

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ Sameer H. Halani _____

_____ 2/21/2018 _____
Date

Multi-faceted computational assessment of risk and progression in oligodendroglioma uncovers
crucial role of Notch and PI3K pathways

By

Sameer H. Halani
M.D., Emory University, 2018
M.S., Tufts University, 2012
B.S., Emory University, 2010

Clinical Research

Daniel J. Brat, MD, PhD
Advisor

Amita Manatunga, PhD
Committee Member

Igho Ofotokun, MD
Committee Member

Lee Cooper, PhD
Advisor

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Multi-faceted computational assessment of risk and progression in oligodendroglioma
uncovers crucial role of Notch and PI3K pathways

By

Sameer H. Halani
M.D., Emory University, 2018
M.S., Tufts University, 2012
B.S., Emory University, 2010

Advisors: Daniel J. Brat, MD, PhD
Lee A.D. Cooper, PhD

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Clinical Research
2018

ABSTRACT

Multi-faceted computational assessment of risk and progression in oligodendroglioma uncovers crucial role of Notch and PI3K pathways

By
Sameer H. Halani

Oligodendroglioma are diffusely infiltrative gliomas defined by IDH-mutation and co-deletion of 1p/19q. They have highly variable clinical courses, with survivals ranging from 6 months to over 20 years, but little is known regarding the pathways involved with their progression or optimal markers for stratifying risk. We utilized machine-learning approaches with genomic data from The Cancer Genome Atlas (TCGA) to objectively identify molecular factors associated with clinical outcomes of oligodendroglioma.

We identified 169 patients with confirmed diagnosis of molecular oligodendroglioma that were collected and followed from 2006 to 2016, with whole-genome sequencing performed through TCGA. Deep learning neural networks were used to leverage the breadth of data of TCGA and *NOTCH1* mutations were the 5th strongest predictor of poor outcomes; these mutations are exclusively found in oligodendroglioma. These findings were extended to study signaling pathways implicated in oncogenesis and clinical endpoints associated with glioma progression. Inhibition of the entire canonical Notch pathway was found to be associated with poor clinical outcomes, thereby suggesting that global inactivation of this pathway is associated with a more aggressive subtype of oligodendroglioma. Investigation of more clinically relevant features of disease progression revealed *NOTCH1* mutations were enriched in tumors that exhibited

features of radiographic disease progression ($P = 0.008$); had greater tumor cell density ($P = 0.0015$); and had greater rates of malignant cell proliferation based on *MKI67* gene expression ($P = 0.095$).

Beyond the *NOTCH1* mutations, expression of downstream targets of the Notch pathway, including *HES* and *HEY*, were reduced in CE+ tumors ($P = 0.016$ and 0.050 , respectively), was negatively correlated with tumor cell density, and negatively correlated with cellular proliferation ($P < 0.05$ for both). Traditional survival analysis showed increased Notch pathway signaling was protective for both overall survival (HR = 0.34; 95% CI 0.18 to 0.64) and progression-free survival (HR = 0.41; 95% CI 0.23 to 0.72).

Our findings that Notch pathway inactivation is associated with advanced disease and survival risk will pave the way for clinically relevant markers of disease progression and therapeutic targets to improve clinical outcomes. Furthermore, our approach demonstrates the strength of machine learning and computational methods for identifying genetic events critical to disease progression in the era of big data and precision medicine.

Multi-faceted computational assessment of risk and progression in oligodendroglioma
uncovers crucial role of Notch and PI3K pathways

By

Sameer H. Halani
M.D., Emory University, 2018
M.S., Tufts University, 2012
B.S., Emory University, 2010

Advisors: Daniel J. Brat, MD, PhD
Lee A.D. Cooper, PhD

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Clinical Research
2018

ACKNOWLEDGEMENTS

I would like to thank my mentors, Dr. Brat and Dr. Cooper for their outstanding support throughout this project. Additionally, our entire research group within the Department of Pathology and Biomedical Informatics were integral parts in completing this project and without them, this project would not have been possible. Additionally, special thanks to the Tissue Procurement Service and the Research Pathology Laboratory of the Cancer Tissue and Pathology Shared Resource, as well as the Proteomics Shared Resource, at the Winship Cancer Institute, supported by the NCI Cancer Center Support Grant.

TABLE OF CONTENTS

Introduction.....	1
Background.....	3
Methods.....	5
Results.....	11
Discussion.....	18
References.....	23
Tables / Figures.....	26

INTRODUCTION

Diffuse lower-grade gliomas (LGGs; astrocytomas and oligodendrogliomas, grades II and III) are infiltrative brain tumors that arise most often in the cerebral hemispheres of adults.(1-3) Historically, these were classified by histology, but are now considered discrete, molecularly defined disease subtypes based on mutational status of *isocitrate dehydrogenase (IDH)* and co-deletion of chromosomes 1p and 19q.(1, 3-5) Molecular signatures of LGG include: *IDH* wild type (wt) astrocytomas, which behave most aggressively; *IDH* mutant astrocytomas, which also harbor *TP53* mutations and alpha thalassemia/mental retardation syndrome X-linked (*ATRX*) alterations; and oligodendrogliomas, defined by *IDH* mutations with 1p/19q co-deletion. The shift towards molecular profiling enables precise and reproducible diagnosis, better prediction of clinical course, and a solid platform for future research.

We employed machine-learning approaches to identify molecular features associated with clinical outcomes of oligodendroglioma using The Cancer Genome Atlas (TCGA) LGG dataset. We advanced and translated these findings using neuroimaging and pathology imaging features of oligodendroglioma progression to identify factors most closely related to advanced disease status, as defined by: 1) contrast-enhancement on magnetic resonance (MR) imaging; 2) high cellular density in digitized histopathologic images; and 3) increased cellular proliferation.(12-14) In addition, our approach enabled us to determine alterations in signaling pathways in patients with more aggressive disease. Our multifaceted, computational approach confirmed the association of NOTCH1 mutations with disease progression and shorter survival in oligodendroglioma,

and further uncovered aberrant regulation of Notch and PI3K as pathways most strongly associated with advanced disease.

BACKGROUND

Diffuse lower-grade gliomas (LGGs; astrocytomas and oligodendrogliomas, grades II and III) are infiltrative brain tumors that arise most often in the cerebral hemispheres of adults.(1-3) Historically, these were classified by histology, but are now considered discrete, molecularly defined disease subtypes based on mutational status of *isocitrate dehydrogenase (IDH)* and co-deletion of chromosomes 1p and 19q.(1, 3-5) Molecular signatures of LGG include: *IDH* wild type (wt) astrocytomas, which behave most aggressively; *IDH* mutant astrocytomas, which also harbor *TP53* mutations and alpha thalassemia/mental retardation syndrome X-linked (*ATR*X) alterations; and oligodendrogliomas, defined by *IDH* mutations with 1p/19q co-deletion. The shift towards molecular profiling enables precise and reproducible diagnosis, better prediction of clinical course, and a solid platform for future research.

Oligodendrogliomas account for 16-22% of all adult gliomas and have the least aggressive clinical course, yet display widely variable outcomes—some patients survive 6 months while others live over 15 years.(1, 3, 6) Patterns of progression in this genetically defined disease have not been established and the mechanisms and biomarkers that drive them or could stratify risk have not been explored; however, morphological markers of prognosis from the pre-*IDH* era are not optimal.(7) Oligodendrogliomas harbor frequent mutations, including: capicua transcriptional repressor (*CIC*) (62%), far upstream element binding protein 1 (*FUBP1*) (27-29%), *NOTCH1* (18-31%), catalytic and regulatory subunits of phosphoinositide-3-kinase (PI3K; *PIK3CA* (15-20%) and *PIK3R1* (7-9%), respectively), and others.(1, 8, 9) Now that LGGs are understood in objective, molecular terms, mechanisms of progression and

targets of therapy are being evaluated in a pure cohort, without the confounding contamination of dissimilar tumor types. Recent investigations of oligodendroglioma by Aoki et al.(10) have uncovered key genetic alterations that are present in oligodendroglioma, namely NOTCH1 mutations, and have demonstrated these mutations are associated with poor clinical outcomes and demonstrate radiographic features of advanced disease. The impact these mutations have on key signaling pathways has yet to be elucidated.

With the tremendous expansion of genomic data available for both investigation and potential clinical care, the need for novel computational approaches to further investigate risk factors in a highly multidimensional and interdependent space has grown in parallel.(11) Machine-learning approaches are capable of modeling clinical outcomes and identifying risk factors using large genomic datasets in a manner that adds value to traditional risk modeling by taking an objective, unbiased approach to identifying key prognostic features from tens of thousands of variables at once.

METHODS

Specific Aims

Aim 1: Develop and employ a machine learning algorithm based on neural networks to identify genetic alterations that are predictive of poor clinical course in molecular oligodendroglioma(hypothesis-generating aim).

Aim 2: Determine which key genetic alterations, that are specific to oligodendroglioma, are enriched in tumors exhibiting features of advanced disease (hypothesis-testing aim). The features focused on as surrogates for aggressive and advanced disease included: A.) Contrast-enhancement present on radiographic imaging; B.) increased tumor cell density; and C.) high tumor cell proliferation.

Study design

Our investigation utilized the clinical and genomic data from the Open Access Data Tier of the TCGA LGG dataset for 169 molecularly defined (IDH mutant, 1p/19q co-deleted) oligodendroglioma (<http://cancergenome.nih.gov/>; last accessed September 7th, 2016). Clinical variables consisted of age, gender, extent of resection, survival time, survival status, event-free survival time, and event status; tumor characteristics included tumor location, and WHO tumor grade.

Neural network survival model

We trained nonlinear Cox proportional hazard models using neural networks to maximize Cox likelihood using Theano software. Two models were constructed: 1) a genetic-protein model based on clinical features (radiation therapy, histologic grade), age, gender, frequent mutations, frequent copy number events, and reverse phase protein array (RPPA) profiles, and 2) a transcriptional model based on mRNA sequencing features alone. Frequent mutations and CNAs were defined using MutSig P -value threshold of 0.1, and Genomic Identification of Significant Targets in Cancer (GISTIC) P -value threshold of 0.25.(15, 16) The prognostic significance of each input feature in each model was assessed by using mathematical derivatives to evaluate the sensitivity of predicted risk to changes in feature values. The prognostic significance weights of features in the mRNA model were used to perform pathway analysis to identify pathways enriched with either good or poor prognosis transcripts. Pathway analysis was performed with GSEA using the Canonical Pathways gene set from the MSigDB curated gene sets.

Mutations and chromosomal aberrations

Gene expression, mutation, and copy number data were obtained from the TCGA data portal (<https://tcga-data.nci.nih.gov>). Key genetic alterations were identified based on their frequency of occurrence in oligodendroglioma; genes mutated in at least 5% of patients were included and particular attention was paid to mutations that were specific to oligodendroglioma (**Table S1A**). Variants were considered mutants if there was an amino acid change and genes were filtered using $q \leq 0.05$ in MutSig analysis. Mutations were then converted into dichotomous variables (mutation and wild-type) for further analysis.

Arm level copy number was obtained from GDAC GISTIC hosted analysis results (<https://gdac.broadinstitute.org/>). Values of chromosomal arm gain or loss were listed as a fraction of the chromosomal arm, where gains were positive values and losses were negative values. A threshold absolute value of 0.10 of the fraction of the chromosomal arm was used to signify chromosomal gain or loss. Frequency of chromosomal gains and losses are summarized in **Table S1B**.

Radiographic imaging review

Corresponding MR imaging studies for TCGA patients were obtained from TCIA (<http://www.cancerimagingarchive.net/>; last accessed February 8th, 2016) for 55 patients. Only untreated patients were included. Institutional neuroradiologists and neurosurgeons reviewed MR images to identify the presence of unequivocal contrast-enhancement on pre-operative MR images.

Quantification of cellular density and nearest-neighbor analysis

Whole-slide digital pathology images (n=142) were available from the CDSA (<http://cancer.digitalslidearchive.net/>; last accessed August 11th, 2016). Digital whole-slide images were analyzed using a nearest-neighbor *in silico* analysis algorithm to quantify cellular density in areas of tumor infiltration. Images were analyzed at 20X objective magnification to delineate individual cell nuclei using a published algorithm.(17) Cellular density was assessed by analyzing the centroids of each delineated nucleus. Tissue sections in LGGs often contain a mixture of dense, highly-cellular tumor regions and sparse adjacent normal tissue. We first used KD-trees to

calculate nearest-neighbor distances for each cell in a slide (a slide can contain up to a million or more cells). We model the incidence of cell nuclei as a Poisson point process, where the distance between neighboring cells follows a Poisson distribution. To model the mixture of tumor/normal populations, we model the distribution of nearest-neighbor distances in each slide using a Poisson mixture model with parameters $\pi_{\text{tumor}} / \lambda_{\text{tumor}}$, $\pi_{\text{normal}} / \lambda_{\text{normal}}$, where π are the mixture parameters and λ are the exponential Poisson distribution parameters. The tumor density of a slide is then represented as $\lambda_{\text{tumor}}^{-1}$. To improve robustness, we performed this analysis using the 5th nearest neighbor instead of the 1st, as this measure is less sensitive to errors in delineating nuclei boundaries and is a better representation of overall density. The median numerical score was used to delineate “less dense” and “more dense”. Cell density was also analyzed visually by a neuropathologist (JV), blinded to nearest neighbor analysis, and scored as: ‘low’, ‘intermediate’, or ‘high’. The algorithm-calculated nearest neighbor distances were then compared to visual scoring, using t-tests and one-way ANOVA (for multiple groups) tests for statistical significance (<0.05 level).

Gene expression of MKI67 as a marker for cellular proliferation

Gene expression for *MKI67* was extracted from the TCGA data portal. ‘High’ *MKI67* was defined as ≥ 700 , correlating with a 15% MIB-1/Ki-67 labeling index by extrapolating from a linear regression model.(13) Samples with *MKI67* < 700 were designated ‘low’.

Survival analysis

Clinical data was obtained from the TCGA data portal (last accessed: January 22nd, 2016). OS was defined as months from initial diagnosis to death for uncensored patients and months to last follow-up for right-censored patients. Survival curves were estimated using the Kaplan-Meier method; log-rank tests were used to compare curves between groups.

PFS was defined as months from initial diagnosis to radiographic disease progression or death for uncensored patients and months to last follow-up for right-censored patients. PFS curves were estimated using the Kaplan-Meier method; log-rank tests were used to compare curves between groups.

Single and multiple-predictor models were fit using Cox regression under the proportional hazards assumption for OS and PFS.

Statistics

Associations between contrast-enhancement and mutational status were calculated using the χ^2 test for independence; for expected counts less than 5, Fisher's exact test was used. Statistical associations between 2 groups of continuous or ordinal variables, such as the cellular density calls, were calculated using t-tests. The Pearson correlation coefficient was used to measure the linear dependence between continuous variables.

All *P*-values reported are two-sided and regarded as statistically significant if $P < 0.05$.

The software used for statistical analysis and graphical representations include: SPSS v23 (SPSS Statistics, IBM Corp., NY) and R Studio v0.99.

