Bioinformatics and Prediction Analysis of Long Non-coding RNAs in the Regulation of Glioblastoma Cell Migration and Invasion

Dexter Hoi Long Leung

Monash University Malaysia

Siti Ayuni Hassanudin

Monash University Malaysia

Muhamad N.A Kamarudin

Monash University Malaysia

Shogo Moriya

Chiba University Graduate School of Medicine

Tomoko Soga

Monash University Malaysia

Stephen Navendran Ponnampalam

Hospital KL

Manjeevan Seera

Monash University Malaysia

Ishwar Parhar (✉ ishwar@monash.edu)

Monash University Malaysia

Research Article

Keywords: cell-invasion, long-ncRNA, glioblastoma, microarray, bioinformatics

Posted Date: July 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-665258/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Long non-coding RNA (lncRNA) has been identified in the regulation of cell-invasion via various cellular processes. Research on lncRNAs is however lacking in the context of Glioblastoma (GBM), which held-up identification of GBM-related lncRNAs via bioinformatic databases. This study intends to identify potential lncRNA candidates in the regulation of GBM cell-invasion based on target miRNA and mRNA/proteins. Microarray was carried out to identify upregulated lncRNA in GBM and astrocytoma cells, followed by bioinformatics-based prediction analysis to identify lncRNAs involved in cellular processes related to GBM cell-invasion. A total of 372 and 806 upregulated lncRNA were identified in GBM and astrocytoma cells. From these lncRNAs, 8 lncRNAs were predicted based on functions of target proteins; LINC00221 associated with EMT, LINC01564 and LINC00265 with angiogenesis, and LOC100240735 with cell-migration. Network analysis of lncRNA based on function of target miRNAs predicted LINC00999, ALOX12-AS1, CHKB-AS1, and LINC01588 as regulators of angiogenesis, with LINC00482, LINC00239, and LINC01003 for regulation of EMT. Based on the network analysis, three novel lncRNA-miRNA-mRNA interacting axis from NEAT1, CRNDE, and SNHG1 were predicted to regulate GBM invasion. These findings have allowed prediction of lncRNA as regulators of GBM cell-invasion which can be considered as potential candidates for experimental studies as a GBM prognostic biomarkers.

Introduction

Glioblastoma (GBM) is one of the most aggressive types of malignant brain tumor, where approximately 3 to 7% of patients recorded overall-survival more than five years¹. The hallmark of GBM is its high invasiveness, which leads to low overall-survival and high recurrence post-therapy in GBM patients. The cause for high malignancy in GBM has been widely researched and non-coding RNAs (ncRNAs) have been identified to play specific roles in the development and progression of cancers, including GBM². A class of ncRNAs, long non-coding RNAs (lncRNAs) is established to influence the malignancy processes by acting as signal, decoys, guides, or scaffolds molecules at different cellular and molecular level which regulate protein translation and activity³.

Malignancy in GBM can be defined via the tumors ability to migrate and invade the surrounding cells. With ever-expanding number of lncRNAs currently being identified and the complex nature of tumorigenesis, questions remain unanswered in how these lncRNAs affect the different cellular and molecular activities in GBM. With the advancement of sequencing technologies, more non-coding RNA sequences, including lncRNAs, and microRNA (miRNAs), have been identified within the human genome. In addition, lncRNAs could be prognostic and therapeutic markers of cancer as they act as molecular sponges with miRNA, which affects miRNA to not being able to translationally silences oncogenic mRNA (Fig. 1). One of the more studied lncRNA in cancer and GBM is MALAT1, which plays multiple biological roles leading to cells exhibiting cancer characteristics⁴. While studies show that lncRNA plays a role in cell proliferation and apoptosis in GBM, they also play a key role in influencing GBM cell migration and invasion in GBM^{5,6}. Nevertheless, more in-depth studies focused on the intrinsic mechanisms of lncRNA

could contribute to the high invasiveness observed in GBM, which would allow them to be considered as beneficial potential GBM biomarkers.

Invasive and migrative capabilities of GBM are associated with several cell features and tumour extracellular matrix (ECM) through the remodelling of the cell cytoskeleton⁷. GBM tumours are also highly vascularised, requiring the involvement of angiogenesis during tumour cell invasion. Angiogenesis can be triggered by hypoxia, and the lack of oxygen will trigger the release of pro-angiogenic factors⁸. Release of angiogenic factors vascular endothelial growth factor (VEGF) and endothelial growth factor (EGF) can activate their respective receptors, which subsequently modulate the Phosphoinositide 3-kinase (PI3K) and Mitogen-activated protein kinases (MAPKs) pathways⁹. The overactivation of these pathways can lead to the promotion of hypoxia-induced angiogenesis, which has been observed in GBM cell migration and invasion¹⁰. Hypoxia has been established to promote GBM invasion and is one of the major drivers for GBM progression. The low tumour oxygenation (hypoxia) stimulates GBM tumours to escape this adverse tumour microenvironment and invade nearby healthy brain tissues. In addition, tumour neovascularisation and expansion can increase the oxygen diffusion distance, creating a hypoxic microenvironment that further promotes GBM tumour invasion. Hypoxia-inducible factor (HIF) is released in response to the hypoxic environment to mediate hypoxia-associated cellular responses¹¹.

