Entropy Analysis of m6A Transcriptomic Profile in Entacapone-Treated Glioblastoma Stem Cells

Brown University Computational Biology Undergraduate Senior Honors Thesis

Author: Jonathan Arditi

Advisor: Nikos Tapinos, MD, PhD

2022

Abstract

Introduction: Glioblastoma multiforme (GBM) is an invariably fatal malignancy of the brain. Tumors are highly migratory and resistant to treatment due to the presence of glioma stem cells (GSCs) in the tumor mass. The RNA modification N^6-methyladenosine (m6A), and the enzymes that remove it, have recently been implicated in GSC tumorigenesis, self-renewal and proliferation. Entacapone, an FDA-approved catechol-*O*-methyltransferase inhibitor for Parkinson's disease, has recently been shown to inhibit FTO, an m6A demethylase, and ameliorate the malignant properties of GSCs. The effects of FTO inhibition with entacapone on specific transcripts, and the mechanism relating these changes to the phenotypic transformation in GSCs, is not yet known.

Methods: GSCs were isolated and cultured from GBM tissue samples resected during surgery. Cells were treated with either entacapone or dimethylsulfoxide (DMSO) as an inert control, and RNA was isolated from the treated cells. The locations of all m6A sites in the transcriptome along with their change in abundance were determined with single-nucleotide resolution m6A sequencing, and transcripts were characterized by their change in m6A entropy, a novel analytical technique developed for this study.

Results: Due to the small site abundance values in this dataset, changes in entropy accurately capture changes in m6A abundance. The vast majority of sites occurred near stop codons, though the transcripts with most extreme entropy change were relatively enriched for sites in the 5'UTR. Gene set enrichment analyses of these transcripts showed a preponderance of nuclear-localizing, DNA-binding, transcription-regulating genes, i.e., transcription factors. Important genes and pathways in cancer were found to have large changes in entropy, including Notch1 and related ligands. Transcriptional complexes mediated by Notch1 also experienced notable methylation changes.

Discussion: Inhibiting FTO with entacapone causes transcript-specific m6A changes, altering stability of certain gene-regulatory molecules. Disruption of Notch1 and its signaling pathway may result in a change in expression of specific transcripts that cause reduced GSC invasion and self-renewal by altering stability of related transcription factors.

Acknowledgements

It takes a small army to make scientific discoveries, and it takes a significantly larger army to teach a student to make them. Thanking appropriately each of the many mentors and genuine saints who have helped me on my journey to produce this report may be the most daunting task I've taken on during this project, but nevertheless I will now endeavor to do it.

First and foremost, I am extremely grateful for the opportunity provided to me by the Synthego Genetic Engineering Corporation, whose Boston conference I attended on a whimsical LinkedIn-prompted misadventure as naive first year. In short order I found myself in rapturous conversation with some dudes in a brain cancer genomics lab doing the most extraordinary, cutting-edge, translational research in RNA epigenetics. Their names were David Karambizi and John Zepecki, and I was thrilled to learn their PI, Dr. Tapinos, was in fact based at Brown and Rhode Island Hospital. I spent an entire summer at home, feeling rather trapped, chomping at the bit to get my feet wet in this lab.

It was well worth the excitement.

I was astounded at how I was immediately respected as a budding scientist, at a time in my career when most students are scrubbing test tubes and counting flies in squares. Giving an undergrad real scientific responsibility is an enormous risk for a PI, and yet Dr. Tapinos gladly and enthusiastically integrated me into many projects in the lab, supported me through trial and tribulation, and allowed me to spin off my own independent project that became this thesis. John Zepecki is a font of biological and methodological insight who is generously and unassumingly willing to promptly answer questions via text or email at all hours of the day or night, no matter the condition I catch him in. And David Karambizi, who graciously worked shoulder-to-shoulder with me for three years in the lab, taught me the ropes of cell culture and data analysis. David would put everything on hold to catch me up on the latest discoveries, explain the purpose and schedule of a season's slate of experiments, and even give me rides back to the dorm late at night. The support and camaraderie that boils from each and every member of the Laboratory for Cancer Epigenetics and Plasticity is the utmost blessing, and keeps me leaping out of bed in the morning excited to contemplate biology and data.

I have also benefited immensely from my many brilliant and passionate professors. Dr. Harrison, with whom I never actually took a course, provided crucial insight as I developed the mathematical methods in this paper, and encouraged me to think holistically about what analyses are appropriate for the biology rather than get buried in numbers and equations. Dr. Lopatto, another professor who never taught me, also provided invaluable initial insight.

Beyond this project, I am grateful to Dr. Johnson, who met me as a frazzled first-year taking the advanced BIOL0470 Genetics course knowing only that DNA was cool. Through his course and later as a TA for it I learned to be comfortable with and critical of genetic research and methods, as well as to keep an eye open toward how genetics is used beyond academia. Dr. Stein of the physics department taught a biophysics course that seemed to connect the whole universe together in my mind, and I've been grateful for his willingness to field my biophysics-related questions that have emerged in other classes since. It was in his class, as well as a population genetics course taught by Dr. Weinreich, that first exposed me to the idea of a truly quantifiable epigenetic entropy landscape, which has since informed my main academic and research interest and

provided tangential inspiration for using entropy in this application. Finally, Ching-Peng Huang, a graduate student in applied mathematics, provided me two generous semesters of mentorship in mathematical inference for biology amidst completing his own thesis, a feat I wish upon no one but which he took on with grace and compassion.

I have a tremendous support network of friends who will not hesitate to help me through thick and thin. Thanks to Emilija and Deniz and Saradha and Alan from the lab, to Mary and Henry for their consolation in CS courses, to Jason and Jacob for their piano appreciation, Summer and Dorrie for their inhuman ability to emanate cheer and good feeling and for enjoying BSO concerts, to Hossam for deep conversations about everything from blockchain to metaphysics, to Lee for somehow having private jokes with everyone and never making anyone feel excluded, to Joe for his aggressive communism and very unrecognized leadership skills, to Jeremy for being a genius, to George for his intellectual and well-developed takes on everything, to Stephen for his mentorship in deep learning and his amusing enjoyment of a video game I will never understand, to Haisong for always saying the wrong thing and Bunlong for telling it like it is, and to James Lu whose full name is one word to me, and who managed to save a seat on the main green for me every single day during the second COVID summer, without trying.

I am greatly influenced and inspired by my parents, who will drop everything to support me at a moment's notice and teach me that I can have it all if I can earn it all; and my sister, who has so little time or patience for my antics that she keeps me on my toes and gives the most straightforward advice.

Finally, I am immeasurably grateful to the glioblastoma patients who contributed the tumor samples that make the research in the Laboratory for Cancer Epigenetics and Plasticity possible. The courageous journeys of brain tumor patients are an inspiration to live each moment with relish and maximum impact, and underscore the immense privilege and responsibility that is biomedical research.

. . . . . .
Author's signature

Contents

List of Figures

1. Introduction

Glioblastoma multiforme (GBM) is the most common primary brain tumor, accounting for 70% of all gliomas [1, 2]. It is also the most lethal, with median survival time of 14.6 months and nearly 100% final mortality rate [3, 4]. The current method of treatment is maximal safe resection and then radiation alongside temozolomide chemotherapy [5]. Despite this multimodal therapy, tumor recurrence is inevitable due in large part to the presence of glioma stem cells (GSCs) [6]. These cells are genetically characterized by specific lesions and DNA methylation patterns[7, 8]. GSCs are incredibly heterogeneous and highly plastic, able to transition reversibly between stem-like and differentiated states and dynamically express new phenotypic markers in new tumor microenvironments [9, 10, 11, 12]. Such dynamic cell phenotypes may be a result of RNA modifications such as N^6-methyladenosine (m6A).

Eukaryotic messenger RNAs containing the consensus sequence GAC (70%) or AAC (30%) may be methylated on the central A, and this defines the common m6A RNA modification [13]. Writer enzymes Mettl3/14 or eraser enzymes FTO and ALKBH5 respectively methylate or demethylate those sites[14]. Transcripts are usually modified with m6A inside large exons and around stop codons, and such action has been shown to mark the transcript for degradation[15, 16]. This suggests m6A modulates translational efficiency across the transcriptome, and m6A presence on a transcript serves to "destabilize" it[16].

Studies have shown that GSC self-renewal and tumorgenicity are regulated by specific methylation activity of m6A readers and writers [17, 18, 19]. The m6A eraser FTO has been shown to regulate translation of a specific set of mRNAs, including cancer stemness genes such as GSK3B, PTCH1, Notch1, ALCAM, and DLL1 [20, 21]. FTO is a known regulator of cancer stemness, and inhibition blocks self-renewal in GSCs [18, 22]. The drug entacapone, a known inhibitor of catechol-*O*-methyltransferase and an FDA-approved therapy for Parkinson's disease, was recently shown to inhibit FTO and increase m6A levels in human cell lines [23, 24]. Higher-resolution data on m6A abundance across the transcriptome and transcript regions, and a way to quantify that abundance at a transcript-specific level, will enable better understanding of the mechanism by which entacapone affects GSCs.

This report uses single-nucleotide resolution m6A sequencing to compare m6A profiles of entacapone-treated and untreated (DMSO-treated) GSCs. I introduce transcriptomic m6A entropy as a quantitative way to assess changes in m6A regulation and show that it is proportional to m6A abundance changes in this dataset. I show that the transcripts associated with the most extreme methylation change are transcription factors and other central regulatory molecules, as well as key transcripts driving cancer stemness, including Notch1 and related transcripts. Further, I present evidence supporting Notch1-mediated shifts in transcriptional profiles via change of transcription factors along CSL genes. Finally, I present novel ontological and transcriptome-wide patterns in methylation and propose a model to explain them.

