THE REGULATORY LANDSCAPE OF THE GLIOMA-ASSOCIATED TRANSCRIPTION FACTOR CAPICUA

by

Juan Marlo R. Firme

B.Sc., The University of British Columbia, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Genome Science and Technology)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

December 2014

Abstract

The metazoan developmental gene *capicua transcriptional repressor* (*CIC*) encodes a transcription factor that transduces receptor tyrosine kinase signaling into gene expression changes. Aberrant CIC function is implicated in oligodendroglioma (ODG) development since one *CIC* allele is lost while the other is mutated in ~70% of ODGs. We therefore investigated how CIC can affect gene expression at a genome-wide level by inactivating CIC in HEK293a cells and subsequently measuring gene expression changes using microarrays. From this, gene expression changes spanning entire chromosomes were detected. Additionally, 24 candidate CIC-regulated genes were identified in HEK293a cells that also have evidence of CIC-dependent regulation in ODGs sequenced by The Cancer Genome Atlas (TCGA). Of these 24 genes, 5 genes (*CNTFR*, *DUSP6*, *GPR3*, *SHC3*, and *SPRY4*) with reported functions in mitogen-activated protein kinase (MAPK) signaling and central nervous system (CNS) development were further validated to undergo CIC-dependent regulation in HeLa cells. Finally, investigating how different *CIC* mutations affect gene expression revealed that different types of ODG-associated *CIC* mutations either abrogated or potentially preserved CIC's transcriptionally repressive activity. These findings shed insight into possible roles for CIC in regulating gene expression at a chromosome-wide scale, MAPK signaling, CNS development, and ODG development.

Preface

Chapter 1. Data for Figure 1 was gathered from The Cancer Genome Atlas (TCGA) though the Cancer Genomics cBioPortal[1,2] and from other applicable sources[3–7].

Chapter 2. The data presented are primarily based on work performed at the BC Cancer Agency by Marlo Firme, Dr. Suganthi Chittaranjan, Susanna Chan, and Jeungeun Song in the Marco Marra laboratory at Canada's Michael Smith Genome Sciences Centre, by Emma Laks in the Sam Aparicio laboratory at the Department of Molecular Oncology, and by Amy Lum at the Centre for Translational and Applied Genomics (CTAG). Sections 2.1.2, 2.1.3, and 2.1.10 of the Materials and Methods section have been directly quoted from previously published material of which I am a coauthor[8]. Data on Figure 2b were generated by Jeungeun Song. Data on Figure 3a, 3b, and 3c were generated by Emma Laks and Marlo Firme. Data on Figure 3d were generated by Amy Lum. Data from Figures 4b and 8 were gathered from TCGA through the Cancer Genomics cBioPortal[1,2] and the Cancer Browser from the University of California, Santa Cruz[9]. The Flag-CIC constructs indicated in Figure 9 were designed by Dr. Suganthi Chittaranjan and generated by Susanna Chan and Marlo Firme. The luciferase reporter construct depicted in Figure 9a was a gift from Dr. Takuro Nakamura[10]. The rest of the experiments were designed and performed by Marlo Firme under the counsel of Dr. Marco Marra.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

1p/19q codeletion	deletion of chromosomal arms 1p and 19q
2HG	2-hydroxyglutarate
αKG	alpha-ketoglutarate
ADG	adult diffuse glioma
C1	CIC's highly conserved C-terminal motif
ChIP-seq	chromatin immunoprecipitation followed by next-generation sequencing
CIC	*capicua transcriptional repressor*
CIC inact	1p/19q-codeleted ADGs with predicted CIC-inactivating mutations
CIC C1	1p/19q-codeleted ADGs with single amino acid mutations targeting the C1 motif
CIC HMG	1p/19q-codeleted ADGs with predicted CIC-inactivating mutations targeting the HMG box domain
cic^{WT}	parental HEK293a clone
cic^{ZFN1}	cic^{WT}-derived clone with reduced CIC expression
cic^{ZFN2}	cic^{ZFN1}-derived clone with essentially undetectable CIC expression
CIC-L	long CIC isoform
CIC-S	short CIC isoform
EGFR	epidermal growth factor receptor
ETS	E-twenty-six
ETV1/4/5	*ETV1*, *ETV4*, and *ETV5*
FDR	false discovery rate
FISH	fluorescence *in situ* hybridization

FUBP1	*far upstream element binding protein 1*
GBM	glioblastoma
HMG	high mobility group
IDH1/2	*isocitrate dehydrogenase 1* or *2*
MAPK	mitogen-activated protein kinase
ODG	oligodendroglioma
Q564X	Gln564* nonsense mutation
R1515H	Arg1515His missense mutation
R201W	Arg201Trp missense mutation
RTK	receptor tyrosine kinase
SCA1	spinocerebellar ataxia type I
TBP	*TATA-binding protein*
TCGA	The Cancer Genome Atlas
WHO	World Health Organization
ZFN	zinc finger nuclease

Acknowledgements

Thank you to Ms. Donna Anderson for your generosity and support of the work in this thesis. Thank you to Dr. Takuro Nakamura and Dr. Carol MacKintosh, who graciously provided us with DNA constructs. Thank you to Emma Laks and Amy Lum for performing experiments in this study. Thank you to Dr. Suganthi Chittaranjan and Susanna Chan for assistance in the lab and for guiding and supporting me. Thank you to Diane Trinh, Julia Pon, Dr. Alessia Gagliardi, Ryan Huff, Veronique LeBlanc, Jeungeun Song, Angelica Lee, and Dr. Isabel Serrano for your assistance with reagents and helpful wet lab advice. Thank you to Dr. Jill Mwenifumbo, Olena Morozova, Emilia Lim, Dr. Farah Zahir, Rodrigo Goya, and Elizabeth Chun for your assistance in bioinformatics. Thank you to Lulu Crisostomo for always helping out with administrative tasks. Thank you to the faculty and staff of the Genome Sciences and Technology Program at the University of British Columbia, especially Sharon Ruschkowski, Dr. Stephen Withers, Dr. Rosie Redfield, and Dr. Phil Hieter for helping to further my learning. Thank you to my committee members Dr. Stephen Yip and Dr. Gregory Cairncross for your enthusiasm and support in this research. Finally, thank you to Dr. Marco Marra for leading by example to teach students like me to become better and wiser.

Chapter 1: Introduction

1.1 Traditional Histological Classification of Adult Diffuse Gliomas

Adult Diffuse Gliomas (ADGs), accounting for ~85% of all central nervous tumours, are the most common type of primary brain tumour[11]. To date, ADGs are considered malignant and incurable. This is largely due to their infiltrative and migratory nature, making them difficult to manage surgically[12]. Less aggressive or lower-grade ADGs also typically progress to higher-grade lesions over time[13].

ADGs can be highly heterogeneous with respect to the clinical courses that they follow[13,14]. The clinical courses of ADGs, however, generally correlate with certain microscopic characteristics. For this reason, the World Health Organization (WHO) has classified ADGs according to their histology (resemblance to specific cell types) and their grade, determined by characteristics such as the presence of anaplastic changes (loss of structural differentiation) and necrosis[13]. Classifying ADGs this way allows doctors to better predict patient survival and aids in decision-making for choosing between different treatment modalities[13–15]. ADGs range in grade from II to IV and the three main histological types of ADG are oligodendroglioma, astrocytoma, and oligoastrocytoma[13].

Oligodendrogliomas (ODGs) are slow-growing neoplasms that comprise ~5-6% of central nervous system gliomas[11,16]. ODGs get their name from their resemblance to oligodendrocytes, non-neuronal brain cells that myelinate or insulate the axons of neurons[17]. With overall survival rates of ~70-80% over 5 years, ODGs generally have the most favourable prognosis of the three ADG types[13,18,19]. ODGs are also characteristically responsive to chemotherapy but can still recur and progress to higher grades[12,18,20].

Astrocytomas, comprising ~75% of central nervous system gliomas, are the most common and generally most aggressive histological type of ADG[11]. Astrocytoma cells resemble astrocytes, highly abundant star-shaped glial cells that are important in maintaining brain structure, nutrition and homeostasis[17]. 5 year overall survival remains at ~47% and ~27% for WHO grade II and III astrocytomas, respectively[11]. Grade IV astrocytomas, more commonly known as glioblastomas (GBMs), have an overall survival rate of just ~5% over 5 years[11]. GBMs can either arise spontaneously (primary GBM) or from the progression of a lower-grade glioma (secondary GBM)[13]. Relative to secondary GBMs, primary GBMs have a overall higher age of onset[21], indicating that primary and secondary GBMs are distinct biological entities.

Oligoastrocytomas are classified as having cells of both oligodendrocytic and astrocytic histologies[13]. Oligoastrocytomas comprise ~3% of central nervous system gliomas diagnosed in the United States[11] but have also been diagnosed at frequencies

of up to 9.2% of intracranial gliomas[22] and 19% of supratentorial low-grade gliomas[23]. This apparent inconsistency suggests that the oligoastrocytomas are difficult to distinguish between other ADG histological types. As a group, oligoastrocytomas have an intermediate prognosis (~61% overall survival over 5 years) relative to ODGs and astrocytomas[11]. However, inconsistent diagnosis of oligoastrocytoma may limit the utility of this classification group in predicting clinical outcomes.

While the histological features of ADGs are generally able to predict clinical outcomes, they may mask the true biological heterogeneity that underlies this group of tumours. In addition, the lack of effective therapies for ADGs prompts us to better understand these tumours at the molecular level. In the recent years, genetic profiling of ADGs has proven useful in refining their classification into clinically and biologically meaningful groups, with the added benefit of providing insight into the molecular underpinnings of ADGs[18,24–29].

1.2 Genetics of Adult Diffuse Gliomas

It is widely accepted that cancer is a genetic disease. That is, changes in the genome and in the epigenome are able to cause gene expression changes that deregulate normal cellular pathways to promote uncontrolled cell growth[30]. Through next-generation sequencing technologies, we are now able to sequence the genome of a tumour within days and for only a few thousand dollars, with costs continuing to

decrease[31]. This cost decrease has increased the feasibility of whole-genome studies, allowing us to uncover the complex landscapes of genomic and epigenomic aberrations that underpin tumour development. Whole-genome and transcriptome profiling techniques are also allowing for the classification of tumours into molecular subtypes, which can often predict clinical outcomes and uncover molecular heterogeneity that we cannot detect using histology alone[6,25,27,29,32–35].

ADGs can now be broadly classified into 3 molecular subtypes according to a specific set of genetic alterations, summarized in Table 1. The different ADG molecular subtypes (IDH1/2-wild type, ATRX/TP53-mutated, and 1p/19q-codeleted) correlate well with different clinical outcomes, indicating they are biologically distinct groups[6,27,28]. The different subtypes also correlate well, albeit not perfectly, with the different ADG histologies, and especially highlight molecular heterogeneity within tumours classified as oligoastrocytomas and GBMs[6,27,28].

Notably, some of the genetic alterations that characterize the ADG molecular subtypes can be shared between subtypes[6,27,28], alluding to shared mechanisms in ADG development. In addition, some of these genetic alterations exhibit distinct patterns of anticorrelation or even mutual exclusivity between subtypes, suggesting redundant mechanisms in ADG development. Also, some of these genetic alterations often co-occur within the same tumour[6,27,28,36], indicating important interactions between

genetic alterations. In the following paragraphs, the nature of these genetic alterations and their possible contributions to promoting ADG development will be discussed.

Isocitrate dehydrogenase 1 or *2* (*IDH1/2*) mutations are common to both the ATRX/TP53-mutated and 1p/19q-codeleted subtypes. *IDH1* and *IDH2* encode for a cytosolic form and the mitochondrial form, respectively, of enzymes that normally produce alpha-ketoglutarate (αKG) in the citric acid cycle[37,38]. ADG-associated *IDH1/2* mutations are highly recurrent point mutations that substitute a single amino acid within the enzymes' catalytic binding cleft. These mutations change their enzymatic activity to instead catalyze the production of 2-hydroxyglutarate (2HG) with αKG as a substrate[39]. The resulting shift in metabolic equilibrium within the cell interferes with the normal functioning of DNA and histone methylation enzymes, resulting in widespread changes in the epigenome[40–42].

The observation that *IDH1/2* mutations are present in both the ATRX/TP53-mutated and 1p/19q-codeleted ADG subtypes[6,27,28] suggests that *IDH1/2* mutations are initiating mutations that arise early on in the tumourigenic process, with successive mutations further contributing to malignancy. Notably, *IDH1/2* mutations have also been observed in acute myeloid leukemia, chondrosarcoma, cholangiocarcinoma, and angioimmunoblastic T-cell lymphoma[37].

ATRX and *TP53* mutations characterize the ATRX/TP53-mutated subtype, which is characteristic of the majority of oligoastrocytomas, grade II and III astrocytomas, and secondary glioblastomas[6,27,28]. *TP53* is a well-studied tumour suppressor gene that is mutated in many cancer types and has been termed the "guardian of the genome[43]." Its gene product, p53, is a DNA-binding protein that can arrest cell cycle progression if DNA is damaged, initiate DNA repair, and initiate apoptosis if the damage is not repaired[44–46]. Meanwhile, *ATRX* encodes a chromatin remodeling protein whose inactivation is linked to a mechanism of maintaining telomere length termed alternative lengthening of telomeres or ALT[47–49]. Maintaining telomere length is an important hallmark of cancers because telomere shortening from successive replications eventually induces replicative senescence[50].

TERT promoter mutations are common to most IDH1/2-wild type and 1p/19q-codeleted ADGs[24,26–28,36]. These mutations can generate a motif within the *TERT* promoter that may be recognized by members of the E-twenty-six (ETS) family of transcription factors[36,51,52]. These mutations therefore likely activate *TERT* expression by allowing members of the ETS transcription factor family to bind to the *TERT* promoter and promote *TERT* transcription. *TERT* encodes the catalytic subunit of the telomerase enzyme, with a canonical function of maintaining telomere length[53,54]. Therefore, like *ATRX* mutations, *TERT* promoter mutations can constitute an important mechanism in a tumour's escape from replicative senescence induced by telomere shortening[54]. Notably, mutations in the *TERT* promoter are present in a number of cancer types that arise from tissues that typically do not self-renew[36].

Currently, one of the most routinely used molecular markers for the prognosis of ADGs is a deletion of the chromosomal arms 1p and 19q (1p/19q codeletion)[55]. 1p/19q codeletions correlate well with increased survival and chemosensitivity and characterize the 1p/19q-codeleted subtype, considered to be "classical" ODG[18]. However, until next-generation sequencing, specific gene targets of these chromosomal arm deletions had remained enigmatic. In recent years, several sequencing studies have unveiled *capicua transcriptional repressor* (*CIC*), located on 19q, and *far upstream element binding protein 1* (*FUBP1*), located on 1p, as being mutated in ~70% and 25-40% of 1p/19q-codeleted ADGs, respectively[3–5]. This implicates tumour suppressive roles for both *CIC* and *FUBP1* since one allele is lost from the 1p/19q codeletion while the other allele is often mutated in these ADGs.

Both *CIC* and *FUBP1* encode transcription factors for which several target genes have been identified. The FUBP1 protein is required for maximal expression of the well-known proto-oncogene *MYC,* whose gene product also functions as a transcription factor that can regulate hundreds of cellular genes[56]. FUBP1 can also bind several mRNA species, both viral and endogenous, to alter the translation or splicing of these mRNA species[57]. Meanwhile, the established target genes for the CIC protein are the 3 members of an oncogenic subfamily of ETS transcription factors *ETV1, ETV4,* and *ETV5* (*ETV1/4/5*)[10,58]. Overexpression of *ETV1/4/5* have been associated with melanoma, breast, and prostate cancers[59]. To our knowledge, however, no genome-

wide studies to comprehensively identify genes regulated by endogenous CIC and FUBP1 in human cells have been reported. How FUBP1 and CIC affect gene expression in human cells has therefore remained largely unexplored, making the roles of *CIC* and *FUBP1* mutations in ADG development unclear. Since *CIC* is more frequently mutated than *FUBP1* in ADG, the focus of this thesis is on CIC and the consequences of *CIC* mutations on gene expression in human cells.

1.3 CIC in Cell Signaling, Development, and Disease

The gene *CIC* encodes a transcription factor that transduces receptor tyrosine kinase (RTK) signaling into changes in gene expression to direct developmental processes such as differentiation and proliferation[60–62]. The functional domains of CIC in different metazoan species are highly conserved[63], implying evolutionary conservation of CIC's biochemical mechanisms. Consistent with this, CIC has invariably been observed in *Drosophila,* mice, and human cells to mediate RTK signaling through a mechanism of default repression[10,60,64–67]. That is, instead of directly activating target gene transcription upon the input of a RTK signal, CIC represses transcription until a RTK signal inhibits CIC's repressive activity[58,60,68]. This mechanism of default repression can facilitate the establishment of sharp boundaries in which genes are turned "off" or "on" in a spatial context in the developing *Drososphila* embryo[69,70]. Notably, Cic functions repeatedly throughout *Drosophila* development to regulate the expression of different target genes through a RTK-MAPK-Cic signaling axis[60,61,64,71].

The RTK-MAPK-CIC signaling axis is a conserved mechanism of gene regulation in human cells[58]. Using human melanoma and HEK293 cell line models, Kumara Dissanayake *et al.* described two biochemical mechanisms linking MAPK activation to the inhibition of CIC's transcriptionally repressive activity[58]. First, activated MAPK can promote an interaction between CIC and 14-3-3 chaperone proteins. This CIC-14-3-3 interaction inhibits CIC binding to an octameric DNA motif (5'-TGAATGAA-3'). CIC normally binds this motif at the promoters or enhancers of CIC target genes[10] to repress transcription when MAPK is not activated[68]. The second potential mechanism of CIC regulation by MAPK is through an interaction of CIC with the nuclear import protein importin α4/karyopherin α3. This interaction is inhibited by the activation of MAPK, thus potentially providing a partial mechanism for compartmentalizing CIC within the cytosol (and out of the nucleus) upon MAPK activation.

CIC has also been studied in several mammalian development and disease contexts. In a rare and aggressive subtype of Ewing family tumours, chimeric forms of *CIC* have been found to be fused to either *DUX4*, *DUX4L,* or *FOXO4*[10,72–74]. The tumour-associated CIC-DUX4 fusion protein retains CIC's known functional domains while acquiring a relatively small portion of the DUX4 C-terminal end[10]. This chimeric form of CIC activates *ETV1/4/5* transcription whereas the wild type form of CIC normally represses *ETV1/4/5* transcription[10,75]. Aberrant CIC function is also implicated in spinocerebellar ataxia type I (SCA1), in which an aberrant, polyglutamine-expanded

form of the ATAXIN-1 protein causes neurodegeneration[66,67]. Polyglutamine-expanded ATAXIN-1 modulates the transcriptionally repressive activity of CIC, and reduced CIC expression can mitigate the disease phenotypes of SCA1[66,67]. Finally, the transient expression of murine Cic in developing cerebellar granule neurons has also implicated CIC in neurogenesis[63].

There are two known main isoforms of CIC, the short (CIC-S) and long (CIC-L) form[3]. CIC-S and CIC-L differ in their N-terminal portions but share two highly conserved domains: a DNA-binding high mobility group (HMG) box domain and a C-terminal motif (C1) that is necessary for repression[60,63]. While the functional differences between these isoforms remain to be elucidated, our group has previously reported that CIC-L and CIC-S predominantly localize within the nucleus and cytoplasm, respectively, with the short isoform also in close proximity to the mitochondria[8].

