OMICS RESEARCH SYMPOSIUM

Understanding The Expression of LncRNAs in Glioblastoma Multiforme Using Gepia

Amity Institute of Biotechnology, Amity University Uttar Pradesh, Noida

Hardik Gupta[1] , Sohini Singh[2]

Introduction

The least studied transcript among non-coding RNA is lncRNA, whose lengths vary from 200 nucleotides to 100 kb. Many lncRNAs are found in the human genome and these are classified into 5 major types depending on the genomic location at which they are found. Those present between 2 adjacent genes is named intergenic lncRNAs while the lncRNAs are present in between the 2 exons of a particular gene or we can say if it is present completely in 1 intron of the protein-coding gene then it is known as an intronic lncRNA. Sense lncRNAs are present on the sense strand coding gene. These may be overlapping the associated exons of the gene. Contrary to this is the antisense lncRNAs. The fifth type of long non-coding RNA found is called bidirectional lncRNA and they are found to have sequences present on the opposite strands of the adjacent genes and the start codons of these genes separated by less than 1000 base pairs.

Molecular studies of the lncRNAs and their associated interactions showed that the main reason behind the role of lncRNA in the cellular mechanism is due to the interaction of these non-coding RNAs and other biomolecules which are coding transcripts (mRNA), proteins, DNA, and other non-coding RNAs (snRNA, piRNA). In lncRNA – mRNA interaction, lncRNA can be affecting the expression of the RNA directly or indirectly, they come in direct proximity to the mRNA and are seen to affect the expression and functionality of the genes. The interaction may be also indirect in which lncRNA and miRNA compete with each other in binding with their common mRNA. One of the major ways by which they interact indirectly is by forming a competing endogenous RNA system also known as ceRNA and maintain an intelligent competing system by which the gene expression (like muscle differentiation) is regulated.

Another important interaction happening is between lncRNA and DNA, in this lncRNA may be forming an R-loop with the DNA which changes the basic orientation of the DNA sequence causing the changes in gene expression. It is also seen that a Triplex is formed between the DNA and corresponding lncRNA causing changes in the gene expression.

Alongside this interaction of the lncRNA is also seen to be present with the protein binding to the DNA which has been seen to increase the association of the protein with their corresponding gene. The next interaction which takes place is between lncRNA and other non-coding RNAs. The lncRNA has been found to attach with other ncRNAs like snoRNA, miRNA, etc.

Micro RNA has been found to directly impact the stability of the lncRNA. miRNA like miR-34a has been found to directly interact with HOTAIR which thus impacts the expression of HOTAIR. Lastly interaction between the proteins and lncRNA has also been found profoundly. These lncRNAs are acting as the decoy, signal, scaffold, or guides for the proteins. lncRNA CCND1 is found to be induced in case of DNA damage and this directly recruits the TLS protein on the promoter of the CCND1. The stability of lncRNAs has also been checked by the proteins, for example, PTBP1 stabilizes the lncRNA MACC1-AS1. Interactions between these 2 have been found to regulate gene expression as shown in a study by Daneshvar K, Ardehali MB, Klein IA, et al. Interaction of the protein-lncRNA forming a ceRNA system had also been studied which impact the expression of mRNA

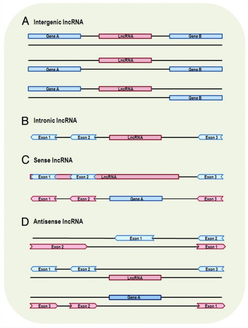


Figure 1: Different types of the lncRNAs classified based on their position with respect to their associated mRNA. (a) Intergenic lncRNA (b) Intronic lncRNA (c) Sense lncRNA (d) Antisense lncRNA
(Ma, Lina et at. 2013)