2. Analyses

2.1 m6A Entropy

The GSC transcriptome consists of a population of RNA transcripts, each of which may have a single copy or many copies. Each copy of a given transcript has some number of m6A sites which are either occupied or not on each copy. Thus, rather than consider m6A abundance as a deterministic quantity, it is more biologically accurate to consider it a stochastic signal. Any copy of a single RNA transcript contains some combination of occupied and unoccupied m6A sites. I define the m6A site abundance as the percentage of occupied sites out of the total number of transcript copies (input). Thus, site abundance is a probability of a binary state: a Bernoulli random variable, which has average information content, or Shannon entropy, given by:

$$\mathrm{H}(p_s) = p_s \log(p_s) + (1 - p_s) \log(1 - p_s) \tag{2.1}$$

where p_s is the m6A site abundance, the probability of site s being occupied.

Entropy should be interpreted here as the amount of "uncontrolledness", or "freedom" a site experiences: if a site has abundance close to 0.5, the entropy is maximum, and there is likely minimal regulatory control occurring. The site is as likely methylated as not, and therefore is maximally "free". If a site has a more extreme probability of occupancy, such as 0.1 or 0.9, then entropy is lower than maximum, and the cell is more closely regulating or controlling the amount of methylation at that site; the methylation at that site is less freely determined. Sites with occupancy probability of 0 or 1 are under complete control and therefore have 0 entropy. This is visualized in Figure 2.1.

Since the vast majority of sites on the same transcript are separated by at least 100 bp, each site is assumed to be independent. So the m6A entropy of an entire transcript is just the sum of its site entropies:

$$\mathrm{H}_t = \sum_{s=1}^{n} \mathrm{H}(p_s) \tag{2.2}$$

where p_s is the probability of a methyl group occupying site s (the abundance of site s), and n is the total number of sites.

Finally, to avoid entropy calculations being dominated by transcripts with large numbers of sites as Figure 2.1 demonstrates, equation (2.2) is adjusted by dividing by the number of sites on the transcript:

$$\mathrm{H}_t = \frac{1}{n} \sum_{s=1}^{n} \mathrm{H}(p_s) \tag{2.3}$$

Thus, m6A entropy represents the average amount of regulation of m6A occurring over the entire transcript.

The m6A entropy change of a transcript was calculated as the entropy of the transcript

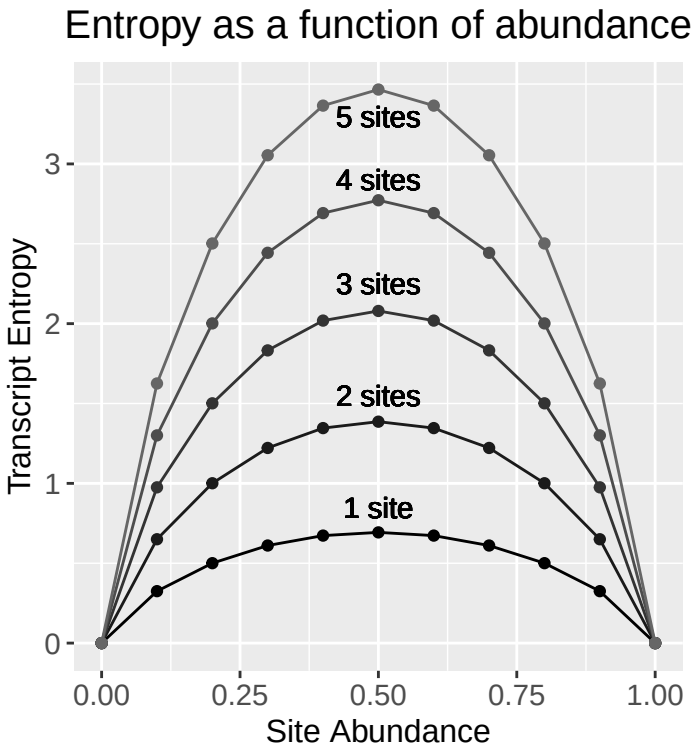

Figure 2.1: The relationship between a transcript's entropy, the number of sites it has, and the abundance of those sites, assuming all sites have the same abundance.

with entacapone treatment transcript minus the entropy of the transcript with DMSO treatment:

$$\Delta \mathrm{H}_t = \mathrm{H}_{t,E} - \mathrm{H}_{t,D} \tag{2.4}$$

The m6A entropy change can be considered the "change in regulation" a transcript's m6A sites undergo upon entacapone treatment

2.2 m6A Euclidean Distance

Euclidean distance is a familiar metric derived from the Pythagorean theorem and often used to quantify read counts in RNA sequence analysis. The Euclidean distance between each entacapone-treated and DMSO-treated transcript was also calculated from m6A abundances, to provide a more straightforward metric of m6A abundance change on a transcript.

The Euclidean distance is given by:

$$\ell^2 = \sqrt{\frac{1}{n}\sum_{s=1}^{n}(p_s - \hat{p}_i)^2} \tag{2.5}$$

where p_s is the entacapone-treated abundance at site s, $\hat{p}_s$ is the DMSO-treated abundance at site s, and n is the number of sites on the transcript.

Since Euclidean distance is a proper measure it is always nonnegative. For the purpose of comparing changes in m6A abundance, which can be positive or negative, I calculated

the Euclidean distance and multiplied it by a sign derived according to the following formula:

$$\ell^{2\prime} = \operatorname{sgn}\left(\sum_{s=1}^{n}(p_s - \hat{p}_s)\right)\sqrt{\frac{1}{n}\sum_{s=1}^{n}(p_s - \hat{p}_i)^2} \tag{2.6}$$

2.3 Visualization

Pathway and gene ontological category enrichment, as well as related plots, were generated using the enrichR library from CRAN. Transcript plots were created by extracting the CDS start and end regions and the transcript length from the provided annotation file for each transcript and then plotting the log2 fold change in site abundance against the index of the site. Plots of multiple transcripts were created with the same process, but then normalizing the lengths of each of the 5' untranslated region (5'UTR), coding sequence (CDS), and 3' UTR. This allowed multiple transcripts of different length and region sizes to be visualized together.

All scripts used for analyses are available for download at this GitHub repository.

3. Results

3.1 Entropy scales with abundance change

The data contained 8,845 m6A sites across 4,451 transcripts for 4,397 genes. No analyzed site had abundance greater than 0.4 (Figure 3.1A). Since abundance was always less than 0.5, entropy is expected to be always proportional to the Euclidean distance in this dataset, since entropy is monotonic increasing until abundance exceeds 0.5 (Figure 2.1). Indeed, the relationship between Euclidean distance and entropy change for the entire transcriptome was roughly linear, with $R^2 = 0.59$ (Figure 3.1B). Wilcox's rank-sum test showed that the differences in abundance captured by changes in entropy are in fact distinct ($p <$ 5e-14), whereas differences in abundance on transcripts chosen by entropy versus Euclidean distance are not ($p > 0.09$) (Table 5.1). These findings indicate that entropy, chosen as a way to quantify the amount of regulation experienced by the m6A sites on a transcript, also directly quantifies m6A in this dataset.

There was no net gain or loss of m6A in this dataset. Of the 4,451 transcripts, 2,064 (46%) increased entropy and 2,229 (50%) had a positive signed Euclidean distance. Of the 8,845 sites, 4,313 (49%) increased abundance.

Transcripts with the most extreme changes in entropy also showed the most direct correlation between m6A entropy and Euclidean distance (Figure 3.1B). These transcripts were examined for trends in site location and gene ontology.

3.2 m6A sites are sparse across the transcriptome but precisely targeted

Of the 4,451 transcripts in the data, 4,346 (98%) were protein-coding mRNAs, and an equal proportion of sites (8,663, 98%) were distributed across them (Figure 3.1A). Most transcripts contained a single m6A site, though the mean number of sites per transcript was 2.01, and some transcripts had as many as 26 sites (AHNAK and MKI67)(Figure 3.1B). Sites were enriched near the stop codon or in the CDS but just 260 (3%) are in the 5' UTR (Figure 3.4).

Transcripts exhibiting the most extreme entropy changes were relatively enriched for sites in the 5' UTR (Figure 3.3). The most extreme 600 transcripts, containing just 7.8% of all sites on coding transcripts, encompassed 22% of all 5' UTR sites. The same trend was not observed in the most extreme 800 sites by fold change, nor in the middle 1600 sites by fold change, all of which contained just 2-4% of all 5'UTR sites (Figure 3.5).

