Differential long non-coding RNA and mRNA expression in differentiated human glioblastoma stem cells

HAO LI1,2, HAOWEN LI3, YAJING HAO4, YUMING JIAO1,2, ZHICEN LI1,2, HAIYAN YUE1, ZHE XU3, SHUO WANG1,2, YONG CAO1,2 and JIZONG ZHAO1,2

1Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University; 2China National Clinical Research Center for Neurological Diseases; 3Laboratory of Clinical Medicine Research, Beijing Tiantan Hospital, Capital Medical University, Beijing 100050; 4Laboratory of Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Beijing 100101, P.R. China

Received August 23, 2015; Accepted May 4, 2016

DOI: 10.3892/mmr.2016.5505

Abstract. Differentiation of glioblastoma stem cells (GSCs) may lead to inhibition of their self-renewing ability and tumorigenic potential, as well as increasing their sensitivity to treatment. The critical role of long non-coding RNAs (lncRNAs) in numerous biological processes has been revealed. However, the involvement of lncRNAs in GSC differentiation remains to be elucidated. In the present study, GSCs were isolated from patient samples and differentiation was induced. Using a high-throughput microarray, the present study identified a profile of 1,545 lncRNAs and 2,729 mRNAs that differed between GSCs and their non-differentiated counterparts. To ascertain the association between the altered lncRNAs and mRNAs, a co-expression network was constructed in which 1,087 lncRNAs and 1,928 mRNAs altered upon GSC differentiation formed a total of 19,642 lncRNA-mRNA pairs. Based on the co-expression network, the lncRNA functions were additionally predicted by a cis- or trans- targeting program. Furthermore, three pairs of lncRNAs and their nearby target mRNAs were selected [ENSG00000261924.1-regulatory associated protein of MTOR complex 1, ENSG00000235427.1-caveolin 1 and Tax1 binding protein 3 (TAX1BP3)-purinergic receptor P2X 5 (P2RX5)-TAX1BP3] and their expression levels were validated by reverse transcription-quantitative polymerase chain reaction. The altered lncRNAs were also regulated by various pluripotency transcription factors (POU domain, class 3, transcription factor, sex determining region Y-box 2, spalt-like transcription factor 2 and oligodendrocyte lineage transcription factor 2). In conclusion, the results of the present study revealed that lncRNAs may function in GSC differentiation by regulating their target mRNAs, and a set of lncRNAs were identified as candidates for further study concerning the future treatment of GSCs.

Introduction

Glioblastoma multiforme (GBM), which accounts for ~40% of all primary brain tumors, is one of the most malignant tumors in humans, with patients having a median overall survival time of 12-15 months (1). Previous studies have demonstrated that tumor cell populations in GBM are heterogeneous in terms of morphology and differentiation status (2). A highly tumorigenic and self-renewing subpopulation of cells, which display stem-like behavior, is the glioblastoma stem cells (GSCs) (2,3). Notably, the property of ‘stemness’ in GSCs is believed to be key for tumor formation, differentiation, proliferation and resistance to chemo- and radiotherapy, possibly explaining the high frequency of treatment failure and tumor relapse observed in glioblastoma (4,5). The differentiation of GSCs may lead to the inhibition of their self-renewing ability and tumorigenic potential, as well as increasing their sensitivity to treatment. The promotion of differentiation is considered to be a potential strategy to eradicate GSCs, and therefore it is important to understand the detailed molecular mechanisms involved in GSC differentiation (6,7).

Advances in next-generation deep sequencing technologies have identified a large number of non-coding RNAs (ncRNAs) termed long non-coding RNAs (lncRNAs). These molecules are >200 nucleotides in length and are a class of single-stranded RNAs, which lack protein-coding ability (8,9). Rather than being irrelevant transcriptional noise, studies have revealed that lncRNAs possess critical regulatory roles in numerous biological processes, including immune responses, cellular metabolism, stem cell differentiation and tumorigenesis (10,11). Furthermore, certain studies have demonstrated that lncRNA signatures correlate with glioma malignancy grade, histological differentiation and prognosis (12,13). The diverse mechanisms underlying the regulatory roles of lncRNAs include interactions with mRNAs, functioning as
miRNA sponges and acting as tethers, guides, decoys and scaffolds for proteins (14,15).