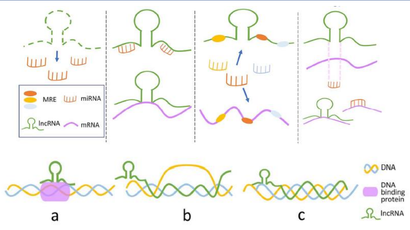


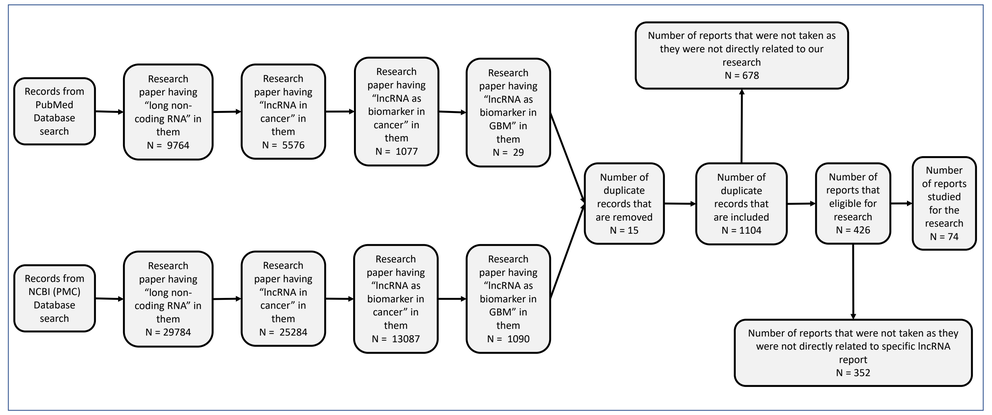
Figure 2: The different types of the lncRNA – RNA interaction (above) and lncRNA–DNA Interaction (below) are depicted in this. (Zhang et al. 2020)

Methods

A present study on long non-coding RNA and Glioblastoma Multiforme (GBM) was performed. Inclusion criteria set included the lncRNA having association with the GBM. Another inclusion criterion was that the lncRNA must be having significance in working as a biomarker. We searched NCBI and PubMed for the articles on "long non-coding" and a total of more than 25,000 results were found. Further, we narrowed our search by sorting the article which we having keywords as "long non-coding RNA" and "cancer". After this, the search was further focused using the inclusion criteria set for the study and we considered the results in which the lncRNA given had a role as Biomarkers. Further, we specified cancer to the focused cancer of our study which was Glioblastoma Multiforme. We got approximately 1000 results from NCBI (PMC) and 29 results from PubMed. We further removed the articles which we there in both the searches and then sorted the results based on their abstract and keywords. At last, we got down our collection of papers to 30 articles.

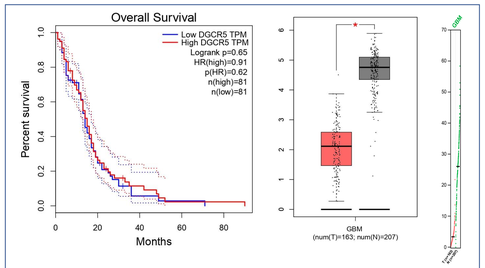
We used the GEPIA portal to produce various plots and graphs for the analysis. GEPIA is an online interactive web server that works on a vast dataset taken from TCGA and GTEx projects. It takes their various normal and tumor sample for the analyses of the RNA expression. The data from these two sources is then analyzed based on the various commands which are been provided by the researcher on the portal of GEPIA. GEPIA allows the user to receive an output based on a wide variety of customization tools that are being provided at the portal. The customizations that are being provided by GEPIA are like normal/tumor differential illustration, the type of cancer or disease-dependent outlining, correlation study, dimension reduction analysis, comparable gene discovery, and patient survival study. We generated results for the five lncRNAs of interest. These lncRNAs were HOTAIR, DGCR5, DLEU1, AGAP2-AS1, and MIAT. We used the TCGA-GBM project's cohort which contained a total of 207 normal tissue entries and along with them, 163 were of the tumorous tissue sample. GTEx dataset had a vast dataset of 945 entries which were targeted towards different parts of the brain. All the transcripts of interest were run separately on the portal with the desired conditions applied for the analysis. Graphical representation of the transcript's Gene expression was obtained for various cancers for normal and tissue samples. Along with this, their particular gene expression profile was also obtained which showed the presence of the transcripts per million (TPM) which were observed in the tumor and normal samples. A box plot was also generated for all the samples depicting an average value of the gene expression. Then a survival plot of the affected patients was generated to understand the role of the transcript in the prognosis of the disease and to understand how the gene is affecting the patients directly in regards to their mean expected survival. The survival plot was obtained depicting the cases showing high or low TPM values.