In 1p/19q-codeleted ADGs, roughly half of *CIC* mutations are frameshifting, nonsense, and splice site mutations that are spread throughout the CIC protein (Figure 1)[3–6]. These mutations are predicted to delete at least a portion of CIC to confer loss-of-function. This pattern of mutations typically characterizes classic tumour suppressor genes such as *retinoblastoma 1*[76,77], in which inactivation of both copies of the gene promotes malignancy[76]. However, the other half of *CIC* mutations are missense mutations and single, in-frame amino acid deletions that localize within and around the HMG domain and C1 motif and can be recurrent (Figure 1)[3–6]. This pattern, in contrast,

is more reminiscent of gain-of-function mutations such as the aforementioned *IDH1/2* mutations, which keep the protein product intact but alter its function[37]. Overall, the mutational spectrum of *CIC* therefore confounds its role in ODGs since there seems to be selection for both mutations that disrupt the CIC protein structure as well as specific, single amino acid mutations that may fundamentally retain CIC's protein structure.

1.4 Thesis Investigation Overview

A mechanistic understanding of the role of *CIC* mutations in ODGs may shed insight into ODG development. This may lead to the identification of molecular targets for the effective treatment of ODGs. Therefore, the goal for this thesis was to study how *CIC* mutations can affect gene expression in an experimentally tractable human cell system.

We hypothesized that ODG-associated *CIC* mutations confer a loss of CIC's repressive activity and deregulate the expression of CIC target genes other than *ETV1/4/5*. To test this, we inactivated CIC in HEK293a cells and subsequently measured resulting gene expression changes using microarrays. From this, gene expression changes spanning entire chromosomes were detected. Additionally, 24 candidate CIC-regulated genes were identified in HEK293a cells that also have evidence of CIC-dependent regulation in 1p/19q-codeleted gliomas of The Cancer Genome Atlas (TCGA). Of these 24 genes, 5 genes (*CNTFR*, *DUSP6*, *GPR3*, *SHC3*,

and *SPRY4*) with reported functions in mitogen-activated protein kinase (MAPK) signaling and central nervous system (CNS) development were further validated to undergo CIC-dependent regulation in HeLa cells. Finally, investigating how different *CIC* mutations affect gene expression revealed that different types of ODG-associated *CIC* mutations either abrogated or potentially preserved CIC's transcriptionally repressive activity. These findings shed insight into possible roles for CIC in regulating gene expression at a chromosome-wide scale, MAPK signaling, CNS development, and ODG development.

Table 1. Molecular subtypes of adult diffuse glioma

Molecular Subtype	**Prognosis**	**Predominant Histologies**	Defining Genetic Alterations[6,26–28]			
			TERT **promoter mut**	***IDH1/2*** **mut**	***ATRX*** **mut** ***TP53*** **mut**	**1p/19q codeletion** ***CIC*** **mut** ***FUBP1*** **mut**
IDH1/2-wild type*	1 year	Primary GBM	P	NP	NP	NP
ATRX/TP53-mutated	5 years	Oligoastrocytoma Astrocytoma Secondary GBM	NP	P	P	NP
1p/19q-codeleted	8 years (chemo-sensitive)	Oligodendroglioma	P	P	NP	P

*IDH1/2-wild type tumours are also often characterized by alterations of *EGFR* and *CDKN2A*, as well as chromosome 7 gains and chromosome 10 deletions[28].

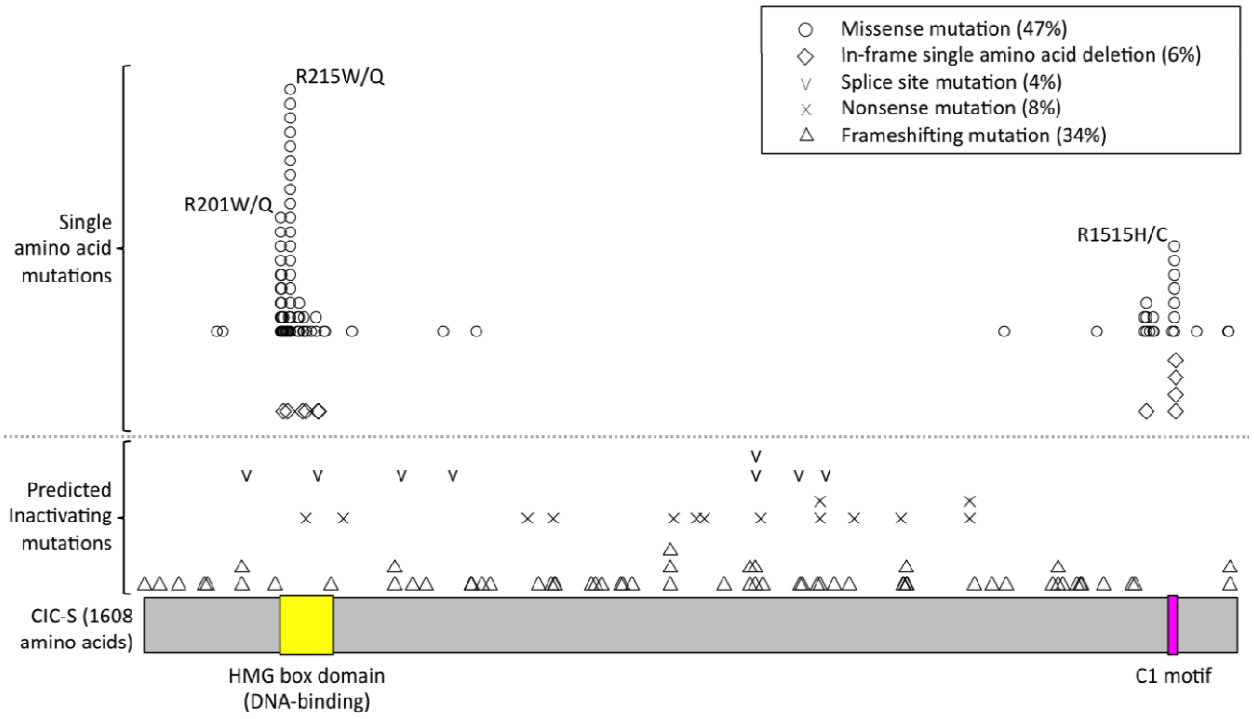


Figure 1. ODG-associated mutations mapped to the predicted CIC short isoform protein.

CIC-S: short isoform of CIC. HMG: high-mobility group. C1 motif: transcriptionally repressive c-terminal domain. Recurrently detected mutations are directly stacked. Frequencies of different mutation types are given in parentheses. Mutational data were gathered from references 1-7.

Chapter 2: Investigation of *CIC* Mutations on Gene Expression

2.1 Materials and Methods

2.1.1 Cell Culture and Conditions

All cells used in this study were incubated at 37℃ with 5% CO_2 in DMEM (Gibco) supplemented with 10% FBS (Gibco). Unless otherwise stated, at ~70-90% confluency, cells were washed with PBS, trypsinized, and either passaged or harvested by pelleting. Pellet storage was at -80℃. All cell lines and their derived clones were passaged from between 2 to 8 times from their initial storage in liquid nitrogen before transfection or harvesting.

2.1.2 Whole Cell Lysate Protein Extraction

Cells were thawed on ice and resuspended in 5X packed-cell volume of ice-cold EDTA-free RIPA lysis buffer (20mM Tris-HCl pH 7.5, 150mM NaCl, 0.1% NP-40, and 0.25% sodium deoxycholate freshly supplemented with 1mM sodium orthovanadate, 1mM NaF, and 1X EDTA-free protease inhibitor from Roche). Cell pellets were homogenized by passing 5-10 times through a 21-gauge needle then mixed for 20 minutes at 4ºC on an automatic rotator. Insoluble cellular debris was pelleted using centrifugation at 13000 x g for 10 minutes at 4ºC.

2.1.3 Western Blot Protein Detection

Extracted protein samples were subjected to gel-electrophoresis on NuPage 3-8% Tris Acetate pre-cast mini-gels (Invitrogen) with 1X MOPs buffer (Invitrogen) for 55 min at 150V. Separated proteins were transferred onto a methanol-activated PVDF membrane (Bio-Rad) for 60 minutes at 100V in 1X transfer buffer (Invitrogen) with 20% (v/v) methanol. Membranes were then incubated with anti-CIC (A301-204A, Bethyl Laboratories), anti-FLAG (F3165, Sigma), anti-tubulin (sc-9104, Santa Cruz), or anti-beta actin (ab8227, Abcam) at a 1:1000 dilution at 4^{o}C overnight. Membranes were then incubated with either goat anti-mouse HRP-IgG (Santa Cruz) or goat anti-rabbit IgG-HRP (dilution 1:5000, Santa Cruz) for 1 hour at room temperature, followed by three PBST washes before application of either ECL substrate (GE Healthcare or Bio-Rad) or SuperSignal West Femto substrate (Thermo Scientific). Images were captured using a LAS-4000 imager (FujiFilm) or ChemiDoc™ MP Imager (Bio-Rad).

2.1.4 Zinc Finger Nuclease-mediated CIC Inactivation

HEK293a (*cic*^{WT}) cells were co-transfected with 5 µL of custom-designed CIC specific CompoZr^{TM} custom zinc finger nuclease (ZFN) construct mRNA (Sigma-Aldrich) with a target site of GCCTCCAACCAGAGCaaaggtGAGGGCTGGTGGGGACTG (with the two ZFN-binding half sites capitalized), along with 250 ng of a Hygro RS reporter

construct (Toolgen) harbouring the same ZFN target site to enrich for cells in which ZFNs are active[78]. Transfection was performed by seeding 4 x 10^5 cells in a 6-well plate ~24 hours prior to transfection and performing lipid-based transfection using the Trans-IT-mRNA transfection reagent and Trans-IT Boost reagent (MIR 2225, Mirius Bio) according to Sigma's recommendations for CompoZr^{TM} custom ZFN transfection. Enrichment for cells with active ZFNs was carried out by treating cells with 0.5 mg/mL hygromycin at ~72 - 96 hours post-transfection. From the resulting enriched cell population, single clones were isolated using a limiting dilution method. These clones were then screened for reduced CIC expression using Western blots. A clone with reduced CIC expression (*cic*^{ZFN1}) was selected and brought through an additional round of transfection, enrichment, limiting dilution, and screening to obtain a clone with essentially absent CIC expression (*cic*^{ZFN2}).

2.1.5 siRNA-mediated Knockdown of CIC Expression

Cells were plated to ~80 % confluency and transfected with CIC-specific Stealth siRNA (HSS118258, Life Technologies) or nonspecific negative control siRNA (12935-200, Life Technologies). Transfection was carried out using Lipofectamine® RNAiMAX Transfection Reagent (13778030, Life Technologies) according to the manufacturer's recommendations. Cells were harvested at ~72 hours post-transfection.

2.1.6 mRNA Quantification by RT-qPCR

RNA extraction was performed using the RNeasy Plus Mini Kit (74136, Qiagen) according to the manufacturer's recommendations. RT-qPCR was carried out using 50 ng of template RNA with the *Power* SYBR® Green RNA-to-C_t™ *1-Step* Kit (4389986, Life Technologies) according to the manufacturer's recommended reaction component amounts and cycling conditions. A 7900 HT Sequence Detection System with SDS 2.2 software was used for temperature cycling and the generation of C_T values. Analysis of relative mRNA expression was performed using the $2\text{-}\Delta\Delta C_T$ method with *TATA-box binding protein* (*TBP*) expression as an endogenous control. Sequences of the RT-qPCR primers used are listed in Table 2.

2.1.7 Sequencing of the ZFN Target Locus

Genomic DNA extraction from cic^{WT}, cic^{ZFN1}, and cic^{ZFN2} cells was performed using the DNeasy Blood & Tissue Kit (69506, Qiagen) according to the manufacturer's recommendations. Amplicons to be sequenced were assigned unique molecular barcodes and adapted for MiSeq flow-cell NGS sequencing chemistry using PCR. The PCR step was performed on 200 ng of template genomic DNA with Jumpstart Taq Polymerase (D9307, Sigma), a forward primer with sequence5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCCTTAGTCCCCTTCCTGG-3', and a reverse primer with sequence 5'-

GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGGTTCTGGGGACACAGAGG-3'. Cycling conditions for the PCR amplification were (1) an initial denaturation at 95 ℃ for 1 min, (2) 35 cycles of denaturation at 95 ℃ for 30 sec, annealing at 58℃ for 30 sec, and extension at 72℃ for 1 min, and (3) a final extension at 72 ℃ for 1 min.

The barcoded amplicon libraries were pooled and purified using conventional preparative agarose gel electrophoresis. Library quality and quantitation was performed using a 2100 Bioanalyzer with DNA 1000 chips (Agilent Technologies) and a Qubit 2.0 Fluorometer (Life Technologies). High-throughput DNA sequencing was conducted using a MiSeq sequencer according to the manufacturer's recommendations (Illumina).

2.1.8 Fluorescence *in situ* Hybridization (FISH)

Fresh HEK293a cells were fixed onto slides using methanol and acetic acid in a 3 to 1 ratio. Slides were then aged in 2X SSC at 37 ℃ for 30 minutes followed by dehydration through an ethanol series. Vysis 1p36/1q25 and 19q13/19p13 FISH probes were applied to the cells and co-denatured at 73 ℃ for 5 minutes and hybridized for 18 hours at 37 ℃. Post-hybridization, cells were washed in 0.4X SSC/0.3% NP-40 for 2 minutes at 73 ℃ and air dried prior to adding DAPI counterstain.

2.1.9 Microarray Expression Profiling

Genome-wide mRNA expression profiling on 3 consecutive passages of *cic*^{WT}, *cic*^{ZFN1}, and *cic*^{ZFN2} cells was carried out using the GeneChip® Human Gene 2.0 ST array (Affymetrix) at The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada. The resulting chp files were analyzed using the RMA algorithm and differential gene expression analysis was carried out using the Transcriptome Analysis Console (Affymetrix) in which Tukey's bi-weight average is calculated for each group.

2.1.10 Flag-CIC Construct Generation

A tandem 3x FLAG sequence was PCR-amplified from plasmid pAFW1111 (*Drososphila* Genomics Resources Centre, Indiana, USA) using KAPA HiFi DNA polymerase (KAPA Biosystems). A forward primer with sequence 5'-ATCGAAGCTTACCATGGACTACAAAGACCATGACG-3' and reverse primer with sequence 5'-ATCGGGATCCCTTGTCATCGTCATCCTTGTA-3' were used generate tandem FLAG amplicons with *Hin*dIII and *Bam*HI restriction sites. A pcDNA™4/TO vector (Invitrogen) was modified by inserting the 3x FLAG sequence at *Hin*dIII/*Bam*HI sites to create N-terminus 3xFLAG expression vector (pcDNA™4/TO/N-FLAG).

A pcDNA5 FRT/TO GFP-CIC construct (wild type) was obtained from University of Dundee, Scotland[58] . Site-directed mutagenesis was performed to create mutant

constructs pcDNA5 FRT/TO GFP-CIC_R1515H and pcDNA5 FRT/TO GFP-CIC_R201W using the QuikChange II XL Kit (Agilent). Mutant constructs were sequence-verified. Using a forward primer with the sequence 5'-ATCGGAATTCAAGATCTATGTATTCGGCCCACAGGCCC-3' and reverse primer with the sequence 5'-AAGGCCAGGTCGTGGTACTTCT-3', and TaKaRa LA Taq (Clontech), the first 1496 bp of CIC-S (referred to as 5' CIC-S), including a *Bgl*II restriction site just before the start codon, was amplified from pcDNA5 FRT/TO GFP-CIC. The 5' CIC-S amplicon was sequence-verified and restriction digestion was performed using *Bgl*II (NEB).

A 4436 bp region of CIC from pcDNA5 FRT/TO GFP-CIC was obtained using *Stu*I (NEB) and *Not*I (NEB) enzymes (referred to as 3' CIC). This 3' CIC fragment was ligated to the 5' CIC-S amplicon using Quick T4 DNA ligase (NEB) to yield full length CIC-S. CIC-S was cloned into *Bam*HI/*Not*I sites of pcDNA™4/TO/N-FLAG to create the pcDNA4/TO/FLAG-CIC-S (referred to as Flag-CIC) plasmid construct. The pcDNA4/TO/FLAG-CIC_R1515H and pcDNA4/TO/FLAG-CIC_R201W constructs were created as described for FLAG-CIC-S.

2.1.11 Luciferase Assay

~24 hrs prior to transfection, 1.4 x 10^5 cells per transfection were seeded in 24-well plates. Transfection was carried out using the Turbofect transfection reagent

(R0531, Thermo Scientific) according to the manufacturer's recommendations. The following DNA amounts were transfected into each well: 0.6 µg of a pGL3 reporter vector with a cloned promoter sequence of *ETV5* (99 to 996 base pairs upstream of the transcription start site[10]), 40 ng of the indicated flag-CIC construct or corresponding flag vector-only control, and 0.1 µg of a pRL-CMV vector (E2261, Promega) as a transfection control. At ~72 hours post-transfection, luciferase expression was measured using the Dual Glo luciferase assay system (E2920, Promega) according to the manufacturer's recommendations and with a Perkin Elmer Wallac 1420 Victor2 microplate reader.

2.1.12 Statistical Analysis

2 sample t-test was used to compare the means of relative mRNA expression, as measured by RT-qPCR and microarrays, between the HEK293a clones. Wilcoxon rank sum test was used to compare transcript abundance of the different genes, as measured by RNA-seq and given in RSEM, between the different groups of 1p/19q-codeleted TCGA gliomas. 1-sided Fisher's exact test was used to assess the positive association between the number of candidate CIC-regulated genes and the number of measured genes within a chromosome for each chromosome. For CIC-specific siRNA-treated HeLa cells, 1-sample t-test was used to compare the mean of relative mRNA expression (normalized to nonspecific siRNA-treated conditions) to a mean of 1.0. Luciferase expression was normalized to expression in cic^{WT} cells transfected with the flag only vector control. 1 sample t-test was used to compare the mean of luciferase

expression in *cic*^{ZFN2} cells transfected with a vector-only control to a mean of 1.0. 2 sample t-test was used for all other comparisons between means of luciferase expression. The Benjamini-Hochberg procedure was used to calculate all false discovery rates to correct for multiple hypothesis testing.

2.2 Results

Zinc finger nuclease treatment of HEK293a cells produces isogenic clones with CIC loss-of-function phenotypes.

In order to investigate CIC's effects on gene expression in cells, CIC was first inactivated in HEK293a cells (*cic*^{WT}) with a zinc finger nuclease (ZFN), an engineered protein that can produce insertions and deletions at unique target sites within the genome. The ZFN used was specific for a unique genomic sequence within *CIC* that is common to both CIC isoforms and maps N-terminal to CIC's DNA-binding domain and repressive C-terminal motif (Figure 2a). This insures that frameshifting insertions and deletions introduced by the ZFN could abrogate CIC function.

After treating *cic*^{WT} cells with the CIC-specific ZFN, a clone was isolated (*cic*^{ZFN1}) with reduced CIC-L and CIC-S expression, as detected using Western blots (Figure 2b). *cic*^{ZFN1} cells were subjected to another round of ZFN treatment and monoclonal isolation from which an additional clone (*cic*^{ZFN2}) with essentially undetectable levels of CIC-L

and CIC-S was obtained (Figure 2b). The specificity of the antibody used in the Western blot detection of CIC was confirmed using siRNA-mediated knockdown of CIC (Figure 2b). Finally, mRNA expression of CIC's known target genes *ETV1/4/5*, as measured by RT-qPCR, all increased by ~2-fold in cic^{ZFN1} cells and by at least 10-fold in cic^{ZFN2} cells relative to cic^{WT} cells (Figure 2c).

