Genetic factors that correlates with glioma aggression within G-CIMP cancer subtypes

Mohamed Elsharnouby, *mohamedelsharnouby@std.sehir.edu.tr*
Nur Betül Yerli, *nuryerli@std.sehir.edu.tr*
Istanbul Sehir University

***Abstract*—Glioblastoma (GBM) is the most aggressive brain tumor, it occurs mostly as a primary tumor and it is grade IV glioma according to WHO classification. Low grade glioma (LGG) is grades II and III of glioma and they are much less aggressive than GBM. The data provided by TCGA has showed that a genomic analysis of it could yield results and insights that guides future research and treatment methodologies of the disease. In this paper we analyzed the data available on TCGA to analyze the genes contributing to the aggressiveness of glioma and what are the possible genes contributing to the progression of LGG to GBM. We also ran a gene enrichment analysis to get the affected pathways contributing to this aggressiveness.**

***Index Terms*—Glioblastoma Multiforme, GBM, Low Grade Glioma, LGG, genomics.**

I. Introduction

GLIOMA is a heterogeneous type of cancer that affects the brain and it differs histologically according to the type of the glial cells the start the tumor. Gliomas are classified by WHO (World Health Organization) into four grades I, II, III and IV according to pathological state of the tumor.

Low grade gliomas (LGG) are grade II and III in WHO classification while Glioblastoma multiforme (GBM) is the most aggressive glioma (grade IV). The patients diagnosed with GBM tend to survive for approximately one year even with aggressive treatment. Gliomas were extensively studied by TCGA (The Cancer Genome Atlas) research network and the first publication about GBM from TCGA [1] explored 206 GBM cases and did genomic and transcriptomic analysis for those cases. This study made it apparent that a systematic genomic analysis of a statistically powered cohort was able to provide knowledge about the disease's core biological pathways and generate significant insights that could be used in future research and treatment methodologies.

Glioblastoma occurs most frequently as a primary tumor, forming directly in the affected area, growing and expanding rapidly to occupy its own place in the brain causing the signs and symptoms to appear on the patient. The signs and symptoms might include headaches, personality changes, nausea, vomiting and even unconsciousness. This form is very aggressive that it leads to a very fast deterioration of patient's clinical state and it is the most common form. The patient that is diagnosed with this form approximately live for one year after diagnosis. The other form of glioblastoma is being a secondary tumor which means it started as a low grade astrocytoma or anaplastic astrocytoma and this constitutes around 10% of glioblastomas. Glioblastoma has four subtypes [2] which are characterized by their underlying genetic abnormalities. They are Classical, Mesenchymal, Proneural and Neural subtypes. The proneural subtype has been recently found [3] to contain two subsets, GCIMP positive and negative subsets.

The availability of datasets as the TCGA's LGG and GBM datasets allowed researchers to relate genomic and DNA methylation with prognosis. For example, the presence of Isocitrate Dehydrogenase genes 1 and 2 mutations (IDH1/IDH2) showed a distinct hypermethylation of a certain number of loci identifying a CpG island methylator phenotype (GCIMP) in favorable outcomes for patients of GBM. These patients were generally younger and have higher survival rate. On the other hand, the absence of IDH mutation (IDH wild type) in LGG tumors makes the prognosis of patients less favorable [4].

As there could be other genetic factors that could contribute to the severity of the glioma GCIMP subtypes, it is a necessity to explore all of the contributing factors and their correlation with the aggression of the glioma. Such study will also show the significance of statistical analysis of available datasets in deeper understanding of the mechanisms underlying cancer severity and outcomes.

II. Materials and Methods

A. Samples

We used TCGA dataset (http://cancergenome.nih.gov/) as the main datasource of samples. The number of samples and their classification is shown in table I. We downloaded the clinical and gene expression data of these cases from the TCGA website including all preservation types (Frozen and FFPE).

	GCIMP+	GCIMP-	Total
GBM	31	368	**399**
LGG	25	2	**27**
Total	**56**	**370**	**426**

TABLE I
Number of samples

B. Data Sets

1) TCGA gene expression dataset: We downloaded the level three of the "AgilentG4502A_07" platform dataset for both LGG and GBM tumors. We imported it into Python using Pandas library.

2) G-CIMP classification dataset: We use the supervised DNA methylation classification in [5] as a G-CIMP classification label for the samples we downloaded form the TCGA and their gene expression data.