Validation set

A cohort of 51 patients with primary oligodendroglioma (*IDH* mutant, 1p/19q co-deleted) with pre-operative MR imaging at Emory University Hospital was identified. Pre-operative MRIs were reviewed by a neuroradiologist for the presence of unequivocal contrast enhancement. Histologic slides from formalin-fixed, paraffin-embedded tissues derived from neurosurgical resection specimens were reviewed by two neuropathologists (DJB and JV). Immunohistochemical (IHC) staining was performed on unstained formalin fixed paraffin-embedded slides for Ki-67; a proliferation index was quantified using digital image analysis in representative regions of interest with Aperio Positive Pixel count software. The 3+ and 2+ cells were summed for a Ki-67 percentage in regions of interest. Cellular density was also calculated using this method in the same regions; all cells were counted and divided by the area of the region of interest (mm²). IHC for Notch signaling was assessed using anti-HEY2 rabbit polyclonal antibody (catalog #AB5716, Millipore, 1:100) and for PI3K using anti-pAkt (S473) rabbit monoclonal antibody (#EP2109Y, Abcam, 1:100). Two neuropathologists (DJB and JV) reviewed HEY2 and pAkt IHC slides blinded to clinical data and samples were separated based on staining intensity. Tissue indicative of tumor bulk was cored for each of the 51 patients. Tumor samples underwent DNA isolation and focused sequencing of the NOTCH1 gene was performed using Sanger sequencing. Sequencing regions of the NOTCH1 gene included the epidermal-growth-factor-like domain (EGF-like) spanning amino acids 300 to 500. Fluidigm sequencing was also performed using a glioma gene panel.

RESULTS

Patient and Tumor Characteristics

The clinical factors from the 169 oligodendroglioma patients included in our study are presented in **Table 1**. *TERT* promoter mutations were present in 98% (86 of 88).¹⁴

Neural network analyses identifies molecular factors associated with outcomes

Analysis of the genetic-protein neural network model revealed multiple mutations, CNAs, and proteins associated with overall survival in oligodendrogliomas including (see **Fig. 1A**). *NOTCH1* (rank #5), *BCOR* (rank #4), and *ZBTB20* (rank #1) mutations were among the most highly ranked factors associated with poor prognosis, along with loss of 15q (rank #3). Interestingly, both Notch1 mutations and 15q loss occur in a substantial subset of oligodendrogliomas and have been suggested as markers of poor prognosis in traditional risk models. The complete list of ranked factors is in the Supplementary Materials (**Data file S1**). Among these factors, we focused on the Notch pathway since *NOTCH1* mutations are relatively specific to oligodendroglioma among diffuse gliomas; occur in a substantial subset (18-31%) compared to *BCOR* and *ZBTB20*; and represent one component of the Notch signaling network that could be more generally relevant to disease progression. PI3K pathway subunit mutations were also of interest since they were heavily enriched among highly ranked negative prognostic factors (*PIK3R1*, #30; *PIK3CA*, #193).

Similar analysis of the gene expression neural network model was performed to determine the prognostic importance of mRNA transcripts, and a gene-set-enrichment

analysis (GSEA) was then used to identify molecular pathways enriched with prognostic transcripts. GSEA identified the NOTCH1 Intracellular Domain Regulates Transcription pathway ($P = 0.004$) as highly enriched in transcripts associated with better prognosis, suggesting that Notch pathway inactivation is associated with poor outcomes (**Fig. 1B**). Regulation of KIT Signaling was also significantly enriched with positive prognosis transcripts ($P = 0.002$). The P38 / MKK3 ($P < 0.05$) and SMAD2 / SMAD3 pathways ($P = 0.002$) were also significantly enriched in transcripts associated with a poor prognosis, and notably, both interface directly with the PI3K pathway.^{15,16}

Radiographic and pathologic features are associated with aggressive clinical behavior

We next focused on mutations and CNAs with a >5% incidence to assess their association with radiographic and pathologic measures of disease progression, including: mutations of *CIC* (ranked #107; 61.5% incidence) *NOTCH1* (ranked #5; 18.9%), *FUBP1* (ranked #20; 27.2%), both *PIK3* subunits (*PIK3RI* ranked #30 and *PIK3CA* ranked #193; 23.1%), and CNA's including gain of chromosomal arms 7p (ranked #300; 8.9%) and 11p (ranked #153; 11.2%), as well as loss of 14q (ranked #310; 11.8%) and 15q (ranked #3; 16.6%) (**Fig. S1; Table S1**).

Contrast-enhancement observed on MRI is a well-known marker of higher-grade disease (**Fig. 2A**). Among 55 patients with MRI images available, contrast-enhancing (CE+) tumors (n=35) had worse overall survival (OS) (median, 154.3 vs. 62.0 months; $P = 0.10$) and progression-free survival (PFS) (median, 97.3 vs. 63.8 months; $P = 0.029$) compared to those that lacked enhancement (CE-) (n=20) (**Fig. 2B-C**). CE+ was highly enriched for histologic grade III tumors; 24 of 25 grade III tumors were CE+ ($P < 0.0001$).

Since cell density increases with disease progression, we used a computational nearest-neighbor analysis to quantify cellular density in tissue sections from 142 cases (**Fig. 2D**). Higher cell density trended towards worse OS (mean 152.8 vs. 126.1 months; $P = 0.076$) and worse PFS (median 142.8 vs. 95.9 months; $P = 0.14$) (**Fig. 2E-F**). High cell density were also enriched for histologic grade III tumors; 44 of 58 high density tumors were WHO grade III ($P < 0.0001$).

As a measure of proliferation, *MKI67* mRNA expression was analyzed for 169 tumors. *MKI67* expression was strongly correlated with the Ki-67/MIB-1 proliferation indices based on measurements performed by IHC and listed in TCGA pathology reports ($P < 0.0001$) (**Fig. 2G-H**). Patients with high cellular proliferation ($n=31$) had worse OS (median 154.3 vs. 62.0 months; $P = 0.001$); no significant difference was noted in PFS ($P=0.38$) (**Fig. 2I-J**). Highly proliferative tumors were also enriched for histologic grade III tumors; 21 of 28 high proliferation tumors were WHO grade III ($P = 0.001$).

Genetic alterations associated with radiographic contrast enhancement, cellular density, and MKI67 expression

Among 55 patients with MR imaging (**Table S2**), *NOTCH1* mutations were most strongly associated with CE+ tumors, with 13 of 14 *NOTCH1* mutants being CE+ ($P = 0.008$) (**Fig. 3A**). The combined *PI3K* group mutants were mostly CE+ (14 of 18; $P = 0.054$), and a similar trend was found among *FUBP1* mutants (14 of 17; $P = 0.13$). All 9 tumors with 11p gain were CE+ ($P = 0.019$). Although 5 of 5 tumors demonstrating loss

of 14q were CE+, this did not reach statistical significance ($P = 0.15$). Similar trends were found with 15q loss (9 of 10 CE+; $P = 0.075$) and 7p gain (6 of 6 CE+; $P = 0.076$).

NOTCH1 mutant oligodendrogliomas (n=26) had higher cellular density than *NOTCH1* wild-type tumors (n=126) and this difference was the most significant among all mutations and CNAs ($P = 0.0015$) (**Fig. 3B**). *FUBP1* mutants (n=40) trended toward a higher cellular density compared to wild-type (n=102; $P = 0.10$), and *CIC* (n=88) and *PIK3* (n=33) mutants did not show increased cell density (**Fig. S2**). Gains of 7p (n=12) or 11p (n=17) were significantly associated with higher cell densities ($P = 0.006$ and 0.03 , respectively), and loss of 15q (n=21) trended towards higher cellular density as well ($P = 0.19$) (**Fig. 3B**).

NOTCH1 mutants (n=32) had higher *MKI67* expression and this association was the strongest among all mutations and CNAs tested ($P = 0.095$) (**Fig. 3C**). *FUBP1*, *CIC*, and *PIK3* mutations were not strongly related to *MKI67* expression (**Fig. S3**). Although gain of 7p and 11p, and loss of 14q and 15q trended towards higher cellular proliferation, none reached statistical significance.

Inactivation of the canonical Notch pathway is associated with disease progression measures

Since *NOTCH1* mutations were consistently and strongly associated with radiologic, pathologic, and molecular measures of progression, we investigated downstream targets of the canonical Notch pathway, including family members of hairy/enhancer of split 1

(*HES*) and hairy/enhancer of split with YRPW motif (*HEY*). Since nearly all (93%) *NOTCH1* mutations were located within the epidermal growth factor (EGF) like region, where they inhibit Notch activation, we hypothesized these targets would be down regulated in *NOTCH1* mutants^{17,18}. Expression of *HES1*, *HEY1*, and *HEY2* was reduced in CE+ tumors, with *HES1* and *HEY2* reaching statistical significance ($P = 0.016$ and 0.050 , respectively) (**Fig. 4A** and **Fig. S4**). *HEY2* (Pearson correlation = 0.230, $P = 0.006$) was positively correlated with nearest-neighbor distance (**Fig. 4B**) and negatively correlated with cellular proliferation as approximated by *MKI67* expression (Pearson correlation = -0.353, $P < 0.0001$) (**Fig. 4C**). Negative correlations between *MKI67* expression and *HES1* (Pearson correlation = -0.152, $P = 0.048$) and *HEY1* (Pearson correlation = -0.082, $P = 0.288$) were also observed. Thus, among *HES* and *HEY* family members, *HES1*, *HEY1* and *HEY2* showed reduced expression with advanced disease, with *HEY2* showing the most consistent and statistically significant reductions.

Alternate mechanisms of Notch pathway inactivation in oligodendroglioma

Recombinant signal binding protein for immunoglobulin kappa-J region (RBPJ), the nuclear binding partner of activated NOTCH1's intracellular binding domain (NICD), was mutated (n=5) or homodeleted (n=1) in 3% (6 of 169) of oligodendrogliomas. *RBPJ* aberrations were mutually exclusive with *NOTCH1* mutations and were not present in *IDH* mutant or *IDH* wild-type astrocytomas. *RBPJ* altered tumors had greater *MKI67* expression compared to wild-type ($P = 0.001$) and showed a trend toward higher cell density ($P = 0.20$), but were not enriched in CE+ tumors (**Fig. S5A**). When *RBPJ* and *NOTCH1* mutant tumors were grouped (n=38), *MKI67* expression and 1/nearest-neighbor

distance show stronger statistical significance in the combined group compared to *NOTCH1* mutants alone ($P = 0.0030$ and 0.00039 for combined groups, respectively vs. $P = 0.095$ and 0.002 for *NOTCH1* mutants alone) (**Fig. S5B**). Thus, *RBPJ* mutation likely represents an alternative mechanism for Notch pathway inactivation in oligodendroglioma.

Survival analysis reveals PIK3 mutations and reduced Notch target expression are associated with worse prognosis

A comprehensive analysis of clinical and genetic factors associated with survival was performed using a Cox proportional hazards models (**Table 2, 3 and Table S3**).

Univariable analysis revealed age and grade as strong predictors of poor OS (Hazards ratio (HR) 3.64 per 10 years, $P < 0.0001$; HR 6.61, $P = 0.013$, respectively). After adjusting for age and grade, the combination of *PIK3* mutations were found to confer poor prognosis (HR 3.11, $P = 0.045$). Among the downstream Notch target genes, increased *HES5* expression had a significant protective effect (HR 0.74, $P = 0.024$) after accounting for age and grade.

Univariable analysis of PFS uncovered increased risk with grade III relative to grade II (HR 2.24, $P = 0.046$). *PIK3* (HR 1.98, $P = 0.092$) mutations trended toward increased risk of progression after accounting for tumor grade. Loss of 14q (HR 3.90 $P = 0.0035$) predicted more rapid time to progression after adjusting for grade. While *NOTCH1* mutants were not individually predictive of PFS, when combined with *RBPJ* altered tumors, the combined mutants predicted shorter time to first progression (HR 2.47, $P =$

0.021). After adjusting for grade, reduced *HEY1* (HR 0.48, $P = 0.018$) expression had a negative impact on PFS, while *HES5* trended in this direction (HR 0.86, $P = 0.120$).

Complete survival analysis results in **Table S3** and **Fig. S6-S7**.

Translation and Validation in Clinical Cases

We investigated 51 newly diagnosed cases of oligodendroglioma, grades II and III, from our institution. Pre-operative imaging was available for 47. We focused our IHC analysis on HEY2, since its gene expression showed greatest reduction in *NOTCH1* mutants, and pAkt, a downstream marker of PI3K activation (**Fig. 4D-E**).

Thirty-two tumors were WHO grade II and 19 were grade III; 21 tumors were CE- and 26 were CE+. By IHC analysis of HEY2, 20 tumors showed low expression and 31 showed high expression. Fourteen of 19 (73.7%) tumors with low HEY2 were CE+. Tumors with low HEY2 also had greater cell density ($P = 0.014$) and were more proliferative ($P = 0.0096$) than those with increased HEY2 staining (**Fig. 4F**). IHC investigation of pAkt found 27 tumors had low expression; 22 showed high expression; 15 of 20 (75%) tumors with pre-operative imaging and high pAkt expression were CE+. Tumors with high pAkt expression had greater cell density and were more proliferative ($P < 0.0001$, for both) (**Fig. 4F**).

DISCUSSION

We used a multi-faceted, technologically advanced, computational approach to identify molecular events associated with aggressive disease within molecularly defined oligodendroglioma (*IDH* mutant, 1p/19q co-deleted) and uncovered Notch pathway inactivation and PI3K activation as critical events. Our neural network deep learning methodology analyzed multiplatform TCGA molecular data by iteratively modeling the training datasets, based on clinical outcomes, in order to objectively rank risk factors. In concordance with recent investigations(10), *NOTCH1* mutations were identified as one of the most highly weighted risk factors in our deep learning prognostic model, and was the genetic event most associated with disease progression in each endpoint assessed (MRI contrast-enhancement, cell density, and cellular proliferation). Therefore, inactivating point mutations of *NOTCH1* are one of the most clinically meaningful alterations in oligodendroglioma progression and might suggest that inactivation of the Notch pathway is more generally responsible for poor clinical outcomes.

The NOTCH family is an evolutionarily conserved set of transmembrane receptors that regulate numerous critical biological functions. Notch pathway is activated by extracellular ligand binding, followed by γ -secretase cleavage to release an active intracellular domain (NICD), which localizes to the nucleus and binds to its partner RBPJ to initiate transcription of downstream targets, including *HES* and *HEY* family members.(23, 24) Both activating and inactivating *NOTCH1* mutations have been described in cancer, including in oligodendroglioma.(10, 24-27) Inactivating mutations, such as those noted oligodendroglioma and head and neck squamous cell carcinoma, are

enriched within EGF-like regions and interfere with ligand-mediated pathway activation.(24, 26, 28, 29) (21, 30) (1)

Our results suggest inactivation of Notch signaling may be more relevant to oligodendroglioma progression than *NOTCH1* mutations alone. For example, reduced expression of Notch targets, namely *HES1*, *HEY1*, and especially *HEY2*, was seen in clinically progressed oligodendroglioma, while *HES5* expression was most associated with shorter survival on multivariable analysis. *HEY2* showed a strong positive correlation with cellular density and proliferation, beyond those of *NOTCH1* mutations alone, suggesting other Notch pathway members might be inactivated and lead to reduced downstream target activation.