The graphical data was then analyzed to understand the expression of the genes and they were compared with each other to find whether there is any relationship amongst the lncRNAs taken up for the study.



Flow chart for study design – Using various inclusion and exclusion criteria research articles were selected and further study on Glioblastoma Multiforme was done on GEPIA.

Results

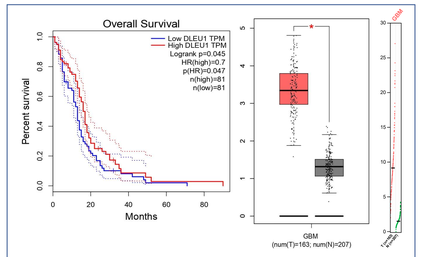


The lncRNA **DGCR5 (DiGeorge syndrome critical region gene 5)** is studied profoundly in association with the Glioblastoma multiforme and shown that the lncRNA expression is down-regulated remarkably.

Survival plot, Box plot and Expression profile of the lncRNA DGCR5

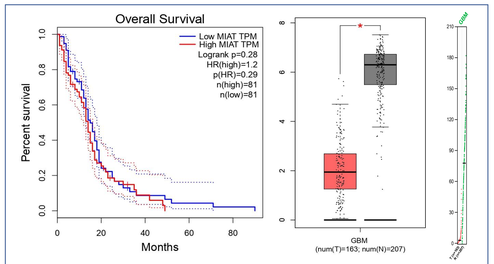
A deeper analysis of the studied articles provided us with a set of five lncRNA which were taken for our study. These lncRNAs were HOTAIR, DGCR5, DLEU1, AGAP2-AS1, and MIAT. These were taken as they qualified all the inclusion criteria which were set up and these were found to have a direct impact on disease. The gene expressions of all these were obtained using the portal of GEPIA.

- DGCR5 and MIAT both have shown down-regulation in the tumor population while DLEU1 and AGAP2-AS1 have shown up-regulation in the case of tumor population.
- HOTAIR where it is shown that the GEPIA portal is unable to calculate any specific significant relation between the expression in the tumor tissue and the expression in the normal conditions.
- Variation in the prognosis predicted (based on the gene expression in the tumor population) was seen in the case of MIAT and DLEU1.
- In the Survival plots of the 5 transcripts studied the initial pattern of the prognosis of these transcripts is the same but all the changes in the curve can be seen in the middle or later stages of prognosis.



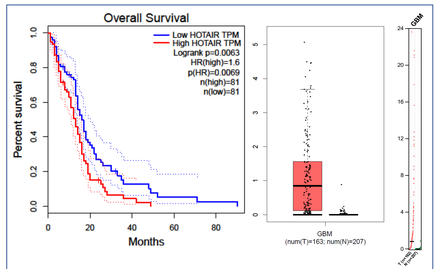
Deleted In Lymphocytic Leukemia 1 (DLEU1) is a long non-coding RNA that has been found as an activator of the proliferation in GBM and it can be put forward as an efficient target for future therapies. It has been suggested as a biomarker of the prognosis of the GBM.