In cic^{ZFN1} and cic^{ZFN2} cells, mutations were likely introduced that abrogate the expression of functional CIC. Furthermore, the observed *ETV1/4/5* expression levels in cic^{ZFN1} and especially in cic^{ZFN2} cells are consistent with a CIC loss-of-function phenotype that would be predicted from the literature. The data therefore indicate the utility of this CIC loss-of-function system in characterizing CIC function and identifying additional CIC target genes.

Frameshifting insertions and deletions were observed in all detected *CIC* alleles of cic^{ZFN1} and cic^{ZFN2} cells.

The locus containing the *CIC*-specific ZFN target site was then amplified and sequenced in cic^{WT}, cic^{ZFN1}, and cic^{ZFN2} cells to identify the nature of the ZFN-induced frameshifting mutations. As expected, all detected alleles of cic^{ZFN2} cells harboured frameshifting mutations while cic^{WT} cells had no detected frameshifting mutations (Figure 3a and 3b). The ratio of the 3 mutant alleles were also detected at roughly a

2:1:2 ratio (Figure 3b), which is consistent with the observation that the HEK293a cells used in this study have up to 5 CIC loci (Figure 3d). Finally, all 3 frameshifting mutations were predicted to add either three or four out-of-frame codons following the mutation before introducing a stop codon to truncate the CIC polypeptide (Figure 3c).

Surprisingly, all sequenced *CIC* alleles of *cic*^{ZFN1} cells also harboured the same frameshifting insertions as *cic*^{ZFN2} cells (Figure 3a and 3b), despite *cic*^{ZFN1} cells having detectable CIC-L and CIC-S protein expression (Figure 2b). Since the observed frameshifting insertions mutations lie close to the intron-exon boundary (Figure 3a), I speculate that a noncanonical splice site upstream of the frameshifting insertions maybe preferentially used in *cic*^{ZFN1} cells relative to *cic*^{ZFN2} cells. This hypothetical splice site, when used instead of the canonical exon 2 splice site in the processing of CIC mRNA, can result in the splicing out the ZFN-induced mutations as intronic sequence. This type of splicing could therefore generate functional transcript in the *cic*^{ZFN1} cells. However, this phenomenon has not yet been investigated further.

Identification of candidate CIC-regulated genes in HEK293a cells and 1p/19q-codeleted ADGs.

To identify genes potentially regulated by CIC, microarray expression profiling was performed to compare gene expression at genome-wide levels between *cic*^{WT} and *cic*^{ZFN2} cells. Consistent with their expression patterns observed from RT-qPCR, *ETV1/4/5* were observed to have increased expression in *cic*^{ZFN2} relative to *cic*^{WT} cells within a calculated false discovery rate (FDR) value of 0.1 or 10% (Figure 3a). For this reason, the 598 genes with differential expression between *cic*^{WT} and *cic*^{ZFN2} cells within a FDR threshold of 0.1 were considered candidate CIC-regulated genes in HEK293a cells (Figure 4a).

24 of the 598 candidate CIC-regulated genes in HEK293a cells also had differential expression (FDR < 0.1) in a data set obtained and derived from the Cancer Genome Atlas (TCGA)[1,2,9]. This data set compared gene expression, as measured by RNA-sequencing, between 21 CIC-inactivated and 43 CIC-wild type 1p/19q-codeleted ADGs (Figure 4b and 4c, and Table 3). These 24 genes included *ETV1/4/5* and were therefore considered to be strong candidates for CIC-regulated genes in both HEK293a cells and 1p/19q-codeleted ADGs.

Detection of possible CIC-mediated chromosome-wide changes in gene expression.

Unexpectedly, a significant enrichment of candidate CIC-regulated genes in HEK293a cells was observed in chromosomes 2, 6, 19, and 20 (FDR < 0.05, Figures 5a and 5b). Most if not all the candidate CIC-regulated genes in these chromosomes span either several Giemsa bands (chromosomes 2 and 6) or the entire length of the chromosome (chromosomes 19 and 20) (Figure 5a). These results suggest that CIC inactivation can mediate changes in gene expression at a chromosome-wide scale.

Assuming that CIC only functions to directly repress transcription, the widespread decreased expression of genes within chromosomes 2, 19, and 20 upon CIC inactivation is suggestive of an indirect mechanism of regulation by CIC. Furthermore, the vast majority of candidate CIC-regulated genes in these chromosomes exhibited either increased expression (chromosome 6) or decreased expression (chromosomes 2, 19, and 20) upon CIC inactivation (Figure 5b). I therefore speculate that CIC may mediate large-scale changes in chromosomal architecture to either limit or enhance the accessibility of genes to transcriptional machinery. Consistent with this speculation, we note several genes implicated in chromatin remodeling through the alteration of DNA methylation or nucleosome occupancy (histone cluster 1, *HAT1, SMARCB1*, *SMARCA4*, *MBD3*, *UHRF1*, and *DNMT1*) were among the candidate CIC-regulated genes in HEK293a cells (Appendix 1 and 2).

Verification and Validation of CIC-regulated genes.

Of the 24 consistently detected candidate CIC-regulated genes, *CNTFR*, *DUSP6*, *GPR3*, *SHC3*, and *SPRY4,* were among the genes with the highest expression in cic^{ZFN2} relative to cic^{WT} cells. These 5 genes also had increased expression in CIC-inactivated relative to CIC-wild type 1p/19q-codeleted ADGs (Table 3). Interestingly, all 5 genes also have reported roles in neuronal cell development and MAPK signaling regulation (Table 3), possibly indicating that CIC regulates these processes by regulating the expression of these 5 genes. Consistent with their observed expression changes in the microarray data, the expression of *DUSP6*, *GPR3*, *SHC3*, and *SPRY4* was verified using RT-qPCR to be increased in cic^{ZFN2} relative to cic^{WT} cells ($p < 0.01$, Figure 6). The same trend was also apparent for *CNTFR* ($p = 0.11$, Figure 6). For all 5 genes, there was also a trend of increased expression in cic^{ZFN1} cells relative to cic^{WT} cells (Figure 6).

To further assess CIC-dependent regulation of *CNTFR*, *DUSP6*, *GPR3*, *SHC3*, and *SPRY4*, expression of these genes was quantified using RT-qPCR upon siRNA-mediated knockdown of CIC in HeLa cells (Figure 7). As expected, CIC-S and CIC-L protein expression decreased (Figure 7a) and *ETV1/4/5* mRNA expression increased (Figure 7b) in these cells upon treatment with CIC-specific siRNA. There was also a significant increase in mRNA expression of *DUSP6* and *SHC3* ($p < 0.05$) and a trend of increased expression of *SPRY4* ($p = 0.13$), *GPR3* ($p = 0.06$), and *CNTFR* ($p = 0.15$)

upon treatment with CIC-specific siRNA (Figure 7c). These results, taken together with all the heretofore presented evidence, strongly implicate that *CNTFR*, *DUSP6*, *GPR3*, *SHC3*, and *SPRY4* are CIC-regulated genes.

Different *CIC* mutation types confer different effects on the expression of CIC-regulated genes.

ODG-associated *CIC* mutations not only consist of mutations predicted to disrupt the CIC protein (e.g. frameshifting mutations), but also single amino acid mutations targeting CIC's HMG box domain and C1 motif that may preserve CIC's protein structure. We therefore investigated whether different *CIC* mutation types confer different effects on the expression of CIC-regulated genes by grouping 82 TCGA 1p/19q-codeleted ADGs into 4 groups according to their *CIC* status (Figure 8a): (1) those that harbour no detectable *CIC* mutations, (2) frameshifting, nonsense or splice site mutations (CIC inact), (3) HMG box-targeting single amino acid mutations (CIC HMG), and (4) C1 motif-targeting single amino acid mutations (CIC C1). mRNA expression of the 8 CIC-regulated genes (*ETV1/4/5*, *CNTFR*, *DUSP6*, *GPR3*, *SHC3*, and *SPRY4*), as measured by RNA-sequencing, was then compared between these 4 groups. From this analysis, it was observed that the CIC inact group generally had the highest expression of all 8 genes (Figure 8c). For *ETV4*, *ETV5*, *GPR3*, and *DUSP6*, however, expression was significantly lower in the CIC C1 group relative to the CIC inact group. Additionally, For *ETV4* and *GPR3,* expression was also significantly lower

in the CIC HMG group relative to CIC inact group. These results indicate ODG-associated *CIC* mutations targeting the HMG box domain and C1 motif may retain CIC's transcriptionally repressive activity on some genes.

To further assess the extent to which types of *CIC* mutations could repress target gene expression, we reintroduced wild type and representative mutant forms of CIC-S into cic^{ZFN2} cells (Figure 9a). At the same time, a reporter construct harbouring a luciferase cassette under the control of the *ETV5* promoter (Figure 9b) was also introduced into these cells. In this system, luciferase expression hence served as an indicator of the repressive activity of different forms of CIC on the *ETV5* promoter. From this assay, we observed that Q564X CIC, a truncating nonsense mutation, had no significant repressive activity ($p = 0.48$, Figure 9c). Meanwhile, wild type CIC, R201W CIC (a recurrent HMG box domain missense mutation), and R1515H CIC (a recurrent C1 motif missense mutation) could repress luciferase expression ($p < 0.05$).

2.3 Discussion

Summary and Comparison to other Studies

In investigating the effects of CIC on gene expression, it was observed that CIC inactivation possibly mediates chromosome-wide gene expression changes. This investigation has also shown that the expression of *CNTFR, DUSP6, GPR3, SPRY4,* and *SHC3* can be deregulated by *CIC* inactivation in multiple cell contexts. Finally, we observe that ODG-associated single amino acid mutations targeting CIC's highly conserved domains have the potential to preserve at least some of CIC's repressive activity while CIC-inactivating mutations such as the Q564X mutation can completely abrogate CIC's repressive function.

To my knowledge, there are two other studies in which CIC-regulated genes were identified in mammalian cells. One was in investigating how expression of the oncogenic CIC-DUX4 fusion protein affects gene expression[10]. *LBH*, a consistently identified candidate CIC-regulated gene in this study, was also among the genes reported to be activated by CIC-DUX4. In a study of the neurodegenerative disease SCA1 (in which aberrant CIC function is implicated), 16 genes, including another consistently identified candidate CIC-regulated gene in this study (*DUSP4*), were reported as targets of Cic repression in mice[67].

Collectively, there are now 10 genes with reported evidence of CIC-dependent regulation in at least two mammalian cell contexts (*ETV1/4/5*, *CNTFR, DUSP6, GPR3, SHC3, SPRY4*, *LBH*, and *DUSP4*). Consistent with CIC's known function as a transcriptional repressor, all 10 genes were consistently observed to have increased expression in CIC-inactivated relative to CIC-wild type conditions in this study. However, we also observed that other candidate CIC-regulated genes can have decreased expression upon CIC inactivation, indicating that CIC also directly or indirectly positively regulates gene expression.

Limitations of Study

One potential limitation and simultaneous strength of this investigation is that the overall impact of CIC inactivation on gene expression was measured regardless of whether or not these gene expression changes were directly or indirectly due to a loss of CIC's repressive activity. For example, while the identified CIC-regulated gene *DUSP6* increased in expression upon CIC inactivation, this change could be due to either *DUSP6* being a direct target of CIC-mediated repression or to *DUSP6* being positively regulated by other CIC targets ETV1/4/5, as has been reported in zebrafish[79]. To better understand the transcriptional network of CIC, the DNA elements in the genome that CIC binds can potentially be identified in future work using chromatin precipitation followed by next-generation sequencing (ChIP-seq). In ChIP-seq, CIC and CIC-bound DNA regions can be isolated using a CIC-specific antibody. The isolated

CIC-bound DNA regions would then be sequenced and mapped to the genome to identify possible genes directly regulated by CIC. This would shed insight into how CIC directly or indirectly influences the transcription of CIC-regulated genes.

We observed that CIC may mediate chromosome-wide gene expression changes. However, there is an important caveat to this argument. This caveat arises from the observation that CIC was nearly undetectable in *cic*^{ZFN2} cells while consistently detectable in *cic*^{ZFN1} cells, despite both these clones harbouring the exact same frameshifting *CIC* mutations at similar frequencies. This suggests that between *cic*^{ZFN2} and *cic*^{ZFN1} cells, there is change in the regulation of CIC, perhaps at the epigenetic or mRNA splicing level. This change in CIC regulation could also be causing the chromosome-wide gene expression changes observed between *cic*^{ZFN2} and *cic*^{WT} cells. In future work, we can address whether the chromosome-wide gene expression changes are CIC-mediated by ectopically reintroducing CIC into *cic*^{ZFN2} cells. If this ectopic reintroduction of CIC reverses the chromosome-wide gene expression changes observed between *cic*^{ZFN2} *and cic*^{WT} cells, then these changes would be considered CIC-mediated.

Possible Tumourigenic Roles of CIC

Tumours such as ADGs develop from the deregulation of normal cellular growth and proliferation pathways[50]. Examples of such pathways include RTK signaling pathways, which interpret growth factor signaling into a wide variety of cellular responses[80,81]. Since CIC is a downstream, negatively-regulated component of RTK signaling[58,62], deregulation of CIC target genes may recapitulate some effects of increased RTK signaling. The identification of novel CIC-regulated genes may therefore provide us with insight into how *CIC* mutations can contribute to ADG development.

Interestingly, the 5 identified CIC-regulated genes in this study, along with the consistently identified CIC-regulated gene *DUSP4*, have all been reported to be involved in regulating MAPK signaling. These genes either promote (*CNTFR*[82], *GPR3*[83], and *SHC3*[84]) or inhibit (*DUSP4*[85,86], *DUSP6*[87] and *SPRY4*[88]) MAPK activation. Notably, CIC itself is negatively regulated by MAPK through RTK[58]. Taken together, these observations suggest that CIC may be positioned in regulatory feedback loops within the MAPK signaling network. CIC may therefore normally function to keep MAPK-mediated RTK signaling from becoming deregulated.

CNTFR, DUSP6, GPR3, SHC3, and *SPRY4* have also been reported to either function in neuronal cell development and/or have brain specific expression[82,84,89–92].

This indicates CIC may be important in the proliferation or differentiation of cells of the CNS. Consistent with this, several reports implicate CIC in CNS development. First, the murine Cic protein has been found to be predominantly expressed in the brain, with expression in the hippocampus and olfactory bulb and transient expression in the developing cerebellum[63,93]. Second, *Drosophila* Cic directly represses the transcription factor *intermediate neuroblasts defective*, a gene essential in the development of part of the *Drosophila* CNS[94]. Third, CIC is implicated in glioma and spinocerebellar ataxia type I, two pathologies of the CNS[3,4,67]. Finally, CIC is negatively regulated by EGFR signaling[58], an essential signaling pathway in regulating the proliferation and differentiation of neural stem cells[95–97]. CIC may therefore normally function to somehow regulate the proliferation or differentiation of cells along the neural stem cell lineage to contribute to CNS development. An aberrant regulation of this process through *CIC* mutations may contribute to the transformation of normal cells to ODG.

This investigation may also provide insight into the biology of CIC as a cancer-associated gene. Cancer-associated genes can typically be classified as either tumour suppressive (i.e. loss of activity drives malignancy) or oncogenic (i.e. increased activity drives malignancy)[30,77]. For *CIC* in ODGs, the loss of one allele through the 1p/19q codeletion and mutation of the other allele in 70% of cases[3,4,6] suggest a tumour suppressive role. However, we observed that different *CIC* mutations likely preserve some of CIC's transcriptionally repressive activity on some genes while other mutations likely do not. This observation, taken together with the observation that about half of CIC mutations are single amino acid mutations that target CIC's HMG box domain and

C1 motif[3–6], indicates selective advantage for CIC mutations that can preserve repressive activity in some ODGs.

Possible Synergistic Interactions of *CIC* Mutations

CIC mutations may not only act individually but also synergistically with other co-occurring mutations in ODG. One possible interaction of CIC mutations is with *IDH1/2* mutations, which cause widespread gene expression changes through the epigenetic alteration of chromatin architecture[41,42]. In this study, it was observed that CIC inactivation possibly mediates chromosome-wide gene expression changes, possibly by altering chromatin architecture as well. It would therefore be unsurprising if *IDH1/2* mutations and *CIC* mutations, when co-occurring in the same cell, had synergistic effects in altering gene expression through the alteration of chromatin architecture.

Another possible synergistic interaction of *CIC* mutations is with *TERT* promoter mutations. As mentioned earlier, *TERT* promoter mutations likely activate *TERT* transcription by allowing members of the ETS family of transcription factors to bind to the *TERT* promoter[51,52]. In a *TERT* promoter-mutated context, *CIC* mutations may potentiate *TERT* transcription by derepressing the expression of *ETV1/4/5,* members of the ETS family of transcription factors[10,58].

TERT promoter mutations are also common in *IDH1/2*-wild type ADGs[28,36]. In addition to *TERT* promoter mutations, RTK signaling is also commonly upregulated in this ADG subtype through aberrations such as *EGFR* amplifications[6,98] and the overexpression of RTK genes[99]. This upregulation of RTK signaling may in turn inactivate CIC's transcriptionally repressive activity to derepress *ETV1/4/5* expression. Therefore, CIC may also play a key role in a synergistic interaction between upregulated RTK signaling and *TERT* promoter mutations to potentiate *TERT* transcription. In support of this, it has been observed that 92% of *EGFR* amplifications have been observed to co-occur with *TERT* promoter mutations in ADG[98].

Conclusion

The CIC loss-of-function system generated in this study has revealed aspects of CIC's biology that seem to be conserved in ODGs. Future work with this system may therefore provide further insight into the mechanisms of ODG-associated *CIC* mutations. In the future, comprehensive knowledge of the mechanisms of *CIC* and other ODG-associated mutations may give us an unprecedented understanding of the molecular biology of ODGs. Powered with this understanding, we may one day be able to rationalize effective treatments for ODGs.