Furthermore, we found mutations and deletions of *RBPJ*, the nuclear binding partner of NOTCH1 and a member of the canonical Notch pathway, are linked to advanced disease, providing additional evidence that Notch pathway inactivation may be a general progression mechanism. RBPJ normally recruits corepressor proteins and suppresses transcription of downstream targets, whereas active NOTCH1 binds RBPJ and initiates transcription.(31) Genetic aberrations of *RBPJ* likely prevent active NOTCH1 from binding to the transcriptional complex. However, Notch-independent functions of RBPJ have also been described.(31) *RBPJ* was mutated in 3% of our cohort and homozygously deleted in another case, which is relatively low, but consistent with other forms of cancer.(22, 32) Importantly, *RBPJ* alterations were mutually exclusive from *NOTCH1* mutations, showed strong trends of association with features of disease progression, and

had reduced downstream target expression when considered independently. When cases with either *NOTCH1* mutations or *RBPJ* alterations were considered together, the combined group was more strongly associated with disease progression and pathway inactivation than either one alone, and was strongly associated with worse PFS, again raising the possibility that Notch pathway inactivation by multiple mechanisms may be associated with oligodendroglioma progression.

Other prognostically-significant chromosomal aberrations associated with disease progression uncovered by our analysis, including losses of 14q and 15q and gains of 7p, also harbor Notch pathway members, and may be mechanistically relevant to pathway inactivation and disease progression, but will require further investigation. Chromosome 14q contains genes that encode presenilin-1 (PSEN1), a component of the γ -secretase that activates Notch; NUMB, a Notch inhibitor; and jagged-2 (JAG2), a NOTCH receptor ligand. 15q, whose loss was nearly mutually exclusive with *NOTCH1* and *RBPJ* aberrations, contains genes coding for Delta-like 4 (DLL4), a NOTCH ligand; a disintegrin and metalloproteinase domain-containing protein 10 (ADAM10), a controller of NOTCH cleavage; and APH1B, a γ -secretase of NOTCH.(33) Chromosome 7 contains the gene encoding lunatic fringe (LFNG), a key Notch signaling repressor, such that its overexpression could suppress Notch signaling.(33)

Mutations of *PIK3* subunits were highly weighted negative prognostic markers in our neural network analysis; were enriched in a subset of our endpoints of advanced disease; and were markers of shorter survival on multivariable analysis. Mutations of *PIK3CA* are

activating, while those of *PIK3R* are inactivating, and both result in enhanced PI3K activity, with downstream activation of Akt and mammalian target of rapamycin (mTOR), which are associated with aggressive clinical behavior in many cancers.(34) Our neural network identified INPP4B, a known suppressor of PI3K signaling,(35) as a protein whose increased expression was strongly associated with improved outcome. The PI3K pathway also strongly converges with SMAD2/3 and P38/MKK3 pathways, which were identified as among the most enriched with negative prognostic transcripts in our neural network.(19, 20) Lastly, our IHC analysis indicated pAkt expression was associated with higher-grade features and may have utility as a prognostic marker.

Importantly, our identification of Notch and PI3K pathways' association with survival risk and disease progression does not demonstrate a causal or temporal relationship, and represents an inherent limitation of our study. We cannot prove *NOTCH1* or *PIK3* subunit mutations evolved temporally from a lower grade tumor, causing its progression. It is entirely possible oligodendrogliomas with Notch inactivation and PI3K activation are in fact distinct genetic subsets at their initiation and these tumors are more rapidly progressive. Furthermore, in addition to the retrospective nature of this study, the survival data in the TCGA is widely known to be incomplete and may affect traditional survival models. The major strengths of this study include its novel methodology, as well as the further investigation of the Notch pathway using both technologically advanced methods with clinical correlates and validation both demonstrate utility in the world of molecular glioma. Longitudinal investigation of patient cohorts with primary and recurrent tumors is needed to identify temporal

evolution.(36, 37) Future investigation will also require the elucidation of downstream targets of Notch and PI3K pathways that may drive glioma progression and gliomagenesis.

REFERENCES

1. Cancer Genome Atlas Research N, Brat DJ, Verhaak RG, et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N Engl J Med* 2015;372(26):2481-98.
2. Ostrom QT, Gittleman H, Xu J, et al. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009–2013. *Neuro-Oncology* 2016;18(suppl 5):v1-v75.
3. Louis DN, Oh W, Wiestler OD, Cavenee WK. *World Health Organization Histological Classification of Tumours of the Central Nervous System*. France: International Agency for Research on Cancer; 2016.
4. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* 2008;321(5897):1807-12.
5. Yan H, Parsons DW, Jin G, et al. IDH1 and IDH2 mutations in gliomas. *The New England journal of medicine* 2009;360(8):765-73.
6. Eckel-Passow JE, Lachance DH, Molinaro AM, et al. Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors. *The New England journal of medicine* 2015;372(26):2499-508.
7. Olar A, Wani KM, Alfaro-Munoz KD, et al. IDH mutation status and role of WHO grade and mitotic index in overall survival in grade II-III diffuse gliomas. *Acta neuropathologica* 2015.
8. Bettgowda C, Agrawal N, Jiao Y, et al. Mutations in CIC and FUBP1 contribute to human oligodendroglioma. *Science* 2011;333(6048):1453-5.
9. Ceccarelli M, Barthel FP, Malta TM, et al. Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. *Cell* 2016;164(3):550-63.
10. Aoki K, Nakamura H, Suzuki H, et al. Prognostic relevance of genetic alterations in diffuse lower-grade gliomas. *Neuro-oncology* 2017.
11. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *The New England journal of medicine* 2016;375(13):1216-9.
12. Reyes-Botero G, Dehais C, Idbah A, et al. Contrast enhancement in 1p/19q-codeleted anaplastic oligodendrogliomas is associated with 9p loss, genomic instability, and angiogenic gene expression. *Neuro Oncol* 2014;16(5):662-70.
13. Trembath D, Miller CR, Perry A. Gray zones in brain tumor classification: evolving concepts. *Advances in anatomic pathology* 2008;15(5):287-97.
14. Wesseling P, van den Bent M, Perry A. Oligodendroglioma: pathology, molecular mechanisms and markers. *Acta neuropathologica* 2015;129(6):809-27.
15. Beroukhi R, Getz G, Nghiemphu L, et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proceedings of the National Academy of Sciences of the United States of America* 2007;104(50):20007-12.
16. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013;499(7457):214-8.

17. Cooper LA, Kong J, Gutman DA, et al. Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images. *Laboratory investigation; a journal of technical methods and pathology* 2015;95(4):366-76.
18. Louis DN, Ohgaki, H., Wiestler, O.D., Cavenee, W.K. *WHO Classification of Tumours of the Central Nervous System*. 4th ed. Lyon: Intl. Agency for Research; 2007.
19. Singh AM, Reynolds D, Cliff T, et al. Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. *Cell stem cell* 2012;10(3):312-26.
20. Locatelli SL, Careddu G, Stirparo GG, et al. Dual PI3K/ERK inhibition induces necroptotic cell death of Hodgkin Lymphoma cells through IER3 downregulation. *Scientific reports* 2016;6:35745.
21. Wang NJ, Sanborn Z, Arnett KL, et al. Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* 2011;108(43):17761-6.
22. Cerami E, Gao J, Dogrusoz U, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer discovery* 2012;2(5):401-4.
23. Kopan R, Ilagan MX. The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell* 2009;137(2):216-33.
24. Yap LF, Lee D, Khairuddin A, et al. The opposing roles of NOTCH signalling in head and neck cancer: a mini review. *Oral diseases* 2015;21(7):850-7.
25. Rampias T, Vgenopoulou P, Avgeris M, et al. A new tumor suppressor role for the Notch pathway in bladder cancer. *Nature medicine* 2014;20(10):1199-205.
26. Agrawal N, Frederick MJ, Pickering CR, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* 2011;333(6046):1154-7.
27. Radtke F, Raj K. The role of Notch in tumorigenesis: oncogene or tumour suppressor? *Nat Rev Cancer* 2003;3(10):756-67.
28. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 2014;505(7484):495-501.
29. Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science* 2011;333(6046):1157-60.
30. Rebay I, Fleming RJ, Fehon RG, et al. Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor. *Cell* 1991;67(4):687-99.
31. Xie Q, Wu Q, Kim L, et al. RBPJ maintains brain tumor-initiating cells through CDK9-mediated transcriptional elongation. *The Journal of clinical investigation* 2016;126(7):2757-72.
32. Kulic I, Robertson G, Chang L, et al. Loss of the Notch effector RBPJ promotes tumorigenesis. *The Journal of experimental medicine* 2015;212(1):37-52.
33. UniProt: a hub for protein information. *Nucleic acids research* 2015;43(Database issue):D204-12.

34. Thorpe LM, Yuzugullu H, Zhao JJ. PI3K in cancer: divergent roles of isoforms, modes of activation and therapeutic targeting. *Nat Rev Cancer* 2015;15(1):7-24.
35. Gewinner C, Wang ZC, Richardson A, et al. Evidence that Inositol polyphosphate 4-phosphatase type II is a tumor suppressor that inhibits PI3K signaling. *Cancer cell* 2009;16(2):115-25.
36. Kim H, Zheng S, Amini SS, et al. Whole-genome and multisector exome sequencing of primary and post-treatment glioblastoma reveals patterns of tumor evolution. *Genome research* 2015;25(3):316-27.
37. Johnson BE, Mazor T, Hong C, et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* 2014;343(6167):189-93.

TABLES / FIGURES

Fig. 1. A. Neural network risk factors. A nonlinear Cox proportional hazards model was trained using a neural network to model survival in oligodendrogliomas using clinical, genetic and proteomic factors. Prognostic significance of each feature was assessed by determining how its changes impact prognosis. Positive scores indicate a negative impact on survival (red) while negative scores (blue) suggest a positive impact. The boxplot contains the top 10 factors ranked by median prognostic importance; complete results in Datafile S1. **B. Gene set enrichment analysis of Notch pathway members.** A separate model based on mRNA expression weighed the prognostic significance of individual transcripts and used this data in a gene-set-enrichment analysis to identify pathways associated with prognosis. The canonical Notch pathway was highly enriched with significantly negatively scored transcripts (i.e. darker blue signifies negative scores). Increased expression of downstream targets, including *HES1*, *HES5*, and *HEY1*, were associated with improved prognosis. This model demonstrates Notch signaling inactivation is associated with poor prognosis.

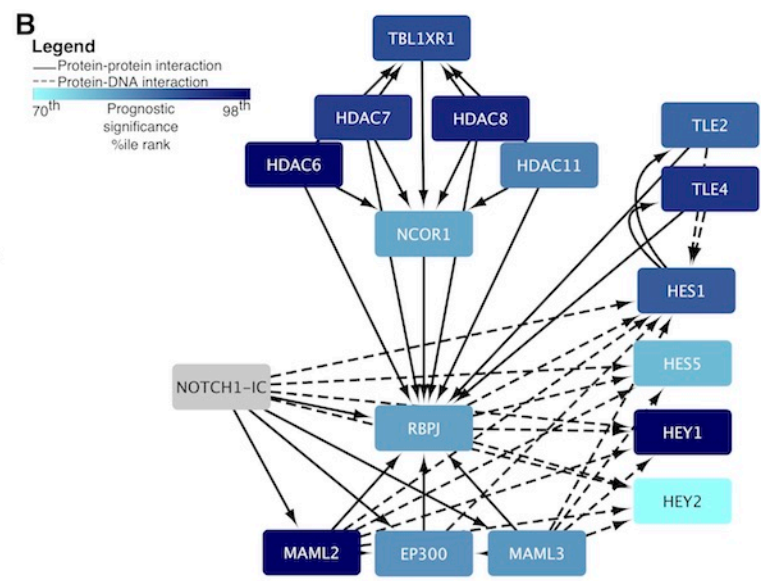
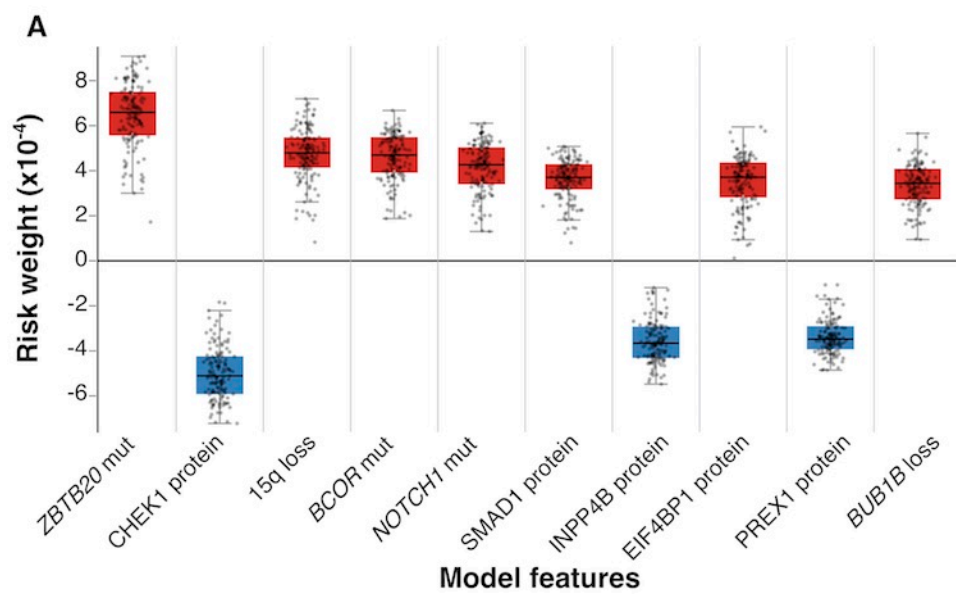


Fig. 2. Markers of disease progression in oligodendroglioma **A.** T1-weighted axial MR images with gadolinium contrast demonstrating CE- (left) and CE+ (right) features of oligodendroglioma from The Cancer Imaging Archive. **B.** Kaplan-Meier plots of overall survival (OS) for CE- vs. CE+. **C.** Progression-free survival (PFS) for CE- vs. CE+. **D.** Visual representation of a tumor heatmap showing regions of interest of cell density, with a schematic diagram of the nearest-neighbor algorithm. **E.** OS for cellular density (less vs. more dense). **F.** PFS for less vs. more dense. **G.** High Ki-67 proliferation index visualized with IHC. **H.** Linear regression of *MKI67* expression and Ki-67 proliferation index approximated by IHC. **I.** OS for high vs. low *MKI67*. **J.** PFS for high vs. low *MKI67*. *P* values for survival plots determined using log-rank tests.