This study aims to predict novel lncRNAs that could potentially play a specific role in contributing to GBM cell migration and invasion via various cellular processes. The prediction is based on the function of the target miRNA or proteins with which they were reported to interact. Microarray experiments were carried out on GBM and astrocytoma cell lines along with normal astrocytes to identify the upregulated lncRNAs. The microarray data was used along with bioinformatic databases to identify potential lncRNAs candidates specific to cell migration and invasion, and novel signaling molecules.

Methods

Cell Culture and RNA extraction

Normal Human Astrocyte cells (NHA, CC- 2565, Lonza US) were cultured using complete Astrocyte Basal Medium (ABM) (AGM, Lonza USA) supplemented with growth supplements 0.1% rhEGF, 0.25% Insulin, 0.1% Ascorbic Acid, 0.1% GA- 1000, 1% L- Glutamine, 3% FBS and with 1% penicillin/ streptomycin. SW 1783 cells (HTB-13, ATCC, US). These astrocytes, originated from a 68-year-old Caucasian male, were grown in Leibovitz's L- 15 medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/ streptomycin. LN18 cells were grown and maintained in RPMI medium supplemented with 10% FBS and 1% penicillin/ streptomycin. RNA was extracted and purified using the Monarch Total RNA Miniprep Kit (T2010S, New England Biolabs, USA) according to the manufacturer's instruction.

Microarray analysis

Total RNAs were subjected to *in vitro* amplification and labeled with Low Input Quick Amp Labeling Kit (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's instruction. Each labeled

RNA was mixed with the control RNA and hybridized to SurePrint G3 Mouse GE 8X60K (Agilent Technologies). Microarray was performed by Hokkaido System Science (Hokkaido, Japan). The microarrays were scanned using SureScan Microarray Scanner (Agilent Technologies). The scanned images were quantitatively analyzed with Feature Extraction Software v10.7.3.1 (Agilent Technologies). The expression data from the microarray were normalized using Linear and Locally Weighted Scatterplot Smoothing (LOWESS) method for p -value estimation for each spot. The q -value (false discovery rate, FDR) was calculated based on the p -value, with a threshold of less than 0.01.

Bioinformatic tools and databases

Transcripts from microarray were selected for lncRNA, performed through biomaRt¹². LNCipedia¹³ was used to check for alternate names and other additional information used for these lncRNA transcripts. These transcripts were analyzed for their protein-coding potential using the CPAT tool¹⁴. All the functional and cluster analysis of gene ontology and KEGG pathway were performed using Enrichr¹⁵. RAID database was utilized to obtain all relevant RNA interactions (lncRNA, miRNA, mRNA, proteins)¹⁶. TAM 2.0 database was utilized to obtain miRNAs which are associated with cellular processes leading to cell migration and invasion¹⁷. Cytoscape¹⁸, an open-source software platform, was used to visualize complex networks.

Analysis of lncRNA and prediction of lncRNA-Protein and lncRNA-miRNA interactions

The upregulated genes identified from the microarray were sorted out for lncRNA. The potential proteins, established to be affected by the lncRNAs, were obtained from the database and clustered based on Gene Ontology (GO) biological process annotations containing keywords associated with cell migration, angiogenesis, hypoxia, extracellular matrix, and epithelial-to-mesenchymal transition. The GO terms that were chosen to sort out the related proteins are available in **Supplementary S1**. A list of lncRNAs that interact with these proteins was obtained and subsequently filtered out with the microarray results.

Prediction of lncRNA-miRNA-protein interaction through LncTar

For lncRNAs that were identified to interact with any of the proteins associated with the GO terms, their sequence was first retrieved from Lncipedia, followed by retrieval of the sequence of the miRNA associated with the respective biological process from miRNAmeConverter¹⁹. The interaction of these lncRNAs and potential miRNAs were predicted through LncTar²⁰.

Prediction of lncRNA-miRNA-mRNA interaction

The information of the miRNAs that interacts with both lncRNAs and proteins from the lncRNA-mRNA interaction identified previously through Enrichr, were retrieved from TargetScan microRNA and miRTarBase database. These miRNAs were filtered with a list of miRNAs obtained from TAM2.0 with the following features: angiogenesis, cell adhesion, cell migration, cell motility, EMT, ECM, neuron

development, neural stem cell proliferation, and onco-miRNA. Python-based programming was used to find a list of predicted lncRNA-miRNA-protein interactions (**Supplementary S2**), lncRNA-miRNA interactions (**Supplementary S3**), and lncRNA-miRNA-mRNA interactions. lncRNA-miRNA-protein interaction output were then inserted into Cytoscape to look for distinct lncRNA-miRNA-protein signatures.

To summarize the general experimental workflow performed, the microarray was performed first, followed by identification of upregulated lncRNA. The functions of lncRNAs were predicted based on the biological roles of the proteins and miRNAs using bioinformatics databases. To further establish the lncRNA-miRNA-mRNA/protein interaction network, databases containing both lncRNA-miRNA and lncRNA-mRNA/protein interactions were utilised. For lncRNA without information regarding their interactions, lncTAR was utilised to predict interaction with miRNA responsible for cell migration and invasion (Fig. 2).

