E}}$, in human cancers is estimated at approximately 7% [21].

Another critical signaling pathway downstream of EGFR activation is the Phosphatidylinositol 3-kinase (PI3K)/Protein kinase B (PKB/Akt) pathway. PI3K contains two subunits: p85 which is responsible for the anchorage to the receptor docking sites, and p110 which is a catalytic subunit that generates phosphatidylinositol 3,4,5-triphosphate (PIP3). PIP3 is responsible for activation of the Akt kinase through recruitment to the plasma membrane, leading to Akt activation by phosphorylation at two different sites [22]. Akt is an essential kinase that has been shown to interact with well over 100 substrates involved in a variety of cellular functions includ-

ing survival, growth, proliferation, angiogenesis, metabolism, and migration [23]. The activity of PIP3 is negatively regulated by dephosphorylation by the tumor suppressor Phosphatase and tensin homolog (PTEN) [24]. Again highlighting the importance of this pathway is the prevalence of alterations, however unlike the Ras/Raf/MAPK pathway, alterations in all the major elements of this pathway have been found mutated or amplified in a broad range of cancers. For example, the p110 catalytic subunit of PI3K has been found to be mutated in up to 30% of breast cancers, and between 10-30% of endometrial, colorectal, urinary tract and ovarian cancers [20]. Loss-of-function mutations or deletion of the tumor suppressor PTEN is also common across a variety of cancers, including approximately 45% of glioblastomas (GBM) [25, 26].

1.2 EGFR in Cancer

Aberrant EGFR signaling is implicated in many human diseases, including many different types of solid tumors. Across all the samples in the The Cancer Genome Atlas (TCGA), which includes over 9,000 tumors and 33 cancer types, the prevalence of EGFR alterations is approximately 4%, making it the third most common alteration in the RTK-RAS pathway [27]. In that same cohort, EGFR alterations were most common in glioblastoma (GBM), low grade glioma (LGG), head and neck cancer (HNSCC), lung adenocarcinoma, and esophagogastric squamous carcinoma, present in 50%, 52%, 13%, 13%, and 14% of these tumors respectively [27]. Alterations within EGFR can take many forms, including amplification, in-frame deletions, and kinase domain mutations. Universally activating, these mutations have variable downstream effects based on different factors including tumor type, intensity of amplification, type of mutation, and previous exposure to EGFR targeted therapy.

1.3 EGFR in Glioblastoma

From 2011-2015, malignant gliomas of the brain including glioblastoma, astrocytoma and oligodendroglioma, accounted for approximately 86,000 of the 110,000 diagnosed tumors of neuroepithelial tissue. This represents approximately 21.7% of all tumors diagnosed in this

time period [28]. Malignant gliomas are among the deadliest of human cancers because they are highly invasive and neurologically destructive [29]. The median survival of patients with the most common and most aggressive of these, grade IV glioblastoma (GBM), is 12- 15 months, with a 5-year survival rate that remains at less than 5% despite the use of intensive treatments [29]. The World Health Organization (WHO) classification of tumors of the Central Nervous System (CNS) distinguishes gliomas based on their histological appearances and molecular compositions, where the grade indicates the level of malignancy [30]. Gliomas, including both high grade GBM and LGG, are the prototypical tumor type which possess *EGFR* alterations at a high level. *EGFR* amplification occurs in approximately 60% of primary GBMs compared to 8% of secondary GBM patients [26] and is a hallmark of the classical subtype, found in approximately 95% of tumors classified in this manner. In contrast, the other molecular subtypes are associated with reduced rates of *EGFR* amplifications at 29% (mesenchymal), 67% (neural), and 17% (proneural) [31]. GBM possesses a strong correlation between *EGFR* copy number and expression. Over 90% of tumors which contain *EGFR* mutation and/or amplification overexpress the protein [32]. *EGFR* amplification in GBM occurs almost exclusively in the form of circularized double-stranded extrachromosomal DNA fragments known as double minutes (DM) [33]. The absence of a centromere in DMs results in a random segregation between daughter cells [34], resulting in highly variable numbers of DMs in a population of GBM cells which can drive tumor evolution and heterogeneity [35].

EGFR Extracellular Domain Deletion Mutations in GBM.

In GBM there are five common deletion mutants, EGFR variants I-V (EGFRvI-EGFRvV). Mutants EGFRvII, EGFRvIII and EGFRvIV have entire exons deleted, while EGFRvI and EGFRvV have NH2-terminal and COOH-terminal truncations, respectively [36]. EGFRvI and EGFRvIV are rare, while EGFRvII and EGFRvV are marginally more common, each accounting for about 10% of all GBM-associated *EGFR* mutations [37]. The most common deletion mutant of EGFR is EGFRvIII. EGFRvIII is a highly tumor specific [38] extracellular domain mutation which shows constitutive tyrosine kinase activity [39], conferring increased tumorigenicity [40]. EGFRvIII results from in-frame deletion of 801 base pairs spanning exons 2-7

of the coding sequence, removing 267 amino acids from the extracellular domain and creating a junction site between exons 1 and 8 and a new glycine residue [41, 42]. Prevalence of EGFRvIII in GBM is in the range of 30-60% depending on the method of identification [43, 44], with recent large-scale genomic analysis indicating an overall prevalence of approximately 40% [26]. EGFRvIII protein has been identified in other tumor types including HNSCC [45], although genomic deletion of exons 2-7 has not been detected. Though EGFRvIII cannot bind ligand and exhibits lower levels of tyrosine kinase activity than activated wild-type EGFR [39], it is able to evade Cbl-mediated receptor internalization [46], rendering the protein highly stable at the cell surface and drastically increasing its tumorigenicity [40]. EGFRvIII is always amplified and exclusively present on ecDNA [47] and exhibits dynamic regulation by integrating and dissociating from the chromosome in response to anti-EGFR therapies [48]. Interestingly, EGFRvIII is heterogeneously expressed among cells within a given tumor, being detected at both very high and very low protein levels in GBM [49]. Previous work from our lab has demonstrated that interclonal cooperation between populations of tumor cells is a predominant mechanism by which GBM aggressive growth is maintained [50]. Mutant EGFRvIII was shown to drive this process via a cytokine circuit to the less aggressive neighboring cells, the majority of which expressed amplified wild-type EGFR (wtEGFR) [50].

Targeting EGFR with Monoclonal Antibodies.

Because of the high prevalence of wtEGFR/EGFRvIII amplification, often present within the same tumor, targeting EGFR and its variants has been a focus in the GBM field. One approach to inhibit EGFR-mediated signaling is to disrupt receptor-activating ligand binding [51]. Monoclonal antibodies, both unconjugated and conjugated, directed towards wtEGFR and EGFRvIII have been developed for therapeutic use in GBM. The most developed of the unconjugated antibodies is cetuximab which functions to prevent EGFR-mediated signal transduction by interfering with ligand binding and EGFR extracellular dimerization [51]. Additionally, cetuximab is believed to trigger EGFR receptor internalization and destruction [52]. Mouse xenograft studies demonstrated that treatment with cetuximab decreases tumor proliferation and increases cell death and overall survival [53], however in clinical trials cetuximab

failed to demonstrate efficacy either as a single agent for recurrent glioblastoma [54], in combination with other reagents [55], or combined with radiation [56]. Some monoclonal antibodies have been engineered to specifically target EGFRvIII, with the hypothesis that these antibodies will be highly tumor specific and avoid off-target side effects. One such antibody is mAb806, which attenuates receptor autophosphorylation by binding to the short cysteine loop of the extracellular domain that is always exposed in EGFRvIII, but may also weakly target amplified wtEGFR, which transiently exposes this epitope during the switch from the inactive to the ligand-activated conformation [57, 58]. Pre-clinical data showed that mAb806 strongly inhibits the growth of tumor xenografts that express EGFRvIII and as expected, more weakly those that express wtEGFR [59]. To enhance the anti-tumor effect of this antibody, a microtubule depolymerization agent monomethyl auristatin F (MMAF) conjugated version was developed and termed ABT-414 [60]. Though it showed initial promise in phase I/II clinical trials [61, 62], use as a monotherapy ultimately failed phase III clinical trials for subjects with newly diagnosed GBM with EGFR amplification [63].

Targeting EGFR with Tyrosine Kinase Inhibitors.

The most clinically advanced, yet altogether disappointing, strategy for targeting amplified *EGFR* in GBM is through the use of tyrosine kinase inhibitors (TKI). Many agents have been developed to target the kinase domain of EGFR, and development has gone through multiple classes of drugs over approximately two decades. Mechanistically, these inhibitors compete with adenosine triphosphate (ATP) for binding to the tyrosine kinase domain of EGFR [64] inhibiting activity and downstream signaling [65]. The leading TKI representatives for GBM include erlotinib, gefitinib, and lapatinib, all of which have shown to be ineffective as both single agents and in combination. Erlotinib showed no efficacy and significant side effects as a single agent in newly diagnosed GBM [66] and also failed in combination with mechanistic target of rapamycin (mTOR) inhibitors [67] or anti-angiogenesis drugs (bevacizumab) [68]. Similarly, gefitinib did not improve overall survival after radiation therapy [69] or administered concurrently in combination with radiation [70] in newly diagnosed GBM. Additionally, although lapatinib is also able to inhibit HER2 [71], it has shown limited efficacy both as a single agent [72] or in

combination with anti-angiogenic agents [73]. Many factors are hypothesized to contribute to the limited efficacy of these TKIs. Evidence supports lack of brain penetrance as a major obstacle to EGFR TKI therapy, with studies showing in vivo EGFR phosphorylation is not significantly blocked in patients receiving these inhibitors [74, 75]. Tumors which are exposed to sufficient doses of TKI therapy are also often able to become resistant through mechanisms including inactivation of PTEN [76], co-activation of other RTKs [77], and signaling pathways such as the TNF-JNK-Axl-ERK pathway [78]. Designing smaller molecules which can more easily pass the blood brain barrier (BBB) [79,80] and rational drug therapy combinations represent novel angles for treating EGFR-positive glioma, however the lack of clinical success thus far by these strategies underscores the challenges in treating GBM by small molecule.

EGFR Extracellular Domain Missense Mutations.

In addition to structural variation at the *EGFR* locus, extracellular domain (ECD) missense mutations also represent a significant proportion of *EGFR* mutations in GBM. Approximately 10-15% of EGFR-altered tumors have these ECD mutations [26, 81] and they confer a significant negative survival effect, lowering the median overall survival (OS) and 2 year survival rate from 15 months and 21% to 6 months and 12% respectively [82]. Although there are many ECD mutations which have been identified, the most common and clinically relevant mutations are A289D/T/V, R108G/K, and G598V which were found in 6%, 3% and 2% of a cohort of 411 GBM cases [82]. Expression of these mutants is associated with increased downstream signaling [81] through ligand independent constitutive signaling [81,83], leading to a more invasive and proliferative phenotype [82]. Importantly, ECD mutation is also highly correlated with *EGFR* amplification [26]. One critical characteristic of the ECD mutations is their ability to induce conformational changes in EGFR on the cell surface. Crystal structures of the EGFR catalytic domain have identified different receptor conformations [84] which significantly affect how these receptors can be targeted. In the presence of ECD mutations, the receptor adopts an inactive symmetric kinase domain (sKD) state in contrast to ligand activated wtEGFR which adopts an asymmetric kinase domain (aKD) [75]. GBMs expressing ECD mutations have been shown to respond to TKIs, however the responses are more robust when using class II TKIs, including

lapatinib, due to preferential binding to the inactive, type II conformation [75,85]. Recent data also indicates the importance of this conformation change in the response to treatments using targeted monoclonal antibodies. In addition to the aKD and sKD conformational states, research using the mAb806 indicates a third ECD transitional state, characterized by exposure of a cryptic epitope [86]. Surprisingly, mAb806 was also found to bind wtEGFR when it is present at high levels on the cell surface [87], however the epitope is buried when the receptor is bound by EGF [88]. Previous research from our lab has shown the enhanced ability of mAb806 to bind to GBM cells expressing $EGFR^{A289V}$ when compared to wtEGFR, nearly to the same degree as it binds EGFRvIII. This binding efficiency led to a significant reduction in both tumor growth and significant enhancement of survival in a mouse xenograft model [82]. These results indicate additional patient stratification may further enhance the clinical potential of the weaponized mAb806 antibody ABT-414.

1.4 EGFR in Non-Small Cell Lung Cancer

Lung cancer is the most prevalent malignancy in the world with an estimated global incidence of approximately 1.6 million and 1.4 million deaths from this disease annually [89]. Lung cancers are classified clinically into two major groups: non-small cell lung cancer (NSCLC), which accounts for about 85% of all lung cancers, and small-cell lung cancer (SCLC) which accounts for the remainder [90]. The most common types of NSCLC include squamous cell carcinoma, large cell carcinoma, and adenocarcinoma which are responsible for approximately 35%, 10% and 45% of NSCLC respectively [91]. Alterations in EGFR are present in a large subset of lung cancers, occurring almost exclusively in NSCLC and primarily those which are of the adenocarcinoma histological subtype [90]. Interestingly, there are significant differences in EGFR alteration prevalence in adenocarcinomas of different populations with Western populations exhibiting approximately 19% EGFR-altered tumors, while Asian populations exhibited EGFR-altered tumors with as much as 48% frequency [92]. Because of this high prevalence of EGFR alteration and the observation that as much as 60% of NSCLCs overexpress EGFR as measured by immunohistochemistry (IHC) [90] anti-EGFR therapies were quickly applied to the

treatment of these tumors. Today, multiple EGFR TKIs are approved by the Food and Drug Administration (FDA) for the treatment of EGFR-positive lung tumors, the first of which was Gefitinib in 2003. Gefitinib was approved as a monotherapy for locally advanced or metastatic NSCLC after failure of both platinum-based and docetaxel chemotherapies, and had an observed objective response rate in about 10% of patients [93]. Interestingly, there was no correlation between EGFR expression and response to gefitinib [94] so to stratify patients more accurately responders and non-responders were compared. Three independent groups identified EGFR kinase domain mutations as the primary predictive marker of response to EGFR TKI including both gefitinib and erlotinib [95–97].

EGFR Kinase Domain Mutations.

Although over 200 EGFR mutations have been described in NSCLC [98], approximately 90% of the activating mutations are one of two mutations: in-frame deletion of five exon-19 residues (exon 19 deletion) and the exon-21 substitution of an arginine for leucine (L858R) [99]. These mutations increase the kinase activity of EGFR, leading to the hyperactivation of downstream pro-survival and growth pathways [100]. Though these mutations have been shown to activate EGFR to a similar degree [101], exon 19 deletion mutations seem to respond better to gefitinib and erlotinib than tumors with the L858R mutation [102]. Perplexingly, the predictive value of EGFR activating mutations is controversial, as approximately 10-20% of patients who show a partial response to gefitinib do not have identifiable EGFR mutations, indicating they are not the sole determinants of TKI response [103]. In contrast to GBM, the role of amplification of *EGFR* in response to TKI therapy is unclear in NSCLC in part due to inconsistencies in EGFR copy number prediction [103]. The frequency of use and initial effectiveness of EGFR TKI therapies in NSCLC spawn inevitable resistance mutations which arise after a median duration of 10-13 months [104]. The most common resistance mechanism by which NSCLC tumors are able to overcome EGFR TKI is the T790M mutation in exon 20 which occurs in 50-60% of patients whose tumors progress on TKI therapy [105]. This mutation replaces threonine with the larger methionine residue near the ATP-binding site, rendering ATP able to compete more effectively with TKIs for binding and decreasing the inhibitory effect of the drugs [106]. To over-

come this common resistance mechanism, mutant-selective third generation EGFR TKIs such as rociletinib [107] and osimertinib [108] were developed. FDA approval through the Breakthrough Therapy designation of osimertinib for patients with metastatic NSCLC whose disease has progressed on EGFR targeted therapy and harbor a T790M mutation speaks to the power of these third generation drugs, however even tumors which initially respond to these treatments eventually develop resistance through additional mutations in exon 20 including C797S. This mutation abolishes the covalent bonding of osimertinib to EGFR [109] and has a prevalence range between 22-40% in patients who have progressed on osimertinib [110]. The ability of NSCLC to adapt and overcome targeted therapy is a testament to the determination of cancer to continue to persevere and the subsequent perseverance of researchers continuing to develop personalized medicines.

1.5 EGFR in Head and Neck Squamous Cell Carcinoma

Although commonly referred to as head and neck cancer, squamous cell carcinomas of the head and neck (HNSCC) represent a diverse set of tumors from various regions of the body including the larynx, throat, lips, mouth, nose, and salivary glands. Worldwide, HNSCC is the seventh most common cancer type with a yearly incidence rate of approximately 600,000 cases, with 40-50% mortality [111]. Although the tumors originate in the epithelial cells of the mucosal linings of the upper airway and food passages, the disease is remarkably heterogeneous at the molecular level [112]. The major known risk factors for development of HNSCC are environmental exposures to tobacco and alcohol, as well as infection with human papillomaviruses (HPV) which is now recognized as the strongest prognostic marker [113]. Because of the heterogeneous nature of this cancer type, significant efforts have been made to identify clinically actionable molecular subtypes and biomarkers [114,115]. Interestingly, these studies have identified distinct genetic signatures between HPV$^-$ and HPV$^+$ tumors whereby the driver genes of HPV$^+$ tumors are largely a mystery while HPV$^-$ tumors have high rates of known oncogenic mutations, including EGFR [116]. One of the largest and most comprehensive datasets of both HPV$^-$ and HPV$^+$ HNSCC is the molecular profiling of 279 tumors by the TCGA [117]. Based on this analysis, between 50 and 100 genes are indicated to be substantially mutated and are considered candidate

cancer driver genes. Many of the genes implicated as drivers for GBM are also frequently altered in HNSCC, including TP53 (84%), CDKN2A (57%), PIK3CA (34%), and EGFR (15%) [117]. Paradoxically, while amplification of EGFR is relatively uncommon in comparison to other driver genes, studies have identified protein overexpression in 47-84% of HNSCC tumors [118–120]. This range may be due to differences in expression of EGFR in different sites of HNSCC [121]. In contrast to both GBM and NSCLC, alteration of EGFR in the form of either KD or ECD mutations is rare. Although high rates of EGFRvIII have been reported in HNSCC [45], updated molecular profiling by TCGA identified only 1 of 279 tumors with EGFRvIII by RNA-seq (0.4%) [117]. Additionally, kinase domain mutations were only identified in 117 out of 4122 patients (2.8%) with missense mutations in exons 18-21 occurring in 73% of the 117 [122].

EGFR Targeted Therapies in HNSCC.

Of the myriad anti-EGFR therapies discussed, only cetuximab has been approved for the treatment of advanced HNSCC as either a single agent [123], in combination with radiation [124], or in combination with platinum based chemotherapeutics [125]. Although it was able to significantly improve survival statistics, response rates as a single agent remain low at approximately 13% [124]. This relatively low response rate does not seem to be due to lack of EGFR expression in these tumors, rather investigation into the relation between EGFR expression and the outcome of HNSCC patients treated with cetuximab have shown that EGFR levels detected by IHC have no impact on response [126, 127]. However, other studies have reported a positive relationship between EGFR expression and anti-EGFR (cetuximab, nimotuzumab) therapy responses [128,129]. These disparities underscore the importance of tumor heterogeneity and careful patient curation in trials assessing the treatment of HNSCC tumors with anti-EGFR therapies. HNSCC tumors commonly escape anti-EGFR therapy by upregulating pathways which can compensate for reduced EGFR signaling and/or modulate EGFR-dependent signaling. One of the primary escape mechanisms in response to cetuximab therapy is upregulation of signaling by other growth factor receptors. The most extensively studied of these receptors in this capacity are HER2/ErbB2 and HER3/ErbB3. HER2 alterations have been directly linked to resistance to EGFR-targeted therapy, with expression of phospho-HER2 and HER2 expression

significantly correlating with protein expression of EGFR [130]. HER3 activity is dependent on EGFR and HER2, and the HER2/HER3 heterodimerization is one of the main forms of cetuximab resistance in HNSCC [131, 132]. In light of these observations, simultaneous blocking of EGFR with either HER2 or HER3 has been proposed. Indeed, combinatorial treatment of cetuximab with pertuzimab (HER2 monoclonal antibody) or seribantumab (HER3 monoclonal antibody) results in potent tumor suppression [131, 132]. Likewise, EGFR TKIs like lapatinib (EGFR/HER2) [133] or dacomitinib (pan-HER) [134] have shown some preclinical benefit. The lack of reliable biomarkers for response to anti-EGFR therapies underscores the need for novel avenues of research on the role of EGFR in HNSCC.

1.6 Summary

EGFR was the first RTK to be identified and labeled as an oncogene, and as such it has become one of the most thoroughly studied proteins in biomedical research. Canonical EGFR signaling is involved in many signaling pathways which are widely recognized as critical for the growth and survival of cancer cells, making EGFR itself and its downstream effectors attractive targets for therapy. Indeed, EGFR has been the target of a myriad of treatments including but not limited to monoclonal antibodies, TKIs, RNA interference [135], vaccines [136], and immune therapy [137]. While some successes have been achieved, including approvals for the monoclonal antibody Cetuximab in HNSCC [123] and TKIs Erlotinib [138] and Gefitinib [93] in NSCLC, the clinical outcomes for many patients with EGFR alterations, particularly those with GBM, remain dismal as evidenced by the most recent phase III failure of an EGFRvIII-specific monoclonal antibody therapy [63]. Thus, while knowledge of inherent genetic alterations is pertinent in determining rational therapeutic targets, the monotherapies that have emerged from this knowledge are inadequate for generating a durable response in many patients. These observations underscore the critical need for new paradigms of thinking for targeting EGFR-positive malignancies. Perhaps though the EGFR protein and its domains have been the focus for anti-EGFR therapies in the past, it is time to look to the EGFR gene for clues to guide the next generation of therapy.

Chapter 2

Introduction to epigenetics and its role in cancer

2.1 Epigenetics

Epigenetics as a term was first used to categorize all of the regulated developmental processes that, beginning with genetic material, lead to a mature organism [139]. The present definition of epigenetics better reflects the understanding that although DNA is essentially the same in all of an organisms somatic cells, patterns of gene expression differ greatly among different cell types, and these patterns can be inherited. Thus, a succinct definition of epigenetics was proposed as the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence [140]. Since the fully completed sequencing of the human genome in 2001, it is known that only approximately 2% of the human genome encodes for protein [141], however large scale functional mapping suggests that more than 80% of the human genome participates in at least one RNA- or chromatin-associated event in at least one cell type [142]. This is concrete evidence that what was once thought of as junk DNA is in fact playing a critical role in the function of the human genome.

2.2 The Nucleosome

The most fundamental unit of epigenetic gene regulation is the nucleosome. The nucleosome is made up of 147 DNA base pairs wrapped around a histone octamer composed of two subunits of each histone H2A, H2B, H3 and H4. The wrapping of histones by DNA allows for the packaging of DNA into highly compacted structures, folding the DNA further into tighter and tighter structures, which when taken together comprise the entire DNA/protein structure within the cell known as chromatin [143]. Nucleosomes are powerful regulatory units due to their ability to be modified at specific residues within the unstructured N-terminal tails of core histones. To date, there are 16 identified classes of histone modifications [144], with the covalent modifications acetylation, methylation, and phosphorylation encompassing most of the known functions. These modifications can alter chromatin structure by modifying noncovalent interactions within and between nucleosomes. They also serve as docking sites for specialized proteins with unique domains that specifically recognize these modifications, which then recruit additional chromatin modifiers and remodeling enzymes, serving as effectors of the modifications [145]. These histone modifications are regulated as part of a highly dynamic process, as their deposition and removal at specific genomic regions is constantly changing through the activity of large classes of specialized histone modifier proteins [143]. The fluidity of this process is underscored by the observation that all of these modifications are reversible [146].

2.3 DNA Methylation

The DNA around the histone is also subject to modifications. To date, four different DNA modifications have been identified [147], however methylation of the 5-carbon on cytosine residues (5mC) in CpG dinucleotides was the first described covalent modification of DNA and is perhaps the most extensively characterized modification of chromatin [148]. To date there are three DNA methyltransferases (DNMT) that have been identified to methylate DNA directly. DNMT1 recognizes DNA bases whose partners have yet to be methylated during DNA synthesis (hemimethylated), and methylates those newly synthesized CpG dinucleotides [149]. DNMT3a

and DNMT3b are also capable of recognizing hemimethylated DNA, however their primary functions are as *de novo* methyltransferases to establish DNA methylation during embryogenesis [150]. Importantly, there are no known mammalian direct DNA de-methylases, however demethylation of mammalian DNA can occur both passively and actively. Passive DNA demethylation can occur through successive cycles of DNA replication in the absence of functional DNMT1, therefore diluting the total number of methylated cytosines [151]. In contrast, active demethylation requires multiple steps and the activities of several proteins. A key intermediate of the demethylation process, 5-hydroxymethylcytosine (5hmC), was discovered by Tahiliani and colleagues [152]. 5hmC occurs through the activity of Ten-eleven translocation methylcytosine dioxygenase 1 (TET1), which oxidizes the methyl group of 5mC. This oxidation process was also discovered to be iterative, yielding further oxidized products such as 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [153]. These oxidized bases are recognized by thymine DNA glycosylase (TDG), excised from the DNA strand, and then repaired by base excision repair (BER) to the correct cytosine base [154]. Methylation represses transcription directly by inhibiting the binding of specific transcription factors, and by recruiting methylated-CpG-binding domain (MBD) proteins and their associated repressive chromatin remodeling activities [155]. MeCP2 is a well-known methyl-CpG binding domain protein, which contains an MBD domain and a transcriptional repression domain that acts to recruit the Sin3 HDAC complex. MeCP2 is capable of binding to a single methylated CpG and facilitates the methylation of lysine 9 in H3, possibly serving as a bridge between DNA methylation and histone methylation [156]. CpG dinucleotides are overall underrepresented in the genome, but are locally enriched in shorter stretches of DNA known as CpG islands [157]. An estimated 60% of human genes contain a CpG island, and these islands are most often associated with a transcriptional control region known as the promoter [158].

2.4 Non-coding Regulatory Elements

Complex organisms are made up of a wide variety of cell types, yet each of these cell types contains the same genome. This incredible diversity can be explained by the presence of non-coding regulatory elements which govern when, where, and to what level each gene will be

expressed.

Promoters.

Gene expression starts with transcription, a process which initiates at a defined position, known as the transcription start site (TSS), embedded within a core promoter sequence. The core promoter serves as a binding platform for the transcription machinery, comprising RNA polymerase II (Pol II) and other associated general transcription factors (GTF) [159]. Because of their critical role in the control of gene expression, the promoter is a tightly regulated region of DNA. As mentioned earlier, CpG islands are often found at gene promoters and are often hypomethylated [160], and the methylation of CpG islands results in stable silencing of gene expression [161]. In addition to promoter methylation as a mechanism of transcriptional control, the histone tails at promoter regions can also be modified and be indicative of activation or repression. Active core promoters are surprisingly known to be devoid of histones, and are instead comprised of non-canonical or partial nucleosomes [162], making them accessible and allowing for the assembly of the pre-initiation complex (PIC) [163]. It is instead the nucleosomes flanking the active promoter which are modified post-translationally.

Identification of these post-translational modifications was only recently applied genome-wide. The current gold-standard for identifying the binding positions of post-translationally modified histones is chromatin immunoprecipitation with sequencing (ChIP-seq) [164]. By crosslinking protein to DNA, shearing the crosslinked DNA by physical or chemical methods, and immunoprecipitating fragmented DNA using antibodies specific to a histone modification of interest one is left with a library of DNA fragments in direct contact with the factor of interest. Subjecting these DNA fragments to next generation sequencing methods allows for the identification of DNA regions which are enriched for the protein of interest. Active promoters are most often marked by tri-methylation of histone 3 at lysine 4 (H3K4me3) and acetylation of histone 3 at lysine 27 (H3K27Ac) [165]. Though these marks are strongly correlated with transcriptional activity, the mechanism is unclear. H3K4me3 is proposed to provide a memory of recent transcriptional activity, facilitating new rounds of transcription [166] while H3K27Ac appears to work through decreasing the affinity of DNA for nucleosomes, promoting open chromatin [167]

and providing binding sites for cofactors which bind acetylated lysines such as BRD4 [168]. The tri-methylation of histone 3 at lysine 27 (H3K27me3) is also often found at promoters, and is methylated by the Enhancer of Zeste 2 (EZH2) subunit of the polycomb repressive complex 2 (PRC2) [169]. H3K27me3 is know to have dual roles in the control of gene transcription. Broad enrichment of H3K27me3 in the gene body is known to be indicative of transcriptional repression [170], however bivalent presence of both H3K27me3 and H3K4me3 at the promoter is a marker of poised genes which can be rapidly transcribed in response to stimuli [171].

Enhancers.

While the promoter is immediately adjacent to the gene it affects, enhancers are additional key regulatory elements which can control cell-type-specific spatiotemporal gene expression irrespective of genomic location [172]. The initial discovery of these elements was a 72-bp sequence of the SV40 virus genome which could enhance transcription of a reporter gene by several hundred fold [173]. Subsequently, enhancers have been found to be highly abundant in the human genome, with some groups estimating >400,000 regions with enhancer-like features [142]. These regions are identified largely through the identification of enriched histone post-translational modifications including mono-methylation of histone H3 at lysine 4 (H3K4me1) and H3K27Ac [174]. Different combinations of these marks in combination with H3K27me3 define various enhancer-states. Neutral or intermediate enhancers are marked by H3K4me1 alone, while active enhancers are marked by H3K4me1 in combination with H3K27Ac [175]. Similarly to promoters, a poised state exists marked specifically by a combination of H3K4me1 and H3K27me3 [176]. Although these are widely used marks for identifying enhancers, there is no consensus on which histone marks should be used. This stems from the fact that none of the known histone modifications correlates perfectly with enhancer activity [177] and the functional roles for most of the chromatin modifications associated with active enhancers is unknown. Indeed, mutation of H3K4 or H3K27 such that they can no longer be modified are still compatible with gene transcription, suggesting these modifications are neither necessary, sufficient, or mechanistically involved in transcription [178]. Enhancer prediction based on these histone marks is not sufficient to ascribe functionality, so additional methods have been utilized to determine the activity of spe-

cific enhancers including enhancer activity assays [173], next generation sequencing (NGS)-based techniques [179], and CRISPR/Cas9-based techniques [180].

Another hypothesis for how enhancers can be identified and are able to influence transcription was proposed following the discovery of extragenic transcripts at the locus control regions (LCR), later understood to be enhancers, of a few genes [181, 182]. These transcripts were expressed in a cell type manner and correlated with LCR functionality, suggesting their transcription was linked to LCR activity [181]. Definitive evidence of enhancer transcription came with the discovery of pervasively transcribed putative enhancers, which produced largely non-polyadenylated non-coding RNAs (ncRNA) henceforth known as enhancer RNA (eRNA) [183, 184]. eRNA transcription appears to be a strong indicator of highly functional enhancers [185], however the specific function of eRNA has been the subject of much research. eRNAs generally display low stability and abundance [186] but are able to respond rapidly, often before promoters, to stimuli [187]. Studies of individual genes have ascribed direct consequences on transcription for specific eRNAs [188, 189], however the direct mechanism is still a mystery. Hypotheses for eRNA function include enhancing chromatin accessibility [186], formation and stabilization of chromatin organization [188], Pol II pause release [190], and stabilization of TF binding [191].

2.5 3D Genome Organization

DNA Interaction.

The intense compaction of DNA into chromatin fueled the hypothesis that a physical interaction between enhancers and promoters facilitated the transcriptional enhancement activity of enhancers. Recent evidence has solidified the validity of this hypothesis through forced chromatin looping between the mouse *Hbb* gene promoter and its enhancer, leading to strong transcriptional activation even in the absence of other transcriptional co-activators [192]. Development of chromosome conformation capture (3C) [193, 194] technologies has further strengthened our understanding of which enhancers and promoters interact. 3C and 3C derivatives coupled with NGS enabled the mapping of genome wide enhancer-promoter contacts [194] and

has facilitated the segmentation of the genome into distinct units such as active chromatins hubs (ACH) [195] and topologically associating domains (TAD) [196]. The discovery of the TAD as a unit of chromatin organization brought together many hypotheses on how enhancers and promoters are able to interact specifically with only a few genomic partners. TADs are megabase scale structures that function as a large DNA loop, pinched off at the base by CCCTC-binding factor (CTCF) and cohesion [197, 198], within which interactions are preferentially occurring and are largely independent of interaction with other TADs [196]. Within a single TAD can exist multiple protein coding genes and their associated promoters, enhancers, and associated transcriptional co-factors, forming an ACH [195, 199, 200]. Though significant evidence suggests genes within TADs form preferential contacts with each other, the boundary insulation is not absolute. High-resolution interaction profiling has shown about one-third of significant interactions crossed TAD boundaries [201] and significant interactions can occur between elements located across large genomic distances, often bypassing more proximally located genes [202]. These observations further demonstrate the highly complex nature of chromatin organization and underscore the need for functional validation of identified transcriptional control units.

Nucleosome Positioning and Open Chromatin.

The likelihood of a chromatin-binding protein mediating its effects on DNA or histones is drastically increased in open or accessible regions of the genome. The opening of chromatin through nucleosome eviction or destabilization at promoters and enhancers results from the binding of specific regulatory factors responsible for transcriptional activation [203]. Open or accessible regions of the genome are thus regarded as the primary positions for regulatory elements [204]. Transposase-accessible chromatin using sequencing (ATAC-seq) is the current gold-standard for assaying chromatin accessibility [205]. In the ATAC-seq protocol, unfixed nuclei are tagged *in vitro* with sequencing adapters by purified Tn5 transposase. Due to steric hindrance the majority of adapters are integrated into regions of accessible chromatin that are subsequently PCR amplified for library construction followed by paired-end NGS [205]. ATAC-seq data can be immensely valuable for identifying nucleosome-TF spacing patterns and TF occupancy [206], enabling identification of TFs within enhancer and promoter regions.

2.6 Alterations in Epigenetic Elements

Mutations in Chromatin Proteins.

Although cancer is typically considered a genetic disease, epigenetic alterations are nearly ubiquitous. Approximately 50% of human cancers harbor mutations in chromatin proteins [207] and many tumors exhibit markers of epigenetic reprogramming [208]. Non-neoplastic cells maintain chromatin homeostasis through the interplay of many different families of proteins repressing, activating, or remodeling chromatin at the appropriate time and in response to the appropriate signals. Recurrent mutations in the genes encoding these factors thus are likely to disrupt this homeostasis. As introduced earlier, the PRC2 complex is a critical mediator of gene repression through the activity of EZH2, a catalyzer of histone methylation. Dysregulation of EZH2 is frequent in several tumor types [209] and can occur in the form of gain-of-function (GOF) [210] and loss-of-function (LOF) mutations [211], overexpression [212], mutations in its demethylase UTX [213], and mutations in the SWI-SNF chromatin remodeling complex [214]. EZH2 can function as a potent oncogene, with forced expression causing neoplastic transformation of human breast epithelial cells [215]. The primary residue which is mutated in EZH2 across cancer types is Y641, a critical residue within the C-terminal catalytic SET domain, which leads to hypermethylation of H3K27me2 and a shifting of the steady state of H3K27 modification to trimethylation thus repressing polycomb targets [216]. Loss of UTX activity by mutation also has a phenotypically identical effect of EZH2 GOF mutations, as UTX catalyzes the removal of methyl groups from H3K27me3. Importantly, the SWI-SNF chromatin remodeling complex has been found to be mutated in approximately 20% of all human cancers [214], and performs a critical role in opposing the activity of PRC2. Mutations in critical SWI-SNF subunits, including SNF5/SMARCB1, are known driver mutations in malignant rhabdoid tumors [217] and can form synthetic lethal interactions with EZH2 in other SWI-SNF-mutant cancers [218]. Given the evidence for EZH2 enzymatic gain of function being a cancer driver, the development of EZH2-specific inhibitors has been an active area of investigation. Potent and selective inhibitors of EZH2 were first identified through high-throughput biochemical screens, the first iterations of which are S-adenosylmethionine (SAM)-competitive inhibiting compounds which bind wild-type and Y641

mutant EZH2 and inhibit EZH2-mediated H3K27me3 deposition [219]. SMARCB1-mutant preclinical models showed dose dependent regression [220], and EZH2 inhibitors in clinical trials and have shown some encouraging response in hematological malignancies [221]. Phase I clinical trial results in solid tumors is less robust, with initial reports showing tumor response in only 2 of 43 (5%) SMARCB1/SMARCA4-negative tumors [222]. As with many treatments, resistance emerges in the form of secondary mutations (Y111L and Y661D) following prolonged exposure to EZH2 inhibitors [223], underscoring the tumor dependency on EZH2 and the challenges involved in precision medicine.

Aberrant DNA Methylation.

Another mechanism of epigenetic alteration in cancer is aberrant DNA methylation, either in the form of hypo- or hyper-methylation depending on the tumor type and associated mutations. In hematological malignancies like acute myeloid leukemia (AML) and T cell acute lymphoblastic leukemia (T-ALL), LOF mutations of the methyltransferase DNMT3A are present at a frequency of approximately 25% [224, 225]. These mutations lead to hypomethylation of cancer-associated genes and aberrant activation of enhancers that drive oncogenic gene expression patterns [226, 227]. Interestingly, the opposite phenotype can also be tumorigenic. *De novo* methylation has been shown to occur at tumor suppressor genes in many tumor types [228] and may be a byproduct of the aging process [229], the most important demographic risk factor for cancer [230]. Some common genes are specifically methylated across tumor types. One example is the gene encoding the DNA repair enzyme O6-methylguanine-DNA methyltransferase (MGMT). MGMT is silenced by hypermethylation events in many cancers, and is often an early event in tumorigenesis [231]. The inactivation of MGMT is primarily epigenetic, evidenced by the higher percentage of MGMT-negative tumors with methylation than mutation [232, 233]. In central nervous system (CNS) malignancies, methylation is now being used to stratify tumor types and inform diagnosis and treatment [234]. In particular, MGMT promoter methylation status has been used to indicate the efficacy of treatment [235] and testing for promoter methylation is becoming part of the current standard of care [236]. Aside from prognostic value, direct inhibition of DNA methylation has shown some effectiveness for the treatment of cancer. The

classical demethylating agents comprise the analogs of deoxycytidine and were developed initially as cytotoxic drugs, however they were found to be effective DNA methylation inhibitors [237]. Incorporation of these nucleoside analogs, including 5-azacytidine (5-aza), inhibits methyltransferases at low doses and at high doses result in cell death [238]. Initial clinical trials with these drugs in solid tumors produced negligible effects [239], however subsequent trials have shown the demethylase activity to be effective in reactivating the expression of previously methylated tumor suppressor genes [240]. While promising, this reactivating effect has yet to demonstrate a clinical response either on its own or in combination with classical cytotoxic therapies [241].

Another example of common methylation-associated mutations are alterations in the genes encoding isocitrate dehydrogenase (IDH). GOF mutations in IDH are frequent initiating events in glioma, leukemia and other less common tumor types [242]. Normal IDH1 and IDH2 are located in the cytoplasm and peroxisomes respectively, and catalyze a redox reaction that converts isocitrate to α-ketoglutarate (α-KG) and reduces NADP to NADPH. NADPH supplies reducing power to key reactions in a number of macromolecular biosynthetic pathways, and to systems that defend against the oxidative stress imposed by reactive oxygen species (ROS). Therefore, mutations that disrupt the normal functions of IDH1 and IDH2 have significant consequences on cellular redox balance [242]. Additionally, mutant IDH was discovered to convert α-KG to 2-hydroxyglutarate (2-HG) in a reaction that consumes, rather than produces, NADPH [243]. 2-HG has received significant attention as an onco-metabolite as it has been shown to inhibit the TET proteins involved in demethylation of DNA, leading to a hypermethylation signature characteristic of stem/progenitor cells [244]. Low grade gliomas in particular are susceptible to IDH mutation, with mutation occurring in 70-90% of all adult grade 2/3 astrocytomas, oligodendrogliomas, and secondary gliomas, but only 5-15% of primary glioblastoma [245]. So strong is this association that IDH mutation is used in the WHO classification of gliomas [30], and is an effective favorable prognostic biomarker of survival when compared to IDH wild-type gliomas [246]. Targeting mutant IDH with specific inhibitors showed promise in preclinical models of glioma [247] and AML [248], and has recently led to FDA approval of IDH1 or IDH2 specific inhibitors for the treatment of newly diagnosed IDH-mutant AML [249].

Genomic Variation.

Beyond chemical modification of DNA or histones, genomic variation within regulatory elements and transcription factors provide a significant risk factor for the development of disease. Recent studies have shown that enhancer-like regions contain a high density of genomic variants [142, 250]. To date, hundreds of genome wide association studies (GWAS) have been conducted spanning diverse diseases and phenotypes, and the majority (~93%) of the identified disease- and trait-associated variants lie within noncoding sequence [251]. The exact mechanism for how non-coding variants can alter gene expression is unknown, however some studies have shown that putative causative single nucleotide polymorphisms (SNP) interfere with recognizable TF binding sites, however this only explains the mechanism behind 10-20% of the variants [252, 253]. Additional proposed mechanisms include post-transcriptional processes like mRNA splicing [254] and stability [255], altered DNA methylation [256], and changes in DNase hypersensitivity [257].

Mutations in genes encoding enhancer binding proteins and proteins involved in mediated enhancer-promoter specificity represent a second class of disease-associated variants. Beyond SNPs in regulatory regions, there are a few examples of direct causal variations in epigenetic-associated proteins in developmental disorders and cancer. Cornelia de Lange syndrome is most likely caused by defects in the subunits of the cohesion complex, leading to alterations in cohesins role in enhancer-promoter communication [258]. Similarly, Kabuki syndrome is caused by missense or nonsense mutations in MLL4, a member of the complex including the H3K27 demethylase UTX, or direct mutation in the *UTX* gene [259]. Gene fusions with MLL are known to be expressed in leukemias, and alter transcriptional elongation and proper H3K4 methylation [260]. Additionally, mutations in the genes encoding CBP or P300, major histone acetyltransferases (HAT), can cause the developmental disorder Rubinstein-Taybi syndrome [261]. Perhaps the most widely studied oncogenic transcription factor is c-Myc. *MYC* is the most frequently amplified oncogene, and the elevated expression of its gene product c-Myc can promote tumorigenesis in a wide range of tissues [262, 263]. Rather than binding and regulating a new set of genes when overexpressed, the transcription factor is known to accumulate in the promoter regions

of most active genes, causing transcriptional amplification of the existing gene expression program [262, 263]. These observations make c-Myc an enticing target for therapeutics, however direct targeting of c-Myc has proven exceedingly difficult. Indirect targeting strategies including disrupting stabilization [264], disruption of Pol II activity [265], and targeting of transcriptional activation partners [266] have shown some promise, however no FDA approved therapeutics have emerged.