Numerous studies of GSC differentiation have revealed certain underlying molecular mechanisms involved in this process. Certain transcription factors (TFs), protein-coding genes and ncRNAs have been proposed to be involved in the maintenance of GSC stemness, and the regulation of these factors may facilitate GSC differentiation (16-18). However, the role of IncRNAs in the GSC differentiation process remains to be elucidated.

In the present study, using patient-derived GSC lines and their differentiation-induced GSC (DGSC) counterparts, a high-throughput microarray analysis was performed to identify the profile changes of IncRNAs and mRNAs. Using bioinformatics analysis, aberrant IncRNA and mRNA expression was integrated and potential regulatory roles of IncRNAs in GSC differentiation were proposed.

Materials and methods

GSC culture and differentiation. GSCs (G0, G1, G2, G3, G4 and G5) were isolated from neurosurgical samples from six patients with GBM who were hospitalized in the Department of Neurosurgery of Beijing Tiantan Hospital (Beijing, China). Informed consent was obtained from the patients and approval was given from the Ethics Committee of Beijing Tiantan Hospital (KY2014-021-02; Beijing, China). Tumor tissue was sectioned into 1-mm³ fragments using scissors and washed with Dulbecco's modified Eagle's medium/nutrient mixture F12 (DMEM/F12; Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA) three times. The fragments were then digested with 0.02% trypsin for 20 min and dissociated into single cells. The cells were resuspended and maintained in GSC-propagating medium composed of 2 mM GlutaMAX (Invitrogen; Thermo Fisher Scientific, Inc.), 20 ng/ml recombinant human epidermal growth factor (R&D Systems, Inc., Minneapolis, MN, USA), 20 ng/ml basic fibroblast growth factor (Invitrogen; Thermo Fisher Scientific, Inc.), N2 supplement (Invitrogen; Thermo Fisher Scientific, Inc.) and B27 supplement (Invitrogen; Thermo Fisher Scientific, Inc.) in DMEM/F12. The above procedure was completed within 1 h of the surgical removal of tissue. GSCs (2x10⁴/mouse; 6 mice) were subcutaneously injected into the left hind flank of 6-week-old female non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice (VitalStar Biotechnology Co., Ltd., Beijing, China). The mice were sacrificed by decapitation 40 days later, and tumors were harvested and stained with hematoxylin and eosin. The animal use was approved from the Ethics Committee of Beijing Tiantan Hospital (KY2014-021-02; Beijing, China).

Immunofluorescence. GSCs grown as neurospheres were collected by centrifugation at 150 g for three min at room temperature, then the cells were transferred to microscope slides, while the DGSCs were grown and stained in 24-well plates. The cells were fixed with 4% formaldehyde, permeabilized with 0.1% Triton X-100 for 10 min and then blocked with 10% normal goat serum (Jackson Immunoresearch Laboratories, Inc., West Grove, PA, USA) for 5 min. The cells were then incubated at 4°C overnight with the following primary antibodies: Mouse anti-cluster of differentiation (CD) 133 (dilution 1:500; catalog no., MAB4399; EMD Millipore, Billerica, MA, USA); rabbit anti-nestin (dilution 1:250; catalog no., ab82375; Abcam, Cambridge, UK); rabbit anti-glial fibrillary acidic protein (GFAP; dilution 1:500; catalog no., ab33922; Abcam); rabbit anti-βIII tubulin (TuJ1; dilution 1:500; catalog no., ab18207; Abcam); mouse anti-04 (dilution 1:200; catalog no., MAB345; EMD Millipore). Following washing with phosphate-buffered saline 3 times, the cells were incubated with fluorescein isothiocyanate (FITC)-conjugated anti-mouse (dilution 1:1,000, catalog no., F-11,021; Thermo Fisher Scientific, Inc.) and anti-rabbit (dilution 1:2,000; catalog no., ab66101; Thermo Fisher Scientific, Inc.) secondary antibodies and Texas Red (TR) -conjugated anti-mouse (dilution 1:2,000, catalog no., T-6390; Thermo Fisher Scientific, Inc.) and anti-rabbit (dilution 1:1,000, catalog no., T-2767; Thermo Fisher Scientific, Inc.) secondary antibodies. The cell nuclei were counterstained with DAPI (Sigma-Aldrich, St. Louis, MO, USA). Images were captured using a DMI 4,000 Leica fluorescent microscope and processed using the Leica Application Suite version 4.2 Imaging System software (Leica Microsystems GmbH, Wetzlar, Germany).