Results

Summary of microarray experiment results

The comparative analyses of LN18, SW 1783, and NHA cell lines with $p < 0.05$ and fold change > 2 provided 372 and 806 upregulated lncRNA transcripts (**Supplementary S4**), respectively, as shown in Fig. 3. The top-10 upregulated lncRNAs for both cell lines (LN18 and SW1783) are listed in Table 1. Next, the top 10 lncRNA genes that were significantly upregulated in LN18 against SW 1783, were identified and compared against NHA cells (Table 2, **Supplementary S5**).

Table 1
Top-10 upregulated lncRNA gene in LN18 and SW 1783 cell lines

Gene Name	Fold Change	P-value	FDR P-value
Glioblastoma			
LINC00221	127.40	3.52×10^{-23}	3.18×10^{-20}
ENST00000446495	108.40	8.12×10^{-23}	3.37×10^{-20}
LOC340340	66.28	4.80×10^{-22}	7.03×10^{-20}
FENDRR	60.29	9.01×10^{-23}	3.43×10^{-20}
DSCR8	57.91	9.74×10^{-23}	3.46×10^{-20}
lnc-EIF3M-2	55.03	1.44×10^{-21}	1.46×10^{-19}
LINC00470	42.68	1.18×10^{-20}	7.38×10^{-19}
LINC00871	41.45	1.63×10^{-20}	9.65×10^{-19}
LINC01296	30.48	8.47×10^{-19}	2.92×10^{-17}
ERICH2	29.43	1.46×10^{-18}	4.68×10^{-17}
Astrocytoma			
UCA1	54.98	1.64×10^{-21}	4.05×10^{-19}
SFTA1P	49.94	5.00×10^{-22}	1.77×10^{-19}
ENST00000412115	46.28	7.37×10^{-21}	1.27×10^{-18}
PLCE1-AS1	39.78	5.58×10^{-20}	6.66×10^{-18}
LINC01551	39.50	1.98×10^{-22}	1.07×10^{-19}
LINC01468	36.86	1.03×10^{-21}	2.94×10^{-19}
LOC100506827	36.63	4.05×10^{-22}	1.58×10^{-19}
FENDRR	35.69	1.21×10^{-19}	1.25×10^{-17}
C7orf69	35.25	2.91×10^{-19}	2.71×10^{-17}
HOTAIRM1	32.11	3.71×10^{-22}	1.51×10^{-19}

Table 2
Differences in fold-change of top 10 upregulated lncRNA identified in LN18 and SW 1783 cells

Gene Name	LN18	SW 1783	Difference in fold-change
LOC6444656	26.43	2.71	9.74
LOC101928837	20.14	2.26	8.92
THC2666657	14.73	2.85	5.16
DSCR8	57.91	12.01	4.82
ADIRF-AS1	13.11	3.87	3.39
LINC01510	12.80	4.01	3.19
NEAT1	13.31	4.51	2.95
LOC101929243	6.38	2.21	2.89
ENST00000513179	11.78	4.17	2.82
LOC254896	12.06	4.34	2.78

Prediction of lncRNA function based on proteins and miRNA associated with cell-invasion processes

The functions of lncRNAs identified from microarray were predicted based on the functions of the proteins and miRNAs they interact with, which is available in bioinformatic databases. Gene enrichment analysis was performed on all 372 transcripts through Enrichr to determine its biological function based on GO terms. Only one lncRNAs, *SRGAP3*, was annotated with GO terms of interest, which is "regulation of cell migration (GO:0030334)", "negative regulation of cell migration (GO:0030336)", and "negative regulation of cell motility (GO:2000146)". To further characterize transcripts that were not found in databases, predicted lncRNA-protein interactions and lncRNA-miRNA for all lncRNAs transcripts were obtained from the RAID bioinformatic database. Proteins were clustered based on GO terms specified previously, and miRNA was selected based on reported biological function in TAM database. From the 372 upregulated lncRNAs identified in LN18 cells, 35 lncRNAs are associated with angiogenesis, 32 with cell migration, 35 with epithelial-to-mesenchymal transition, 27 with hypoxia and 10 lncRNAs are predicted to be involved in ECM based on protein functions. In addition, 65 lncRNAs are predicted to be involved in angiogenesis, 36 with cell migration, 51 with cell motility, and 70 with EMT based on miRNAs function (Fig. 4A).

Network analysis was performed to allow visualization of large amount of interaction networks which allow visualization of potential lncRNAs based on their predicted protein function obtained through gene enrichment analysis. Eight lncRNAs (*HOTAIR*, *OLMALINC*, *SNHG1*, *FENDRR*, *LOC440028*, *MALAT1*, *UCA1*, and *ZFAS1*) were predicted to interact with all the proteins that were annotated with GO terms specified.

Four specific lncRNA were identified ; i) LINC00221 associated with EMT, ii) LINC01564 and LINC00265 with angiogenesis, and iii) LOC100240735 with cell migration (Fig. 4B). From the network analysis of lncRNAs predicted through miRNA functions, 35 lncRNAs were predicted to be involved to regulate all the specified cellular processes, and four specific lncRNAs in regulating specific cellular process was predicted. LINC00999, ALOX12-AS1, CHKB-AS1, and LINC01588 were predicted to regulate angiogenesis; LINC00482, LINC00239, and LINC01003 were predicted in regulating the EMT process (Fig. 4C). All related lncRNA-miRNA interactions are available in **Supplementary S6**.