2.7 Summary

Though the idea of epigenetics has been around for at least as long as we've known DNA is the carrier of genetic information, its functional application to the expression of genes is a relatively new field of study. In this relatively short time period, epigenetics has become one of the fundamental facets of biomedical research due to its clear role in the myriad of processes described above. The power in the study of epigenetics is that it is a fluid process, unlike the rigid genetic code of DNA. Through the dynamic deposition and removal of various DNA and histone modifications, huge families of proteins can associate, dissociate, and interact in every imaginable combination to produce the huge variety of cell types present in our bodies. The mobile nature of these modifications also provide a significant challenge for the study of epigenetics.

Most assays take a snapshot in time, a static picture of where modifications and proteins happened to be sitting at the arbitrary point in time we decided to fix them. It is therefore currently impossible to get a full understanding of the epigenetic changes undergoing in a cell during development or tumorigenesis. It is nonetheless important to study these snapshots to begin to understand the fundamental processes that underly gene expression in these important cell types. We can utilize the knowledge we gain to come up with hypotheses for rational treatment of diseases like cancer. The validity of this process is exemplified in the drugs that have already been developed to treat cancers like lymphoma [221] and leukemia [249]. Clearly alterations in DNA methylation, histone post-translational modification, and mutation in chromatin modifiers have significant effects on the progression of disease as discussed above, thus underscoring the rationale behind targeted therapies to reduce the malicious effects of these changes.

Continued study of the epigenetic drivers behind disease progression is an important endeavor, and will continue to produce significant breakthroughs as evidenced by the positive results thus far generated.

Chapter 3

Identification of putative active enhancers in the first intron of *EGFR*

3.1 Introduction

Recently, a new subset of enhancers was discovered when it was observed that there are unusual co-occurrences of clusters of multiple enhancers spread throughout the genome. These enhancer groups, initially called stretch enhancers and now commonly referred to as super enhancers (SE), have unique properties which distinguish them from common enhancers and justify their super moniker. The initial discovery of super enhancers stemmed from the observation that known master transcription factors OCT4, SOX2 and Nanog are critical for the embryonic stem cell (ESC) state, and they bind enhancers along with the Pol II co-activator Mediator [267]. ChIP-seq for these factors identified two groups of enhancers. First, small enhancers with low levels of Mediator, and much larger clusters of enhancers which were found within the same TADs as critical ESC identity genes which they dubbed super enhancers [267]. Specific knockdown of TFs enriched at these SEs had larger effects at genes nearby the SEs, indicating the increased role the SEs play in gene expression. Subsequent studies have laid out many themes

which distinguish SEs from other regulatory elements [267].

SEs are densely occupied by H3K27ac and H3K4me1 enhancer marks, the Mediator complex (MED1) and bromodomain-containing protein 4 (BRD4) [268, 269]. They contain cell type-specific TFs and constituent enhancers that exhibit an order-of-magnitude higher abundance of enhancer associated chromatin marks and TFs as compared to the composition of regular enhancers [267–269]. The genes driven by SEs are expressed at significantly higher levels than are genes under the control of regular enhancers [267–269]. Within a single SE, individual constituent enhancers are capable of increased transcriptional activation as compared to regular enhancers, and exhibit some functional redundancy but can also interact with additional genes independently of each other [268, 270–272]. Additionally, massively parallel constituent enhancer silencing identified only 1-2 constituent enhancers within a SE that can significantly affect gene expression [273]. Importantly, SEs are characterized by differential binding of tissue-specific or disease-specific TFs [269, 270]. For example, in many cancers SEs are marked by lineage-specific TFs at oncogenes including neuroblastoma [274], medulloblastoma [275], esophageal cancer [276], and melanoma [277]. These specific signatures open up the possibility of tumor-specific treatments which target SEs, and CDK inhibitors [278] or bromodomain inhibitors [274] have proven effective at inhibiting SE-specific gene expression. In contrast to the myriad of genetic and epigenetic changes that have been observed between tumor cells *in vitro* and *in vivo*, thus far established cancer cell lines have recapitulated the SE profiles identified in primary human tumors, supporting the hypothesis that SEs are key to the maintenance of tumor identity [278, 279] (Figure 3.2).

In spite of the prevalence of EGFR dependency in solid tumors, few studies have attempted to elucidate the mechanisms of transcriptional control of the gene. Early studies have primarily focused on the *EGFR* promoter, a 36bp element upstream of the AUG translation initiation codon [280]. The promoter does not contain TSS consensus sequences, such as TATA or CAAT box, thus transcription of the gene starts at multiple initiation sites [281]. Later studies attempted to identify other *cis*-acting elements which act in concert with the promoter to drive *EGFR* expression. These studies identified CA dinucleotide repeats [282–284], intron 1 DNase

I hypersensitive sites [285], and cooperative promoter-upstream and intron 1 enhancers [286] as important regulatory mechanisms. Lack of access to or utilization of NGS techniques limited the scope of these studies and argues for a larger breadth of analysis. Recently, *EGFR* super enhancers were identified in several cancer cell lines of different origins [287–291], however detailed mapping of these super enhancers was lacking. Additionally, while some of these studies identified TF binding to *EGFR* enhancers [287, 288] they either failed to demonstrate any functional effect of loss of these transcription factors on *EGFR* expression or did not include *EGFR* as part of their detailed analysis [289–291].

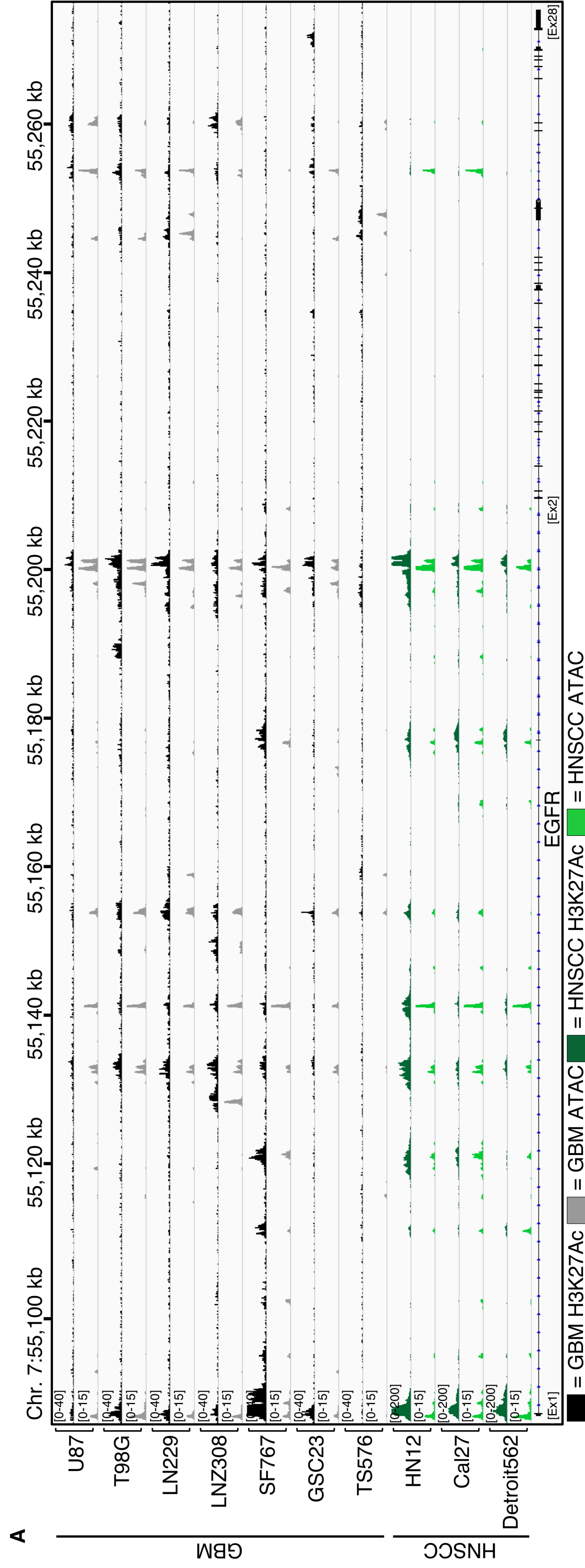
As described in Chapter 2, enhancer prediction based on histone marks is not sufficient to ascribe functionality. The gold standard for validating enhancers is the reporter assay. These assays rely on the ability of enhancers to increase the expression of a gene, independent of its orientation and flexible with respect to its position relative to the transcriptional start site [173,292]. The standard approach is to relocate the candidate enhancer sequence to a reporter vector, adjacent to a minimal promoter driving expression of a reporter gene, e.g. luciferase or green fluorescent protein (GFP). Many groups have attempted to functionally annotate predicted enhancers through a multitude of genome wide methods [180, 293]. The most comprehensive method for functionally testing predicted enhancers is self-transcribing active regulatory region sequencing (STARR-seq). This strategy was first validated in *Drosophila melanogaster* DNA, where initial enhancer prediction identified 50,000-100,000 enhancers. Brute force methods to validate these enhancers required analyzing fly embyos by *in situ* hybridization for 7,705 reporter constructs, requiring massive time and effort resulting in the identification of 3,557 active enhancers [294]. STARR-seq in these fly cells required only a single sequencing experiment to cover 96% of the non-repetitive genome and ascribe functionality to 5,499 regions which passed the threshold of activation to be considered enhancers. Validation of 77 randomly chosen peaks by reporter assays showed that 81% of the peaks could be functionally annotated, bringing the validated total strikingly close to low-throughput methods [179]. Ultimately as these studies have demonstrated, validation of enhancers still falls to the reporter assay for the final confirmation of activity.

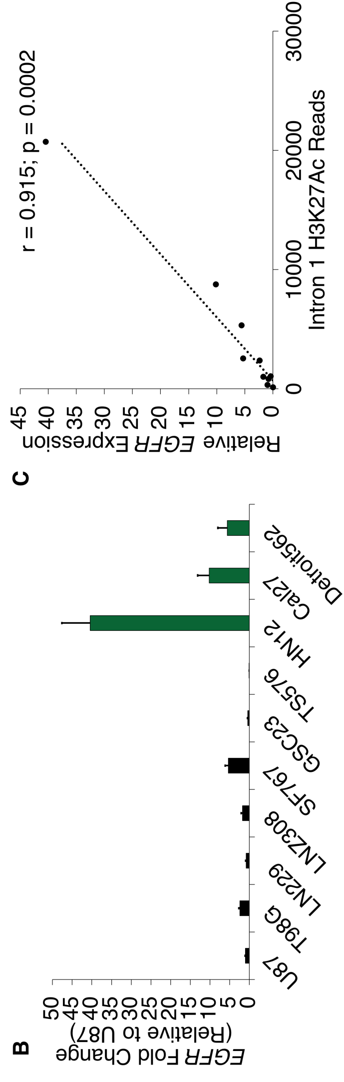
3.2 Results

EGFR Intron 1 Contains Open Chromatin Regions Containing Histone Marks Indicative of Enhancers

To gain further insight into the mechanisms responsible for *EGFR* transcriptional control in HNSCC and GBM we performed ChIP-seq for the enhancer-marking histone modification H3K27Ac and ATAC-seq in 7 GBM and 3 HNSCC cell lines with non-amplified *EGFR* copy number and varying levels of *EGFR* activation. Cell lines with different *EGFR* expression levels were chosen to select for signatures which would preferentially identify regulatory elements which are responsible for driving *EGFR* expression. Overlay of IGV tracks of all 10 cell lines showed conservation of H3K27Ac intensity and open chromatin regions throughout intron 1, indicating the presence of enhancers in these regions (Figure 3.1A).

Figure 3.1: Chromatin landscape of wild-type *EGFR*. **A:** IGV snapshots showing H3K27Ac ChIP-seq (dark) and ATAC-seq (light) read densities at the *EGFR* locus in GBM and HNSCC cell lines. **B:** *EGFR* expression in 7 GBM and 3 HNSCC cell lines were analyzed by RT-qPCR. *EGFR* transcript level was first normalized to GAPDH and subsequently calculated as fold change relative to U87. **C:** Total aligned H3K27Ac ChIP-seq reads were calculated and plotted against the relative *EGFR* expression fold change. The relationship was analyzed using Spearmans rank order correlation.


A
GBM
HNSCC
U87
T98G
LN229
LNZ308
SF767
GSC23
TS576
HN12
Cal27
Detroit562
Chr. 7:55,100 kb
55,120 kb
55,140 kb
55,160 kb
55,180 kb
55,200 kb
55,220 kb
55,240 kb
55,260 kb
EGFR
[Ex1]
[Ex2]
[Ex28]
= GBM H3K27Ac
= GBM ATAC
= HNSCC H3K27Ac
= HNSCC ATAC


B
EGFR Fold Change
(Relative to U87)
50
45
40
35
30
25
20
15
10
5
0
U87
T98G
LN229
LNZ308
SF767
GSC23
TS576
HN12
Cal27
Detroit562
C
Relative EGFR Expression
r = 0.915; p = 0.0002
Intron 1 H3K27Ac Reads
0
10000
20000
30000

Since SEs are defined as large clusters of transcriptional enhancers that drive expression of genes that define cell identity, and are often found at oncogenes [268] we identified SEs using the ROSE (Rank Ordering of Super-Enhancers) algorithm [267, 269] (Figure 3.2A). The ROSE algorithm was specifically written to identify super enhancers utilizing sequencing data from enhancer marking histone modifications and TFs. ROSE requires two inputs: a file containing previously identified constituent enhancers and sequencing reads for the factor of interest with a control. From these data sets ROSE is able to rank the constituent enhancers based upon their specific enrichment of reads for the factor of interest over control, then stitch the highly ranking constituent enhancers together based upon given separation requirements. The default distance, and the distance used in this study, is 12.5kb. In the initial description of this algorithm, this distance was found to be optimal for stitching together the closely spaced enriched regions with very high signal while not being so large as to stitch together the more widely spaced regions with lower signal [267]. Finally, the constituent enhancers and their associated factor enrichment are plotted in relationship to each other, and the data is scaled such that the x and y axis are from 0-1. The x-axis point for which a line with a slope of 1 is tangent to the curve is identified, and enhancers above this point are defined as SEs, and enhancers below that point are defined as typical enhancers. Using this algorithm, we discovered cell-line specific SEs in the first intron of *EGFR*, many of which rank highly amongst all the identified SEs in several of our cell lines (Figure 3.2B). Interestingly, the location and size of these SEs varied and were dependent upon the local enrichment of H3K27Ac ChIP-seq signal (Figure 3.2A).

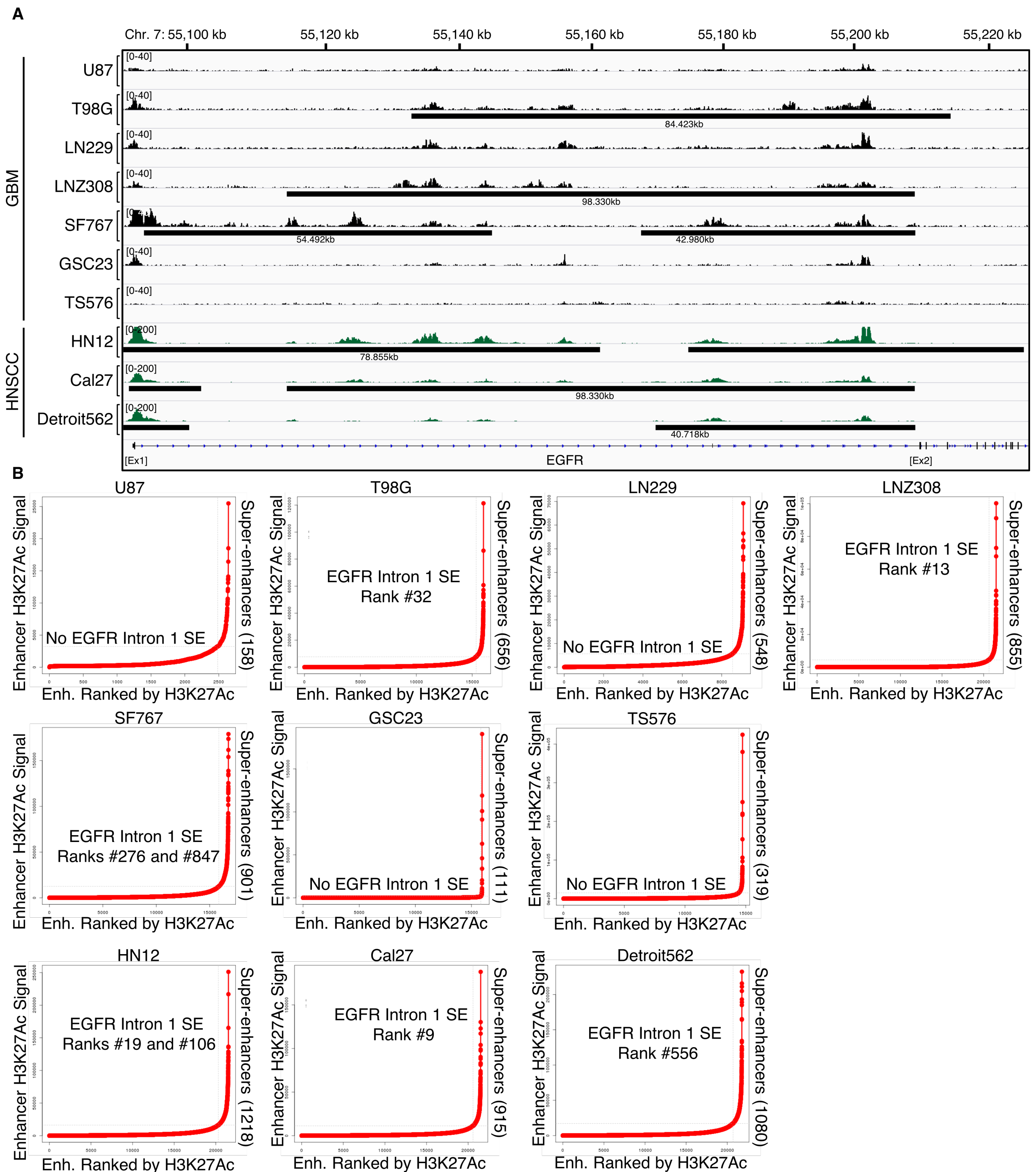


Figure 3.2: Identification of Super enhancers by ROSE. **A:** Super enhancer tracks in IGV overlaid with H3K27Ac tracks from measured cell lines at the EGFR gene. Super enhancers were identified by stitching together peaks of H3K27Ac histone marks with the ROSE software package. (Black bars) Super enhancers as called by ROSE. **B:** The EGFR intron 1 super enhancers were ranked based on H3K27Ac signal at the shown intron 1 super enhancers in the indicated GBM and HNSCC cell lines using the ROSE software package.

To determine if presence of putative enhancers in *EGFR* intron 1 was predictive of *EGFR* expression we first measured *EGFR* transcript levels in each GBM and HNSCC cell line (Figure 3.1B). Relative to the U87 glioma cell line, stark differences in *EGFR* expression were found between each measured cell line. Overall, measured glioma cell lines expressed significantly lower *EGFR* transcript than HNSCC cell lines. Among glioma cell lines, SF767 had the highest *EGFR* expression levels, while the HNSCC cell line HN12 exhibited the highest *EGFR* transcript amongst HNSCC lines and overall (Figure 3.1B). Because of these significant differences in *EGFR* transcript and the presence of *EGFR* SEs in intron 1 as identified by H3K27Ac enrichment, we hypothesized that *EGFR* transcript level may be correlated with SE-associated H3K27Ac enrichment. Due to cell line-specific presence and location of predicted *EGFR* super enhancers, we used total number of intron-1 mapped H3K27Ac ChIP-seq reads as a measure of enhancer presence and plotted these values against the relative fold change in *EGFR* expression. Analyzing the relationship by Spearman's correlation showed a highly significant correlation ($p = 0.0002$) (Figure 3.1C). To further confirm the conserved nature of these enhancers we utilized a recently published data set consisting of 44 patient-derived glioblastoma stem cells (GSCs) and 50 primary tumors [289]. The majority of the samples had gene expression, whole exome, copy number profiles, DNA methylomes and histone modification profiling completed, thus integration of these data sets was possible. We curated RNA sequencing (RNA-seq) and H3K27Ac ChIP-seq data for the 87 samples and ordered their acetylation tracks by the expression of *EGFR* as represented by fragments per kilobase of transcript per million (FPKM) (Figure 3.3). Those samples which expressed GBM to the highest degree are depicted at the top of the diagram, and maintain a consistent H3K27Ac signature reminiscent of the cell lines we have measured. As *EGFR* expression levels decrease, so too do intron 1 H3K27Ac reads (Figure 3.3). This extensive dataset further confirms the conservation of putative enhancers within *EGFR* intron 1 and argues for their regulatory potential both *in vivo* and *in vitro*. Together, these results identify regions containing characteristics of SEs in the first intron of *EGFR* and suggests that activity of these putative enhancers is important for high levels of *EGFR* transcript.

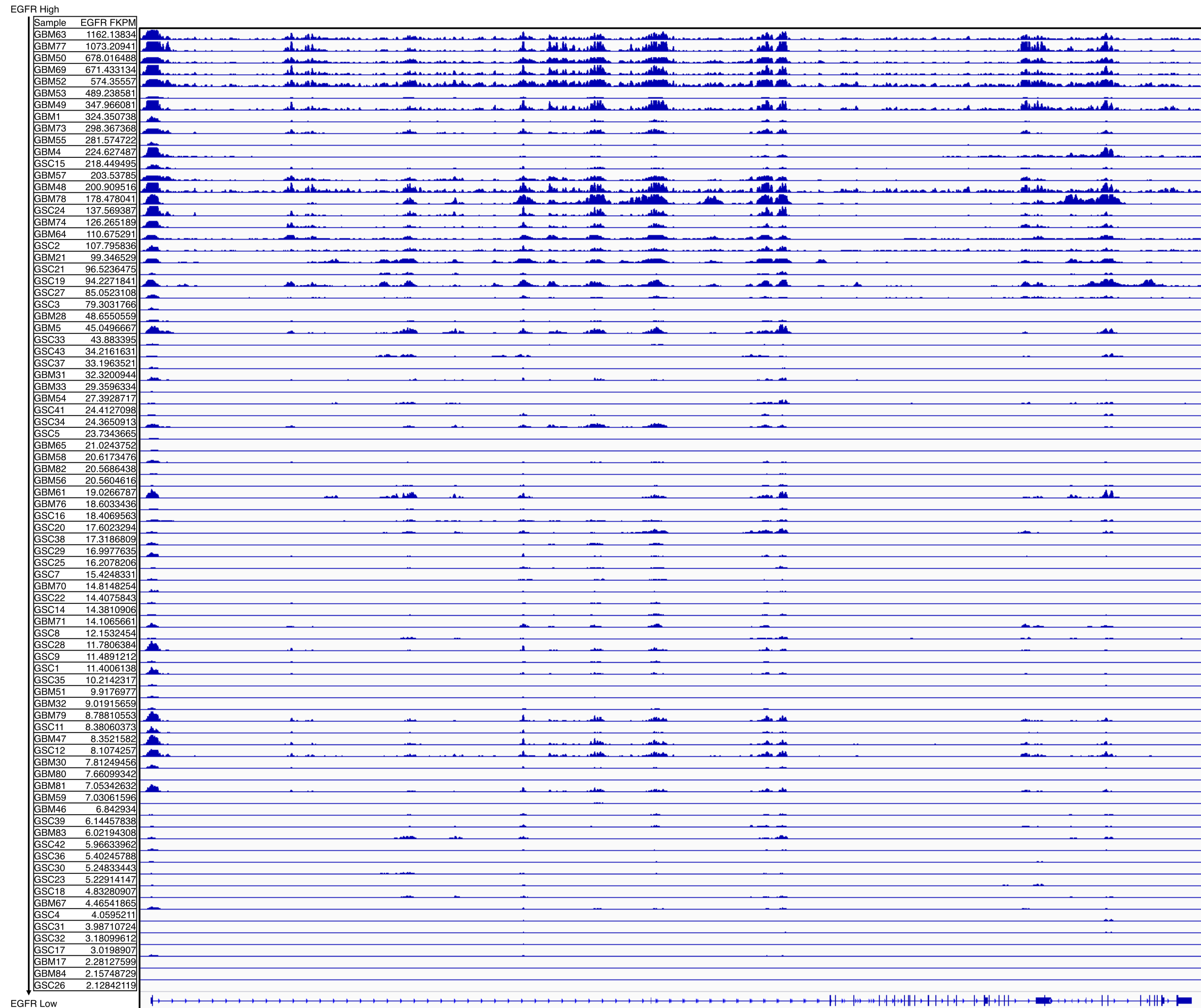


Figure 3.3: GSCs and GBM primary tumors confirm putative *EGFR* intron 1 enhancer presence. H3K27Ac ChIP-seq and RNA-seq from 87 *in vivo* and *in vitro* GBM samples were accessed from publicly available databases. ChIP-seq tracks were ordered from most to least *EGFR* expression as measured by FPKM determined by RNA-seq from matched tumor samples.

Two Critical Constituent Enhancers Reside in the First Intron of *EGFR*

As introduced earlier, it has been reported that SEs are congregations of active constituent enhancers (CE). To determine which CEs of the identified SEs are active, we performed a luciferase reporter assay utilizing regions which were identified as active by H3K27Ac. We segmented regions which exhibited highly conserved H3K27Ac ChIP-seq and ATAC-seq signals into 2kb segments (Figure 3.4A, Figure 3.5A). Each segment was then measured for enhancer activity by *in vitro* bioluminescence in each of the 10 cell lines (Figure 3.4B, Figure 3.5B). Measuring activity in each cell line is critical due to differential binding of tissue-specific or disease-specific TFs which may be active or inactive within each individual cell line [269, 270]. Regions which exhibited conserved luciferase expression included 1, 3, and 16-19 (Figure 3.4B, Figure 3.5B). Interestingly, activity of segments 8 and 9 were HNSCC specific (Figure 3.5B), while segments 1 and 3 were more GBM specific (Figure 3.4B). Segments 16-19 consistently enhanced luciferase activity in both tumor models (Figure 3.4B, Figure 3.5B). Each luciferase reporter segment resides in regions which have been identified as enhancers, yet some segments enhance luciferase expression to a larger extent than others. We hypothesized that the specific enrichment of H3K27Ac at an enhancer segment would correlate to luciferase expression in the matched cell line. Indeed, plotting the normalized luciferase intensity against the average H3K27Ac read intensity for each segment reveals a highly significant relationship ($p < 0.001$, Spearman's correlation) for both GBM and HNSCC cell lines (Figure 3.4C, Figure 3.5C). Combining H3K27Ac presence by ChIP-seq, open chromatin accessible regions by ATAC-seq, and functional activity as defined by our luciferase system in both tumor types, we define two distinct CEs. Specifically, we combined segments 1-3 into an approximately 6kb region which we have termed Constituent Enhancer 1 (CE1), and combined segments 16-19 into an approximately 8kb region which we have termed Constituent Enhancer 2 (CE2) (Figure 3.4A, Figure 3.5A). These results define two conserved putative CEs within *EGFR* intron 1 and establish a relationship between histone acetylation at these regions and enhancer activity as measured by *in vitro* bioluminescence.

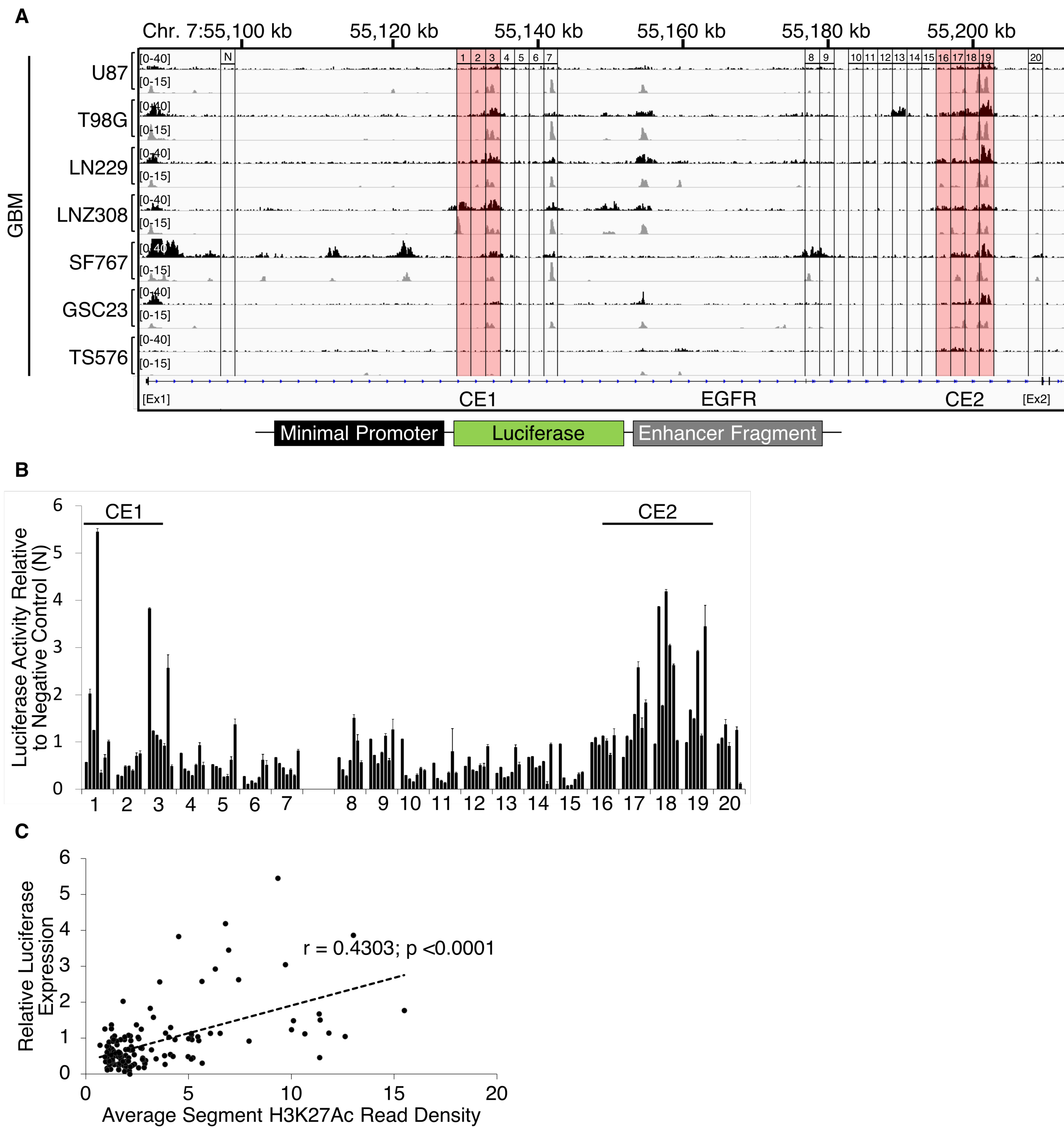


Figure 3.4: Identification of critical constituent enhancers in *EGFR* intron 1 of GBM. **A:** (Top) Schematic of positioning of enhancer segments 1-20 and the negative control region analyzed in the pGL4.24 vector. CE/CE2 (pink outline) are highlighted. (Bottom) Schematic of positioning of the enhancer segments in the pGL4.24 construct. **B:** Luciferase activity in GBM (cell lines left to right: U87, T98G, LN229, LNZ308, SF767, GSC23, TS576) cell lines after transfection with pGL4.24 constructs containing cloned fragments of *EGFR* intron 1. A negative control region 10kb downstream of the *EGFR* promoter was used for normalization. **C:** Relative luciferase expression for P1-20 was plotted against the average H3K27Ac read density for each individual segment. The relationship was analyzed using Spearman's rank order correlation.

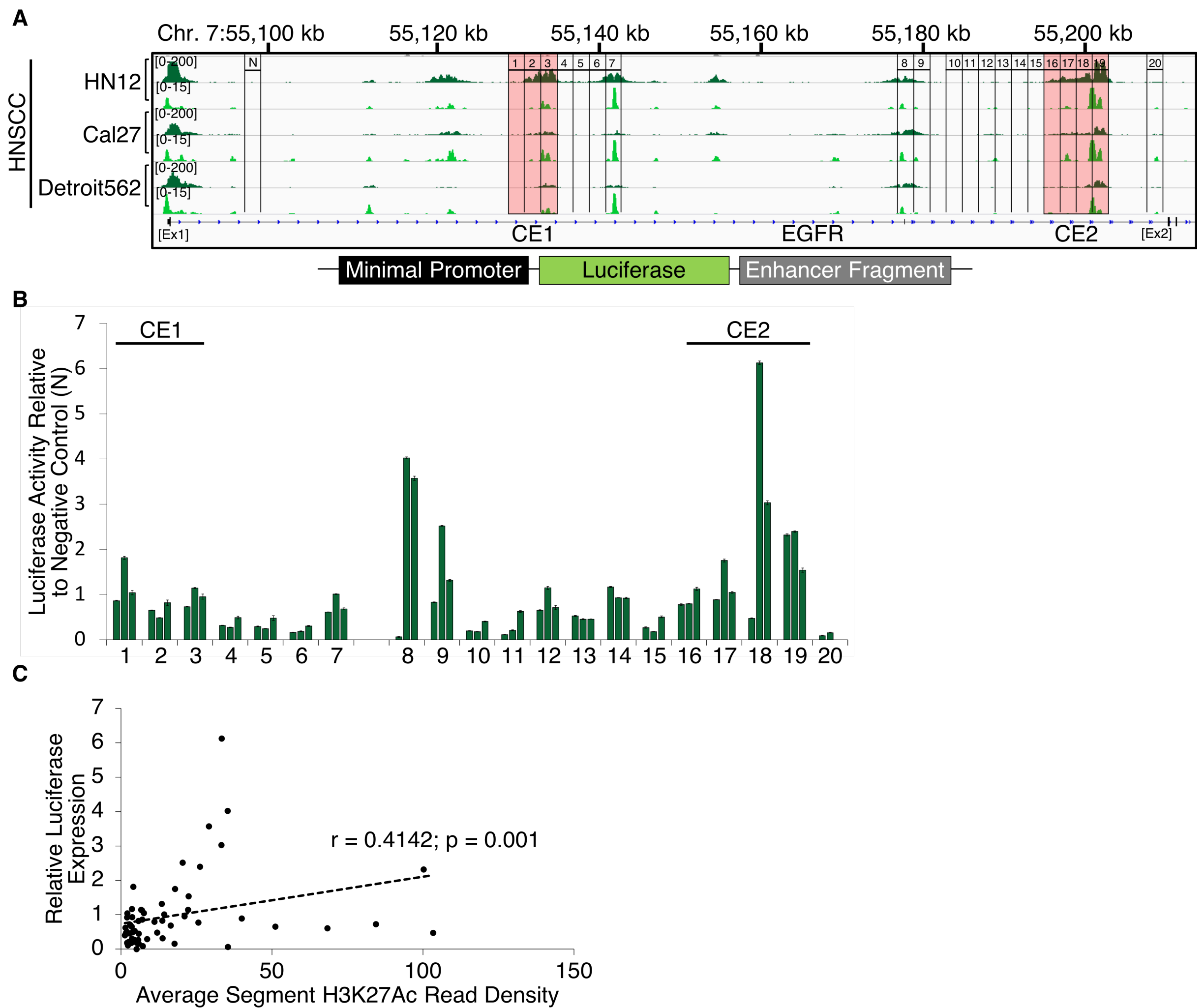


Figure 3.5: Identification of critical constituent enhancers in *EGFR* intron 1 of HNSCC. **A:** (Top) Schematic of positioning of enhancer segments 1-20 and the negative control region analyzed in the pGL4.24 vector. CE/CE2 (pink outline) are highlighted. (Bottom) Schematic of positioning of the enhancer segments in the pGL4.24 construct. **B:** Luciferase activity in HNSCC cell lines (left to right: HN12, Cal27, Detroit562) after transfection with pGL4.24 constructs containing cloned fragments of *EGFR* intron 1. A negative control region 10kb downstream of the *EGFR* promoter was used for normalization. **C:** Relative luciferase expression for P1-20 was plotted against the average H3K27Ac read density for each individual segment. The relationship was analyzed using Spearman's rank order correlation.

Significant Interactions Between CE1, CE2 and the *EGFR* Promoter

As introduced earlier, a critical characteristic of enhancers is their ability to make contact with one or more gene promoters through long-range interactions [268]. To test if the enhancer regions CE1 or CE2 interact with the *EGFR* promoter we performed a simple 3C assay. As opposed to high throughput methods which measure the interactions of a single region with the rest of the genome (circular chromosome conformation capture [4C] [295]) or of all regions of the genome with each other (Hi-C) [296], 3C measures the relative interaction frequency of two specific regions. 3C is based on formaldehyde crosslinking of interacting chromatin segments, followed by restriction digestion and ligation of crosslinked fragments. Ligation products are subsequently analyzed by PCR using primers specific for the restriction fragments of interest. Detection of a ligation product between two segments does not reveal much, instead to identify a chromatin loop between two segments, it needs to be demonstrated that they interact more frequently with each other than with neighboring DNA fragments [297]. To accurately identify 3D interactions between CE1 and CE2 and the *EGFR* promoter, qPCR primers were designed around the EcoRI sites in CE1 (Figure 3.6A, F1 to F5) and CE2 (Figure 3.6A, F6 to F10) and nearby the *EGFR* promoter. Since we hypothesized that an increased interaction frequency would correlate with increased transcript levels, we chose the cell lines with the highest *EGFR* expression levels in both tumor types and compared them against a cell line with no *EGFR* expression. Compared to TS576, SF767 cells exhibited low levels of interaction with CE1 and CE2 (Figure 3.6B); however, significantly stronger interactions were identified at F2 and F4 of CE1 and 4 out of 5 regions of CE2 in the HN12 cell line (Figure 3.6B). The primers which produced the strongest signals of interaction reside in highly acetylated regions in HN12 cells (Figure 3.6A), thus we hypothesized that H3K27Ac enrichment and 3C-interaction were correlated. Correspondingly, we identified a significant correlation between H3K27Ac peak intensity and interaction frequency at these measured regions (Figure 3.6C). Together, the data thus far indicates that CE1 and CE2 have characteristics ascribing them to active enhancers: surrounded by nucleosomes with high H3K27Ac, open chromatin, transcriptional enhancement, and interaction with a promoter.

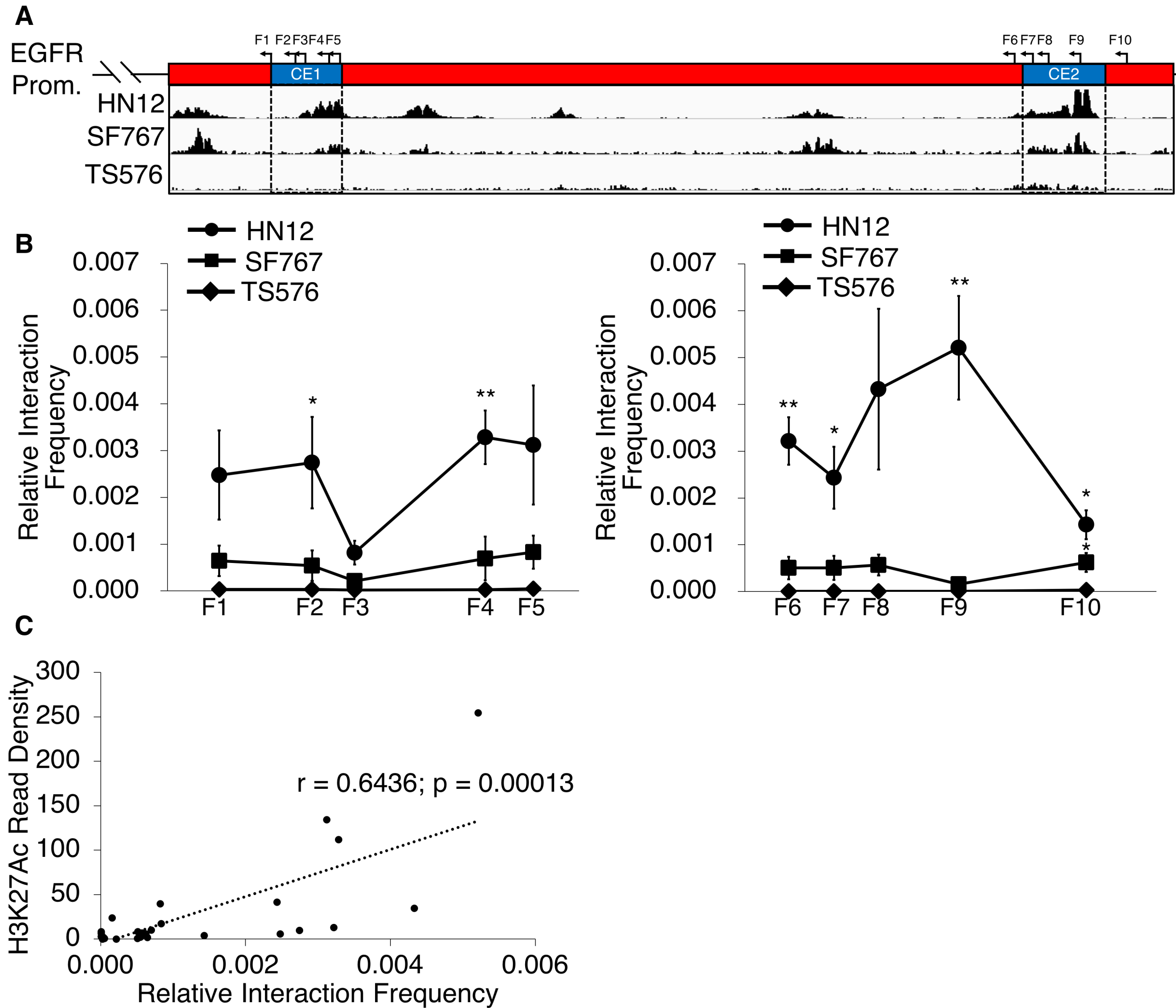


Figure 3.6: Enhancer-promoter interaction by chromatin conformation capture (3C). **A:** Schematic showing position of 3C primers relative to CE1 and CE2. 3C qPCR was done in combination with a forward primer in the *EGFR* promoter region. **B:** Relative interaction frequency of each restriction fragment (F1-10) was calculated as described in the experimental procedures and was plotted against genomic location of the cutting site of each fragment. Values reported were derived from three biological repeats (*$p < 0.05$, Student's t test). **C:** H3K27Ac read density at the primer site was plotted against the relative interaction frequency. The relationship was analyzed using Spearman's rank order correlation.