RNA extraction and microarray analysis. Total RNA was extracted from GSCs and DGSCs using TRizol® reagent (Sigma-Aldrich) according to the manufacturer's instructions. RNA purity and concentration was assessed using a NanoDrop ND-1,000 spectrophotometer (Thermo Fisher Scientific, Inc.). RNA integrity was determined by formaldehyde denaturing gel electrophoresis. Microarray hybridization was performed by CapitalBio Corporation (Beijing, China). Briefly, each RNA sample was amplified and transcribed into double-stranded complementary DNA (cDNA). The labeled cDNA was then hybridized to the IncRNA-mRNA Human Gene Expression Microarray version 4.0, 4x180K chip. Data normalization, quality control and the calculation of differences in gene expression were performed with the GeneSpring GX version 11.5.1 software package (Agilent Technologies, Inc., Santa Clara, CA, USA). IncRNAs and mRNAs were defined as differentially expressed (DE) if the fold-change values were >2.0, or if P<0.05. The microarray data was uploaded to the Gene Expression Omnibus (GEO) database (GSE68343; www.ncbi.nlm.nih.gov/geo/). In addition, hierarchical clustering with average linkage was used for calculating the distinguishable IncRNA and mRNA expression patterns with Cluster version 3.0 software (bonsai.hgc.jp/~mdehoon/software/cluster/software.htm#ctv).
Co-expression network construction and lncRNA target prediction. The lncRNA-mRNA co-expression network was constructed based on associations between DE lncRNAs and DE mRNAs during GSC differentiation. Using the open source bioinformatics software Cytoscape (Institute of Systems Biology, Seattle, WA, USA), a network of the selected lncRNA-mRNA pairs was drawn (Pearson correlation coefficient >0.99 or < -0.99; and P<0.05). Subsequently, the lncRNA-mRNA co-expression network was analyzed using a cis- or trans-regulatory prediction program. A target gene was defined as cis if the mRNA transcribed was within a 10 kilobase (kb) window upstream or downstream of the lncRNA genomic location, using gene annotations from the University of California, Santa Cruz (Santa Cruz, CA, USA; genome.ucsc.edu/). As lncRNAs may act as competing endogenous RNAs, which regulate the mRNA transcripts by competing for shared microRNAs, the trans- predictions were primarily executed by searching the pairs of lncRNAs and mRNAs that had similar sequences in the 3' untranslated region using the BLAST-Like Alignment Tool (default parameter settings) software (genome.ucsc.edu/cgi-bin/hgBlat).

Reverse transcription-polymerase chain reaction (RT-qPCR) validation. Total RNA was reverse transcribed to cDNA using the PrimeScript RT Reagent kit (Takara Bio, Inc., Otsu, Japan). Using a 7500 Real Time PCR System (Applied Biosystems; Thermo Fisher Scientific, Inc.) with SYBR® Premix Ex Taq™ II (Tli RNaseH Plus; Takara Bio, Inc.), RT-qPCR was performed according to the manufacturer's instructions. The specific primers for each gene are presented in Table 1. All experiments were performed in triplicate. Gene expression was normalized to GAPDH in each sample.