Establishing lncRNA-miRNA-protein network through LncTar

To predict the miRNA target of the four lncRNAs identified previously based on protein functions (LINC00221, LINC00265, LINC001564, and LOC100240735 (also known as lnc-SMUG1), LncTAR was utilized to predict the binding potential of miRNAs specific to cell migration and invasion (**Supplementary S7**). Based on summarized LncTAR prediction analysis, 58 miRNAs were predicted to bind with LINC00221, 50 for both LINC001564 and LINC00265, and 23 for LOC100240735. To further establish the lncRNA-miRNA-protein interacting network, mRNAs that will bind to these miRNAs were obtained from the miRNA-protein interaction database. Following this, they were further filtered to include proteins associated with GBM, which was found in the disease database through cluster analysis. These interactions were then analyzed using Cytoscape to visualize distinct signatures (Fig. 4D). Results through LncTAR predictions were able further to visualize the distinct and potential interacting pathways of these four lncRNAs. LINC00221 was predicted to bind with miRNAs, affecting biological pathways at different axis/locuses at *TP53*, *EGF*, *TERT*, and *MDM2*; lncSMUG1 with miR-584 will interact with *BRCA1*, *STAT1*, *PDGFRA*, and *MDM2*; and both LINC01564 and LINC00265 with *PTEN*, *PTG52*, and *EGFR*.

Establishing lncRNA-miRNA-protein/mRNA network using bioinformatics database

A complete lncRNA-miRNA-protein interaction network for all the potential lncRNAs identified in microarray analysis based on protein functions was established. First, a list of miRNAs that interacts with proteins annotated with cell migration and invasion identified previously were obtained from the miRNA-protein interaction database. This was followed by obtaining a list of miRNAs that will interact with these lncRNAs from the lncRNA-miRNA interaction database. These three interaction networks (lncRNA-proteins, miRNA-proteins, and lncRNA-miRNA) were then analyzed using Cytoscape (Fig. 4E) (**Supplementary S8**). Based on the network analysis, several distinct lncRNA signatures were identified from eight lncRNAs (*CRNDE*, *SNHG1*, *NEAT1*, *MALAT1*, *HOTAIR*, *IPP*, *UCA1* and *DLEU1*) which predicted interacting miRNAs and proteins. Three distinct lncRNA signatures from lncRNA, (i) *NEAT1*(*NEAT1*-miR449b-SMAD4 and *NEAT1*-miR124-LOXL4), (ii) *CRNDE*(*CRNDE*-miR145-BTG1), and (iii) *SNHG1*(*SNHG1*-miR101-FBN2) were identified. For lncRNAs predicted based on miRNAs function, mRNA target for these miRNAs was obtained from the miRNA-mRNA interaction database.

Discussions

The roles of lncRNAs in GBM, especially in cell migration and invasion, have not been extensively explored. Recent studies reported that LINC00470 sponges with miR-101, which causes high expression and epigenetic silencing of *ELFN2*, leading to decreased GBM cell autophagy⁵. In another study by the same authors demonstrated that the same lncRNA interacts with the protein FUS which activates the AKT pathway, leading to inhibited GBM cell autophagy⁶. This suggests that lncRNAs may operate simultaneously via different targets and mechanisms within a single biological process that can alter the cell characteristics. Studies have demonstrated that the progression of lower-grade gliomas to GBM can be regulated by the AKT pathway²¹. This has driven the need to validate if lncRNAs could contribute towards GBM invasion by activating cell-migration and invasion related biological pathways through their targets such as proteins and/or miRNAs.

Our microarray data indicates the expression of multiple lncRNAs is significantly upregulated in LN18 cells than in anaplastic astrocytoma cells (SW 1783) compared to normal astrocytes. Among them, LOC644656 and LOC101928837 demonstrated high difference in fold-change between LN18 and SW 1783. However, there is limited information available on LOC644656 and LOC101928837, and we could not find any involvement of both lncRNAs with cell invasion and migration. Future studies could investigate the reasons for their high expression (difference in fold-change) observed since GBM is more invasive than anaplastic astrocytoma. Studying the different upregulated lncRNAs in both cell lines could shed light on the possible biological pathways that are potentially regulated and influence the GBM characteristics. However, most of these lncRNAs are uncharacterised; hence, further bioinformatics analysis is not possible. From the list of lncRNAs identified from microarray, only one lncRNA, SRGAP3, is found in the database to be associated with cell migration and invasion. SRGAP3 is predicted to be associated with several cancers such as lung adenocarcinoma²², pancreatic ductal adenocarcinoma²³, and breast cancer²⁴. However, there are no available literatures on SRGAP3 related to GBM or molecular function that could be found. Hence, no further information could be drawn from functional analysis as most of the lncRNAs are uncharacterised.