3.3 Discussion

The rules governing the location of an enhancer with respect to its target promoter has been a topic in which there is little agreement. Some studies indicate that the closer an enhancer is to a promoter, the more likely it is to interact with and enhance transcription of that gene. As such, many studies use the most proximally located gene to putatively identify the interaction partner of an enhancer in genome-wide analyses [183, 298]. Still other studies have shown no preference for proximity [299] and that interactions can occur across TAD boundaries [201], at vast distances [202], and even inter-chromosomally [300]. What is not up for debate, however, is the relationship between transcriptional enhancers and the first intron of a gene. A multitude of studies have identified enhancer elements within the first intron of a gene, and have shown them to be functionally relevant to the activity of the gene in which they are located [301–308]. Additionally, there appears to be a correlation between intron size and *cis*-regulatory DNA with the first intron often being the largest and most predisposed to having regulatory DNA contained within [309–312]. Though a direct relationship between intron size and gene expression could not be confirmed genome wide, 5' UTR length was positively correlated with expression in specific gene families [313], and as many as 35% of genes have introns in their 5' UTR [314]. Whole genome analysis of enhancer elements has further strengthened the relationship between introns and regulatory DNA, showing over 55% of functional enhancer elements in *Drosophila Melanogaster* are located in introns and over 37% in the first intron [179]. Thus, it is not surprising that there is evidence for potent enhancers in the first intron of *EGFR* in the data we have presented here.

The characterization of super enhancers that are critical for the expression of oncogenes or key development genes spawned a new appreciation for the impact of transcriptional enhancers, and rapidly expanded the research into drug development to target this small subset of enhancers. The identification of these SEs is far from standardized. The initial SE study used a combination of OCT4, SOX2, Nanog and Mediator [315] binding to define a highly enriched enhancer population in ESCs, however subsequent studies have shown other combinations of factors to be sufficient. Mediator and BRD4 plus H3K4me1 and H3K27Ac was applied widely to cancer cells

for SE identification [269], however the algorithms written to identify SEs only require input from a single factor (e.g. H3K27Ac) and SEs are still able to be robustly identified from this single factor [316]. This discordance in the field on the necessary factors makes defining a SE difficult and potentially diminishes the relevance of a study if the improper factor is used. Regardless, extensive functional validation is still required to ascribe a true function to a SE.

3.4 Materials and Methods

Cell Culture.

GBM cell lines U87 (RRID:CVCL_0022), T98G (RRID:CVCL_0556), and LN229 (RRID:CVCL_0393) were purchased from ATCC. LNZ308 (RRID:CVCL_0394) was provided by Dr. Erwin Van Meir (Emory University, Atlanta, GA). SF767 (RRID:CVCL_6950) was provided by Dr. Mitch Berger (UCSF Brain Tumor Center, San Francisco, CA). Head and neck cell lines HN12 (RRID:CVCL_5518), Cal27 (RRID:CVCL_1107) and Detroit562 (RRID:CVCL_1171) were provided by Dr. Silvio Gutkind (UCSD Moores Cancer Center, San Diego, CA). Adherent cell lines were maintained in DMEM (Hyclone, #SH30022.01) supplemented with 10% fetal bovine serum (Atlanta Biologicals, #S12450) and 1% penicillin-streptomycin (Gibco, #15140-122) and grown as adherent cultures. GBM neurosphere cell lines GSC23 (RRID:CVCL_DR59, Dr. Fred Lang, University of Texas MD Anderson Cancer Center, Houston, TX) and TS576 (Dr. Cameron Brennan, Memorial Sloan Kettering Cancer Center, New York, NY) were maintained in DMEM/F12 (Gibco, #11320-033) supplemented with B27 supplement (Gibco, #12504-044) and 1% penicillin-streptomycin and grown in suspension. Mycoplasma testing was performed with the PlasmoTest kit (InvivoGen, #rep-pt1) and found to be negative. All experiments are performed within 10 passages of the original frozen stock or post-manipulation.

Luciferase Reporter Assays.

DNA fragments tested in the luciferase reporter assay were cloned from human genomic DNA. PCR products were cloned downstream of firefly luciferase in the pGL4.24 minimal promoter vector (Promega, #E8421) using the SalI (NEB, #R3138S) site. Constructs were sequence

confirmed by Sanger sequencing using the pGL4.24-R primer. pMIEG3-JunDN was a gift from Alexander Dent (RRID: Addgene_40350). pMIEG3-Empty was created by removing the JunDN sequence by EcoRI digestion. For each transfection reaction, 100ng control plasmid expressing Renilla luciferase (Promega, #E2241) and 1µg Firefly luciferase construct were co-transfected with Lipofectamine 2000 (ThermoFisher, #11668030) into 2 x 10^5 cells in a 12-well plate well. After 24 hours, cells were collected in 1X PLB. Luciferase activity was measured by the Dual-Luciferase Reporter Assay System (Promega, #E1910) on a Tecan Spark 10M with injection control. Transfection efficiency was controlled for by dividing Firefly luminescence by Renilla luminescence, and final activity was normalized to a negative control.

Quantitative real-time PCR.

RNA was extracted with the RNeasy Plus kit (Qiagen, #74134) according to the manufacturers instructions. Reverse transcription of mRNA was performed using 1µg RNA with the iScript Reverse Transcription Supermix (BioRad, #1708841). For real-time PCR analysis, 5µl of cDNA (50ng of starting RNA) was amplified per reaction using the iTaq Universal SYBR Green Supermix (Bio-Rad, #1725124) and the Bio-Rad CFX96 qPCR system. Fold change analysis was performed using the $2^{-\Delta\Delta Ct}$ method and normalized as indicated.

Chromatin Conformation Capture (3C).

The experiment was performed as described [297] with the following modifications. Nuclei were treated with 1000U EcoRI (NEB, #R3101S) at 37°C overnight. 100U T4 enzyme (NEB, #M0202S) was added to digested nuclei and incubated at 16°C for 4 hours. Another 100U T4 enzyme was added and nuclei were incubated with rotation at 4°C overnight. 150ng of ligated DNA was quantified in triplicate by TaqMan real-time PCR using the PrimeTime Gene Expression Supermix (IDT, #1055772). Control 3C template was generated by using two bacterial artificial chromosomes (BAC) encompassing the entire EGFR gene, RP11-159M24 and RP11-148P17, were purchased from the Childrens Hospital Oakland Research Institute (CHORI). Equimolar of the two BACs were digested with EcoRI and ligated. The ligation product from BAC control was used for normalization. The relative interaction frequency was calculated as:

$2^{\text{Ct (BAC)-Ct (3C)}}$.

Chromatin Immunoprecipitation.

Chromatin immunoprecipitation was performed as described previously [317] with the following modifications. Chromatin was sheared in diluted lysis buffer to 200-500bp using a Covaris M220 Focused-Ultrasonicator with the following parameters: 10 minutes, peak incident power 75, duty factor 10%, 200 cycles/burst. Antibodies for ChIP were obtained from commercially available sources: anti-H3K27Ac (Active Motif, #39133). 5% of the chromatin was not exposed to antibody and was used as control (input). For ChIP-seq samples, after DNA purification ChIP-seq DNA libraries were prepared with either the TruSeq ChiP Library Prep Kit (Illumina, #IP-202-1012) or the Accel-NGS 2S Plus DNA Library kit (Swift Bioscience, #21024) and sequenced using 75 bp single-end sequencing on an Illumina Hi-seq 4000.

ChIP-seq analysis.

Raw reads from Illumina Hi-seq 4000 were aligned to the human genome (hg19) using Bowtie2 software [318] with default parameters. Non-uniquely mapped reads were removed with awk command and unique reads were used for secondary analysis. Genome-wide read coverage was calculated by igvtools count and visualized using the Integrative Genomics Viewer (IGV). Peaks were called by first creating TagDirectories of H3K27Ac and Input samples using the makeTagDirectory command in the HOMER (Hypergeometric Optimization of Motif EnRichment) suite of tools [319]. Peaks were identified from the H3K27Ac tag directories using the input directories for normalization with the findPeaks command from HOMER with the following parameters: -style histone -size 250 -minDist 250.

Super enhancer identification.

The ROSE (Rank Ordering of Super-Enhancers) algorithm was downloaded and run from the command line [267, 269]. Bam files of H3K27Ac and Input ChIP-seq reads for each cell line were generated from raw files with bowtie2 and sorted with the samtools command samtools sort. GFF files were created from peak files for each cell line using the awk command. ROSE

was called using the command python ROSE_main.py with the following parameters: -g HG19 -i [GFF_File] -r [H3K27Ac_ChIP.sorted.bam] -o [Output_Folder] -c [Input.sorted.bam] -t 3000.

ATAC-seq.

Approximately 50,000 permeabilized nuclei were transposed using Tn5 transposase (Illumina, #FC-121-1030) as described previously [205]. Libraries were amplified using NEBNext High-Fidelity 2X PCR Master Mix (NEB, #M0541) with primer extension at 72°C for 5 min, denaturation at 98°C for 30s, followed by 8 cycles of denaturation at 98°C for 10s, annealing at 63°C for 30s and extension at 70°C for 60s. Each library was size selected and sequenced on an Illumina NextSeq500 or HiSeq4000 to a depth of $\geq$ 20 million usable reads pairs. Sequencing runs that did not meet the read pair threshold were sequenced again, and all replicates were pooled for analysis.

ATAC-seq analysis.

Nextera adapters were trimmed from the raw fastq files by using cutadapt [320] with parameters -m 5 -e 0.10 -a CTGTCTCTTATA -A CTGTCTCTTATA and then aligned to human reference genome hg38 using bowtie2 with parameters -X2000 –mm –local. Next, the improperly mapped, poorly mapped and unmatched reads were filtered from the resultant raw bam files using samtools view with parameters -F 1804 -f 2 -q30. Duplicates were marked with Picard with the command picard MarkDuplicate and removed by samtools view. Final bam files were generated after removing mitochondrial reads by awk command. Replicate Bam files were merged using samtools merge and converted to tagAlign format using the bedtools command bamtobed with parameters -bedpe -mate1 -I and awk. To account for the cutting offset of Tn5 transposase, mapping position was shifted using awk. Peaks were called from tagAlign files using the Model-based Analysis of ChIP-Seq (MACS2) [321] callpeak command with parameters -g hs -p 0.01 –nomodel –shift -75 –extsize 150 -B –SPMR –keep-dup all –call-summits. Bedgraph files output from MACS2 were converted to BigWig files using bedGraphToBigWig and visualized using IGV.

Data Access.

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO) under accession number GSE128275. Accession numbers for publicly available data accessed are as follows: ChIP-seq (GSE119755), RNA-seq (GSE119834).

3.5 Acknowledgements

Chapter 3 in part, has been submitted for publication of the material as it may appear in Molecular Cancer Research; **Jameson, N.M.**, Ma, J., Benitez, J., Izurieta, A., Han, J.Y., Mendez, R., Parisian, A., Furnari, F., AACR Publications, 2019. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Genetic perturbation of *EGFR* intron 1 enhancers

4.1 Introduction

As has been thoroughly established previously, it is essential that putative transcriptional enhancers be extensively validated to determine their true function in a particular cell type. Many of the most powerful and versatile techniques utilize the relatively new discovery of the Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR) mechanism and the CRISPR-associated (Cas) proteins. The mechanism by which CRISPR/Cas works was first described in the prokaryotic *Streptococcus thermophilus*. This study showed experimentally that CRISPR systems are prokaryotic adaptive immune responses which work by capturing small DNA sequences from invading viruses and bacteriophages, and use them to create DNA segments known as CRISPR arrays [322]. The CRISPR array serves as a memory bank within the bacteria, allowing it to produce RNA segments from the array which recognize and target invading virus DNA through double strand breaks (DSB) carried out by the nuclease ability of a Cas protein [323, 324]. While interesting in the context of bacterial immunity, it was the adaptation of CRISPR/Cas to eukaryotic cells which provided the platform for the CRISPR revolution in biomedical research. Two studies published at nearly the same time discovered the ability to

guide the *Streptococcus pyogenes* Cas9 (spCas9) protein to a genomic region using a CRISPR guide RNA (gRNA) [325], which recognizes a protospacer-adjacent motif (PAM) and cuts specifically at that PAM sequence [326, 327]. Targeting this system to a genomic region facilitated a double strand break at single nucleotide resolution, and resulted in DNA repair through either non-homologous end joining (NHEJ) or homology directed repair (HDR) [327].

CRISPR-mediated HDR or NHEJ.

The utilization of either of these repair pathways has advantages depending on the purpose of the experiment. In most eukaryotic cells, the NHEJ pathway generates insertions and deletions during DSB repair. However, in the presence of a DNA template with homology to the sequences flanking the DSB location, HDR can seal the DSB in an error-free manner [328]. The efficiency of HDR is determined by the concentration of donor DNA present at the time of repair, the length of the homology arms of the donor DNA, the cell cycle, and the activity of the endogenous repair systems [329]. HDR methods allow for very precise mutations, and shows incredible promise for introducing single base substitutions to correct disease phenotypes [330,331]. Though CRISPR/Cas9 editing is relatively easy and incredibly powerful, significant concerns have been raised about its use in a clinical setting due to the evidence of off-target effects [332, 333]. The ethics of using CRISPR in a human setting is under hot debate, and its misuse cannot be better represented than recent reports of so-called CRISPR-babies in China. Though based on science supporting the mutation of *CCR5* can promote resistance to human immunodeficiency virus (HIV) infection [334], editing of this gene in human children resulted in considerable backlash from the scientific community as being premature and irresponsible [335, 336]. Though not as precise, NHEJ-mediated repair is particularly useful when a precise mutation is not necessary as it is an error-prone repair mechanism that often leads to insertions or deletions (indel). These indels can cause frameshift mutations, premature stop codons, and/or nonsense-mediated decay (NMD) to the target gene, which results in loss-of-function. Additionally, NHEJ allows for multiple gRNAs to be introduced at the same time to make genomic deletions [337]. This process however is relatively inefficient, and there is an inverse relationship between deletion frequency and deletion size [338].

CRISPR/Cas9 for gene regulation.

In addition to its nuclease activity, Cas9 can serve as a unique platform to recruit protein and RNA factors to a targeted DNA site, and it has been engineered into powerful tools for sequence-specific gene regulation. To achieve this, transcriptional activators and repressors are fused to a catalytically dead Cas9 (dCas9) [339]. dCas9 maintains its ability to bind both the gRNA and targeted DNA, but it lacks nuclease activity and can thus serve as a sequence-specific RNA-guided DNA-binding platform. In bacterial cells, but less so in mammalian cells, dCas9 alone can efficiently inhibit the transcription of targeted genes through steric hindrance of transcriptional machinery in a process known as CRISPR interference (CRISPRi) [339]. Earlier studies using other DNA targeting proteins fused to transcription effector domains [340] showed proof-of-principle that targeting these domains could affect transcription either positively or negatively depending on the domain. Due to the ease of fusion of dCas9 to these effector domains, it was a natural progression to apply this technology to CRISPRi and CRISPR activation (CRISPRa). The first study to describe CRISPRi and CRISPRa utilized dCas9 fused to the Krüppel-associated box (KRAB) domain of Kox1 (dCas9-KRAB) to mediate gene repression, and dCas9 fused to four copies of the transcriptional activator VP16 (dCas9-VP64) to mediate gene activation [341]. Highly specific and stable activation or repression of the target gene was observed, and a linear relationship between the level of expression of the gRNA and the level of gene activation/repression was discovered [341]. Genome wide application of these techniques have also allowed for comprehensive mapping of complex pathways [342]. Taken together, the CRISPR toolkit represents a powerful methodology for performing functional genomics.

4.2 Results

Deletion of CE1 and CE2 by CRISPR/Cas9 Results in Reduced *EGFR* Expression

As introduced in Chapter 2, the functional testing of enhancers is critical for determining their phenotypic effects in a cell type of interest. Many groups have chosen to use the CRISPR/Cas9 system to carry out these functional testing assays. Examples of this strategy are abundant in both mouse and human models. The simultaneous use of two gRNAs to delete

defined genomic regions has been used, for example, to measure the regulatory contribution of enhancers near oncogenes [269, 315, 316], at the Sox2 locus in ESCs [343, 344], at the Myc locus [345], at the HER2 locus [346], at the MGMT locus [347] and at the androgen receptor locus [348]. To directly assess if CE1 and CE2 were essential for *EGFR* expression, we used the CRISPR/Cas9 system to delete CE1 (hg38, chr7: 55,060,994-55,066,815) and CE2 (hg38, chr7: 55,127,646-55,135,347). To minimize clonal effects, we selected at minimum 2 homozygous clones for each deletion and mixed them at equal numbers (Figure 4.1).

To test how enhancer loss effects cells with different *EGFR* expression levels we made these deletions in two GBM cell lines and two HNSCC cell lines which had either high (SF767, HN12) or low (U87, Cal27) relative *EGFR* expression. Single enhancer deletions were generated with a dual-guide deletion strategy (Figure 4.1) and the presence of editing at none, one, or both alleles was confirmed by genotyping PCR using a combination of primers within and outside the deleted region (Figure 4.1, Figure 4.2A). Compared to parental cell lines, *EGFR* transcript was significantly decreased in each deletion (Figure 4.2B). Interestingly, deletions of either CE1 or CE2 had different effects depending on the cell line. Though each cell line exhibited the strongest enrichment for H3K27Ac in CE2 (Figure 3.1A), SF767 and Cal27 showed significantly more repression of *EGFR* transcript with the loss of CE1. These cell-type specific differences in transcript levels between CE1 and CE2 indicate there may be differential utilization of either CE1 or CE2 in different cell lines.

Figure 4.1: CRISPR/Cas9-mediated deletion of CE1 and CE2. Schematic outlining the deletion strategy. Guides were designed flanking the intended deletion regions. PCR primers were designed inside and outside the deletion region to test for the presence of editing at neither, one, or both alleles.

Previous studies which deleted individual constituent enhancers within SEs revealed that enhancer activity is mostly dependent on a few constituents that activate transcription [270, 273, 316, 349]. Additionally, partial redundant control of a gene by multiple CEs within a single SE has been observed [349]. To evaluate if there was partially redundant control of *EGFR* transcription by either CE1 or CE2, we performed a second round of CRISPR/Cas9 editing on the homozygous edited populations. These deletions utilized the same set of gRNAs used for single deletion, and again a minimum of 2 homozygous clones for each deletion were isolated and mixed at equal numbers (Figure 4.3A). Compared to parental cell lines, *EGFR* transcript levels were most significantly decreased with loss of both enhancers (Figure 4.3B). Notably, the amplitude of *EGFR* transcript loss was greater in CE1$^{-/-}$ + CE2$^{-/-}$ when compared against CE1$^{-/-}$ or CE2$^{-/-}$ alone. Together, these results demonstrate both cell type-specific CE utilization as well as a cooperative relationship between the CE1 and CE2 whereby double enhancer deletion results in more significant deleterious effects than single deletions alone.

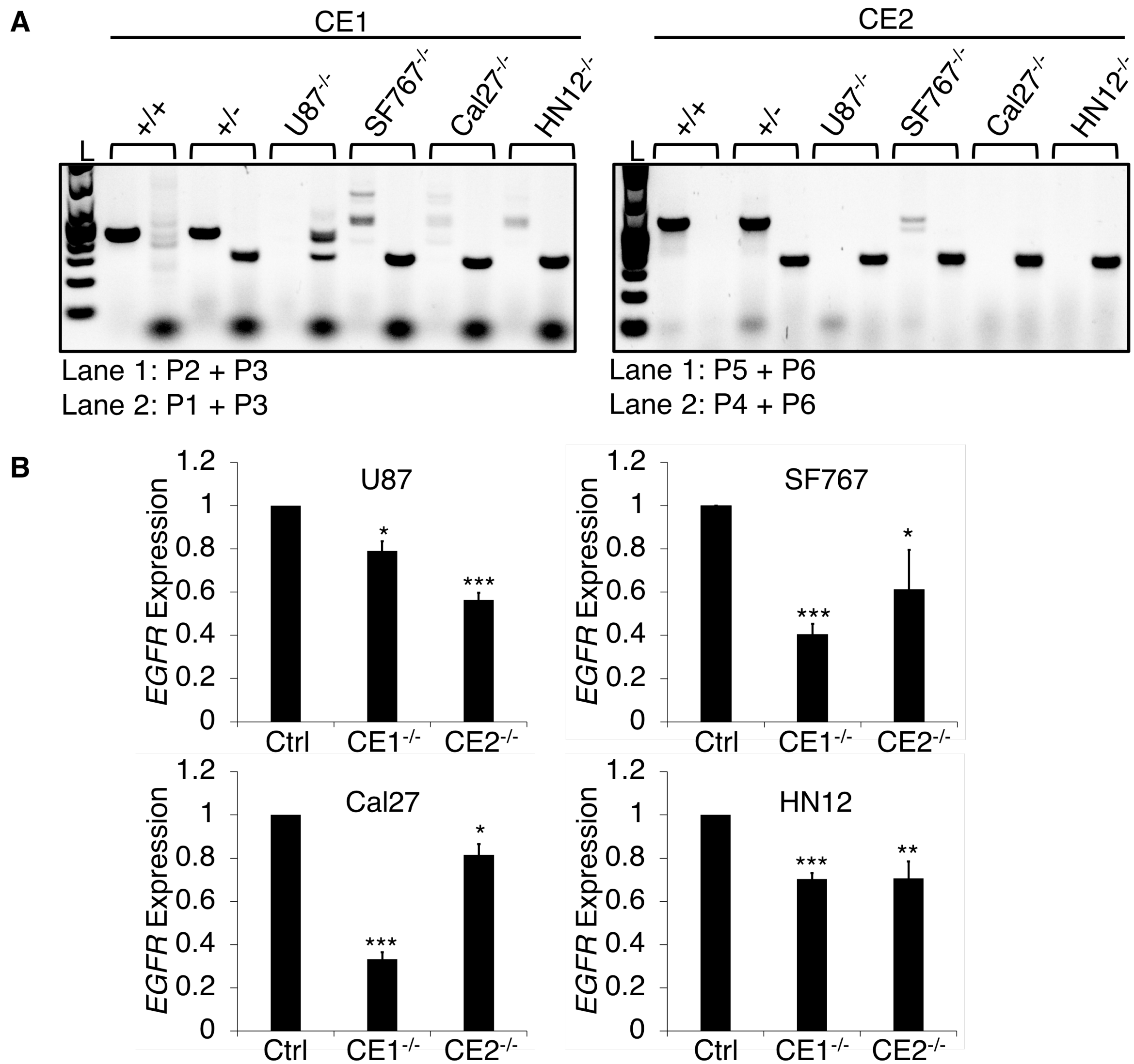


Figure 4.2: CRISPR/Cas9-mediated deletion of CE1 and CE2 in GBM and HNSCC cell lines reduces *EGFR* expression. **A:** Genotyping PCR for (left) CE1 and (right) CE2. Homozygous parental (lanes 1, 2) and heterozygous deleted (lanes 3, 4) are shown as PCR controls. Homozygous enhancer deletion (lanes 5-12) is shown for clone mixtures. **B:** *EGFR* expression in deleted cell lines were analyzed by RT-qPCR. *EGFR* transcript level was first normalized to *GAPDH* and subsequently calculated as fold change relative to parental. **B:** (*$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, $n = 3$ independent experiments, Students t test)

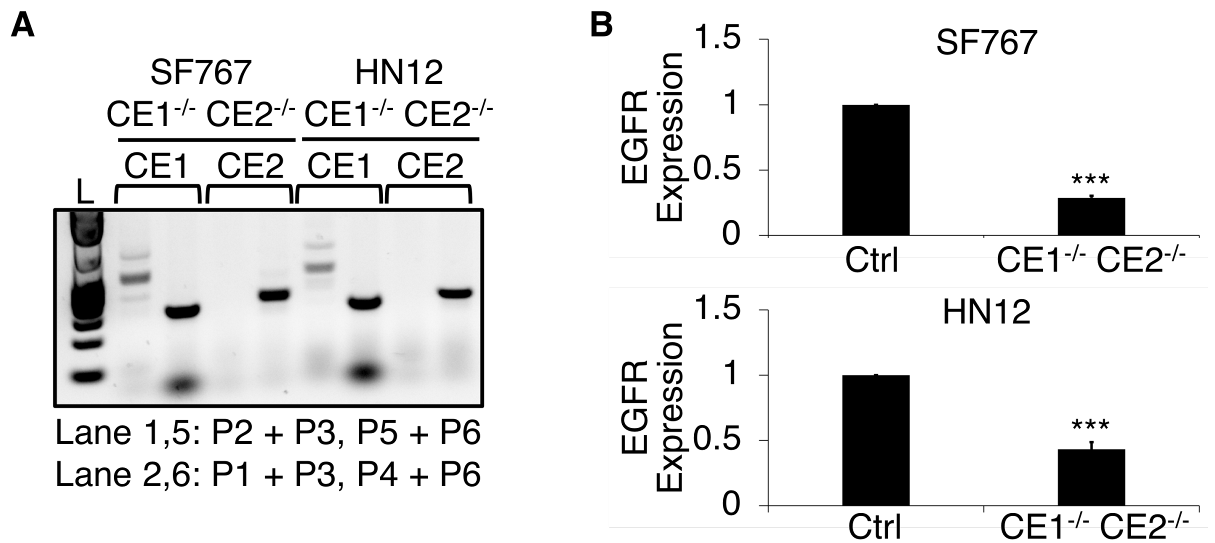


Figure 4.3: CRISPR/Cas9-mediated double deletion of CE1 and CE2 in GBM and HNSCC cell lines further reduces *EGFR* expression. **A:** Genotyping PCR for CE1 and CE2 deletion in HN12 and SF767 cells. Homozygous enhancer deletions are shown for clone mixtures. **B:** CE2 deletions were layered on top of CE1$^{-/-}$ cells in SF767 and HN12 cells. A minimum of 2 double-deleted homozygous clones were combined for downstream analysis. *EGFR* expression in double-deleted cell lines was analyzed by RT-qPCR. *EGFR* transcript level was first normalized to *GAPDH* and subsequently calculated as fold change relative to parental (*p < 0.05, **p < 0.005, ***p $<$ 0.0005, n = 3 independent experiments, Students *t* test).

Repression of H3K27Ac by dCas9-KRAB Decreases *EGFR* Expression

Recently, a mechanism known as enhancer hijacking has demonstrated that structural variation can significantly alter the enhancer-promoter contacts within a cell. For example, duplication, inversion or deletion of genomic regions containing oncogenes or regulatory elements in group 3 and 4 medulloblastoma rearranges the local chromatin structure such that oncogenes and enhancers that were once located distally are now located proximally to each other [350]. This phenomenon is further enforced by the identification of *de novo* 3D contact domains within the IGF2 gene following recurrent tandem duplications intersecting with a TAD boundary, mediating the formation of a novel interaction with a lineage-specific SE and facilitating high level gene activation [351]. To eliminate the possibility of structural variation being the root cause of *EGFR* expression loss in CRISPR/Cas9 deleted clones, we aimed to perform a CRISPRi experiment which would repress enhancer activity without genomic deletion. We hypothesized that histone de-acetylation would be sufficient for *EGFR* transcriptional repression due to the relationship between H3K27Ac and enhancer activity as demonstrated in Figure 3.1, Figure 3.4 and Figure 3.5. dCas9-KRAB is known to recruit endogenous chromatin modifying complexes to de-acetylate histones [352], therefore to test our hypothesis we targeted dCas9-KRAB to four regions within CE1 and five regions within CE2 with specific gRNAs in HN12 (Figure 4.4A) and SF767 (Figure 4.5A) cell lines. Additionally, we targeted a non-acetylated region within *EGFR* intron 1 as a negative control (Figure 4.4A, Figure 4.5A).

We first transduced each cell line with a lentivirus engineered to integrate dCas9-KRAB and a blasticidin selection gene [353]. Following selection, HN12 cells expressed high levels of dCas9 (Figure 4.4B), while SF767 cells expressed relatively less protein (Figure 4.5B). Because of the linear relationship between gRNA expression and target repression [341], we chose to utilize a lentivirus engineered to integrate a puromycin resistance gene as well as each specific gRNA driven by a U6 promoter [354]. Cell lines were selected for gRNA expressing cells with puromycin, creating cell lines with stable repression of the target region. To ensure successful targeting of dCas9-KRAB we confirmed de-acetylation at each region. Compared to an enhancer off-target (O-T) gRNA, the targeting of dCas9-KRAB resulted in significant decreases in H3K27Ac in

HN12 cells (Figure 4.4C). Interestingly, in HN12 cells compared to the O-T control, CE1.4 and CE2.4 gRNA targeted regions also had significantly decreased enrichment of H3K27Ac at the opposing targeted region (e.g. H3K27Ac loss at CE1.4 when 2.4 was targeted) (Figure 4.4C). This phenotype was not present in SF767 cells targeted with dCas9-KRAB. This cross-repression is likely not due to off-target effects or due to the large size of dCas9-KRAB protein complex [352], but may in fact be due to the cooperation between CE1 and CE2 at the *EGFR* promoter in the 3D chromosome, a hypothesis supported by the ACH hypothesis introduced in Chapter 3 [195]. To assess functional impact of de-acetylation at these regions, we measured *EGFR* transcript levels in each engineered cell line and observed significant decreases in transcript in 73% (8/11) of targeted regions in HN12 cells (Figure 4.4D). Significant repression of *EGFR* transcript levels was observed in 45% (5/11) of targeted regions in SF767 (Figure 4.5D), though less significantly, likely due to the decreased expression of dCas9-KRAB and subsequent reduction in repressive activity. Interestingly, in HN12 cells, targeting CE2 had an overall stronger repressive effect on *EGFR* transcript levels, with repression by gRNA targeting CE2.4 achieving a greater than 3-fold decrease in expression. In SF767 cells there was no observable preference for CE2. We performed double dCas9-KRAB repression by adding a second gRNA (CE1.4 + CE2.4) by a second round of lentiviral transduction and observed the most significant decrease in *EGFR* transcript levels, achieving an 8-fold decrease in expression compared to the off-target gRNA (Figure 4.4D). Similarly, in SF767 cells the strongest repression was observed by combining CE1.4 and CE2.4 (Figure 4.5D).

Finally, the effect of dCas9-KRAB-mediated *EGFR* repression on cell proliferation was assessed in HN12 cells by ATPlite assay. To enhance the effect of *EGFR* repression and demonstrate *EGFR* dependence, we chose to perform this assay at low serum. Serum is known to contain many growth factors which can activate other EGFR-parallel growth pathways and mask the effects of *EGFR* depletion. At low (0.5%) serum, the relative proliferation of all on-target *EGFR*-repressed cells was significantly inhibited (Figure 4.4E). Importantly, cell lines with stronger repression of *EGFR* exhibited the most significant negative effect on proliferation over time. These results indicate that *EGFR* transcriptional changes in enhancer deleted regions (Figure 4.2, Figure 4.3) are

not due solely to structural alteration within the first intron. Moreover, these results demonstrate that loss of H3K27Ac at the identified *EGFR* enhancers is sufficient for significant decreases in *EGFR* transcript levels.

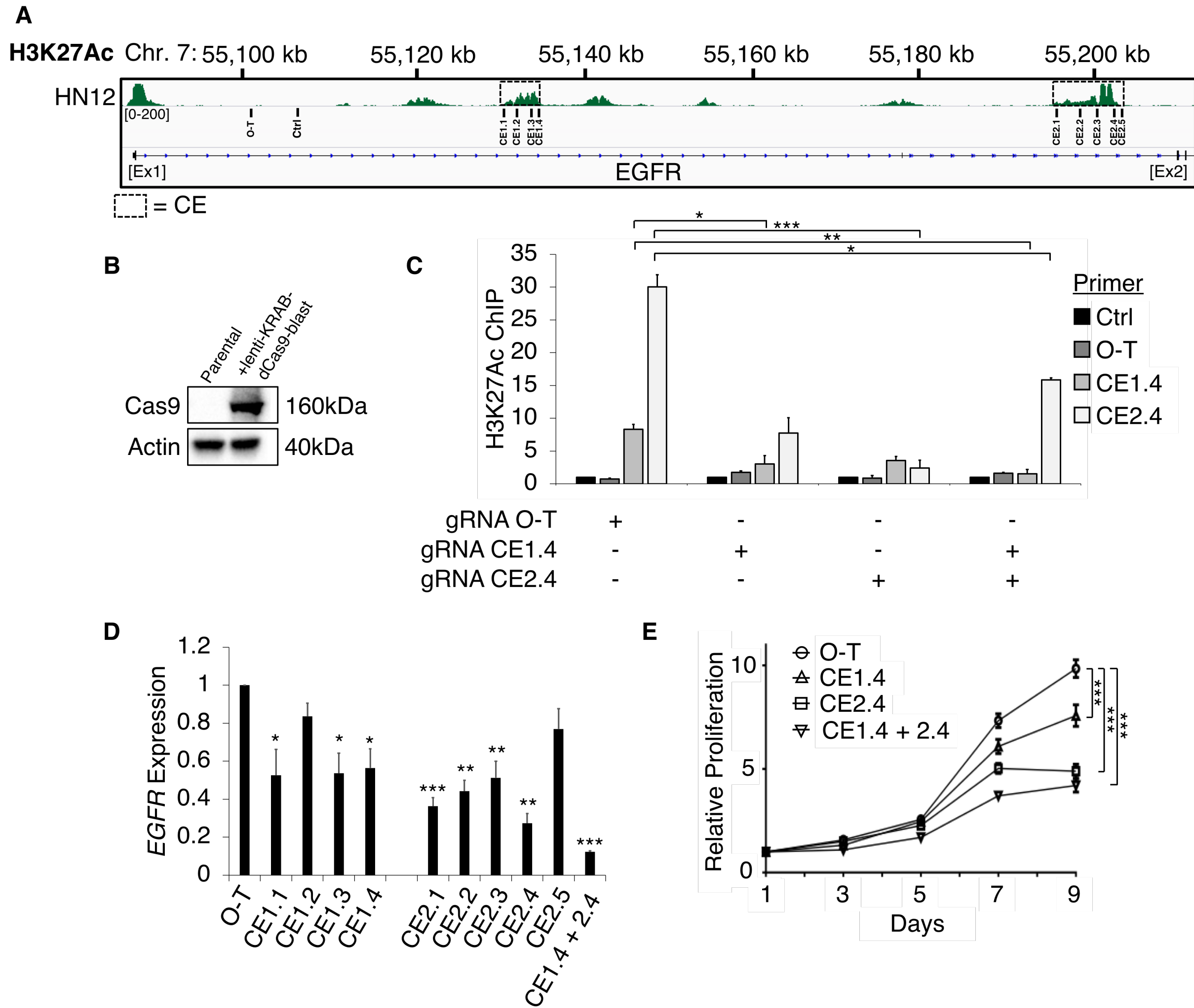


Figure 4.4: Targeting dCas9-KRAB to the CEs decreases *EGFR* gene transcription and reduces proliferation in HN12 cells. **A:** H3K27Ac IGV track of HN12 cells showing the position of gRNAs targeting the *EGFR* intron 1 enhancers and off-target (O-T) control. **B:** Western blot of dCas9-KRAB expression in HN12 cells after transduction with lenti-dCas9-KRAB-blast. β-Actin was used as a loading control. **C:** H3K27Ac enrichment at the targeted enhancer regions before and after dCas9-KRAB targeting was analyzed by ChIP-qPCR. Primers were designed around the targeted regions as well as a PCR negative control (Ctrl) from a H3K27Ac negative region of *EGFR* intron 1. A primer within a gene desert in chromosome 12 was used for normalization. **D:** *EGFR* expression in dCas9-KRAB expressing cell lines was analyzed by RT-qPCR. *EGFR* transcript level was first normalized to *GAPDH* and subsequently calculated as fold change relative to off-target control. **E:** Cell proliferation curves were generated by measuring ATP levels every two days over 9 days. Significance is measured relative to O-T. **C-E:** (*$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, $n = 3$ independent experiments, Students *t* test)

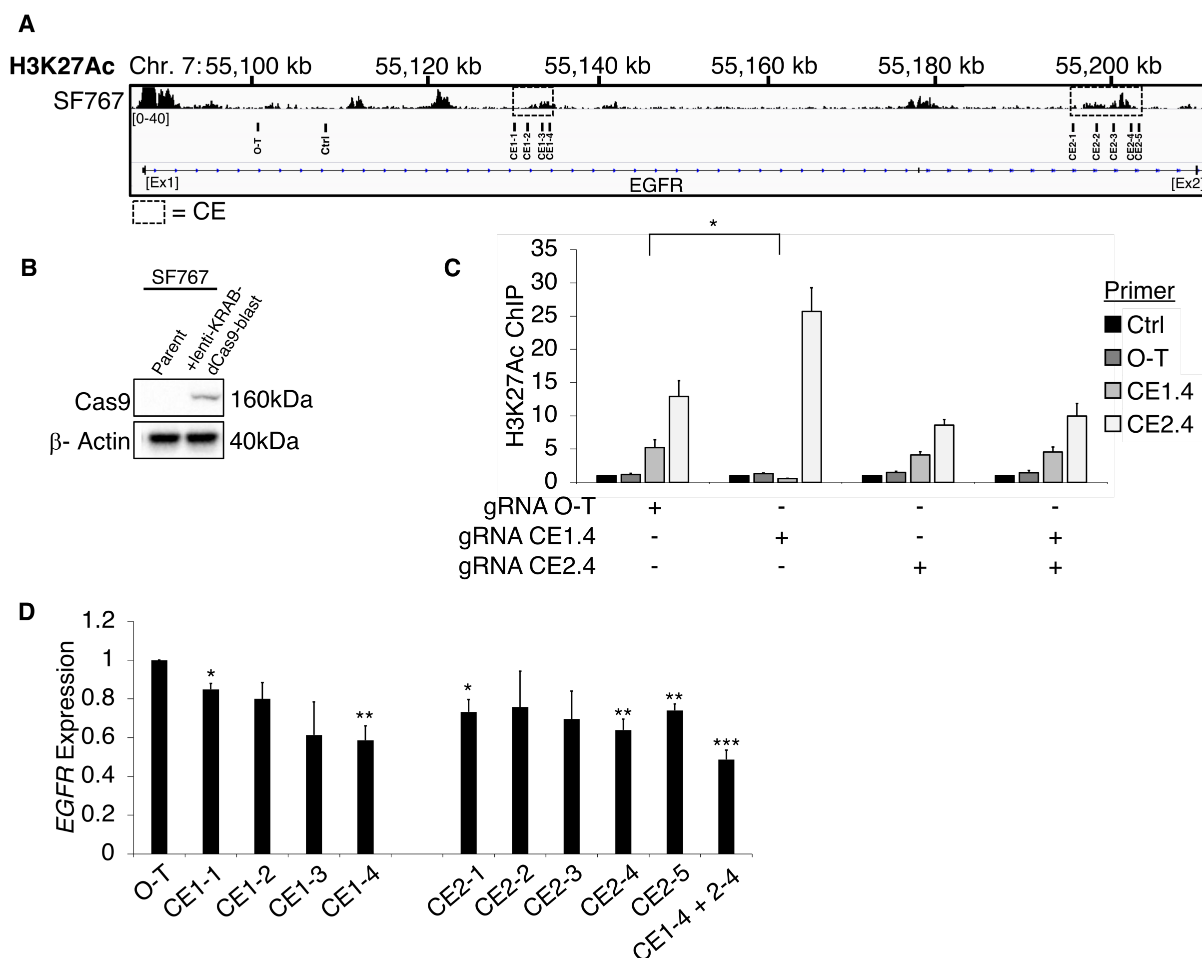


Figure 4.5: Targeting dCas9-KRAB to the CEs decreases *EGFR* gene transcription SF767 cells. **A:** H3K27Ac IGV track of HN12 cells showing the position of gRNAs targeting the *EGFR* intron 1 enhancers and off-target (O-T) control. **B:** Western blot of dCas9-KRAB expression in SF767 cells after transduction with lenti-dCas9-KRAB-blast. β-Actin was used as a loading control. **C:** H3K27Ac enrichment at the targeted enhancer regions before and after dCas9-KRAB targeting was analyzed by ChIP-qPCR. Primers were designed around the targeted regions as well as a PCR negative control (Ctrl) from a H3K27Ac negative region of *EGFR* intron 1. A primer within a gene desert in chromosome 12 was used for normalization. **D:** *EGFR* expression in dCas9-KRAB expressing cell lines was analyzed by RT-qPCR. *EGFR* transcript level was first normalized to *GAPDH* and subsequently calculated as fold change relative to off-target control. **C-D:** (*$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, $n = 3$ independent experiments, Students t test)

4.3 Summary

The discovery of CRISPR/Cas9 as a tool for genetic editing and functional genomics has revolutionized many fields of biomedical research, and the field of epigenetics has certainly benefitted from the tools which have been developed. In particular, the ability to functionally validate the activity of an enhancer has been expanded far beyond the simple reporter assay. Making a deletion of a putative enhancer, even one which has been validated by a reporter assay, allows for the native functional effect of that enhancer to be determined. Reporter assays suffer from many caveats including artificially high copy number, non-native cell contexts, and a lack of chromatin and DNA which surrounds the enhancer in the chromosome. CRISPR/Cas9 allows us to bypass many of these caveats. Through direct deletion, we can see how a regulatory element was influencing gene expression in the cell type in which it was deleted. Dose dependency of an enhancer cab be measured through incomplete (heterozygous) or complete (homozygous) editing. Enhancer-promoter contacts and how they are disrupted (or formed) can additionally be measured.

Although deletion holds a significant amount of power, the question remains whether structural variation affects local and distal gene expression control. Recent studies have identified methods such as enhancer hijacking which indicate that altering genomic regions through structural variation can have significant effects on the local chromatin structure, leading to aberrant gene expression [350, 351]. Additionally, these deletions are irreversible so further perturbation of other pathways can only inform you as to their effects in that particular context. For these reasons, the discovery of a mutation in Cas9 which destroys its nuclease function has been critical for creating other less permanent methods for targeted gene expression alterations [339]. dCas9-KRAB in particular has been widely utilized for repressing transcription as it has the ability to recruit repressive chromatin complexes, including histone deacetylase (HDAC) and histone methyltransferase (HMT) proteins [352].

In this study we performed both of these analyses on the *EGFR* gene in glioma and head and neck cancer cell lines which either express or do not express the gene. Single enhancer deletions through a dual gRNA strategy resulted in significant decreases in *EGFR* expression

in every case. Deletions of either CE1 or CE2 had different effects depending on the cell line. Though each cell line exhibited the strongest enrichment for H3K27Ac in CE2 (Figure 3.1A), SF767 and Cal27 showed significantly more repression of *EGFR* transcript with the loss of CE1. These cell-type specific differences in transcript levels between CE1 and CE2 indicate there may be differential utilization of either CE1 or CE2 in different cell lines. This suggests there may be a more complicated mechanism which determines which CE is used in each cell line, and may be due to differential expression of enhancer-specific transcription factors in those cell lines. Because partial redundant control of a gene by multiple CEs within a single SE has been observed [349], double enhancer deletion in these cells was tested. *EGFR* transcript levels were most significantly decreased with loss of both enhancers (Figure 4.3B) and the amplitude of loss was greater in $CE1^{-/-}$ + $CE2^{-/-}$ when compared against $CE1^{-/-}$ or $CE2^{-/-}$ alone. These data confirm the previous observations that enhancers can compensate for each other when necessary, and suggests that complete loss of enhancer-associated *EGFR* transcription would require targeting of at least these two, and potentially additional, enhancers.