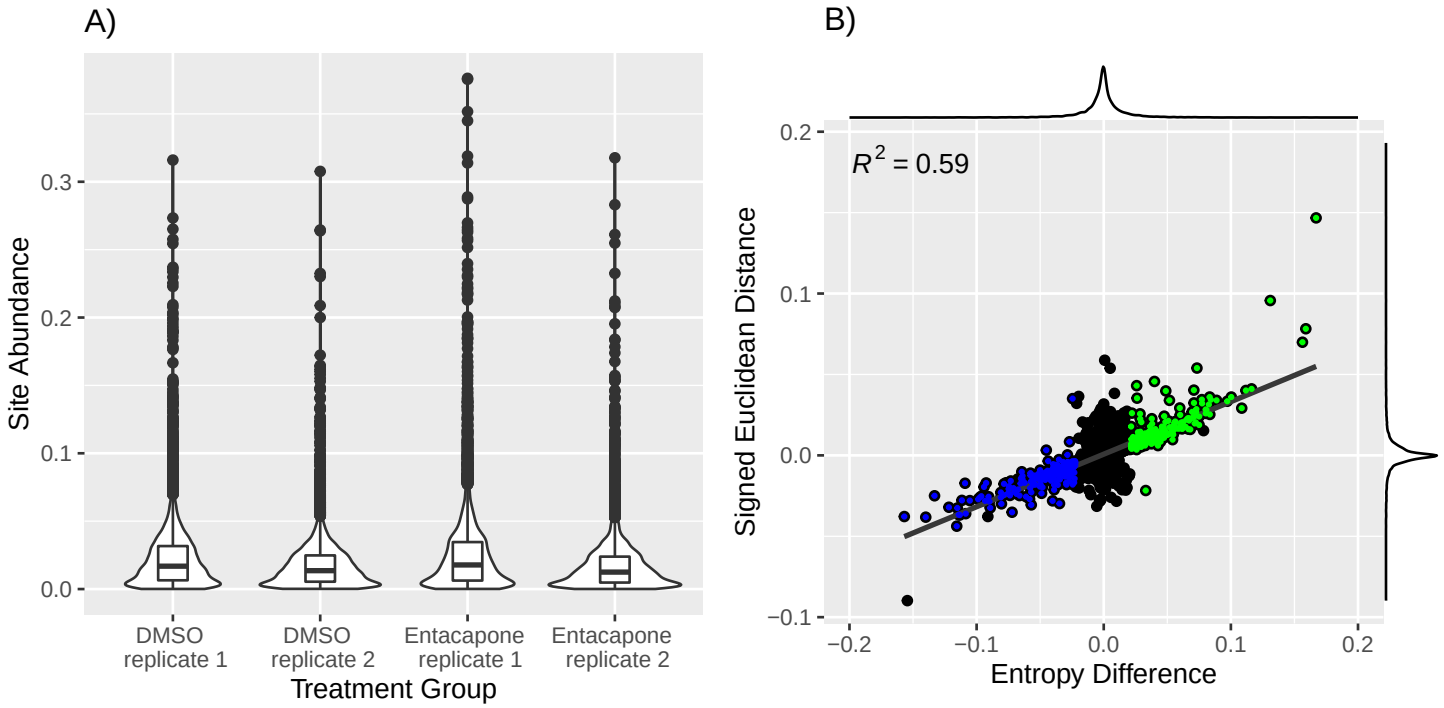

Figure 3.1: A) Distribution of m6A site abundance values in each treatment group. B) Correlation between site entropy and Euclidean distance in the data. Top 300 transcripts by entropy increase are labelled in green, and the top 300 by entropy decrease in blue. These transcripts were selected for further analyses.

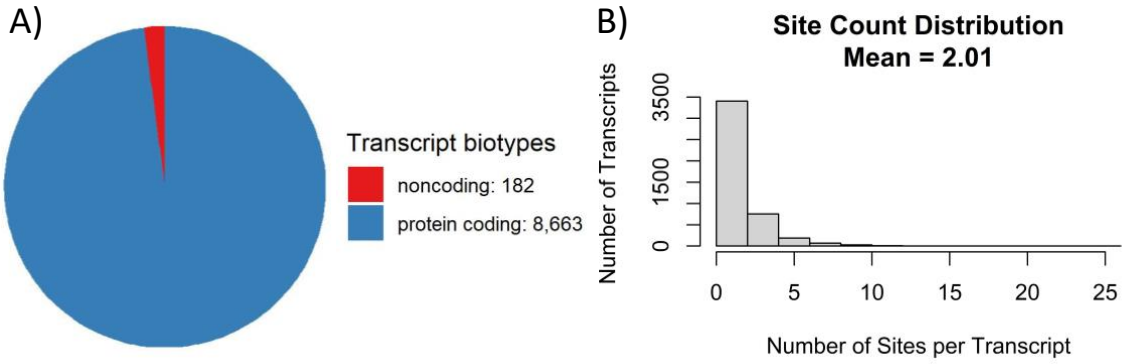

Figure 3.2: Properties of the observed m6A distribution over transcripts. A) Proportion of sites on protein-coding and noncoding transcripts. B) Distribution of site numbers over transcripts. The vast majority of transcripts had a single site, but two had as many as 26 sites.

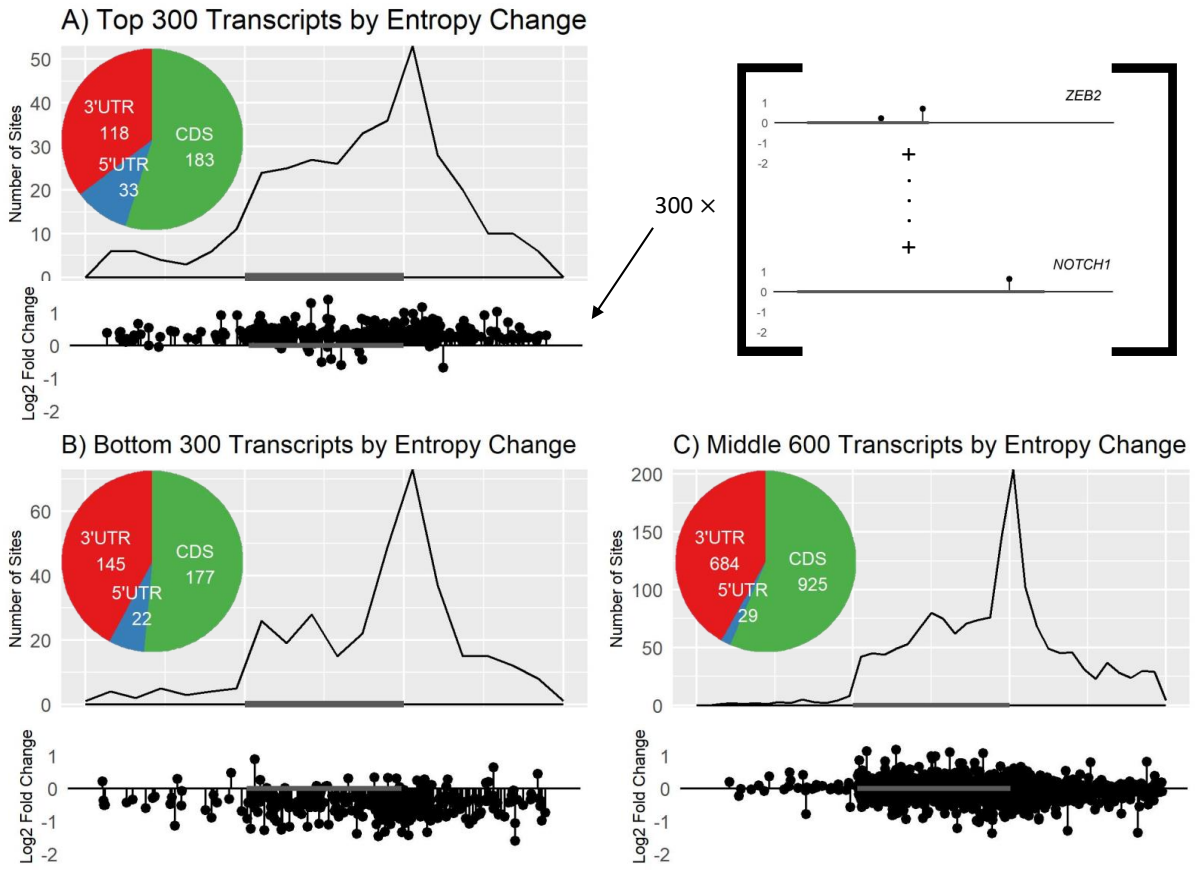

Figure 3.3: Map of m6A sites in transcripts selected by A) greatest entropy increase, B) greatest entropy decrease, and C) minimal change in entropy. Diagrams were made by normalizing all transcript regions to a standard length and plotting all the sites together. Grey bars represent the CDS.

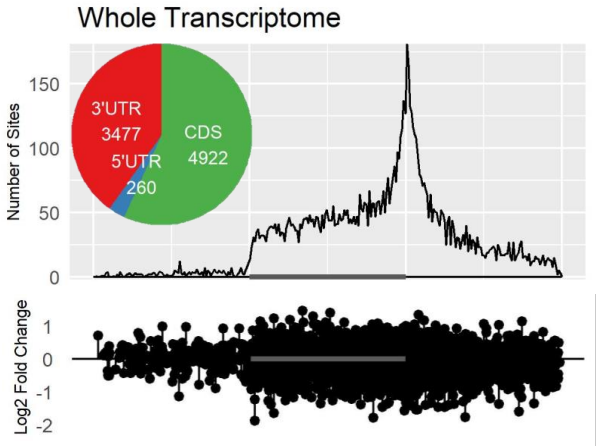

Figure 3.4: Map of all m6A sites in the transcriptome. Diagram was made the same way as Figure 3.3 but summed over the entire set of methylated transcripts. Grey bar represents the CDS.