At the current stage, most of the functions of the lncRNAs, specifically in GBM, are still undefined at both the cellular and molecular levels. There are very limited studies that have extensively studied any specific lncRNAs that can allow the use of GO-annotated lncRNA database to perform cluster analysis as compared with other types of cancers. Hence, to predict the lncRNA functions, both miRNA and proteins that interact with these lncRNAs were first obtained and subsequently clustered to roughly predict the specific lncRNA's function based on the miRNA or proteins they are reported to interact with, in the processes relevant to GBM migration and invasion. Angiogenesis is a critical event in the progression of GBM since the tumours exhibits a high degree of vascular proliferation and endothelial cell hyperplasia²⁵. Endothelial cells-associated with angiogenesis are among the critical inducers of cell invasion²⁶. On the other hand, hypoxia can trigger the invasive phenotype of GBM by upregulating the levels of invasion proteins that drive the degradation and remodelling of the extracellular matrix and EMT in GBM²⁷.

Four lncRNAs, LINC00221, LINC00265, LINC001564, and LOC100240735, were predicted to be associated with these specific processes. Among them, LINC00221 promotes cisplatin resistance in non-small-cell lung cancer, leading to non-effective chemotherapy²⁸. Additionally, LINC00221 suppresses the malignancy in acute lymphoblastic leukaemia (ALL) where its overexpression leads to an anti-proliferative and pro-apoptotic effect in ALL cells²⁹. Within the context of GBM, the literature regarding these four lncRNAs is very scarce. This raises the need to investigate the role of these lncRNAs in GBM. Among the lncRNAs predicted based on miRNA functions, CHKB-AS1 has been reported as a potential oncogenic lncRNA³⁰ while LINC00482 regulates angiogenesis in bladder cancer through FOXA1³¹. LINC00239 promotes cell migration and invasion through regulation of PI3K/Akt pathway and enhancing KLF12 expression via sponging with mi-484 in acute myeloid leukaemia and colorectal cancer, respectively^{32,33}. LINC1003 predicted to be involved in the EMT process in this study, was previously reported as oncogene and tumour-suppressor in colorectal cancer and multiple myeloma^{34,35}.

CRNDE has been reported to promote malignancy in GBM through miR-384/PIWIL₄/STAT₃ axis³⁶. The analysis performed in our study also shows that CRNDE can act on miR-145-5p, which subsequently regulates tumour suppressor BTG1. Studies demonstrated that PUM2 could promote GBM cell migration through repression of BTG1³⁷. This can be further validated if the lncRNA CRNDE could play any role in this interaction. Our data indicates that NEAT1 is significantly upregulated in GBM cells than in astrocytoma, with a difference in fold-change than normal astrocytes at 2.95 times. In GBM, NEAT1 promotes its progression through the WNT/ β -catenin pathway³⁸, and upregulation of SOX2 expression³⁹. Our analysis predicts that NEAT1 affects SMAD4 and LOXL2, which were both established to be upregulated in GBM^{40,41}. The miRNA miR124 was predicted to interact with NEAT1, and will subsequently target LOXL2, which also has been reported to target SMAD4⁴². SMAD4 is an important mediator for transforming growth factor β (TGF β) signalling, where its abnormal activity often leads to tumourigenesis⁴³. Similarly, the current data predicts NEAT1 to regulate CITED1 and SMAD4, which is part of TGF β -signalling. Our prediction analysis suggests that both lncRNAs and miRNAs interaction is essential as they may play significant roles in regulating cell migration and invasion, where multiple proteins/mRNA involved are targeted simultaneously. Several studies have demonstrated SNHG1 contributes to GBM malignancy through several mechanisms. SNHG1 binds with miRNA-154-5p, affecting FOXP2, leading to enhanced expression of oncogene KDM5B⁴⁴. Another study shows that SNHG1 upregulation promotes glioma progression through sponging with miR-194 that regulates PHLDA1⁴⁵. Recently, SNHG1 is shown to regulate PI3K/AKT pathway and miR140, leading to GBM malignant progression⁴⁶. In this study, we predict that SNHG1 in the regulation of FBN2 expression through miR101. FBN2 overexpression is recently studied in lung cancer tissues, contributing to the proliferation, invasion, and migration abilities of lung cancer cells⁴⁷.

The complete lncRNA-miRNA-protein network were only able to be established from the lncRNAs predicted based on protein function. This is because there are various lncRNA-miRNA interactions and, subsequently, miRNA-mRNA interactions in the bioinformatics database. Establishing the complete

network for the upregulated lncRNAs identified from microarray would be a huge challenge as there are millions of miRNA-mRNA associated interactions with these lncRNA. Thus, to perform network analysis on such a huge dataset would be impossible to visualise for a specific biological axis. In this predictive analysis, there will be loss of potential lncRNA candidates as predictions are based on established interactions from other cancer studies compiled into the bioinformatic databases. Thus, where lncRNA-miRNA-protein/mRNA interactions are concerned, only lncRNA which was shown in both lncRNA-miRNA and lncRNA-protein/mRNA interaction databases, will be shown within the study output. lncRNAs interactions reported in only one of either network will not yield results in the complete network analysis. In this study, our prediction of the lncRNAs in GBM migration and invasion is limited to lncRNA and based on the available experimental data in other types of cancers. Additionally, the prediction depends on the new available literature and the frequency the database is updated. Although our protocol prediction study may identify novel interaction/axis, with the limitation of information, further novel interactions between lncRNAs and their target with no prior studies cannot be predicted.