To eliminate the possibility of structural variation being the root cause of *EGFR* transcriptional changes in response to CRISPR/Cas9 editing, we performed a dCas9-KRAB repressive experiment. Targeted repression, as confirmed by loss of H3K27Ac levels at the intended regions, was successful in significantly reducing *EGFR* expression in two EGFR-high cell lines. Interestingly, some important differences were observed between CRISPR/Cas9 deletion and dCas9-KRAB repression. First, in contrast to deletion in HN12 cells in which loss of either CE had equally deleterious effects (Figure 4.2B), there appeared to be a significant preference for repression of CE2 to downregulate *EGFR* transcript levels (Figure 4.4D). Our hypothesis is that the complete loss of CE2 in the deleted cells allowed for a restructuring of the 3D organization of the *EGFR* gene, leading to more efficient compensation of expression by CE1. In contrast, dCas9-KRAB repression of CE2 was able to silence this compensatory effect by DNA bending, bringing the repressive KRAB in close contact with CE1 and the *EGFR* promoter. CE1 deletion in (Figure 4.2B) perhaps indicates the true repressive effect of CE1 loss, as CE1 repression by dCas9-KRAB produces similar fold changes in *EGFR* transcript (Figure 4.4D). Although these

phenotypes are not maintained in SF767 cells (Figure 4.5), this may be due to poor expression of dCas9-KRAB protein (Figure 4.5B) or gRNA through inefficient transduction by lentivirus in these cells. However, as with our deletion approach, repression of both CE1 and CE2 in combination produced strong inhibition of *EGFR* expression (Figure 4.4D, Figure 4.5D), further confirming their cooperativity and strengthening the hypothesis that complete enhancer blockade would be critical for full *EGFR* repression.

4.4 Materials and Methods

Cell Culture.

GBM cell line SF767 (RRID:CVCL_6950) was provided by Dr. Mitch Berger (UCSF Brain Tumor Center, San Francisco, CA). Head and neck cell line HN12 (RRID:CVCL_5518) was provided by Dr. Silvio Gutkind (UCSD Moores Cancer Center, San Diego, CA). Cells were maintained in DMEM (Hyclone, #SH30022.01) supplemented with 10% fetal bovine serum (Atlanta Biologicals, #S12450) and 1% penicillin-streptomycin (Gibco, #15140-122) and grown as adherent cultures. Mycoplasma testing was performed with the PlasmoTest kit (InvivoGen, #rep-pt1) and found to be negative. All experiments are performed within 10 passages of the original frozen stock or post-manipulation.

Guide RNA design.

Guide RNAs were designed using the MIT CRISPR Design website (http://crispr.mit.edu). To minimize potential off-target effects of guides, only high-score guide RNAs (score >80) were used. Guide RNAs were annealed and diluted 1:200 in ddH2O and used for downstream applications.

CRISPR/Cas9-mediated Genomic Deletion.

Guide RNAs were cloned into pX330-BFP (from Dr. Tim Fenton) for upstream guides or pX458-GFP (Addgene, Plasmid #48138) for downstream guides. For pX330-BFP constructs, annealed gRNA duplexes were first cloned into pX335-U6-Chimeric_BB-CBh-hSpCas9n (D10A) (Addgene, #42335). pX335 was digested at 37°C for 1 hour with BbsI (NEB, #R0539S). Lin-

earized plasmid was run on a 1% agarose gel, excised and purified with the QIAquick Gel Extraction Kit (Qiagen, #28704). Annealed gRNA duplexes were ligated into the linearized plasmid with the DNA Ligation Kit Version 2.1 (TaKara, #6022). Ligation products were transformed into XL1-Blue *E. Coli* and single clones were purified with the QIAprep Spin Miniprep Kit (Qiagen, #27104) and sent for sequencing (IDT) with the hU6-F sequencing primer. Sequencing-confirmed PX335-gRNA and PX330-BFP-empty were double digested with KpnI-HF (NEB, #R3142S) and FspI (NEB, #R0135S) and run on a 1% agarose gel. The 1kb digested piece from pX335-gRNA was extracted and purified with the QIAquick Gel Extraction Kit. The larger digested construct from PX330-empty was extracted and purified as well. The DNA fragment containing the gRNA and the PX330-empty digested construct were ligated. Ligated constructs were transformed into XL1-Blue and single clones were mini-prepped. Presence of guides in the PX330-BFP constructs was confirmed by Sanger sequencing (IDT). Constructs were maxi prepped and co-transfected with Lipofectamine 3000 (ThermoFisher, #L3000015) into 4.5 x 10^5 of the indicated cell line in a 6-well plate well. After 24 hours cells were collected and the top 1% of BFP+/GFP+ cells were sorted using the SH800S Cell Sorter (Sony Biotechnology). Single cells were plated in 96 well plates and grown for 2-3 weeks. Single clones were screened using PCR with primers described in Figure 4.2. A minimum of 2 homozygous clones were mixed at equal ratios and used for downstream analysis.

Enhancer Silencing by CRISPR/dCas9-KRAB.

SF767 and HN12 cells were transduced with maxi-prepped Lenti-dCas9-KRAB-blast (from Dr. Paul Mischel) and selected with 10µg/ml blasticidin for 72 hours post transduction. Lentivirus was generated by transfecting 293T cells with the plasmid together with the packaging plasmids pVSVG and pΔ8.9 using Lipofectamine 3000. Supernatant containing virus was harvested 24 and 48 hours after transfection, filtered through a 40µm filter, and used to transduce the indicated cell lines. Viral transductions were performed in the presence of 4mg/ml polybrene. Medium was changed 24 hours after transduction and replaced with medium containing 10µg/ml blasticidin for 3 days. Guide RNAs were cloned into the lentiGuide-Puro vector (Addgene, Plasmid #52963). LentiGuide-Puro was first digested with BsmBI (NEB, #R0580S) and run on

a 1% agarose gel. The upper, larger band was excised and purified using the QIAquick Gel Extraction Kit. Purified products were added to a ligation reaction with the following components: 1µl BsmBI digested plasmid (50ng), 1µl diluted oligo duplex, 5µl 2X Quick Ligase Buffer (NEB), 3µl ddH2O, 1µl Quick Ligase (NEB, #M2200S) and incubated at room temperature for 10 minutes. 1uµl ligation product was transformed into Stbl3 bacteria. Individual colonies were mini-prepped and confirmed for gRNA sequence by Sanger sequencing (IDT) using the hU6-F sequencing primer. Confirmed constructs were maxi-prepped and made into virus as described above. Individual gRNA-expressing viruses were transduced into cells expressing dCas9-KRAB. Medium was changed 24 hours after transduction and replaced with medium containing 1µg/ml puromycin for 3 days. After assessing *EGFR* transcript levels by RT-qPCR, one highly effective CE1 guide was selected for double gRNA expression. One stable line expressing the highly effective CE1 guide was transduced a second time with the complementary CE2 guide. Enhancer activity was assessed by H3K27Ac ChIP-qPCR.

Western Blotting.

Protein samples were collected in SDS sample buffer, separated using gel electrophoresis and transferred via wet transfer onto a PVDF membrane. The membrane was blocked with 5% milk in TBST and probed with primary antibodies at 1:1000 dilution overnight at 4°C and secondary horeseradish peroxidase (HRP) antibodies at 1:2000 for 1 hour at room temperature. Signal was assessed via chemiluminescence with the SuperSignal West Pico PLUS substrate (ThermoFisher, #34580) and visualized on a ChemiDoc MP system (Bio-Rad). anti-β-actin (Sigma, #A3854) and Anti-Cas9 (Cell Signal, #14697) antibodies were used for analysis.

Quantitative real-time PCR.

RNA was extracted with the RNeasy Plus kit (Qiagen, #74134) according to the manufacturers instructions. Reverse transcription of mRNA was performed using 1µg RNA with the iScript Reverse Transcription Supermix (BioRad, #1708841). For real-time PCR analysis, 5µl of cDNA (50ng of starting RNA) was amplified per reaction using the iTaq Universal SYBR Green Supermix (Bio-Rad, #1725124) and the Bio-Rad CFX96 qPCR system. Fold change

analysis was performed using the $2^{-\Delta\Delta Ct}$ method and normalized as indicated.

Chromatin Immunoprecipitation.

Chromatin immunoprecipitation was performed as described previously [317] with the following modifications. Chromatin was sheared in diluted lysis buffer to 200-500bp using a Covaris M220 Focused-Ultrasonicator with the following parameters: 10 minutes, peak incident power 75, duty factor 10%, 200 cycles/burst. Antibodies for ChIP were obtained from commercially available sources: anti-H3K27Ac (Active Motif, #39133). 5% of the chromatin was not exposed to antibody and was used as control (input). For ChIP-qPCR analysis DNA quantity for each ChIP sample was normalized against input DNA.

Cell growth analysis.

$5x10^2$ HN12 cells expressing the indicated dCAs9-KRAB gRNA were seeded in black, clear bottom 96 well plate in 6 replicate wells in complete media. After 24 hours, complete media was removed and 100µl of 10µg/ml blasticidin and 1µg/ml puromycin in DMEM + 0.5% were added to each well. Baseline luminescence was measured at day 1 with the ATPlite 1step Luminescence Assay System (PerkinElmer, #6016731) on a Tecan Spark 10M. Luminescence measurements were obtained at every other day for 9 days and plotted using GraphPad Prism.

Data Access.

All raw and processed sequencing data generated in this study have been submitted to NCBI GEO under accession number GSE128275.

4.5 Acknowledgements

Chapter 4 in part, has been submitted for publication of the material as it may appear in Molecular Cancer Research; **Jameson, N.M.**, Ma, J., Benitez, J., Izurieta, A., Han, J.Y., Mendez, R., Parisian, A., Furnari, F., AACR Publications, 2019. The dissertation author was the primary investigator and author of this paper.

Chapter 5

Identification and characterization of transcription factors critical for intron 1-mediated *EGFR* expression

5.1 Introduction

Sequence-specific transcription factors (TF) interpret the signals encoded within regulatory DNA. Many TFs function as master regulators of specific cell types, including stem cells [355] or muscle cells [356], or signaling pathways like immune response [357]. Their power can be directly demonstrated through the ability to force terminally differentiated fibroblasts to a stem-cell state through only the expression of a small set of stem-cell-specific master TFs [358]. The initial discovery of human-specific TFs was aided by the discovery of Deoxyribonuclease I (DNase I) footprinting [359], a process which takes advantage of the remodeling of nucleosomes after the binding of TFs to regulatory DNA regions. This remodeling exposes DNA, normally protected by nucleosome interactions, to cleavage by DNase I. This cleavage is not uniform, as regulatory DNA currently bound by TFs is protected, leaving footprints that mark TF occupancy at high resolution [360]. These footprints are enriched with DNA sequences specific to particular TFs, known as motifs. Many TFs mediate their effects through direct binding to DNA via

these motifs, utilizing DNA binding domains (DBD) of which there are three major classes: zinc fingers, homeodomains, and helix-loop-helix domains [361]. TFs can have 1,000-fold or greater preference for these specific binding motifs relative to other sequences [362], however the overlap between experimentally determined binding sites and sequence matching the motif is not perfect. TF-binding motifs are small (6-12 bases) and flexible. This flexibility is represented by sequence logos in which the relative preference for the TF for each base in the binding site is represented through the size of the base relative to other bases at that position [363]. The identification of these motifs is critical for assigning function to a particular TF, as motifs can be identified at regulatory regions of DNA and the role of the TF at that region can then be assessed.

Systems-wide approaches have been undertaken to identify the transcription-factor-binding patterns in unicellular organisms, and these studies have been successful in identifying how limited cohorts of TFs are able to organize the large diversity of gene expression patterns [364, 365]. Similar studies in more complex organisms are challenging due to the size of the TF repertoire and genome, however TF regulatory networks have been successfully built in a few cell types, including embryonic stem cells (ESC) [355]. The challenge of these studies is further exacerbated by the observation that the same TF can regulate different genes in different cell types [366], and their regulatory patterns can change when mutated [367]. ChIP-seq has revolutionized the study of TF-binding sites by enabling the genome-wide identification of regions occupied by a TF of interest. Using ChIP-seq for transcription factor identification has several caveats. TF binding is a dynamic process, and the use of crosslinkers for ChIP does not enable the measurement of dynamic TF binding. ChIP data is also highly dependent on antibody quality, and the cross reaction of some antibodies eliminates the specificity of a particular TF binding site. Finally, ChIP-seq can detect indirect binding, which can lead to identification of binding sites for proteins other than the ChIPped protein [368]. This final caveat, while presenting a significant challenge for some experiments, also provides an opportunity for studying TF cooperativity. TFs are known to collaborate in a myriad of ways, including aiding each other in binding DNA, impacting chromatin state (synergistic regulation), binding cooperatively as homo- and hetero-dimers, or higher order structures [369]. Cooperative binding of TFs is often mediated

by protein-protein interactions, which confer additional stability when two or more interacting proteins bind DNA. This cooperation can impact the sequence preferences of TF complexes and constrain the DNA sequence in between two binding sites [370]. In some cases, TFs are able to enhance the binding of other TFs at a particular location through DNA-mediated cooperative binding. For example, the IFN-beta enhancer recruits eight DNA binding TFs which then allow for the recruitment of three non-DNA binding cofactors. Structural analysis reveals few contacts among the TFs, with stability instead being conferred by changes in DNA structure and interactions with cofactors [371].

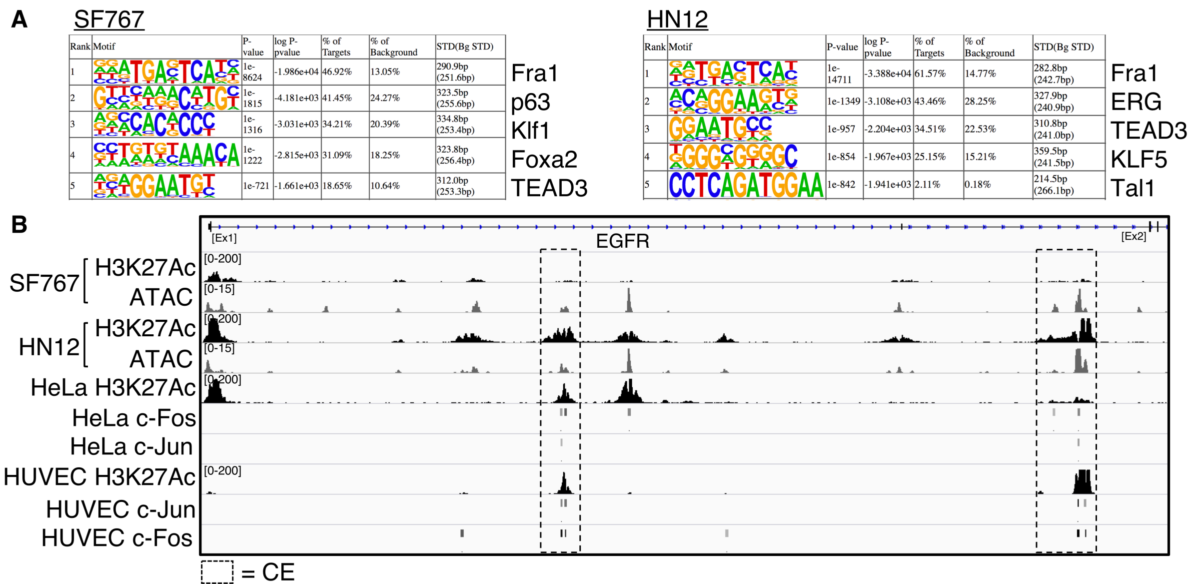
TFs can vary dramatically in how they impact transcription. Some TFs can directly recruit Pol II, provide steric hindrance by blocking binding of other proteins [372], or act by recruiting cofactors including coactivators and corepressors [373]. These cofactors are frequently large multi-subunit protein complexes or multi-domain proteins that regulate transcription through nucleosome remodeling, covalent modifications of histones or other proteins, or recruitment of Pol II [374]. Through their transcriptional activation activities, whether direct or indirect, TFs are critical for driving gene expression in a cell-type-specific and timing-dependent manner. Thus, identification of transcription factors critical for the activity of particular genes is an important step in understanding the regulation of that gene and recognizing how to restore transcriptional balance to aberrantly activated or repressed genes in diseased states.

5.2 Results

AP-1 Family Transcription Factors Bind to and Influence *EGFR* Intron 1 Enhancers.

The primary function of enhancers is to serve as a binding site for different TFs, which can recruit cofactors to mediate the recruitment of Pol II at core promoters. Every enhancer has a slightly different set of factors it can bind to, and this is influenced by many things including enhancer size, motifs present, and proximity to other TF motifs. We wanted to identify the critical motifs in CE1 and CE2 in order to find the crucial TFs mediating the interaction between each CE and the *EGFR* promoter. To begin identifying critical motifs we further analyzed our H3K27Ac ChIP-seq and ATAC-seq data in SF767 and HN12 cells. We utilized a recent paper

which identified critical TFs in macrophages as a guideline for this process [375]. In brief, Lavin et al. performed H3K27Ac/H3K4me1 ChIP-seq, ATAC-seq, and RNA-seq in purified macrophages from fresh mouse tissues. Enhancers were identified by overlapping H3K4me1 and H3K27Ac signals, and a strong correlation between ATAC-seq peaks and identified enhancers was identified, with the vast majority of enhancer regions containing at least one ATAC-seq peak. Because identification of small TF motifs would not be informative from broad enhancer peaks, they overlaid ATAC-seq peaks from corresponding cells with identified enhancers to narrow the search regions to the likely sites of TF binding. DNA lifted from overlaid ATAC-seq peaks could then be subjected to motif-finding algorithms to find highly enriched motifs. We utilized this strategy with some adjustment to find more enhancer-specific transcription factors. To eliminate non-enhancer regulatory regions (e.g. promoters), we intersected ATAC-seq peaks with enhancer peaks from H3K27Ac and kept only the TSS-distal (+/- 2.5kb) ATAC-seq peaks which mapped within an enhancer. Performing *de novo* motif analysis on these peaks in *EGFR* expressing cells (SF767 and HN12) identified an Activating Protein 1 (AP-1) transcription factor motif as the most significantly enriched motif (Figure 5.1A).



A

SF767

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	
1		1e-8624	-1.986e+04	46.92%	13.05%	290.9bp (251.6bp)	Fra1
2		1e-1815	-4.181e+03	41.45%	24.27%	323.5bp (255.6bp)	p63
3		1e-1316	-3.031e+03	34.21%	20.39%	334.8bp (253.4bp)	Klf1
4		1e-1222	-2.815e+03	31.09%	18.25%	323.8bp (256.4bp)	Foxa2
5		1e-721	-1.661e+03	18.65%	10.64%	312.0bp (253.3bp)	TEAD3

HN12

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	
1		1e-14711	-3.388e+04	61.57%	14.77%	282.8bp (242.7bp)	Fra1
2		1e-1349	-3.108e+03	43.46%	28.25%	327.9bp (240.9bp)	ERG
3		1e-957	-2.204e+03	34.51%	22.53%	310.8bp (241.0bp)	TEAD3
4		1e-854	-1.967e+03	25.15%	15.21%	359.5bp (241.5bp)	KLF5
5		1e-842	-1.941e+03	2.11%	0.18%	214.5bp (266.1bp)	Tal1

Figure 5.1: AP-1 family transcription factors are identified as possible modulators of *EGFR* intron 1 enhancers. **A:** Top 5 HOMER *de novo* motif search results in ATAC-seq peaks located within TSS-distal enhancers. % of Targets = Percentage of input peaks containing the indicated motif. % of Background = Percentage of randomly generated background sequences which contain the indicated motif. The number of background regions is 2x the total number of input peaks. STD(Bg STD) = average distance from the peak center where the motif was identified. **B:** Tracks of ENCODE ChIP-seq peaks from c-Jun and c-Fos ChIP-seq experiments in HeLa and HUVEC cells overlaid with tracks for H3K27Ac and ATAC-seq in glioma and HNSCC cell lines. Darker bars indicate stronger binding intensity.

AP-1 is a collective term referring to dimeric transcription factor composed of Jun, Fos, or ATF family subunits. A common feature of the proteins within these families is the conserved basic region leucine zipper (bZIP) DNA-binding domain, first identified in the initially discovered c-Jun protein [376]. The leucine zipper component of the bZIP domain is responsible for dimerization, which is a prerequisite for DNA binding mediated by the basic domain of bZIP. The primary AP-1 proteins in mammalian cells are c-Fos and c-Jun, however these proteins can form many combinations of both homo- and hetero-dimers depending on the cellular context, and these different dimerization patterns can have different effects on the genes that are regulated by AP-1 [377]. Different dimers are known to bind specific motifs, and this can drastically affect their transcriptional transactivation potential. The primary consensus motif, known as the TPA-responsive element (TRE), has the consensus sequence 5'-TGAG/CTCA-3'. Dimers of several AP-1 proteins, including c-Fos, c-Jun and FosB, can efficiently transform cells in culture and have potent transactivation domains [378]. Other AP-1 proteins, including JunB, JunD, Fra1 and Fra2, exhibit only weak transactivation potential and may even act as repressors of AP-1 activity by competing for binding sites or forming inactive dimers [379, 380]. Depending on the cellular context, the same AP-1 protein can have both pro- and anti-oncogenic activities [381, 382], thus it is critical to assess the phenotypic effects of perturbing AP-1 in our cell line models. Importantly, studies have shown the *EGFR* promoter to be a direct target of c-Jun transcriptional activity [383, 384]. Additionally, a previous study which identified an *EGFR* SE showed the binding of the AP-1 family member JunB to the identified SE in A549 NSCLC cells [287].

To validate the TF motifs identified in silico, we examined AP-1 family transcription factor ChIP-seq data deposited by the Encyclopedia of DNA Elements (ENCODE) consortium (Figure 5.1B). To compare ENCODE-generated ChIP-seq to data generated in-house, we overlaid tracks of H3K27Ac ChIP-seq and ATAC-seq from SF767 and HN12 with available H3K27Ac/c-Fos/c-Jun ChIP-seq data from HeLa and HUVEC cells. Importantly, H3K27Ac ChIP-seq in HeLa and HUVEC also shows high levels of enhancer marks in the AP-1 marked regions (Figure 5.1B). Using this approach multiple c-Jun and c-Fos peaks were identified within the CE1 and CE2 of

EGFR in HN12 and SF767 cells (Figure 5.2A). Of note, these c-Fos and c-Jun peaks appear to be enhancer specific within *EGFR* as no significant peaks were identified at the *EGFR* promoter in either HeLa or HUVEC cells (Figure 5.1B). For further analysis we chose c-Jun and c-Fos as the prototype AP-1 factors, as most AP-1 heterodimers contain at least one of c-Fos or c-Jun [385]. We validated and quantified c-Fos and c-Jun enrichment at the CE1 and CE2 regions in *EGFR* expressing (SF767 and HN12) and non-expressing (TS576) cells. We performed ChIP-qPCR at four regions within each constituent enhancer and identified significant fold enrichment of c-Fos and c-Jun within CE1 and CE2 in *EGFR* expressing cells over a negative control region in a gene desert on chromosome 12 (Figure 5.2B). Additionally, we identified significantly increased binding of c-Fos at the CE1-AP1-3 and c-Jun in CE1-AP1-3 and CE2-AP1-4 sub-regions in HN12 cells, indicating these regions may be important for the increased *EGFR* expression levels in these cells (Figure 5.2B).

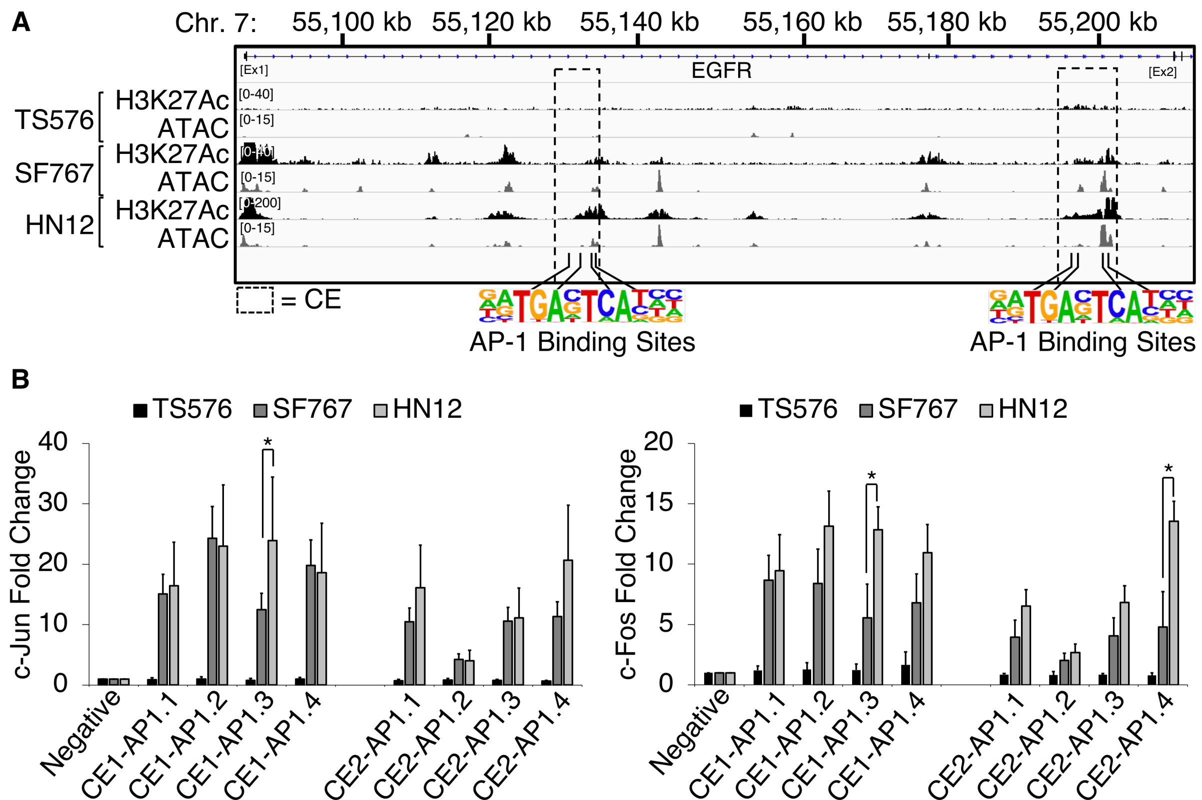


Figure 5.2: AP-1 family members bind to *EGFR* Intron 1. **A:** Schematic of positions of AP-1 binding positions based on ENCODE ChIP-seq data, shown relative to ChIP-seq and ATAC-seq peak density. CE1 and CE2 are highlighted. **B:** Analysis of (left) c-Jun and (right) c-Fos occupancy at the indicated sites. Transcription factor binding is represented as fold change over a negative control region located in Chr12. (*$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, $n = 3$ independent experiments, Students t test)

To validate the role of AP-1 transcription factors in *EGFR* transcription, we utilized a dominant-negative version of c-Jun (JunDN) [386]. The dominant negative was prepared as described in [386] and is missing the first 122 amino acids consisting of the N-terminal activation domain. JunDN can dimerize with other AP-1 family members and bind DNA, however the transcriptional activation capability is eliminated. A similar construct known as JunbZIP has been shown to downregulate *EGFR* expression through disruption of c-Jun in the fibrosarcoma cell line HT-1080 [387], however the effectiveness of a dominant negative c-Jun has not been shown in GBM or HNSCC. Constitutive expression of JunDN through transduction with a pMIEG3-JunDN retrovirus in HN12 cells showed decreased EGFR protein levels (Figure 5.3A), thus supporting a role for c-Jun heterodimers in the regulation of *EGFR* transcription. Additionally, we detected a decrease in c-Jun levels when JunDN was present, likely due to autoregulation of the *JUN* promoter by c-Jun heterodimers [388] (Figure 5.3A). To confirm the effect of JunDN was due to reduced c-Jun heterodimer activity, we utilized a luciferase reporter containing a trimerized AP-1 binding motif [389]. Transfection of the reporter construct in combination with a *Renilla* control plasmid in HN12 cells significantly decreases reporter activity when JunDN is present compared to the empty vector control (Figure 5.3B). These results confirm the specific transcriptional activation potential of JunDN is abrogated. Taken together, these data suggest that AP-1 family members specifically bind to *EGFR* enhancer regions and are critical for fine-tuned regulation of *EGFR* expression. Perturbation of this AP-1 transactivation effect by expression of a dominant negative results in a significant repression of c-Jun heterodimer targets including *EGFR* and *JUN*, confirming the role of this family of transcription factors in intron 1-mediated *EGFR* expression.

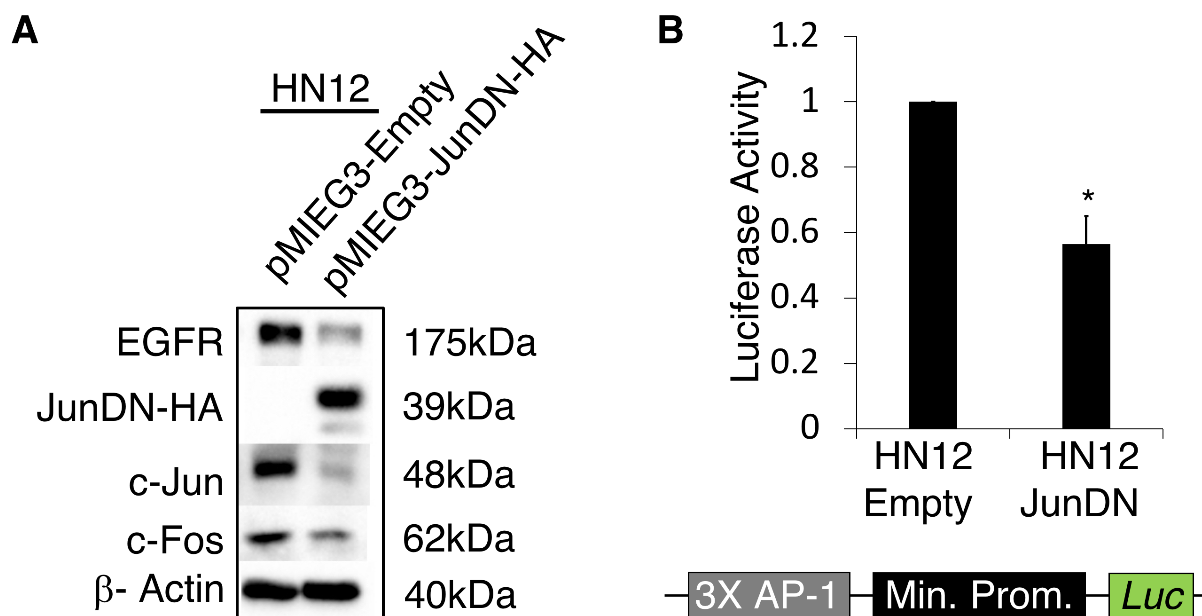


Figure 5.3: AP-1 family members modulate *EGFR* expression. **A:** Analysis of EGFR, JunDN-HA, c-Jun, and c-Fos protein expression in HN12 cells by western blotting after transduction with pMIEG3-JunDN-HA. β-Actin was used as a loading control. **B:** Analysis of JunDN efficacy on a luciferase reporter containing a trimerized AP-1 binding site. Relative luciferase activity after expression of JunDN is normalized to the empty vector. Relative luciferase expression is normalized against *Renilla* luciferase expression to control for transfection efficiency (*$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, $n = 3$ independent experiments, Student's t test)

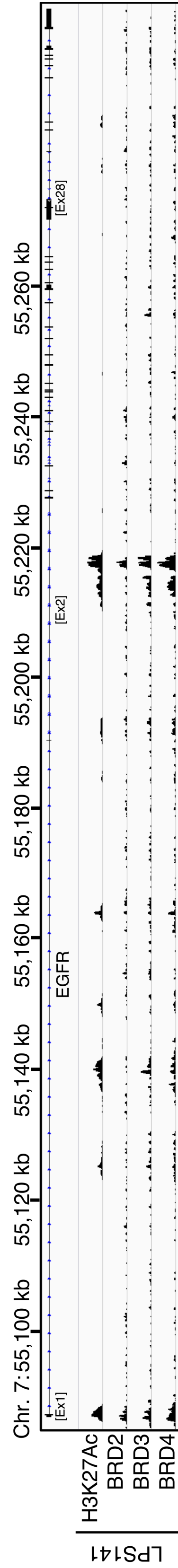
Treatment with JQ1 Reduces *EGFR* Expression by Modulation of TF Activity.

The number of proteins that have been found to interact with AP-1 components and regulate their transcriptional ability is both large and constantly increasing. One example of this interaction is the cooperative relationship between AP-1 and the glucocorticoid receptor (GR). GR can tether to AP-1 complexes without the need for DNA contact and was initially found to downregulate the activity of AP-1 [390], however subsequent genome-wide studies have indicated that the interaction is enriched at regulatory DNA and facilitates the binding and activity of GR through local relaxing of chromatin by AP-1 [391]. Another interaction exemplifying the crosstalk between AP-1 and other TFs is the relationship between AP-1 and NF-kB. NF-kB and AP-1 are activated by the same multitude of stimuli [392], and many genes are reported to require the combinatorial activation of AP-1 and NF-kB, suggesting they work cooperatively [393]. The bZIP domains of c-Fos and c-Jun have been shown to physically interact with the p65 subunit of NF-kB, and the transcriptional effect of AP-1 binding is robustly enhanced when p65 is already present at that region [394]. These results demonstrate that AP-1 has dual roles as both a DNA binding transcriptional activator through opening chromatin, as well as a cofactor with the ability to recruit additional DNA-binding or chromatin reader proteins to regulatory regions through protein-protein interaction.

Recently, AP-1 has been discovered to co-occupy enhancers with the bromodomain and extraterminal domain (BET) protein BRD4 [395]. Bromodomains occur in a variety of nuclear proteins that function as chromatin-binding proteins. The BET family of bromodomain proteins are distinguished by the presence of 2 bromodomains and at ET domain, and are known to bind to acetylated histones [168]. BET proteins including BRD2/3/4 have been shown to be important for enhancer function. BRD2/3 have been shown to bind to hyperacetylated regions and allow for the activity of RNA Pol II [396], and BRD4 in particular has been the subject of many studies. BRD4 is enriched at enhancers [269] and is a critical oncoprotein in AML [397], midline carcinomas [398], and GBM [399] and plays a role in autoimmunity and inflammatory diseases through its interactions with NF-kB [400]. The nature of the interaction between BET proteins and AP-1 is yet unknown, however previous research from our lab has shown treatment of mice

harboring GBM neurosphere PDX models with the pan-BET protein inhibitor JQ1 significantly prolongs survival [401], and combination of JQ1 with anti-EGFR therapy further increases this effect [399]. JQ1 functions by competing for the acetyl-lysine binding pocket of the bromodomain, displacing BET proteins from chromatin and altering the transcriptional activity of the target gene [401]. Additionally, recent data has shown that treatment of childhood sarcomas with JQ1 dramatically reduces AP-1 levels and activity through the suppression of the AP-1 family member FOSL1 [402], and reduced levels of both epidermal growth factor receptor (EGFR) and its ligand, heparin-binding EGF (HB-EGF), seem to be responsible for the severe proliferation defect in keratinocytes lacking c-Jun [317, 384].

Figure 5.4: BET bromodomain proteins bind in *EGFR* intron 1 acetylated regions. BET bromodomain protein and histone acetylation ChIP-seq tracks from the liposarcoma cell line LPS141.


Chr. 7:55,100 kb
55,120 kb
55,140 kb
55,160 kb
55,180 kb
55,200 kb
55,220 kb
55,240 kb
55,260 kb
[Ex1]
EGFR
[Ex2]
[Ex28]
H3K27Ac
BRD2
BRD3
BRD4
LPS141

To determine if the anti-tumorigenic effect of JQ1 as observed by [399] is partially attributable to downregulation of *EGFR* transcription by disruption of AP-1 and BET proteins, we treated HN12 and SF767 cells with JQ1 and measured the effects on *EGFR* expression. Interestingly, after 24 hours of 0.5µM JQ1 EGFR protein and transcript levels were decreased in both HN12 (Figure 5.5A, Figure 5.5B) and SF767 (Figure 5.6A, Figure 5.6B) cell lines. Additionally, BET family proteins BRD2 and BRD4 expression remained unchanged in both cell lines, however BRD3 expression increased in response to JQ1 in HN12 (Figure 5.5A).

To determine if JQ1 treatment was affecting EGFR levels through reduced activity of BRD family members, we first looked for evidence of binding of these factors to CE1 and CE2. Recent data in the liposarcoma cell line LPS141 [403] shows presence of H3K27Ac in CE1 and CE2, and has binding of BRD family members BRD2, BRD3 and BRD4 in those regions (Figure 5.4). To confirm binding of these factors in GBM and HNSCC and to interrogate their relationship with AP-1, we performed ChIP-qPCR for c-Fos, c-Jun, BRD2, BRD3, BRD4 and H3K27Ac at regions of open chromatin in CE1 and CE2. In HN12 cells, treatment with JQ1 significantly reduces occupancy of H3K27Ac at all measured regions, and significantly reduces binding of BET and AP-1 family transcription factors to CE1 and CE2 (Figure 5.5C). Interestingly, in contrast to steady state (Figure 5.2B) which suggests CE2-AP1-4 is a critical c-Fos and c-Jun binding site, treatment with JQ1 only affects binding of BRD4 at that region (Figure 5.5C). In SF767 cells, significant reductions in TF occupancy were observed primarily in CE2, with only BRD4 showing a significant reduction in binding to CE1 (Figure 5.6C).

To determine the specific BET protein critical for maintenance of *EGFR* expression we performed siRNA knockdowns of BRD2/3/4 individually as well as in combination. In HN12 cells single knockdown of BRD2 or BRD4, but not BRD3, resulted in downregulation of EGFR protein. Complete loss of BRD2/3/4 protein by combination siRNA treatment resulted in the strongest downregulation of EGFR protein (Figure 5.5D). In SF767 cells single knockdown of BRD2 produced the strongest downregulation of EGFR protein Figure 5.6D), indicating BRD2 and BRD4 activity at *EGFR* may be cell-type specific. In agreement with Figure 5.5A, knockdown of specific BRD proteins often resulted in upregulation of other BET family members.

These results are suggestive of some functional redundancy in these proteins, a result that is supported by previous data indicating partial redundancy of BRD2 and BRD3 in erythroid cells [404]. Despite this redundancy, increased expression of BRD3 after JQ1 treatment (Figure 5.5A) or RNAi (Figure 5.5D) fails to rescue the expression of EGFR, though triple siRNA treatment produced the strongest loss of EGFR protein. Thus, BRD3 alone is not sufficient for *EGFR* repression, but has some functional redundancy when other BET proteins are lost. These results argue against a role for BRD3 and implicate a role for BRD2 and BRD4 in cooperation with AP-1, in the maintenance of *EGFR* expression in GBM and HNSCC.