C. Statistical t-test

We used t-test to compare the values of genes expression in between the independent groups of GBM and LGG inside the two groups of GCIMP. We split the data first into GCIMP+ and GCIMP- data within each of these 2 groups of samples we split them further into GBM samples and LGG samples. After this we run the t-test to compare the gene expression data of GBM and LGG samples with each group. This is to compare the genes that contribute to the aggressiveness of the tumor. We choose the genes with significant change in expression (with p-value less than or equal to 0.01). After this we combine both sets of genes from the GCIMP+ and the GCIMP- groups into one single set of shared genes that have significant change in expression in both groups.

D. Logistic regression classifier

After getting the set of genes affecting the aggressiveness of glioma, we used a logistic regression classifier to rank them according to the most probable gene contributing to being a GBM instead of a LGG. As a logistic regression classifier result of a single sample could be interpreted as the probability that this gene belongs to either of the classes, we can use the weights that results inside that classifier for each feature (gene) as a probability that this gene is contributing more for a sample to be of one of the classes (GBM or LGG). This could be done if we train the logistic regression classifier to classify a sample as a GBM or a LGG using the expression of this set of genes of each sample as a feature vector. After training the classifier we used the weights of each gene that are used by the classifier to rank the genes as contributing more for the probability of a sample to be GBM. Negative weights of genes could be interpreted as the genes that are contributing more for the probability of a sample to be LGG.

E. Gene ontology analysis

We use the gene ontology analysis tool [6] on (http://amigo.geneontology.org/rte) to analyze the set of genes that we came up with from the previous step to further analyze and understand the affected pathways.

F. Software Tools

We used Python, Pandas and Scipy library to import the downloaded clinical and gene expression data and to run our analysis using the included statistics package.

III. Results

For the analysis of GCIMP+ group 7509 genes were found to be significantly different in expression in between the GBM and LGG subgroups. The genes were ranked according to the p-value of the t-test that was run on the samples for each gene. The top 10 genes according to the p-value have the upcoming functionality:

1) CCDC115 (p-value=2.17E-23): the protein encoded by this gene has been observed to localize to the endoplasmic reticulum (ER)-Golgi Intermediate Compartment (ERGIC) and Coat Protein Complex I (COPI) vesicles in some human cells. Defects in this gene are a cause of congenital disorder of glycosylation, type IIo in humans.
2) PFDN6 (p-value=2.71E-23): chains pending their transfer to the cytosolic chaperonin containing TCP1 complex
3) ATG2A (p-value=9.75E-23): the protein encoded by this gene is essential for autophagosome formation and important for regulation of size and distribution of lipid droplets. It is also required for localization to both the autophagic membrane and lipid droplet and is also essential for autophagy. Depletion of both Atg2A and Atg2B causes clustering of enlarged lipid droplets in an autophagy-independent manner.
4) RIC8A (p-value=3.94E-21): guanine-exchange factor RIC8A binds to the calcium ion sensor NCS-1 to regulate synapse number and neurotransmitter release
5) CCDC71 (p-value=4.86E-21)
6) HNRPCL1 (p-value=5.89E-21): The gene ontology of this gene is in the nucleic acid binding group
7) LIMK1 (p-value=3.59E-20): LIMK1 is a serine/threonine kinase that regulates actin polymerization via phosphorylation and inactivation of the actin binding factor cofilin. This protein is ubiquitously expressed during development and plays a role in many cellular processes associated with cytoskeletal structure. This protein also stimulates axon growth and may play a role in brain development. LIMK1 hemizygosity is implicated in the impaired visuospatial constructive cognition of Williams syndrome
8) MRFAP1 (p-value=3.62E-20): this gene encodes an intracellular protein that interacts with members of the MORF4/MRG (mortality factor on chromosome 4/MORF4 related gene) family and the tumor suppressor Rb (retinoblastoma) protein. The protein may play a role in senescence, cell growth and immortalization
9) MAP3K7IP1 (p-value=6.47E-20): the protein encoded by this gene was identified as a regulator of a protein known to mediate various intracellular signaling pathways, such as those induced by TGF beta, interleukin 1, and WNT-1. This protein plays an important role in skin homeostasis, wound repair and oncogenesis
10) FAM134A (p-value=7.63E-20): it belongs in gene ontology to integral component of membrane group