Table 1. **RT-qPCR primers used.**

Gene	**Forward Primer (5' to 3')**	**Reverse Primer (5' to 3')**
ETV1	TGTCCTCCTCGTTGATGTGACG	TGGGGCATTCAGAAAAACAGG
ETV4	CAGGCGGAGGTTGAAGAAAGG	AAGCGCAGAAGAAAGGCAAAGG
ETV5	AGGGAAATCTCGATCTGAGGAATG	GCTAACCAAGCCTCTTGAAGTTGAC
CNTFR	GAGGAGGAGGAGGACATTGA	AGGAGCAGCCATCTCTTCAC
GPR3	CTCCACGGTTCCAGAATGT (Figure 6) TGAGCGGTACCATGATGTG (Figure 7	GGGAGAAGGCTCTGGTTTCT (Figure 6) GATGATGGCCACCACTAGC (Figure 7)
DUSP6	GCAACAGACTCGGATGGTAG	TGCGTTCTCAAAGAGATTCG
SHC3	ATTACCAGGGAAGCCATCAG	CGTGGAGATGGTCAGAGAGA
SPRY4	GTGACCAGGATGTCACCCAC	GACCACCTTGGGCTGGATG
TBP	CAGCTCTTCCACTCACAGACT	GTGCAATGGTCTTTAGGTCAA

Table 2. Consistently identified candidate CIC-regulated genes

Gene Symbol	*cic*^{ZFN2} vs. *cic*^{WT} HEK293a cells			CIC-inactivated vs. CIC-wild type 1p/19q-codeleted ADGs		Description
	Rank by fold change	Log_2 fold change	FDR q-value	Log_2 fold change	FDR q-value	
ETV4	1	0.644	0.0621	0.882	1.28E-05	ets variant 4
SPRY4*	2	0.517	0.0427	0.217	0.0073	sprouty homolog 4 (*Drosophila*)
ETV5	3	0.504	0.0641	0.245	0.0011	ets variant 5
GPR3*	4	0.400	0.0427	0.320	0.0138	G protein-coupled receptor 3
ETV1	5	0.358	0.0524	0.171	0.0462	ets variant 1
DUSP6*	6	0.355	0.0631	0.151	0.0030	dual specificity phosphatase 6
SHC3*	7	0.208	0.0666	0.276	0.0194	SHC (Src homology 2 domain containing) transforming protein 3
DUSP4	8	0.183	0.0852	0.543	0.0011	dual specificity phosphatase 4
CNTFR*	9	0.145	0.0441	0.060	0.0820	ciliary neurotrophic factor receptor
BACH2	10	0.122	0.0822	0.093	0.0073	BTB and CNC homology 1, basic leucine zipper transcription factor 2
LBH	11	0.089	0.0727	0.129	0.0227	limb bud and heart development homolog (mouse)
HSF2BP	12	0.082	0.0793	0.223	0.0021	heat shock transcription factor 2 binding protein
TJAP1	13	0.044	0.0648	-0.068	0.0226	tight junction associated protein 1 (peripheral)
PTPN9	14	0.043	0.0842	0.055	0.0837	protein tyrosine phosphatase, non-receptor type 9

Asterisk: genes involved in MAPK regulation and neuronal cell development or have brain-specific expression. Red and green values: evidence of increased expression and decreased expression, respectively, when CIC is inactivated.

	cic^{ZFN2} vs. cic^{WT} HEK293a cells			CIC-inactivated vs. CIC-wild type 1p/19q-codeleted ADGs		
Gene Symbol	**Rank by fold change**	**Log_2 fold change**	**FDR q-value**	**Log_2 fold change**	**FDR q-value**	**Description**
PAQR4	15	-0.045	0.0228	-0.070	0.0826	progestin and adipoQ receptor family member IV
UHRF1	16	-0.061	0.0841	0.128	0.0495	ubiquitin-like with PHD and ring finger domains 1
MRI1	17	-0.068	0.0822	-0.150	0.0436	methylthioribose-1-phosphate isomerase homolog (S. cerevisiae)
CIC	18	-0.071	0.0375	-0.066	0.0073	capicua homolog (*Drosophila*)
GNG7	19	-0.076	0.0470	-0.064	0.0842	guanine nucleotide binding protein (G protein), gamma 7
SIX1	20	-0.083	0.0736	0.406	0.0490	SIX homeobox 1
SLC35F1	21	-0.197	0.0641	0.117	0.0073	solute carrier family 35, member F1
GPRIN3	22	-0.219	0.0736	0.232	0.0366	GPRIN family member 3
GALC	23	-0.297	0.0899	-0.143	0.0366	galactosylceramidase
ITGA4	24	-0.309	0.0441	0.231	0.0523	integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)

Asterisk: genes involved in MAPK regulation and neuronal cell development or have brain-specific expression. Red and green values: evidence of increased expression and decreased expression, respectively, when CIC is inactivated.

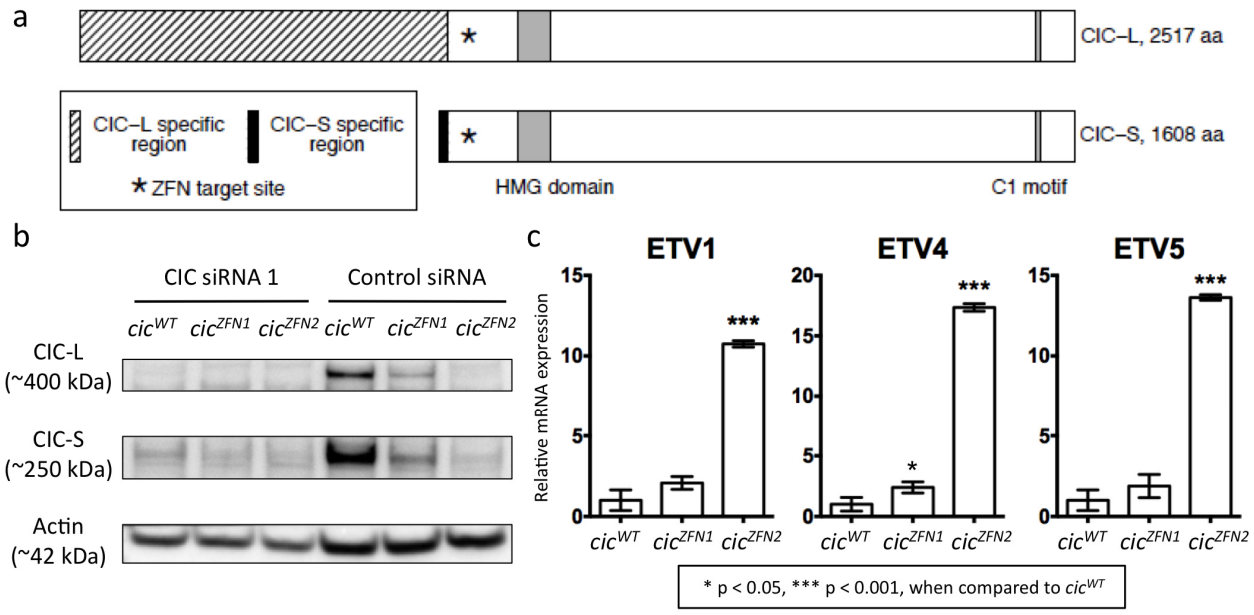
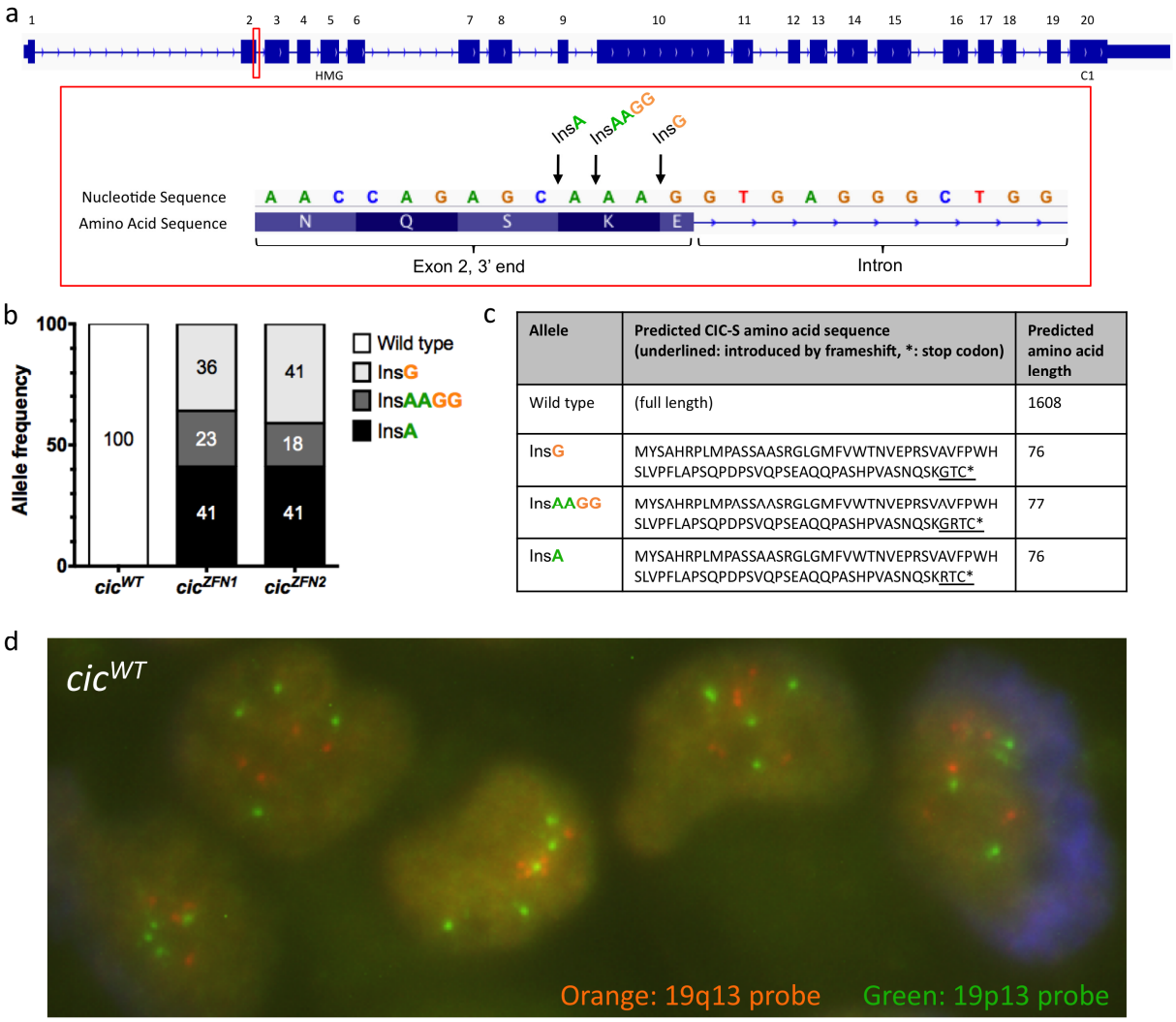


Figure 2. Zinc finger nuclease-mediated mutation of endogenous *CIC* produces a model for hypomorphic CIC function.

(**a**) Protein structure of long (CIC-L) and short (CIC-S) CIC isoforms. CIC domains highly conserved between metazoan species are shaded in gray. ZFN: zinc finger nuclease. aa: amino acids. HMG: DNA-binding high mobility group box domain. C1: repressive C-terminal motif. (**b**) Western blot detection of CIC, with actin as an endogenous control, in parental (cic^{WT}) and ZFN-treated (cic^{ZFN1} and cic^{ZFN2}) HEK293a clones. Cells were treated with either CIC-specific siRNA or nonspecific siRNA to confirm antibody specificity. (**c**) Relative mRNA expression, measured by RT-qPCR, of known CIC target genes *ETV1/4/5* in HEK293a clones specified in (b). Error bars indicate standard deviations over 3 separate passages.



Allele	Predicted CIC-S amino acid sequence (underlined: introduced by frameshift, *: stop codon)	Predicted amino acid length
Wild type	(full length)	1608
InsG	MYSAHRPLMPASSAASRGLGMFVWTNVEPRSVAVFPWH SLVPFLAPSQPDPSVQPSEAQQPASHPVASNQSKGTC*	76
InsAAGG	MYSAHRPLMPASSAASRGLGMFVWTNVEPRSVAVFPWH SLVPFLAPSQPDPSVQPSEAQQPASHPVASNQSKGRTC*	77
InsA	MYSAHRPLMPASSAASRGLGMFVWTNVEPRSVAVFPWH SLVPFLAPSQPDPSVQPSEAQQPASHPVASNQSKRTC*	76

Figure 3. All sequenced *CIC* alleles of *cic*^{ZFN1} and *cic*^{ZFN2} cells harbour frameshifting insertions.

(**a**) Upper panel: exon structure of CIC short isoform. Exons that encode the highly conserved DNA-binding high mobility group (HMG) box domain (exon 5) and repressive C1 motif (exon 20) are identified. The red box indicates the zoomed-in region in lower panel. Lower panel: detected zinc finger nuclease (ZFN)-mediated insertions (Ins) in alleles of *cic*^{ZFN1} and *cic*^{ZFN2} clones. (**b**) Frequency of wild type and variant alleles indicated in (a) in the HEK293a clones. Bar labels indicate the number of observed reads from a representative 100 mapped reads. (**c**) Predicted CIC short isoform (CIC-S) amino acid sequences from the observed alleles in (b). (**d**) Fluorescence *in situ* hybridization (FISH) detection of chromosome 19 loci in parental (*cic*^{WT}) HEK293a clone to determine *CIC* copy number. Out of 100 counted cells, 4 or 5 probes per cell were typically counted for 19q13, where *CIC* is located.

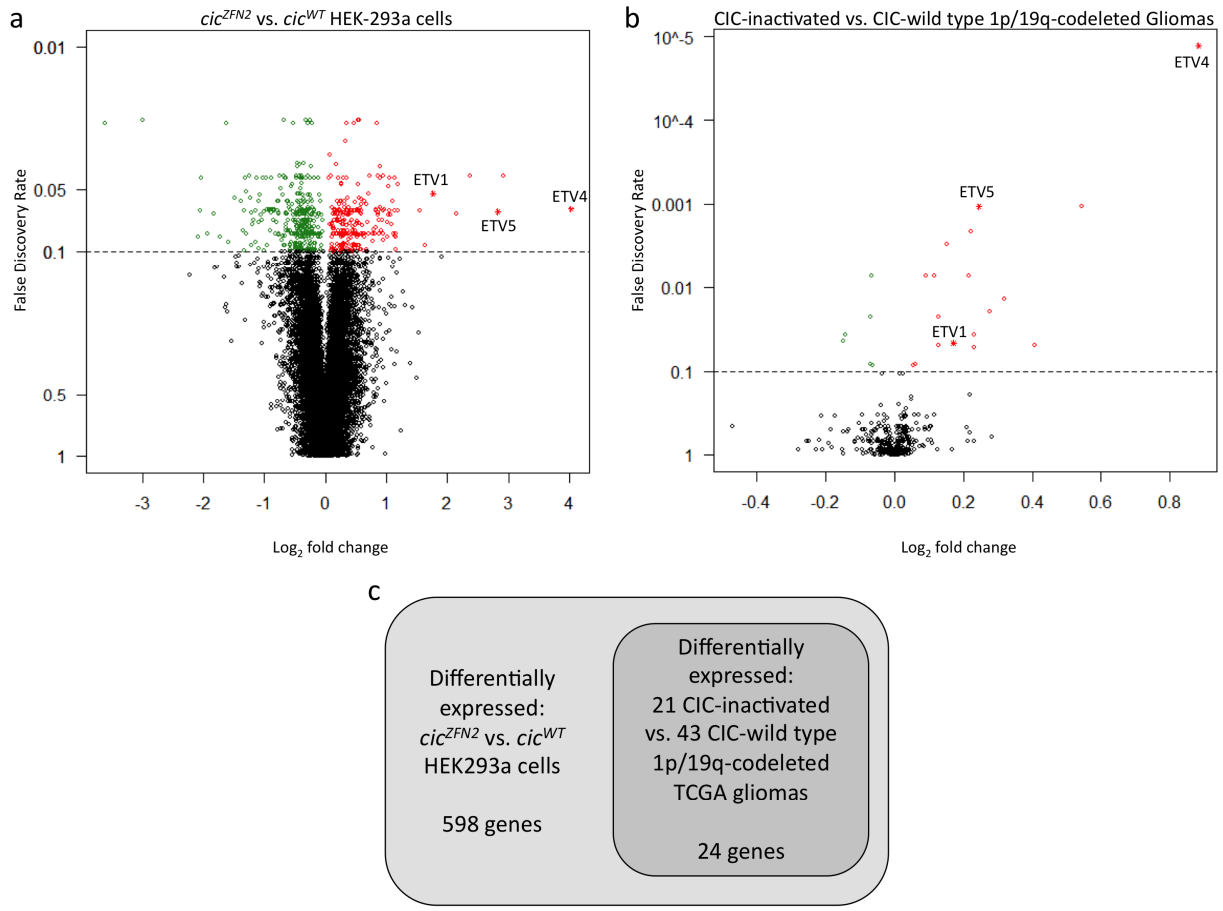


Figure 4. Microarray expression analysis of *cic*^{WT} and *cic*^{ZFN2} cells identifies candidate CIC-regulated genes.

(**a**) Comparison of gene expression between *cic*^{ZFN2} and *cic*^{WT} cells. Red and green data points indicate the 598 candidate CIC-regulated genes with increased and decreased expression, respectively, upon CIC inactivation. (**b**) Comparison of gene expression changes of the candidate CIC-regulated genes between CIC-inactivated and CIC-wild type TCGA 1p/19q-codeleted gliomas. Red and green data points indicate the 24 consistently observed CIC-regulated genes with increased and decreased expression, respectively, upon CIC inactivation in 1p/19q-codeleted gliomas and HEK293a cells(**c**) Schematic for the identification of the candidate CIC-regulated genes in HEK293a cells and in 1p/19q-codeleted ADGs. The nature of the CIC-inactivating mutations is summarized in Figure 8b.

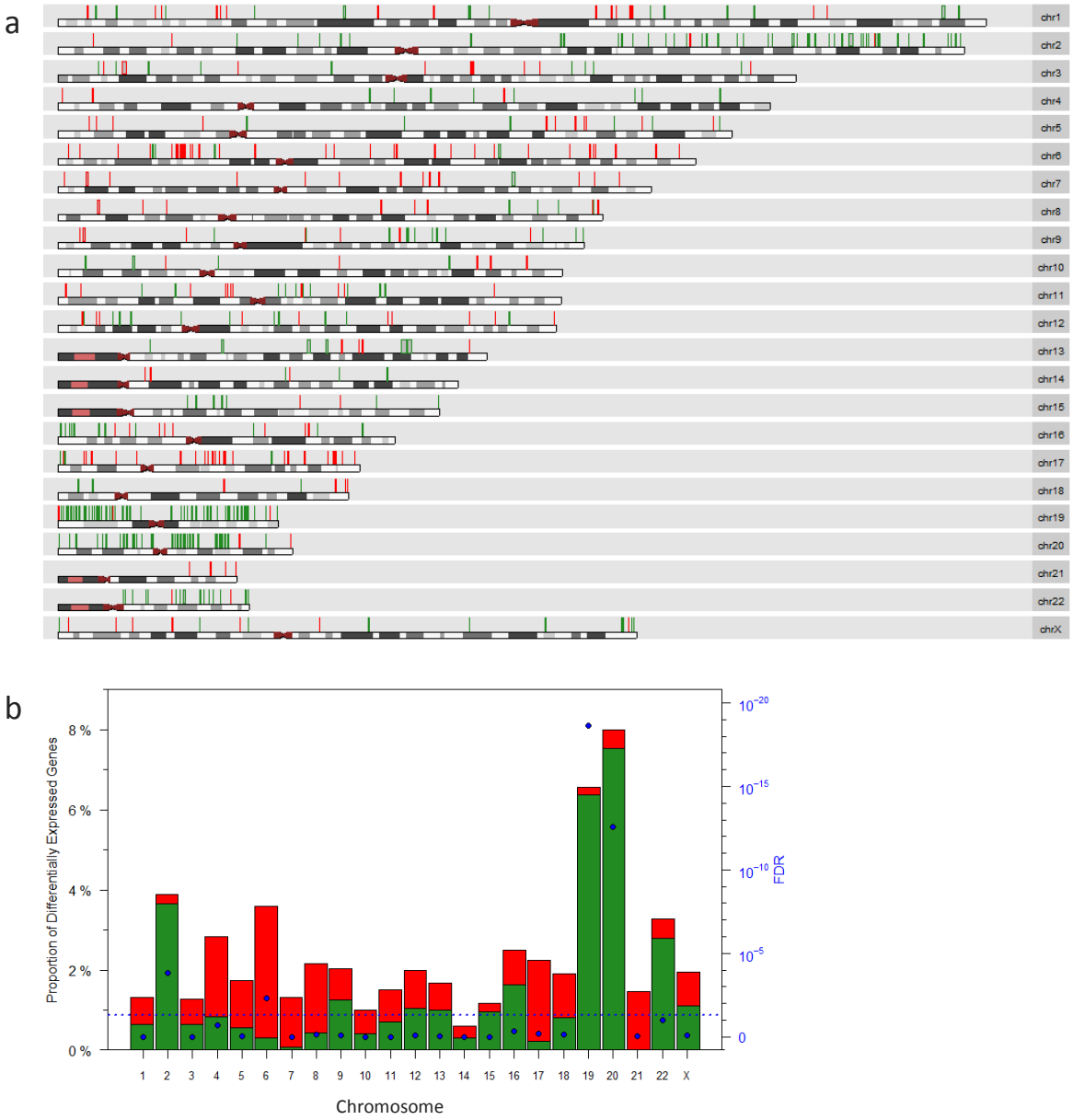


Figure 5. CIC inactivation possibly mediates chromosome-wide gene expression changes in HEK293a cells.