Figure 5.5: JQ1 treatment reduces *EGFR* transcription through inhibition of BRD2 and BRD4 activity in HN12 cells. **A:** Analysis of EGFR, BRD4, BRD3, BRD2 and β-Actin protein expression in HN12 cells by western blotting after treatment with 0.5μM JQ1 for 24 hours. β-Actin was used as a loading control. **B:** *EGFR* expression was analyzed by RT-qPCR in HN12 cells treated with 0.5μM JQ1 for 24 hours. *EGFR* transcript level was first normalized to *GAPDH* and subsequently calculated as fold change relative to DMSO control. **C:** Fold changes in enrichment of the indicated factors after 24 hours of 0.5μM JQ1 was measured at the indicated regions by ChIP-qPCR. ChIP enrichment is normalized to a negative control primer in chr12. **D:** Analysis of EGFR, BRD4, BRD3, BRD2 and β-Actin protein expression in HN12 cells by western blotting after treatment with indicated siRNA. A scrambled siRNA was used as treatment control and β-Actin was used as a loading control. **B-C:** (*$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, $n = 3$ independent experiments, Student's t test)

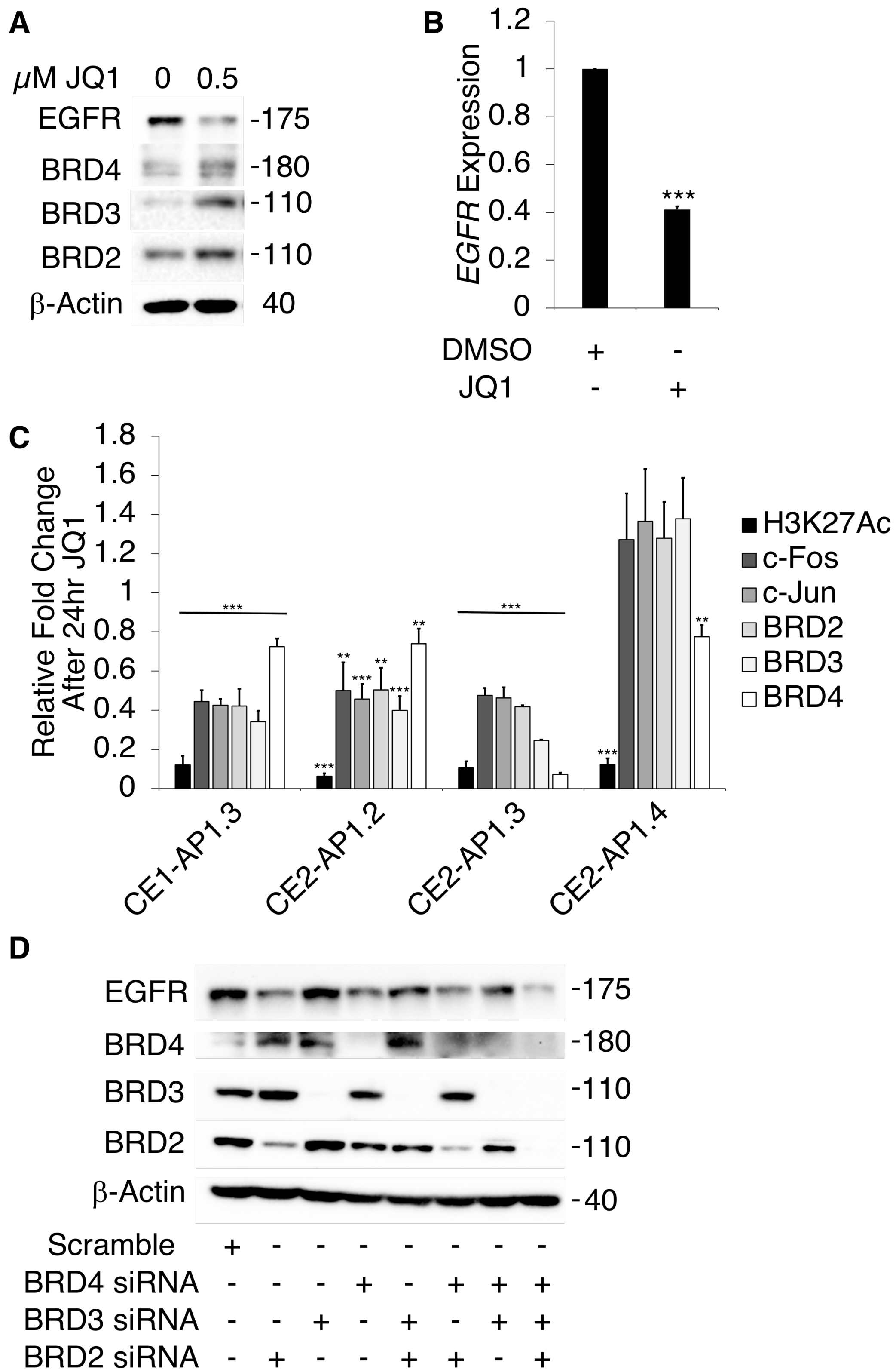
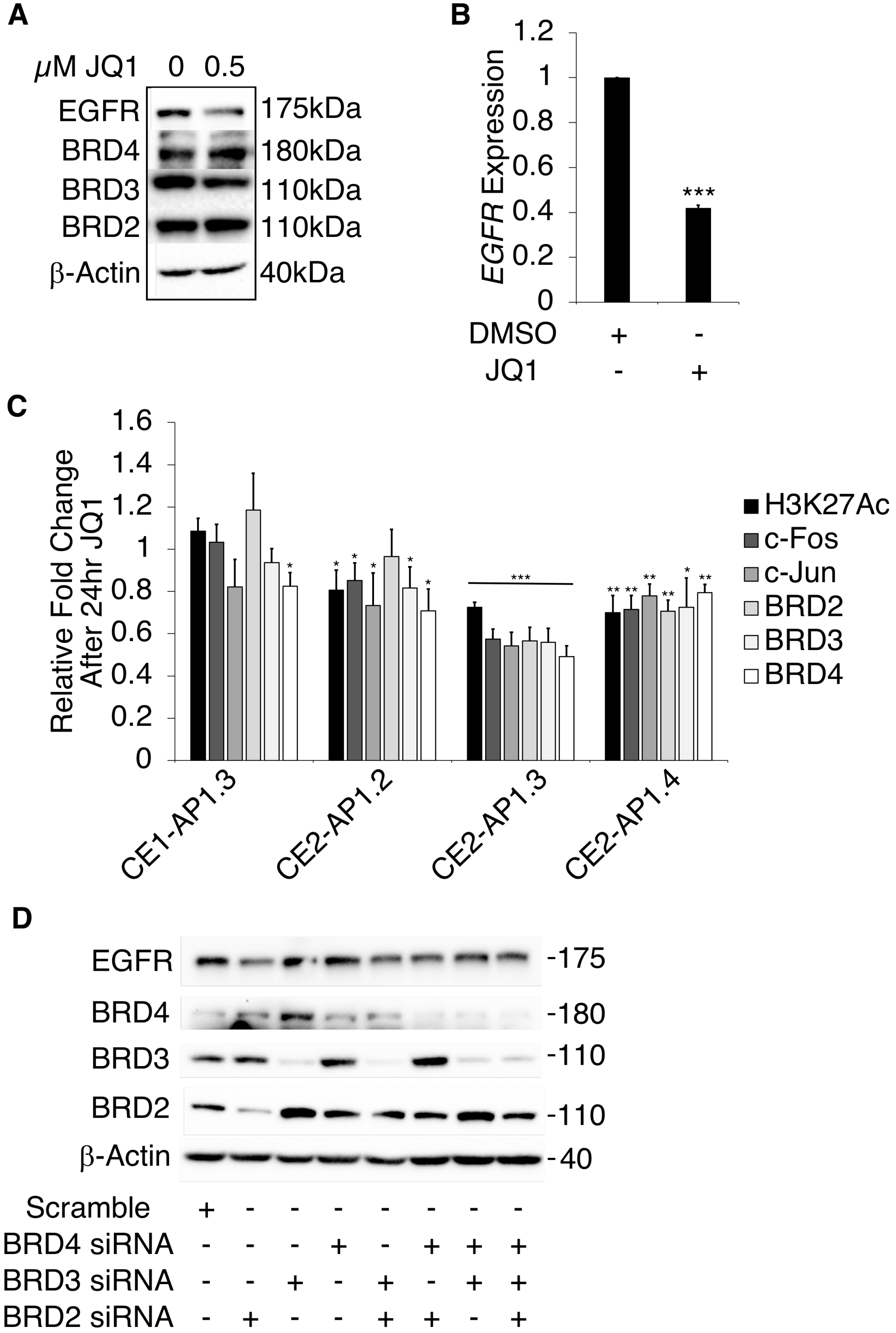

A
μM JQ1 0 0.5
EGFR -175
BRD4 -180
BRD3 -110
BRD2 -110
β-Actin 40
B
EGFR Expression
1.2
1
0.8
0.6
0.4
0.2
0
***
DMSO + -
JQ1 - +
C
Relative Fold Change After 24hr JQ1
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
H3K27Ac
c-Fos
c-Jun
BRD2
BRD3
BRD4
CE1-AP1.3
CE2-AP1.2
CE2-AP1.3
CE2-AP1.4
D
EGFR -175
BRD4 -180
BRD3 -110
BRD2 -110
β-Actin -40
Scramble + - - - - - - -
BRD4 siRNA - - - + - + + +
BRD3 siRNA - - + - + - + +
BRD2 siRNA - + - - + + - +

Figure 5.6: JQ1 treatment reduces *EGFR* transcription through inhibition of BRD2 and BRD4 activity in SF767 cells. **A:** Analysis of EGFR, BRD4, BRD3, BRD2 and β-Actin protein expression in SF767 cells by western blotting after treatment with 0.5μM JQ1 for 24 hours. β-Actin was used as a loading control. **B:** *EGFR* expression was analyzed by RT-qPCR in SF767 cells treated with 0.5μM JQ1 for 24 hours. *EGFR* transcript level was first normalized to *GAPDH* and subsequently calculated as fold change relative to DMSO control. **C:** Fold changes in enrichment of the indicated factors after 24 hours of 0.5μM JQ1 was measured at the indicated regions by ChIP-qPCR. ChIP enrichment is normalized to a negative control primer in chr12. **D:** Analysis of EGFR, BRD4, BRD3, BRD2 and β-Actin protein expression in SF767 cells by western blotting after treatment with indicated siRNA. A scrambled siRNA was used as treatment control and β-Actin was used as a loading control. **B-C:** (*$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$, $n = 3$ independent experiments, Student's *t* test)


A
μM JQ1 0 0.5
EGFR 175kDa
BRD4 180kDa
BRD3 110kDa
BRD2 110kDa
β-Actin 40kDa
B
EGFR Expression
1.2
1
0.8
0.6
0.4
0.2
0
***
DMSO + -
JQ1 - +
C
Relative Fold Change After 24hr JQ1
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
H3K27Ac
c-Fos
c-Jun
BRD2
BRD3
BRD4
CE1-AP1.3
CE2-AP1.2
CE2-AP1.3
CE2-AP1.4
D
EGFR -175
BRD4 -180
BRD3 -110
BRD2 -110
β-Actin -40
Scramble + - - - - - - -
BRD4 siRNA - - - + - + + +
BRD3 siRNA - - + - + - + +
BRD2 siRNA - + - - + + - +

5.3 Summary

Many potent oncogenes are influenced by super enhancers, including *MYC* and *HER2* [316, 346]. These genes and others have been successfully targeted in cancers which express them by inhibiting the BET protein BRD4, a hallmark factor involved in super enhancer identity [269]. The BET-bromodomain inhibitor JQ1 has been shown to inhibit BRD4 at super enhancers [269] and additionally sensitizes *EGFR* amplified GBM [399] and HNSCC [405] cells to EGFR TKI. Our data suggests that this sensitization may also be due in part to JQ1-mediated inhibition of AP-1 and other BET proteins at *EGFR* intron 1 enhancers (Figure 5.5, Figure 5.6). Although site-specific inhibition of TF binding is cell-type dependent, whether targeting TF binding pharmacologically (Figure 5.5A-C, Figure 5.6A-C) or through RNAi (Figure 5.5D, Figure 5.6D), the effects on *EGFR* are consistently negative. These results support the global targeting of AP-1 and BET rather than site-specific repression (e.g. dCas9-KRAB). Additionally, our study shows significant reduction of EGFR protein and transcript after JQ1 treatment independent of gene amplification (Figure 5.5A-C, Figure 5.6A-C), indicating a broader effect of JQ1 on tumor models which express *EGFR* at varying levels. Indeed, *EGFR* transcript levels are reduced to similar levels in SF767 cells which express less *EGFR* and exhibit less significant TF occupancy differences in response to JQ1 treatment (Figure 5.6C). These data combined further support the combination of EGFR TKI and JQ1 as treatment for EGFR-positive malignancies.

It has been previously observed that Ras is a significant activator of AP-1 activity, and has been identified as one of the main cooperating partners of AP-1 [406, 407]. Oncogenic transformation of a cell with Ras or MEK significantly induces AP-1 protein expression [408, 409], specifically the Fra1 and c-Jun proteins [408]. This potent transformation appears to be through the N terminal phosphorylation of c-Jun by the Jun N-terminal Kinase (JNK) proteins [410, 411] and is supported by the observation that Ras-mediated transformation is suppressed in fibroblasts that lack c-Jun [412]. As introduced in chapter 1, Ras is potently activated upon EGFR phosphorylation [15] leading to the activation of many different pathways, including enhanced AP-1 transcription. This pathway indicates a potential positive feedback loop for the continued expression of EGFR. Activated EGFR (following ligand binding or amplification) activates

Ras and MAPK pathways, stimulating AP-1 production and phosphorylation, and leading to increased enhancer activation of *EGFR* transcription. There are many opportunities for pharmacological inhibition of this pathway as discussed earlier, however the wide range of kinases and transcription factors involved also introduces a high likelihood for bypass pathways. Further studies should evaluate the efficacy of combining various upstream and downstream pathway inhibitors, and determine the mechanism by which they are exacting their effects so that inevitable resistance mechanisms can then be targeted in a logical and systematic manner.

5.4 Materials and Methods

Cell Culture.

SF767 (RRID:CVCL_6950) was provided by Dr. Mitch Berger (UCSF Brain Tumor Center, San Francisco, CA). Head and neck cell line HN12 (RRID:CVCL_5518) was provided by Dr. Silvio Gutkind, UCSD Moores Cancer Center, San Diego, CA. Cells were maintained in DMEM (Hyclone, #SH30022.01) supplemented with 10% fetal bovine serum (Atlanta Biologicals, #S12450) and 1% penicillin-streptomycin (Gibco, #15140-122) and grown as adherent cultures.

TF Motif identification.

Tag directories were created from H3K27Ac alignment files using HOMER makeTagDirectory. Peaks from the H3K27Ac ChIP-seq were identified from the tag directories using HOMER findPeaks with parameters -style histone -size 250 -minDist 250. NarrowPeak files from individual ATAC-seq replicates and pooled peak files from MACS2 were sorted for high P-value peaks ($\geq$ 13 -log10 [$p < 0.05$]) with awk. In order to get a set of high P-value, replicated peaks, high P-value pooled peaks were intersected with high P-value peaks from each replicate using bedtools intersect with the -u parameter. In order to exclude promoter proximal peaks, replicated high P-value peaks were sorted for. The Gencode hg38 annotation was downloaded and TSS +/- 2kb were filtered out with awk. High p-value peaks were intersected with the Gencode TSS annotation with bedtools intersect with the -v parameter. H3K27Ac peaks were modified from hg19 to hg38 using the hgLiftOver tool for intersecting with ATAC-seq peaks.

Promoter distal high p-value ATAC peaks were intersected with the H3K27Ac ChIP-seq peaks to identify ATAC peaks located within regions of H3K27Ac using bedtools intersect with the -wa parameter. Motifs were identified from these peaks using HOMER findMotifsGenome with the parameter -size given. De novo motifs were identified from html files output by HOMER.

Data Access.

All raw and processed sequencing data generated in this study have been submitted to NCBI GEO under accession number GSE128275. Accession numbers for publicly available data accessed are as follows: HeLa c-Fos (GSM935317), HeLa c-Jun (GSM935341), HeLa H3K27Ac (GSM733684), HUVEC c-Fos (GSM935585), HUVEC c-Jun (GSM935278), HUVEC H3K27Ac (GSM733691), LPS141 H3K27Ac (GSM3027215), LPS141 BRD2 (GSM3027219), LPS141 BRD3 (GSM3027220), LPS141 BRD4 (GSM3027221).

Quantitative real-time PCR.

RNA was extracted with the RNeasy Plus kit (Qiagen, #74134) according to the manufacturers instructions. Reverse transcription of mRNA was performed using 1µg RNA with the iScript Reverse Transcription Supermix (BioRad, #1708841). For real-time PCR analysis, 5µl of cDNA (50ng of starting RNA) was amplified per reaction using the iTaq Universal SYBR Green Supermix (Bio-Rad, #1725124) and the Bio-Rad CFX96 qPCR system. Fold change analysis was performed using the $2^{-\Delta\Delta Ct}$ method and normalized as indicated.

Luciferase Reporter Assay.

pMIEG3-JunDN (RRID: Addgene_40350) and 3xAP1pGL3 (RRID:Addgene_40342) were gifts from Alexander Dent. pMIEG3-Empty was created by removing the JunDN sequence by EcoRI digestion. For each transfection reaction, 100ng control plasmid expressing Renilla luciferase (Promega, #E2241) and 1µg Firefly luciferase construct were co-transfected with Lipofectamine 2000 (ThermoFisher, #11668030) into 2 x 10^5 cells in a 12-well plate well. After 24 hours, cells were collected in 1X PLB. Luciferase activity was measured by the Dual-Luciferase Reporter Assay System (Promega, #E1910) on a Tecan Spark 10M with injection

control. Transfection efficiency was controlled for by dividing firefly luminescence by *Renilla* luminescence, and final activity was normalized to a negative control.

Western Blotting.

Protein samples were collected in SDS sample buffer, separated using gel electrophoresis and transferred via wet transfer onto a PVDF membrane. The membrane was blocked with 5% milk in TBST and probed with primary antibodies at 1:1000 dilution overnight at 4°C and secondary horeseradish peroxidase (HRP) antibodies at 1:2000 for 1 hour at room temperature. Signal was assessed via chemiluminescence with the SuperSignal West Pico PLUS substrate (ThermoFisher, #34580) and visualized on a ChemiDoc MP system (Bio-Rad). anti-β-actin (Sigma, #A3854), anti-c-Jun (Cell Signal, #9165S), anti-c-Fos (Santa Cruz Biotechnology, #sc-52), anti-HA-HRP (Santa Cruz Biotechnology, #sc-805), anti-BRD4 (Active Motif, #39909), anti-BRD3 (Santa Cruz Biotechnology, #sc-515729), anti-BRD2 (Cell Signal, #5848S) and anti-EGFR (BD Biosciences, #610017) were used for analysis.

Chromatin Immunoprecipitation.

Chromatin immunoprecipitation was performed as described previously [317] with the following modifications. Chromatin was sheared in diluted lysis buffer to 200-500bp using a Covaris M220 Focused-Ultrasonicator with the following parameters: 10 minutes, peak incident power 75, duty factor 10%, 200 cycles/burst. Antibodies for ChIP were obtained from commercially available sources: anti-H3K27Ac (Active Motif, #39133), anti-BRD4 (Active Motif, #39909), anti-c-Jun (Cell Signaling, #9165T), and anti-c-Fos (Santa Cruz Biotechnology, #sc-166940). 5% of the chromatin was not exposed to antibody and was used as control (input). For ChIP-qPCR analysis DNA quantity for each ChIP sample was normalized against input DNA.

siRNA Transfection.

$1x10^5$ SF767 or $5x10^4$ HN12 cells were seeded in 12 well plates and grown overnight. siRNAs were transfected into each well with Lipofectamine 2000 in serum free and antibiotic free DMEM. Media was changed to complete media 6 hours later. Samples were collected in SDS sam-

ple buffer 48-72 hours later. siRNAs used for this study include BRD2 (Ambion, #s12070), BRD3 (Ambion, #s15544), BRD4 (Ambion, #s23902), and scramble control (Invitrogen, #12935-300).

JQ1 Treatment.

SF767 or HN12 cells were treated with 0.5µM JQ1 dissolved in DMSO (MedChemExpress, #HY-13030) for 24 hours. Vehicle control samples were treated with equal volume DMSO for 24 hours. Samples were collected in SDS sample buffer and analyzed.

5.5 Acknowledgements

Chapter 5 in part, has been submitted for publication of the material as it may appear in Molecular Cancer Research; **Jameson, N.M.**, Ma, J., Benitez, J., Izurieta, A., Han, J.Y., Mendez, R., Parisian, A., Furnari, F., AACR Publications, 2019. The dissertation author was the primary investigator and author of this paper.

Chapter 6

Conclusion

In this study we identify regions of epigenetic regulation within the first intron of the *EGFR* gene, characterizing DNA regions which are cell type specific in their H3K27Ac deposition, but contain conserved regions of open chromatin and histone acetylation within EGFR-expressing cells. These regions pass the threshold to be considered super enhancers and contain individual constituents which demonstrate functional attributes of active enhancers, including transcriptional enhancement in reporter assays, 3D interactions with the *EGFR* promoter, and negative regulation of their target gene when removed or repressed. We identify the presence and activity of AP-1 and BET transcription factors in the CE1 and CE2, and when the activity of these transcription factors is eliminated significant effects are seen in expression of target genes including *EGFR* and *JUN*, indicating direct AP-1 and BET dependency of these genes. Pharmacologic disruption of the transcription factor complexes at these enhancers has significant effects on *EGFR* expression, providing a mechanism by which this transcriptional control mechanism may be targeted.

As mentioned throughout, few studies have attempted to elucidate the mechanisms of transcriptional control of *EGFR*, and the few early studies primarily focused on the *EGFR* promoter as the primary unit mediating any control. Though this chapter and this study focus exclusively on CE1 and CE2 for the control of *EGFR* transcription, it is important that the promoters role is not lost as part of this regulatory circuit. Although the dCas9-KRAB

transcriptional repression system has been extensively validated through the silencing of distal enhancer elements [413, 414], its initial utility was described at the promoter regions of genes targeted for silencing [341, 342]. Keeping in mind the ACH hypothesis introduced in Chapter 2 [195] and the fact that enhancers and promoters interact with each other directly through DNA bending in 3D chromosomal space, it could be hypothesized there are transcriptional-silencing mechanisms beyond KRAB-mediated deposition of H3K9me3 or de-acetylation of H3K27 at an enhancer. Targeting a KRAB domain to an enhancer region known to be interacting with a promoter could exact the same KRAB-mediated deposition of H3K9me3 or de-acetylation of H3K27 at the promoter of that gene at the same time. To my knowledge, this has not been experimentally validated, however evidence we have shown here indicates that this may be the case.

In chapter 4 Figure 4.4, in HN12 cells compared to the O-T control, the CE1.4 and CE2.4 gRNA targeted regions had significantly decreased enrichment of H3K27Ac at the opposing targeted region (e.g. H3K27Ac loss at CE1.4 when 2.4 was targeted). If these regions are cooperatively interacting with the *EGFR* promoter, as suggested by the strongest repression occurring when either both enhancers are deleted (Figure 4.3) or repressed (Figure 4.4, Figure 4.5), then the association with dCas9-KRAB at a specific region within CE1 or CE2 could be mediating the identified de-acetylation of H3K27 at the other CE. It stands to reason that this phenomenon could also be occurring at the *EGFR* promoter, and H3K27Ac ChIP-qPCR for regions around the promoter could help validate this hypothesis. Similarly, though significant repression of *EGFR* transcription was mediated by dCas9-KRAB targeting to CE1 and CE2, targeting additional KRAB domains specifically to the *EGFR* promoter could completely shut down the expression of the gene in combination with CE1 and CE2 repression. Further experiments are required to confirm this hypothesis in our cell line models.

Elevated EGFR is a well-established therapeutic target, however responses to EGFR tyrosine kinase inhibitors (TKI) are sporadic [66,70,72,93,133]. The mechanisms behind resistance to EGFR TKI are unique to each tumor type, and secondary resistance mutations are common, particularly in NSCLC [105, 109]. In HNSCC, high *EGFR* copy numbers are statistically associated

with cetuximab and gefitinib resistance [415], and although rare, kinase domain mutations may be associated with altered responses to EGFR inhibitors [416]. In GBM, resistance mechanisms are less well understood with prevailing theories including intratumoral heterogeneity [417], *EGFR* amplified extrachromosomal DNA (ecDNA) [48], and loss of PTEN [76]. With these challenges in mind, this study presents a kinase domain independent mechanism by which *EGFR* expression and activity can be prevented. Recent studies have shown targeted transcription factor blockade can overcome EGFR TKI resistance [418], thus this study presents an additional set of pathways which can be targeted alone or in combination with EGFR TKIs to treat EGFR-positive tumors.

Evidence for the effectiveness of this strategy has been published before by our lab and others. Treatment of mice harboring GBM neurosphere PDX models with the pan-BET protein inhibitor JQ1 significantly prolongs survival, and combination of JQ1 with anti-EGFR TKI further increases this effect [399]. This anti-EGFR effect of JQ1 is at least partially attributable to disruption of genes which associate with enhancers co-occupied by AP-1 and BRD4 [395]. Both BRD4 and AP-1 have been shown to interact with or co-occupy NF-kB targets [392,393,400], thus small molecule inhibitors targeting these molecules should prove effective in EGFR-overexpressing cells. For example, SP100030 is a molecule which shows potent inhibition of NF-kB and AP-1 transcription activation [419] and has been shown to be effective inhibiting AP-1 and NF-kB targets [420]. Other molecules targeting AP-1 binding to DNA [421] have even been utilized in clinical trials [422]. Further studies should investigate the effectiveness of these anti-AP-1 compounds in EGFR-positive tumors, both as single agents and in combination. Further, additional studies should interrogate the CE1 and CE2 in *EGFR*-mutated models to identify if the use of transcription factor blocking molecules can rationally be expected to have increased effects regardless of the status of the kinase domain of the protein.

Though the preclinical evidence for utilizing JQ1 is compelling, JQ1 has limited use in the clinic. JQ1 is a pan-BET inhibitor, and thus is not selective for any BRD proteins [401]. Initial hope for JQ1 was primarily based on its potent inhibition of BRD4, a protein which was known to cooperate with MYC to drive the transcription of oncogenic genes [423]. However, because JQ1 targets BRD2 and BRD3 in addition to BRD4, the off-target effects are significant.

Preclinical studies showed that knocking out BRD4 by RNAi produced widespread stem cell depletion [424], and BRD2 knockout mice are not viable [425]. JQ1 has a very short half-life and dose concentrations required to mediate single agent activity are above physiologic safety levels *in vivo* [426], therefore it has not even been tested in clinical trials. Other pan-BET inhibitors like, ABBV-075 [427], I-BET151 [428], CPI203 [429], and PFI-1 [430] function in a similar anti-MYC manner and have shown preclinical benefit, however clinical trials featuring these inhibitors have generally not passed phase I as single therapies. With the significant off-target effects of pan-BET inhibitors in mind, more selective compounds have begun to be developed. OTX015 is a BRD2/4 inhibitor [431], and RVX-208 is a BRD3 specific inhibitor [432]. OTX015 has progressed furthest and shown the most effectiveness overall of these inhibitors [433, 434], but still was withdrawn at Phase II. Toxicities for these drugs are often found in the stem cell populations, most often leading to thrombocytopenia (platelet loss) and anemia [435–438].

Additionally, it appears that BRD2/3/4 are not fully functionally redundant [425, 439] meaning pan-disruption is likely causing widespread toxicities. Because of their lack of selectivity, actual pan-BET bromodomain inhibitors constitute a major limiting issue for the elucidation of specific functions of BRD2, BRD3 or BRD4. Solving this problem by providing selective compounds of individual BET proteins is a crucial challenge to understand their biological roles, to unravel the molecular mechanisms of their signaling, and to develop relevant targeted therapeutic strategies. For the purposes of *EGFR* transcription, it will be important for future experiments to further specify which BRD protein is critical for driving enhancer activity and utilizing BRD protein-specific inhibitors, possibly in combination with EGFR TKIs, to identify truly effective combination therapies.

The mechanisms for *EGFR* transcriptional control which we have described here may have implications for *EGFR* mutations, specifically the EGFRvIII ED mutation. EGFRvIII is a critical oncogene in GBM, present in 30-60% of primary GBMs depending on the method of identification [43, 44]. As mentioned earlier, EGFRvIII results from in-frame deletion of 801 base pairs spanning exons 27 of the coding sequence [41, 42]. EGFRvIII shows constitutive tyrosine kinase activity [39] and is highly tumorigenic [40]. In GBM every incidence of *EGFRvIII*

has unique genomic breakpoints within intron 1 and intron 8, removing exons 2-7 including large fragments of intron 1 depending on breakpoint location, often including regions which we have identified as enhancers [440] (Figure 3.1). It is important to note the majority of identified *EGFRvIII* breakpoints occur nearer the 3 end of *EGFR* intron 1, often removing the region which we have identified as CE2 [440]. Interestingly, in some of our cell line models the data indicates a more significant role for CE2 in interaction with the *EGFR* promoter by 3C (Figure 3.6) and influence on the expression of EGFR by dCas9-KRAB (Figure 4.4, Figure 4.5). In other models the utilization of each CE is either equal (Figure 4.5) or demonstrates a CE1 bias Figure 4.2B, SF767/Cal27). But, targeting both CEs always produces a combinatorial effect (Figure 4.3B, Figure 4.4D, Figure 4.5D). These data demonstrate an intricate regulatory system that is cell type-specific and makes predicting the functional consequences of intron 1-loss on EGFRvIII expression difficult.

Further research should interrogate directly how extent of intron 1-loss affects EGFRvIII through analysis of patient derived and/or CRISPR/Cas9-generated EGFRvIII+ models. Engineered cells which express homozygous EGFRvIII in an isogenic background would provide a few advantages to directly study the effect of intron 1 loss in *EGFR*. First, progressively larger deletions could be made and compared against an EGFRvIII missing the minimal amount of intron 1. Second, each engineered cell line would have the same driver and passenger mutations, ensuring no compensatory pathways can activate the intron 1 enhancers. We would expect that loss of CE2 alone would either lead to increased activity of CE1, or overall decreased transcriptional activation of *EGFR* expression. This could be measured by ChIP-seq for H3K27Ac at the remaining enhancer regions paired with qPCR to assess *EGFR* expression levels. However, to make any clinically relevant conclusions, experiments would need to be further validated using neurospheres or clinical samples expressing EGFRvIII.

This study focuses exclusively on non-amplified *EGFR*, however significant fractions of both HNSCC and GBM tumors have high copy numbers of the gene [26, 118]. In GBM, *EGFR* can be amplified both as a homogeneously staining region (HSR) or on extrachromosomal DNA (ecDNA) [48], both of which include the entire *EGFR* gene and surrounding regions. Ampli-

fied enhancer regions maintain their enhancer signatures [348] and focal amplifications of super enhancers have been shown to drive transcription [353]. Previous data in our lab has shown this to be significant specifically in GBM neurospheres which contain ecDNA with EGFRvIII, as treatment with JQ1 significantly increases survival [399]. To most accurately establish a link between JQ1 and *EGFR* transcriptional inhibition in amplified *EGFR/EGFRvIII* cases, ChIP-seq for direct JQ1 targets (e.g. BRD4) should be done in treatment naïve and treated neurospheres. Because *EGFR* is often amplified on ecDNA and ecDNA is often stitched together from many non-adjacent genomic regions [441, 442], it is conceivable that other amplified genes on ecDNA are directly able to increase transcription through epigenetic interactions. Further analysis of ecDNA in a wider range of GBM clinical samples should be done to confirm whether chromatin-altering proteins are amplified on ecDNA at any meaningful rate. These data support a role for *EGFR* intron 1 enhancers in the control of amplified *EGFR* on ecDNA, and future studies should further explore the chromatin landscape of ecDNA and utilize treatment strategies which have already been discussed for disrupting the transcriptional activation of *EGFR* by intron 1 enhancers.

In conclusion, we found that *EGFR* expression is maintained in part through presence and activity of critical constituent enhancers present in intron 1 of the gene. Characterization of CE1 and CE2 in multiple cell line systems identified a novel role for BET transcriptional co-activators and AP-1 transcription factors in these enhancers, and provided the rationale for therapeutic targeting of *EGFR* through perturbation of BET and AP-1 in EGFR-positive malignancies. Therapeutic targeting strategies for mitigating the oncogenic effects of active EGFR, though numerous and wide in scope, have proven largely ineffective in many cancer types including HNSCC and GBM. Thus, we present a previously uncharacterized *EGFR* expression mechanism which we believe is an important modality which can be targeted in EGFR-positive malignancies (Figure 6.1).

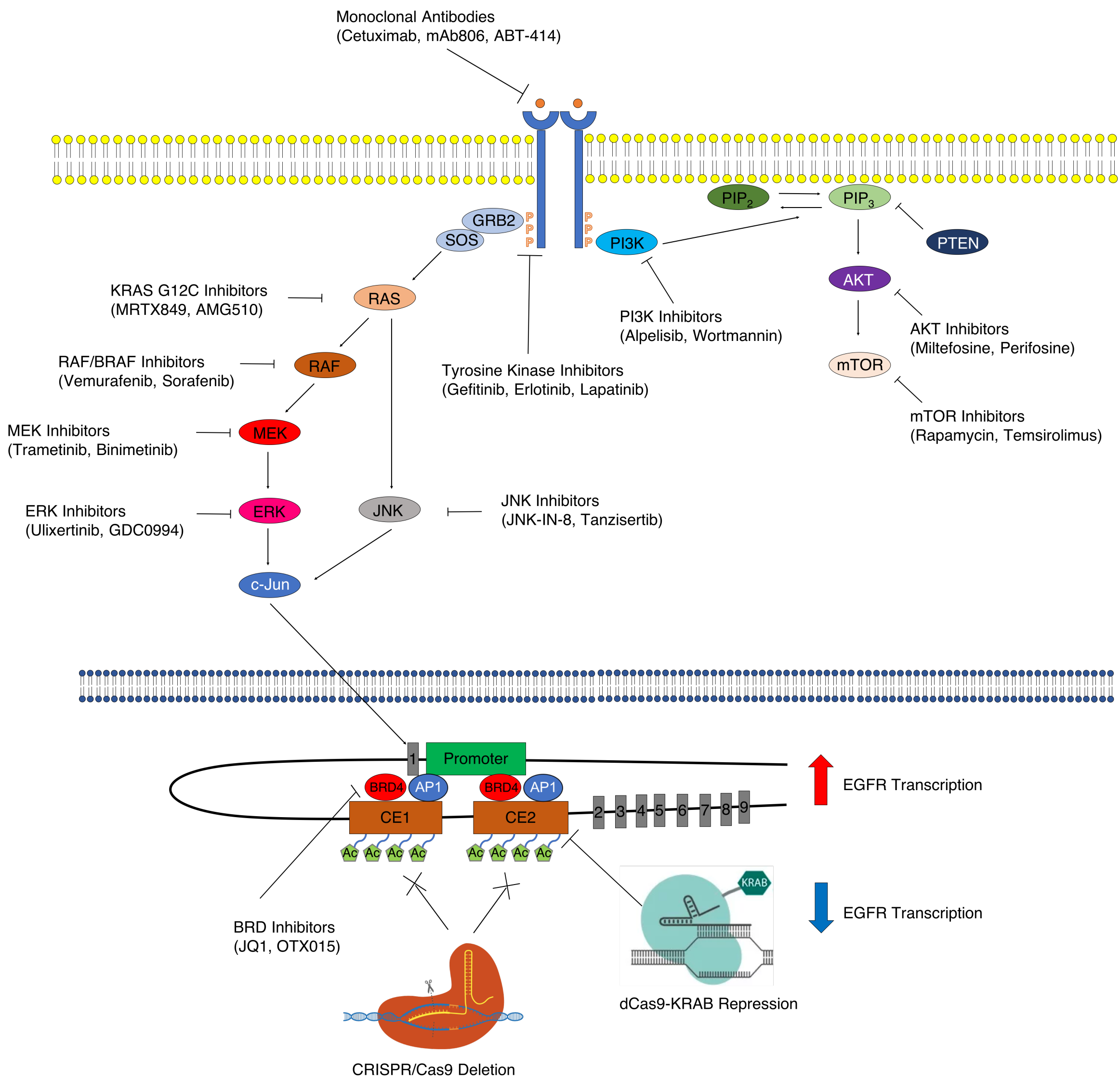


Figure 6.1: Summary of current anti-EGFR strategies and findings from this study.

Bibliography

[1] Carpenter G, King L, Cohen S (1978) Epidermal growth factor stimulates phosphorylation in membrane preparations in vitro [21]. Nature 276: 409–410.

[2] Schlessinger J (2000) Cell Signaling by Receptor Tyrosine Kinases. Cell 103: 211–225.

[3] Cohen S (1960) PURIFICATION OF A NERVE-GROWTH PROMOTING PROTEIN FROM THE MOUSE SALIVARY GLAND AND ITS NEURO-CYTOTOXIC ANTI-SERUM. Proceedings of the National Academy of Sciences of the United States of America 46: 302–11.

[4] Massague J (1990) Transforming growth factor-α. A model for membrane-anchored growth factors. Technical Report 35. URL `http://www.jbc.org/`.

[5] Shoyab M, McDonald VL, Bradley JG, Todaro GJ (1988) Amphiregulin: a bifunctional growth-modulating glycoprotein produced by the phorbol 12-myristate 13-acetate-treated human breast adenocarcinoma cell line MCF-7. Proceedings of the National Academy of Sciences 85: 6528–6532.

[6] Toyoda H, Komurasaki T, Uchida D, Takayama Y, Isobe T, Okuyama T, Hanada K (1995) Epiregulin. Journal of Biological Chemistry 270: 7495–7500.

[7] Strachan L, Murison JG, Prestidge RL, Sleeman MA, Watson JD, Kumble KD (2001) Cloning and Biological Activity of Epigen, a Novel Member of the Epidermal Growth Factor Superfamily. Journal of Biological Chemistry 276: 18265–18271.

[8] Higashiyama S, Lau K, Besner GE, Abraham JA, Klagsbrun M (1992) Structure of heparin-binding EGF-like growth factor. Multiple forms, primary structure, and glycosylation of the mature protein. The Journal of biological chemistry 267: 6205–12.

[9] Schlessinger J (1988) Signal transduction by allosteric receptor oligomerization. Trends in Biochemical Sciences 13: 443–447.

[10] Pawson T, Schlessingert J (1993) SH2 and SH3 domains. Current Biology 3: 434–442.

[11] Olayioye MA, Neve RM, Lane HA, Hynes NE (2000) The ErbB signaling network: receptor heterodimerization in development and cancer. The EMBO Journal 19: 3159–3167.

[12] Jo M, Stolz DB, Esplen JE, Dorko K, Michalopoulos GK, Strom SC (2000) Cross-talk between Epidermal Growth Factor Receptor and c-Met Signal Pathways in Transformed Cells. Journal of Biological Chemistry 275: 8806–8811.

[13] Graus-Porta D, Beerli RR, Daly JM, Hynes NE (1997) ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. The EMBO journal 16: 1647–55.

[14] Pinkas-Kramarski R, Soussan L, Waterman H, Levkowitz G, Alroy I, Klapper L, Lavi S, Seger R, Ratzkin BJ, Sela M, Yarden Y (1996) Diversification of Neu differentiation factor and epidermal growth factor signaling by combinatorial receptor interactions. The EMBO journal 15: 2452–67.

[15] Lowenstein E, Daly R, Batzer A, Li W, Margolis B, Lammers R, Ullrich A, Skolnik E, Bar-Sagi D, Schlessinger J (1992) The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. Cell 70: 431–442.

[16] Liebmann C (2001) Regulation of MAP kinase activity by peptide receptor signalling pathway: paradigms of multiplicity. Cellular signalling 13: 777–85.

[17] Hallberg B, Rayter SI, Downward J (1994) Interaction of Ras and Raf in intact mammalian cells upon extracellular stimulation. The Journal of biological chemistry 269: 3913–6.

[18] Hill CS, Treisman R (1995) Transcriptional regulation by extracellular signals: mechanisms and specificity. Cell 80: 199–211.

[19] Rajalingam K, Schreck R, Rapp UR, Albert (2007) Ras oncogenes and their downstream targets. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1773: 1177–1195.

[20] Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, Teague JW, Campbell PJ, Stratton MR, Futreal PA (2011) COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Research 39: D945-D950.

[21] Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, Ewing R, Floyd Y, Gray K, Hall S, Hawes R, Hughes J, Kosmidou V, Menzies A, Mould C, Parker A, Stevens C, Watt S, Hooper S, Wilson R, Jayatilake H, Gusterson BA, Cooper C, Shipley J, Hargrave D, Pritchard-Jones K, Maitland N, Chenevix-Trench G, Riggins GJ, Bigner DD, Palmieri G, Cossu A, Flanagan A, Nicholson A, Ho JWC, Leung SY, Yuen ST, Weber BL, Seigler HF, Darrow TL, Paterson H, Marais R, Marshall CJ, Wooster R, Stratton MR, Futreal PA (2002) Mutations of the BRAF gene in human cancer. Nature 417: 949–954.

[22] Vivanco I, Sawyers CL (2002) The phosphatidylinositol 3-KinaseAKT pathway in human cancer. Nature Reviews Cancer 2: 489–501.

[23] Manning BD, Cantley LC (2007) AKT/PKB Signaling: Navigating Downstream. Cell 129: 1261–1274.

[24] Ali IU, Schriml LM, Dean M (1999) Mutational Spectra of PTEN/MMAC1 Gene: a Tumor Suppressor With Lipid Phosphatase Activity. JNCI Journal of the National Cancer Institute 91: 1922–1932.

[25] Choe G, Horvath S, Cloughesy TF, Crosby K, Seligson D, Palotie A, Inge L, Smith BL, Sawyers CL, Mischel PS (2003) Analysis of the phosphatidylinositol 3'-kinase signaling pathway in glioblastoma patients in vivo. Cancer research 63: 2742–6.

[26] Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, Zheng S, Chakravarty D, Sanborn JZ, Berman SH, Beroukhim R, Bernard B, Wu CJ, Genovese G, Shmulevich I, Barnholtz-Sloan J, Zou L, Vegesna R, Shukla SA, Ciriello G, Yung WK, Zhang W, Sougnez C, Mikkelsen T, Aldape K, Bigner DD, Van Meir EG, Prados M, Sloan A, Black KL, Eschbacher J, Finocchiaro G, Friedman W, Andrews DW, Guha A, Iacocca M, O'Neill BP, Foltz G, Myers J, Weisenberger DJ, Penny R, Kucherlapati R, Perou CM, Hayes DN, Gibbs R, Marra M, Mills GB, Lander E, Spellman P, Wilson R, Sander C, Weinstein J, Meyerson M, Gabriel S, Laird PW, Haussler D, Getz G, Chin L, Network TR (2013) The somatic genomic landscape of glioblastoma. Cell 155: 462–477.

[27] Sanchez-Vega F, Mina M, Armenia J, Ciriello G, Sander C, Schultz N, Chatila WK, Luna A, La KC, Dimitriadoy S, Liu DL, Kantheti HS, Saghafinia S, Chakravarty D, Daian F, Gao Q, Bailey MH, Liang WW, Foltz SM, Shmulevich I, Ding L, Heins Z, Ochoa A, Gross B, Gao J, Zhang H, Kundra R, Kandoth C, Bahceci I, Dervishi L, Dogrusoz U, Zhou W, Shen H, Laird PW, Way GP, Greene CS, Liang H, Xiao Y, Wang C, Iavarone A, Berger AH, Bivona TG, Lazar AJ, Hammer GD, Giordano T, Kwong LN (2018) Oncogenic Signaling Pathways in The Cancer Genome Atlas Article Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell 173: 321–337.

[28] Ostrom QT, Gittleman H, Truitt G, Boscia A, Kruchko C, Barnholtz-Sloan JS (2018) CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 20112015. Neuro-Oncology 20: iv1-iv86.

[29] Furnari FB, Fenton T, Bachoo RM, Mukasa A, Stommel JM, Stegh A, Hahn WC, Ligon KL, Louis DN, Brennan C, Chin L, DePinho RA, Cavenee WK (2007) Malignant astrocytic glioma: genetics, biology, and paths to treatment. Genes & Development 21: 2683–2710.

[30] Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW (2016) The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta neuropathologica 131: 803–20.

[31] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O'Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, Hodgson JG, James CD, Sarkaria JN, Brennan C, Kahn A, Spellman PT, Wilson RK, Speed TP, Gray JW, Meyerson M, Getz G, Perou CM, Hayes DN, Cancer Genome Atlas Research Network (2010) Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17: 98–110.

[32] Shinojima N, Tada K, Shiraishi S, Kamiryo T, Kochi M, Nakamura H, Makino K, Saya H, Hirano H, Kuratsu J, Oka K, Ishimaru Y, Ushio Y (2003) Prognostic value of epidermal growth factor receptor in patients with glioblastoma multiforme. Cancer Res 63: 6962–6970.

[33] Bigner SH, Mark J, Bigner DD (1990) Cytogenetics of human brain tumors. Cancer genetics and cytogenetics 47: 141–54.

[34] You J, Croyle JL, Nishimura A, Ozato K, Howley PM (2004) Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. Cell 117: 349–60.

[35] Turner KM, Deshpande V, Beyter D, Koga T, Rusert J, Lee C, Li B, Arden K, Ren B, Nathanson DA, Kornblum HI, Taylor MD, Kaushal S, Cavenee WK, Wechsler-Reya R, Furnari FB, Vandenberg SR, Rao PN, Wahl GM, Bafna V, Mischel PS (2017) Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature 543: 122–125.

[36] Nicholas MK, Lukas RV, Jafri NF, Faoro L, Salgia R (2006) Epidermal Growth Factor Receptor-Mediated Signal Transduction in the Development and Therapy of Gliomas. Clinical Cancer Research 12: 7261–7270.

[37] Kastenhuber ER, Huse JT, Berman SH, Pedraza A, Zhang J, Suehara Y, Viale A, Cavatore M, Heguy A, Szerlip N, Ladanyi M, Brennan CW (2014) Quantitative assessment of intragenic receptor tyrosine kinase deletions in primary glioblastomas: their prevalence and molecular correlates. Acta Neuropathologica 127: 747–759.

[38] Jungbluth AA, Stockert E, Huang HJ, Collins VP, Coplan K, Iversen K, Kolb D, Johns TJ, Scott AM, Gullick WJ, Ritter G, Cohen L, Scanlan MJ, Cavenee WK, Old LJ, Cavanee WK (2003) A monoclonal antibody recognizing human cancers with amplification/overexpression of the human epidermal growth factor receptor. Proc Natl Acad Sci U S A 100: 639–644.