For the GCIMP- group 1274 genes were found to be have significantly different gene expression in between GBM and LGG subgroups. The genes were ranked as well according

to the p-value of the t-test that was run on the samples for each gene. The top 10 genes according to the p-value have the upcoming functionality:

1) C1orf142 (p-value=2.57E-14): also known as SNAP47 and it encodes for one of the SNARE proteins which are an essential components of the intracellular fusion machinery. It is thought that they form a tight four-helix complex between membranes, in effect initiating fusion. It also has a relation to lung cancer [7]
2) KCTD21 (p-value=6.56E-13): in gene ontology it encodes a protein of homooligomerization which is the process of creating protein oligomers, compounds composed of a small number, usually between three and ten, of identical component monomers
3) KIAA1183 (p-value=2.38E-11): also known as PNMAL2
4) MTCH1 (p-value=2.96E-10): this gene encodes a member of the mitochondrial carrier family. The encoded protein is localized to the mitochondrion inner membrane and induces apoptosis independent of the proapoptotic proteins Bax and Bak
5) CCDC115 (p-value=3.09E-10): mentioned before in GCIMP+ group
6) CRY2 (p-value=3.05E-09): this gene encodes a flavin adenine dinucleotide-binding protein that is a key component of the circadian core oscillator complex, which regulates the circadian clock. This gene is upregulated by CLOCK/ARNTL heterodimers but then represses this upregulation in a feedback loop using PER/CRY heterodimers to interact with CLOCK/ARNTL. Polymorphisms in this gene have been associated with altered sleep patterns. The circadian gene CRY2 is associated with breast cancer aggressiveness possibly via epigenomic modifications [8]
7) PARP1 (p-value=5.93E-09): this gene encodes a chromatin-associated enzyme, poly(ADP-ribosyl)transferase, which modifies various nuclear proteins by poly(ADP-ribosyl)ation. The modification is dependent on DNA and is involved in the regulation of various important cellular processes such as differentiation, proliferation, and tumor transformation and also in the regulation of the molecular events involved in the recovery of cell from DNA damage. In addition, this enzyme may be the site of mutation in Fanconi anemia, and may participate in the pathophysiology of type I diabetes
8) FAM120B (p-value=1.09E-08): in gene ontology it is involved in cell differentiation, regulation of transcription and DNA-templated transcription
9) SLC22A6 (p-value=1.62E-08): the protein encoded by this gene is involved in the sodium-dependent transport and excretion of organic anions, some of which are potentially toxic. The encoded protein is an integral membrane protein and may be localized to the basolateral membrane
10) GSX1 (p-value=1.72E-08): its expression defines neurons required for prepulse inhibition in schizophrenia [9]

Taking the intersection of the two sets of genes that were significantly different in expression from both groups resulted in 1078 shared genes in between the two groups of GCIMP+ and GCIMP-. We used this set of shared genes to run our gene ontological analysis which resulted in 133 pathways that are affected by these genes.

The found pathways were organized hierarchically into main pathways using the tool of gene ontology analysis, here is the list of the main significantly affected pathways (p-value ¡ 0.05):

1) Detection of chemical stimulus involved in sensory perception of smell (GO:0050911): The series of events involved in the perception of smell in which an olfactory chemical stimulus is received and converted into a molecular signal
2) G-protein coupled receptor signaling pathway (GO:0007186): A series of molecular signals that proceeds with an activated receptor promoting the exchange of GDP for GTP on the alpha-subunit of an associated heterotrimeric G-protein complex
3) transcription, DNA-templated (GO:0006351): The cellular synthesis of RNA on a template of DNA
4) Axonogenesis (GO:0007409): De novo generation of a long process of a neuron, that carries efferent (outgoing) action potentials from the cell body towards target cells. Refers to the morphogenesis or creation of shape or form of the developing axon
5) Golgi vesicle transport (GO:0048193): The directed movement of substances into, out of or within the Golgi apparatus, mediated by vesicles
6) negative regulation of Wnt signaling pathway (GO:0030178): Any process that stops, prevents, or reduces the frequency, rate or extent of the Wnt signaling pathway
7) Nucleosome assembly (GO:0006334): The aggregation, arrangement and bonding together of a nucleosome, the beadlike structural units of eukaryotic chromatin composed of histones and DNA
8) Protein transport (GO:0015031): The directed movement of proteins into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore
9) Cellular protein localization (GO:0034613): Any process in which a protein is transported to, and/or maintained in, a specific location at the level of a cell
10) Protein modification by small protein conjugation or removal (GO:0070647): A protein modification process in which one or more groups of a small protein, such as ubiquitin or a ubiquitin-like protein, are covalently attached to or removed from a target protein