Conclusions

By utilizing bioinformatics databases, four lncRNA, LINC00221, LINC00265, LINC001564, and LOC100240735 from our microarray data were predicted to promote cell proliferation invasion and migration in GBM based on the function of the proteins they interact. The analysis of lncRNA based on miRNA functions predicted LINC00999, ALOX12-AS1, CHKB-AS1, and LINC01588 as potential regulators GBM cell migration and invasion. Using the prediction model, novel lncRNA-miRNA-mRNA interacting networks involving three lncRNA, NEAT1, CRNDE, and SNHG1, are established in GBM cell invasion. Hence, by utilizing the protocol established in this study, lncRNA functions can be predicted based on the miRNA and protein with which they were previously reported to interact. Considering the empirical validation of this prediction, this protocol is limited to lncRNA, which has been experimentally studied in other types of cancers and therefore, further biological assays involving GBM models are warranted to validate the characterization using this prediction model. Nonetheless, future empirical studies on these lncRNAs as prognostic biomarkers could be beneficial to the future development of therapeutic strategies in treating and managing GBM. With the establishment of various interaction databases in relation to lncRNAs, protocols developed through this study can be utilized in future lncRNA studies to predict the potential regulatory pathways and not only within the context of cancer.

Declarations

Conflict of Interest

The authors declare that there are no competing interests as defined by Nature Research, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Author Contributions

DHLL: Conceptualization, protocol development, data analysis, original draft writing and editing.

SAH: Cell-culture & RNA extraction for microarray.

MNAK: Project administration, review & editing.

SM: Methodology, data analysis of microarray.

TS: Funding acquisition and project administration.

SNP: Funding acquisition and project administration.

MS: Review & editing.

IP: Conceptualization, Project administration, Funding acquisition, Writing – review & editing.

All authors read and approved the final manuscript.

Funding

No funds were required

Acknowledgement

The authors would like to thank Dr. Satoshi Ogawa and Dr. Mageswary Sivalingam from BRIMS, Monash University Malaysia, for their support given throughout the study period. The authors would also like to confirm that the data supporting the findings of this study are available within the article and or in its supplementary materials. The lncRNAs predicted in this study, using bioinformatic analysis, were obtained from a microarray dataset. The complete microarray data is not publicly available because this data forms part of a separate project funded by the National Institutes of Health (NIH)/IMR (grant number NMRR 17-1064-36200) but available upon reasonable request from IP and SNP. DHLL and SAH are PhD candidates supported by scholarships from Monash University Malaysia.

References

1. World Health Organization, W. *World Cancer Report 2014*. (2014).
2. Yuan, C., Ning, Y. & Pan, Y. Emerging roles of HOTAIR in human cancer. *Journal of Cellular Biochemistry* **121**, 3235-3247, doi:<https://doi.org/10.1002/jcb.29591> (2020).
3. Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol Cell* **43**, 904-914, doi:10.1016/j.molcel.2011.08.018 (2011).
4. Baspinar, Y., Elmaci, I., Ozpinar, A. & Altinoz, M. A. Long non-coding RNA MALAT1 as a key target in pathogenesis of glioblastoma. Janus faces or Achilles' heel? *Gene* **739**, 144518,

doi:<https://doi.org/10.1016/j.gene.2020.144518> (2020).

5. Liu, C.*et al.* LINC00470 Coordinates the Epigenetic Regulation of ELFN2 to Distract GBM Cell Autophagy. *Mol Ther* **26**, 2267-2281, doi:10.1016/j.ymthe.2018.06.019 (2018).
6. Liu, C.*et al.* A cytoplasmic long noncoding RNA LINC00470 as a new AKT activator to mediate glioblastoma cell autophagy. *J Hematol Oncol* **11**, 77, doi:10.1186/s13045-018-0619-z (2018).
7. Vollmann-Zwerenz, A., Leidgens, V., Feliciello, G., Klein, C. A. & Hau, P. Tumor Cell Invasion in Glioblastoma. *Int J Mol Sci* **21**, 1932, doi:10.3390/ijms21061932 (2020).
8. da Ponte, K. F.*et al.* In Vivo Relationship Between Hypoxia and Angiogenesis in Human Glioblastoma: A Multimodal Imaging Study. *Journal of Nuclear Medicine* **58**, 1574, doi:10.2967/jnumed.116.188557 (2017).
9. Arcaro, A. & Guerreiro, A. S. The phosphoinositide 3-kinase pathway in human cancer: genetic alterations and therapeutic implications. *Curr Genomics* **8**, 271-306, doi:10.2174/138920207782446160 (2007).
10. Nicolas, S., Abdellatef, S., Haddad, M. A., Fakhoury, I. & El-Sibai, M. Hypoxia and EGF Stimulation Regulate VEGF Expression in Human Glioblastoma Multiforme (GBM) Cells by Differential Regulation of the PI3K/Rho-GTPase and MAPK Pathways. *Cells* **8**, doi:10.3390/cells8111397 (2019).
11. Zhang, B.*et al.* Regulation of branched-chain amino acid metabolism by hypoxia-inducible factor in glioblastoma. *Cell Mol Life Sci* **78**, 195-206, doi:10.1007/s00018-020-03483-1 (2021).
12. Smedley, D.*et al.* The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Research* **43**, W589-W598, doi:10.1093/nar/gkv350 (2015).
13. Volders, P.-J.*et al.* LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Research* **47**, D135-D139, doi:10.1093/nar/gky1031 (2018).
14. Wang, L.*et al.* CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res* **41**, e74, doi:10.1093/nar/gkt006 (2013).
15. Chen, E. Y.*et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128, doi:10.1186/1471-2105-14-128 (2013).
16. Yi, Y.*et al.* RAID v2.0: an updated resource of RNA-associated interactions across organisms. *Nucleic Acids Res* **45**, D115-d118, doi:10.1093/nar/gkw1052 (2017).
17. Li, J.*et al.* TAM 2.0: tool for MicroRNA set analysis. *Nucleic Acids Res* **46**, W180-w185, doi:10.1093/nar/gky509 (2018).
18. Shannon, P.*et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).
19. Haunsberger, S. J., Connolly, N. M. C. & Prehn, J. H. M. miRNAmeConverter: an R/bioconductor package for translating mature miRNA names to different miRBase versions. *Bioinformatics* **33**, 592-593, doi:10.1093/bioinformatics/btw660 (2016).
20. Li, J.*et al.* LncTar: a tool for predicting the RNA targets of long noncoding RNAs. *Brief Bioinform* **16**, 806-812, doi:10.1093/bib/bbu048 (2015).