[39] Huang HS, Nagane M, Klingbeil CK, Lin H, Nishikawa R, Ji XD, Huang CM, Gill GN, Wiley HS, Cavenee WK (1997) The enhanced tumorigenic activity of a mutant epidermal growth factor receptor common in human cancers is mediated by threshold levels of constitutive tyrosine phosphorylation and unattenuated signaling. J Biol Chem 272: 2927–2935.

[40] Nishikawa R, Ji XD, Harmon RC, Lazar CS, Gill GN, Cavenee WK, Huang HJ (1994) A mutant epidermal growth factor receptor common in human glioma confers enhanced tumorigenicity. Proceedings of the National Academy of Sciences of the United States of America 91: 7727–7731.

[41] Humphrey PA, Wong AJ, Vogelstein B, Friedman HS, Werner MH, Bigner DD, Bigner SH (1988) Amplification and expression of the epidermal growth factor receptor gene in human glioma xenografts. Cancer research 48: 2231–8.

[42] Yamazaki H, Ohba Y, Tamaoki N, Shibuya M (1990) A deletion mutation within the ligand binding domain is responsible for activation of epidermal growth factor receptor gene in human brain tumors. Jpn J Cancer Res 81: 773–779.

[43] Aldape KD, Ballman K, Furth A, Buckner JC, Giannini C, Burger PC, Scheithauer BW, Jenkins RB, James CD (2004) Immunohistochemical Detection of EGFRvIII in High Malignancy Grade Astrocytomas and Evaluation of Prognostic Significance. Journal of Neuropathology & Experimental Neurology 63: 700–707.

[44] Saikali S, Avril T, Collet B, Hamlat A, Bansard JY, Drenou B, Guegan Y, Quillien V (2006) Expression of nine tumour antigens in a series of human glioblastoma multiforme: interest of EGFRvIII, IL-13Rα2, gp100 and TRP-2 for immunotherapy. Journal of Neuro-Oncology 81: 139–148.

[45] Sok JC, Coppelli FM, Thomas SM, Lango MN, Xi S, Hunt JL, Freilino ML, Graner MW, Wikstrand CJ, Bigner DD, Gooding WE, Furnari FB, Grandis JR (2006) Mutant Epidermal Growth Factor Receptor (EGFRvIII) Contributes to Head and Neck Cancer Growth and Resistance to EGFR Targeting. Clinical Cancer Research 12: 5064–5073.

[46] Schmidt MHH, Furnari FB, Cavenee WK, Bogler O (2003) Epidermal growth factor receptor signaling intensity determines intracellular protein interactions, ubiquitination, and internalization. Proceedings of the National Academy of Sciences 100: 6505–6510.

[47] Lopez-Gines C, Gil-Benso R, Ferrer-Luna R, Benito R, Serna E, Gonzalez-Darder J, Quilis V, Monleon D, Celda B, Cerdá-Nicolas M (2010) New pattern of EGFR amplification in glioblastoma and the relationship of gene copy number with gene expression profile. Modern Pathology 23: 856–865.

[48] Nathanson Da, Gini B, Mottahedeh J, Visnyei K, Koga T, Gomez G, Eskin A, Hwang K, Wang J, Masui K, Paucar A, Yang H, Ohashi M, Zhu S, Wykosky J, Reed R, Nelson SF, Cloughesy TF, James CD, Rao PN, Kornblum HI, Heath JR, Cavenee WK, Furnari FB, Mischel PS (2014) Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. Science (New York, NY) 343: 72–6.

[49] Bonavia R, Inda MdM, Cavenee WK, Furnari FB (2011) Heterogeneity maintenance in glioblastoma: a social network. Cancer research 71: 4055–4060.

[50] Inda MdM, Bonavia R, Mukasa A, Narita Y, Sah DWY, Vandenberg S, Brennan C, Johns TG, Bachoo R, Hadwiger P, Tan P, Depinho RA, Cavenee W, Furnari F (2010) Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. Genes & development 24: 1731–1745.

[51] Weiner LM, Surana R, Wang S (2010) Monoclonal antibodies: versatile platforms for cancer immunotherapy. Nature Reviews Immunology 10: 317–327.

[52] Rivera F, Vega-Villegas ME, López-Brea MF (2008) Cetuximab, its clinical use and future perspectives. Anti-Cancer Drugs 19: 99–113.

[53] Eller JL, Longo SL, Hicklin DJ, Canute GW (2002) Activity of Anti-epidermal Growth Factor Receptor Monoclonal Antibody C225 against Glioblastoma Multiforme. Neurosurgery 51: 1005–1014.

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

[55] Hasselbalch B, Lassen U, Hansen S, Holmberg M, Sørensen M, Kosteljanetz M, Broholm H, Stockhausen MT, Poulsen HS (2010) Cetuximab, bevacizumab, and irinotecan for patients with primary glioblastoma and progression after radiation therapy and temozolomide: a phase II trial. Neuro-Oncology 12: 508–16.

[56] Combs SE, Heeger S, Haselmann R, Edler L, Debus J, Schulz-Ertner D (2006) Treatment of primary glioblastoma multiforme with cetuximab, radiotherapy and temozolomide (GERT) phase I/II trial: study protocol. BMC Cancer 6: 133.

[57] Mishima K, Johns TG, Luwor RB, Scott AM, Stockert E, Jungbluth AA, Ji XD, Suvarna P, Voland JR, Old LJ, Huang HJ, Cavenee WK (2001) Growth suppression of intracranial xenografted glioblastomas overexpressing mutant epidermal growth factor receptors by systemic administration of monoclonal antibody (mAb) 806, a novel monoclonal antibody directed to the receptor. Cancer research 61: 5349–54.

[58] Gan HK, Lappas M, Cao DX, Cvrljevdic A, Scott AM, Johns TG (2009) Targeting a unique EGFR epitope with monoclonal antibody 806 activates NF-κB and initiates tumour vascular normalization. Journal of Cellular and Molecular Medicine 13: 3993–4001.

[59] Luwor RB, Johns TG, Murone C, Huang HJ, Cavenee WK, Ritter G, Old LJ, Burgess AW, Scott AM (2001) Monoclonal antibody 806 inhibits the growth of tumor xenografts expressing either the de2-7 or amplified epidermal growth factor receptor (EGFR) but not wild-type EGFR. Cancer research 61: 5355–61.

[60] Phillips AC, Boghaert ER, Vaidya KS, Mitten MJ, Norvell S, Falls HD, DeVries PJ, Cheng D, Meulbroek JA, Buchanan FG, McKay LM, Goodwin NC, Reilly EB (2016) ABT-414, an Antibody-Drug Conjugate Targeting a Tumor-Selective EGFR Epitope. Molecular Cancer Therapeutics 15: 661–669.

[61] van den Bent M, Gan HK, Lassman AB, Kumthekar P, Merrell R, Butowski N, Lwin Z, Mikkelsen T, Nabors LB, Papadopoulos KP, Penas-Prado M, Simes J, Wheeler H, Walbert T, Scott AM, Gomez E, Lee HJ, Roberts-Rapp L, Xiong H, Bain E, Ansell PJ, Holen KD, Maag D, Reardon DA (2017) Efficacy of depatuxizumab mafodotin (ABT-414) monotherapy in patients with EGFR-amplified, recurrent glioblastoma: results from a multi-center, international study. Cancer Chemotherapy and Pharmacology 80: 1209–1217.

[62] Gan HK, Fichtel L, Lassman AB, Merrell R, Van Den Bent MJ, Kumthekar P, Scott AM, Pedersen M, Gomez E, Fischer J, Ames W, Xiong H, Dudley MW, Munasinghe W, Roberts-Rapp L, Ansell P, Holen KD, Reardon DA (2014) A phase 1 study evaluating ABT-414 in combination with temozolomide (TMZ) for subjects with recurrent or unresectable glioblastoma (GBM). Journal of Clinical Oncology 32: 2021–2021.

[63] AbbVie Provides Update on Depatuxizumab Mafodotin (Depatux-M), an Investigational Medicine for Newly Diagnosed Glioblastoma, an Aggressive Form of Brain Cancer — AbbVie News Center. URL `https://news.abbvie.com/news/press-releases/abbvie-provides-update-on-depatuxizumab-mafodotin-depatux-m-an-investigational-medicine-for-newly-diagnosed-glioblastoma-an-aggressive-form-brain-cancer.htm`.

[64] Mischel PS, Cloughesy TF (2006) Targeted Molecular Therapy of GBM. Brain Pathology 13: 52–61.

[65] Shawver LK, Slamon D, Ullrich A (2002) Smart drugs: tyrosine kinase inhibitors in cancer therapy. Cancer cell 1: 117–23.

[66] Peereboom DM, Shepard DR, Ahluwalia MS, Brewer CJ, Agarwal N, Stevens GHJ, Suh JH, Toms SA, Vogelbaum MA, Weil RJ, Elson P, Barnett GH (2010) Phase II trial of erlotinib with temozolomide and radiation in patients with newly diagnosed glioblastoma multiforme. Journal of Neuro-Oncology 98: 93–99.

[67] Wen PY, Chang SM, Lamborn KR, Kuhn JG, Norden AD, Cloughesy TF, Robins HI, Lieberman FS, Gilbert MR, Mehta MP, Drappatz J, Groves MD, Santagata S, Ligon AH, Yung WKA, Wright JJ, Dancey J, Aldape KD, Prados MD, Ligon KL (2014) Phase I/II study of erlotinib and temsirolimus for patients with recurrent malignant gliomas: North American Brain Tumor Consortium trial 04-02. Neuro-Oncology 16: 567–578.

[68] Raizer JJ, Giglio P, Hu J, Groves M, Merrell R, Conrad C, Phuphanich S, Puduvalli VK, Loghin M, Paleologos N, Yuan Y, Liu D, Rademaker A, Yung WK, Vaillant B, Rudnick J, Chamberlain M, Vick N, Grimm S, Tremont-Lukats IW, De Groot J, Aldape K, Gilbert MR, Brain Tumor Trials Collaborative (2016) A phase II study of bevacizumab and erlotinib after radiation and temozolomide in MGMT unmethylated GBM patients. Journal of Neuro-Oncology 126: 185–192.

[69] Uhm JH, Ballman KV, Wu W, Giannini C, Krauss J, Buckner JC, James C, Scheithauer BW, Behrens RJ, Flynn PJ, Schaefer PL, Dakhill SR, Jaeckle KA (2011) Phase II Evaluation of Gefitinib in Patients With Newly Diagnosed Grade 4 Astrocytoma: Mayo/North Central Cancer Treatment Group Study N0074. International Journal of Radiation Oncology*Biology*Physics 80: 347–353.

[70] Chakravarti A, Wang M, Robins HI, Lautenschlaeger T, Curran WJ, Brachman DG, Schultz CJ, Choucair A, Dolled-Filhart M, Christiansen J, Gustavson M, Molinaro A, Mischel P, Dicker AP, Bredel M, Mehta M (2013) RTOG 0211: A Phase 1/2 Study of Radiation Therapy With Concurrent Gefitinib for Newly Diagnosed Glioblastoma Patients. International Journal of Radiation Oncology*Biology*Physics 85: 1206–1211.

[71] Gril B, Palmieri D, Bronder JL, Herring JM, Vega-Valle E, Feigenbaum L, Liewehr DJ, Steinberg SM, Merino MJ, Rubin SD, Steeg PS (2008) Effect of Lapatinib on the Outgrowth of Metastatic Breast Cancer Cells to the Brain. JNCI Journal of the National Cancer Institute 100: 1092–1103.

[72] Thiessen B, Stewart C, Tsao M, Kamel-Reid S, Schaiquevich P, Mason W, Easaw J, Belanger K, Forsyth P, McIntosh L, Eisenhauer E (2010) A phase I/II trial of GW572016 (lapatinib) in recurrent glioblastoma multiforme: clinical outcomes, pharmacokinetics and molecular correlation. Cancer Chemotherapy and Pharmacology 65: 353–361.

[73] Reardon DA, Groves MD, Wen PY, Nabors L, Mikkelsen T, Rosenfeld S, Raizer J, Barriuso J, McLendon RE, Suttle AB, Ma B, Curtis CM, Dar MM, de Bono J (2013) A Phase I/II Trial of Pazopanib in Combination with Lapatinib in Adult Patients with Relapsed Malignant Glioma. Clinical Cancer Research 19: 900–908.

[74] Lassman AB, Rossi MR, Raizer JJ, Razier JR, Abrey LE, Lieberman FS, Grefe CN, Lamborn K, Pao W, Shih AH, Kuhn JG, Wilson R, Nowak NJ, Cowell JK, DeAngelis LM, Wen P, Gilbert MR, Chang S, Yung WA, Prados M, Holland EC (2005) Molecular study of malignant gliomas treated with epidermal growth factor receptor inhibitors: tissue analysis from North American Brain Tumor Consortium Trials 01-03 and 00-01. Clin Cancer Res 11: 7841–7850.

[75] Vivanco I, Robins HI, Rohle D, Campos C, Grommes C, Nghiemphu PL, Kubek S, Oldrini B, Chheda MG, Yannuzzi N, Tao H, Zhu S, Iwanami A, Kuga D, Dang J, Pedraza A, Brennan CW, Heguy A, Liau LM, Lieberman F, Yung WKA, Gilbert MR, Reardon DA, Drappatz J, Wen PY, Lamborn KR, Chang SM, Prados MD, Fine HA, Horvath S, Wu N, Lassman AB, DeAngelis LM, Yong WH, Kuhn JG, Mischel PS, Mehta MP, Cloughesy TF, Mellinghoff IK (2012) Differential Sensitivity of Glioma- versus Lung CancerSpecific EGFR Mutations to EGFR Kinase Inhibitors. Cancer Discovery 2: 458–471.

[76] Mellinghoff IK, Wang MY, Vivanco I, Haas-Kogan DA, Zhu S, Dia EQ, Lu KV, Yoshimoto K, Huang JH, Chute DJ, Riggs BL, Horvath S, Liau LM, Cavenee WK, Rao PN, Beroukhim R, Peck TC, Lee JC, Sellers WR, Stokoe D, Prados M, Cloughesy TF, Sawyers CL, Mischel PS (2005) Molecular Determinants of the Response of Glioblastomas to EGFR Kinase Inhibitors. New England Journal of Medicine 353: 2012–2024.

[77] Stommel JM, Kimmelman AC, Ying H, Nabioullin R, Ponugoti AH, Wiedemeyer R, Stegh AH, Bradner JE, Ligon KL, Brennan C, Chin L, DePinho RA (2007) Coactivation of Receptor Tyrosine Kinases Affects the Response of Tumor Cells to Targeted Therapies. Science 318: 287–290.

[78] Guo G, Gong K, Ali S, Ali N, Shallwani S, Hatanpaa KJ, Pan E, Mickey B, Burma S, Wang DH, Kesari S, Sarkaria JN, Zhao D, Habib AA (2017) Primary resistance to EGFR inhibition in glioblastoma is mediated by a TNF-JNK-Axl-ERK signaling axis. Nature neuroscience 20: 1074.

[79] Yoshida Y, Ozawa T, Yao TW, Shen W, Brown D, Parsa AT, Raizer JJ, Cheng SY, Stegh AH, Mazar AP, Giles FJ, Sarkaria JN, Butowski N, Nicolaides T, James CD (2014) NT113, a Pan-ERBB Inhibitor with High Brain Penetrance, Inhibits the Growth of Glioblastoma Xenografts with EGFR Amplification. Molecular Cancer Therapeutics 13: 2919–2929.

[80] Wang Z, Dove P, Rosa D, Marshall A, Catalano T, Bossen B, Choi J, Wang C, Wong M, Winston J, Slassi M, Petrova PS, Uger RA (2017) DDIS-11. TTI-2341: ANOVEL, ORALLY BIOAVAILABLE, BRAIN-PENETRANT, COVALENT EPIDERMAL GROWTH FACTOR RECEPTOR (EGFR) INHIBITOR FOR TREATMENT OF GLIOBLASTOMA MULTIFORME (GBM) AND BRAIN METASTASES OF NON-SMALL CELL LUNG CANCER (NSCLC). Neuro-Oncology 19: vi61-vi61.

[81] Lee JC, Vivanco I, Beroukhim R, Huang JHY, Feng WL, DeBiasi RM, Yoshimoto K, King JC, Nghiemphu P, Yuza Y, Xu Q, Greulich H, Thomas RK, Paez JG, Peck TC, Linhart DJ, Glatt KA, Getz G, Onofrio R, Ziaugra L, Levine RL, Gabriel S, Kawaguchi T, O'Neill K, Khan H, Liau LM, Nelson SF, Rao PN, Mischel P, Pieper RO, Cloughesy T, Leahy DJ, Sellers WR, Sawyers CL, Meyerson M, Mellinghoff IK (2006) Epidermal

Growth Factor Receptor Activation in Glioblastoma through Novel Missense Mutations in the Extracellular Domain. PLoS Medicine 3: e485.

[82] Binder ZA, Thorne AH, Bakas S, Wileyto EP, Bilello M, Akbari H, Rathore S, Ha SM, Zhang L, Ferguson CJ, Dahiya S, Bi WL, Reardon DA, Idbaih A, Felsberg J, Hentschel B, Weller M, Bagley SJ, Morrissette JJ, Nasrallah MP, Ma J, Zanca C, Scott AM, Orellana L, Davatzikos C, Furnari FB, O'Rourke DM (2018) Epidermal Growth Factor Receptor Extracellular Domain Mutations in Glioblastoma Present Opportunities for Clinical Imaging and Therapeutic Development. Cancer Cell 34: 163–177.

[83] Ymer SI, Greenall SA, Cvrljevic A, Cao DX, Donoghue JF, Epa VC, Scott AM, Adams TE, Johns TG, Ymer SI, Greenall SA, Cvrljevic A, Cao DX, Donoghue JF, Epa VC, Scott AM, Adams TE, Johns TG (2011) Glioma Specific Extracellular Missense Mutations in the First Cysteine Rich Region of Epidermal Growth Factor Receptor (EGFR) Initiate Ligand Independent Activation. Cancers 3: 2032–2049.

[84] Jura N, Zhang X, Endres N, Seeliger M, Schindler T, Kuriyan J (2011) Catalytic Control in the EGF Receptor and Its Connection to General Kinase Regulatory Mechanisms. Molecular Cell 42: 9–22.

[85] Wood ER, Truesdale AT, McDonald OB, Yuan D, Hassell A, Dickerson SH, Ellis B, Pennisi C, Horne E, Lackey K, Alligood KJ, Rusnak DW, Gilmer TM, Shewchuk L (2004) A Unique Structure for Epidermal Growth Factor Receptor Bound to GW572016 (Lapatinib). Cancer Research 64: 6652–6659.

[86] Scott AM, Lee FT, Tebbutt N, Herbertson R, Gill SS, Liu Z, Skrinos E, Murone C, Saunder TH, Chappell B, Papenfuss AT, Poon AMT, Hopkins W, Smyth FE, MacGregor D, Cher LM, Jungbluth AA, Ritter G, Brechbiel MW, Murphy R, Burgess AW, Hoffman EW, Johns TG, Old LJ (2007) A phase I clinical trial with monoclonal antibody ch806 targeting transitional state and mutant epidermal growth factor receptors. Proceedings of the National Academy of Sciences 104: 4071–4076.

[87] Gan HK, Burgess AW, Clayton AHA, Scott AM (2012) Targeting of a Conformationally Exposed, Tumor-Specific Epitope of EGFR as a Strategy for Cancer Therapy. Cancer Research 72: 2924–2930.

[88] Walker F, Orchard SG, Jorissen RN, Hall NE, Zhang HH, Hoyne PA, Adams TE, Johns TG, Ward C, Garrett TPJ, Zhu HJ, Nerrie M, Scott AM, Nice EC, Burgess AW (2004) CR1/CR2 Interactions Modulate the Functions of the Cell Surface Epidermal Growth Factor Receptor. Journal of Biological Chemistry 279: 22387–22398.

[89] Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D (2011) Global cancer statistics. CA: A Cancer Journal for Clinicians 61: 69–90.

[90] Herbst RS, Heymach JV, Lippman SM (2008) Lung Cancer. New England Journal of Medicine 359: 1367–1380.

[91] Collins LG, Haines C, Perkel R, Enck RE (2007) Lung cancer: diagnosis and management. American family physician 75: 56–63.

[92] Dearden S, Stevens J, Wu YL, Blowers D (2013) Mutation incidence and coincidence in non small-cell lung cancer: meta-analyses by ethnicity and histology (mutMap). Annals of Oncology 24: 2371–2376.

[93] Cohen MH, Williams GA, Sridhara R, Chen G, McGuinn WD, Morse D, Abraham S, Rahman A, Liang C, Lostritto R, Baird A, Pazdur R (2004) United States Food and Drug Administration Drug Approval summary: Gefitinib (ZD1839; Iressa) tablets. Clinical cancer research : an official journal of the American Association for Cancer Research 10: 1212–8.

[94] Bailey R, Kris M, Wolf M, Kay A, Averbuch S, Askaa J, Janas M, Schmidt K, Fukuoka M (2003) O-242 Gefitinib (Iressa, ZD1839) monotherapy for pretreated advanced non-small-cell lung cancer in IDEAL 1 and 2: tumor response is not clinically relevantly predictable from tumor EGFR membrane staining alone. Lung Cancer 41: S71.

[95] Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, Harris PL, Haserlat SM, Supko JG, Haluska FG, Louis DN, Christiani DC, Settleman J, Haber DA (2004) Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of NonSmall-Cell Lung Cancer to Gefitinib. New England Journal of Medicine 350: 2129–2139.

[96] Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE, Meyerson M (2004) EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy. Science 304: 1497–1500.

[97] Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L, Mardis E, Kupfer D, Wilson R, Kris M, Varmus H (2004) EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proceedings of the National Academy of Sciences 101: 13306–13311.

[98] Massarelli E, Johnson FM, Erickson HS, Wistuba II, Papadimitrakopoulou V (2013) Uncommon Epidermal Growth Factor Receptor mutations in non-small cell lung cancer and their mechanisms of EGFR tyrosine kinase inhibitors sensitivity and resistance. Lung Cancer 80: 235–241.

[99] Sharma SV, Bell DW, Settleman J, Haber DA (2007) Epidermal growth factor receptor mutations in lung cancer. Nature Reviews Cancer 7: 169–181.

[100] Sordella R, Bell DW, Haber DA, Settleman J (2004) Gefitinib-Sensitizing EGFR Mutations in Lung Cancer Activate Anti-Apoptotic Pathways. Science 305: 1163–1167.

[101] Greulich H, Chen TH, Feng W, Jänne PA, Alvarez JV, Zappaterra M, Bulmer SE, Frank DA, Hahn WC, Sellers WR, Meyerson M (2005) Oncogenic Transformation by Inhibitor-Sensitive and -Resistant EGFR Mutants. PLoS Medicine 2: e313.

[102] Riely GJ, Pao W, Pham D, Li AR, Rizvi N, Venkatraman ES, Zakowski MF, Kris MG, Ladanyi M, Miller VA (2006) Clinical Course of Patients with Non-Small Cell Lung Cancer

and Epidermal Growth Factor Receptor Exon 19 and Exon 21 Mutations Treated with Gefitinib or Erlotinib. Clinical Cancer Research 12: 839–844.

[103] Bell DW, Lynch TJ, Haserlat SM, Harris PL, Okimoto RA, Brannigan BW, Sgroi DC, Muir B, Riemenschneider MJ, Iacona RB, Krebs AD, Johnson DH, Giaccone G, Herbst RS, Manegold C, Fukuoka M, Kris MG, Baselga J, Ochs JS, Haber DA (2005) Epidermal Growth Factor Receptor Mutations and Gene Amplification in NonSmall-Cell Lung Cancer: Molecular Analysis of the IDEAL/INTACT Gefitinib Trials. Journal of Clinical Oncology 23: 8081–8092.

[104] Paz-Ares L, Soulières D, Melezínek I, Moecks J, Keil L, Mok T, Rosell R, Klughammer B (2010) Clinical outcomes in non-small-cell lung cancer patients with EGFR mutations: pooled analysis. Journal of Cellular and Molecular Medicine 14: 51–69.

[105] Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, Bergethon K, Shaw AT, Gettinger S, Cosper AK, Akhavanfard S, Heist RS, Temel J, Christensen JG, Wain JC, Lynch TJ, Vernovsky K, Mark EJ, Lanuti M, Iafrate AJ, Mino-Kenudson M, Engelman JA (2011) Genotypic and Histological Evolution of Lung Cancers Acquiring Resistance to EGFR Inhibitors. Science Translational Medicine 3: 26–75.

[106] Yun CH, Mengwasser KE, Toms AV, Woo MS, Greulich H, Wong KK, Meyerson M, Eck MJ (2008) The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proceedings of the National Academy of Sciences 105: 2070–2075.

[107] Sequist LV, Goldman JW, Wakelee HA, Camidge DR, Yu HA, Varga A, Solomon B, Oxnard GR, Ou SHI, Papadimitrakopoulou V, Chao BH, Liu SV, Reckamp KL, Spira AI, Piotrowska Z, Despain D, Karlovich CA, Yurasov S, Soria JC (2015) Efficacy of rociletinib (CO-1686) in plasma-genotyped T790M-positive non-small cell lung cancer (NSCLC) patients (pts). Journal of Clinical Oncology 33: 8001–8001.

[108] Soria JC, Ohe Y, Vansteenkiste J, Reungwetwattana T, Chewaskulyong B, Lee KH, Dechaphunkul A, Imamura F, Nogami N, Kurata T, Okamoto I, Zhou C, Cho BC, Cheng Y, Cho EK, Voon PJ, Planchard D, Su WC, Gray JE, Lee SM, Hodge R, Marotti M, Rukazenkov Y, Ramalingam SS, FLAURA Investigators (2018) Osimertinib in Untreated ¡i¿EGFR¡/i¿ -Mutated Advanced NonSmall-Cell Lung Cancer. New England Journal of Medicine 378: 113–125.

[109] Ercan D, Choi HG, Yun CH, Capelletti M, Xie T, Eck MJ, Gray NS, Janne PA (2015) EGFR Mutations and Resistance to Irreversible Pyrimidine-Based EGFR Inhibitors. Clinical Cancer Research 21: 3913–3923.

[110] Thress KS, Paweletz CP, Felip E, Cho BC, Stetson D, Dougherty B, Lai Z, Markovets A, Vivancos A, Kuang Y, Ercan D, Matthews SE, Cantarini M, Barrett JC, Jänne PA, Oxnard GR (2015) Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. Nat Med 21: 560–562.

[111] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F (2015) Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. International Journal of Cancer 136: E359-E386.

[112] Walter V, Yin X, Wilkerson MD, Cabanski CR, Zhao N, Du Y, Ang MK, Hayward MC, Salazar AH, Hoadley KA, Fritchie K, Sailey CG, Weissler MC, Shockley WW, Zanation AM, Hackman T, Thorne LB, Funkhouser WD, Muldrew KL, Olshan AF, Randell SH, Wright FA, Shores CG, Hayes DN, Hayes DN (2013) Molecular Subtypes in Head and Neck Cancer Exhibit Distinct Patterns of Chromosomal Gain and Loss of Canonical Cancer Genes. PLoS ONE 8: e56823.

[113] Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, Westra WH, Chung CH, Jordan RC, Lu C, Kim H, Axelrod R, Silverman CC, Redmond KP, Gillison ML (2010) Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer. New England Journal of Medicine 363: 24–35.

[114] Chung CH, Parker JS, Karaca G, Wu J, Funkhouser WK, Moore D, Butterfoss D, Xiang D, Zanation A, Yin X, Shockley WW, Weissler MC, Dressler LG, Shores CG, Yarbrough WG, Perou CM (2004) Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. Cancer cell 5: 489–500.

[115] Keck MK, Zuo Z, Khattri A, Stricker TP, Brown CD, Imanguli M, Rieke D, Endhardt K, Fang P, Bragelmann J, DeBoer R, El-Dinali M, Aktolga S, Lei Z, Tan P, Rozen SG, Salgia R, Weichselbaum RR, Lingen MW, Story MD, Ang KK, Cohen EEW, White KP, Vokes EE, Seiwert TY (2015) Integrative Analysis of Head and Neck Cancer Identifies Two Biologically Distinct HPV and Three Non-HPV Subtypes. Clinical Cancer Research 21: 870–881.

[116] Leemans CR, Snijders PJF, Brakenhoff RH (2018) The molecular landscape of head and neck cancer. Nature Reviews Cancer 18: 269–282.

[117] Network TCGA (2015) Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517: 576–582.

[118] Maiti GP, Mondal P, Mukherjee N, Ghosh A, Ghosh S, Dey S, Chakrabarty J, Roy A, Biswas J, Roychoudhury S, Panda CK (2013) Overexpression of EGFR in head and neck squamous cell carcinoma is associated with inactivation of SH3GL2 and CDC25A genes. PLoS One 8: e63440.

[119] Ongkeko WM, Altuna X, Weisman RA, Wang-Rodriguez J (2005) Expression of Protein Tyrosine Kinases in Head and Neck Squamous Cell Carcinomas. American Journal of Clinical Pathology 124: 71–76.

[120] Bei R, Budillon A, Masuelli L, Cereda V, Vitolo D, Di Gennaro E, Ripavecchia V, Palumbo C, Ionna F, Losito S, Modesti A, Kraus MH, Muraro R (2004) Frequent overexpression of multiple ErbB receptors by head and neck squamous cell carcinoma contrasts with rare antibody immunity in patients. The Journal of Pathology 204: 317–325.

[121] Takes RP, Baatenburg de Jong RJ, Schuuring E, Litvinov SV, Hermans J, Van Krieken JH (1998) Differences in expression of oncogenes and tumor suppressor genes in different sites of head and neck squamous cell. Anticancer research 18: 4793–800.

[122] Perisanidis C (2017) Prevalence of EGFR tyrosine Kinase domain mutations in head and neck squamous cell carcinoma: Cohort study and systematic review. In Vivo 31: 23–34.

[123] Bonner JA, Harari PM, Giralt J, Azarnia N, Shin DM, Cohen RB, Jones CU, Sur R, Raben D, Jassem J, Ove R, Kies MS, Baselga J, Youssoufian H, Amellal N, Rowinsky EK, Ang KK (2006) Radiotherapy plus Cetuximab for Squamous-Cell Carcinoma of the Head and Neck. New England Journal of Medicine 354: 567–578.

[124] Vermorken JB, Trigo J, Hitt R, Koralewski P, Diaz-Rubio E, Rolland F, Knecht R, Amellal N, Schueler A, Baselga J (2007) Open-Label, Uncontrolled, Multicenter Phase II Study to Evaluate the Efficacy and Toxicity of Cetuximab As a Single Agent in Patients With Recurrent and/or Metastatic Squamous Cell Carcinoma of the Head and Neck Who Failed to Respond to Platinum-Based Therapy. J Clin Oncol 25: 2171–2177.

[125] Vermorken JB, Mesia R, Rivera F, Remenar E, Kawecki A, Rottey S, Erfan J, Zabolotnyy D, Kienzer HR, Cupissol D, Peyrade F, Benasso M, Vynnychenko I, De Raucourt D, Bokemeyer C, Schueler A, Amellal N, Hitt R (2008) Platinum-Based Chemotherapy plus Cetuximab in Head and Neck Cancer. New England Journal of Medicine 359: 1116–1127.

[126] Ang KK, Zhang Q, Rosenthal DI, Nguyen-Tan PF, Sherman EJ, Weber RS, Galvin JM, Bonner JA, Harris J, El-Naggar AK, Gillison ML, Jordan RC, Konski AA, Thorstad WL, Trotti A, Beitler JJ, Garden AS, Spanos WJ, Yom SS, Axelrod RS (2014) Randomized Phase III Trial of Concurrent Accelerated Radiation Plus Cisplatin With or Without Cetuximab for Stage III to IV Head and Neck Carcinoma: RTOG 0522. Journal of Clinical Oncology 32: 2940–2950.

[127] Hitt R, Irigoyen A, Cortes-Funes H, Grau JJ, García-Sáenz JA, Cruz-Hernandez JJ, Spanish Head and Neck Cancer Cooperative Group (TTCC) (2012) Phase II study of the combination of cetuximab and weekly paclitaxel in the first-line treatment of patients with recurrent and/or metastatic squamous cell carcinoma of head and neck. Annals of Oncology 23: 1016–1022.

[128] Wheeler S, Siwak DR, Chai R, LaValle C, Seethala RR, Wang L, Cieply K, Sherer C, Joy C, Mills GB, Argiris A, Siegfried JM, Grandis JR, Egloff AM (2012) Tumor Epidermal Growth Factor Receptor and EGFR PY1068 Are Independent Prognostic Indicators for Head and Neck Squamous Cell Carcinoma. Clinical Cancer Research 18: 2278–2289.

[129] Basavaraj C, Sierra P, Shivu J, Melarkode R, Monte E, Nair P (2010) Nimotuzumab with chemoradiation confers a survival advantage in treatment-naïve head and neck tumors over expressing EGFR. Cancer Biology & Therapy 10: 673–681.

[130] Pollock NI, Grandis JR (2015) HER2 as a Therapeutic Target in Head and Neck Squamous Cell Carcinoma. Clinical Cancer Research 21: 526–533.

[131] Wang D, Qian G, Zhang H, Magliocca KR, Nannapaneni S, Amin AR, Rossi M, Patel M, El-Deiry M, Wadsworth JT, Chen Z, Khuri FR, Shin DM, Saba NF, Chen ZG (2017) HER3 Targeting Sensitizes HNSCC to Cetuximab by Reducing HER3 Activity and HER2/HER3 Dimerization: Evidence from Cell Line and Patient-Derived Xenograft Models. Clinical Cancer Research 23: 677–686.

[132] Wheeler DL, Huang S, Kruser TJ, Nechrebecki MM, Armstrong EA, Benavente S, Gondi V, Hsu KT, Harari PM (2008) Mechanisms of acquired resistance to cetuximab: role of HER (ErbB) family members. Oncogene 27: 3944–3956.

[133] Wilson T, Lee D, Berry L, Shames D, Settleman J (2011) Neuregulin-1-Mediated Autocrine Signaling Underlies Sensitivity to HER2 Kinase Inhibitors in a Subset of Human Cancers. Cancer Cell 20: 158–172.

[134] Ather F, Hamidi H, Fejzo MS, Letrent S, Finn RS, Kabbinavar F, Head C, Wong SG (2013) Dacomitinib, an Irreversible Pan-ErbB Inhibitor Significantly Abrogates Growth in Head and Neck Cancer Models That Exhibit Low Response to Cetuximab. PLoS ONE 8: e56112.

[135] Kang CS, Zhang ZY, Jia ZF, Wang GX, Qiu MZ, Zhou HX, Yu SZ, Chang J, Jiang H, Pu PY (2006) Suppression of EGFR expression by antisense or small interference RNA inhibits U251 glioma cell growth in vitro and in vivo. Cancer Gene Therapy 13: 530–538.

[136] Heimberger AB, Sampson JH (2009) The PEPvIII-KLH (CDX-110) vaccine in glioblastoma multiforme patients. Expert Opinion on Biological Therapy 9: 1087–1098.

[137] Li H, Huang Y, Jiang DQ, Cui LZ, He Z, Wang C, Zhang ZW, Zhu HL, Ding YM, Li LF, Li Q, Jin HJ, Qian QJ (2018) Antitumor activity of EGFR-specific CAR T cells against non-small-cell lung cancer cells in vitro and in mice. Cell Death & Disease 9: 177.

[138] Cohen MH, Johnson JR, Chen YF, Sridhara R, Pazdur R (2005) FDA Drug Approval Summary: Erlotinib (Tarceva(R)) Tablets. The Oncologist 10: 461–466.

[139] Waddington CH (1953) Epigenetics and evolution. Symp Soc Exp Biol 7: 186–199.

[140] Riggs AD RV Martienssen RA (1996) Introduction. In: Epigenetic mechanisms of gene regulation, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. pp. 1–4.

[141] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, Fitzhugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, Levine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Hong ML, Dubois J, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Roe BA, Chen F, Pan H, Ramser

J, Lehrach H, Reinhardt R, McCombie WR, De La Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Patrinos A, Morgan MJ (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921.

[142] Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, Epstein CB, Frietze S, Harrow J, Kaul R, Khatun J, Lajoie BR, Landt SG, Lee BK, Pauli F, Rosenbloom KR, Sabo P, Safi A, Sanyal A, Shoresh N, Simon JM, Song L, Trinklein ND, Altshuler RC, Birney E, Brown JB, Cheng C, Djebali S, Dong X, Ernst J, Furey TS, Gerstein M, Giardine B, Greven M, Hardison RC, Harris RS, Herrero J, Hoffman MM, Iyer S, Kellis M, Kheradpour P, Lassmann T, Li Q, Lin X, Marinov GK, Merkel A, Mortazavi A, Parker SC, Reddy TE, Rozowsky J, Schlesinger F, Thurman RE, Wang J, Ward LD, Whitfield TW, Wilder SP, Wu W, Xi HS, Yip KY, Zhuang J, Bernstein BE, Green ED, Gunter C, Snyder M, Pazin MJ, Lowdon RF, Dillon LA, Adams LB, Kelly CJ, Zhang J, Wexler JR, Good PJ, Feingold EA, Crawford GE, Dekker J, Elnitski L, Farnham PJ, Giddings MC, Gingeras TR, Guigó R, Hubbard TJ, Kent WJ, Lieb JD, Margulies EH, Myers RM, Stamatoyannopoulos JA, Tenenbaum SA, Weng Z, White KP, Wold B, Yu Y, Wrobel J, Risk BA, Gunawardena HP, Kuiper HC, Maier CW, Xie L, Chen X, Mikkelsen TS, Gillespie S, Goren A, Ram O, Zhang X, Wang L, Issner R, Coyne MJ, Durham T, Ku M, Truong T, Eaton ML, Dobin A, Tanzer A, Lagarde J, Lin W, Xue C, Williams BA, Zaleski C, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Batut P, Bell I, Bell K, Chakrabortty S, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Li G, Luo OJ, Park E, Preall JB, Presaud K, Ribeca P, Robyr D, Ruan X, Sammeth M, Sandhu KS, Schaeffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Hayashizaki Y, Reymond A, Antonarakis SE, Hannon GJ, Ruan Y, Carninci P, Sloan CA, Learned K, Malladi VS, Wong MC, Barber GP, Cline MS, Dreszer TR, Heitner SG, Karolchik D, Kirkup VM, Meyer LR, Long JC, Maddren M, Raney BJ, Grasfeder LL, Giresi PG, Battenhouse A, Sheffield NC, Showers KA, London D, Bhinge AA, Shestak C, Schaner MR, Kim SK, Zhang ZZ, Mieczkowski PA, Mieczkowska JO, Liu Z, McDaniell RM, Ni Y, Rashid NU, Kim MJ, Adar S, Zhang Z, Wang T, Winter D, Keefe D, Iyer VR, Zheng M, Wang P, Gertz J, Vielmetter J, Partridge EC, Varley KE, Gasper C, Bansal A, Pepke S, Jain P, Amrhein H, Bowling KM, Anaya M, Cross MK, Muratet MA, Newberry KM, McCue K, Nesmith AS, Fisher-Aylor KI, Pusey B, DeSalvo G, Parker SL, Balasubramanian S, Davis NS, Meadows SK, Eggleston T, Newberry JS, Levy SE, Absher DM, Wong WH, Blow MJ, Visel A, Pennachio LA, Petrykowska HM,

Abyzov A, Aken B, Barrell D, Barson G, Berry A, Bignell A, Boychenko V, Bussotti G, Davidson C, Despacio-Reyes G, Diekhans M, Ezkurdia I, Frankish A, Gilbert J, Gonzalez JM, Griffiths E, Harte R, Hendrix DA, Hunt T, Jungreis I, Kay M, Khurana E, Leng J, Lin MF, Loveland J, Lu Z, Manthravadi D, Mariotti M, Mudge J, Mukherjee G, Notredame C, Pei B, Rodriguez JM, Saunders G, Sboner A, Searle S, Sisu C, Snow C, Steward C, Tapanari E, Tress ML, Van Baren MJ, Washietl S, Wilming L, Zadissa A, Zhang Z, Brent M, Haussler D, Valencia A, Addleman N, Alexander RP, Auerbach RK, Balasubramanian S, Bettinger K, Bhardwaj N, Boyle AP, Cao AR, Cayting P, Charos A, Cheng Y, Eastman C, Euskirchen G, Fleming JD, Grubert F, Habegger L, Hariharan M, Harmanci A, Iyengar S, Jin VX, Karczewski KJ, Kasowski M, Lacroute P, Lam H, Lamarre-Vincent N, Lian J, Lindahl-Allen M, Min R, Miotto B, Monahan H, Moqtaderi Z, Mu XJ, O'Geen H, Ouyang Z, Patacsil D, Raha D, Ramirez L, Reed B, Shi M, Slifer T, Witt H, Wu L, Xu X, Yan KK, Yang X, Struhl K, Weissman SM, Penalva LO, Karmakar S, Bhanvadia RR, Choudhury A, Domanus M, Ma L, Moran J, Victorsen A, Auer T, Centanin L, Eichenlaub M, Gruhl F, Heermann S, Hoeckendorf B, Inoue D, Kellner T, Kirchmaier S, Mueller C, Reinhardt R, Schertel L, Schneider S, Sinn R, Wittbrodt B, Wittbrodt J, Jain G, Balasundaram G, Bates DL, Byron R, Canfield TK, Diegel MJ, Dunn D, Ebersol AK, Frum T, Garg K, Gist E, Hansen RS, Boatman L, Haugen E, Humbert R, Johnson AK, Johnson EM, Kutyavin TV, Lee K, Lotakis D, Maurano MT, Neph SJ, Neri FV, Nguyen ED, Qu H, Reynolds AP, Roach V, Rynes E, Sanchez ME, Sandstrom RS, Shafer AO, Stergachis AB, Thomas S, Vernot B, Vierstra J, Vong S, Wang H, Weaver MA, Yan Y, Zhang M, Akey JM, Bender M, Dorschner MO, Groudine M, MacCoss MJ, Navas P, Stamatoyannopoulos G, Beal K, Brazma A, Flicek P, Johnson N, Lukk M, Luscombe NM, Sobral D, Vaquerizas JM, Batzoglou S, Sidow A, Hussami N, Kyriazopoulou-Panagiotopoulou S, Libbrecht MW, Schaub MA, Miller W, Bickel PJ, Banfai B, Boley NP, Huang H, Li JJ, Noble WS, Bilmes JA, Buske OJ, Sahu AD, Kharchenko PV, Park PJ, Baker D, Taylor J, Lochovsky L (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57–74.

[143] Khorasanizadeh S (2004) The Nucleosome: From Genomic Organization to Genomic Regulation. Cell 116: 259–272.