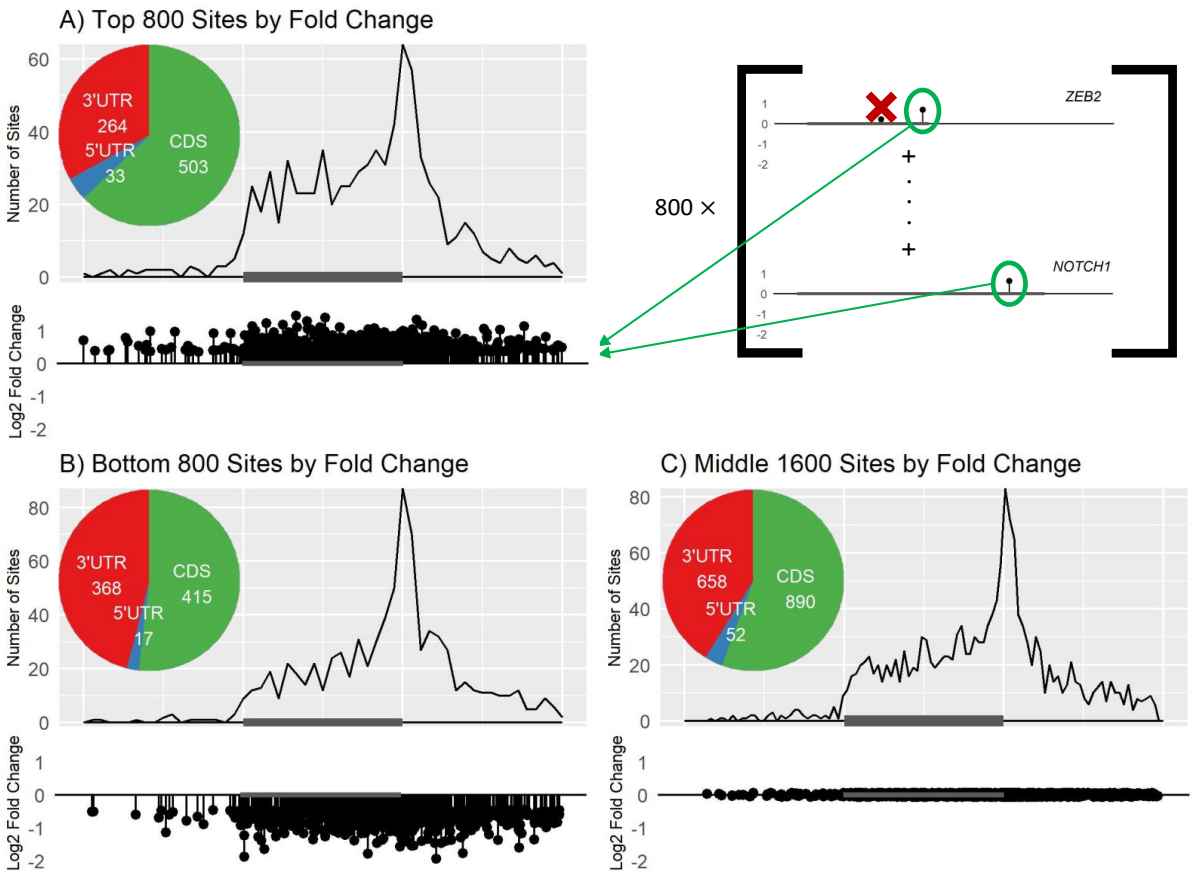

Figure 3.5: Map of m6A sites selected by A) most positive fold change in abundance, B) most negative fold change in abundance, and C) least fold change in abundance. Diagram was made by selecting sites of large fold change, without regard to the transcript it came from or other sites on the transcript. Grey bars represent the CDS.

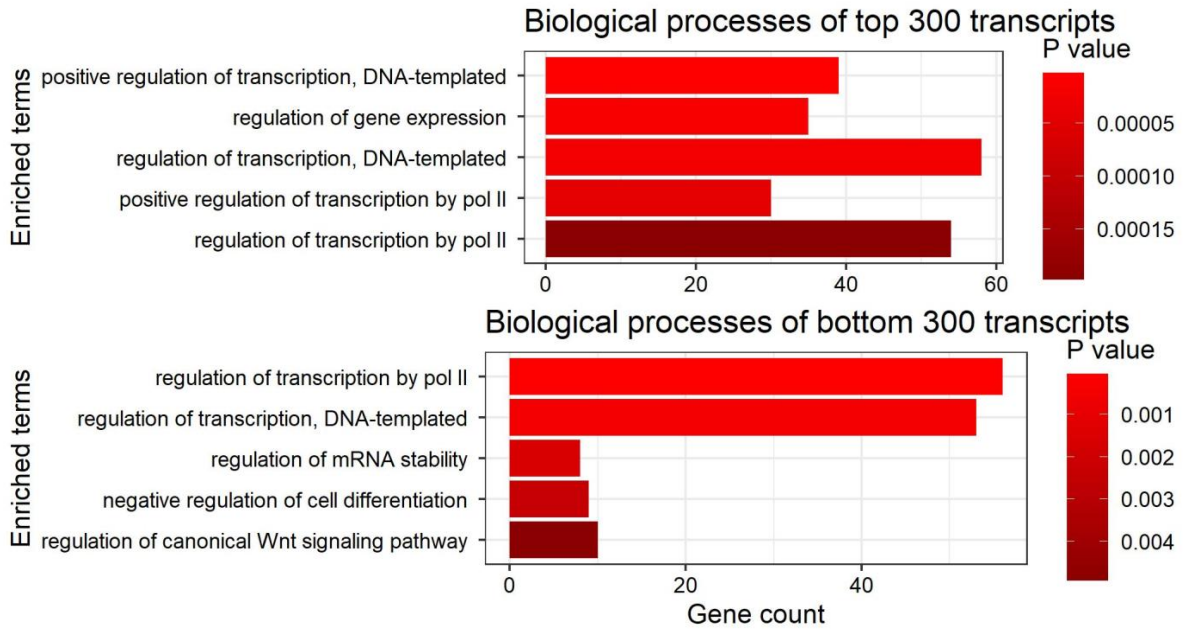

Figure 3.6: Biological processes involving the transcripts with most extreme entropy change.

3.3 Transcription factors and other central regulatory molecules experience extreme m6A abundance changes

Gene set enrichment analyses showed that many of the top 300 transcripts by entropy increase were DNA or protein binding, while the bottom 300 transcripts were enriched for DNA binding but not protein binding (Figure 3.6). Both categories contained high numbers of genes for proteins localized to the nucleus or intracellular membrane-bound organelles (Figure 3.7), and for regulators of transcription and gene expression (Figure 3.8). These findings suggest that m6A sites are enriched on transcripts of transcription factors.

3.4 Key cancer stemness genes, including Notch1 and related genes, show important methylation changes

KEGG pathway analysis indicated that the most extreme transcripts were enriched for signaling pathways associated with cancer when dysregulated (Figure 3.9). These include the Notch[25], Hedgehog[25], TGF-beta[26], ErbB[27], and GnRH[28] pathways, as well as pathways related to cell cycle and, for glioblastoma, axon guidance[29].

28 genes regulating cancer stemness were discovered with m6A sites and are profiled in Figure 3.12. In this subset of transcripts, clusters of sites tend to be in or near the CDS, whereas lone sites could be anywhere. ERBB2 is an obvious exception to this rule, which is notable because genes in the ErbB signaling pathway were enriched among the top 300 transcripts with increased entropy.

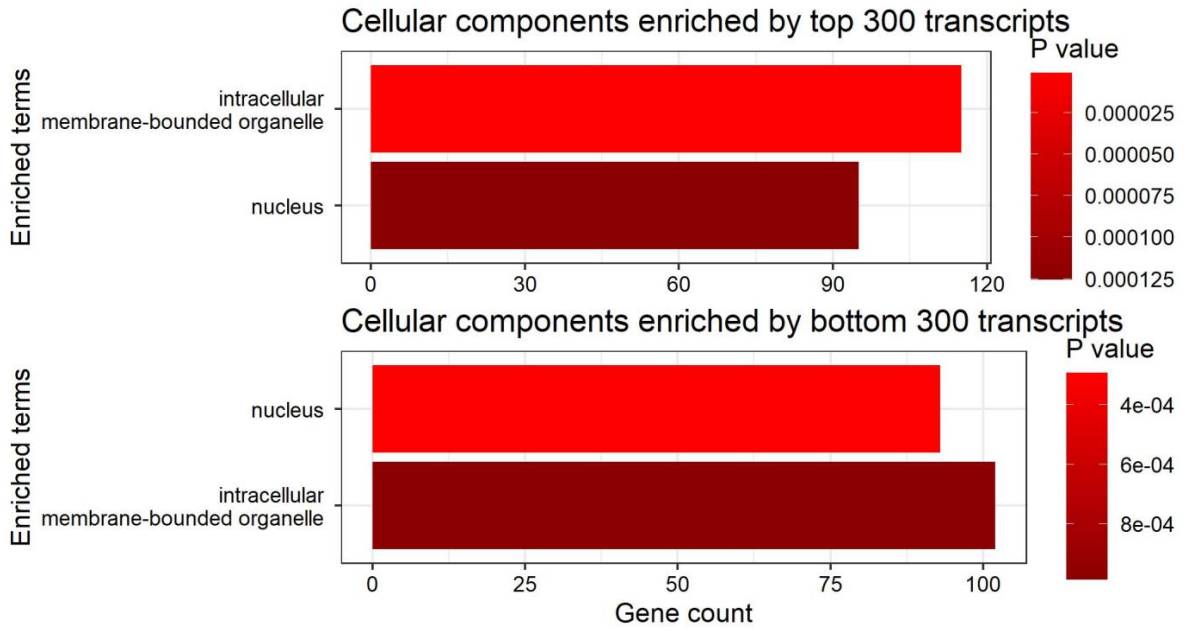

Figure 3.7: Cell components which contain or are comprised of the transcripts with most extreme entropy change.