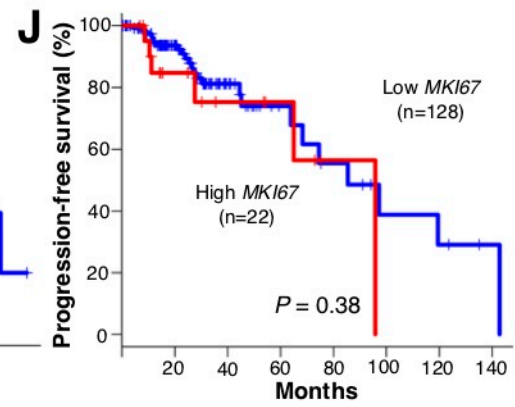
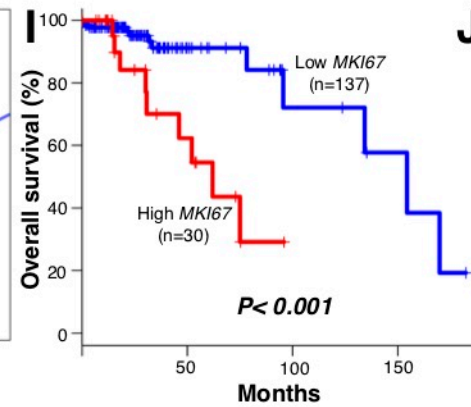
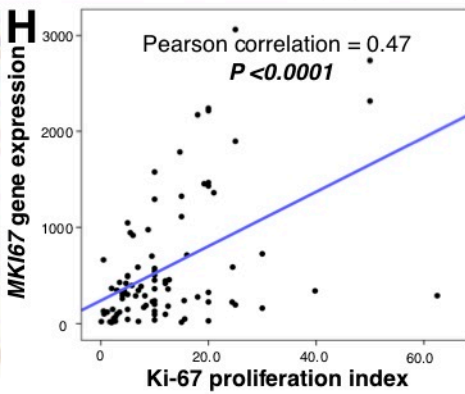
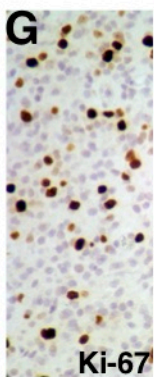
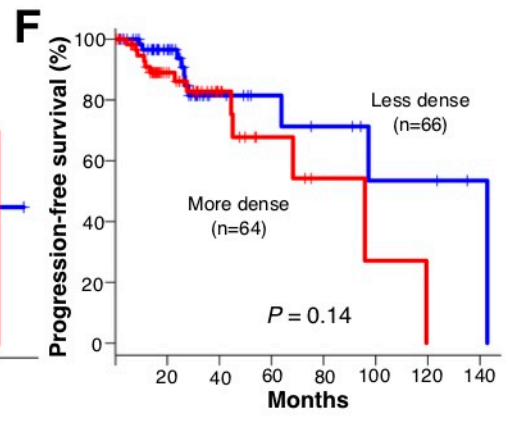
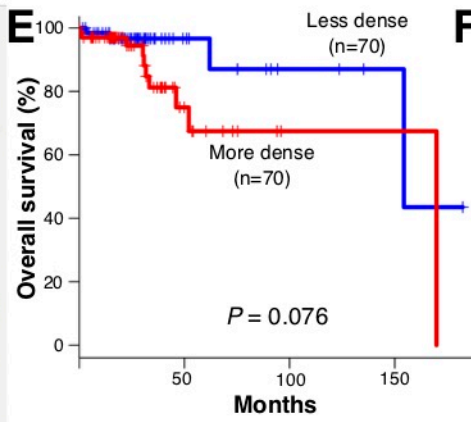
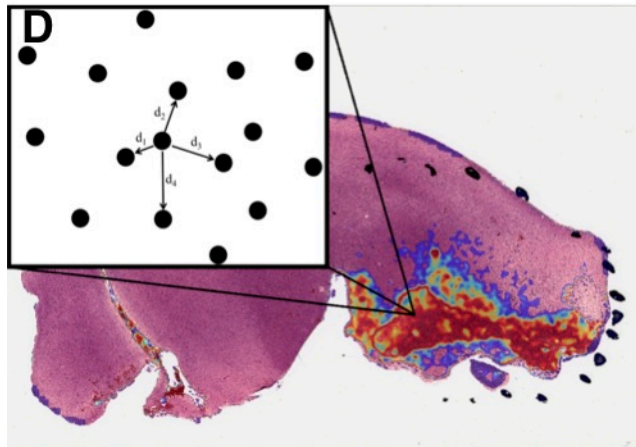
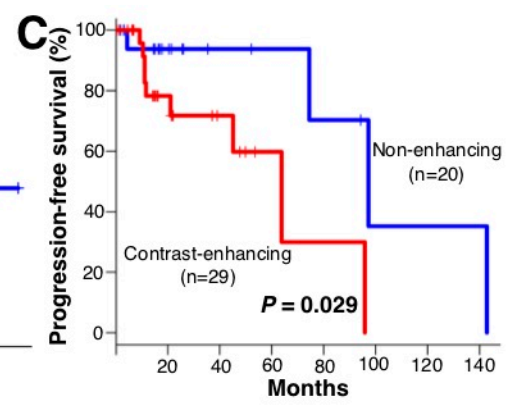
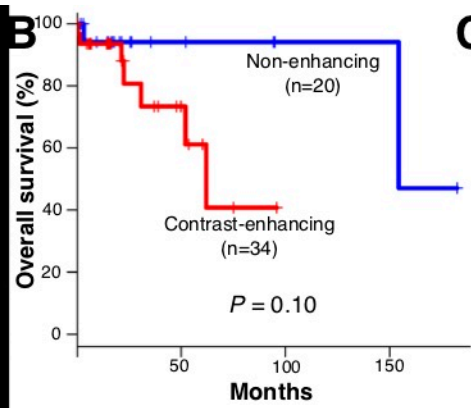
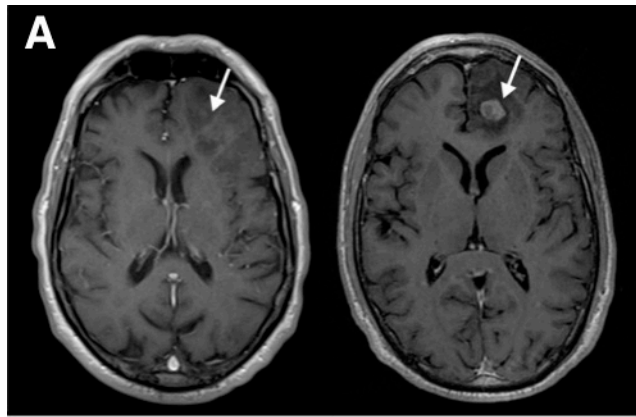


Fig. 3. Genetic alterations associated with advanced disease progression **A.** Waterfall plot illustrating the mutational landscape of oligodendrogliomas based on radiographic features of progression. **B.** Boxplots demonstrating nearest-neighbor validation, and differential 1/nearest-neighbor distances in key genetic alterations of oligodendroglioma. **C.** Boxplots for differential *MKI67* expression in key genetic alterations of oligodendroglioma. *P* values determined using Wilcoxon rank sum tests.

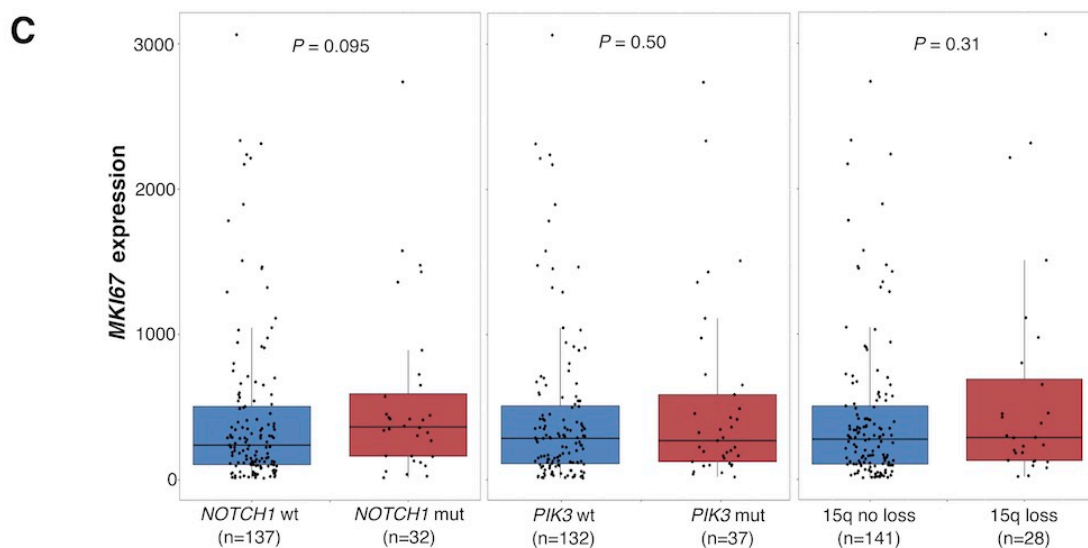
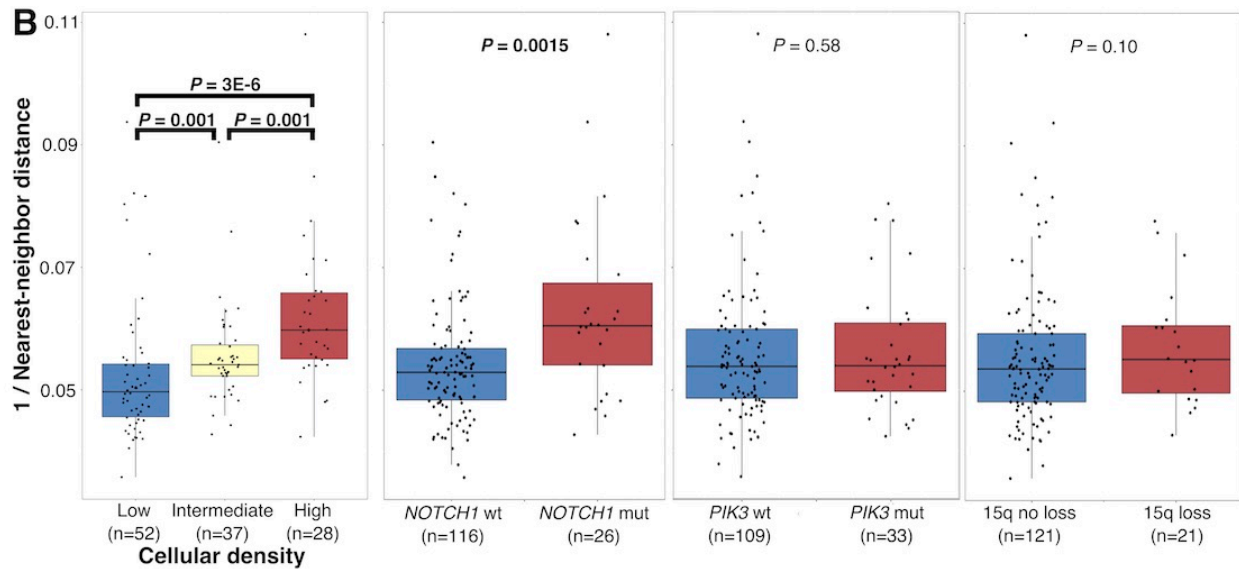
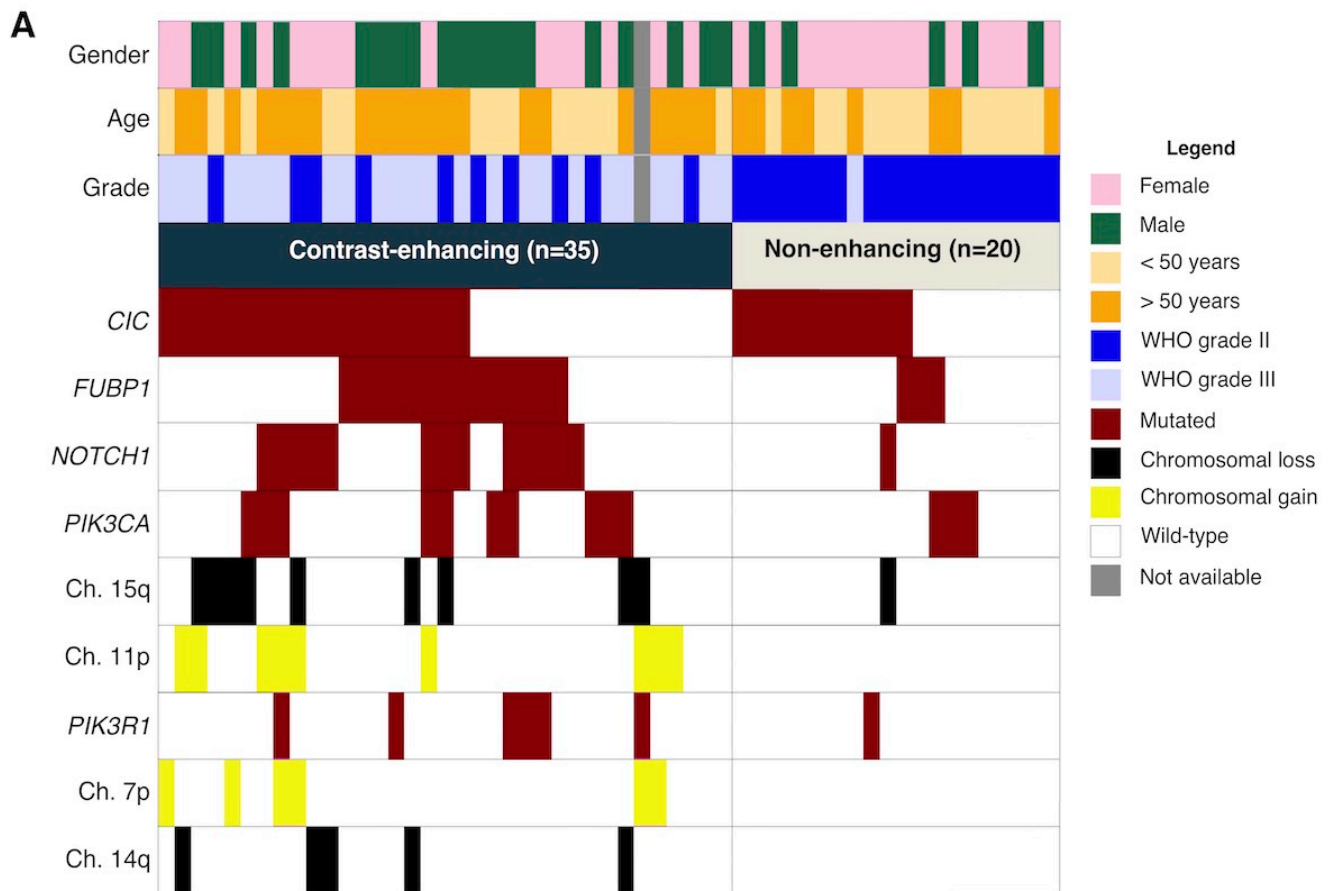
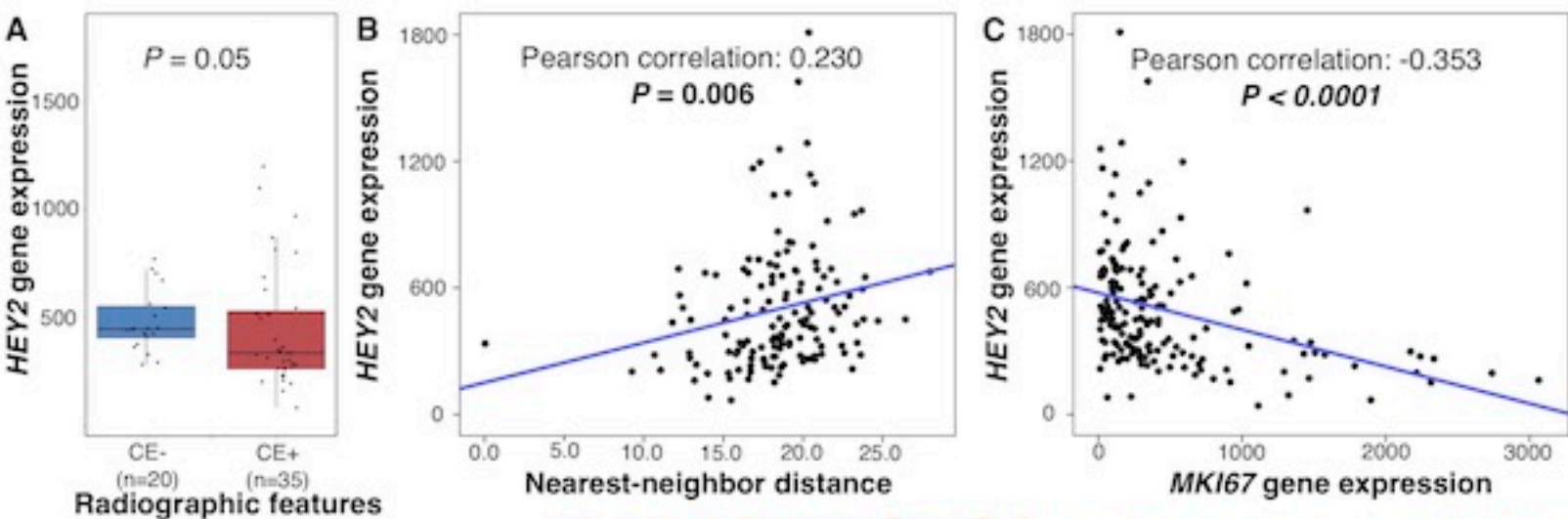


Fig. 4. *HEY2* associations with advanced disease and validation cohort. **A.** Boxplots demonstrating differential *HEY2* gene expression in CE- and CE+; *P* value determined using Wilcoxon rank-sum test. **B.** Linear regression of *HEY2* gene expression and nearest-neighbor distance, demonstrating positive correlation. **C.** Linear regression of *HEY2* and *MKI67* expression, demonstrating negative correlation. *P* values from Pearson correlation. **D.** IHC showing high Ki-67 proliferation index (25%) (bar, 250 μm), with corresponding absent *HEY2* expression (bar, 100 μm) and high pAkt expression (bar 100 μm). **E.** IHC showing low Ki-67 proliferation index (1%) (bar, 250 μm), with corresponding high *HEY2* expression (bar, 100 μm) and absent pAkt expression (bar, 100 μm). **F.** *HEY2* and pAKT IHC intensity as related to cellular density and Ki-67 proliferation indices.