Ranking the shared set of genes according to the logistic regression classifier resulted into these 10 genes as the most contributing ones to be GBM: TAF15, CTSZ, APP, LIMK1, CCDC71, UBE2E1, KIAA0754, C5orf5, SLC39A7, TNAP and the following genes as the most contributing ones for a sample to be a LGG: VEPH1, SFRP2, UNC84B, PPP1R12C, DOHH, TMEM28, NUMA1, RP11-78J21.1, STIP1, ZNF768

IV. Discussion

In this study, we have explored the genes contributing to the aggressiveness of glioma and the possible genes that contribute to the evolution of LGG to GBM. Running t-test on the TCGA data gave lots genes that have significant change of expression in between LGG and GBM that could be explored later to further understand how they contribute to this aggressiveness. The split of the Data into two groups first of GCIMP+ and GCIMP- allowed us to get rid of the change of genes that could contribute to those subtypes in both LGG and GBM as they are shared subtypes of LGG and GBM.

This study gives guidance to what genes contributing more to the aggressiveness of gliomas and the involvement of many of the found genes in other types of cancer as lung cancer (C1orf142) and the presence of some genes as (LIMK1) in both the highest ranking genes from logistic regression and the most significant genes from the GCIM+ group along with its importance for axonal growth could be an indication that those genes are contributing factors in gliomas as well. Those highly ranked genes that are contributing to GBMs more than LGGs would need further genomic studies to prove their involvement in the aggressiveness of gliomas.

References

[1] "Comprehensive genomic characterization defines human glioblastoma genes and core pathways," *Nature*, vol. 455, no. 7216, pp. 1061–1068, Sep. 2008. [Online]. Available: http://dx.doi.org/10.1038/nature07385

[2] R. G. Verhaak, K. A. Hoadley, E. Purdom, V. Wang, Y. Qi, M. D. Wilkerson, C. R. Miller, L. Ding, T. Golub, J. P. Mesirov, G. Alexe, M. Lawrence, M. O'Kelly, P. Tamayo, B. A. Weir, S. Gabriel, W. Winckler, S. Gupta, L. Jakkula, H. S. Feiler, J. G. Hodgson, C. D. James, J. N. Sarkaria, C. Brennan, A. Kahn, P. T. Spellman, R. K. Wilson, T. P. Speed, J. W. Gray, M. Meyerson, G. Getz, C. M. Perou, D. N. Hayes, and Cancer Genome Atlas Research Network, "Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1." *Cancer cell*, vol. 17, no. 1, pp. 98–110, Jan. 2010. [Online]. Available: http://dx.doi.org/10.1016/j.ccr.2009.12.020

[3] H. Noushmehr, D. J. Weisenberger, K. Diefes, H. S. Phillips, K. Pujara, B. P. Berman, F. Pan, C. E. Pelloski, E. P. Sulman, K. P. Bhat, R. G. W. Verhaak, K. A. Hoadley, D. N. Hayes, C. M. Perou, H. K. Schmidt, L. Ding, R. K. Wilson, D. Van Den Berg, H. Shen, H. Bengtsson, P. Neuvial, L. M. Cope, J. Buckley, J. G. Herman, S. B. Baylin, P. W. Laird, and K. Aldape, "Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma," *Cancer Cell*, vol. 17, no. 5, pp. 510–522, Mar. 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ccr.2010.03.017