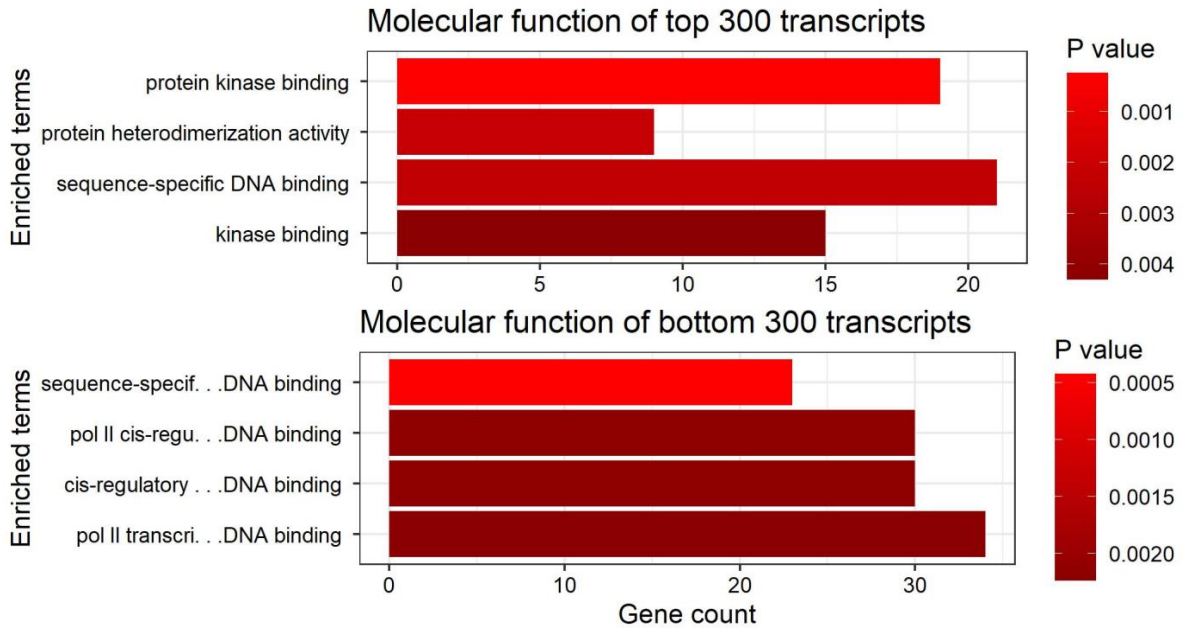

Figure 3.8: Molecular functions of the transcripts with most extreme entropy change.

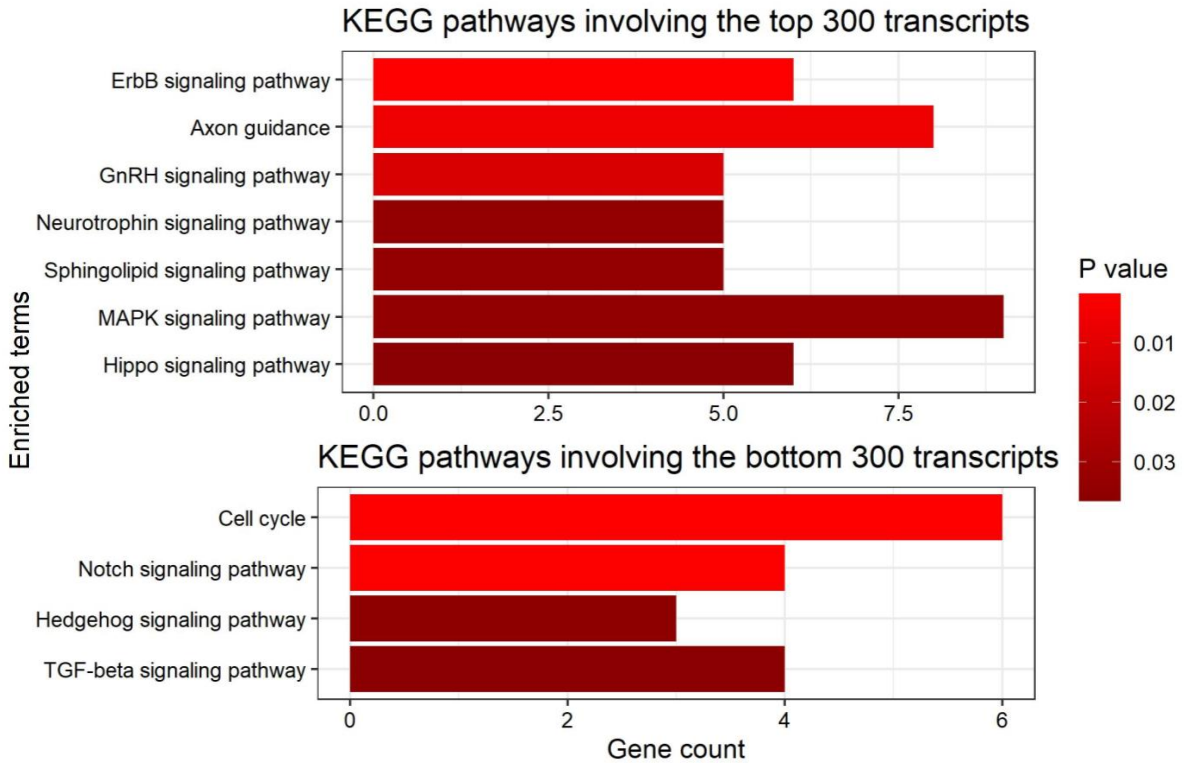

Figure 3.9: KEGG pathway enrichment of the transcripts with most extreme entropy changes.

Notch1 exhibited the sixth highest entropy change and the 12th highest signed Euclidean distance out of all transcripts, and its single site was in the top 2% of abundance increases. Four transcripts in the Notch signaling pathway decreased entropy: ATXN1, DLL1, DVL1, and PSEN1 (Figure 3.11).

3.5 Treatment alters Notch1-mediated transcription regulators

CSL proteins are the primary nuclear effectors of the Notch signaling pathway[30]. They transition from transcriptional repression to transcriptional activation when bound by the Notch1 intracellular domain (Notch1-IC), which must be cleaved from Notch1 protein[31]. This change is accompanied by loss of repressive complex proteins N-CoR and HDAC1 and the recruitment of activator complex proteins MAML1 and p300 alongside Notch1-IC[31]. Transcripts of all these complex proteins possessed m6A sites. Transcripts NCOR1 (for N-CoR) and HDAC1 respectively increased and decreased entropy, and each had one site. Transcripts MAML1 and EP300 (for p300) had six and seven sites, respectively, and both increased entropy(Figure 3.11).

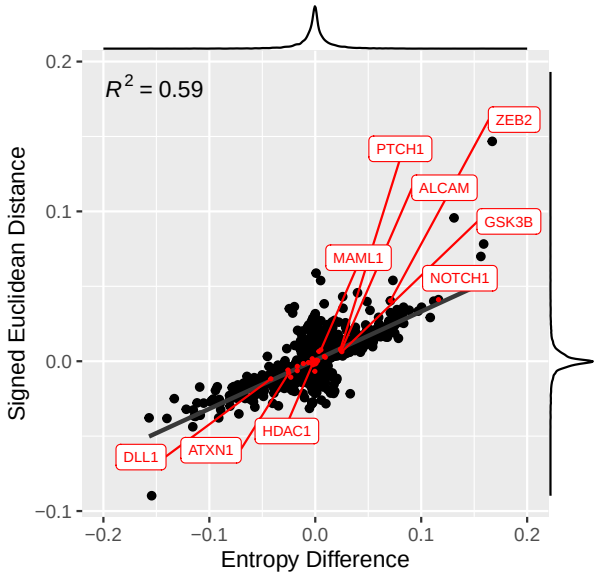

Figure 3.10: Alternate version of Figure 3.1(B), with cancer stemness genes, some of which are labelled, in red.