Analysis of TF binding sites and lncRNA conservation. Genomic annotations of human lncRNA genes and transcripts were downloaded from the GENCODE website (www.gencodegenes.org/). Peak lists of chromatin immunoprecipitation sequencing (ChIP-Seq) datasets performed on 4 TFs [POU domain, class 3, transcription factor (POU3F), sex determining region Y-box 2 (SOX2), spalt-like transcription factor 2 (SALL2) and oligodendrocyte lineage transcription factor (OLIG) 2] were available from the deep sequencing of ChIP-Seq libraries in GSCs (21). Subsequently, the peaks of the TFs were overlapped with the regulatory regions of the lncRNA genes (5 kb upstream from the transcription start site) and the gene body of the lncRNA using the IntersectBed tool from Bedtools (bedtools.readthedocs.io/en/latest/). If the regulatory regions or the gene body of an lncRNA gene had at least one TF peak, the lncRNA gene was determined as having the potential to be regulated by the TF. Conservation analysis was conducted using phastCons data (compgen.cshl.edu/phast/) as described previously (22). If the average phastCons score was >0.1, the lncRNA was defined as conserved.

Statistical analysis. Statistical analyses were performed using SPSS software, version 13.0 (SPSS, Inc., Chicago, IL, USA). Values are presented as the mean ± standard deviation, and Student's t-test was used for comparing sample sets. P<0.05 was considered to indicate a statistically significant difference.

Results

Characterization of GSCs isolated from surgical GBM samples and determination of GSC differentiation. GSCs were cultured in GSC-propagating media, where they formed characteristic renewable neurospheres and were able to proliferate indefinitely (Fig. 1Aa). Immunofluorescence revealed that GSCs exhibited an increased proportion of cells positive for CD133 (Fig. 1Ab), and increased expression of nestin, a neural progenitor cell marker (Fig. 1Ac), compared with DGSCs (Fig. 1Ad and e). The tumorigenesis of GSCs was confirmed by tumor formation in NOD/SCID mice following inoculation with 2x10⁵ GSCs (data not shown).

The differentiation of GSCs was induced in DMEM/F12 medium containing 10% FBS for 4 days. Following serum exposure, the GSCs acquired glial- and neurite-like cell features, with protrusions and adherence to the flask, observed using an optical microscope (Fig. 1Af). In addition, immunofluorescent staining analysis confirmed astrocytic- and neural-cell differentiation, indicated by the positive staining with anti-GFAP and anti-Tuj1 in the majority of DGSCs (Fig. 1Ag and h), but not in GSCs (Fig. 1i and j). Furthermore, it was observed that a small proportion of the cells differentiated into the oligodendrocytic lineage, as demonstrated by positive staining for O4 sulfatides in DGSCs (Fig. 1Ak), but not in GSCs (Fig. 1Al). The results of the present study are consistent with a previous study (16), and indicate that GSCs may be efficiently induced into astrocytic and neural lineages, and partly induced into oligodendrocytic lineages. RT-qPCR detected similar trends in CD133 (P=0.02), nestin (P=0.03), GFAP (P<0.001) and Tuj1 (P<0.001) expression in the GSC differentiation process. In addition, SOX2 (P<0.001), which functions in the maintenance of GSC stemness, was markedly decreased in DGSCs compared with the expression observed in GSCs (Fig. 1B).

lncRNAs and mRNAs are differentially expressed during GSC differentiation. Microarray analysis was performed to investigate the expression of IncRNAs and mRNAs during GSC differentiation. A total of 1,545 lncRNAs were differentially expressed greater than 2-fold (P<0.05) during the GSC differentiation process, of which 650 lncRNAs were upregulated and 895 were downregulated. In addition, 2,729 mRNAs were differentially expressed greater than 2-fold (P<0.05) between GSCs and DGSCs, of which 2,179 were upregulated and 550 were downregulated. Subsequently, Cluster 3.0 software was used for the clustering of the lncRNA and mRNA expression data. As presented in Fig. 2A, the GSCs exhibited similar expression patterns of the IncRNAs and mRNAs, but were distinct from the DGSCs. The results of the present study indicated that the expression patterns of lncRNAs and mRNAs may be associated with stemness, and that lncRNAs and mRNAs may have potential roles in regulating GSC differentiation.