Survival plot, Box plot and Expression profile of the lncRNA DLEU1



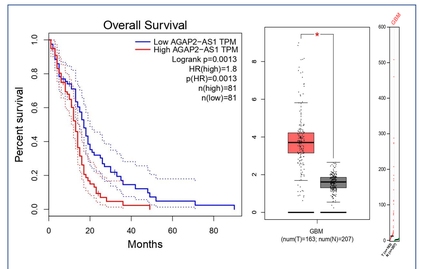
Myocardial Infarction Associated Transcript, MIAT long non-coding RNA whose involvement in various hallmark processes of cancer has been surfaced in few years. It has been believed that the lncRNA can be used as a prognostic biomarker and this is possibly acting as an oncogene in humans

Survival plot, Box plot and Expression profile of the lncRNA MIAT



HOTAIR (HOX Transcript Antisense Intergenic RNA) expression in the tumor has been consider as potent biomarker for the diagnosis of Glioblastoma Multiforme.

Survival plot, Box plot and Expression profile of the lncRNA HOTAIR



AGAP2 Antisense RNA-1 (AGAP-AS1) is a long non-coding RNA that is linked with the AGAP2 protein. This protein is seen to have a significant role in the migration of the cells.

Survival plot, Box plot and Expression profile of the lncRNA AGAP2-AS1

Conclusions

Glioblastoma multiforme is the most aggressive form of brain tumor. In recent times, a significant part of lncRNA in the manifestation and prognosis of the disease is been discovered very intensively. The interaction of the lncRNA and coding RNA (mRNA) was one of the very first steps in grasping the role of these non-coding transcripts in the GBM and further studies found a direct impact of these lncRNAs on tumor angiogenesis, cell migration, metastasis, and proliferation. In our study, we took 5 of these lncRNAs; HOTAIR, DGCR5, DLEU1, AGAP2-AS1, and MIAT; and we studied their gene expression using the GEPIA portal. The lncRNAs have been chosen based on our set inclusion criteria. Our analysis found varied results and some of the interesting questions which must be answered with further studies hypothesizing the newly found ideas.

One of the major outcomes which are shown by the GEPIA in the box plot of the expression of the HOTAIR where it is shown that the GEPIA portal is unable to calculate any specific significant relation between the expression in the tumor tissue and the expression in the normal conditions. During our study, we were not able to find any paper which has signified the result that there might not be a direct relationship among cases of the tumor and the normal population of the HOTAIR patients.

LncRNA MIAT is down-regulated in the tumor population of the GBM patients. A similar result was gathered in our study on GEPIA. This can be thus interpreted that the over-expression of the lncRNA will be helping in the prognosis of the disease and this can be supported also by the fact that it helps in better permeability of the blood-tumor barrier. In contrast to this, an interesting result was seen in the survival plot which has shown that the patients those were having lower levels of the transcript improved their prognosis in the later stages of the treatment. This hints at an association of some other transcript or some other mechanism in which the low levels of MIAT are helping in the late prognosis period. A slightly similar result was seen in the case of DLEU1 too where the lncRNA is up-regulated in the tumoral population of the taken patient dataset on the GEPIA. This up-regulation was shown in the profile plot as well as the cancer distribution plot. But to our surprise, almost a similar prognosis was seen in the low and high levels of DLEU1, and in fact, a better prognosis was seen in the case of high expression of the lncRNA DLEU1.

An observation was seen in our results, in the survival plots of the 5 transcripts studied the initial prognosis of these transcripts was the same but all the survival patterns change in the middle or later stages of life. This hints that the effect of lncRNA expression might be delayed or lower and gradually the expression is accumulated and with time the effect on the prognosis is seen.

The study has shown a significant correlation of the lncRNAs with Glioblastoma Multiforme which were analyzed based on their expression in tumor and normal populations. We found that there is a significant up-regulation of the AGAP2-AS1 and DLEU1 in the tumor population available on the GEPIA. Also, we discovered that a significant down-regulation of the DGCR5 and MIAT is seen in the tumor cells of the GBM. Further, we are focused to understand more about the relation between these pairs of the long non-coding RNAs which have been showing a similar expression in the GBM cases and to find that whether any significant relation is present between the 2 and if there is a relationship, then do they share a similar pathway or are a part of a bigger network of various ncRNA, mRNA, and proteins working together in giving the final effect on the Glioblastoma patients.