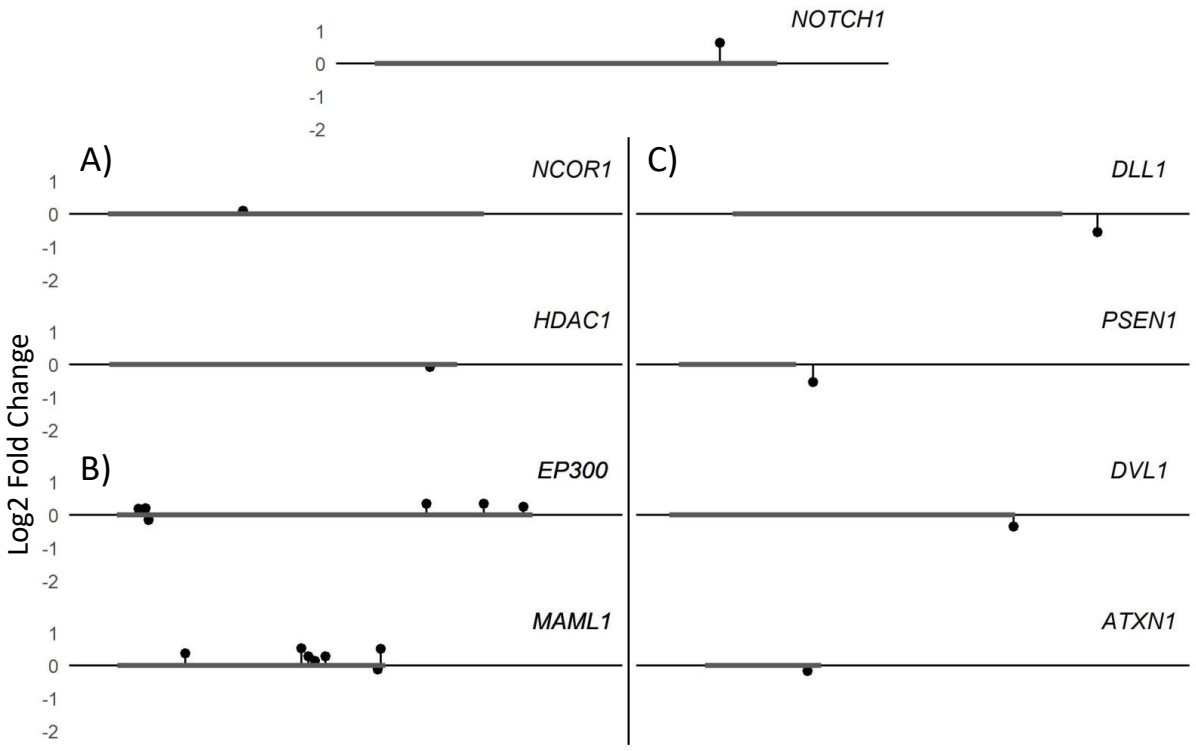

Figure 3.11: Transcripts involved in Notch1 action. A) Transcripts of the CSL repressor complex have a single site in the CDS. B) Transcripts of the CSL activator complex have many sites throughout the CDS and near the stop codon. C) The four enriched transcripts in the Notch1 pathway, ordered by magnitude of abundance change. Most sites in this subset of transcripts are around the stop codon.

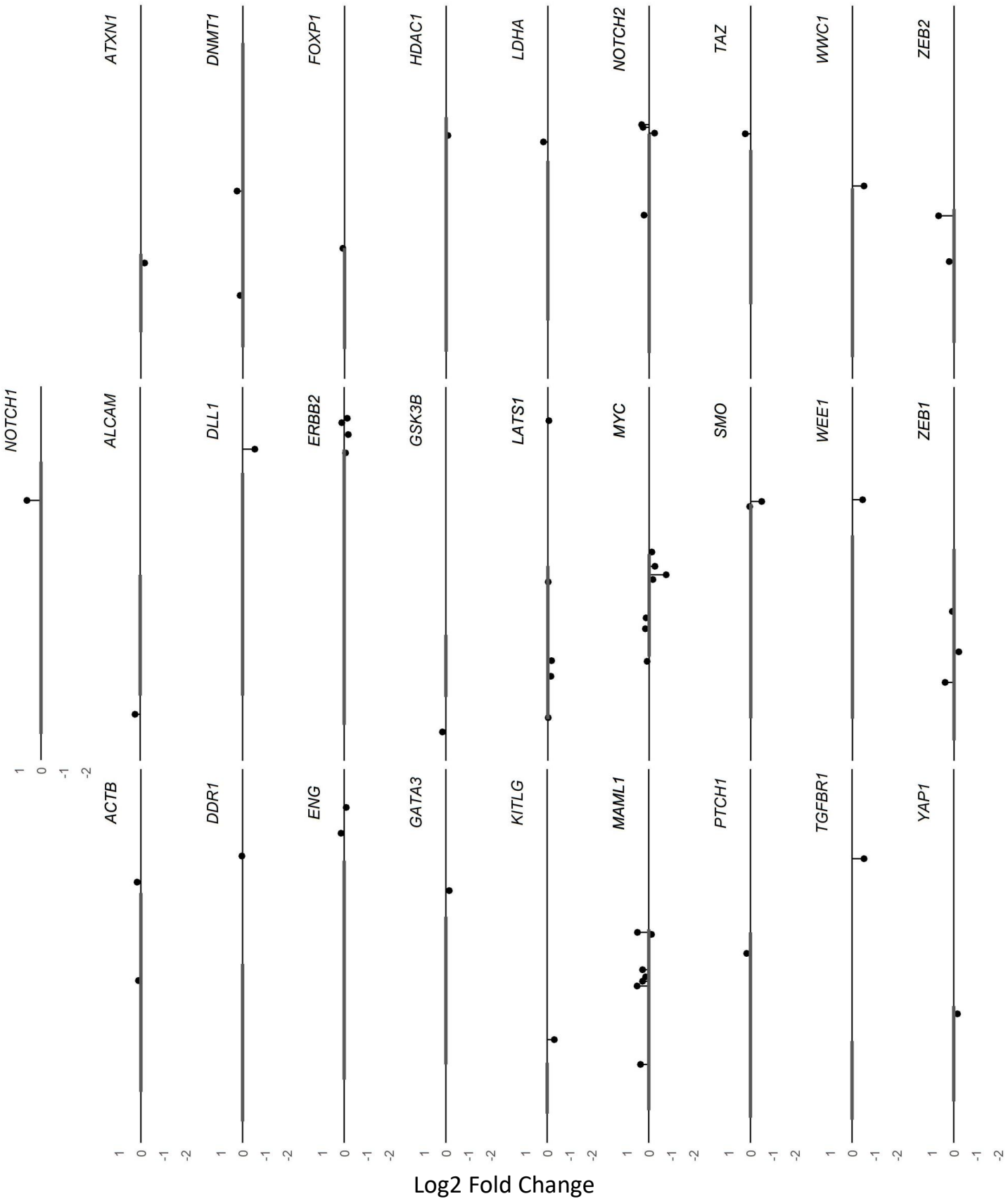

Figure 3.12: Locations of sites on the 28 cancer stemness genes found in the dataset. Height is proportional to the log2 fold change in abundance experienced by that site. Grey box indicates CDS.

4. Discussion and Future Directions

The apparent symmetry of entropy and abundance change is the result of methodological constraint in this study. All site abundances were normalized against the site with maximum abundance; since this site had very large abundance, all other site abundances were forced into the bottom left corner of Figure 2.1, where entropy was positively correlated with abundance. In other datasets with a maximum abundance closer to the median, the relationship between entropy and abundance may not be linear.

Furthermore, RNA transcripts have one or more copies in cells, and single-nucleotide m6A sequence data contains no information about the prevalence of each transcript. The finding that approximately equal numbers of sites and transcripts increased and decreased entropy cannot be taken to imply that global m6A in cells was unchanged with treatment. The possibility that subsets of transcripts are more prevalent with entacapone treatment, and that the m6A abundance on these transcripts is unusually high or low, means that there may be a global increase or decrease in m6A abundance that is not captured by this study.

Enrichment of m6A sites in the 5'UTR on transcripts of extreme entropy change in response to entacapone treatment has not been seen in any previous literature. Whether these transcripts have a special regulatory role or are merely "along for the ride" on transcripts driven by other sites may be answered by comparing the co-occurrence of these sites with other sites of high abundance change or around stop codons. Further investigation is required to determine the consequence of these sites for transcript regulation and GSC physiology.

Transcription factors and signal transduction molecules were found to be prominent among transcripts with extreme entropy change. Considering that the presence of m6A destabilizes transcripts, it appears that entacapone de-enforces certain transcriptional programs by destabilizing transcripts of certain transcription factors, while favoring others by decreasing m6A and increasing stability.

Entacapone uses this mechanism to dysregulate transcripts in critical cancer pathways, usually in the CDS or the very beginning of the 3' UTR. Enrichment of 5' UTR sites was not seen among cancer stemness genes, which suggests that whatever process drives 5' UTR methylation is entirely separate from the stop-codon-centered mechanism regulating transcripts of cancer stemness genes. Notch1 is highly destabilized according to this study, while the genes in the Notch pathway decreased entropy, implying increased stability. Two of these are antagonists of Notch1-IC, including DLL1 and ATXN1 [32, 33]. Another, DLV1, is associated with loss of Notch1 [34]. In contrast, PSEN1 is required for effective Notch1 cleavage and signaling[35]. Nevertheless, these findings suggest that entacapone increases Notch1 signaling by destabilizing Notch1 and generally stabilizing transcripts that negatively regulate Notch1.

The theory that entacapone restores lost Notch1 signaling is also supported by m6A changes on transcripts coding for CSL complex proteins. The large number of sites on activator complex transcripts MAML1 and EP300 suggest that these transcripts play an

important role that demands intense regulation, and their increase in entropy implies they were destabilized, which would favor the repressor complex that occurs without Notch1-IC. Confirming this is the entropy decrease and stabilization of HDAC1. The Notch pathway mediates GSC aggression, invasion, and stemness, so these changes may contribute to the GSC phenotypic change observed with entacapone treatment [21].

4.1 Conclusions

Inhibiting FTO with entacapone causes transcript-specific m6A changes. These m6A changes specifically stabilize or destabilize certain transcripts of gene-regulatory molecules, which alters the GSC transcriptional profile. CSL-regulated genes are repressed by destabilizing Notch1 and stabilizing its negative regulators, resulting in reduced GSC invasion and self-renewal.

5. Appendix

5.1 Laboratory Methods

5.1.1 Ethics statement

The institutional review boards at Rhode Island Hospital and Geisinger Clinic approved the collection of de-identified patient-derived GBM tissue. All participants provided written informed consent for the use of glioblastoma tissue for research purposes.