(**a**) Genomic location of candidate CIC-regulated genes in HEK293a cells (see Figure 4a). Green and red lines or rectangles: genes with decreased and increased expression, respectively, upon CIC inactivation. (**b**) Red and green bars: proportion of differentially expressed genes (given along left y-axis) with increased and decreased expression, respectively, relative to the total number of measured genes in each chromosome. Blue points: False discovery rate (FDR)-corrected p values (given along right y-axis) for the positive association between the number of differentially expressed genes and the number of measured genes within each chromosome. Dashed line: FDR threshold of 0.05.

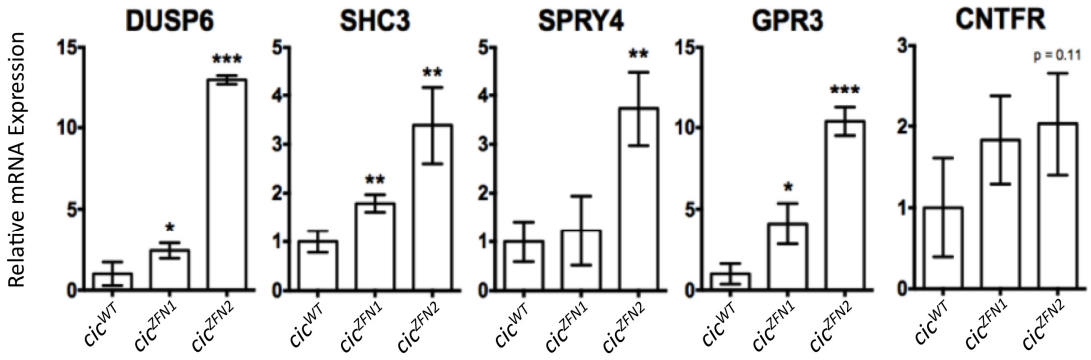


Figure 6. RT-qPCR verifies gene expression changes detected from microarrays.

Relative mRNA expression (y-axes), measured using RT-qPCR, of indicated potential CIC target genes in parental (*cic*^{WT}) and ZFN-treated (*cic*^{ZFN1} and *cic*^{ZFN2}) HEK293a clones. Error bars indicate standard deviations over 3 separate passages. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ when compared to *cic*^{WT} cells.

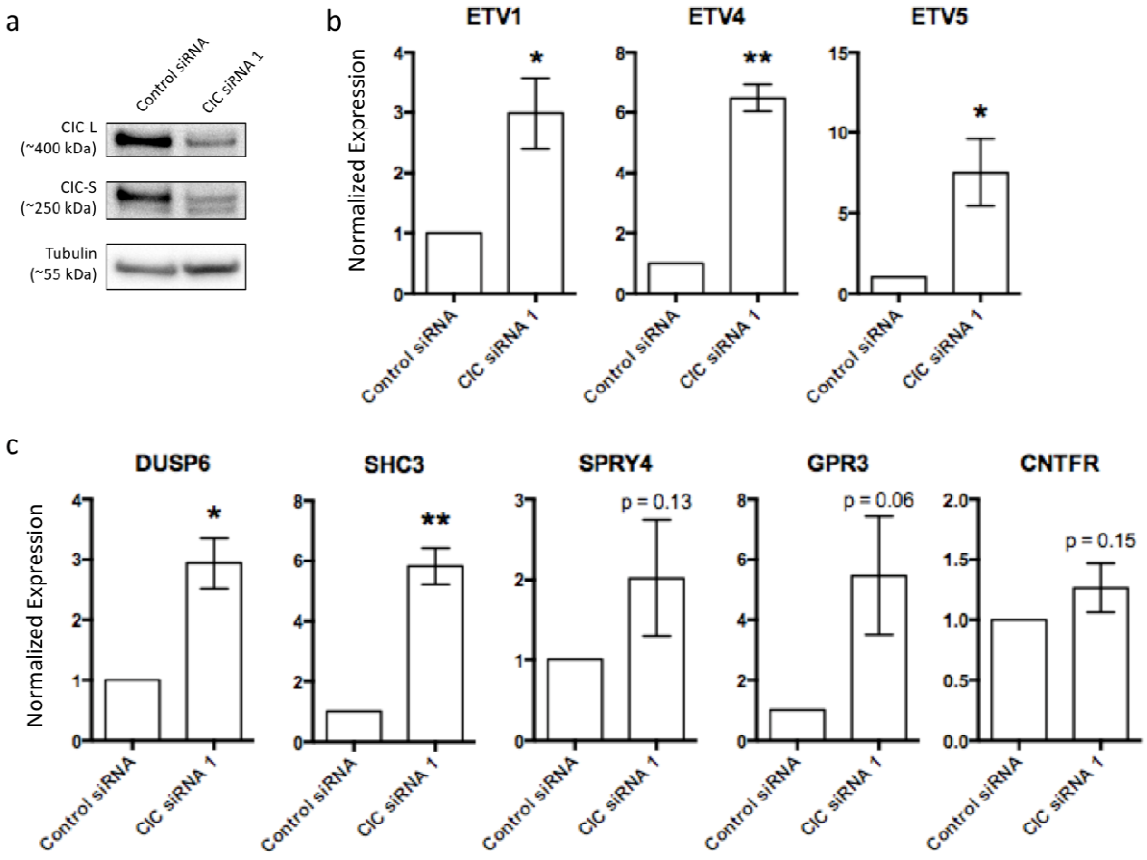


Figure 7. siRNA-mediated CIC knockdown validates genes as CIC-regulated.

(**a**) Representative Western blot expression of long (CIC-L) and short (CIC-S) CIC isoforms, with tubulin as an endogenous control, in HeLa cells treated with CIC-specific (CIC siRNA 1) or nonspecific siRNA (control siRNA). (**b**) and (**c**) Relative mRNA expression, measured using RT-qPCR and normalized to control siRNA treatment, of known CIC-regulated genes (b) and novel CIC-regulated genes (c) in CIC siRNA 1-treated HeLa cells. Error bars: standard deviations between 3 biological replicates. *: $p < 0.05$, **: $p < 0.01$.

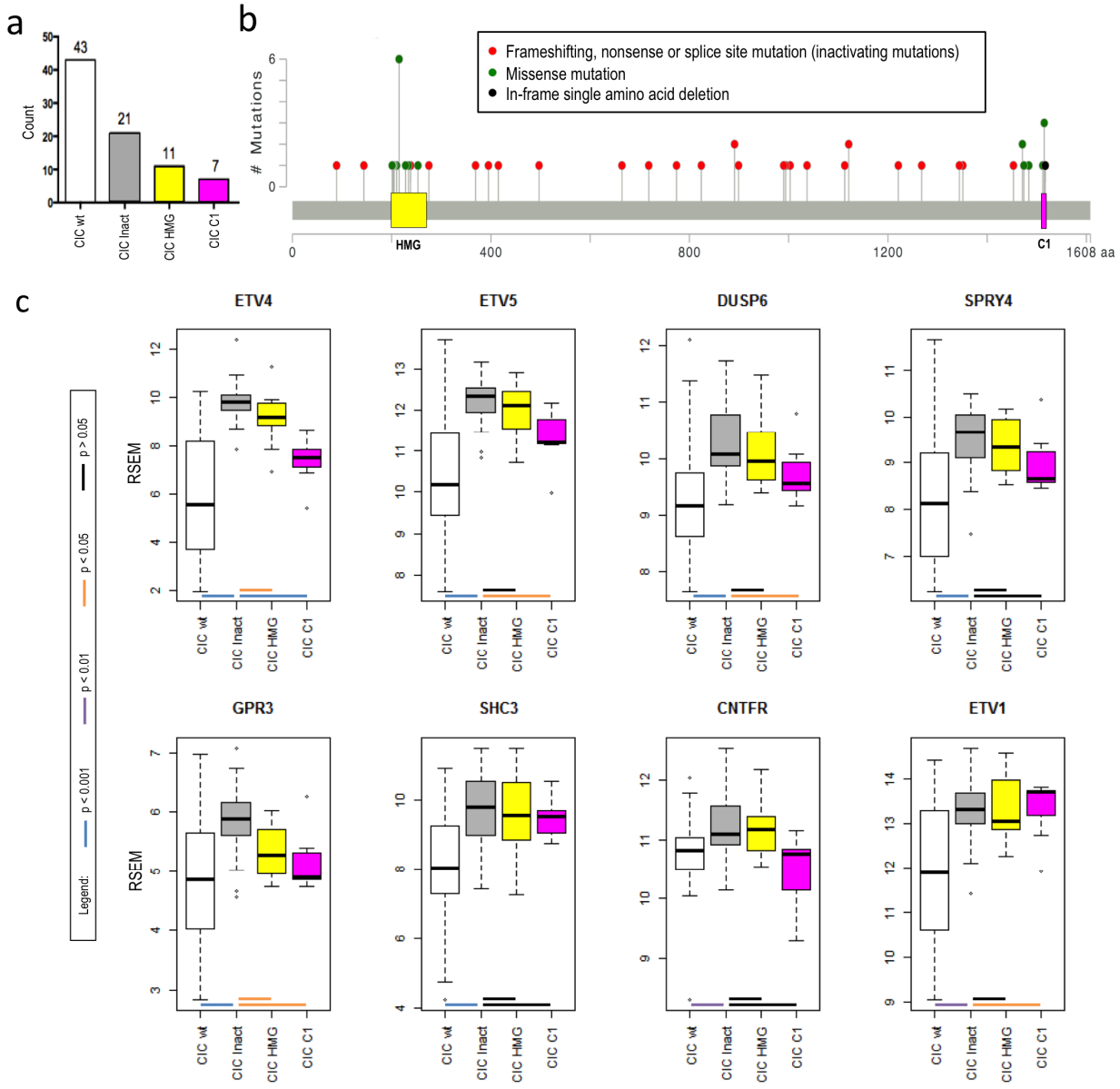


Figure 8. CIC-regulated genes exhibit CIC mutation type-specific expression.

(**a**) Frequency of 1p/19q-codeleted TCGA LGGs with CIC wild type status (CIC wt), predicted inactivating mutations (CIC inact) and single amino acid mutations targeting CIC's DNA-binding HMG domain (CIC HMG) and C1 motif (CIC C1) as shown in (b). (**b**) Observed *CIC* mutations in 39 *CIC*-mutated 1p/19q-codeleted TCGA gliomas. (**c**) Distributions of transcript abundance, measured using RNA-seq and given in RSEM (y-axes), of CIC-regulated genes in the different *CIC* mutation groups shown in (a). The horizontal lines above the x-axis labels correspond to comparisons made between the different *CIC* mutation groups, with the line colours indicating statistical significance of the differences between these groups.

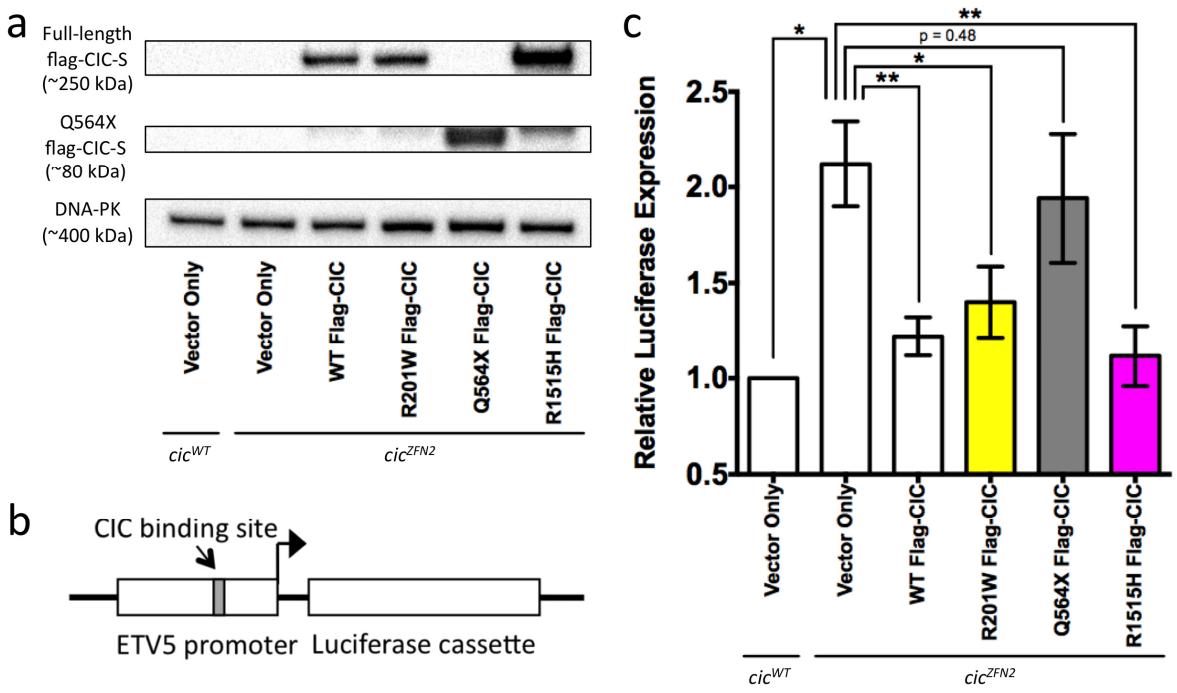


Figure 9. Representative missense mutations preserve CIC's repressive activity while a CIC truncation does not.

(**a**) Representative Western blot detection of ectopically introduced CIC, with DNA-PK as an endogenous control, in *cic*^{WT} and *cic*^{ZFN2} cells transfected with the indicated flag-tagged forms of CIC or flag vector-only controls. (**b**) Diagram of relevant portion of luciferase reporter vector used in all conditions of the luciferase assay. (**c**) Normalized luciferase expression in the conditions indicated in (a). Error bars indicate standard deviations over 3 biological replicates. *: $p < 0.05$, **: $p < 0.01$.

References

1. Cerami, E. *et al.* The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov.* **2,** 401–404 (2012).

2. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6,** pl1 (2013).

3. Yip, S. *et al.* Concurrent CIC mutations, IDH mutations, and 1p/19q loss distinguish oligodendrogliomas from other cancers. *J. Pathol.* **226,** 7–16 (2012).

4. Bettegowda, C. *et al.* Mutations in CIC and FUBP1 contribute to human oligodendroglioma. *Science* **333,** 1453–5 (2011).

5. Sahm, F. *et al.* CIC and FUBP1 mutations in oligodendrogliomas, oligoastrocytomas and astrocytomas. *Acta Neuropathol.* **123,** 853–60 (2012).

6. Jiao, Y. *et al.* Frequent ATRX, CIC, FUBP1 and IDH1 mutations refine the classification of malignant gliomas. *Oncotarget* **3,** 709–22 (2012).

7. Chan, A. K.-Y. *et al.* Loss of CIC and FUBP1 expressions are potential markers of shorter time to recurrence in oligodendroglial tumors. *Mod. Pathol.* **27,** 332–42 (2014).

8. Chittaranjan, S. *et al.* Mutations in CIC and IDH1 cooperatively regulate 2-hydroxyglutarate levels and cell clonogenicity. *Oncotarget* **5,** 7960–79 (2014).

9. Cline, M. S. *et al.* Exploring TCGA Pan-Cancer data at the UCSC Cancer Genomics Browser. *Sci. Rep.* **3,** 2652 (2013).

10. Kawamura-Saito, M. *et al.* Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Hum. Mol. Genet.* **15,** 2125–37 (2006).

11. Ostrom, Q. T. *et al.* CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006-2010. *Neuro. Oncol.* **15 Suppl 2,** ii1–56 (2013).

12. Claes, A., Idema, A. J. & Wesseling, P. Diffuse glioma growth: a guerilla war. *Acta Neuropathol.* **114,** 443–58 (2007).

13. Louis, D. N. *WHO classification of tumours of the central nervous system*. (World Health Organization (WHO), 2007).

14. Wu, W. *et al.* Joint NCCTG and NABTC prognostic factors analysis for high-grade recurrent glioma. *Neuro. Oncol.* **12,** 164–72 (2010).

15. Kufe, D. W. *et al.* Holland-Frei Cancer Medicine. (2003).

16. Ohgaki, H. & Kleihues, P. Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas. *J. Neuropathol. Exp. Neurol.* **64,** 479–89 (2005).

17. Barres, B. A. The mystery and magic of glia: a perspective on their roles in health and disease. *Neuron* **60,** 430–40 (2008).

18. Cairncross, J. G. *et al.* Benefit from procarbazine, lomustine, and vincristine in oligodendroglial tumors is associated with mutation of IDH. *J. Clin. Oncol.* **32,** 783–90 (2014).

19. Van den Bent, M. J. *et al.* Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study 26951. *J. Clin. Oncol.* **31,** 344–50 (2013).

20. Ohno, M. *et al.* Histopathological malignant progression of grade II and III gliomas correlated with IDH1/2 mutation status. *Brain Tumor Pathol.* **29,** 183–91 (2012).

21. Thakkar, J. P. *et al.* Epidemiologic and Molecular Prognostic Review of Glioblastoma. *Cancer Epidemiol. Biomarkers Prev.* (2014). doi:10.1158/1055-9965.EPI-14-0275

22. Helseth, A., Langmark, F. & Mørk, S. J. Neoplasms of the central nervous system in Norway. I. Quality control of the registration in the Norwegian Cancer Registry. *APMIS* **96,** 1002–8 (1988).

23. Jaskólsky, D., Zawirski, M., Papierz, W. & Kotwica, Z. Mixed gliomas. Their clinical course and results of surgery. *Zentralbl. Neurochir.* **48,** 120–3 (1987).

24. Nonoguchi, N. *et al.* TERT promoter mutations in primary and secondary glioblastomas. *Acta Neuropathol.* **126,** 931–7 (2013).

25. Appin, C. L. & Brat, D. J. Molecular genetics of gliomas. *Cancer J.* **20,** 66–72 (2014).

26. Chan, A. K.-Y. *et al.* TERT promoter mutations contribute to subset prognostication of lower-grade gliomas. *Mod. Pathol.* (2014). doi:10.1038/modpathol.2014.94

27. Killela, P. J. *et al.* Mutations in IDH1, IDH2, and in the TERT promoter define clinically distinct subgroups of adult malignant gliomas. *Oncotarget* **5,** 1515–25 (2014).