21. Sonoda, Y.*et al.* Akt Pathway Activation Converts Anaplastic Astrocytoma to Glioblastoma Multiforme in a Human Astrocyte Model of Glioma. *Cancer Research* **61**, 6674-6678 (2001).
22. Jin, D., Song, Y., Chen, Y. & Zhang, P. Identification of Three lncRNAs as Potential Predictive Biomarkers of Lung Adenocarcinoma. *BioMed Research International* **2020**, 7573689, doi:10.1155/2020/7573689 (2020).
23. Song, J.*et al.* Five key lncRNAs considered as prognostic targets for predicting pancreatic ductal adenocarcinoma. *Journal of Cellular Biochemistry* **119**, 4559-4569, doi:<https://doi.org/10.1002/jcb.26598> (2018).
24. Tuersong, T., Li, L., Abulaiti, Z. & Feng, S. Comprehensive analysis of the aberrantly expressed lncRNA-associated ceRNA network in breast cancer. *Mol Med Rep* **19**, 4697-4710, doi:10.3892/mmr.2019.10165 (2019).
25. Onishi, M., Ichikawa, T., Kurozumi, K. & Date, I. Angiogenesis and invasion in glioma. *Brain Tumor Pathol* **28**, 13-24, doi:10.1007/s10014-010-0007-z (2011).
26. Vollmann-Zwerenz, A., Leidgens, V., Feliciello, G., Klein, C. A. & Hau, P. Tumor Cell Invasion in Glioblastoma. *Int J Mol Sci* **21**, doi:10.3390/ijms21061932 (2020).
27. Monteiro, A. R., Hill, R., Pilkington, G. J. & Madureira, P. A. The Role of Hypoxia in Glioblastoma Invasion. *Cells* **6**, 45, doi:10.3390/cells6040045 (2017).
28. Tang, H., Han, X., Li, M., Li, T. & Hao, Y. Linc00221 modulates cisplatin resistance in non-small-cell lung cancer via sponging miR-519a. *Biochimie* **162**, 134-143, doi:<https://doi.org/10.1016/j.biochi.2019.04.019> (2019).
29. Huang, M.*et al.* LINC00221 suppresses the malignancy of children acute lymphoblastic leukemia. *Bioscience Reports* **40**, doi:10.1042/bsr20194070 (2020).
30. Wang, Z.-L.*et al.* Integrative analysis reveals clinical phenotypes and oncogenic potentials of long non-coding RNAs across 15 cancer types. *Oncotarget* **7**, 35044-35055, doi:10.18632/oncotarget.9037 (2016).
31. Wang, Y.*et al.* Silencing LINC00482 inhibits tumor-associated inflammation and angiogenesis through down-regulation of MMP-15 via FOXA1 in bladder cancer. *Aging (Albany NY)* **13**, 2264-2278, doi:10.18632/aging.202247 (2020).
32. Yang, Y., Dai, W., Sun, Y. & Zhao, Z. Long non-coding RNA linc00239 promotes malignant behaviors and chemoresistance against doxorubicin partially via activation of the PI3K/Akt/mTOR pathway in acute myeloid leukaemia cells. *Oncol Rep* **41**, 2311-2320, doi:10.3892/or.2019.6991 (2019).
33. Luo, X., Yue, M., Li, C., Sun, D. & Wang, L. Long Non-Coding RNA LINC00239 Functions as a Competitive Endogenous RNA by Sponging microRNA-484 and Enhancing KLF12 Expression to Promote the Oncogenicity of Colorectal Cancer. *Onco Targets Ther* **13**, 12067-12081, doi:10.2147/OTT.S278582 (2020).
34. Ji, H.*et al.* FOXP1-Induced Long Non-Coding RNA linc01003 Regulating Proliferation and Apoptosis of Colorectal Cancer Cells Through EZH2/HKDC1 and miR-106b-5p/DIC Regulation Axis. *Social Science Research Network*, doi:<http://dx.doi.org/10.2139/ssrn.3421592> (2019).