Radiographic features

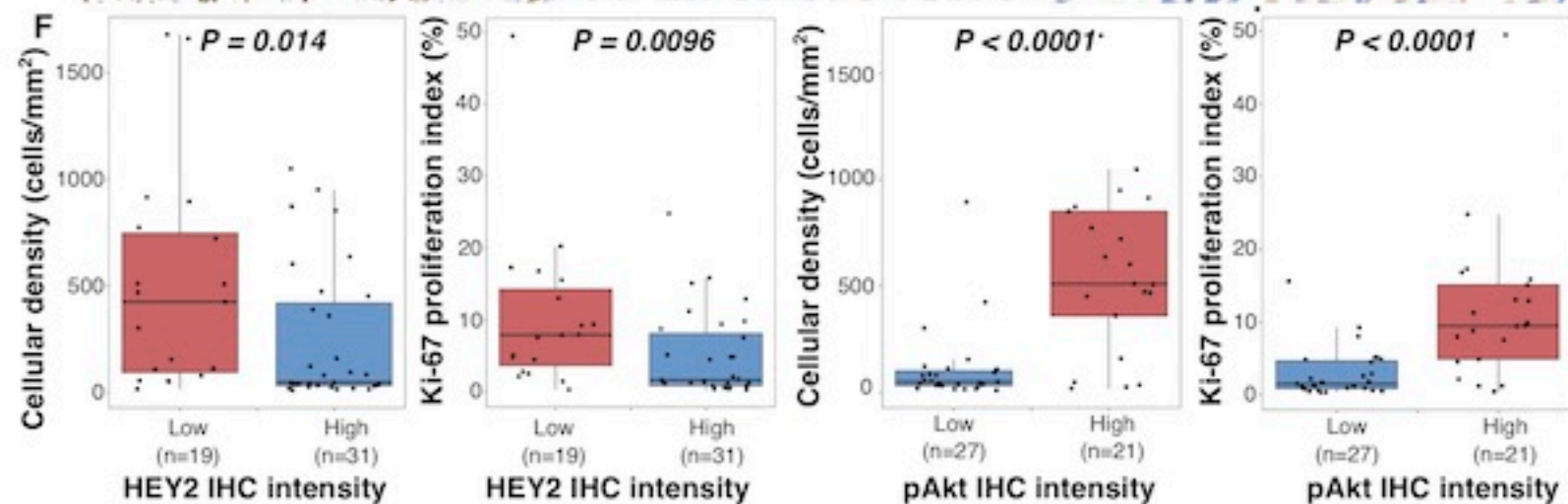
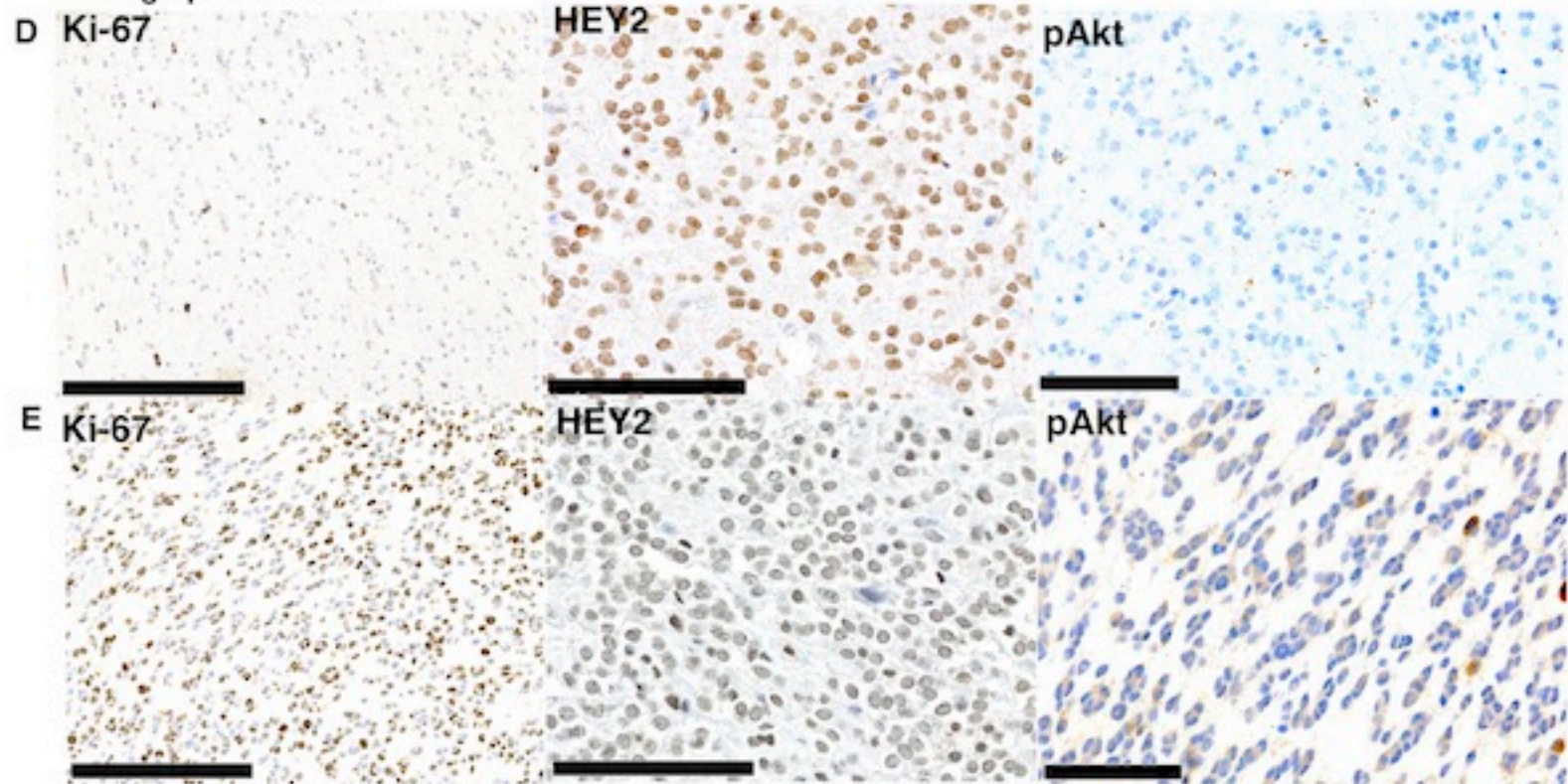


Table 1. Patient demographics. Clinical characteristics of patients from The Cancer Genome Atlas database with confirmed diagnosis of oligodendroglioma (i.e. *IDH*-mutant, 1p19q co-deleted glioma).

Characteristic	Total (N=169)
Original histologic diagnosis (WHO 2007) – no. (%)	
Oligodendroglioma	
Grade II	62 (36.7)
Grade III	55 (32.5)
Oligoastrocytoma	
Grade II	17 (10.1)
Grade III	13 (7.7)
Astrocytoma	
Grade II	2 (1.2)
Grade III	2 (1.2)
Age at diagnosis (yrs)	
Mean \pm SD	45.8 \pm 12.8
Range	17–75
Male sex – no. (%)	84 (49.7)
White race – no./total no. (%)	155/164 (94.5)
Extent of resection – no./total no. (%)	
Open biopsy	1/164 (0.6)
Subtotal resection	59/164 (36.0)
Gross total resection	104/164 (63.4)
Tumor location – no./total no. (%)	
Frontal lobe	122/166 (73.5)
Occipital lobe	3/166 (1.8)
Parietal lobe	14/166 (8.4)
Temporal lobe	27/166 (16.3)
Laterality – no/total no. (%)	
Left	79/168 (47.0)
Midline	3/168 (1.8)
Right	86/168 (51.2)

Table 2. Overall survival tables. Cox proportional hazard models for overall survival (OS). Multivariable analysis of OS adjusted for grade and age.

Legend: Mut: mutation; exp.: expression; *Significant on univariate analysis; †significant on multivariate analysis; ‡gene expression on a log2 scale, such that the hazard ratio is for each doubling of gene expression.

Predictor	OS Hazard Ratio (95% conf. interval)	P-value	Adjusted OS Hazard Ratio (95% conf. interval)	P-value
*Age (per 10 yrs)	3.64 (2.16-6.11)	<0.0001	—	—
*Grade III (vs. II)	6.61 (2.08-20.95)	0.013	—	—
**‡ <i>MKI67</i> exp.	1.58 (1.17-2.14)	0.0029	1.12 (0.84-1.50)	0.42
<i>NOTCH1</i> mut.	1.71 (0.65-4.50)	0.28	1.10 (0.37-3.27)	0.87
† <i>PIK3</i> mut.	1.97 (0.78-4.97)	0.15	3.11 (1.02-9.47)	0.045
<i>RBPJ</i> + <i>NOTCH1</i> mut.	1.81 (0.71-4.61)	0.210	0.85 (0.30-2.40)	0.76
‡ <i>HES1</i> exp.	0.60 (0.35-1.05)	0.071	0.86 (0.47-1.56)	0.611
‡ <i>HES2</i> exp.	1.01 (0.93-1.10)	0.76	1.00 (0.89-1.12)	0.954
†‡ <i>HES5</i> exp.	0.82 (0.65-1.03)	0.086	0.74 (0.57-0.96)	0.024
**‡ <i>HEY1</i> exp.	0.34 (0.18-0.64)	0.0009	0.86 (0.38-1.95)	0.72
**‡ <i>HEY2</i> exp.	0.35 (0.21-0.60)	0.0001	0.79 (0.37-1.68)	0.54

Table 3. Progression-free survival. Cox proportional hazard models for progression-free survival (PFS). Multivariable analysis of PFS adjusted for grade.

Legend: Mut: mutation; exp.: expression; *Significant on univariate analysis; †significant on multivariate analysis; ‡gene expression on a log2 scale, such that the hazard ratio is for each doubling of gene expression.

Predictor	PFS Hazard Ratio (95% conf. interval)	P-value	Adjusted PFS Hazard Ratio (95% conf. interval)	P-value
*Grade III (vs. II)	2.24 (1.01-4.94)	0.046	—	—
<i>NOTCH1</i> mut.	2.07 (0.93-4.60)	0.091	1.52 (0.66-3.53)	0.33
<i>PIK3</i> mut.	1.91 (0.87-4.23)	0.11	1.98 (0.89-4.40)	0.092
*<i>RBPJ</i> + <i>NOTCH1</i> mut.	2.47 (1.14-5.34)	0.021	1.86 (0.82-4.20)	0.13
‡ <i>HES1</i> exp.	0.75 (0.49-1.15)	0.18	0.70 (0.45-1.10)	0.118
‡ <i>HES2</i> exp.	1.01 (0.96-1.08)	0.67	1.02 (0.95-1.10)	0.638
†‡ <i>HES5</i> exp.	0.83 (0.68-1.02)	0.076	0.86 (0.71-1.04)	0.120
*†‡<i>HEY1</i> exp.	0.41 (0.23-0.72)	0.0022	0.475 (0.26-0.88)	0.018
‡ <i>HEY2</i> exp.	0.87 (0.53-1.44)	0.59	0.96 (0.57-1.61)	0.869

Table S1A. Gene names and frequency of mutations for oligodendroglioma in the TCGA database.

Gene Name	Frequency of Mutations
IDH1	154
CIC	104
FUBP1	45
PI3K comb.	37
NOTCH1	32
PIK3CA	26
ZBTB20	16
IDH2	15
PIK3R1	12
ARID1A	11
NF1	8
SMARCA4	7
TP53	7
CREBZF	6
TCF12	6
ANKRD30A	5
RBPJ	5
ATRX	4
MYH4	4
SOX4	4
CDKN2C	3
EGFR	3
EMG1	3
MYH8	3
SLC6A3	3
ARL6	2
CXorf22	2
KEL	2
KPRP	2
KRT15	2
LHFPL1	2

Gene Name	Frequency of Mutations
NRAS	2
OR4S2	2
OR5D18	2
PDGFRA	2
SEMA3E	2
SPANXD	2
TRPA1	2
UGT2A3	2
VSIG4	2
ZNF709	2
CDH18	1
DCAF12L2	1
KRAS	1
LZTR1	1
MED9	1
MGAT4C	1
OPRK1	1
OR2A12	1
OR2A25	1
OR4P4	1
OR8K3	1
OR9G1	1
P2RY11	1
PTEN	1
PTPN11	1
RAP2C	1
RB1	1
RPL5	1
SPTA1	1
SSTR4	1

Gene Name	Frequency of Mutations
ST3GAL6	1
STK19	1
TEAD3	1
TREML2	1
TRIM58	1
TYRP1	1
ZPBP	1

Supplemental Table 1B. Gain and loss of whole chromosome arms of oligodendroglioma within the TCGA database.