[4] Cancer Genome Atlas Research Network, D. J. Brat, R. G. Verhaak, K. D. Aldape, A. K. Yung, S. R. Salama, L. A. Cooper, E. Rheinbay, C. R. Miller, M. Vitucci, O. Morozova, A. G. Robertson, H. Noushmehr, P. W. Laird, A. D. Cherniack, R. Akbani, J. T. Huse, G. Ciriello, L. M. Poisson, J. S. Barnholtz-Sloan, M. S. Berger, C. Brennan, R. R. Colen, H. Colman, A. E. Flanders, C. Giannini, M. Grifford, A. Iavarone, R. Jain, I. Joseph, J. Kim, K. Kasaian, T. Mikkelsen, B. A. Murray, B. P. P. O'Neill, L. Pachter, D. W. Parsons, C. Sougnez, E. P. Sulman, S. R. Vandenberg, E. G. Van Meir, A. von Deimling, H. Zhang, D. Crain, K. Lau, D. Mallery, S. Morris, J. Paulauskis, R. Penny, T. Shelton, M. Sherman, P. Yena, A. Black, J. Bowen, K. Dicostanzo, J. Gastier-Foster, K. M. Leraas, T. M. Lichtenberg, C. R. Pierson, N. C. Ramirez, C. Taylor, S. Weaver, L. Wise, E. Zmuda, T. Davidsen, J. A. Demchok, G. Eley, M. L. Ferguson, C. M. Hutter, K. R. Mills Shaw, B. A. Ozenberger, M. Sheth, H. J. Sofia, R. Tarnuzzer, Z. Wang, L. Yang, J. C. C. Zenklusen, B. Ayala, J. Baboud, S. Chudamani, M. A. Jensen, J. Liu, T. Pihl, R. Raman, Y. Wan, Y. Wu, A. Ally, J. T. Auman, M. Balasundaram, S. Balu, S. B. Baylin, R. Beroukhim, M. S. Bootwalla, R. Bowlby, C. A. Bristow, D. Brooks, Y. Butterfield, R. Carlsen, S. Carter, L. Chin, A. Chu, E. Chuah, K. Cibulskis, A. Clarke, S. G. Coetzee, N. Dhalla, T. Fennell, S. Fisher, S. Gabriel, G. Getz, R. Gibbs, R. Guin, A. Hadjipanayis, D. N. Hayes, T. Hinoue, K. Hoadley, R. A. Holt, A. P. Hoyle, S. R. Jefferys, S. Jones, C. D. Jones, R. Kucherlapati, P. H. Lai, E. Lander, S. Lee, L. Lichtenstein, Y. Ma, D. T. Maglinte, H. S. Mahadeshwar, M. A. Marra, M. Mayo, S. Meng, M. L. Meyerson, P. A. Mieczkowski, R. A. Moore, L. E. Mose, A. J. Mungall, A. Pantazi, M. Parfenov, P. J. Park, J. S. Parker, C. M. Perou, A. Protopopov, X. Ren, J. Roach, T. S. Sabedot, J. Schein, S. E. Schumacher, J. G. Seidman, S. Seth, H. Shen, J. V. Simons, P. Sipahimalani, M. G. Soloway, X. Song, H. Sun, B. Tabak, A. Tam, D. Tan, J. Tang, N. Thiessen, T. Triche, D. J. Van Den Berg, U. Veluvolu, S. Waring, D. J. Weisenberger, M. D. Wilkerson, T. Wong, J. Wu, L. Xi, A. W. Xu, L. Yang, T. I. Zack, J. Zhang, B. A. Aksoy, H. Arachchi, C. Benz, B. Bernard, D. Carlin, J. Cho, D. DiCara, S. Frazer, G. N. Fuller, J. Gao, N. Gehlenborg, D. Haussler, D. I. Heiman, L. Iype, A. Jacobsen, Z. Ju, S. Katzman, H. Kim, T. Knijnenburg, R. B. B. Kreisberg, M. S. Lawrence, W. Lee, K. Leinonen, P. Lin, S. Ling, W. Liu, Y. Liu, Y. Liu, Y. Lu, G. Mills, S. Ng, M. S. Noble, E. Paull, A. Rao, S. Reynolds, G. Saksena, Z. Sanborn, C. Sander, N. Schultz, Y. Senbabaoglu, R. Shen, I. Shmulevich, R. Sinha, J. Stuart, S. O. Sumer, Y. Sun, N. Tasman, B. S. Taylor, D. Voet, N. Weinhold, J. N. Weinstein, D. Yang, K. Yoshihara, S. Zheng, W. Zhang, L. Zou, T. Abel, S. Sadeghi, M. L. Cohen, J. Eschbacher, E. M. Hattab, A. Raghunathan, M. J. Schniederjan, D. Aziz, G. Barnett, W. Barrett, D. D. Bigner, L. Boice, C. Brewer, C. Calatozzolo, B. Campos, C. G. G. Carlotti, T. A. Chan, L. Cuppini, E. Curley, S. Cuzzubbo, K. Devine, F. DiMeco, R. Duell, J. B. Elder, A. Fehrenbach, G. Finocchiaro, W. Friedman, J. Fulop, J. Gardner, B. Hermes, C. Herold-Mende, C. Jungk, A. Kendler, N. L. Lehman, E. Lipp, O. Liu, R. Mandt, M. McGraw, R. Mclendon, C. McPherson, L. Neder, P. Nguyen, A. Noss, R. Nunziata, Q. T. Ostrom, C. Palmer, A. Perin, B. Pollo, A. Potapov, O. Potapova, W. K. Rathmell, D. Rotin, L. Scarpace, C. Schilero, K. Senecal, K. Shimmel, V. Shurkhay, S. Sifri, R. Singh, A. E. Sloan, K. Smolenski, S. M. Staugaitis, R. Steele, L. Thorne, D. P. Tirapelli, A. Unterberg, M. Vallurupalli, Y. Wang, R. Warnick, F. Williams, Y. Wolinsky, S. Bell, M. Rosenberg, C. Stewart, F. Huang, J. L. Grimsby, A. J. Radenbaugh, and J. Zhang, "Comprehensive, integrative genomic analysis of diffuse Lower-Grade gliomas." *The New England journal of medicine*, vol. 372, no. 26, pp. 2481–2498, Jun. 2015. [Online]. Available: http://view.ncbi.nlm.nih.gov/pubmed/26061751