To aid the interpretation of the functionality of the lncRNAs, the DE lncRNAs were classified into five categories: Antisense, intergenic, intronic, divergent and others, based on their locations relative to nearby protein-coding genes (23). The anatomy of the lncRNA loci implied that the lncRNAs have potential functions in regulating their neighboring
protein-coding genes. According to this categorization method, nearly half of the lncRNAs (593/1,545) (Fig. 2B) belonged to the intergenic lncRNA category, of which a large proportion may have important functions due to their clear conservation across mammalian species and their association with chromatin-modifying complexes and gene expression (24).

**lncRNAs are significantly correlated with mRNAs.** While the differential expression of lncRNAs and protein-coding mRNAs indicated that this subset of lncRNAs may be associated with the local transcriptional activity of protein-coding mRNAs, a lncRNA-mRNA co-expression network was constructed to ascertain the correlation between DE lncRNAs and DE mRNAs in GSC differentiation. A total of 19,642 lncRNA-mRNA pairs, composed of 1,087 DE lncRNAs and 1,928 DE mRNAs, were identified. In the network, single lncRNAs were associated with multiple (between 1 and 10) mRNAs and vice versa (data not shown). The lncRNA-mRNA network was large and complex; therefore, to present the association between lncRNAs and mRNAs more clearly, DE lncRNAs and DE mRNAs with fold-change values >3.0 and P<0.05 were selected to generate a ‘core’ sub-network map (data not shown). A total of 30 lncRNA-mRNA matched pairs were fitted with cis-regulatory effects and 13 lncRNA-mRNA matched pairs were fitted with trans-regulatory effects (Fig. 4). Among the matched pairs, lncRNA purinergic receptor P2X 5 (P2RX5)-Tax1 binding protein 3 (TAX1BP3) had cis-or trans-regulatory effects. From the core sub-network, a schematic of eight RNA hub nodes (RNA nodes were designated as hub nodes if they were associated with multiple RNAs, implying their important functional roles) and their associated RNAs were drawn (Fig. 3A). Hub nodes were selected based on the degree distribution graph (Fig. 3B; degree was used to depict the number of RNAs associated with one RNA node, and one hub node was defined when its degree was ≥10). Among them, C20orf27 was a protein-coding mRNA, and the remaining 7 hub nodes (uc.359-, XLOC_006350, ENSG00000232956.3, ENSG00000172965.7, ENSG00000242808.2, ENSG00000172965.7 and XLOC_003079) were lncRNAs, suggesting that lncRNAs were more likely to be executing functions via the regulation of other RNAs due to their multiple connections with other RNAs. These lncRNA hub nodes were associated with protein-coding mRNAs, including leukemia inhibitory factor, OLIG1 and fibroblast growth factor 2, which have key roles in the regulation of biological processes associated with stem cells. Together, these results suggest that DE lncRNAs may potentially be involved in GSC differentiation via associations with protein-coding mRNAs.