Chromosome	Loss	Gain
1p	166	0
1q	12	1
4p	45	0
4q	50	0
7p	0	16
7q	0	24
9p	21	2
9q	15	3
11p	0	19
11q	0	24
12p	5	4
13q	28	0
14q	20	1
15q	27	0
17p	2	7
17q	1	8
18p	31	0
18q	33	0
19p	7	57
19q	167	0
21q	2	9
22q	4	17
Xp	33	0
Xq	32	1

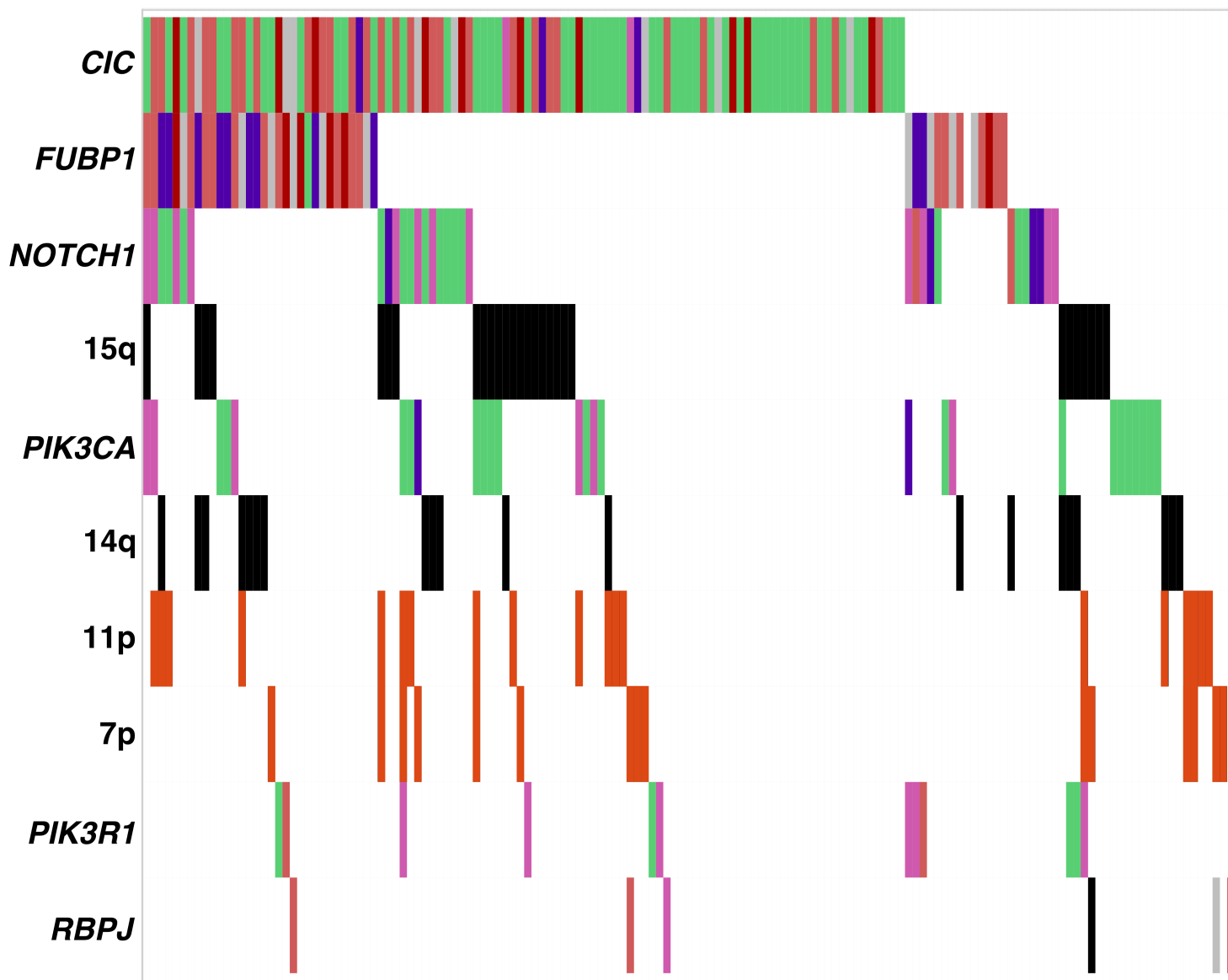
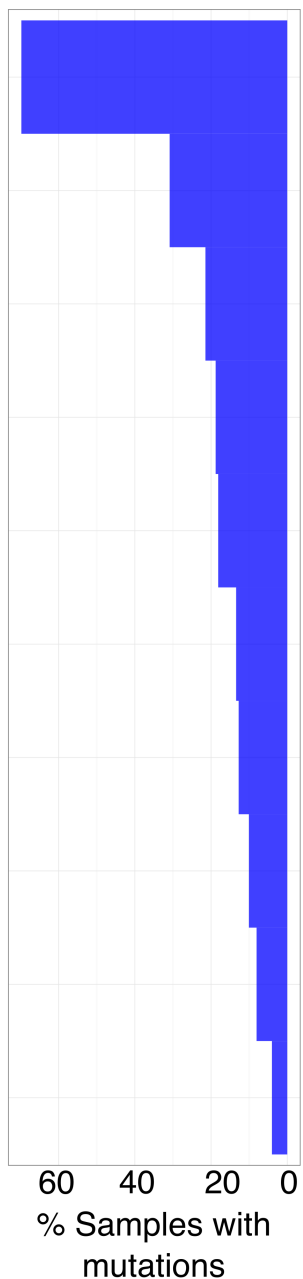
Table S2. Patient demographics. Clinical characteristics of patients with pre-operative radiographic imaging from The Cancer Genome Atlas database, with confirmed diagnosis of oligodendroglioma (i.e. IDH-mutant, 1p19q co-deleted glioma).

Characteristic	Total (N=55)
Histological subtype – no. (%)	
Oligodendroglioma	
Grade II	23 (41.8)
Grade III	24 (43.6)
Oligoastrocytoma	
Grade II	5 (9.1)
Grade III	1 (1.8)
Astrocytoma	
Grade II	1 (1.8)
Grade III	0 (0.0)
Age at diagnosis (yrs)	
Mean	49.4 ± 13.2
Range	20–75
Radiographic features present – no./total no. (%)	
- Contrast-enhanced	20/55 (36.4)
+ Contrast-enhanced	35/55 (63.6)
Tumor location – no./total no. (%)	
Frontal lobe	40/55 (72.7)
Parietal lobe	7/55 (12.7)
Temporal lobe	3/55 (5.5)
Other	5/55 (9.1)

Predictor	OS Hazard Ratio (95% conf. interval)	P-value	Adjusted OS Hazard Ratio (95% conf. interval)	P-value
*Age (per 10 yrs)	3.64 (2.16-6.11)	<0.0001	—	—
Histologic astrocytoma vs. oligodendroglioma	2.94 (0.046-2.70)	0.31	—	—
Extent of resection (GTR vs. less than GTR)	0.65 (0.24-1.68)	0.37	—	—
*Grade III (vs. II)	6.61 (2.08-20.95)	0.013	—	—
**MKI67 exp.	1.58 (1.17-2.14)	0.0029	1.12 (0.84-1.50)	0.42
<i>ATRX</i> mut.	NA	NA	<0.0001	0.998
<i>CIC</i> mut.	0.65 (0.27-1.55)	0.33	0.44 (0.17-1.15)	0.095
<i>FUBP1</i> mut.	1.70 (0.63-4.53)	0.31	2.60 (0.84-8.00)	0.096
<i>NOTCH1</i> mut.	1.71 (0.65-4.50)	0.28	1.10 (0.37-3.27)	0.87
†PIK3 mut.	1.97 (0.78-4.97)	0.15	3.11 (1.02-9.47)	0.045
<i>RBPJ</i> + <i>NOTCH1</i> mut.	1.81 (0.71-4.61)	0.210	0.85 (0.30-2.40)	0.76
<i>TP53</i> mut.	1.36 (0.18-10.33)	0.77	1.16 (0.14-9.65)	0.893
7p gain	1.90 (0.42-8.52)	0.44	1.23 (0.26-5.68)	0.795
11p gain	2.08 (0.58-7.48)	0.26	0.60 (0.16-2.33)	0.463
14q loss	3.04 (0.94-9.77)	0.063	1.49 (0.39-5.62)	0.559
*15q loss	3.52 (1.41-8.82)	0.007	1.48 (0.51-4.30)	0.47
‡ <i>HES1</i> exp.	0.60 (0.35-1.05)	0.071	0.86 (0.47-1.56)	0.611
‡ <i>HES2</i> exp.	1.01 (0.93-1.10)	0.76	1.00 (0.89-1.12)	0.954
†‡ <i>HES5</i> exp.	0.82 (0.65-1.03)	0.086	0.74 (0.57-0.96)	0.024
*‡HEY1 exp.	0.34 (0.18-0.64)	0.0009	0.86 (0.38-1.95)	0.72
*‡HEY2 exp.	0.35 (0.21-0.60)	0.0001	0.79 (0.37-1.68)	0.54

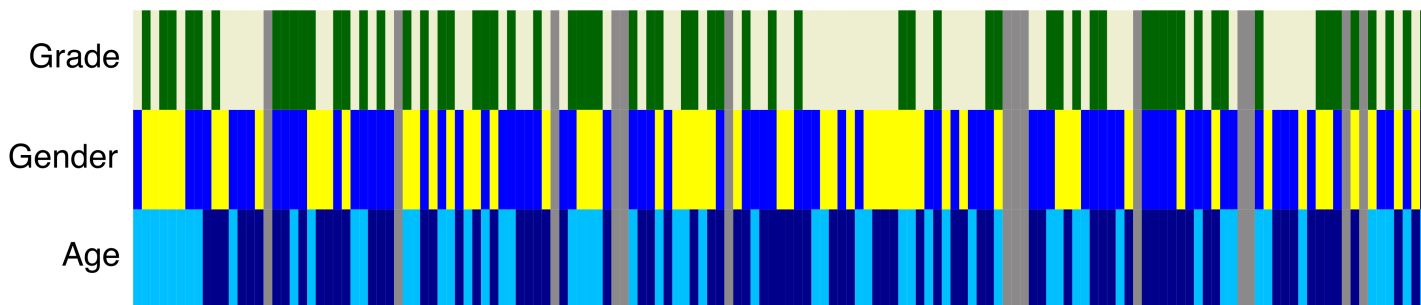
	PFS Hazard Ratio (95% conf. interval)	P-value	Adjusted PFS Hazard Ratio (95% conf. interval)	P-value
Age (per 10 yrs)	1.12 (0.86-1.66)	0.28	—	—
*Histologic astrocytoma vs. oligodendroglioma	4.77 (1.09-20.78)	0.038	—	—
Extent of resection (GTR vs. less than GTR)	0.66 (0.30-1.44)	0.30	—	—
*Grade III (vs. II)	2.24 (1.01-4.94)	0.046	—	—
‡ <i>MKI67</i> exp.	1.04 (0.84-1.29)	0.71	0.97 (0.78-1.21)	0.81
<i>ATRX</i> mut.	2.93 (0.38-22.67)	0.30	2.09 (0.27-16.3)	0.483
<i>CIC</i> mut.	0.61 (0.29-1.27)	0.18	0.58 (0.28-1.20)	0.140
*<i>FUBP1</i> mut.	2.48 (1.14-5.38)	0.022	2.14 (0.98-4.69)	0.058
<i>NOTCH1</i> mut.	2.07 (0.93-4.60)	0.091	1.52 (0.66-3.53)	0.33
<i>PIK3</i> mut.	1.91 (0.87-4.23)	0.11	1.98 (0.89-4.40)	0.092
*<i>RBPJ</i> + <i>NOTCH1</i> mut.	2.47 (1.14-5.34)	0.021	1.86 (0.82-4.20)	0.13
<i>TP53</i> mut.	2.57 (0.60-11.03)	0.21	2.57 (0.59-11.26)	0.211
7p gain	0.55 (0.07-4.11)	0.56	0.42 (0.06-3.19)	0.404
11p gain	1.31 (0.38-4.44)	0.67	1.19 (0.35-4.08)	0.779
*†14q loss	3.70 (1.49-9.20)	0.010	3.90 (1.56-9.74)	0.0035
15q loss	1.75 (.073-4.19)	0.21	1.72 (0.70-4.22)	0.239
‡ <i>HES1</i> exp.	0.75 (0.49-1.15)	0.18	0.70 (0.45-1.10)	0.118
‡ <i>HES2</i> exp.	1.01 (0.96-1.08)	0.67	1.02 (0.95-1.10)	0.638
†‡ <i>HES5</i> exp.	0.83 (0.68-1.02)	0.076	0.86 (0.71-1.04)	0.120
*†‡<i>HEY1</i> exp.	0.41 (0.23-0.72)	0.0022	0.475 (0.26-0.88)	0.018
‡ <i>HEY2</i> exp.	0.87 (0.53-1.44)	0.59	0.96 (0.57-1.61)	0.869

Table S3. Complete survival tables. Cox proportional hazard models for overall survival (OS) and progression-free survival (PFS); multivariate OS adjusted for grade and age; multivariate PFS adjusted for grade. GTR: gross total resection; mut: mutation; exp.: expression; *Significant on univariate analysis; †significant on multivariate analysis; ‡gene expression on a log₂ scale, such that the hazard ratio is for each doubling of gene expression.



Alteration Type

- Nonsense
- Frame shift ins
- Frame shift del
- InFrame del
- Missense
- Splice site
- Gain
- Loss