[5] M. Ceccarelli, F. P. Barthel, T. M. Malta, T. S. Sabedot, S. R. Salama, B. A. Murray, O. Morozova, Y. Newton, A. Radenbaugh, S. M. Pagnotta, S. Anjum, J. Wang, G. Manyam, P. Zoppoli, S. Ling, A. A. Rao, M. Grifford, A. D. Cherniack, H. Zhang, L. Poisson, C. G. Carlotti, D. P. Tirapelli, A. Rao, T. Mikkelsen, C. C. Lau, W. K. A. Yung, R. Rabadan, J. Huse, D. J. Brat, N. L. Lehman, J. S. Barnholtz-Sloan, S. Zheng, K. Hess, G. Rao, M. Meyerson, R. Beroukhim, L. Cooper, R. Akbani, M. Wrensch, D. Haussler, K. D. Aldape, P. W. Laird, D. H. Gutmann, H. Noushmehr, A. Iavarone, and R. G. W. Verhaak, "Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma," *Cell*, vol. 164, no. 3, pp. 550–563, Jan. 2016. [Online]. Available: http://dx.doi.org/10.1016/j.cell.2015.12.028

[6] S. Carbon, A. Ireland, C. J. Mungall, S. Shu, B. Marshall, S. Lewis, the AmiGO Hub, and the Web Presence Working Group, "AmiGO: online access to ontology and annotation data," *Bioinformatics*, vol. 25, no. 2, pp. 288–289, Jan. 2009. [Online]. Available: http://dx.doi.org/10.1093/bioinformatics/btn615

[7] W. Sun, C. Guo, X. Meng, Y. Yu, Y. Jin, D. Tong, J. Geng, Q. Huang, J. Qi, A. Liu, R. Guan, L. Xu, D. Sun, W. Ji, P. Liu, F. Liu, H. Sun, G. Ji, S. Fu, and J. Bai, "Differential expression of PAI-RBP1, c1orf142, and COTL1 in non-small cell lung cancer cell lines with different tumor metastatic potential." *Journal of investigative medicine : the official publication of the American Federation for Clinical Research*, vol. 60, no. 4, pp. 689–694, Apr. 2012. [Online]. Available: http://view.ncbi.nlm.nih.gov/pubmed/22373659

[8] Y. Mao, A. Fu, A. E. Hoffman, D. I. Jacobs, M. Jin, K. Chen, and Y. Zhu, "The circadian gene CRY2 is associated with breast cancer aggressiveness possibly via epigenomic modifications." *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*, vol. 36, no. 5, pp. 3533–3539, May 2015. [Online]. Available: http://view.ncbi.nlm.nih.gov/pubmed/25740058

[9] S. A. Bergeron, N. Carrier, G. H. Li, S. Ahn, and H. A. Burgess, "Gsx1 expression defines neurons required for prepulse inhibition." *Molecular psychiatry*, vol. 20, no. 8, pp. 974–985, Aug. 2015. [Online]. Available: http://view.ncbi.nlm.nih.gov/pubmed/25224259