**lncRNA functional prediction via cis- and trans-target prediction.** To improve the prediction accuracy of the functions of the lncRNAs, cis- and trans-target prediction programs were utilized based on the lncRNA-mRNA co-expression network (containing 19,642 lncRNA-mRNA pairs, as described previously) obtained. A total of 30 lncRNA-mRNA matched pairs were fitted with cis-regulatory effects and 13 lncRNA-mRNA matched pairs were fitted with trans-regulatory effects (Fig. 4). Among the matched pairs, lncRNA purinergic receptor P2X 5 (P2RX5)-Tax1 binding protein 3 (TAX1BP3) had 3 cis- or trans-genes (TAX1BP3, derlin 2 and proteasome 26S subunit, non-ATPase 12), lncRNA ENSG00000261924.1 had 2 cis-or trans-genes [kinesin family member 1C and regulatory-associated protein of MTOR complex 1 (RPTOR)], and the remaining lncRNAs had 1 cis- or trans-gene each. The functional roles of the target protein-coding mRNAs, including RPTOR, in stem cell biology or glioma malignancy (25,26) raised the possibility that lncRNAs regulate GSC differentiation through regulation of their target mRNAs. Thus, the results identified the most plausible functional lncRNAs and their target mRNAs through a series of bioinformatics filter strategies.

**Validation of the lncRNAs and their predicted protein-coding mRNA genes.** To evaluate the consistency of the microarray and confirmexpression of lncRNAs and their target mRNAs, RT-qPCR was performed in samples G0-G5 and their differentiated counterparts. A total of three pairs [RPTOR-ENSG00000261924.1, TAX1BP3, Tax1 binding protein 3; CAV1, caveolin 1; RPTOR, regulatory associated protein of MTOR complex 1; P2RX5, purinergic receptor P2X 5; CD133, cluster of differentiation 133; Tuji, βIII tubulin; GFAP, glial fibrillary acidic protein; SOX2, sex determining region Y-box 2.]

### Table I. Reverse transcription-quantitative polymerase chain reaction primers.

<table>
<thead>
<tr>
<th>Name</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSG00000235427.1</td>
<td>5'-AAAACCACCTGAGACACGGAGGACG-3'</td>
<td>5'-CCAGGGACAGGCAGCATCA-3'</td>
</tr>
<tr>
<td>TAX1BP3</td>
<td>5'-CCTCACATCCGGGGCCAGC-3'</td>
<td>5'-CACCTCTAATGCGTGAACATC-3'</td>
</tr>
<tr>
<td>CAV1</td>
<td>5'-CGACCCTAAACACCTCAAGCA-3'</td>
<td>5'-GGCAGACAGCAAGCCTGATAA-3'</td>
</tr>
<tr>
<td>RPTOR</td>
<td>5'-GTGTTGAGCTGGGAGCCAGGAGA-3'</td>
<td>5'-TCAGGGCCATATGGGACAGCA-3'</td>
</tr>
<tr>
<td>ENSG00000261924.1</td>
<td>5'-AGGAATATGACATACACGAGGGA-3'</td>
<td>5'-TGTTGTAAGGAGGAGAAGCAGAGG-3'</td>
</tr>
<tr>
<td>P2RX5-TAX1BP3</td>
<td>5'-CGAGGGCAAGCGTGGGA-3'</td>
<td>5'-GCGCCCAATAGCGTGAACAGT-3'</td>
</tr>
<tr>
<td>CD133</td>
<td>5'-TACCAAGGAACAGCGGTGAGA-3'</td>
<td>5'-GCTGAAGCTTCTCAAGGTCGGT-3'</td>
</tr>
<tr>
<td>Nestin</td>
<td>5'-CGTTGGGAACAGAGGTTGGAG-3'</td>
<td>5'-TAAGGAAAGGGCCTGACAGGT-3'</td>
</tr>
<tr>
<td>Tuji</td>
<td>5'-GTACGAAGAGACGAGCAGGAGG-3'</td>
<td>5'-GCTCAGGGAGCTGCAATAAGAC-3'</td>
</tr>
<tr>
<td>GFAP</td>
<td>5'-GTCTCATGTTGAGCTTGACG-3'</td>
<td>5'-GCAGGTCAGAGGACTGCAACT-3'</td>
</tr>
<tr>
<td>SOX2</td>
<td>5'-ATGCACAATCGGAGATCG-3'</td>
<td>5'-TATATCCGGGTGCTCTTCT-3'</td>
</tr>
<tr>
<td>GAPDH</td>
<td>5'-CTATAAATGGACCCGAGCC-3'</td>
<td>5'-GCCGCCCAATAGCGAACAAT-3C'</td>
</tr>
</tbody>
</table>