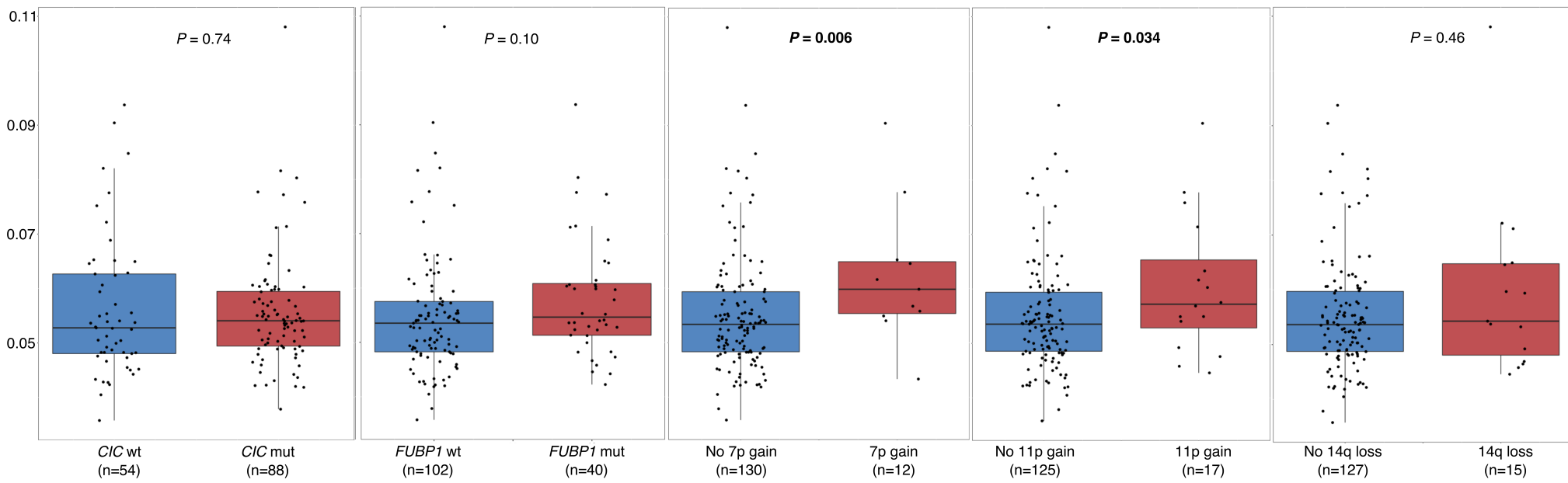


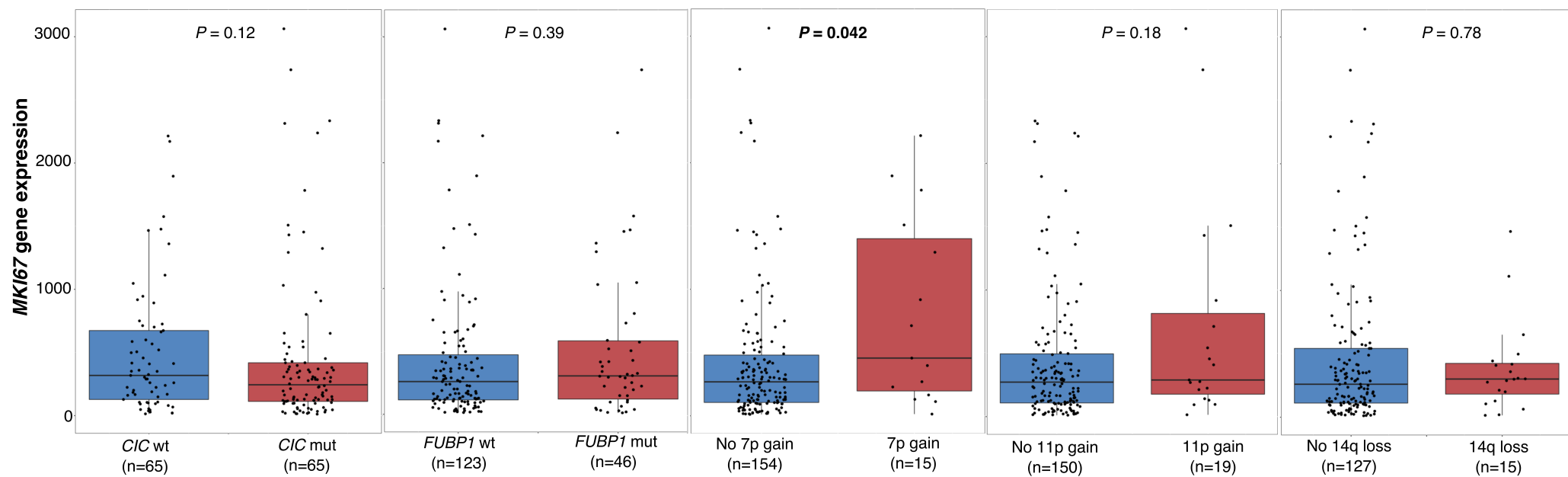
Clinical Data

- Grade II
- Grade III
- Male
- Female
- <50 years
- >50 years
- NA

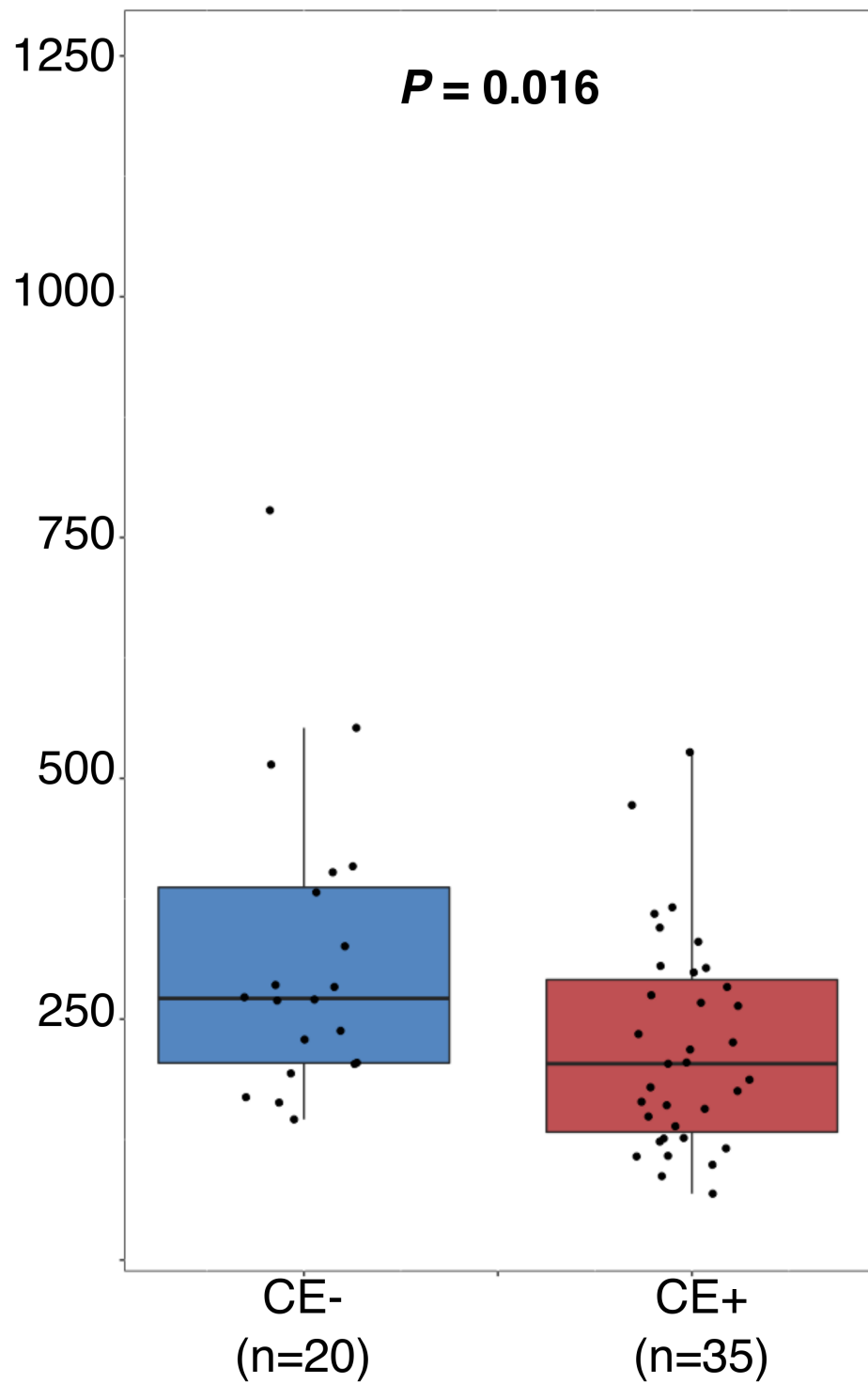
Total samples n = 149

1 / Nearest-neighbor distance

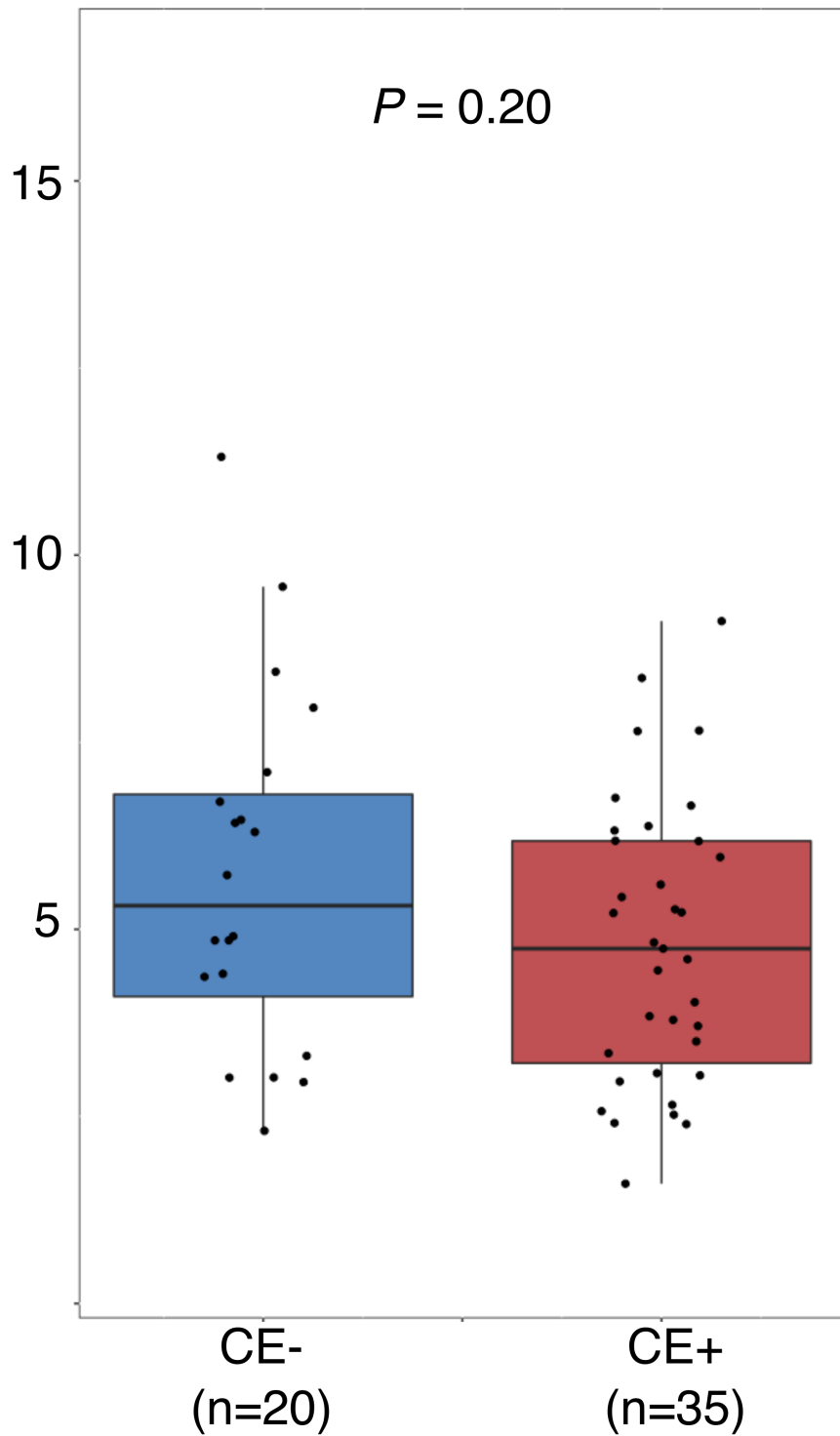


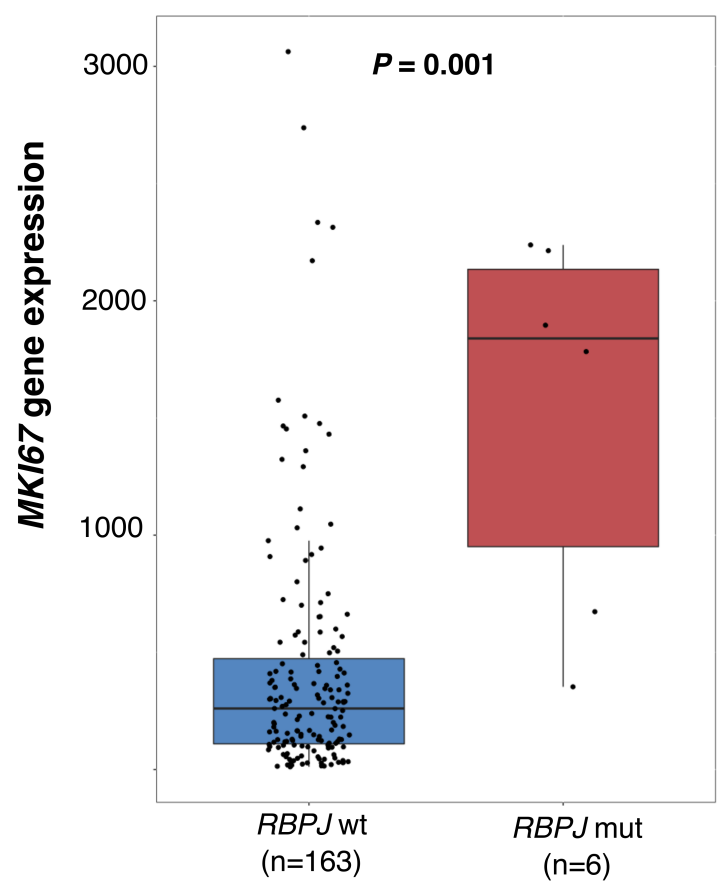
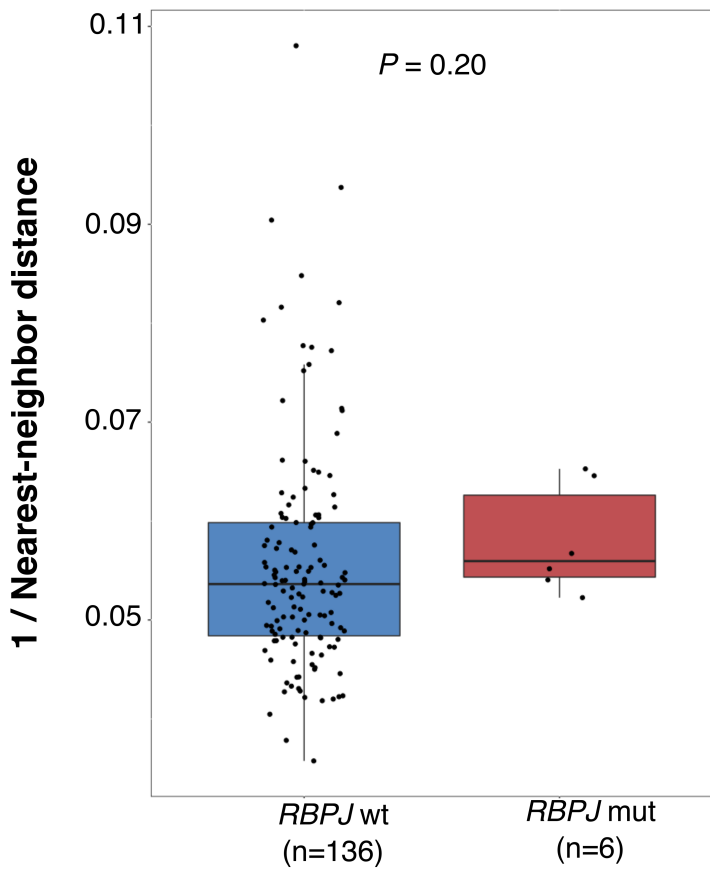
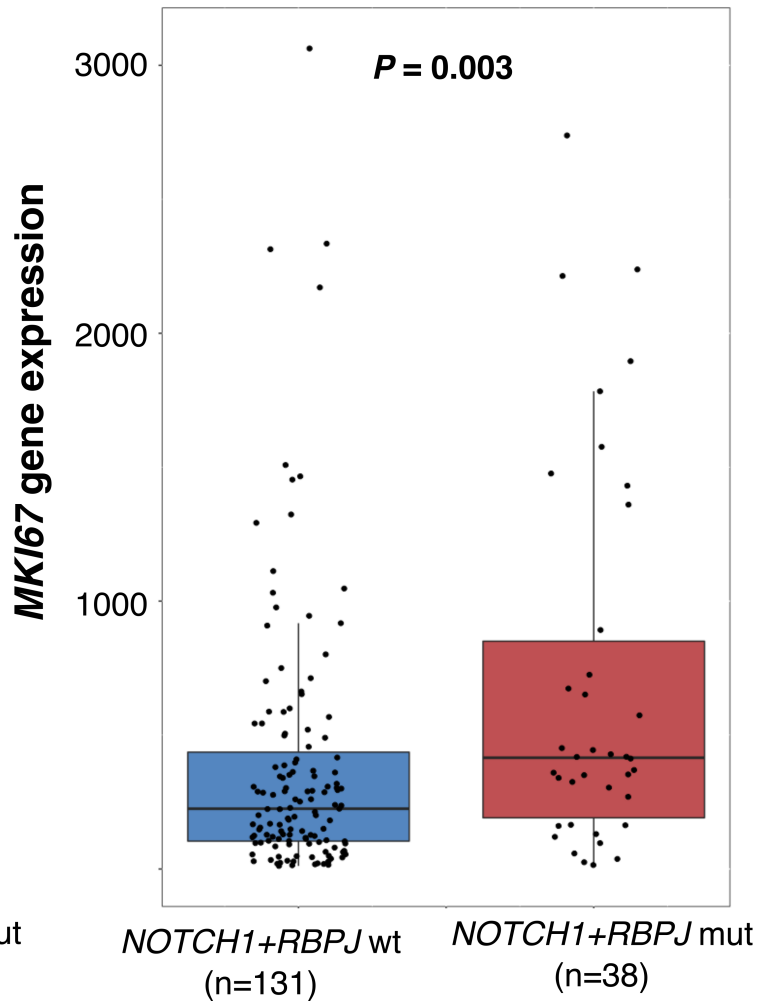
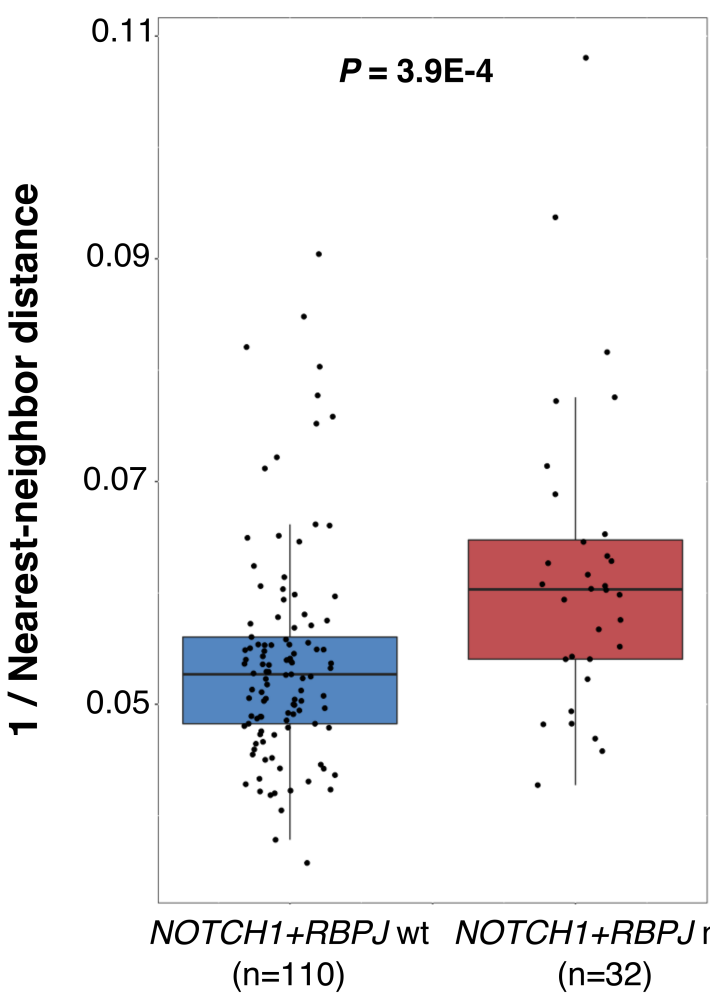