In our study, we got a few novel observations in our study which we intend to look upon in further greater detail. We aim to study each of the taken transcripts individually and study them with a wider population of datasets. We also plan to further investigate the observations which have been seen in the survival plots and to find the reason behind the delayed effect of the long non-coding RNA expression in the survival of the patients with Glioblastoma multiforme. Our study has put forward some very interesting questions which must be pondered upon and further more studies are required in this field and the novel observations found in the study must be investigated which might help us improve our knowledge in the manifestation and prognosis of cancer in general. This will help us in coming up with newer more effective targets for the treatment of the disease in the future.

References

- Francis, C. (1970). Central dogma of molecular biology. Nature, 227(5258), 561-563.
- Grisolia S. (1989). Mapping the human genome. The Hastings Center report, 19(4), S18–S19.
- Hauptman, N., & Glavač, D. (2013). Long non-coding RNA in cancer. International journal of molecular sciences, 14(3), 4655–4669.
- Ponting, C. P., Oliver, P. L., & Reik, W. (2009). Evolution and functions of long noncoding RNAs. Cell, 136(4), 629-641.
- Ma, L., Bajic, V. B., & Zhang, Z. (2013). On the classification of long non-coding RNAs. RNA biology, 10(6), 924-933.
- Novikova, I. V., Hennelly, S. P., & Sanbonmatsu, K. Y. (2012). Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. Nucleic acids research, 40(11), 5034-5051.
- Pickard, M. R., & Williams, G. T. (2015). Molecular and cellular mechanisms of action of tumour suppressor GAS5 LncRNA. Genes, 6(3), 484-499.
- Chiyomaru, T., Yamamura, S., Fukuhara, S., Yoshino, H., Kinoshita, T., Majid, S., ... & Dahiya, R. (2013). Genistein inhibits prostate cancer cell growth by targeting miR-34a and oncogenic HOTAIR. PloS one, 8(8), e70372.
- Yang, L., Tang, Y., Xiong, F., He, Y., Wei, F., Zhang, S., ... & Zeng, Z. (2018). LncRNAs regulate cancer metastasis via binding to functional proteins. Oncotarget, 9(1), 1426.
- Zhang, Y., Jia, C., & Kwoh, C. K. (2021). Predicting the interaction biomolecule types for lncRNA: an ensemble deep learning approach. Briefings in Bioinformatics, 22(4), bbaa228.
- Lee, S. H., & Jun, B. H. (2019). Silver nanoparticles: synthesis and application for nanomedicine. International journal of molecular sciences, 20(4), 865.
- Qiu, X., Lei, Z., Wang, Z., Xu, Y., Liu, C., Li, P., ... & Gong, Z. (2019). Knockdown of LncRNA RHPN1-AS1 inhibits cell migration, invasion and proliferation in head and neck squamous cell carcinoma. Journal of Cancer, 10(17), 4000.
- He, B., Peng, F., Li, W., & Jiang, Y. (2019). Interaction of lncRNA-MALAT1 and miR-124 regulates HBx-induced cancer stem cell properties in HepG2 through PI3K/Akt signaling. Journal of cellular biochemistry, 120(3), 2908-2918.
- Zhang, Q., Li, T., Wang, Z., Kuang, X., Shao, N., & Lin, Y. (2020). lncRNA NR2F1-AS1 promotes breast cancer angiogenesis through activating IGF-1/IGF-1R/ERK pathway. Journal of cellular and molecular medicine, 24(14), 8236-8247.