TAX1BP3, Tax1 binding protein 3; CAV1, caveolin 1; RPTOR, regulatory associated protein of MTOR complex 1; P2RX5, purinergic receptor P2X 5; CD133, cluster of differentiation 133; Tuji, βIII tubulin; GFAP, glial fibrillary acidic protein; SOX2, sex determining region Y-box 2.
caveolin 1 (CAV1)-ENSG00000235427.1 and TAX1BP3-P2RX5-TAX1BP3] were selected for investigation, according to fold-change values and reported gene functions. Whilst the fold-change values of the lncRNAs and mRNAs varied between samples, the expression trends were similar to the microarray results. These data, presented in Fig. 5, verify the consistency of the microarray and confirm the expression of lncRNAs and their target mRNAs despite the tumor-specific genetic background. The P-values for the expression level of DGSCs compared to that of GSCs for each transcript in each sample were as follows: RPTOR G0, P=0.008; G1, P=0.012; G2, P=0.04; G3, P=0.04; G4, P=0.02; G5, P<0.001; CAV1 G0, P<0.001; G1, P<0.001; G2, P=0.001; G3, P<0.001; G4, P<0.001; G5, P<0.001; ENSG00000235427.1 G0, P<0.001; G1, P<0.001; G2, P<0.001; G3, P<0.001; G4, P<0.001; G5, P<0.001; CA V1 G0, P<0.001; G1, P<0.001; G2, P<0.001; G3, P<0.001; G4, P<0.001; G5, P<0.001; ENSG00000261924.1 G0, P=0.008; G1, P=0.012; G2, P=0.04; G3, P=0.04; G4, P=0.02; G5, P<0.001; SOX2, P<0.001; ESG00000261924.1 G0, P=0.008; G1, P=0.012; G2, P=0.04; G3, P=0.04; G4, P=0.02; G5, P<0.001; CAV1 G0, P<0.001; G1, P<0.001; G2, P=0.001; G3, P<0.001; G4, P<0.001; G5, P<0.001; ENSG00000235427.1 G0, P<0.001; G1, P<0.001; G2, P<0.001; G3, P=0.003; G4, P<0.001; G5, P=0.08; TAX1BP3 G0, P<0.001; G1, P=0.008; G2, P=0.01; G3, P<0.001; G4, P<0.001; G5, P<0.001; P2RX5-TAX1BP3 G0, P<0.001; G1, P<0.001; G2, P=0.004; G3, P=0.02; G4, P=0.01; G5, P<0.001.

Identification of DE lncRNAs regulated by a set of core TFs. A set of core TFs (POU3F, SOX2, SALL2 and OLIG2) have previously been demonstrated to bind to, and activate, GSC-specific
Figure 2. lncRNA and mRNA profile signatures in GSCs and DGSCs. (A) Cluster heat maps of lncRNA and mRNA for GSCs and DGSCs. (B) Subgroups of differentially expressed lncRNAs according to genomic location and associations with nearby protein-coding genes. lncRNA, long non-coding RNA; GSC, glioblastoma stem cell; DGSC, differentiated glioblastoma stem cell.