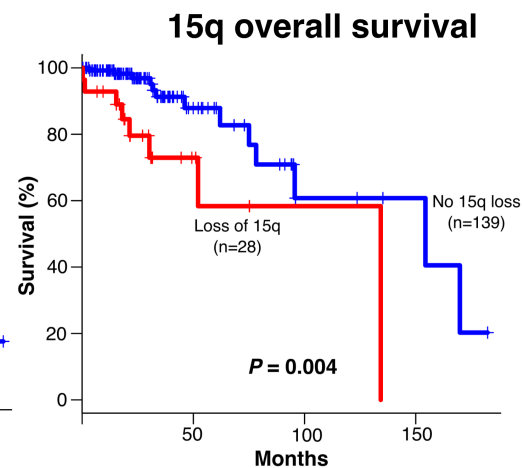
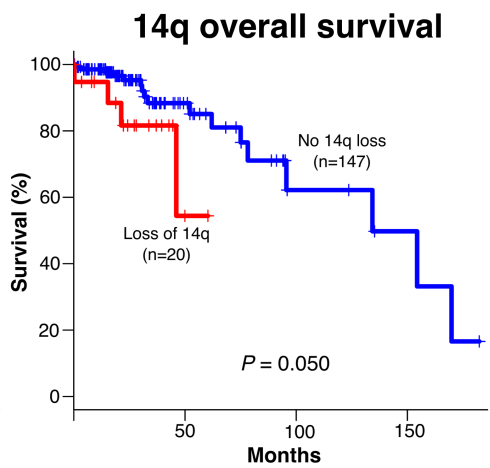
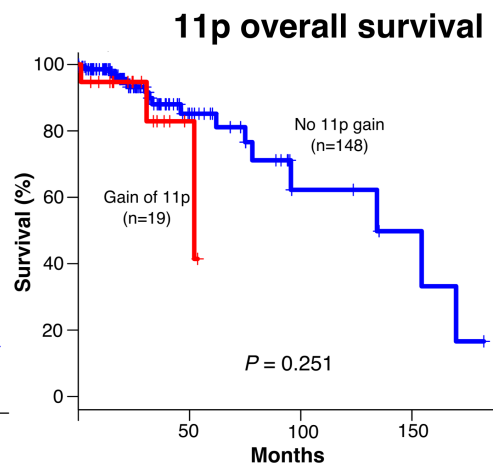
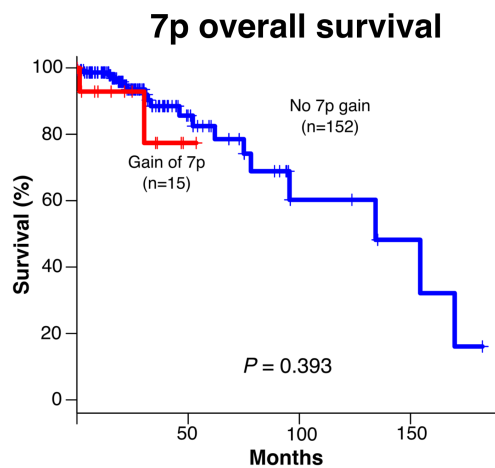
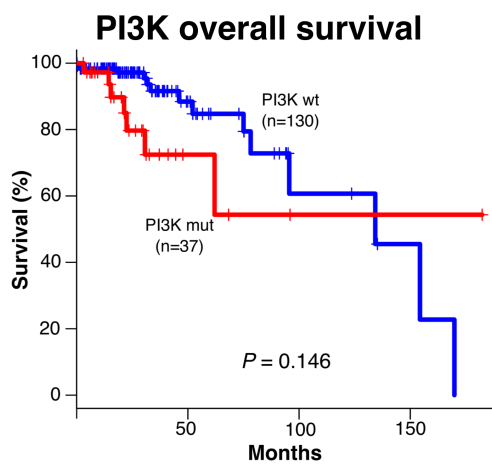
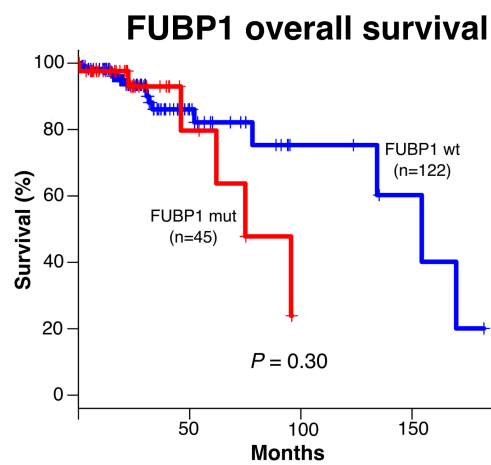
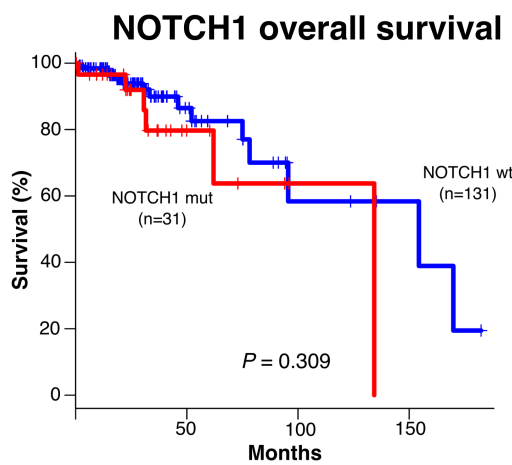
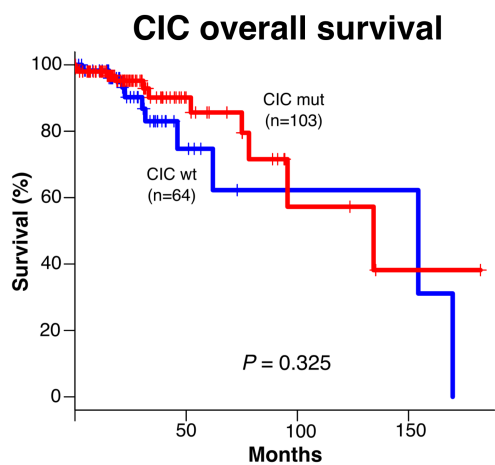
***HES1* gene expression**



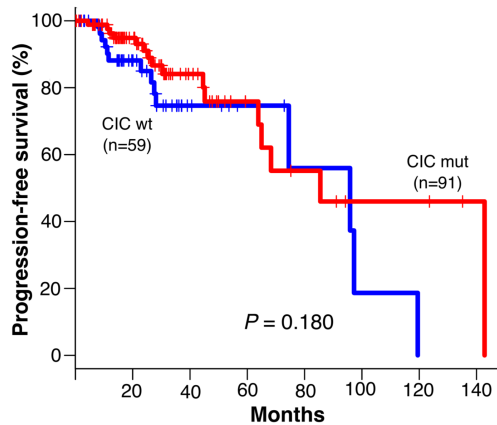
***HEY1* gene expression ($\times 10^3$)**



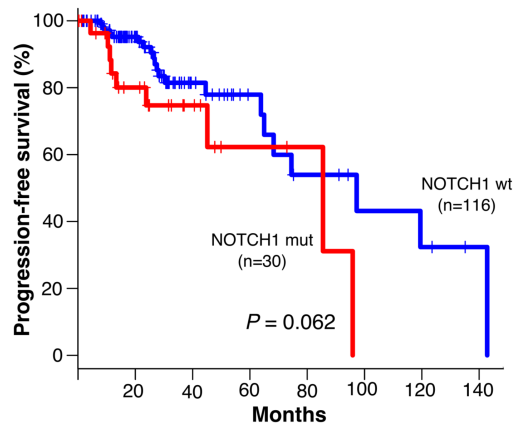
A**B**



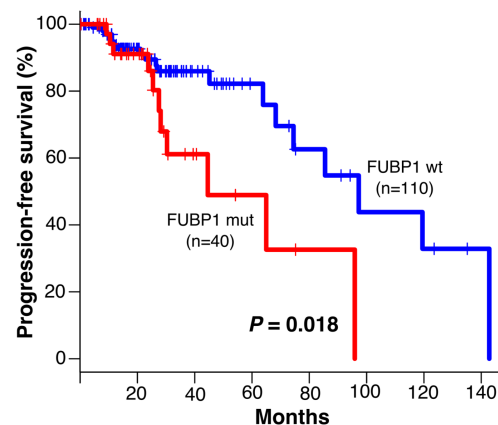
CIC PFS



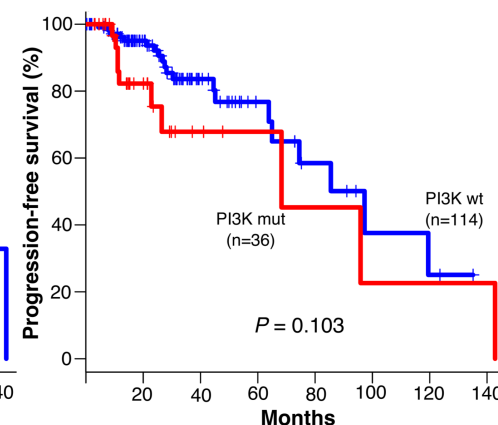
NOTCH1 PFS



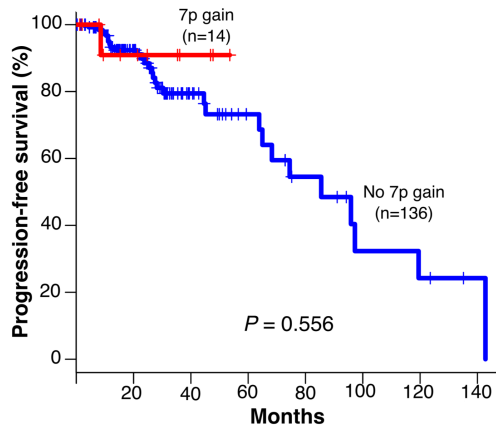
FUBP1 PFS



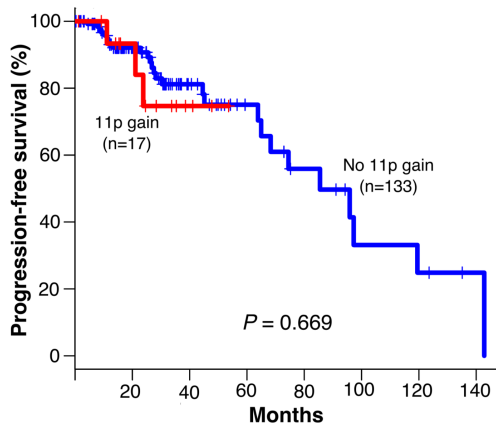
PI3K PFS



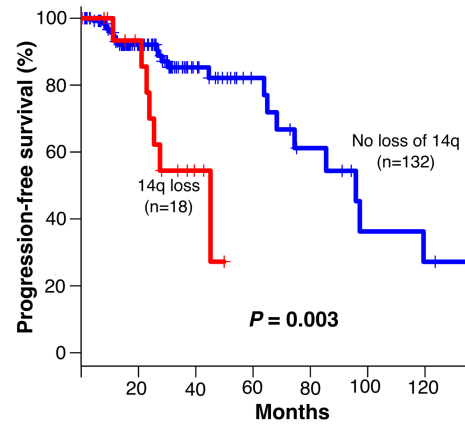
7p PFS



11p PFS



14q PFS



15q PFS