28. Labussière, M. *et al.* TERT promoter mutations in gliomas, genetic associations and clinico-pathological correlations. *Br. J. Cancer* **111,** 2024–32 (2014).

29. Goodenberger, M. L. & Jenkins, R. B. Genetics of adult glioma. *Cancer Genet.* **205,** 613–21 (2012).

30. Weinberg, R. A. *The Biology of Cancer, Volume 2*. 796 (Garland Science, 2007). at <http://books.google.com/books?id=cr5rAAAAMAAJ&pgis=1>

31. Hayden, E. C. Technology: The $1,000 genome. *Nature* **507,** 294–5 (2014).

32. Northcott, P. A. *et al.* Medulloblastoma comprises four distinct molecular variants. *J. Clin. Oncol.* **29,** 1408–14 (2011).

33. Schmadeka, R., Harmon, B. E. & Singh, M. Triple-negative breast carcinoma: current and emerging concepts. *Am. J. Clin. Pathol.* **141,** 462–77 (2014).

34. Mack, S. C. *et al.* Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. *Nature* **506,** 445–50 (2014).

35. Schroeder, K. M., Hoeman, C. M. & Becher, O. J. Children are not just little adults: recent advances in understanding of diffuse intrinsic pontine glioma biology. *Pediatr. Res.* **75,** 205–9 (2014).

36. Killela, P. J. *et al.* TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 6021–6 (2013).

37. Cairns, R. A. & Mak, T. W. Oncogenic isocitrate dehydrogenase mutations: mechanisms, models, and clinical opportunities. *Cancer Discov.* **3,** 730–41 (2013).

38. Zhang, C., Moore, L. M., Li, X., Yung, W. K. A. & Zhang, W. IDH1/2 mutations target a key hallmark of cancer by deregulating cellular metabolism in glioma. *Neuro. Oncol.* **15,** 1114–26 (2013).

39. Gross, S. *et al.* Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. *J. Exp. Med.* **207,** 339–44 (2010).

40. Noushmehr, H. *et al.* Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* **17,** 510–22 (2010).

41. Duncan, C. G. *et al.* A heterozygous IDH1R132H/WT mutation induces genome-wide alterations in DNA methylation. *Genome Res.* **22,** 2339–55 (2012).

42. Turcan, S. *et al.* IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483,** 479–83 (2012).

43. Lane, D. P. Cancer. p53, guardian of the genome. *Nature* **358,** 15–6 (1992).

44. Guimaraes, D. P. & Hainaut, P. TP53: a key gene in human cancer. *Biochimie* **84,** 83–93 (2002).

45. Muller, P. A. J. & Vousden, K. H. Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer Cell* **25,** 304–17 (2014).

46. Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **88,** 323–31 (1997).

47. Heaphy, C. M. *et al.* Altered telomeres in tumors with ATRX and DAXX mutations. *Science* **333,** 425 (2011).

48. Lovejoy, C. A. *et al.* Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS Genet.* **8,** e1002772 (2012).

49. Schwartzentruber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482,** 226–31 (2012).

50. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144,** 646–674 (2011).

51. Huang, F. W. *et al.* Highly recurrent TERT promoter mutations in human melanoma. *Science* **339,** 957–9 (2013).

52. Horn, S. *et al.* TERT Promoter Mutations in Familial and Sporadic Melanoma. *Science (80-.).* **339,** 959–961 (2013).

53. Kim, N. W. *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* **266,** 2011–5 (1994).

54. Shay, J. W. & Bacchetti, S. A survey of telomerase activity in human cancer. *Eur. J. Cancer* **33,** 787–91 (1997).

55. Ma, R., de Pennington, N., Hofer, M., Blesing, C. & Stacey, R. Diagnostic and prognostic markers in gliomas - an update. *Br. J. Neurosurg.* **27,** 311–5 (2013).

56. Duncan, R. *et al.* A sequence-specific, single-strand binding protein activates the far upstream element of c-myc and defines a new DNA-binding motif. *Genes Dev.* **8,** 465–80 (1994).

57. Zhang, J. & Chen, Q. M. Far upstream element binding protein 1: a commander of transcription, translation and beyond. *Oncogene* **32,** 2907–16 (2013).

58. Dissanayake, K. *et al.* ERK/p90(RSK)/14-3-3 signalling has an impact on expression of PEA3 Ets transcription factors via the transcriptional repressor capicúa. *Biochem. J.* **433,** 515–25 (2011).

59. Oh, S., Shin, S. & Janknecht, R. ETV1, 4 and 5: an oncogenic subfamily of ETS transcription factors. *Biochim. Biophys. Acta* **1826,** 1–12 (2012).

60. Jimenez, G., Guichet, A., Ephrussi, A. & Casanova, J. Relief of gene repression by Torso RTK signaling: role of capicua in Drosophila terminal and dorsoventral patterning. *Genes & Dev.* **14,** 224–231 (2000).

61. Tseng, A.-S. K. *et al.* Capicua regulates cell proliferation downstream of the receptor tyrosine kinase/ras signaling pathway. *Curr. Biol.* **17,** 728–33 (2007).

62. Jiménez, G., Shvartsman, S. Y. & Paroush, Z. The Capicua repressor--a general sensor of RTK signaling in development and disease. *J. Cell Sci.* **125,** 1383–91 (2012).

63. Lee, C.-J. *et al.* CIC, a member of a novel subfamily of the HMG-box superfamily, is transiently expressed in developing granule neurons. *Brain Res. Mol. Brain Res.* **106,** 151–6 (2002).

64. Goff, D. J., Nilson, L. A. & Morisato, D. Establishment of dorsal-ventral polarity of the Drosophila egg requires capicua action in ovarian follicle cells. *Development* **128,** 4553–4562 (2001).

65. Lee, Y. *et al.* ATXN1 protein family and CIC regulate extracellular matrix remodeling and lung alveolarization. *Dev. Cell* **21,** 746–57 (2011).

66. Fryer, J. D. *et al.* Exercise and genetic rescue of SCA1 via the transcriptional repressor Capicua. *Science* **334,** 690–3 (2011).

67. Lam, Y. C. *et al.* ATAXIN-1 interacts with the repressor Capicua in its native complex to cause SCA1 neuropathology. *Cell* **127,** 1335–47 (2006).

68. Ajuria, L. *et al.* Capicua DNA-binding sites are general response elements for RTK signaling in Drosophila. *Development* **138,** 915–24 (2011).

69. Kim, Y. *et al.* Context-dependent transcriptional interpretation of mitogen activated protein kinase signaling in the Drosophila embryo. *Chaos* **23,** 025105 (2013).

70. Chen, H., Xu, Z., Mei, C., Yu, D. & Small, S. A system of repressor gradients spatially organizes the boundaries of Bicoid-dependent target genes. *Cell* **149,** 618–29 (2012).

71. Roch, F., Jiménez, G. & Casanova, J. EGFR signalling inhibits Capicua-dependent repression during specification of Drosophila wing veins. *Development* **129,** 993–1002 (2002).

72. Mariño-Enríquez, A. & Fletcher, C. D. M. Round cell sarcomas-biologically important refinements in subclassification. *Int. J. Biochem. Cell Biol.* (2014). doi:10.1016/j.biocel.2014.04.022

73. Italiano, A. *et al.* High prevalence of CIC fusion with double-homeobox (DUX4) transcription factors in EWSR1-negative undifferentiated small blue round cell sarcomas. *Genes. Chromosomes Cancer* **51,** 207–18 (2012).

74. Sugita, S. *et al.* A Novel CIC-FOXO4 Gene Fusion in Undifferentiated Small Round Cell Sarcoma: A Genetically Distinct Variant of Ewing-like Sarcoma. *Am. J. Surg. Pathol.* (2014). doi:10.1097/PAS.0000000000000286

75. Specht, K. *et al.* Distinct transcriptional signature and immunoprofile of CIC-DUX4 fusion-positive round cell tumors compared to EWSR1-rearranged ewing sarcomas: further evidence toward distinct pathologic entities. *Genes. Chromosomes Cancer* **53,** 622–33 (2014).

76. Richter, S. *et al.* Sensitive and efficient detection of RB1 gene mutations enhances care for families with retinoblastoma. *Am. J. Hum. Genet.* **72,** 253–69 (2003).

77. Liu, H., Xing, Y., Yang, S. & Tian, D. Remarkable difference of somatic mutation patterns between oncogenes and tumor suppressor genes. *Oncol. Rep.* **26,** 1539–1546 (2011).

78. Kim, H. *et al.* Magnetic separation and antibiotics selection enable enrichment of cells with ZFN/TALEN-induced mutations. *PLoS One* **8,** e56476 (2013).

79. Znosko, W. A. *et al.* Overlapping functions of Pea3 ETS transcription factors in FGF signaling during zebrafish development. *Dev. Biol.* **342,** 11–25 (2010).

80. Choura, M. & Rebaï, A. Receptor tyrosine kinases: from biology to pathology. *J. Recept. Signal Transduct. Res.* **31,** 387–94 (2011).

81. Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141,** 1117–34 (2010).

82. Bhattacharya, S., Das, A. V, Mallya, K. B. & Ahmad, I. Ciliary neurotrophic factor-mediated signaling regulates neuronal versus glial differentiation of retinal stem cells/progenitors by concentration-dependent recruitment of mitogen-activated protein kinase and Janus kinase-signal transducer and activator of. *Stem Cells* **26,** 2611–24 (2008).

83. Tanaka, S. *et al.* Developmental expression of GPR3 in rodent cerebellar granule neurons is associated with cell survival and protects neurons from various apoptotic stimuli. *Neurobiol. Dis.* (2014). doi:10.1016/j.nbd.2014.04.007

84. Nakamura, T. *et al.* N-Shc: a neural-specific adapter molecule that mediates signaling from neurotrophin/Trk to Ras/MAPK pathway. *Oncogene* **13,** 1111–21 (1996).

85. Balko, J. M. *et al.* Activation of MAPK pathways due to DUSP4 loss promotes cancer stem cell-like phenotypes in basal-like breast cancer. *Cancer Res.* **73,** 6346–58 (2013).

86. Chu, Y., Solski, P. A., Khosravi-Far, R., Der, C. J. & Kelly, K. The mitogen-activated protein kinase phosphatases PAC1, MKP-1, and MKP-2 have unique substrate specificities and reduced activity in vivo toward the ERK2 sevenmaker mutation. *J. Biol. Chem.* **271,** 6497–501 (1996).

87. Caunt, C. J. & Keyse, S. M. Dual-specificity MAP kinase phosphatases (MKPs): shaping the outcome of MAP kinase signalling. *FEBS J.* **280,** 489–504 (2013).

88. Sasaki, A. *et al.* Mammalian Sprouty4 suppresses Ras-independent ERK activation by binding to Raf1. *Nat. Cell Biol.* **5,** 427–32 (2003).

89. Muda, M. *et al.* MKP-3, a novel cytosolic protein-tyrosine phosphatase that exemplifies a new class of mitogen-activated protein kinase phosphatase. *J. Biol. Chem.* **271,** 4319–26 (1996).

90. Zhang, B.-L., Gao, D.-S. & Xu, Y.-X. [G protein-coupled receptor 3: a key factor in the regulation of the nervous system and follicle development]. *Yi Chuan* **35,** 578–86 (2013).

91. Yu, T., Yaguchi, Y., Echevarria, D., Martinez, S. & Basson, M. A. Sprouty genes prevent excessive FGF signalling in multiple cell types throughout development of the cerebellum. *Development* **138,** 2957–68 (2011).

92. Shimazaki, T., Shingo, T. & Weiss, S. The ciliary neurotrophic factor/leukemia inhibitory factor/gp130 receptor complex operates in the maintenance of mammalian forebrain neural stem cells. *J. Neurosci.* **21,** 7642–53 (2001).

93. Chen, T. *et al.* Characterization of Bbx, a member of a novel subfamily of the HMG-box superfamily together with Cic. *Dev. Genes Evol.* (2014). doi:10.1007/s00427-014-0476-x

94. Lim, B. *et al.* Kinetics of gene derepression by ERK signaling. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 10330–5 (2013).

95. Aguirre, A., Rubio, M. E. & Gallo, V. Notch and EGFR pathway interaction regulates neural stem cell number and self-renewal. *Nature* **467,** 323–7 (2010).

96. Xian, C. J. & Zhou, X.-F. EGF family of growth factors: essential roles and functional redundancy in the nerve system. *Front. Biosci.* **9,** 85–92 (2004).

97. Wong, R. W. C. & Guillaud, L. The role of epidermal growth factor and its receptors in mammalian CNS. *Cytokine Growth Factor Rev.* **15,** 147–56

98. Arita, H. *et al.* Upregulating mutations in the TERT promoter commonly occur in adult malignant gliomas and are strongly associated with total 1p19q loss. *Acta Neuropathol.* **126,** 267–76 (2013).

99. Nakada, M. *et al.* Receptor tyrosine kinases: principles and functions in glioma invasion. *Adv. Exp. Med. Biol.* **986,** 143–70 (2013).

Appendix 1

Candidate CIC-regulated genes in HEK-293a cells with increased expression in *cic*^{ZFN2} cells relative to *cic*^{WT} cells.

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
GPR3	16661429	5.19	4.10E-05	0.042743	1
HIST2H2BC	16692624	2.19	0.001547	0.083881	1
S100A16	16693474	2.17	0.001212	0.081541	1
HIST2H2AA4, HIST2H2AA3	16692620	1.91	0.001727	0.088225	1
RN5S46	16663446	1.9	0.000789	0.070737	1
PLK3	16664005	1.81	0.000698	0.066595	1
TNFRSF9	16681288	1.46	0.001503	0.082641	1
LOC100132999	16669626	1.43	0.000458	0.063141	1
LOC646626	16666730	1.23	0.000657	0.06624	1
LOC646471	16683703	1.23	0.000816	0.071158	1
CDC14A	16667662	1.22	8.10E-05	0.044115	1
TPM3	16693722	1.16	0.000111	0.044115	1
ADIPOR1	16698122	1.07	0.000931	0.073565	1
FOXJ3	16685958	1.05	2.10E-05	0.033825	1
IAH1	16876907	2.08	7.90E-05	0.044115	2
NOSTRIN	16887194	1.56	0.001627	0.085229	2
LBH	16878676	1.42	0.00089	0.072746	2
TMBIM1	16908338	1.29	0.000684	0.066568	2
ETV5	16962380	7.04	0.000526	0.064093	3
TIMP4	16950825	1.55	0.000248	0.056937	3
PHLDB2, PLCXD2	16943819	1.49	0.001241	0.081882	3
PLXND1	16959007	1.26	0.001376	0.082201	3
DCBLD2	16956714	1.23	0.001139	0.079304	3
TBC1D5	16951357	1.18	0.00136	0.082201	3
SNX4	16958487	1.1	0.000796	0.070737	3
PVRL3	16943763	1.1	0.001062	0.077753	3
NME6	16953303	1.07	5.40E-05	0.042743	3
USP17, USP17L6P, USP17L2, LOC728419, LOC100287205, LOC100287441, LOC100287478, LOC100287178, LOC100287364, LOC100287404, LOC100287513, LOC100288520, LOC100287238, LOC100287327, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965021	2.19	0.000472	0.063141	4
USP17, USP17L6P, USP17L2, LOC728419, LOC100287205, LOC100287441, LOC100287478, LOC100287178, LOC100287364, LOC100287404, LOC100287513, LOC100288520, LOC100287238, LOC100287327, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965023	2.19	0.000472	0.063141	4
USP17, USP17L6P, USP17L2, LOC728419, LOC100287205, LOC100287441, LOC100287478, LOC100287178, LOC100287364, LOC100287404, LOC100287513, LOC100288520, LOC100287238, LOC100287327, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965025	2.19	0.000472	0.063141	4
USP17, USP17L6P, USP17L2, LOC728419, LOC100287205, LOC100287441, LOC100287478, LOC100287178, LOC100287364, LOC100287404, LOC100287513,	16965029	2.19	0.000472	0.063141	4

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
LOC100288520, LOC100287238, LOC100287327, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369					
USP17, USP17L6P, USP17L2, LOC728419, LOC100287205, LOC100287441, LOC100287478, LOC100287178, LOC100287364, LOC100287404, LOC100287513, LOC100288520, LOC100287238, LOC100287327, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965031	2.19	0.000472	0.063141	4
USP17, USP17L6P, USP17L2, LOC728419, LOC100287205, LOC100287441, LOC100287478, LOC100287178, LOC100287364, LOC100287404, LOC100287513, LOC100288520, LOC100287238, LOC100287327, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965033	2.19	0.000472	0.063141	4
USP17, USP17L6P, USP17L2, LOC728419, LOC100287205, LOC100287441, LOC100287478, LOC100287178, LOC100287364, LOC100287404, LOC100287513, LOC100288520, LOC100287238, LOC100287327, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965037	2.19	0.000472	0.063141	4
USP17, USP17L2, LOC100287404, LOC100287364, LOC100287178, LOC100287205, LOC100287478, LOC100287441, LOC100287513, LOC100288520, LOC728419, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965011	2.13	0.000911	0.073368	4
USP17, USP17L6P, USP17L2, LOC100287364, LOC100287178, LOC100287404, LOC100287441, LOC100287513, LOC100287205, LOC100287478, LOC100288520, LOC728419, LOC100287327, LOC100287144, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369, LOC100287238	16965009	2.08	0.000688	0.066568	4
LOC100287205, LOC100287478, LOC100287441, LOC100287178, LOC100287404, LOC100287513, LOC100287364, LOC100288520	16965002	2.03	0.000864	0.072128	4
USP17, LOC100287205, LOC100287478, LOC100287441, LOC100287178, LOC100287404, LOC100287513, LOC100287364, LOC100288520, LOC728419, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965015	2.03	0.000864	0.072128	4
USP17, USP17L6P, USP17L2, LOC100287513, LOC100287178, LOC100287441, LOC100287205, LOC100287364, LOC100287478, LOC100288520, LOC100287404, LOC728419, LOC100287327, LOC100287238, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965017	1.94	0.00039	0.062854	4
USP17, USP17L6P, USP17L2, LOC100287441, LOC100287178, LOC100287205, LOC100287478, LOC100287513, LOC100287364, LOC100288520, LOC100287404, LOC728419, LOC100287327, LOC100287238, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965013	1.92	0.000435	0.063141	4
USP17, USP17L6P, USP17L2, LOC100287178, LOC100287364, LOC100287441, LOC100287513, LOC100287205, LOC100287404, LOC100287478, LOC100288520, LOC728419, LOC100287327, LOC100287238, LOC100287144, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965000	1.87	0.000324	0.061185	4
USP17L6P, LOC649352	16965039	1.83	0.000373	0.062854	4
USP17, USP17L6P, USP17L2, USP17L8, USP17L3, USP17L4, USP17L1P, USP17L7, LOC100287327, LOC100287238, LOC100287513, LOC100287478, LOC728419, LOC100287144, LOC649352, USP17L5, LOC728405, LOC728400, LOC728393, LOC728379, LOC728373, LOC728369	16965007	1.82	0.000548	0.064753	4
CTBP1-AS1	16963913	1.44	0.001528	0.083346	4
FLJ45340	16970094	1.33	0.001663	0.085756	4
SPRY4	17001063	7.58	6.30E-05	0.042743	5
ROPN1L	16983236	2.26	0.000137	0.046989	5
OSMR	16984244	1.55	0.001439	0.082201	5
PCDH1	17001005	1.53	0.00152	0.083075	5
TCF7	16989202	1.38	0.001463	0.082201	5
ARL10	16992796	1.27	0.001345	0.082201	5
FAM105B	16983388	1.26	0.001446	0.082201	5
DNAJC18	17000618	1.21	0.001494	0.082311	5
CXXC5	16989897	1.14	0.001349	0.082201	5
CYFIP2	16991527	1.11	0.001379	0.082201	5
C5orf65	17000605	1.09	0.001007	0.075338	5
MIR4458, LOC100505738	16983117	1.08	0.000315	0.061173	5
RN5S206	17017695	2.22	0.001287	0.082201	6
GPSM3	17017805	1.78	1.10E-05	0.023731	6

