HOTAIR, a cell cycle–associated long noncoding RNA and a strong predictor of survival, is preferentially expressed in classical and mesenchymal glioma

Jun-Xia Zhang†, Lei Han†, Zhao-Shi Bao†, Ying-Yi Wang†, Lu-Yue Chen, Wei Yan, Shi-Zhu Yu, Pei-Yu Pu, Ning Liu, Yong-Ping You, Tao Jiang, and Chun-Sheng Kang, for the Chinese Glioma Cooperative Group

Department of Neurosurgery, Tianjin Medical University General Hospital, Laboratory of Neuro-Oncology, Tianjin Neurological Institute, Key Laboratory of Post-trauma Neuro-repair and Regeneration in Central Nervous System, Ministry of Education, Tianjin Key Laboratory of Injuries, Variations and Regeneration of Nervous System, Tianjin, China (J.-X.Z., L.H., L.-Y.C., P.-Y.P., C.-S.K.); Department of Neurosurgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China (J.-X.Z., Y.-Y.W., N.L., Y.-P.Y.); Beijing Neurosurgical Institute, Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China (Z.-S.B., W.Y., T.J.); Department of Neuropathology, Tianjin Neurological Institute, Tianjin Medical University General Hospital, Tianjin, China (S.-Z.Y.)

Background. Long noncoding RNA Hox transcript antisense intergenic RNA (HOTAIR) has been characterized as a negative prognostic factor in breast and colon cancer patients. The clinical significance and function of HOTAIR in glioma remains unclear.

Methods. We analyzed the clinical significance of HOTAIR in 3 different glioma cohorts with gene expression data, including correlation with tumor grade, prognosis, and molecular subtype. The function of HOTAIR in glioma was explored by performing gene set enrichment analysis and in vitro and in vivo experiments.

Results. HOTAIR expression was closely associated with glioma grade and poor prognosis. Multivariate Cox regression analysis revealed that HOTAIR was an independent prognostic factor in glioblastoma multiforme patients. HOTAIR expression correlated with glioma molecular subtype, including those of The Cancer Genome Atlas. HOTAIR was preferentially expressed in the classical and mesenchymal subtypes compared with the neural and proneural subtypes. A gene set enrichment analysis designed to show gene set differences between patients with high and low HOTAIR expression indicated that HOTAIR expression was associated with gene sets involved in cell cycle progression. HOTAIR reduction induced colony formation suppression, cell cycle G0/G1 arrest, and orthotopic tumor growth inhibition.

Conclusion. Our data establish that HOTAIR is an important long noncoding RNA that primarily serves as a prognostic factor for glioma patient survival, as well as a biomarker for identifying glioma molecular subtypes, a critical regulator of cell cycle progression.

Keywords: cell cycle, glioma, HOTAIR, molecular subtype, survival.

Over the last few decades, researchers have been exploring novel, noncoding RNAs (ncRNAs) to characterize their potential roles in biological processes and disease development.1–4 The human genome includes a number of ncRNAs, such as microRNAs (miRNAs), long noncoding RNAs (long ncRNAs, lncRNAs), Piwi-interacting RNAs, and small
nucleolar RNAs. LncRNAs are nonprotein coding transcripts longer than 200 nucleotides and are implicated in a number of important events, such as epigenetic regulation, transcriptional regulation, and posttranscriptional regulation.\(^5,6\) LncRNAs exhibit unique profiles in various human cancers, which reflect disease progression and serve as a predictor of patient outcomes.\(^7-9\) It was recently discovered that lncRNAs function in various aspects of cell biology and can potentially contribute to tumor development.

Gliomas are the most frequent primary tumors in the brain. Despite recent advances in cancer treatment, this statistic has not changed significantly.\(^10,11\) Therefore, it is essential to investigate the mechanism involved in the development and progression of glioma. In the current study, we profiled the lncRNA Hox transcript antisense intergenic RNA (HOTAIR), which was closely correlated with tumor grade, poor prognosis, and molecular subtype in glioma. Thus, we clarified the clinical significance and function of HOTAIR in glioma by analyzing clinical and molecular pathology features and performing gene set enrichment analysis (GSEA) and in vitro and in vivo experiments.

Materials and Methods

Patients and Samples

In total, 295 glioma samples from the Chinese Glioma Genome Atlas (CGGA, http://www.cgcg.org.cn/) were included in this study. The first cohort (CGGA1) comprised 58 astrocytomas, 17 oligodendroglialomas, 22 oligoastrocytomas (OAs), 8 anaplastic astrocytomas (AAs), 11 anaplastic oligodendroglialomas (AOs), 15 anaplastic oligoastrocytomas (AOAs), 89 tumors of glioblastoma multiforme (GBM), and 5 normal brain tissue samples. The second cohort (CGGA2) comprised 8 astrocytomas, 10 oligodendroglialomas, 8 OAs, 7 AAs, 5 AOs, 3 AOAs, and 34 GBM. Anaplastic glial (AG) includes AA, AO, and AOA. High-grade glioma (HGG), includes AG and GBM. Further information is described in the Supplementary material. Glioma gene expression datasets are deposited at the Repository of Molecular Brain Neoplasia Data (REMBRANDT; http://caintegrator.nci.nih.gov/rembrandt/) and the Gene Expression Omnibus Web site (http://www.ncbi.nlm.nih.gov/geo/, accession nos. GSE4290 and GSE7181).