5.1.2 Cell Culture

Primary GSC spheres were cultured from human glioblastoma samples according to an established protocol[36]. The GSC strain used in this study was authenticated by the American Type Culture Collection using Short Tandem Repeat (STR) analysis. All human primary cells used were between passages 5–20. All cultures were routinely tested for mycoplasma contamination using the LookOut Mycoplasma PCR Detection kit (Sigma).

GSCs were cultured on fibronectin-coated plates 10 µg/mL, MilliporeSigma) in a medium of 1X Neurobasal Medium (Thermo Fisher), 50X serum-free supplement, minus Vitamin A (Thermo Fisher), 100X L-glutamine, 2 mg/mL Heparin (STEMCELL Technologies), 20 ng/mL epidermal growth factor (Thermo Fisher), 20 ng/mL basic-fibroblast growth factor (Thermo Fisher), and 1% Antibiotic-Antimycotic (Thermo Fisher). To induce differentiation, GSCs were cultured for 7 days in a medium of 1X Neurobasal Medium (Thermo Fisher), 50X serum-free supplement, minus Vitamin A (Thermo Fisher), 100X L-glutamine, 10X fetal bovine serum (Gemini), and 1% Antibiotic-Antimycotic (Thermo Fisher).

5.1.3 mRNA isolation

GSC spheres were treated either with entacapone (40 µM × 2 replicates) or dimethylsulfoxide (DMSO, 40 µM × 2 replicates) for 48 hours. Total RNA was isolated from both groups using RNeasy Mini Kit (Qiagen), and subsequently treated with RiboMinus (Thermo Fisher) to remove ribosomal RNA. From this, mRNA was purified with Dynabeads mRNA purification kit (Invitrogen).

5.2 Sequencing

5.2.1 Single-Nucleotide m6A Sequencing

Isolated RNA samples were sent to Arraystar, Inc for sequencing according to an established method [37, 38, 39]. Briefly, each RNA sample was divided into two fractions, one of which was treated with MazF enzyme, and the other not. MazF enzyme cleaves RNA

at unmethylated ACA sites. Since m6A only occurs at the consensus ACA site, cleavage is conditional on m6A presence at that site. These separate MazF-treated and MazF-untreated groups were amplified and treated respectively with Cy5 or Cy3 dye-labelled RNAs. The fractions were then combined, hybridized on a m6A Single Nucleotide Array slide, and scanned using an Agilent Scanner G2505C.

5.2.2 Calculating m6A Abundance

At Arraystar, Agilent Feature Extraction software (version 11.0.1.1) was used to analyze array images. Raw intensities of MazF-Digested (MazF digested RNA, Cy5-labelled) and MazF-Undigested (total RNA, Cy3-labelled) were normalized with average of log2-scaled Spike-in RNA intensities. m6A site abundance was calculated for the m6A site methylation amount based on the MazF-Digested (Cy5-labelled) normalized intensities. These data were annotated with genome build hg19. It is critical for experiments validating transcript diagrams generated in this study that the experimenter use the same annotation file used by the sequencing firm.

Thus the data I used for analyses were prepared. The data consisted of a list of sites, where each line contained a refGene transcript ID, the integer indicating the index in the transcript containing the m6A site, and a series of float values indicating the m6A abundance in each replicate of entacapone-treated and DMSO-treated GSCs. Other lower-level information, such as the signal intensity from the sequencing array, were included in the datafile from the sequencing company but ignored. Before entropy analysis, I scaled all m6A site abundances by the maximum observed abundance out of all the groups, to avoid the computational and interpretational consequences of positive logarithms and negative entropies while preserving differences between replicates, or treatment and control.

Group 1			Group 2			Wilcox Rank-Sum Test Result	
Sites in…	Treatment	Median m6A abundance	Sites in…	Treatment	Median m6A abundance	W	P-value
Top 300 transcripts by entropy change	Entacapone	0.0373	Top 300 transcripts by entropy change	DMSO	0.0278	74048	1.837e-14
Bottom 300 transcripts by entropy change	Entacapone	0.0202	Bottom 300 transcripts by entropy change	DMSO	0.0302	35123	<2.2e-16
Top 300 transcripts by entropy change	Entacapone	0.0373	Bottom 300 transcripts by entropy change	Entacapone	0.0302	91769	<2.2e-16
Top 300 transcripts by entropy change	DMSO	0.0278	Bottom 300 transcripts by entropy change	DMSO	0.0302	52543	0.09429
Top 300 transcripts by Euclidean distance	Entacapone	0.0364	Top 300 transcripts by Euclidean distance	DMSO	0.0299	339070	2.291e-09
Bottom 300 transcripts by Euclidean distance	Entacapone	0.0249	Bottom 300 transcripts by Euclidean distance	DMSO	0.0336	129250	<2.2e-16
Top 300 transcripts by Euclidean distance	Entacapone	0.0364	Bottom 300 transcripts by Euclidean distance	Entacapone	0.0249	302795	<2.2e-16
Top 300 transcripts by Euclidean distance	DMSO	0.0299	Bottom 300 transcripts by Euclidean distance	DMSO	0.0336	204169	0.0005205
Top 300 transcripts by entropy change	Entacapone	0.0373	Top 300 transcripts by Euclidean distance	Entacapone	0.0364	134022	0.09369
Top 300 transcripts by entropy change	DMSO	0.0278	Top 300 transcripts by Euclidean distance	DMSO	0.0299	122135	0.4203
Bottom 300 transcripts by entropy change	Entacapone	0.0202	Bottom 300 transcripts by Euclidean distance	Entacapone	0.0249	84887	5.201e-06
Bottom 300 transcripts by entropy change	DMSO	0.0302	Bottom 300 transcripts by Euclidean distance	DMSO	0.0336	94656	0.0326

Figure 5.1: Results of Wilcox's rank-sum test. P-value indicates probability that the m6A site abundances encompassed by the two tested groups are from the same distribution.

Bibliography

[1] Tracy Seymour, Anna Nowak, and Foteini Kakulas. "Targeting Aggressive Cancer Stem Cells in Glioblastoma". In: *Frontiers in Oncology* 5 (2015), p. 159. ISSN: 2234-943X. DOI: 10.3389/fonc.2015.00159.

[2] David N. Louis et al. "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary". In: *Acta Neuropathologica* 131.6 (June 1, 2016), pp. 803–820. ISSN: 1432-0533. DOI: 10.1007/s00401-016-1545-1. URL: https://doi.org/10.1007/s00401-016-1545-1 (visited on 03/24/2022).

[3] Quinn T. Ostrom et al. "CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2008-2012". In: *Neuro-Oncology* 17 (Suppl 4 Oct. 2015), pp. iv1–iv62. ISSN: 1522-8517. DOI: 10.1093/neuonc/nov189. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4623240/ (visited on 03/24/2022).

[4] Taylor A. Wilson, Matthias A. Karajannis, and David H. Harter. "Glioblastoma multiforme: State of the art and future therapeutics". In: *Surgical Neurology International* 5 (2014), p. 64. ISSN: 2229-5097. DOI: 10.4103/2152-7806.132138.

[5] Jennifer Clarke, Nicholas Butowski, and Susan Chang. "Recent advances in therapy for glioblastoma". In: *Archives of Neurology* 67.3 (Mar. 2010), pp. 279–283. ISSN: 1538-3687. DOI: 10.1001/archneurol.2010.5.

[6] Roger Stupp and Damien C. Weber. "The role of radio- and chemotherapy in glioblastoma". In: *Onkologie* 28.6 (June 2005), pp. 315–317. ISSN: 0378-584X. DOI: 10.1159/000085575.

[7] "Comprehensive genomic characterization defines human glioblastoma genes and core pathways". In: *Nature* 455.7216 (Oct. 23, 2008), pp. 1061–1068. ISSN: 0028-0836. DOI: 10.1038/nature07385. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2671642/ (visited on 03/24/2022).

[8] Cameron W. Brennan et al. "The Somatic Genomic Landscape of Glioblastoma". In: *Cell* 155.2 (Oct. 10, 2013), pp. 462–477. ISSN: 0092-8674. DOI: 10.1016/j.cell.2013.09.034. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3910500/ (visited on 03/24/2022).

[9] Philippe L. Bedard et al. "Tumour heterogeneity in the clinic". In: *Nature* 501.7467 (Sept. 2013). Number: 7467 Publisher: Nature Publishing Group, pp. 355–364. ISSN: 1476-4687. DOI: 10.1038/nature12627. URL: https://www.nature.com/articles/nature12627 (visited on 03/24/2022).

[10] Anoop P. Patel et al. “Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma”. In: *Science (New York, N.Y.)* 344.6190 (June 20, 2014), pp. 1396–1401. ISSN: 0036-8075. DOI: 10.1126/science.1254257. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4123637/ (visited on 03/24/2022).

[11] Ittai Ben-Porath et al. “An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors”. In: *Nature genetics* 40.5 (May 2008), pp. 499–507. ISSN: 1061-4036. DOI: 10.1038/ng.127. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2912221/ (visited on 03/24/2022).

[12] Xun Jin et al. “Targeting Glioma Stem Cells through Combined BMI1 and EZH2 Inhibition”. In: *Nature medicine* 23.11 (Nov. 2017), pp. 1352–1361. ISSN: 1078-8956. DOI: 10.1038/nm.4415. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5679732/ (visited on 03/24/2022).