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
NRN1	17015324	1.7	0.001448	0.082201	6
HIST1H2BC, HIST1H2BI, HIST1H2BE, HIST1H2BF, HIST1H2BG	17005582	1.63	0.000168	0.052357	6
BACH2	17021738	1.62	0.001386	0.082201	6
BVES	17022139	1.55	0.000231	0.056937	6
FUCA2	17024374	1.51	0.000436	0.063141	6
TRAF3IP2-AS1	17011735	1.46	5.00E-06	0.022818	6
TAP1	17017979	1.46	0.002053	0.093692	6
MAPK13	17007910	1.44	0.000139	0.046989	6
TEAD3	17018459	1.43	0.001326	0.082201	6
GMPR	17005094	1.38	1.30E-05	0.023731	6
NEU1	17017495	1.38	0.002049	0.093692	6
RRAGD	17021596	1.34	0.000493	0.063141	6
PHF1	17007475	1.34	0.00126	0.082201	6
PLAGL1	17024394	1.33	0.001511	0.082748	6
GPR126	17013126	1.31	0.000313	0.061173	6
MDC1	17016966	1.29	0.000197	0.055038	6
LEMD2	17018309	1.28	0.000429	0.063141	6
ZFAND3	17008196	1.28	0.000601	0.065859	6
MTO1	17010316	1.25	0.000293	0.058961	6
LOC100289495	17014622	1.25	0.001182	0.080511	6
SLC35D3	17012904	1.24	0.000806	0.070737	6
C6orf62	17016205	1.23	0.001176	0.080511	6
PPP1R10	17016888	1.23	0.001412	0.082201	6
GRM4	17018347	1.23	0.001814	0.090974	6
MRPS18B	17006324	1.22	0.001912	0.092499	6
ASCC3	17022035	1.22	0.001953	0.093309	6
TJAP1	17009008	1.21	0.00054	0.064753	6
ICK	17020118	1.21	0.001118	0.079104	6
RDBP, MIR1236	17017575	1.21	0.001539	0.083613	6
TAB2	17013567	1.2	0.00013	0.046585	6
PFDN6	17007446	1.2	0.000138	0.046989	6
DOPEY1	17010639	1.19	0.00077	0.070722	6
BAG6	17017244	1.17	0.001328	0.082201	6
TRMT11	17012350	1.16	0.000358	0.062438	6
PGM3	17021188	1.16	0.000715	0.067744	6
IGF2R	17014364	1.14	0.000763	0.070722	6
OGFRL1	17010175	1.14	0.001231	0.081748	6
HCG25	17007417	1.13	0.001458	0.082201	6
ETV1	17055354	3.4	0.000173	0.052357	7
PARP12	17063480	1.81	4.20E-05	0.042743	7

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
ELFN1	17042925	1.77	0.000505	0.063141	7
ARHGEF35	17063996	1.71	0.001435	0.082201	7
LOC644794	17118262	1.59	0.000282	0.058913	7
KCNH2	17064299	1.39	0.001899	0.092499	7
SPDYE2, SPDYE2L, SPDYE5, SPDYE1, SPDYE6, LOC100509023	17049852	1.32	0.00089	0.072746	7
SPDYE2L, SPDYE2, SPDYE5, SPDYE1, SPDYE7P, SPDYE6, LOC100509023	17049869	1.31	0.000636	0.065859	7
C7orf57	17045904	1.27	0.000962	0.074411	7
PMS2P1	17060545	1.12	0.000364	0.062636	7
LOC729852	17043573	1.12	0.000459	0.063141	7
TECPR1	17060098	1.08	0.000896	0.072802	7
ZSCAN21	17049090	1.05	0.000915	0.073474	7
REXO1L2P, REXO1L1, LOC100288562	17070480	2.2	0.000106	0.044115	8
REXO1L2P	17078754	2.2	0.000106	0.044115	8
DUSP4	17075973	2	0.001631	0.085229	8
REXO1L2P, REXO1L1, LOC100288562	17070478	1.93	6.00E-05	0.042743	8
REXO1L2P	17078758	1.93	6.00E-05	0.042743	8
REXO1L2P	17078760	1.93	6.00E-05	0.042743	8
REXO1L2P	17078762	1.93	6.00E-05	0.042743	8
REXO1L2P	17078764	1.93	6.00E-05	0.042743	8
ESRP1	17070949	1.89	0.001099	0.078786	8
REXO1L2P	17078756	1.76	0.000402	0.062854	8
RHOBTB2	17066791	1.72	0.00175	0.089068	8
NIPAL2	17079448	1.5	0.000295	0.058961	8
EPPK1	17082362	1.39	0.00027	0.057502	8
XKR6	17074589	1.19	0.000835	0.071674	8
MAF1	17073577	1.13	2.90E-05	0.037813	8
SHC3	17095566	2.05	0.000666	0.066556	9
CNTFR	17093463	1.86	0.000104	0.044115	9
PTGER4P2	17085429	1.42	0.001397	0.082201	9
LHX2	17088916	1.3	0.00143	0.082201	9
KIAA2026	17092233	1.25	1.70E-05	0.029093	9
ANXA1	17085901	1.24	0.00115	0.079318	9
KIAA1161	17093417	1.19	0.000258	0.056937	9
KDM4C	17083492	1.16	0.001113	0.079104	9
DUSP5	16709128	2.23	0.001653	0.085401	10
AFAP1L2	16718592	1.45	0.001934	0.093214	10
BBIP1, LOC100130175	16718395	1.17	0.000294	0.058961	10
NDST2	16715642	1.17	0.000422	0.063141	10
WAC	16703520	1.17	0.001956	0.093309	10
CHST15	16719217	1.11	0.001233	0.081748	10

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
FOSL1	16740630	3.12	0.001937	0.093214	11
MDK	16724346	1.85	0.00078	0.070722	11
CCKBR	16721355	1.49	0.000563	0.065003	11
PAMR1	16737344	1.43	0.000827	0.07121	11
INS-IGF2, IGF2, INS	16734371	1.37	0.001492	0.082311	11
BACE1	16744822	1.32	0.00089	0.072746	11
CHST1	16737614	1.3	0.001704	0.087376	11
CAPN5	16729336	1.29	0.001108	0.079008	11
TP53I11	16737543	1.22	0.001239	0.081882	11
SLC25A45	16740378	1.2	0.001623	0.085229	11
LRP4-AS1	16724432	1.18	0.00139	0.082201	11
GDPD5	16742244	1.16	0.00198	0.093692	11
DUSP6	16768297	2.93	0.000502	0.063141	12
LPAR5	16760516	2.15	0.000249	0.056937	12
RHEBL1	16764220	2.05	0.000144	0.048085	12
COX6A1	16757886	1.44	0.000489	0.063141	12
LOC100509976	16747257	1.42	0.000397	0.062854	12
CLEC1A	16761259	1.41	0.001991	0.093692	12
RN5S379	16772702	1.4	0.001118	0.079104	12
TAS2R31, TAS2R45	16761518	1.28	0.002047	0.093692	12
C12orf66	16767009	1.21	0.000943	0.073565	12
ANKRD13A	16756865	1.14	0.000816	0.071158	12
SPRY2	16780069	1.64	5.00E-05	0.042743	13
IRS2	16780917	1.47	0.000882	0.072744	13
LMO7	16775434	1.25	0.000242	0.056937	13
LINC00564	16780082	1.08	0.001872	0.092359	13
LRP10	16782207	1.3	0.000677	0.066568	14
CBLN3	16791393	1.2	0.001312	0.082201	14
RBPMS2	16810572	1.85	3.30E-05	0.038791	15
PTPN9	16811816	1.23	0.001568	0.084186	15
MMP2	16819064	2.27	0.000464	0.063141	16
SYT17	16816424	1.75	0.001473	0.082201	16
IL4R	16817254	1.49	0.000341	0.06215	16
LOC100505865	16819813	1.45	0.001532	0.083398	16
KIAA0895L	16827255	1.44	0.001639	0.085401	16
ATP6V0D1	16827366	1.22	0.000559	0.064974	16
MIR3180-4, MIR3180-5	16824191	1.08	0.001483	0.082201	16
SULT1A1	16825391	1.04	0.000114	0.044115	16
ETV4	16845410	16.23	0.000342	0.06215	17
CD68	16830577	2.07	0.000468	0.063141	17

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
SLC43A2	16839425	1.73	0.001073	0.078141	17
MRC2	16836824	1.53	0.000957	0.074234	17
GPRC5C	16837571	1.5	0.000209	0.055562	17
FAM100B	16838049	1.47	5.00E-06	0.022818	17
NTN1	16831013	1.41	0.000504	0.063141	17
FMNL1	16834931	1.39	0.000498	0.063141	17
MFSD11	16838116	1.35	0.001993	0.093692	17
PGS1	16838392	1.34	0.000482	0.063141	17
ARSG	16837308	1.3	0.001399	0.082201	17
ARL17A, ARL17B, LOC100294341	16846006	1.27	1.30E-05	0.023731	17
TMEM132E	16833246	1.27	0.001247	0.081883	17
LOC100505873	17117728	1.26	0.00199	0.093692	17
ARL17A, ARL17B, LOC100294341	16845996	1.25	0.000323	0.061185	17
SLC16A13	16830295	1.25	0.00129	0.082201	17
G6PC3	16834700	1.25	0.001382	0.082201	17
KRT37	16844684	1.22	0.000821	0.07121	17
SHBG	16830607	1.19	0.000461	0.063141	17
UBE2Z	16835497	1.18	0.000431	0.063141	17
WBP2	16848938	1.18	0.00042	0.063141	17
NKIRAS2	16834252	1.15	0.000618	0.065859	17
HGS	16838855	1.15	0.001448	0.082201	17
PCGF2	16843987	1.07	0.00161	0.084995	17
HSBP1L1	16853325	1.54	0.001338	0.082201	18
KATNAL2	16852241	1.46	0.000503	0.063141	18
ZNF236	16853171	1.23	0.00093	0.073565	18
NFATC1	16853277	1.21	0.001952	0.093309	18
OR7C1	16869710	4.44	0.000634	0.065859	19
LOC386758	16865886	1.4	0.000967	0.074587	19
LOC100653348, LOC100506347	16866389	1.35	0.002016	0.093692	19
LOC284751	16914830	1.37	0.000995	0.075206	20
SNAI1	16914791	1.22	0.001833	0.091091	20
DNAJC5	16916146	1.1	0.001416	0.082201	20
HSF2BP	16926241	1.36	0.001128	0.079304	21
C21orf58	16926725	1.25	0.001292	0.082201	21
MCM3AP	16926679	1.09	0.001061	0.077753	21
LIF	16933760	1.48	0.000201	0.055038	22
SLC2A11	16928064	1.23	0.00095	0.0739	22
C22orf26	16936070	1.11	0.000379	0.062854	22
IL2RG	17111895	1.92	0.001224	0.081613	X
GYG2	17101231	1.85	0.001371	0.082201	X

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
L1CAM	17115271	1.77	0.001032	0.076522	X
RN5S514	17114190	1.62	0.000857	0.072128	X
CA5BP1	17101732	1.51	0.000608	0.065859	X
MAP7D2	17109680	1.35	0.000779	0.070722	X
CA5BP1	17101726	1.26	0.001378	0.082201	X
TFE3	17110745	1.14	0.000218	0.05685	X
TAB3	17109901	1.12	0.000594	0.065859	X
TAP1	17027144	1.46	0.002053	0.093692	6_apd_hap1
NEU1	17026875	1.38	0.002049	0.093692	6_apd_hap1
MDC1	17026706	1.28	8.10E-05	0.044115	6_apd_hap1
TAP1	17029788	1.46	0.002053	0.093692	6_cox_hap2
MDC1	17028857	1.27	9.20E-05	0.044115	6_cox_hap2
PPP1R10	17028781	1.24	0.001363	0.082201	6_cox_hap2
MRPS18B	17027506	1.22	0.001912	0.092499	6_cox_hap2
PFDN6	17028495	1.18	0.000262	0.056937	6_cox_hap2
VPS52	17028466	1.13	0.001458	0.082201	6_cox_hap2
RNF5	17028297	1.1	0.000622	0.065859	6_cox_hap2
TAP1	17032476	1.46	0.002053	0.093692	6_dbb_hap3
NEU1	17032134	1.38	0.002049	0.093692	6_dbb_hap3
MDC1	17031635	1.28	8.10E-05	0.044115	6_dbb_hap3
PPP1R10	17031560	1.24	0.001363	0.082201	6_dbb_hap3
MRPS18B	17030351	1.22	0.001912	0.092499	6_dbb_hap3
PFDN6	17031309	1.18	0.000262	0.056937	6_dbb_hap3
VPS52	17031280	1.13	0.001458	0.082201	6_dbb_hap3
TAP1	17034791	1.46	0.002053	0.093692	6_mann_hap4
NEU1	17034509	1.38	0.002049	0.093692	6_mann_hap4
MDC1	17034090	1.27	0.000446	0.063141	6_mann_hap4
PPP1R10	17034014	1.23	0.001412	0.082201	6_mann_hap4
MRPS18B	17033056	1.22	0.001912	0.092499	6_mann_hap4
VPS52	17033765	1.13	0.001458	0.082201	6_mann_hap4
TAP1	17037271	1.46	0.002053	0.093692	6_mcf_hap5
NEU1	17036841	1.38	0.002049	0.093692	6_mcf_hap5
MDC1	17036357	1.27	9.20E-05	0.044115	6_mcf_hap5
PPP1R10	17036281	1.23	0.001412	0.082201	6_mcf_hap5
MRPS18B	17035176	1.22	0.001912	0.092499	6_mcf_hap5
PFDN6	17036058	1.18	0.000262	0.056937	6_mcf_hap5
BRD2	17035950	1.15	0.00191	0.092499	6_mcf_hap5
VPS52	17036029	1.13	0.001458	0.082201	6_mcf_hap5
TAP1	17039977	1.46	0.002053	0.093692	6_qbl_hap6
NEU1	17039585	1.38	0.002049	0.093692	6_qbl_hap6

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
MDC1	17039133	1.27	0.000121	0.044176	6_qbl_hap6
PPP1R10	17039057	1.24	0.001363	0.082201	6_qbl_hap6
MRPS18B	17037835	1.22	0.001912	0.092499	6_qbl_hap6
PFDN6	17038779	1.18	0.000262	0.056937	6_qbl_hap6
VPS52	17038750	1.14	0.000736	0.069255	6_qbl_hap6
PHF1	17041368	1.35	0.000569	0.06519	6_ssto_hap7
MDC1	17041700	1.27	9.00E-05	0.044115	6_ssto_hap7
PPP1R10	17041641	1.24	0.001363	0.082201	6_ssto_hap7
MRPS18B	17040549	1.22	0.001912	0.092499	6_ssto_hap7
VPS52	17041336	1.13	0.001458	0.082201	6_ssto_hap7
ARL17B, ARL17A, LOC100294341	16850322	1.29	0.001817	0.090974	17_ctg5_hap1

Appendix 2

Candidate CIC-regulated genes in HEK-293a cells with decreased expression in *cic*^{ZFN2} cells relative to *cic*^{WT} cells.

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
PLA2G4A	16675197	-4.14	0.000417	0.063141	1
IFI16	16672390	-2.55	0.001142	0.079304	1
DDR2	16673075	-2.12	0.00023	0.056937	1
RYR2	16679142	-1.78	0.000387	0.062854	1
TDRD5	16674414	-1.54	0.001097	0.078786	1
TSPAN2	16691314	-1.44	0.000521	0.064093	1
OPN3	16701092	-1.39	0.001331	0.082201	1
RCC1, SNHG3, SNORA73A	16661589	-1.25	0.001143	0.079304	1
ORC1	16687188	-1.24	0.00085	0.072128	1
UAP1	16673056	-1.19	0.000328	0.061516	1
EFHD2	16659605	-1.19	0.001144	0.079304	1
UBE4B	16658758	-1.16	0.001651	0.085401	1
FH	16701077	-1.12	0.001341	0.082201	1
GNAI3	16668272	-1.06	0.000234	0.056937	1
FAM40A	16668464	-1.06	0.001395	0.082201	1
LOC440894	16884280	-8.04	1.00E-06	0.022818	2
IGFBP5	16908197	-3.84	0.001452	0.082201	2
LOC151009, LOC440894	16901683	-3.3	0.001648	0.085401	2
ITGA4	16888270	-2.49	9.70E-05	0.044115	2
ERBB4	16907863	-2.46	0.000161	0.051262	2
HOXD13	16887917	-2.33	0.000565	0.065003	2
SNAR-H	16899476	-2.2	0.000428	0.063141	2
DLX2	16905108	-2.06	0.000651	0.066021	2
IGFBP2	16890675	-1.74	0.000867	0.072159	2
GLS	16888865	-1.72	0.000493	0.063141	2
MAP2	16890207	-1.7	0.00056	0.064974	2
RPRM	16903863	-1.69	4.90E-05	0.042743	2
IFIH1	16904365	-1.59	8.90E-05	0.044115	2
B3GALT1	16887171	-1.56	8.10E-05	0.044115	2
GRB14	16904425	-1.55	0.000933	0.073565	2
MTX2	16887993	-1.48	0.000146	0.048166	2
ZAK	16887810	-1.48	0.001149	0.079318	2
ASNSD1	16888708	-1.48	0.001764	0.089614	2
NHEJ1, SLC23A3	16908557	-1.45	0.000295	0.058961	2
BCS1L	16890970	-1.45	0.001365	0.082201	2