35. Wu, L.*et al.* Long non-coding RNA LINC01003 suppresses the development of multiple myeloma by targeting miR-33a-5p/PIM1 axis. *Leukemia Research* **106**, 106565, doi:<https://doi.org/10.1016/j.leukres.2021.106565> (2021).
36. Zheng, J.*et al.* CRNDE Promotes Malignant Progression of Glioma by Attenuating miR-384/PIWIL4/STAT3 Axis. *Mol Ther* **24**, 1199-1215, doi:10.1038/mt.2016.71 (2016).
37. Wang, Y.*et al.* PUM2 Promotes Glioblastoma Cell Proliferation and Migration via Repressing BTG1 Expression. *Cell Structure and Function* **44**, 29-39, doi:10.1247/csf.18030 (2019).
38. Chen, Q.*et al.* Long Noncoding RNA NEAT1, Regulated by the EGFR Pathway, Contributes to Glioblastoma Progression Through the WNT/ β -Catenin Pathway by Scaffolding EZH2. *Clin Cancer Res* **24**, 684-695, doi:10.1158/1078-0432.Ccr-17-0605 (2018).
39. Zhou, K.*et al.* Knockdown of long non-coding RNA NEAT1 inhibits glioma cell migration and invasion via modulation of SOX2 targeted by miR-132. *Molecular Cancer* **17**, 105, doi:10.1186/s12943-018-0849-2 (2018).
40. Xiong, Y. & Wang, Q. STC1 regulates glioblastoma migration and invasion via the TGF- β /SMAD4 signaling pathway. *Mol Med Rep* **20**, 3055-3064, doi:10.3892/mmr.2019.10579 (2019).
41. Du, X.-G. & Zhu, M.-J. Clinical relevance of lysyl oxidase-like 2 and functional mechanisms in glioma. *OncoTargets and therapy* **11**, 2699-2708, doi:10.2147/OTT.S164056 (2018).
42. Zhang, Z.*et al.* MicroRNA-124 inhibits the proliferation of C6 glioma cells by targeting Smad4. *Int J Mol Med* **40**, 1226-1234, doi:10.3892/ijmm.2017.3088 (2017).
43. Zhao, M., Mishra, L. & Deng, C.-X. The role of TGF- β /SMAD4 signaling in cancer. *Int J Biol Sci* **14**, 111-123, doi:10.7150/ijbs.23230 (2018).
44. Li, H.*et al.* SNHG1 promotes malignant biological behaviors of glioma cells via microRNA-154-5p/miR-376b-3p-FOXP2-KDM5B participating positive feedback loop. *Journal of Experimental & Clinical Cancer Research* **38**, 59, doi:10.1186/s13046-019-1063-9 (2019).
45. Liu, L.*et al.* The long non-coding RNA SNHG1 promotes glioma progression by competitively binding to miR-194 to regulate PHLDA1 expression. *Cell Death & Disease* **10**, 463, doi:10.1038/s41419-019-1698-7 (2019).
46. Cai, R.-D.*et al.* SNHG1 Promotes Malignant Progression of Glioma by Targeting miR-140-5p and Regulating PI3K/AKT Pathway. *Cancer Manag Res* **12**, 12011-12020, doi:10.2147/CMAR.S269572 (2020).
47. Hong, Q., Li, R., Zhang, Y. & Gu, K. Fibrillin 2 gene knockdown inhibits invasion and migration of lung cancer cells. *Cell Mol Biol (Noisy-le-grand)* **66**, 190-196 (2020).

Figures

Absence of lncRNA

Absence of lncRNA

lncRNA regulation

Figure 1

A typical model depicting lncRNA sponging with miRNA, allowing translation of mRNAs, which will regulate downstream processes which may further promote oncogenic phenotype .

Figure 2

Overall workflow for the prediction of potential lncRNA candidates in the regulation of cellular processes leading to cell migration and invasion in GBM.

A **B**

Figure 3

Upregulated lncRNAs identified from LN18 (A) and SW 1783 (B) cell lines compared against NHA cells

Figure 4

(A) Bar-chart for the distribution of lncRNA in respective of biological processes based on the miRNA and protein which they interact with. (B) Network analysis of lncRNAs that had been clustered based on selected GO terms associated with cell migration and invasion. (C) Network analysis of lncRNA based on the functions of the miRNA in which they were reported to interact with, lncRNA highlighted in purple was predicted with involvement in all 4 processes, while those highlighted in yellow was predicted to only

involve in one specific cellular process associate with cell-invasion. (D) Network analysis of lncRNA-miRNA-protein interactions through lncTAR prediction. (E) Network analysis of lncRNA-miRNA-protein interactions for the 8 predicted lncRNA(UCA1, MALAT1, SNHG1, DLEU1, IPP, CRNDE, HOTAIR, and NEAT1). There is distinct interaction network for SNHG1, CRNDE, NEAT, and UCA1. These distinct signatures are associated with miRNA and proteins/mRNA which is involved in regulation of cell-invasion.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.pdf](#)
- [SupplementaryS6.xlsx](#)
- [SupplementaryS7.xlsx](#)
- [SupplementaryS8.xlsx](#)