Figure 3. RNA hub nodes co-expressed with multiple RNAs. (A) Schematic of the eight RNA hub nodes, which are co-expressed with multiple RNAs. Pink and green hexagons indicate upregulated and downregulated PC mRNAs, respectively. Pink and green rhombuses indicate upregulated and downregulated lncRNAs, respectively. Pink and blue lines indicate the positive and negative co-expression association between the two RNAs, respectively. (B) The degree distribution graph depicting the number of RNA nodes (y-axis) and the number of RNAs associated with each (x-axis). lncRNA, long non-coding RNA; PC, protein-coding.
regulatory elements, and thus determine and reprogram the differentiation status of the GSCs (21). In the present study, microarray and RT-qPCR revealed that the expression of all four TFs decreased upon differentiation of GSCs to DGSCs, which was consistent with a previous study (21). A total of 262 lncRNAs were revealed to contain at least one TF binding site in their regulatory regions or gene body, and 21 lncRNAs contained binding sites for all four TFs. The results are summarized as a Venn diagram in Fig. 6A. In addition, a representative schematic of the binding of the four TFs in the ENSG00000242808 regions is presented in Fig. 6B. Furthermore, as evidence has suggested that conservation often indicates functionality (27), the sequence conservation of the 21 lncRNAs was then evaluated with the phastCons score, with 10 lncRNAs confirmed to be conserved among vertebrates (data not shown). Taken together, the present results suggest that the DE lncRNAs are regulated by a set of core TFs, and that they may be integrated into known pluripotency TF networks regulating GSC differentiation.

Discussion

There is a growing understanding of the importance of lncRNAs in the regulation of pluripotency and differentiation in stem cells, including cancer stem cells (14,28). In addition,
certain classical lncRNAs, including HOX transcript antisense RNA (HOTAIR), metastasis associated lung adenocarcinoma transcript 1 (MALAT-1) and long intergenic non-protein coding RNA, regulator of reprogramming (Linc-RoR), whose roles in numerous biological processes have been widely investigated, have been reported to possess regulatory roles in stemness maintenance and differentiation in various cancer stem cells (29-31). However, no significant alterations in HOTAIR, MALAT-1 and Linc-RoR expression were observed in the present study, which may be partly explained by the differences between cell lines of different types of cancer. Based on transcriptome microarray analysis and deep sequencing technologies, a number of novel lncRNAs were observed to regulate stem cell stemness and differentiation in the present study, notably LncTCF7 and pnky. Previously, LncTCF7 was reported to maintain the stemness of human liver cancer stem cells via the activation of Wnt signaling (32). Pnky, a conserved lncRNA, has been demonstrated to interact with and regulate the neuronal differentiation of embryonic and postnatal neural stem cells (33).

Despite a number of lncRNAs being examined in a variety of cancer stem cells, few studies have reported lncRNA profile and their roles in GSC differentiation. In the current study, the use of an lncRNA-mRNA human gene expression microarray platform identified for the first time a total of 1,545 lncRNA profile changes during the differentiation of patient-derived GSCs. Using a series of bioinformatics strategies for lncRNA prediction, it was observed that DE lncRNAs may interact with protein-coding mRNAs by associating with them and in addition through cis- or trans-targeting. Certain protein-coding mRNAs that are regulated by lncRNAs possess significant biological functions in cancer or stem...
of the present study not only reveal that a large proportion of DE lncRNAs are regulated by TFs, but in addition provide an alternative method to interpret the functions of the core TFs that bind to and activate DE lncRNAs and thus direct the GSC differentiation process.

In conclusion, the present study identified a lncRNA and mRNA expression profile that clearly distinguished GSCs from DGSCs. Using a series of bioinformatics strategies for lncRNA prediction, it was observed that DE lncRNAs were able to regulate the GSC differentiation process through an association with protein-coding mRNAs and additionally through cis- or trans-targeting. Furthermore, the DE lncRNAs were regulated by a set of core pluripotency TFs. The results of the present study provide a potential direction for future research via the comprehensive integration of lncRNAs and mRNAs, and the identification of a set of lncRNAs as candidates for further study.

Acknowledgements

The authors would like to thank Dr Anchen Guo at the Laboratory of Clinical Medicine Research, Beijing Tiantan Hospital for providing the GSC-propagating medium. The present study was supported by a grant from the China National Clinical Research Center for Neurological Diseases, the Training Plan for Beijing High-Level Healthcare Personnel (grant no. 2011-3-28).

References