- Zhao, J., Jin, W., Yi, K., Wang, Q., Zhou, J., Tan, Y., ... & Kang, C. (2021). Combination LSD1 and HOTAIR-EZH2 inhibition disrupts cell cycle processes and induces apoptosis in glioblastoma cells. Pharmacological Research, 171, 105764.
- Zhang, Y., Zheng, L., Xu, B. M., Tang, W. H., Ye, Z. D., Huang, C., ... & Wang, Q. (2018). Lnc RNA-RP 11-714G18. 1 suppresses vascular cell migration via directly targeting LRP 2 BP. Immunology and cell biology, 96(2), 175-189.
- Faghihi, M. A., Zhang, M., Huang, J., Modarresi, F., Van der Brug, M. P., Nalls, M. A., ... & Wahlestedt, C. (2010). Evidence for natural antisense transcript-mediated inhibition of microRNA function. Genome biology, 11(5), 1-13.
- Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., ... & Bozzoni, I. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell, 147(2), 358-369.
- Wang, Y., Hou, J., He, D., Sun, M., Zhang, P., Yu, Y., & Chen, Y. (2016). The emerging function and mechanism of ceRNAs in cancer. Trends in Genetics, 32(4), 211-224.
- Han, M., Gu, Y., Lu, P., Li, J., Cao, H., Li, X., ... & Dong, H. (2020). Exosome-mediated lncRNA AFAP1-AS1 promotes trastuzumab resistance through binding with AUF1 and activating ERBB2 translation. Molecular cancer, 19(1), 1-18.
- Tan, S. K., Pastori, C., Penas, C., Komotar, R. J., Ivan, M. E., Wahlestedt, C., & Ayad, N. G. (2018). Serum long noncoding RNA HOTAIR as a novel diagnostic and prognostic biomarker in glioblastoma multiforme. Molecular cancer, 17(1), 1-7.
- Zhou, X., Ren, Y., Zhang, J., Zhang, C., Zhang, K., Han, L., ... & Kang, C. (2015). HOTAIR is a therapeutic target in glioblastoma. Oncotarget, 6(10), 8353.
- Yang, F., & Huang, Y. L. (2019). DGCR5 suppresses the EMT of pediatric primary glioblastoma multiforme cell and serves as a prognostic biomarker. European review for medical and pharmacological sciences, 23(22), 10024-10034.
- Wu, X., Hou, P., Qiu, Y., Wang, Q., & Lu, X. (2020). Large-scale analysis reveals the specific clinical and immune features of DGCR5 in glioma. OncoTargets and therapy, 13, 7531.
- Bountali, A., Tonge, D. P., & Mourtada-Maarabouni, M. (2019). RNA sequencing reveals a key role for the long non-coding RNA MIAT in regulating neuroblastoma and glioblastoma cell fate. International journal of biological macromolecules, 130, 878-891.
- He, J., Xue, Y., Wang, Q., Zhou, X., Liu, L., Zhang, T., ... & Ma, T. (2020). Long non-coding RNA MIAT regulates blood tumor barrier permeability by functioning as a competing endogenous RNA. Cell death & disease, 11(10), 1-18.
- Wang, J., Quan, X., Peng, D., & Hu, G. (2019). Long non-coding RNA DLEU1 promotes cell proliferation of glioblastoma multiforme. Molecular medicine reports, 20(2), 1873-1882.
- Song, C., Zhang, J., Zhao, Z., Yang, Y., Meng, D., Wang, J., ... & Yuan, C. (2020). DLEU1: a functional long noncoding RNA in tumorigenesis. Current pharmaceutical design, 26(15), 1742-1748.
- Tian, Y., Zheng, Y., & Dong, X. (2019). AGAP2-AS1 serves as an oncogenic lncRNA and prognostic biomarker in glioblastoma multiforme. Journal of cellular biochemistry, 120(6), 9056-9062.
- Luo, W., Li, X., Song, Z., Zhu, X., & Zhao, S. (2019). Long non-coding RNA AGAP2-AS1 exerts oncogenic properties in glioblastoma by epigenetically silencing TFPI2 through EZH2 and LSD1. Aging (Albany NY), 11(11), 3811.

We are thankful to Amity University Uttar Pradesh, Noida for giving us this opportunity to conduct this study.