[13] P. Narayan and F. M. Rottman. “An in vitro system for accurate methylation of internal adenosine residues in messenger RNA”. In: *Science (New York, N.Y.)* 242.4882 (Nov. 25, 1988), pp. 1159–1162. ISSN: 0036-8075. DOI: 10.1126/science.3187541.

[14] Ying Yang et al. “Dynamic transcriptomic m6A decoration: writers, erasers, readers and functions in RNA metabolism”. In: *Cell Research* 28.6 (June 2018). Number: 6 Publisher: Nature Publishing Group, pp. 616–624. ISSN: 1748-7838. DOI: 10.1038/s41422-018-0040-8. URL: https://www.nature.com/articles/s41422-018-0040-8 (visited on 03/24/2022).

[15] Kate D. Meyer et al. “Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3’ UTRs and Near Stop Codons”. In: *Cell* 149.7 (June 22, 2012), pp. 1635–1646. ISSN: 0092-8674. DOI: 10.1016/j.cell.2012.05.003. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3383396/ (visited on 03/24/2022).

[16] Xiao Wang et al. “N6-methyladenosine Modulates Messenger RNA Translation Efficiency”. In: *Cell* 161.6 (June 4, 2015), pp. 1388–1399. ISSN: 0092-8674. DOI: 10.1016/j.cell.2015.05.014. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4825696/ (visited on 03/24/2022).

[17] A. Visvanathan et al. “Essential role of METTL3-mediated m6A modification in glioma stem-like cells maintenance and radioresistance”. In: *Oncogene* 37.4 (Jan. 25, 2018), pp. 522–533. ISSN: 1476-5594. DOI: 10.1038/onc.2017.351.

[18] Qi Cui et al. “m6A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells”. In: *Cell reports* 18.11 (Mar. 14, 2017), pp. 2622–2634. ISSN: 2211-1247. DOI: 10.1016/j.celrep.2017.02.059. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5479356/ (visited on 03/24/2022).

[19] Sicong Zhang et al. “The m6A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-Like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program”. In: *Cancer cell* 31.4 (Apr. 10, 2017), 591–606.e6. ISSN: 1535-6108. DOI: 10.1016/j.ccell.2017.02.013. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5427719/ (visited on 03/24/2022).

[20] John P. Zepecki et al. “miRNA-mediated loss of m6A increases nascent translation in glioblastoma”. In: *PLoS Genetics* 17.3 (Mar. 8, 2021), e1009086. ISSN: 1553-7390. DOI: 10.1371/journal.pgen.1009086. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7971852/ (visited on 03/24/2022).

[21] Li Yi et al. "Notch1 signaling pathway promotes invasion, self-renewal and growth of glioma initiating cells via modulating chemokine system CXCL12/CXCR4". In: *Journal of Experimental & Clinical Cancer Research : CR* 38 (Aug. 5, 2019), p. 339. ISSN: 0392-9078. DOI: 10.1186/s13046-019-1319-4. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6683584/ (visited on 04/06/2022).

[22] Rui Li et al. "Pan-Cancer Prognostic, Immunity, Stemness, and Anticancer Drug Sensitivity Characterization of N6-Methyladenosine RNA Modification Regulators in Human Cancers". In: *Frontiers in Molecular Biosciences* 8 (June 4, 2021), p. 644620. ISSN: 2296-889X. DOI: 10.3389/fmolb.2021.644620. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8211991/ (visited on 04/08/2022).

[23] Mengnuo Chen and Chun-Ming Wong. "The emerging roles of N6-methyladenosine (m6A) deregulation in liver carcinogenesis". In: *Molecular Cancer* 19 (Feb. 28, 2020), p. 44. ISSN: 1476-4598. DOI: 10.1186/s12943-020-01172-y. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7047367/ (visited on 03/24/2022).

[24] Shiming Peng et al. "Identification of entacapone as a chemical inhibitor of FTO mediating metabolic regulation through FOXO1". In: *Science Translational Medicine* 11.488 (Apr. 17, 2019), eaau7116. ISSN: 1946-6242. DOI: 10.1126/scitranslmed.aau7116.

[25] Vivek Kumar et al. "The Role of Notch, Hedgehog, and Wnt Signaling Pathways in the Resistance of Tumors to Anticancer Therapies". In: *Frontiers in Cell and Developmental Biology* 9 (Apr. 22, 2021), p. 650772. ISSN: 2296-634X. DOI: 10.3389/fcell.2021.650772. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8100510/ (visited on 04/06/2022).

[26] Isabel Burghardt et al. "Endoglin and TGF- signaling in glioblastoma". In: *Cell and Tissue Research* 384.3 (2021), pp. 613–624. ISSN: 0302-766X. DOI: 10.1007/s00441-020-03323-5. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8211614/ (visited on 04/06/2022).

[27] Muhammad Nadeem Abbas et al. "Advances in Targeting the Epidermal Growth Factor Receptor Pathway by Synthetic Products and Its Regulation by Epigenetic Modulators as a Therapy for Glioblastoma". In: *Cells* 8.4 (Apr. 12, 2019), p. 350. ISSN: 2073-4409. DOI: 10.3390/cells8040350. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6523687/ (visited on 04/06/2022).

[28] Cai-Ping Chen and Xiang Lu. "Gonadotropin-releasing hormone receptor inhibits triple-negative breast cancer proliferation and metastasis". In: *The Journal of International Medical Research* 50.3 (Mar. 9, 2022), p. 03000605221082895. ISSN: 0300-0605. DOI: 10.1177/03000605221082895. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8918972/ (visited on 04/06/2022).

[29] Yong Huang et al. "Plexin-B2 facilitates glioblastoma infiltration by modulating cell biomechanics". In: *Communications Biology* 4 (Jan. 29, 2021), p. 145. ISSN: 2399-3642. DOI: 10.1038/s42003-021-01667-4. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7846610/ (visited on 04/06/2022).

[30] Eric C. Lai. "Keeping a good pathway down: transcriptional repression of Notch pathway target genes by CSL proteins". In: *EMBO Reports* 3.9 (Sept. 16, 2002), pp. 840–845. ISSN: 1469-221X. DOI: 10.1093/embo-reports/kvf170. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1084223/ (visited on 04/06/2022).

[31] J J Hsieh et al. “Truncated mammalian Notch1 activates CBF1/RBPJk-repressed genes by a mechanism resembling that of Epstein-Barr virus EBNA2.” In: *Molecular and Cellular Biology* 16.3 (Mar. 1996), pp. 952–959. ISSN: 0270-7306. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC231077/ (visited on 04/06/2022).

[32] Jane Jung et al. “Regulation of Notch1 Signaling by Delta-like Ligand 1 Intracellular Domain through Physical Interaction”. In: *Molecules and Cells* 32.2 (Aug. 31, 2011), pp. 161–165. ISSN: 1016-8478. DOI: 10.1007/s10059-011-1046-y. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3887669/ (visited on 04/06/2022).

[33] Xin Tong et al. “Ataxin-1 and Brother of ataxin-1 are components of the Notch signalling pathway”. In: *EMBO Reports* 12.5 (May 1, 2011), pp. 428–435. ISSN: 1469-221X. DOI: 10.1038/embor.2011.49. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3090018/ (visited on 04/06/2022).

[34] D. Cotter et al. “Disturbance of Notch-1 and Wnt signalling proteins in neuroglial balloon cells and abnormal large neurons in focal cortical dysplasia in human cortex”. In: *Acta Neuropathologica* 98.5 (Nov. 1999), pp. 465–472. ISSN: 0001-6322. DOI: 10.1007/s004010051111.

[35] Bart De Strooper et al. “A presenilin-1-dependent -secretase-like protease mediates release of Notch intracellular domain”. In: *Nature* 398.6727 (Apr. 1999). Number: 6727 Publisher: Nature Publishing Group, pp. 518–522. ISSN: 1476-4687. DOI: 10.1038/19083. URL: https://www.nature.com/articles/19083 (visited on 04/06/2022).

[36] Rossella Galli et al. “Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma”. In: *Cancer Research* 64.19 (Oct. 1, 2004), pp. 7011–7021. ISSN: 0008-5472. DOI: 10.1158/0008-5472.CAN-04-1364.

[37] Miki Imanishi et al. “Detection of N6-methyladenosine based on the methyl-sensitivity of MazF RNA endonuclease”. In: *Chemical Communications* 53.96 (Nov. 30, 2017). Publisher: The Royal Society of Chemistry, pp. 12930–12933. ISSN: 1364-548X. DOI: 10.1039/C7CC07699A. URL: https://pubs.rsc.org/en/content/articlelanding/2017/cc/c7cc07699a (visited on 04/05/2022).

[38] Miguel Angel Garcia-Campos et al. “Deciphering the “m6A Code” via Antibody-Independent Quantitative Profiling”. In: *Cell* 178.3 (July 25, 2019). Publisher: Elsevier, 731–747.e16. ISSN: 0092-8674, 1097-4172. DOI: 10.1016/j.cell.2019.06.013. URL: https://www.cell.com/cell/abstract/S0092-8674(19)30676-2 (visited on 04/03/2022).

[39] Hong-Xuan Chen et al. “Mapping single-nucleotide m6A by m6A-REF-seq”. In: *Methods* (June 24, 2021). ISSN: 1046-2023. DOI: 10.1016/j.ymeth.2021.06.013. URL: https://www.sciencedirect.com/science/article/pii/S1046202321001754 (visited on 04/03/2022).