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
MGAT5	16885874	-1.44	0.000529	0.064093	2
HECW2	16906749	-1.43	0.001452	0.082201	2
EEF1B2, SNORA41	16889938	-1.42	0.00182	0.090974	2
RAB3GAP1	16885939	-1.41	0.000802	0.070737	2
GMPPA	16891349	-1.39	7.00E-05	0.044115	2
CREB1	16890067	-1.39	0.001823	0.090974	2
BARD1	16907960	-1.38	0.000547	0.064753	2
RAMP1	16892975	-1.38	0.000595	0.065859	2
CNPPD1	16908604	-1.38	0.000639	0.065859	2
HOXD11	16887927	-1.36	0.000907	0.073261	2
SGOL2	16889251	-1.35	0.000323	0.061185	2
HSPE1-MOB4, MOB4, HSPE1	16889126	-1.33	5.00E-04	0.063141	2
HOXD9	16887945	-1.33	0.000647	0.065859	2
ACVR1C	16903953	-1.33	0.000985	0.075206	2
MMADHC	16903461	-1.32	0.000113	0.044115	2
DNPEP	16908782	-1.29	0.000184	0.053053	2
HAT1	16887635	-1.29	0.000641	0.065859	2
NOP58	16889602	-1.29	0.000632	0.065859	2
CUL3	16909049	-1.29	0.000773	0.070722	2
RND3	16903491	-1.29	0.00113	0.079304	2
PELI1	16898326	-1.29	0.001944	0.093309	2
RQCD1	16890915	-1.28	0.001101	0.078786	2
PECR	16908154	-1.27	0.001332	0.082201	2
SPATS2L	16889218	-1.26	0.000775	0.070722	2
SF3B1	16906921	-1.24	0.000627	0.065859	2
NCL	16909491	-1.24	0.000635	0.065859	2
FBXO11	16897349	-1.23	0.001581	0.084224	2
HDLBP	16910375	-1.22	0.000749	0.070237	2
TWIST2	16893143	-1.2	0.000345	0.06215	2
MFF	16891723	-1.19	0.00122	0.081613	2
NIF3L1	16889375	-1.18	3.40E-05	0.038791	2
NDUFA10	16910081	-1.18	0.00121	0.081541	2
GCFC2	16899429	-1.15	0.000226	0.056937	2
ZFAND2B	16891152	-1.13	0.001828	0.091008	2
YEATS2	16948589	-1.35	0.00182	0.090974	3
SLC6A11	16937626	-1.32	0.000508	0.063228	3
RNF7	16946439	-1.32	0.001588	0.084432	3
XYLB	16939203	-1.28	0.000779	0.070722	3
THRB	16951567	-1.22	0.001397	0.082201	3
PDZRN3	16956285	-1.21	0.000388	0.062854	3

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
DBR1	16959628	-1.19	0.0019	0.092499	3
SFRP2	16980762	-2.51	0.000491	0.063141	4
GPRIN3	16977970	-2.27	0.000936	0.073565	4
QRFPR	16979468	-1.85	0.000551	0.064753	4
VEGFC	16981730	-1.55	0.000251	0.056937	4
GUCY1B3	16971737	-1.38	0.001558	0.084086	4
SCD5	16977396	-1.21	0.000995	0.075206	4
PITX2	16979024	-1.18	0.001837	0.091125	4
LOC100507053	16969157	-1.06	0.001592	0.084481	4
ISL1	16984612	-2.36	6.40E-05	0.042743	5
LOX	16999180	-1.59	0.000111	0.044115	5
FABP6	16991729	-1.51	0.000284	0.058913	5
NR2F1, NR2F2	16987287	-1.26	0.001644	0.085401	5
SLC35F1	17012087	-2.02	0.000524	0.064093	6
C6orf132	17019190	-1.48	0.001458	0.082201	6
HIST1H2AE, HIST1H2AB	17005589	-1.34	0.000489	0.063141	6
LRRC16A	17005420	-1.13	0.000826	0.07121	6
CADPS2	17062321	-3.02	0.001783	0.090244	7
MYC	17072669	-2.78	0.000204	0.055306	8
NDRG1	17081401	-1.71	0.000626	0.065859	8
THEM6	17073234	-1.41	0.000872	0.072355	8
TUBBP5	17091877	-1.85	7.50E-05	0.044115	9
SYK	17086708	-1.34	0.000697	0.066595	9
MRPL50	17096631	-1.34	0.001102	0.078786	9
NCBP1	17087343	-1.32	0.000116	0.044115	9
ANKS6	17096516	-1.32	0.002019	0.093692	9
AUH	17095686	-1.31	7.10E-05	0.044115	9
MGC21881	17085432	-1.26	0.000937	0.073565	9
SEC61B	17087498	-1.24	0.000527	0.064093	9
LOC554249, MGC21881	17085187	-1.24	0.000854	0.072128	9
ZCCHC6	17095461	-1.22	0.000454	0.063141	9
SUSD3	17086845	-1.17	0.000312	0.061173	9
LRSAM1	17089324	-1.13	0.000399	0.062854	9
NACC2	17099816	-1.09	0.001034	0.076522	9
PLXDC2	16703036	-12.14	6.00E-06	0.02347	10
INA	16708796	-1.93	0.001956	0.093309	10
LOC84856	16704107	-1.18	0.001803	0.090971	10
SFMBT2	16711562	-1.13	0.001272	0.082201	10
ELP4	16723228	-1.45	0.000798	0.070737	11
OR4D11	16725097	-1.38	0.001962	0.093317	11

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
TMEM216	16725608	-1.36	0.000363	0.062636	11
RAB38	16743104	-1.36	0.000805	0.070737	11
CTSC	16743111	-1.3	4.50E-05	0.042743	11
UNC93B1	16741173	-1.24	0.000937	0.073565	11
NDUFC2-KCTD14, KCTD14, NDUFC2	16742594	-1.2	0.001015	0.07573	11
PRSS23	16729789	-1.2	0.001225	0.081613	11
DRAP1	16727168	-1.16	1.30E-05	0.023731	11
OTUB1	16726188	-1.04	0.000786	0.070737	11
PLBD1	16761726	-4.09	8.50E-05	0.044115	12
MGST1	16748788	-3.58	6.00E-04	0.065859	12
LGR5	16754134	-3.09	1.10E-05	0.023731	12
LOC400027	16763479	-1.55	0.000557	0.064974	12
LOC100130776	16753158	-1.43	0.001432	0.082201	12
LOC727803	17117588	-1.34	5.80E-05	0.042743	12
PKP2	16763032	-1.17	0.000208	0.055562	12
LRIG3	16766822	-1.17	0.001288	0.082201	12
PTMS	16747394	-1.13	0.000679	0.066568	12
AEBP2	16748939	-1.08	0.001577	0.084224	12
LPCAT3	16760668	-1.04	0.000691	0.066595	12
MLEC	16757969	-1.02	0.00118	0.080511	12
PCDH9	16779667	-2.26	0.000806	0.070737	13
GPC5	16775785	-1.49	0.000402	0.062854	13
DACH1	16779701	-1.32	0.000647	0.065859	13
SPATA13-AS1	16777448	-1.07	0.000188	0.05307	13
GALC	16795508	-2.22	0.001772	0.089854	14
MLH3	16794864	-2.12	0.000845	0.072128	14
SIX1	16793613	-1.27	0.000937	0.073565	14
ALPK3	16804251	-1.82	0.000771	0.070722	15
SORD	16800506	-1.66	0.000153	0.049288	15
CKMT1B, CKMT1A	16800242	-1.58	0.000646	0.065859	15
LPCAT4	16806920	-1.54	0.00076	0.070722	15
OIP5	16807605	-1.27	0.001804	0.090971	15
TM2D3	16813974	-1.24	0.000458	0.063141	15
TYRO3	16799852	-1.24	0.000718	0.067794	15
C15orf41	16799170	-1.16	0.001193	0.081059	15
WDR76	16800355	-1.15	0.001304	0.082201	15
ACSM3	16816604	-1.58	0.00156	0.084086	16
TRAP1	16823413	-1.44	9.00E-06	0.023731	16
TFAP4	16823512	-1.4	0.000574	0.06533	16
PLCG2	16821330	-1.33	0.000238	0.056937	16

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
TOX3	16826510	-1.29	8.50E-05	0.044115	16
MSRB1	16822868	-1.29	0.000859	0.072128	16
DNAJA3	16815513	-1.29	0.001624	0.085229	16
HMOX2	16815528	-1.28	2.60E-05	0.03747	16
THOC6	16815316	-1.27	0.000863	0.072128	16
PAQR4	16815289	-1.25	3.00E-06	0.022818	16
TBL3	16814872	-1.25	0.000487	0.063141	16
CPPED1	16824046	-1.22	0.000266	0.057351	16
NDUFB10	16814854	-1.2	0.000183	0.053053	16
C16orf13	16822466	-1.12	0.000695	0.066595	16
HAGHL	16814510	-1.09	0.001568	0.084186	16
GEMIN4	16839254	-1.26	0.000979	0.075206	17
GDPD1	16836511	-1.23	0.001484	0.082201	17
DSEL	16855781	-2.46	0.001346	0.082201	18
ZNF544	16866232	-4.23	0.001602	0.084847	19
ZNF257	16860221	-2.17	0.000119	0.044176	19
SYT3	16874591	-2.02	7.60E-05	0.044115	19
SNAR-C1, SNAR-C2, SNAR-C5, SNAR-C3, SNAR-C4	16863705	-1.94	0.000356	0.062438	19
SNAR-C1, SNAR-C2, SNAR-C5, SNAR-C3, SNAR-C4	16863711	-1.94	0.000356	0.062438	19
SNAR-C1, SNAR-C2, SNAR-C5, SNAR-C3, SNAR-C4	16863721	-1.94	0.000356	0.062438	19
SNAR-C1, SNAR-C2, SNAR-C3, SNAR-C4, SNAR-C5, SNAR-B1, SNAR-B2	16863715	-1.89	0.000391	0.062854	19
SNAR-C1, SNAR-C2, SNAR-C3, SNAR-C4, SNAR-C5, SNAR-B1, SNAR-B2	16863719	-1.89	0.000391	0.062854	19
SNAR-E	16873645	-1.89	0.000436	0.063141	19
SNAR-B1, SNAR-B2, SNAR-C1, SNAR-C2, SNAR-C3, SNAR-C4, SNAR-C5	16874508	-1.82	0.000404	0.062854	19
SNAR-B1, SNAR-B2, SNAR-C1, SNAR-C2, SNAR-C3, SNAR-C4, SNAR-C5	16874510	-1.82	0.000404	0.062854	19
JUND	16870384	-1.69	0.000262	0.056937	19
ZNF260	16871680	-1.64	0.001718	0.087929	19
QTRT1	16858263	-1.61	4.00E-06	0.022818	19
BCAT2	16874082	-1.58	0.002001	0.093692	19
ZNF529	16871688	-1.56	0.000575	0.06533	19
LONP1	16867511	-1.49	0.000366	0.062636	19
GNG7	16866974	-1.46	0.000136	0.046989	19
GCDH	16858756	-1.45	0.000683	0.066568	19
MRI1	16858849	-1.41	0.001257	0.082201	19
MEGF8	16862721	-1.4	6.30E-05	0.042743	19
LOC100506469	16871411	-1.4	0.000611	0.065859	19
PKN1	16858991	-1.4	0.001002	0.075338	19
LOC728485	16861454	-1.4	0.001326	0.082201	19
ZNF562	16868443	-1.4	0.001867	0.092278	19
ECSIT	16869006	-1.39	0.000112	0.044115	19

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
UHRF1	16857258	-1.39	0.001554	0.084086	19
BCL2L12	16864234	-1.38	0.000248	0.056937	19
PLEKHJ1	16866912	-1.38	0.000543	0.064753	19
NOSIP	16874327	-1.38	0.00059	0.065859	19
CARM1	16858321	-1.38	0.001084	0.078669	19
CIC	16862677	-1.37	2.50E-05	0.03747	19
NDUFA11, FUT5	16867583	-1.37	0.000604	0.065859	19
MARCH2	16857905	-1.36	0.001144	0.079304	19
NRTN	16857389	-1.36	0.001675	0.086212	19
RUVBL2	16863946	-1.35	2.80E-05	0.037813	19
KANK2	16868847	-1.35	0.000273	0.057502	19
ERCC1	16873313	-1.35	0.000941	0.073565	19
QPCTL	16863344	-1.35	0.001207	0.081541	19
PNPLA6	16857630	-1.35	0.001322	0.082201	19
LSM14A	16860678	-1.35	0.001353	0.082201	19
LSM4	16870387	-1.34	0.001434	0.082201	19
USE1	16859437	-1.33	0.000239	0.056937	19
MED25	16864331	-1.33	0.001469	0.082201	19
PEPD	16871239	-1.33	0.001611	0.084995	19
ERF	16872760	-1.33	0.00185	0.091603	19
SMARCA4	16858344	-1.32	0.000152	0.049288	19
AP2A1	16864304	-1.32	0.00017	0.052357	19
FSD1	16857163	-1.32	0.00049	0.063141	19
NR1H2	16864472	-1.32	0.001461	0.082201	19
NR2C2AP	16870581	-1.31	0.000287	0.058961	19
FBL	16872267	-1.31	0.000384	0.062854	19
CDC37	16868732	-1.3	0.001444	0.082201	19
ZNF296	16873221	-1.3	0.001466	0.082201	19
DDX39A	16869624	-1.29	0.000997	0.075206	19
EIF3K	16861841	-1.28	0.000444	0.063141	19
PRMT1	16864244	-1.28	0.000534	0.064414	19
SERTAD3	16872447	-1.27	0.000271	0.057502	19
TBC1D17, MIR4750	16864373	-1.27	0.000591	0.065859	19
TIMM13	16866946	-1.27	0.000612	0.065859	19
PIK3R2	16859740	-1.27	0.001408	0.082201	19
DNMT1	16868576	-1.26	0.000511	0.063313	19
HOOK2	16869299	-1.25	0.000115	0.044115	19
NUMBL, LOC100130713	16872460	-1.25	9.50E-05	0.044115	19
ILVBL	16869740	-1.24	0.00094	0.073565	19
ZNF317	16857950	-1.23	1.20E-05	0.023731	19

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
DAZAP1	16856567	-1.23	0.000377	0.062854	19
MRPL34	16859481	-1.23	0.000824	0.07121	19
ARHGEF18	16857567	-1.23	0.001072	0.078141	19
URI1	16860430	-1.22	7.00E-06	0.023731	19
NAPA	16873751	-1.22	0.000185	0.053053	19
ZNF180	16873183	-1.22	0.000401	0.062854	19
USF2	16860959	-1.21	0.000314	0.061173	19
SUPT5H	16862033	-1.2	0.001741	0.088775	19
NUDT19	16860545	-1.19	5.00E-06	0.022818	19
ZNF791	16858654	-1.19	0.000344	0.06215	19
MBD3	16866709	-1.18	0.000338	0.06215	19
ATG4D	16858195	-1.18	0.000686	0.066568	19
MED16	16866514	-1.18	0.000796	0.070737	19
MAU2	16859990	-1.18	0.001202	0.081468	19
R3HDM4	16866539	-1.18	0.001459	0.082201	19
ZNF57	16856848	-1.16	0.000644	0.065859	19
INSR	16867915	-1.16	0.001086	0.078669	19
XAB2	16867948	-1.15	0.000658	0.06624	19
ILF3	16858235	-1.15	0.001404	0.082201	19
TBXA2R	16867088	-1.14	0.000457	0.063141	19
CLPTM1	16863148	-1.13	0.000101	0.044115	19
KXD1	16859817	-1.13	0.000993	0.075206	19
PPP2R1A	16864778	-1.09	0.001645	0.085401	19
SNORD17	16917529	-2.1	0.000186	0.053053	20
LINC00493	16911780	-2.06	0.000779	0.070722	20
RBBP9	16917602	-1.98	0.000121	0.044176	20
NANP	16918132	-1.86	0.000351	0.062438	20
C20orf3	16918011	-1.86	0.001141	0.079304	20
SNX5	16917504	-1.78	0.000471	0.063141	20
ADA	16919466	-1.76	0.001246	0.081883	20
EYA2	16914478	-1.74	0.000258	0.056937	20
DSTN	16911651	-1.71	0.001511	0.082748	20
PLK1S1	16911923	-1.62	0.001577	0.084224	20
RALGAPA2	16917689	-1.61	5.50E-05	0.042743	20
BTBD3	16911463	-1.61	0.000183	0.053053	20
NAA20	16911853	-1.6	0.001937	0.093214	20
DTD1	16911783	-1.57	0.000551	0.064753	20
CRNKL1	16917655	-1.57	0.001026	0.076342	20
SOX12	16910597	-1.57	0.001092	0.078786	20
ENTPD6	16912140	-1.53	0.000228	0.056937	20

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
SEC23B	16911754	-1.51	0.000318	0.061185	20
NCOA5	16919769	-1.41	0.000875	0.072384	20
CTNNBL1	16913456	-1.41	0.002051	0.093692	20
SNRPB2	16911605	-1.39	0.000663	0.066499	20
DNMT3B	16912597	-1.39	0.001472	0.082201	20
NFS1	16918832	-1.37	3.30E-05	0.038791	20
ABHD12	16918077	-1.37	0.000987	0.075206	20
FAM83D	16913681	-1.36	0.000173	0.052357	20
DHX35	16913689	-1.33	0.001482	0.082201	20
GINS1	16912192	-1.31	0.000131	0.046585	20
POFUT1	16912520	-1.29	0.000212	0.055817	20
ITCH	16912905	-1.28	0.000582	0.065853	20
TRPC4AP	16918609	-1.27	0.001343	0.082201	20
JAG1	16917183	-1.26	0.000894	0.072802	20
DBNDD2, SYS1-DBNDD2, SYS1	16914213	-1.26	0.000983	0.075206	20
SLC35C2	16919804	-1.25	0.001692	0.086924	20
TOX2	16913985	-1.24	0.000854	0.072128	20
SPTLC3	16911493	-1.23	0.001181	0.080511	20
FOXA2	16917822	-1.21	4.40E-05	0.042743	20
TM9SF4	16912492	-1.2	0.000112	0.044115	20
DDRGK1	16916667	-1.19	0.000795	0.070737	20
CDS2	16911151	-1.16	0.001631	0.085229	20
SOGA1	16918996	-1.15	0.00117	0.080495	20
CEP250	16913095	-1.13	0.001824	0.090974	20
C20orf4	16913305	-1.09	0.000607	0.065859	20
BMP7	16920585	-1.08	8.60E-05	0.044115	20
CECR5	16931942	-1.32	0.000337	0.06215	22
SELO	16931662	-1.32	0.001047	0.077141	22
LARGE	16934308	-1.25	0.001006	0.075338	22
MCAT	16935767	-1.22	0.001048	0.077141	22
RANBP1	16927198	-1.21	0.000499	0.063141	22
SHANK3	16931815	-1.21	0.000416	0.063141	22
ZC3H7B	16930598	-1.21	0.000907	0.073261	22
SMARCB1	16928046	-1.21	0.001581	0.084224	22
BCR	16927907	-1.19	0.000103	0.044115	22
ALG12	16936255	-1.14	0.001432	0.082201	22
BCL2L13	16926893	-1.12	0.001363	0.082201	22
JOSD1	16935130	-1.09	0.000711	0.067599	22
DRG1	16929203	-1.08	0.001477	0.082201	22
LOC100130899	16930381	-1.08	0.001486	0.082201	22

Gene Symbol	Transcript Cluster ID	Fold Change	P value	Adjusted P Value	Chromosome
GABRA3	17115014	-2.63	0.000246	0.056937	X
F8A1, F8A3, F8A2	17108528	-2.45	0.00137	0.082201	X
GABRQ	17107855	-2.34	0.000201	0.055038	X
CAPN6	17113362	-2.33	0.001963	0.093317	X
RPS6KA6	17112439	-1.7	0.00122	0.081613	X
SRPX	17110071	-1.48	0.0019	0.092499	X
NUDT11	17111008	-1.37	0.000174	0.052357	X
MIR1184-1, MIR1184-2, MIR1184-3	17108585	-1.35	0.000684	0.066568	X
MIR1184-1, MIR1184-2, MIR1184-3	17115763	-1.35	0.000684	0.066568	X
MIR1184-1, MIR1184-2, MIR1184-3	17115796	-1.35	0.000684	0.066568	X
LOC286467	17114177	-1.26	0.000235	0.056937	X
HLA-DPB1	17026444	-1.06	0.000644	0.065859	6_apd_hap1
HLA-DPB1	17035971	-1.06	0.000644	0.065859	6_mcf_hap5