[144] Tan M, Luo H, Lee S, Jin F, Yang J, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B, Zhao Y (2011) Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification. Cell 146: 1016–1028.

[145] Dawson M, Kouzarides T (2012) Cancer Epigenetics: From Mechanism to Therapy. Cell 150: 12–27.

[146] Bannister AJ, Schneider R, Kouzarides T (2002) Histone methylation: dynamic or static? Cell 109: 801–6.

[147] Baylin SB, Jones PA (2011) A decade of exploring the cancer epigenome biological and translational implications. Nature Reviews Cancer 11: 726–734.

[148] Robertson KD (2005) DNA methylation and human disease. Nature Reviews Genetics 6: 597–610.

[149] Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69: 915–26.

[150] Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99: 247–57.

[151] Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nature Reviews Genetics 11: 204–220.

[152] Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science 324: 930–935.

[153] Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y (2011) Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science 333: 1300–1303.

[154] He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL (2011) Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. Science 333: 1303–1307.

[155] Moore LD, Le T, Fan G (2013) DNA Methylation and Its Basic Function. Neuropsychopharmacology 38: 23–38.

[156] Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T (2003) The Methyl-CpG-binding Protein MeCP2 Links DNA Methylation to Histone Methylation. Journal of Biological Chemistry 278: 4035–4040.

[157] Bird AP (1986) CpG-rich islands and the function of DNA methylation. Nature 321: 209–213.

[158] Antequera F, Bird A (1993) Number of CpG islands and genes in human and mouse. Proceedings of the National Academy of Sciences 90: 11995–11999.

[159] Hampsey M (1998) Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiology and molecular biology reviews : MMBR 62: 465–503.

[160] Bird A, Taggart M, Frommer M, Miller OJ, Macleod D (1985) A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell 40: 91–9.

[161] Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schübeler D (2008) Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. Molecular Cell 30: 755–766.

[162] Fei J, Torigoe SE, Brown CR, Khuong MT, Kassavetis GA, Boeger H, Kadonaga JT (2015) The prenucleosome, a stable conformational isomer of the nucleosome. Genes & development 29: 2563–75.

[163] Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC, Gilmour DS, Albert I, Pugh BF (2008) Nucleosome organization in the Drosophila genome. Nature 453: 358–362.

[164] Kharchenko PV, Tolstorukov MY, Park PJ (2008) Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nature Biotechnology 26: 1351–1359.

[165] Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-Resolution Profiling of Histone Methylations in the Human Genome. Cell 129: 823–837.

[166] Ng HH, Robert F, Young RA, Struhl K (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Molecular cell 11: 709–19.

[167] Tropberger P, Pott S, Keller C, Kamieniarz-Gdula K, Caron M, Richter F, Li G, Mittler G, Liu E, Bühler M, Margueron R, Schneider R (2013) Regulation of Transcription through Acetylation of H3K122 on the Lateral Surface of the Histone Octamer. Cell 152: 859–872.

[168] Dey A, Chitsaz F, Abbasi A, Misteli T, Ozato K (2003) The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. Proceedings of the National Academy of Sciences 100: 8758–8763.

[169] Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y (2002) Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. Science 298: 1039–1043.

[170] Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, ODonovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448: 553–560.

[171] Young MD, Willson TA, Wakefield MJ, Trounson E, Hilton DJ, Blewitt ME, Oshlack A, Majewski IJ (2011) ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. Nucleic acids research 39: 7415–27.

[172] Bulger M, Groudine M (2011) Functional and mechanistic diversity of distal transcription enhancers. Cell 144: 327–339.

[173] Banerji J, Rusconi S, Schaffner W (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell 27: 299–308.

[174] Calo E, Wysocka J (2013) Modification of enhancer chromatin: what, how, and why? Mol Cell 49: 825–837.

[175] Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A 107: 21931–21936.

[176] Zentner GE, Tesar PJ, Scacheri PC (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Research 21: 1273–1283.

[177] Bonn S, Zinzen RP, Girardot C, Gustafson EH, Perez-Gonzalez A, Delhomme N, Ghavi-Helm Y, Wilczyński B, Riddell A, Furlong EEM (2012) Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nature Genetics 44: 148–156.

[178] Pengelly AR, Copur O, Jackle H, Herzig A, Muller J (2013) A Histone Mutant Reproduces the Phenotype Caused by Loss of Histone-Modifying Factor Polycomb. Science 339: 698–699.

[179] Arnold CD, Gerlach D, Stelzer C, Boryn LM, Rath M, Stark A (2013) Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-seq. Science 339: 1074–1077.

[180] Diao Y, Fang R, Li B, Meng Z, Yu J, Qiu Y, Lin KC, Huang H, Liu T, Marina RJ, Jung I, Shen Y, Guan KL, Ren B (2017) A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. Nature Methods 14: 629–635.

[181] Ashe HL, Monks J, Wijgerde M, Fraser P, Proudfoot NJ (1997) Intergenic transcription and transinduction of the human beta -globinlocus. Genes & Development 11: 2494–2509.

[182] Tuan D, Kong S, Hu K (1992) Transcription of the hypersensitive site HS2 enhancer in erythroid cells. Proceedings of the National Academy of Sciences 89: 11219–11223.

[183] Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME (2010) Widespread transcription at neuronal activity-regulated enhancers. Nature 465: 182–187.

[184] De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G (2010) A Large Fraction of Extragenic RNA Pol II Transcription Sites Overlap Enhancers. PLoS Biology 8: e1000384.

[185] Wu H, Nord AS, Akiyama JA, Shoukry M, Afzal V, Rubin EM, Pennacchio LA, Visel A (2014) Tissue-Specific RNA Expression Marks Distant-Acting Developmental Enhancers. PLoS Genetics 10: e1004610.

[186] Mousavi K, Zare H, DellOrso S, Grontved L, Gutierrez-Cruz G, Derfoul A, Hager G, Sartorelli V (2013) eRNAs Promote Transcription by Establishing Chromatin Accessibility at Defined Genomic Loci. Molecular Cell 51: 606–617.

[187] Arner E, Daub CO, Vitting-Seerup K, Andersson R, Lilje B, Drabløs F, Lennartsson A, Rönnerblad M, Hrydziuszko O, Vitezic M, Freeman TC, M N Alhendi A, Arner P, Axton R, Baillie JK, Beckhouse A, Bodega B, Briggs J, Brombacher F, Davis M, Detmar M, Ehrlund A, Endoh M, Eslami A, Fagiolini M, Fairbairn L, Faulkner GJ, Ferrai C, Fisher ME, Forrester L, Goldowitz D, Guler R, Ha T, Hara M, Herlyn M, Ikawa T, Kai C, Kawamoto H, M Khachigian L, Klinken SP, Kojima S, Koseki H, Klein S, Mejhert N, Miyaguchi K, Mizuno Y, Morimoto M, Morris KJ, Mummery C, Nakachi Y, Ogishima S, Okada-Hatakeyama M, Okazaki Y, Orlando V, Ovchinnikov D, Passier R, Patrikakis M, Pombo A, Qin XY, Roy S, Sato H, Savvi S, Saxena A, Schwegmann A, Sugiyama D, Swoboda R, Tanaka H, Tomoiu A, Winteringham LN, Wolvetang E, Yanagi-Mizuochi C,

Yoneda M, Zabierowski S, Zhang P, Abugessaisa I, Bertin N, Diehl AD, Fukuda S, Furuno M, Harshbarger J, Hasegawa A, Hori F, Ishikawa-Kato S, Ishizu Y, Itoh M, Kawashima T, Kojima M, Kondo N, Lizio M, Meehan TF, Mungall CJ, Murata M, Nishiyori-Sueki H, Sahin S, Nagao-Sato S, Severin J, de Hoon MJL, Kawai J, Kasukawa T, Lassmann T, Suzuki H, Kawaji H, Summers KM, Wells C, Hume DA, Forrest ARR, Sandelin A, Carninci P, Hayashizaki Y, Hayashizaki Y (2015) Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. Science 347: 1010–1014.

[188] Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R (2013) Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature 494: 497–501.

[189] Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim HS, Glass CK, Rosenfeld MG (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498: 516–520.

[190] Schaukowitch K, Joo JY, Liu X, Watts J, Martinez C, Kim TK (2014) Enhancer RNA Facilitates NELF Release from Immediate Early Genes. Molecular Cell 56: 29–42.

[191] Sigova AA, Abraham BJ, Ji X, Molinie B, Hannett NM, Guo YE, Jangi M, Giallourakis CC, Sharp PA, Young RA (2015) Transcription factor trapping by RNA in gene regulatory elements. Science 350: 978–981.

[192] Deng W, Lee J, Wang H, Miller J, Reik A, Gregory P, Dean A, Blobel G (2012) Controlling Long-Range Genomic Interactions at a Native Locus by Targeted Tethering of a Looping Factor. Cell 149: 1233–1244.

[193] Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing Chromosome Conformation. Science 295: 1306–1311.

[194] Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science 326: 289–293.

[195] Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. Molecular cell 10: 1453–65.

[196] Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485: 376–380.

[197] Rao SS, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, Huang X, Shamim MS, Shin J, Turner D, Ye Z, Omer AD, Robinson JT, Schlick T, Bernstein BE, Casellas R, Lander ES, Aiden EL (2017) Cohesin Loss Eliminates All Loop Domains. Cell 171: 305–320.

[198] Tang Z, Luo O, Li X, Zheng M, Zhu J, Szalaj P, Trzaskoma P, Magalska A, Wlodarczyk J, Ruszczycki B, Michalski P, Piecuch E, Wang P, Wang D, Tian S, Penrad-Mobayed M, Sachs

L, Ruan X, Wei CL, Liu E, Wilczynski G, Plewczynski D, Li G, Ruan Y (2015) CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell 163: 1611–1627.

[199] Blackledge NP, Ott CJ, Gillen AE, Harris A (2009) An insulator element 3 to the CFTR gene binds CTCF and reveals an active chromatin hub in primary cells. Nucleic Acids Research 37: 1086–1094.

[200] Markova EN, Kantidze OL, Razin SV (2011) Transcriptional regulation and spatial organisation of the human AML1/RUNX1 gene. Journal of Cellular Biochemistry 112: 1997–2005.

[201] Javierre BM, Burren OS, Wilder SP, Kreuzhuber R, Hill SM, Sewitz S, Cairns J, Wingett SW, Várnai C, Thiecke MJ, Burden F, Farrow S, Cutler AJ, Rehnström K, Downes K, Grassi L, Kostadima M, Freire-Pritchett P, Wang F, Stunnenberg HG, Todd JA, Zerbino DR, Stegle O, Ouwehand WH, Frontini M, Wallace C, Spivakov M, Fraser P, Martens JH, Kim B, Sharifi N, Janssen-Megens EM, Yaspo ML, Linser M, Kovacsovics A, Clarke L, Richardson D, Datta A, Flicek P (2016) Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell 167: 1369–1384.

[202] Schoenfelder S, Furlan-Magaril M, Mifsud B, Tavares-Cadete F, Sugar R, Javierre BM, Nagano T, Katsman Y, Sakthidevi M, Wingett SW, Dimitrova E, Dimond A, Edelman LB, Elderkin S, Tabbada K, Darbo E, Andrews S, Herman B, Higgs A, LeProust E, Osborne CS, Mitchell JA, Luscombe NM, Fraser P (2015) The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. Genome Research 25: 582–597.

[203] Henikoff S (2008) Nucleosome destabilization in the epigenetic regulation of gene expression. Nature Reviews Genetics 9: 15–26.

[204] John S, Sabo PJ, Thurman RE, Sung MH, Biddie SC, Johnson TA, Hager GL, Stamatoyannopoulos JA (2011) Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nature Genetics 43: 264–268.

[205] Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10: 1213–1218.

[206] Pique-Regi R, Degner JF, Pai AA, Gaffney DJ, Gilad Y, Pritchard JK (2011) Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Research 21: 447–455.

[207] Shen H, Laird P (2013) Interplay between the Cancer Genome and Epigenome. Cell 153: 38–55.

[208] Suva ML, Riggi N, Bernstein BE (2013) Epigenetic Reprogramming in Cancer. Science 339: 1567–1570.

[209] Kim KH, Roberts CWM (2016) Targeting EZH2 in cancer. Nature Medicine 22: 128–134.

[210] Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, Paul JE, Boyle M, Woolcock BW, Kuchenbauer F, Yap D, Humphries RK, Griffith OL, Shah S, Zhu H, Kimbara M, Shashkin P, Charlot JF, Tcherpakov M, Corbett R, Tam A, Varhol R, Smailus D, Moksa M, Zhao Y, Delaney A, Qian H, Birol I, Schein J, Moore R, Holt R, Horsman DE, Connors JM, Jones S, Aparicio S, Hirst M, Gascoyne RD, Marra MA (2010) Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nature Genetics 42: 181–185.

[211] Ernst T, Chase AJ, Score J, Hidalgo-Curtis CE, Bryant C, Jones AV, Waghorn K, Zoi K, Ross FM, Reiter A, Hochhaus A, Drexler HG, Duncombe A, Cervantes F, Oscier D, Boultwood J, Grand FH, Cross NCP (2010) Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nature Genetics 42: 722–726.

[212] Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienta KJ, Sewalt RGAB, Otte AP, Rubin MA, Chinnaiyan AM (2002) The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature 419: 624–629.

[213] van Haaften G, Dalgliesh GL, Davies H, Chen L, Bignell G, Greenman C, Edkins S, Hardy C, O'Meara S, Teague J, Butler A, Hinton J, Latimer C, Andrews J, Barthorpe S, Beare D, Buck G, Campbell PJ, Cole J, Forbes S, Jia M, Jones D, Kok CY, Leroy C, Lin ML, McBride DJ, Maddison M, Maquire S, McLay K, Menzies A, Mironenko T, Mulderrig L, Mudie L, Pleasance E, Shepherd R, Smith R, Stebbings L, Stephens P, Tang G, Tarpey PS, Turner R, Turrell K, Varian J, West S, Widaa S, Wray P, Collins VP, Ichimura K, Law S, Wong J, Yuen ST, Leung SY, Tonon G, DePinho RA, Tai YT, Anderson KC, Kahnoski RJ, Massie A, Khoo SK, Teh BT, Stratton MR, Futreal PA (2009) Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. Nature Genetics 41: 521–523.

[214] Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, Ranish J, Crabtree GR (2013) Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. Nature Genetics 45: 592–601.

[215] Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA, Ghosh D, Sewalt RGAB, Otte AP, Hayes DF, Sabel MS, Livant D, Weiss SJ, Rubin MA, Chinnaiyan AM (2003) EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proceedings of the National Academy of Sciences 100: 11606–11611.

[216] Sneeringer CJ, Scott MP, Kuntz KW, Knutson SK, Pollock RM, Richon VM, Copeland RA (2010) Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. Proceedings of the National Academy of Sciences 107: 20980–20985.

[217] Biegel JA, Rorke LB, Packer RJ, Emanuel BS (1990) Monosomy 22 in rhabdoid or atypical tumors of the brain. Journal of Neurosurgery 73: 710–714.

[218] Kim KH, Kim W, Howard TP, Vazquez F, Tsherniak A, Wu JN, Wang W, Haswell JR, Walensky LD, Hahn WC, Orkin SH, Roberts CWM (2015) SWI/SNF-mutant cancers depend on catalytic and non-catalytic activity of EZH2. Nature Medicine 21: 1491–1496.

[219] Knutson SK, Wigle TJ, Warholic NM, Sneeringer CJ, Allain CJ, Klaus CR, Sacks JD, Raimondi A, Majer CR, Song J, Scott MP, Jin L, Smith JJ, Olhava EJ, Chesworth R, Moyer MP, Richon VM, Copeland RA, Keilhack H, Pollock RM, Kuntz KW (2012) A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. Nature Chemical Biology 8: 890–896.

[220] Knutson SK, Kawano S, Minoshima Y, Warholic NM, Huang KC, Xiao Y, Kadowaki T, Uesugi M, Kuznetsov G, Kumar N, Wigle TJ, Klaus CR, Allain CJ, Raimondi A, Waters NJ, Smith JJ, Porter-Scott M, Chesworth R, Moyer MP, Copeland RA, Richon VM, Uenaka T, Pollock RM, Kuntz KW, Yokoi A, Keilhack H (2014) Selective Inhibition of EZH2 by EPZ-6438 Leads to Potent Antitumor Activity in EZH2-Mutant Non-Hodgkin Lymphoma. Molecular Cancer Therapeutics 13: 842–854.

[221] Morschhauser F, Salles G, McKay P, Tilly H, Schmitt A, Gerecitano J, Johnson P, Le Gouill S, Dickinson M, Fruchart C, Lamy T, Chaidos A, Jurczak W, Opat S, Radford J, Zinzani P, Assouline S, Cartron G, Clawson A, Picazio N, Ribich S, Blakemore S, Larus J, Miao H, Ho P, Ribrag V (2017) Interim Report From a Phase 2 Multicenter Study of Tazemetostat, an Ezh2 Inhibitor, in Patients With Relapsed or Refractory B-Cell Non-Hodgkin Lymphomas. Hematological Oncology 35: 24–25.

[222] Italiano A, Soria JC, Toulmonde M, Michot JM, Lucchesi C, Varga A, Coindre JM, Blakemore SJ, Clawson A, Suttle B, McDonald AA, Woodruff M, Ribich S, Hedrick E, Keilhack H, Thomson B, Owa T, Copeland RA, Ho PTC, Ribrag V (2018) Tazemetostat, an EZH2 inhibitor, in relapsed or refractory B-cell non-Hodgkin lymphoma and advanced solid tumours: a first-in-human, open-label, phase 1 study. The Lancet Oncology 19: 649–659.

[223] Gibaja V, Shen F, Harari J, Korn J, Ruddy D, Saenz-Vash V, Zhai H, Rejtar T, Paris CG, Yu Z, Lira M, King D, Qi W, Keen N, Hassan AQ, Chan HM (2016) Development of secondary mutations in wild-type and mutant EZH2 alleles cooperates to confer resistance to EZH2 inhibitors. Oncogene 35: 558–566.

[224] Grossmann V, Haferlach C, Weissmann S, Roller A, Schindela S, Poetzinger F, Stadler K, Bellos F, Kern W, Haferlach T, Schnittger S, Kohlmann A (2013) The molecular profile of adult T-cell acute lymphoblastic leukemia: Mutations in RUNX1 and DNMT3A are associated with poor prognosis in T-ALL. Genes Chromosomes and Cancer 52: 410–422.

[225] Shah MY, Licht JD (2011) DNMT3A mutations in acute myeloid leukemia. Nature Genetics 43: 289–290.

[226] Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, Wen B, Wu H, Liu Y, Diep D, Briem E, Zhang K, Irizarry RA, Feinberg AP (2011) Increased methylation variation in epigenetic domains across cancer types. Nature Genetics 43: 768–775.

[227] Yang L, Rodriguez B, Mayle A, Park HJ, Lin X, Luo M, Jeong M, Curry CV, Kim SB, Ruau D, Zhang X, Zhou T, Zhou M, Rebel VI, Challen GA, Göttgens B, Lee JS, Rau R, Li W, Goodell MA (2016) DNMT3A Loss Drives Enhancer Hypomethylation in FLT3-ITD-Associated Leukemias. Cancer Cell 29: 922–934.

[228] Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E, Pikarski E, Young RA, Niveleau A, Cedar H, Simon I (2006) Evidence for an instructive mechanism of de novo methylation in cancer cells. Nature Genetics 38: 149–153.

[229] Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, Campan M, Noushmehr H, Bell CG, Maxwell AP, Savage DA, Mueller-Holzner E, Marth C, Kocjan G, Gayther SA, Jones A, Beck S, Wagner W, Laird PW, Jacobs IJ, Widschwendter M (2010) Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Research 20: 440–446.

[230] Hoeijmakers JH (2009) DNA Damage, Aging, and Cancer. New England Journal of Medicine 361: 1475–1485.

[231] Gao D, Herman JG, Guo M (2016) The clinical value of aberrant epigenetic changes of DNA damage repair genes in human cancer. Oncotarget 7: 37331–37346.

[232] Wang L, Zhu D, Zhang C, Mao X, Wang G, Mitra S, Li BF, Wang X, Wu M (1997) Mutations of O6-methylguanine-DNA methyltransferase gene in esophageal cancer tissues from Northern China. International journal of cancer 71: 719–23.

[233] Zhang YJ, Chen Y, Ahsan H, Lunn RM, Lee PH, Chen CJ, Santella RM (2003) Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation and its relationship to aflatoxin B1-DNA adducts andp53 mutation in hepatocellular carcinoma. International Journal of Cancer 103: 440–444.

[234] Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, Koelsche C, Sahm F, Chavez L, Reuss DE, Kratz A, Wefers AK, Huang K, Pajtler KW, Schweizer L, Stichel D, Olar A, Engel NW, Lindenberg K, Harter PN, Braczynski AK, Plate KH, Dohmen H, Garvalov BK, Coras R, Hölsken A, Hewer E, Bewerunge-Hudler M, Schick M, Fischer R, Beschorner R, Schittenhelm J, Staszewski O, Wani K, Varlet P, Pages M, Temming P, Lohmann D, Selt F, Witt H, Milde T, Witt O, Aronica E, Giangaspero F, Rushing E, Scheurlen W, Geisenberger C, Rodriguez FJ, Becker A, Preusser M, Haberler C, Bjerkvig R, Cryan J, Farrell M, Deckert M, Hench J, Frank S, Serrano J, Kannan K, Tsirigos A, Brück W, Hofer S, Brehmer S, Seiz-Rosenhagen M, Hänggi D, Hans V, Rozsnoki S, Hansford JR, Kohlhof P, Kristensen BW, Lechner M, Lopes B, Mawrin C, Ketter R, Kulozik A, Khatib Z, Heppner F, Koch A, Jouvet A, Keohane C, Mühleisen H, Mueller W, Pohl U, Prinz M, Benner A, Zapatka M, Gottardo NG, Driever PH, Kramm CM, Müller HL, Rutkowski S, von Hoff K, Frühwald MC, Gnekow A, Fleischhack G, Tippelt S, Calaminus G, Monoranu CM, Perry A, Jones C, Jacques TS, Radlwimmer B, Gessi M, Pietsch T, Schramm J, Schackert G, Westphal M, Reifenberger G, Wesseling P, Weller M, Collins VP, Blümcke I, Bendszus M, Debus J, Huang A, Jabado N, Northcott PA, Paulus W, Gajjar A, Robinson GW, Taylor MD, Jaunmuktane Z, Ryzhova M, Platten M, Unterberg A, Wick W, Karajannis MA, Mittelbronn M, Acker T, Hartmann C, Aldape K, Schüller U, Buslei R, Lichter P, Kool M, Herold-Mende C, Ellison DW, Hasselblatt M, Snuderl M, Brandner S, Korshunov A, von Deimling A, Pfister SM (2018) DNA methylation-based classification of central nervous system tumours. Nature 555: 469–474.

[235] Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, Kros JM, Hainfellner JA, Mason W, Mariani L, Bromberg JE, Hau P, Mirimanoff RO, Cairncross JG, Janzer RC,

Stupp R (2005) MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma . New England Journal of Medicine 352: 997–1003.

[236] Fernandes C, Costa A, Osório L, Lago RC, Linhares P, Carvalho B, Caeiro C (2017) Current Standards of Care in Glioblastoma Therapy. Codon Publications. doi:10.15586/CODON.GLIOBLASTOMA.2017.CH11. URL http://www.ncbi.nlm.nih.gov/pubmed/29251860.

[237] Jones PA, Taylor SM (1980) Cellular differentiation, cytidine analogs and DNA methylation. Cell 20: 85–93.

[238] Mummaneni P, Shord SS (2014) Epigenetics and Oncology. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy 34: 495–505.

[239] Abele R, Clavel M, Dodion P, Bruntsch U, Gundersen S, Smyth J, Renard J, van Glabbeke M, Pinedo HM (1987) The EORTC Early Clinical Trials Cooperative Group experience with 5-aza-2'-deoxycytidine (NSC 127716) in patients with colo-rectal, head and neck, renal carcinomas and malignant melanomas. European journal of cancer & clinical oncology 23: 1921–4.

[240] Zambrano P, Segura-Pacheco B, Perez-Cardenas E, Cetina L, Revilla-Vazquez A, Taja-Chayeb L, Chavez-Blanco A, Angeles E, Cabrera G, Sandoval K, Trejo-Becerril C, Chanona-Vilchis J, Duenas-González A (2005) A phase I study of hydralazine to demethylate and reactivate the expression of tumor suppressor genes. BMC Cancer 5: 44.

[241] Fandy TE, Herman JG, Kerns P, Jiemjit A, Sugar EA, Choi SH, Yang AS, Aucott T, Dauses T, Odchimar-Reissig R, Licht J, McConnell MJ, Nasrallah C, Kim MKH, Zhang W, Sun Y, Murgo A, Espinoza-Delgado I, Oteiza K, Owoeye I, Silverman LR, Gore SD, Carraway HE (2009) Early epigenetic changes and DNA damage do not predict clinical response in an overlapping schedule of 5-azacytidine and entinostat in patients with myeloid malignancies. Blood 114: 2764–2773.

[242] Cairns RA, Mak TW (2013) Oncogenic Isocitrate Dehydrogenase Mutations: Mechanisms, Models, and Clinical Opportunities. Cancer Discovery 3: 730–741.

[243] Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, Fantin VR, Jang HG, Jin S, Keenan MC, Marks KM, Prins RM, Ward PS, Yen KE, Liau LM, Rabinowitz JD, Cantley LC, Thompson CB, Vander Heiden MG, Su SM (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 462: 739–744.

[244] Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, Tallman MS, Sun Z, Wolniak K, Peeters JK, Liu W, Choe SE, Fantin VR, Paietta E, Löwenberg B, Licht JD, Godley LA, Delwel R, Valk PJM, Thompson CB, Levine RL, Melnick A (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer cell 18: 553–67.

[245] Ichimura K (2012) Molecular pathogenesis of IDH mutations in gliomas. Brain Tumor Pathology 29: 131–139.

[246] Parsons DW, Jones S, Zhang X, Lin JCH, Leary RJ, Angenendt P, Mankoo P, Carter H, Siu IM, Gallia GL, Olivi A, McLendon R, Rasheed BA, Keir S, Nikolskaya T, Nikolsky Y, Busam DA, Tekleab H, Diaz LA, Hartigan J, Smith DR, Strausberg RL, Marie SKN, Shinjo SMO, Yan H, Riggins GJ, Bigner DD, Karchin R, Papadopoulos N, Parmigiani G, Vogelstein B, Velculescu VE, Kinzler KW (2008) An Integrated Genomic Analysis of Human Glioblastoma Multiforme. Science 321: 1807–1812.

[247] Rohle D, Popovici-Muller J, Palaskas N, Turcan S, Grommes C, Campos C, Tsoi J, Clark O, Oldrini B, Komisopoulou E, Kunii K, Pedraza A, Schalm S, Silverman L, Miller A, Wang F, Yang H, Chen Y, Kernytsky A, Rosenblum MK, Liu W, Biller SA, Su SM, Brennan CW, Chan TA, Graeber TG, Yen KE, Mellinghoff IK (2013) An Inhibitor of Mutant IDH1 Delays Growth and Promotes Differentiation of Glioma Cells. Science 340: 626–630.

[248] Abou Dalle I, DiNardo CD (2018) The role of enasidenib in the treatment of mutant IDH2 acute myeloid leukemia. Therapeutic advances in hematology 9: 163–173.

[249] Stein EM, DiNardo CD, Fathi AT, Mims AS, Pratz KW, Savona MR, Stein AS, Stone RM, Winer ES, Seet CS, Döhner H, Pollyea DA, McCloskey JK, Odenike O, Lowenberg B, Ossenkoppele GJ, Patel PA, Roshal M, Lersch F, Nabhan S, Choe S, Wang H, Hua L, Almon C, Cooper M, Tallman MS (2018) Ivosidenib or Enasidenib Combined with Induction and Consolidation Chemotherapy in Patients with Newly Diagnosed AML with an IDH1 or IDH2 Mutation Is Safe, Effective, and Leads to MRD-Negative Complete Remissions. Blood 132: 560 LP - 560.

[250] Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutyavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R, Stamatoyannopoulos JA (2012) Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. Science 337: 1190–1195.

[251] MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A, Milano A, Morales J, Pendlington ZM, Welter D, Burdett T, Hindorff L, Flicek P, Cunningham F, Parkinson H (2017) The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Research 45: D896-D901.

[252] Ding Z, Ni Y, Timmer SW, Lee BK, Battenhouse A, Louzada S, Yang F, Dunham I, Crawford GE, Lieb JD, Durbin R, Iyer VR, Birney E (2014) Quantitative Genetics of CTCF Binding Reveal Local Sequence Effects and Different Modes of X-Chromosome Association. PLoS Genetics 10: e1004798.

[253] Farh KKH, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shoresh N, Whitton H, Ryan RJH, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE (2015) Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 518: 337–343.

[254] Paraboschi EM, Rimoldi V, Solda G, Tabaglio T, Dall'Osso C, Saba E, Vigliano M, Salviati A, Leone M, Benedetti MD, Fornasari D, Saarela J, De Jager PL, Patsopoulos

NA, D'Alfonso S, Gemmati D, Duga S, Asselta R (2014) Functional variations modulating PRKCA expression and alternative splicing predispose to multiple sclerosis. Human Molecular Genetics 23: 6746–6761.

[255] Zhang L, Liu Y, Song F, Zheng H, Hu L, Lu H, Liu P, Hao X, Zhang W, Chen K (2011) Functional SNP in the microRNA-367 binding site in the 3'UTR of the calcium channel ryanodine receptor gene 3 (RYR3) affects breast cancer risk and calcification. Proceedings of the National Academy of Sciences 108: 13653–13658.

[256] Kaplow IM, MacIsaac JL, Mah SM, McEwen LM, Kobor MS, Fraser HB (2015) A pooling-based approach to mapping genetic variants associated with DNA methylation. Genome Research 25: 907–917.

[257] Degner JF, Pai AA, Pique-Regi R, Veyrieras JB, Gaffney DJ, Pickrell JK, De Leon S, Michelini K, Lewellen N, Crawford GE, Stephens M, Gilad Y, Pritchard JK (2012) DNaseI sensitivity QTLs are a major determinant of human expression variation. Nature 482: 390–394.

[258] Dorsett D, Krantz ID (2009) On the Molecular Etiology of Cornelia de Lange Syndrome. Annals of the New York Academy of Sciences 1151: 22–37.

[259] Miyake N, Koshimizu E, Okamoto N, Mizuno S, Ogata T, Nagai T, Kosho T, Ohashi H, Kato M, Sasaki G, Mabe H, Watanabe Y, Yoshino M, Matsuishi T, Takanashi JI, Shotelersuk V, Tekin M, Ochi N, Kubota M, Ito N, Ihara K, Hara T, Tonoki H, Ohta T, Saito K, Matsuo M, Urano M, Enokizono T, Sato A, Tanaka H, Ogawa A, Fujita T, Hiraki Y, Kitanaka S, Matsubara Y, Makita T, Taguri M, Nakashima M, Tsurusaki Y, Saitsu H, Yoshiura KI, Matsumoto N, Niikawa N (2013) MLL2 and KDM6A mutations in patients with Kabuki syndrome. American Journal of Medical Genetics, Part A 161: 2234–2243.

[260] Smith E, Lin C, Shilatifard A (2011) The super elongation complex (SEC) and MLL in development and disease. Genes & Development 25: 661–672.

[261] Roelfsema JH, White SJ, Ariyürek Y, Bartholdi D, Niedrist D, Papadia F, Bacino CA, den Dunnen JT, van Ommen GJB, Breuning MH, Hennekam RC, Peters DJ (2005) Genetic Heterogeneity in Rubinstein-Taybi Syndrome: Mutations in Both the CBP and EP300 Genes Cause Disease. The American Journal of Human Genetics 76: 572–580.

[262] Lin C, Lovén J, Rahl P, Paranal R, Burge C, Bradner J, Lee T, Young R (2012) Transcriptional Amplification in Tumor Cells with Elevated c-Myc. Cell 151: 56–67.

[263] Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, Wang R, Green D, Tessarollo L, Casellas R, Zhao K, Levens D (2012) c-Myc Is a Universal Amplifier of Expressed Genes in Lymphocytes and Embryonic Stem Cells. Cell 151: 68–79.

[264] Tavana O, Li D, Dai C, Lopez G, Banerjee D, Kon N, Chen C, Califano A, Yamashiro DJ, Sun H, Gu W (2016) HAUSP deubiquitinates and stabilizes N-Myc in neuroblastoma. Nature Medicine 22: 1180–1186.

[265] Delmore J, Issa G, Lemieux M, Rahl P, Shi J, Jacobs H, Kastritis E, Gilpatrick T, Paranal R, Qi J, Chesi M, Schinzel A, McKeown M, Heffernan T, Vakoc C, Bergsagel P, Ghobrial

I, Richardson P, Young R, Hahn W, Anderson K, Kung A, Bradner J, Mitsiades C (2011) BET Bromodomain Inhibition asa Therapeutic Strategy to Target c-Myc. Cell 146: 904–917.

[266] Berg T, Cohen SB, Desharnais J, Sonderegger C, Maslyar DJ, Goldberg J, Boger DL, Vogt PK (2002) Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts. Proceedings of the National Academy of Sciences 99: 3830–3835.

[267] Whyte Wa, Orlando Da, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young Ra (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153: 307–19.

[268] Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, Hoke HA, Young RA (2013) Super-enhancers in the control of cell identity and disease. Cell 155: 934–947.

[269] Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, Young RA (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153: 320–334.

[270] Hay D, Hughes JR, Babbs C, Davies JOJ, Graham BJ, Hanssen LLP, Kassouf MT, Oudelaar AM, Sharpe JA, Suciu MC, Telenius J, Williams R, Rode C, Li PS, Pennacchio LA, Sloane-Stanley JA, Ayyub H, Butler S, Sauka-Spengler T, Gibbons RJ, Smith AJH, Wood WG, Higgs DR (2016) Genetic dissection of the α-globin super-enhancer in vivo. Nature Genetics 48: 895–903.

[271] Osterwalder M, Barozzi I, Tissières V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, Lee EA, Zhu Y, Plajzer-Frick I, Pickle CS, Kato M, Garvin TH, Pham QT, Harrington AN, Akiyama JA, Afzal V, Lopez-Rios J, Dickel DE, Visel A, Pennacchio LA (2018) Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 554: 239–243.

[272] Shin HY, Willi M, Yoo KH, Zeng X, Wang C, Metser G, Hennighausen L (2016) Hierarchy within the mammary STAT5-driven Wap super-enhancer. Nature Genetics 48: 904–911.

[273] Xie S, Duan J, Li B, Zhou P, Hon GC (2017) Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. Molecular Cell 66: 285–299.

[274] Chapuy B, McKeown M, Lin C, Monti S, Roemer M, Qi J, Rahl P, Sun H, Yeda K, Doench J, Reichert E, Kung A, Rodig S, Young R, Shipp M, Bradner J (2013) Discovery and Characterization of Super-Enhancer-Associated Dependencies in Diffuse Large B Cell Lymphoma. Cancer Cell 24: 777–790.

[275] Lin CY, Erkek S, Tong Y, Yin L, Federation AJ, Zapatka M, Haldipur P, Kawauchi D, Risch T, Warnatz HJ, Worst BC, Ju B, Orr BA, Zeid R, Polaski DR, Segura-Wang M, Waszak SM, Jones DTW, Kool M, Hovestadt V, Buchhalter I, Sieber L, Johann P, Chavez L, Gröschel S, Ryzhova M, Korshunov A, Chen W, Chizhikov VV, Millen KJ, Amstislavskiy V, Lehrach H, Yaspo ML, Eils R, Lichter P, Korbel JO, Pfister SM, Bradner JE, Northcott PA (2016) Active medulloblastoma enhancers reveal subgroup-specific cellular origins. Nature 530: 57–62.

[276] Jiang YY, Lin DC, Mayakonda A, Hazawa M, Ding LW, Chien WW, Xu L, Chen Y, Xiao JF, Senapedis W, Baloglu E, Kanojia D, Shang L, Xu X, Yang H, Tyner JW, Wang MR, Koeffler HP (2017) Targeting super-enhancer-associated oncogenes in oesophageal squamous cell carcinoma. Gut 66: 1358–1368.

[277] Zhou B, Wang L, Zhang S, Bennett BD, He F, Zhang Y, Xiong C, Han L, Diao L, Li P, Fargo DC, Cox AD, Hu G (2016) INO80 governs superenhancer-mediated oncogenic transcription and tumor growth in melanoma. Genes & Development 30: 1440–1453.

[278] Chipumuro E, Marco E, Christensen C, Kwiatkowski N, Zhang T, Hatheway C, Abraham B, Sharma B, Yeung C, Altabef A, Perez-Atayde A, Wong KK, Yuan GC, Gray N, Young R, George R (2014) CDK7 Inhibition Suppresses Super-Enhancer-Linked Oncogenic Transcription in MYCN-Driven Cancer. Cell 159: 1126–1139.

[279] Cejas P, Li L, O'Neill NK, Duarte M, Rao P, Bowden M, Zhou CW, Mendiola M, Burgos E, Feliu J, Moreno-Rubio J, Guadalajara H, Moreno V, García-Olmo D, Bellmunt J, Mullane S, Hirsch M, Sweeney CJ, Richardson A, Liu XS, Brown M, Shivdasani RA, Long HW (2016) Chromatin immunoprecipitation from fixed clinical tissues reveals tumor-specific enhancer profiles. Nature Medicine 22: 685–691.

[280] Hudson LG, Thompson KL, Xu J, Gill GN (1990) Identification and characterization of a regulated promoter element in the epidermal growth factor receptor gene. Proceedings of the National Academy of Sciences 87: 7536–7540.

[281] Ishii S, Xu YH, Stratton RH, Roe BA, Merlino GT, Pastan I (1985) Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. Proceedings of the National Academy of Sciences of the United States of America 82: 4920–4.

[282] Brandt B, Meyer-Staeckling S, Schmidt H, Agelopoulos K, Buerger H (2006) Mechanisms of egfr gene transcription modulation: relationship to cancer risk and therapy response. Clin Cancer Res 12: 7252–7260.

[283] Gebhardt F, Zänker KS, Brandt B (1999) Modulation of Epidermal Growth Factor Receptor Gene Transcription by a Polymorphic Dinucleotide Repeat in Intron 1. Journal of Biological Chemistry 274: 13176–13180.

[284] Chi D, Hing A, Helms C, Steinbrueck T, Mishra SK, Donis-Keller H (1992) Two chromosome 7 dinucleotide repeat polymorphisms at gene loci epidermal growth factor receptor (EGFR) and proα2 (1) collagen (COL1A2). Human Molecular Genetics 1: 135–135.

[285] Chrysogelos SA (1993) Chromatin structure of the EGFR gene suggests a role for intron 1 sequences in its regulation in breast cancer cells. Nucleic Acids Res 21: 5736–5741.

[286] Maekawa T, Imamoto F, Merlino GT, Pastan I, Ishii S (1989) Cooperative function of two separate enhancers of the human epidermal growth factor receptor proto-oncogene. Journal of Biological Chemistry 264: 5488–5494.

[287] Chang H, Liu Y, Xue M, Liu H, Du S, Zhang L, Wang P (2016) Synergistic action of master transcription factors controls epithelial-to-mesenchymal transition. Nucleic Acids Research 44: 2514–2527.

[288] Loo JX, Zhao M, Xia S, Li F, Ke H, He C, Shi X, Jiang Y, Zhong S, Chen X, Yang M, Fu L, Xiong W, Guo Y (2018) E6 Protein Expressed by High-Risk HPV Activates Super-Enhancers of the EGFR and c-MET Oncogenes by Destabilizing the Histone Demethylase KDM5C . Cancer Research 78: 1418–1430.

[289] Mack SC, Singh I, Wang X, Hirsch R, Wu Q, Villagomez R, Bernatchez JA, Zhu Z, Gimple RC, Kim LJ, Morton A, Lai S, Qiu Z, Prager BC, Bertrand KC, Mah C, Zhou W, Lee C, Barnett GH, Vogelbaum MA, Sloan AE, Chavez L, Bao S, Scacheri PC, Siqueira-Neto JL, Lin CY, Rich JN (2019) Chromatin landscapes reveal developmentally encoded transcriptional states that define human glioblastoma. The Journal of Experimental Medicine 216: 1071–1090.

[290] Gimple RC, Kidwell RL, Kim LJY, Sun T, Gromovsky AD, Wu Q, Wolf M, Lv D, Bhargava S, Jiang L, Prager BC, Wang X, Ye Q, Zhu Z, Zhang G, Dong Z, Zhao L, Lee D, Bi J, Sloan AE, Mischel PS, Brown JM, Cang H, Huan T, Mack SC, Xie Q, Rich JN (2019) Glioma Stem Cell Specific Super Enhancer Promotes Polyunsaturated Fatty Acid Synthesis to Support EGFR Signaling. Cancer Discovery : 19–0061.

[291] Chu T, Rice EJ, Booth GT, Salamanca HH, Wang Z, Core LJ, Longo SL, Corona RJ, Chin LS, Lis JT, Kwak H, Danko CG (2018) Chromatin run-on and sequencing maps the transcriptional regulatory landscape of glioblastoma multiforme. Nature Genetics 50: 1553–1564.

[292] Moreau P, Hen R, Wasylyk B, Everett R, Gaub M, Chambon P (1981) The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Research 9: 6047–6068.

[293] Korkmaz G, Lopes R, Ugalde AP, Nevedomskaya E, Han R, Myacheva K, Zwart W, Elkon R, Agami R (2016) Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. Nature Biotechnology 34: 192–198.

[294] Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K, Dickson BJ, Stark A (2014) Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature 512: 91–95.

[295] Zhao Z, Tavoosidana G, Sjölinder M, Göndör A, Mariano P, Wang S, Kanduri C, Lezcano M, Sandhu KS, Singh U, Pant V, Tiwari V, Kurukuti S, Ohlsson R (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nature Genetics 38: 1341–1347.

[296] Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J (2012) HiC: A comprehensive technique to capture the conformation of genomes. Methods 58: 268–276.

[297] Hagège H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, de Laat W, Forné T (2007) Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nature Protocols 2: 1722–1733.

[298] Kaikkonen M, Spann N, Heinz S, Romanoski C, Allison K, Stender J, Chun H, Tough D, Prinjha R, Benner C, Glass C (2013) Remodeling of the Enhancer Landscape during Macrophage Activation Is Coupled to Enhancer Transcription. Molecular Cell 51: 310–325.

[299] Zhang Y, Wong CH, Birnbaum RY, Li G, Favaro R, Ngan CY, Lim J, Tai E, Poh HM, Wong E, Mulawadi FH, Sung WK, Nicolis S, Ahituv N, Ruan Y, Wei CL (2013) Chromatin connectivity maps reveal dynamic promoterenhancer long-range associations. Nature 504: 306–310.

[300] Botta M, Haider S, Leung IXY, Lio P, Mozziconacci J (2010) Intra- and inter-chromosomal interactions correlate with CTCF binding genome wide. Molecular Systems Biology 6: 426.

[301] Beaulieu E, Green L, Elsby L, Alourfi Z, Morand EF, Ray DW, Donn R (2011) Identification of a novel cell type-specific intronic enhancer of macrophage migration inhibitory factor (MIF) and its regulation by mithramycin. Clinical & Experimental Immunology 163: 178–188.

[302] Tourmente S, Chapel S, Dreau D, Drake ME, Bruhat A, Couderc JL, Dastugue B (1993) Enhancer and silencer elements within the first intron mediate the transcriptional regulation of the beta 3 tubulin gene by 20-hydroxyecdysone in Drosophila Kc cells. Insect biochemistry and molecular biology 23: 137–43.

[303] Bianchi M, Crinelli R, Giacomini E, Carloni E, Magnani M (2009) A potent enhancer element in the 5-UTR intron is crucial for transcriptional regulation of the human ubiquitin C gene. Gene 448: 88–101.

[304] Bornstein P, McKay J, Liska DJ, Apone S, Devarayalu S (1988) Interactions between the promoter and first intron are involved in transcriptional control of alpha 1(I) collagen gene expression. Molecular and cellular biology 8: 4851–7.

[305] Chan RYY, Boudreau-Lariviere C, Angus LM, Mankal FA, Jasmin BJ (1999) An intronic enhancer containing an N-box motif is required for synapse- and tissue-specific expression of the acetylcholinesterase gene in skeletal muscle fibers. Proceedings of the National Academy of Sciences 96: 4627–4632.

[306] Gaunitz F, Heise K, Schumann R, Gebhardt R (2002) Glucocorticoid induced expression of glutamine synthetase in hepatoma cells. Biochemical and biophysical research communications 296: 1026–32.

[307] Scohy S, Gabant P, Szpirer C, Szpirer J (2000) Identification of an enhancer and an alternative promoter in the first intron of the alpha-fetoprotein gene. Nucleic Acids Research 28: 3743–3751.

[308] Sleckman BP, Gorman JR, Alt FW (2002) ACCESSIBILITY CONTROL OF ANTIGEN-RECEPTOR VARIABLE-REGION GENE ASSEMBLY: Role of cis -Acting Elements . Annual Review of Immunology 14: 459–481.

[309] Bradnam KR, Korf I (2008) Longer First Introns Are a General Property of Eukaryotic Gene Structure. PLoS ONE 3: e3093.

[310] Chamary JV, Hurst LD (2004) Similar Rates but Different Modes of Sequence Evolution in Introns and at Exonic Silent Sites in Rodents: Evidence for Selectively Driven Codon Usage. Molecular Biology and Evolution 21: 1014–1023.

[311] Majewski J (2002) Distribution and Characterization of Regulatory Elements in the Human Genome. Genome Research 12: 1827–1836.

[312] Vinogradov AE (2006) "Genome design" model: Evidence from conserved intronic sequence in human-mouse comparison. Genome Research 16: 347–354.

[313] Cenik C, Derti A, Mellor JC, Berriz GF, Roth FP (2010) Genome-wide functional analysis of human 5' untranslated region introns. Genome Biology 11: R29.

[314] Pesole G, Mignone F, Gissi C, Grillo G, Licciulli F, Liuni S (2001) Structural and functional features of eukaryotic mRNA untranslated regions. Gene 276: 73–81.

[315] Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153: 307–319.

[316] Hnisz D, Schuijers J, Lin CY, Weintraub AS, Abraham BJ, Lee TI, Bradner JE, Young RA (2015) Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. Mol Cell 58: 362–370.

[317] Li Z, Van Calcar S, Qu C, Cavenee WK, Zhang MQ, Ren B (2003) A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. Proceedings of the National Academy of Sciences of the United States of America 100: 8164–9.

[318] Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357–359.

[319] Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38: 576–589.

[320] Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 17: 3.

[321] Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). Genome Biol 9: R137.

[322] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. Science (New York, NY) 315: 1709–12.

[323] Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin EV, van der Oost J (2008) Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. Science 321: 960–964.

[324] Garneau JE, Dupuis M, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468: 67–71.

[325] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337: 816–821.

[326] Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339: 819–823.

[327] Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-Guided Human Genome Engineering via Cas9. Science 339: 823–826.

[328] Iliakis G, Wang H, Perrault A, Boecker W, Rosidi B, Windhofer F, Wu W, Guan J, Terzoudi G, Pantelias G (2004) Mechanisms of DNA double strand break repair and chromosome aberration formation. Cytogenetic and Genome Research 104: 14–20.

[329] Lin S, Staahl BT, Alla RK, Doudna JA (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. eLife 3.

[330] Barrangou R, May AP (2015) Unraveling the potential of CRISPR-Cas9 for gene therapy. Expert Opinion on Biological Therapy 15: 311–314.

[331] Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, Koteliansky V, Sharp PA, Jacks T, Anderson DG (2014) Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nature Biotechnology 32: 551–553.

[332] Kimberland ML, Hou W, Alfonso-Pecchio A, Wilson S, Rao Y, Zhang S, Lu Q (2018) Strategies for controlling CRISPR/Cas9 off-target effects and biological variations in mammalian genome editing experiments. Journal of Biotechnology 284: 91–101.

[333] Schaefer KA, Wu WH, Colgan DF, Tsang SH, Bassuk AG, Mahajan VB (2017) Unexpected mutations after CRISPRCas9 editing in vivo. Nature Methods 14: 547–548.

[334] Ye L, Wang J, Beyer AI, Teque F, Cradick TJ, Qi Z, Chang JC, Bao G, Muench MO, Yu J, Levy JA, Kan YW (2014) Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Δ32 mutation confers resistance to HIV infection. Proceedings of the National Academy of Sciences 111: 9591–9596.

[335] Cyranoski D (2019) The CRISPR-baby scandal: whats next for human gene-editing. Nature 566: 440–442.

[336] Dickenson D, Darnovsky M (2019) Did a permissive scientific culture encourage the CRISPR babies experiment? Nature Biotechnology 37: 355–357.

[337] Bauer DE, Canver MC, Orkin SH (2014) Generation of Genomic Deletions in Mammalian Cell Lines via CRISPR/Cas9. Journal of Visualized Experiments : e52118.

[338] Canver MC, Bauer DE, Dass A, Yien YY, Chung J, Masuda T, Maeda T, Paw BH, Orkin SH (2014) Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells. Journal of Biological Chemistry 289: 21312–21324.

[339] Qi L, Larson M, Gilbert L, Doudna J, Weissman J, Arkin A, Lim W (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell 152: 1173–1183.

[340] Beerli RR, Dreier B, Barbas CF (2000) Positive and negative regulation of endogenous genes by designed transcription factors. Proceedings of the National Academy of Sciences 97: 1495–1500.

[341] Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154: 442–451.

[342] Gilbert L, Horlbeck M, Adamson B, Villalta J, Chen Y, Whitehead E, Guimaraes C, Panning B, Ploegh H, Bassik M, Qi L, Kampmann M, Weissman J (2014) Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell 159: 647–661.

[343] Li Y, Rivera CM, Ishii H, Jin F, Selvaraj S, Lee AY, Dixon JR, Ren B (2014) CRISPR Reveals a Distal Super-Enhancer Required for Sox2 Expression in Mouse Embryonic Stem Cells. PLoS ONE 9: e114485.

[344] Zhou W, Fong MY, Min Y, Somlo G, Liu L, Palomares MR, Yu Y, Chow A, O'Connor STF, Chin AR, Yen Y, Wang Y, Marcusson EG, Chu P, Wu J, Wu X, Li AX, Li Z, Gao H, Ren X, Boldin MP, Lin PC, Wang SE (2014) Cancer-Secreted miR-105 Destroys Vascular Endothelial Barriers to Promote Metastasis. Cancer cell 25: 501–15.

[345] Bahr C, von Paleske L, Uslu VV, Remeseiro S, Takayama N, Ng SW, Murison A, Langenfeld K, Petretich M, Scognamiglio R, Zeisberger P, Benk AS, Amit I, Zandstra PW, Lupien M, Dick JE, Trumpp A, Spitz F (2018) A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. Nature 553: 515–520.

[346] Liu Q, Kulak MV, Borcherding N, Maina PK, Zhang W, Weigel RJ, Qi HH (2018) A novel HER2 gene body enhancer contributes to HER2 expression. Oncogene 37: 687–694.

[347] Chen H, Li C, Peng X, Zhou Z, Weinstein JN, Caesar-Johnson SJ, Demchok JA, Felau I, Kasapi M, Ferguson ML, Hutter CM, Sofia HJ, Tarnuzzer R, Wang Z, Yang L, Zenklusen JC, Zhang JJ, Chudamani S, Liu J, Lolla L, Naresh R, Pihl T, Sun Q, Wan Y, Wu Y, Cho J, DeFreitas T, Frazer S, Gehlenborg N, Getz G, Heiman DI, Kim J, Lawrence MS, Lin P, Meier S, Noble MS, Saksena G, Voet D, Zhang H, Bernard B, Chambwe N, Dhankani V, Knijnenburg T, Kramer R, Leinonen K, Liu Y, Miller M, Reynolds S, Shmulevich I, Thorsson V, Zhang W, Akbani R, Broom BM, Hegde AM, Ju Z, Kanchi RS, Korkut A, Li J, Liang H, Ling S, Liu W, Lu Y, Mills GB, Ng KS, Rao A, Ryan M, Wang J, Weinstein JN, Zhang J, Abeshouse A, Armenia J, Chakravarty D, Chatila WK, de Bruijn I, Gao J, Gross BE, Heins ZJ, Kundra R, La K, Ladanyi M, Luna A, Nissan MG, Ochoa A, Phillips SM, Reznik E, Sanchez-Vega F, Sander C, Schultz N, Sheridan R, Sumer SO, Sun Y, Taylor BS, Wang J, Zhang H, Anur P, Peto M, Spellman P, Benz C, Stuart JM, Wong CK, Yau C, Hayes DN, Parker JS, Wilkerson MD, Ally A, Balasundaram M, Bowlby R, Brooks D, Carlsen R, Chuah E, Dhalla N, Holt R, Jones SJ, Kasaian K, Lee D, Ma Y, Marra MA, Mayo M, Moore RA, Mungall AJ, Mungall K, Robertson AG, Sadeghi S, Schein JE, Sipahimalani

P, Tam A, Thiessen N, Tse K, Wong T, Berger AC, Beroukhim R, Cherniack AD, Cibulskis C, Gabriel SB, Gao GF, Ha G, Meyerson M, Schumacher SE, Shih J, Kucherlapati MH, Kucherlapati RS, Baylin S, Cope L, Danilova L, Bootwalla MS, Lai PH, Maglinte DT, Van Den Berg DJ, Weisenberger DJ, Auman JT, Balu S, Bodenheimer T, Fan C, Hoadley KA, Hoyle AP, Jefferys SR, Jones CD, Meng S, Mieczkowski PA, Mose LE, Perou AH, Perou CM, Roach J, Shi Y, Simons JV, Skelly T, Soloway MG, Tan D, Veluvolu U, Fan H, Hinoue T, Laird PW, Shen H, Zhou W, Bellair M, Chang K, Covington K, Creighton CJ, Dinh H, Doddapaneni H, Donehower LA, Drummond J, Gibbs RA, Glenn R, Hale W, Han Y, Hu J, Korchina V, Lee S, Lewis L, Li W, Liu X, Morgan M, Morton D, Muzny D, Santibanez J, Sheth M, Shinbrot E, Wang L, Wang M, Wheeler DA, Xi L, Zhao F, Hess J, Appelbaum EL, Bailey M, Cordes MG, Ding L, Fronick CC, Fulton LA, Fulton RS, Kandoth C, Mardis ER, McLellan MD, Miller CA, Schmidt HK, Wilson RK, Crain D, Curley E, Gardner J, Lau K, Mallery D, Morris S, Paulauskis J, Penny R, Shelton C, Shelton T, Sherman M, Thompson E, Yena P, Bowen J, Gastier-Foster JM, Gerken M, Leraas KM, Lichtenberg TM, Ramirez NC, Wise L, Zmuda E, Corcoran N, Costello T, Hovens C, Carvalho AL, de Carvalho AC, Fregnani JH, Longatto-Filho A, Reis RM, Scapulatempo-Neto C, Silveira HC, Vidal DO, Burnette A, Eschbacher J, Hermes B, Noss A, Singh R, Anderson ML, Castro PD, Ittmann M, Huntsman D, Kohl B, Le X, Thorp R, Andry C, Duffy ER, Lyadov V, Paklina O, Setdikova G, Shabunin A, Tavobilov M, McPherson C, Warnick R, Berkowitz R, Cramer D, Feltmate C, Horowitz N, Kibel A, Muto M, Raut CP, Malykh A, Barnholtz-Sloan JS, Barrett W, Devine K, Fulop J, Ostrom QT, Shimmel K, Wolinsky Y, Sloan AE, De Rose A, Giuliante F, Goodman M, Karlan BY, Hagedorn CH, Eckman J, Harr J, Myers J, Tucker K, Zach LA, Deyarmin B, Hu H, Kvecher L, Larson C, Mural RJ, Somiari S, Vicha A, Zelinka T, Bennett J, Iacocca M, Rabeno B, Swanson P, Latour M, Lacombe L, Têtu B, Bergeron A, McGraw M, Staugaitis SM, Chabot J, Hibshoosh H, Sepulveda A, Su T, Wang T, Potapova O, Voronina O, Desjardins L, Mariani O, Roman-Roman S, Sastre X, Stern MH, Cheng F, Signoretti S, Berchuck A, Bigner D, Lipp E, Marks J, McCall S, McLendon R, Secord A, Sharp A, Behera M, Brat DJ, Chen A, Delman K, Force S, Khuri F, Magliocca K, Maithel S, Olson JJ, Owonikoko T, Pickens A, Ramalingam S, Shin DM, Sica G, Van Meir EG, Zhang H, Eijckenboom W, Gillis A, Korpershoek E, Looijenga L, Oosterhuis W, Stoop H, van Kessel KE, Zwarthoff EC, Calatozzolo C, Cuppini L, Cuzzubbo S, DiMeco F, Finocchiaro G, Mattei L, Perin A, Pollo B, Chen C, Houck J, Lohavanichbutr P, Hartmann A, Stoehr C, Stoehr R, Taubert H, Wach S, Wullich B, Kycler W, Murawa D, Wiznerowicz M, Chung K, Edenfield WJ, Martin J, Baudin E, Bubley G, Bueno R, De Rienzo A, Richards WG, Kalkanis S, Mikkelsen T, Noushmehr H, Scarpace L, Girard N, Aymerich M, Campo E, Giné E, Guillermo AL, Van Bang N, Hanh PT, Phu BD, Tang Y, Colman H, Evason K, Dottino PR, Martignetti JA, Gabra H, Juhl H, Akeredolu T, Stepa S, Hoon D, Ahn K, Kang KJ, Beuschlein F, Breggia A, Birrer M, Bell D, Borad M, Bryce AH, Castle E, Chandan V, Cheville J, Copland JA, Farnell M, Flotte T, Giama N, Ho T, Kendrick M, Kocher JP, Kopp K, Moser C, Nagorney D, OBrien D, ONeill BP, Patel T, Petersen G, Que F, Rivera M, Roberts L, Smallridge R, Smyrk T, Stanton M, Thompson RH, Torbenson M, Yang JD, Zhang L, Brimo F, Ajani JA, Gonzalez AMA, Behrens C, Bondaruk J, Broaddus R, Czerniak B, Esmaeli B, Fujimoto J, Gershenwald J, Guo C, Lazar AJ, Logothetis C, Meric-Bernstam F, Moran C, Ramondetta L, Rice D, Sood A, Tamboli P, Thompson T, Troncoso P, Tsao A, Wistuba I, Carter C, Haydu L, Hersey P, Jakrot V, Kakavand H, Kefford R, Lee K, Long G, Mann G, Quinn M, Saw R, Scolyer

R, Shannon K, Spillane A, Stretch J, Synott M, Thompson J, Wilmott J, Al-Ahmadie H, Chan TA, Ghossein R, Gopalan A, Levine DA, Reuter V, Singer S, Singh B, Tien NV, Broudy T, Mirsaidi C, Nair P, Drwiega P, Miller J, et al. (2018) A Pan-Cancer Analysis of Enhancer Expression in Nearly 9000 Patient Samples. Cell 173: 386–399.

[348] Takeda DY, Spisák S, Seo JH, Bell C, OConnor E, Korthauer K, Ribli D, Csabai I, Solymosi N, Szállási Z, Stillman DR, Cejas P, Qiu X, Long HW, Tisza V, Nuzzo PV, Rohanizadegan M, Pomerantz MM, Hahn WC, Freedman ML (2018) A Somatically Acquired Enhancer of the Androgen Receptor Is a Noncoding Driver in Advanced Prostate Cancer. Cell 174: 422–432.

[349] Moorthy SD, Davidson S, Shchuka VM, Singh G, Malek-Gilani N, Langroudi L, Martchenko A, So V, Macpherson NN, Mitchell JA (2017) Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes. Genome Res 27: 246–258.

[350] Northcott Pa, Lee C, Zichner T, Stütz AM, Erkek S, Kawauchi D, Shih DJH, Hovestadt V, Zapatka M, Sturm D, Jones DTW, Kool M, Remke M, Cavalli FMG, Zuyderduyn S, Bader GD, VandenBerg S, Esparza LA, Ryzhova M, Wang W, Wittmann A, Stark S, Sieber L, Seker-Cin H, Linke L, Kratochwil F, Jäger N, Buchhalter I, Imbusch CD, Zipprich G, Raeder B, Schmidt S, Diessl N, Wolf S, Wiemann S, Brors B, Lawerenz C, Eils J, Warnatz HJ, Risch T, Yaspo ML, Weber UD, Bartholomae CC, von Kalle C, Turányi E, Hauser P, Sanden E, Darabi A, Siesjö P, Sterba J, Zitterbart K, Sumerauer D, van Sluis P, Versteeg R, Volckmann R, Koster J, Schuhmann MU, Ebinger M, Grimes HL, Robinson GW, Gajjar A, Mynarek M, von Hoff K, Rutkowski S, Pietsch T, Scheurlen W, Felsberg J, Reifenberger G, Kulozik AE, von Deimling A, Witt O, Eils R, Gilbertson RJ, Korshunov A, Taylor MD, Lichter P, Korbel JO, Wechsler-Reya RJ, Pfister SM (2014) Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. Nature .

[351] Weischenfeldt J, Dubash T, Drainas AP, Mardin BR, Chen Y, Stütz AM, Waszak SM, Bosco G, Halvorsen AR, Raeder B, Efthymiopoulos T, Erkek S, Siegl C, Brenner H, Brustugun OT, Dieter SM, Northcott PA, Petersen I, Pfister SM, Schneider M, Solberg SK, Thunissen E, Weichert W, Zichner T, Thomas R, Peifer M, Helland A, Ball CR, Jechlinger M, Sotillo R, Glimm H, Korbel JO (2017) Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. Nature Genetics 49: 65–74.

[352] O'Geen H, Ren C, Nicolet CM, Perez AA, Halmai J, Le VM, Mackay JP, Farnham PJ, Segal DJ (2017) dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. Nucleic Acids Res 45: 9901–9916.

[353] Zhang X, Choi PS, Francis JM, Imielinski M, Watanabe H, Cherniack AD, Meyerson M (2016) Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. Nat Genet 48: 176–182.

[354] Sanjana NE, Shalem O, Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. Nature Methods 11: 783–784.

[355] Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan

Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH (2008) Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. Cell 133: 1106–1117.

[356] Yun K, Wold B (1996) Skeletal muscle determination and differentiation: story of a core regulatory network and its context. Current opinion in cell biology 8: 877–89.

[357] Singh H, Khan AA, Dinner AR (2014) Gene regulatory networks in the immune system. Trends in Immunology 35: 211–218.

[358] Takahashi K, Yamanaka S (2006) Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell 126: 663–676.

[359] Dynan WS, Tjian R (1983) The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell 35: 79–87.

[360] Hesselberth JR, Chen X, Zhang Z, Sabo PJ, Sandstrom R, Reynolds AP, Thurman RE, Neph S, Kuehn MS, Noble WS, Fields S, Stamatoyannopoulos JA (2009) Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. Nature Methods 6: 283–289.

[361] Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM (2009) A census of human transcription factors: function, expression and evolution. Nature Reviews Genetics 10: 252–263.

[362] Geertz M, Shore D, Maerkl SJ (2012) Massively parallel measurements of molecular interaction kinetics on a microfluidic platform. Proceedings of the National Academy of Sciences of the United States of America 109: 16540–5.

[363] Schneider TD, Stephens R (1990) Sequence logos: a new way to display consensus sequences. Nucleic Acids Research 18: 6097–6100.

[364] Jothi R, Balaji S, Wuster A, Grochow JA, Gsponer J, Przytycka TM, Aravind L, Babu MM (2009) Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. Molecular Systems Biology 5: 294.

[365] Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne JB, Volkert TL, Fraenkel E, Gifford DK, Young RA (2002) Transcriptional Regulatory Networks in Saccharomyces cerevisiae. Science 298: 799–804.

[366] Gertz J, Reddy TE, Varley KE, Garabedian MJ, Myers RM (2012) Genistein and bisphenol A exposure cause estrogen receptor 1 to bind thousands of sites in a cell type-specific manner. Genome Research 22: 2153–2162.

[367] Barrera LA, Vedenko A, Kurland JV, Rogers JM, Gisselbrecht SS, Rossin EJ, Woodard J, Mariani L, Kock KH, Inukai S, Siggers T, Shokri L, Gordan R, Sahni N, Cotsapas C, Hao T, Yi S, Kellis M, Daly MJ, Vidal M, Hill DE, Bulyk ML (2016) Survey of variation in human transcription factors reveals prevalent DNA binding changes. Science 351: 1450–1454.

[368] Worsley Hunt R, Wasserman WW (2014) Non-targeted transcription factors motifs are a systemic component of ChIP-seq datasets. Genome Biology 15: 412.

[369] Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT (2018) The Human Transcription Factors. Cell 172: 650–665.

[370] Jolma A, Yin Y, Nitta KR, Dave K, Popov A, Taipale M, Enge M, Kivioja T, Morgunova E, Taipale J (2015) DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature 527: 384–388.

[371] Panne D (2008) The enhanceosome. Current Opinion in Structural Biology 18: 236–242.

[372] Akerblom IE, Slater EP, Beato M, Baxter JD, Mellon PL (1988) Negative regulation by glucocorticoids through interference with a cAMP responsive enhancer. Science (New York, NY) 241: 350–3.

[373] Reiter F, Wienerroither S, Stark A (2017) Combinatorial function of transcription factors and cofactors. Current Opinion in Genetics & Development 43: 73–81.

[374] Frietze S, Farnham PJ (2011) Transcription Factor Effector Domains. In: Sub-cellular biochemistry, volume 52. pp. 261–277. doi:10.1007/978-90-481-9069-0{_}12. URL http://www.ncbi.nlm.nih.gov/pubmed/21557087http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4151296http://www.springerlink.com/index/10.1007/978-90-481-9069-0_12.

[375] Lavin Y, Winter D, Blecher-Gonen R, David E, Keren-Shaul H, Merad M, Jung S, Amit I (2014) Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell 159: 1312–1326.

[376] Vogt PK, Bos TJ (1990) jun: oncogene and transcription factor. Advances in cancer research 55: 1–35.

[377] Chinenov Y, Kerppola TK (2001) Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. Oncogene 20: 2438–2452.

[378] Jochum W, Passegué E, Wagner EF (2001) AP-1 in mouse development and tumorigenesis. Oncogene 20: 2401–2412.

[379] Bergers G, Graninger P, Braselmann S, Wrighton C, Busslinger M (1995) Transcriptional activation of the fra-1 gene by AP-1 is mediated by regulatory sequences in the first intron. Molecular and Cellular Biology 15: 3748–3758.

[380] Vandel L, Pfarr CM, Huguier S, Loiseau L, Sergeant A, Castellazzi M (1995) Increased transforming activity of JunB and JunD by introduction of an heterologous homodimerization domain. Oncogene 10: 495–507.

[381] Agarwal SK, Guru SC, Heppner C, Erdos MR, Collins RM, Park SY, Saggar S, Chandrasekharappa SC, Collins FS, Spiegel AM, Marx SJ, Burns AL (1999) Menin interacts with the AP1 transcription factor JunD and represses JunD-activated transcription. Cell 96: 143–52.

[382] Pfarr CM, Mechta F, Spyrou G, Lallemand D, Carillo S, Yaniv M (1994) Mouse JunD negatively regulates fibroblast growth and antagonizes transformation by ras. Cell 76: 747–60.

[383] Johnson AC, Murphy BA, Matelis CM, Rubinstein Y, Piebenga EC, Akers LM, Neta G, Vinson C, Birrer M (2000) Activator Protein-1 Mediates Induced but not Basal Epidermal Growth Factor Receptor Gene Expression. Molecular Medicine 6: 17–27.

[384] Zenz R, Scheuch H, Martin P, Frank C, Eferl R, Kenner L, Sibilia M, Wagner EF (2003) c-Jun regulates eyelid closure and skin tumor development through EGFR signaling. Developmental cell 4: 879–89.

[385] O'Shea EK, Rutkowski R, Kim PS (1992) Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell 68: 699–708.

[386] Wang ZY, Sato H, Kusam S, Sehra S, Toney LM, Dent AL (2005) Regulation of IL-10 gene expression in Th2 cells by Jun proteins. J Immunol 174: 2098–2105.

[387] Mialon A, Sankinen M, Söderström H, Junttila TT, Holmström T, Koivusalo R, Papageorgiou AC, Johnson RS, Hietanen S, Elenius K, Westermarck J (2005) DNA topoisomerase I is a cofactor for c-Jun in the regulation of epidermal growth factor receptor expression and cancer cell proliferation. Molecular and cellular biology 25: 5040–51.

[388] Angel P, Hattori K, Smeal T, Karin M (1988) The jun proto-oncogene is positively autoregulated by its product, Jun/AP-1. Cell 55: 875–885.

[389] Vasanwala FH, Kusam S, Toney LM, Dent AL (2002) Repression of AP-1 Function: A Mechanism for the Regulation of Blimp-1 Expression and B Lymphocyte Differentiation by the B Cell Lymphoma-6 Protooncogene. The Journal of Immunology 169: 1922–1929.

[390] Jonat C, Rahmsdorf HJ, Park KK, Cato AC, Gebel S, Ponta H, Herrlich P (1990) Antitumor promotion and antiinflammation: down-modulation of AP-1 (Fos/Jun) activity by glucocorticoid hormone. Cell 62: 1189–204.

[391] Biddie S, John S, Sabo P, Thurman R, Johnson T, Schiltz R, Miranda T, Sung MH, Trump S, Lightman S, Vinson C, Stamatoyannopoulos J, Hager G (2011) Transcription Factor AP1 Potentiates Chromatin Accessibility and Glucocorticoid Receptor Binding. Molecular Cell 43: 145–155.

[392] von Knethen A, Callsen D, Brüne B (1999) NF-kappaB and AP-1 activation by nitric oxide attenuated apoptotic cell death in RAW 264.7 macrophages. Molecular biology of the cell 10: 361–72.

[393] Verma IM, Stevenson JK, Schwarz EM, Van Antwerp D, Miyamoto S (1995) Rel/NF-kappa B/I kappa B family: intimate tales of association and dissociation. Genes & Development 9: 2723–2735.

[394] Stein B, Baldwin AS, Ballard DW, Greene WC, Angel P, Herrlich P (1993) Cross-coupling of the NF-kappa B p65 and Fos/Jun transcription factors produces potentiated biological function. The EMBO journal 12: 3879–91.

[395] Najafova Z, Tirado-Magallanes R, Subramaniam M, Hossan T, Schmidt G, Nagarajan S, Baumgart SJ, Mishra VK, Bedi U, Hesse E, Knapp S, Hawse JR, Johnsen SA (2017) BRD4 localization to lineage-specific enhancers is associated with a distinct transcription factor repertoire. Nucleic Acids Res 45: 127–141.

[396] LeRoy G, Rickards B, Flint S (2008) The Double Bromodomain Proteins Brd2 and Brd3 Couple Histone Acetylation to Transcription. Molecular Cell 30: 51–60.

[397] Zuber J, Shi J, Wang E, Rappaport AR, Herrmann H, Sison EA, Magoon D, Qi J, Blatt K, Wunderlich M, Taylor MJ, Johns C, Chicas A, Mulloy JC, Kogan SC, Brown P, Valent P, Bradner JE, Lowe SW, Vakoc CR (2011) RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature 478: 524–528.

[398] French CA, Miyoshi I, Kubonishi I, Grier HE, Perez-Atayde AR, Fletcher JA (2003) BRD4-NUT fusion oncogene: a novel mechanism in aggressive carcinoma. Cancer research 63: 304–7.

[399] Zanca C, Villa GR, Benitez JA, Thorne AH, Koga T, D'Antonio M, Ikegami S, Ma J, Boyer AD, Banisadr A, Jameson NM, Parisian AD, Eliseeva OV, Barnabe GF, Liu F, Wu S, Yang H, Wykosky J, Frazer KA, Verkhusha VV, Isaguliants MG, Weiss WA, Gahman TC, Shiau AK, Chen CC, Mischel PS, Cavenee WK, Furnari FB (2017) Glioblastoma cellular cross-talk converges on NF-κB to attenuate EGFR inhibitor sensitivity. Genes Dev 31: 1212–1227.

[400] Nicodeme E, Jeffrey KL, Schaefer U, Beinke S, Dewell S, Chung Cw, Chandwani R, Marazzi I, Wilson P, Coste H, White J, Kirilovsky J, Rice CM, Lora JM, Prinjha RK, Lee K, Tarakhovsky A (2010) Suppression of inflammation by a synthetic histone mimic. Nature 468: 1119–1123.

[401] Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, Morse EM, Keates T, Hickman TT, Felletar I, Philpott M, Munro S, McKeown MR, Wang Y, Christie AL, West N, Cameron MJ, Schwartz B, Heightman TD, La Thangue N, French CA, Wiest O, Kung AL, Knapp S, Bradner JE (2010) Selective inhibition of BET bromodomains. Nature 468: 1067–1073.

[402] Bid HK, Phelps DA, Xaio L, Guttridge DC, Lin J, London C, Baker LH, Mo X, Houghton PJ (2016) The Bromodomain BET Inhibitor JQ1 Suppresses Tumor Angiogenesis in Models of Childhood Sarcoma .

[403] Chen Y, Xu L, Mayakonda A, Huang ML, Kanojia D, Tan TZ, Dakle P, Lin RYT, Ke XY, Said JW, Chen J, Gery S, Ding LW, Jiang YY, Pang A, Puhaindran ME, Goh BC, Koeffler HP (2019) Bromodomain and extraterminal proteins foster the core transcriptional regulatory programs and confer vulnerability in liposarcoma. Nature Communications 10: 1353.

[404] Stonestrom AJ, Hsu SC, Jahn KS, Huang P, Keller CA, Giardine BM, Kadauke S, Campbell AE, Evans P, Hardison RC, Blobel GA (2015) Functions of BET proteins in erythroid gene expression. Blood 125: 2825–34.

[405] Leonard B, Brand TM, O'Keefe RA, Lee ED, Zeng Y, Kemmer JD, Li H, Grandis JR, Bhola NE (2018) BET Inhibition Overcomes Receptor Tyrosine Kinase-Mediated Cetuximab Resistance in HNSCC. Cancer Res 78: 4331–4343.

[406] Alani R, Brown P, Binétruy B, Dosaka H, Rosenberg RK, Angel P, Karin M, Birrer MJ (1991) The transactivating domain of the c-Jun proto-oncoprotein is required for cotransformation of rat embryo cells. Molecular and cellular biology 11: 6286–95.

[407] Westwick JK, Cox AD, Der CJ, Cobb MH, Hibi M, Karin M, Brenner DA (1994) Oncogenic Ras activates c-Jun via a separate pathway from the activation of extracellular signal-regulated kinases. Proceedings of the National Academy of Sciences 91: 6030–6034.

[408] Mechta F, Lallemand D, Pfarr CM, Yaniv M (1997) Transformation by ras modifies AP1 composition and activity. Oncogene 14: 837–847.

[409] Treinies I, Paterson HF, Hooper S, Wilson R, Marshall CJ (1999) Activated MEK stimulates expression of AP-1 components independently of phosphatidylinositol 3-kinase (PI3-kinase) but requires a PI3-kinase signal To stimulate DNA synthesis. Molecular and cellular biology 19: 321–9.

[410] Binétruy B, Smeal T, Karin M (1991) Ha-Ras augments c-Jun activity and stimulates phosphorylation of its activation domain. Nature 351: 122–127.

[411] Smeal T, Binetruy B, Mercola DA, Birrer M, Karin M (1991) Oncogenic and transcriptional cooperation with Ha-Ras requires phosphorylation of c-Jun on serines 63 and 73. Nature 354: 494–496.

[412] Johnson R, Spiegelman B, Hanahan D, Wisdom R (1996) Cellular transformation and malignancy induced by ras require c-jun. Molecular and cellular biology 16: 4504–11.

[413] Fulco CP, Munschauer M, Anyoha R, Munson G, Grossman SR, Perez EM, Kane M, Cleary B, Lander ES, Engreitz JM (2016) Systematic mapping of functional enhancerpromoter connections with CRISPR interference. Science 354: 769–773.

[414] Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, Reddy TE, Crawford GE, Gersbach CA (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nature Methods 12: 1143–1149.

[415] Erjala K, Sundvall M, Junttila TT, Zhang N, Savisalo M, Mali P, Kulmala J, Pulkkinen J, Grenman R, Elenius K (2006) Signaling via ErbB2 and ErbB3 associates with resistance and epidermal growth factor receptor (EGFR) amplification with sensitivity to EGFR inhibitor gefitinib in head and neck squamous cell carcinoma cells. Clin Cancer Res 12: 4103–4111.

[416] Loeffler-Ragg J, Witsch-Baumgartner M, Tzankov A, Hilbe W, Schwentner I, Sprinzl GM, Utermann G, Zwierzina H (2006) Low incidence of mutations in EGFR kinase domain in Caucasian patients with head and neck squamous cell carcinoma. Eur J Cancer 42: 109–111.

[417] Patel A, Tirosh I, Trombetta J (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science : 1–9.

[418] Yochum ZA, Cades J, Wang H, Chatterjee S, Simons BW, OBrien JP, Khetarpal SK, Lemtiri-Chlieh G, Myers KV, Huang EHB, Rudin CM, Tran PT, Burns TF (2019) Targeting the EMT transcription factor TWIST1 overcomes resistance to EGFR inhibitors in EGFR-mutant non-small-cell lung cancer. Oncogene 38: 656–670.

[419] Moore-Carrasco R, Busquets S, Almendro V, Palanki M, López-Soriano FJ, Argilés JM (2007) The AP-1/NF-kappaB double inhibitor SP100030 can revert muscle wasting during experimental cancer cachexia. International journal of oncology 30: 1239–45.

[420] Shen Q, Uray IP, Li Y, Krisko TI, Strecker TE, Kim HT, Brown PH (2008) The AP-1 transcription factor regulates breast cancer cell growth via cyclins and E2F factors. Oncogene 27: 366–377.

[421] Tsuchida K, Chaki H, Takakura T, Yokotani J, Aikawa Y, Shiozawa S, Gouda H, Hirono S (2004) Design, synthesis, and biological evaluation of new cyclic disulfide decapeptides that inhibit the binding of AP-1 to DNA. Journal of medicinal chemistry 47: 4239–46.

[422] Aikawa Y, Morimoto K, Yamamoto T, Chaki H, Hashiramoto A, Narita H, Hirono S, Shiozawa S (2008) Treatment of arthritis with a selective inhibitor of c-Fos/activator protein-1. Nature Biotechnology 26: 817–823.

[423] Mertz JA, Conery AR, Bryant BM, Sandy P, Balasubramanian S, Mele DA, Bergeron L, Sims RJ (2011) Targeting MYC dependence in cancer by inhibiting BET bromodomains. Proceedings of the National Academy of Sciences 108: 16669–16674.

[424] Bolden J, Tasdemir N, Dow L, vanEs J, Wilkinson J, Zhao Z, Clevers H, Lowe S (2014) Inducible InVivo Silencing of Brd4 Identifies Potential Toxicities of Sustained BET Protein Inhibition. Cell Reports 8: 1919–1929.

[425] Wang F, Liu H, Blanton W, Belkina A, Lebrasseur N, Denis G (2009) Brd2 disruption in mice causes severe obesity without Type2 diabetes . Biochemical Journal 425: 71–85.

[426] Andrieu G, Belkina AC, Denis GV (2016) Clinical trials for BET inhibitors run ahead of the science. Drug discovery today Technologies 19: 45–50.

[427] Bui MH, Lin X, Albert DH, Li L, Lam LT, Faivre EJ, Warder SE, Huang X, Wilcox D, Donawho CK, Sheppard GS, Wang L, Fidanze S, Pratt JK, Liu D, Hasvold L, Uziel T, Lu X, Kohlhapp F, Fang G, Elmore SW, Rosenberg SH, McDaniel KF, Kati WM, Shen Y (2017) Preclinical Characterization of BET Family Bromodomain Inhibitor ABBV-075 Suggests Combination Therapeutic Strategies. Cancer Research 77: 2976–2989.

[428] Long J, Li B, Rodriguez-Blanco J, Pastori C, Volmar CH, Wahlestedt C, Capobianco A, Bai F, Pei XH, Ayad NG, Robbins DJ (2014) The BET Bromodomain Inhibitor I-BET151 Acts Downstream of Smoothened Protein to Abrogate the Growth of Hedgehog Protein-driven Cancers. Journal of Biological Chemistry 289: 35494–35502.

[429] Díaz T, Rodríguez V, Lozano E, Mena MP, Calderón M, Rosiñol L, Martínez A, Tovar N, Pérez-Galán P, Bladé J, Roué G, De Larrea CF (2017) The BET bromodomain inhibitor CPI203 improves lenalidomide and dexamethasone activity in in vitro and in vivo models of multiple myeloma by blockade of ikaros and MYC signaling. Haematologica 102: 1776–1784.

[430] Picaud S, Da Costa D, Thanasopoulou A, Filippakopoulos P, Fish PV, Philpott M, Fedorov O, Brennan P, Bunnage ME, Owen DR, Bradner JE, Taniere P, O'Sullivan B, Muller S, Schwaller J, Stankovic T, Knapp S (2013) PFI-1, a Highly Selective Protein Interaction Inhibitor, Targeting BET Bromodomains. Cancer Research 73: 3336–3346.

[431] Coudé MM, Braun T, Berrou J, Dupont M, Bertrand S, Masse A, Raffoux E, Itzykson R, Delord M, Riveiro ME, Herait P, Baruchel A, Dombret H, Gardin C (2015) BET inhibitor OTX015 targets BRD2 and BRD4 and decreases c-MYC in acute leukemia cells. Oncotarget 6: 17698–712.

[432] Picaud S, Wells C, Felletar I, Brotherton D, Martin S, Savitsky P, Diez-Dacal B, Philpott M, Bountra C, Lingard H, Fedorov O, Muller S, Brennan PE, Knapp S, Filippakopoulos P (2013) RVX-208, an inhibitor of BET transcriptional regulators with selectivity for the second bromodomain. Proceedings of the National Academy of Sciences 110: 19754–19759.

[433] Borthakur G, Wolff JE, Aldoss I, Hu B, Dinh M, Torres A, Chen X, Rizzieri D, Sood A, Odenike O, Jonas BA (2018) First-in-human study of ABBV-075 (mivebresib), a pan-inhibitor of bromodomain and extra terminal (BET) proteins, in patients (pts) with relapsed/refractory (RR) acute myeloid leukemia (AML): Preliminary data. Journal of Clinical Oncology 36: 7019–7019.

[434] Massard C, Soria J, Stathis A, Delord J, Awada A, Peters S, Lewin J, Bekradda M, Rezai K, Zeng Z, Azher H, Perez S, Siu L (2016) A phase Ib trial with MK-8628/OTX015, a small molecule inhibitor of bromodomain (BRD) and extra-terminal (BET) proteins, in patients with selected advanced solid tumors. European Journal of Cancer 69: S2-S3.

[435] Aftimos PG, Bechter O, Awada A, Jungels C, Dumez H, Huyvaert N, Costermans J, Lee C, Meeus MA, Burkard U, Musa H, Zhao Y, Schoffski P (2017) Phase I first-in-man trial of a novel bromodomain and extra-terminal domain (BET) inhibitor (BI 894999) in patients (Pts) with advanced solid tumors. Journal of Clinical Oncology 35: 2504–2504.

[436] Amorim S, Stathis A, Gleeson M, Iyengar S, Magarotto V, Leleu X, Morschhauser F, Karlin L, Broussais F, Rezai K, Herait P, Kahatt C, Lokiec F, Salles G, Facon T, Palumbo A, Cunningham D, Zucca E, Thieblemont C (2016) Bromodomain inhibitor OTX015 in patients with lymphoma or multiple myeloma: a dose-escalation, open-label, pharmacokinetic, phase 1 study. The Lancet Haematology 3: e196-e204.

[437] Blum KA, Abramson J, Maris M, Flinn I, Goy A, Mertz J, Sims R, Garner F, Senderowicz A, Younes A (2018) 41OA phase I study of CPI-0610, a bromodomain and extra terminal protein (BET) inhibitor in patients with relapsed or refractory lymphoma. Annals of Oncology 29.

[438] Hottinger AF, Sanson M, Moyal E, Delord JP, De Micheli R, Rezai K, Leung AC, Perez S, Bekradda M, Lachaux N, Lokiec FM, Chinot OL (2016) Dose optimization of MK-8628 (OTX015), a small molecule inhibitor of bromodomain and extra-terminal (BET) proteins, in patients (pts) with recurrent glioblastoma (GB). Journal of Clinical Oncology 34: e14123-e14123.

[439] Maruyama T, Farina A, Dey A, Cheong J, Bermudez VP, Tamura T, Sciortino S, Shuman J, Hurwitz J, Ozato K (2002) A Mammalian Bromodomain Protein, Brd4, Interacts with Replication Factor C and Inhibits Progression to S Phase. Molecular and Cellular Biology 22: 6509–6520.

[440] Koga T, Li B, Figueroa JM, Ren B, Chen CC, Carter BS, Furnari FB (2018) Mapping of genomic EGFRvIII deletions in glioblastoma: insight into rearrangement mechanisms and biomarker development. Neuro Oncol 20: 1310–1320.

[441] Deshpande V, Luebeck J, Nguyen NPD, Bakhtiari M, Turner KM, Schwab R, Carter H, Mischel PS, Bafna V (2019) Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. Nature Communications 10: 392.

[442] Sanborn JZ, Salama SR, Grifford M, Brennan CW, Mikkelsen T, Jhanwar S, Katzman S, Chin L, Haussler D (2013) Double minute chromosomes in glioblastoma multiforme are revealed by precise reconstruction of oncogenic amplicons. Cancer research 73: 6036–45.