Described previously\(^12\) have been whole genome gene profiling, pyrophosphate sequencing for the isocitrate dehydrogenase 1 (IDH1) mutation (IDH1R132), pyrophosphate sequencing for O\(^6\)-DNA methylguanine-methyltransferase (MGMT) promoter methylation, and immunohistochemistry.

Gene Set Enrichment Analysis With HOTAIR Expression

The gene expression profiles of GBM samples from CGGA1 were analyzed by GSEA.\(^13\) For GSEA, HOTAIR expression was treated as a binary variable divided into low or high HOTAIR expression by a criterion of whether the value was greater than the median. To determine functional gene sets for GSEA, we used gene sets from an analysis of global occupancy of H3K27me3 and enhancer of zeste homolog 2 (EZH2) induced by HOTAIR overexpression in MDA-MB-231 breast cancer cells.\(^14\) As a metric for ranking genes in the GSEA, the difference between the means of samples with low and high HOTAIR expression was used, and the other parameters were set by their default values.

Cell Experiments

Cell culture and transfection, colony formation assay, cell cycle analysis, and western blot analysis are described in the Supplementary material.

Orthotopic Glioma Model and Treatment

Bagg albino (BALB)\(c\) nude mice at 4 weeks of age were purchased from the Animal Center of the Cancer Institute at the Chinese Academy of Medical Science. To establish intracranial gliomas, \(0.5 \times 10^3\) U87 cells were transduced with luciferase lentivirus of small interfering (si)HOTAIR and then implanted stereotactically. Mice were imaged for Flu activity using bioluminescence imaging on days 1, 10, 20, and 30.

Statistical Analysis

Kaplan–Meier survival analysis was used to estimate the survival distributions, and the log-rank test was used to assess the statistical significance between stratified survival groups using the median value as the cutoff. Cox proportional hazards regression analyses were performed using SPSS software for Windows. Pearson correlation was used to determine significant differences. One-way ANOVA was used to test for differences among at least 3 groups, and a least significant difference post-hoc test was used to obtain individual \(P\) values followed by ANOVA. The \(t\) test was used to determine differences in each 2-group comparison. All data are presented as mean ± standard error. A 2-sided \(P\) value of <.05 was regarded as significant.

Results

HOTAIR Expression Correlates With Glioma Grade

First, we analyzed HOTAIR expression level in whole genome gene profiling of 220 glioma and 5 normal tissues. HOTAIR expression was significantly higher in HGG than in low-grade glioma (LGG; \(P < .001\)). Moreover, as shown in Fig. 1A, GBM demonstrated a significant increase in HOTAIR transcript levels, compared with that observed in normal tissues (\(P = .093\)), LGGGs (\(P < .001\)), or AGs (\(P = .011\)). No significant difference in HOTAIR expression levels was observed between LGG and AG (\(P = .326\)). Next, we employed 2 independent
glioma gene expression datasets (REMBRANDT and GSE4290) to examine the association between HOTAIR expression levels and glioma grade (Supplementary Fig. S1A). One-way ANOVA showed that HOTAIR was significantly associated with tumor grade ($P = 0.002$ and $P = 0.001$ for REMBRANDT data and GSE4290 data, respectively), which was consistent with the CGGA1 data. These findings suggest that HOTAIR may play an important role in glioma development.

**HOTAIR Overexpression Confers a Poor Prognosis in Glioma Patients**

Next, we investigated the correlation between HOTAIR expression and overall survival using Kaplan–Meier survival curve analysis with a log-rank comparison. HGG samples expressing higher than median levels of HOTAIR were associated with decreased survival relative to those with HOTAIR levels lower than the median ($P = 0.031$) in the CGGA1 data (Fig. 1B). Further, HOTAIR expression was inversely correlated with overall survival in AG ($P = 0.0284$) and GBM ($P = 0.0077$) (Fig. 1C and D), and similar results were detected in the REMBRANDT data (Supplementary Fig. S1B). Highly statistically significant correlations were observed between overall survival and the expression levels of HOTAIR ($P < 0.0001$ for HGG; $P = 0.0091$ for AG); however, the $P$ value for GBM ($P = 0.0759$) did not reach statistical significance. To further confirm these results, we performed microarray analysis to examine HOTAIR levels in another independent cohort of Chinese glioma (CGGA2). As shown in Fig. 1E and F, HOTAIR expression was significantly increased in HGG compared with LGG ($P < 0.001$), and cases of GBM that were highly positive for HOTAIR had a markedly worse outcome ($P = 0.0088$). Overall, these data indicate that HOTAIR...
overexpression correlates with a significantly worse survival outcome.

**HOTAIR is an Independent Prognostic Factor in GBM Patients**

High expression of HOTAIR was associated with older age at diagnosis (P = .012), nonmutated IDH1 (P < .001), unmethylated MGMT promoter (P = .027), and high expression of epidermal growth factor receptor (EGFR; P = .005) (Table 1). Next, we conducted univariate Cox regression analysis using clinical and genetic variables for 89 GBM patients from the CGGA1 cohort and found that high expression of HOTAIR, high KPS score, and total resection were statistically associated with overall survival, while IDH1 mutation and MGMT promoter methylation were not associated with overall survival (Table 2). Then we evaluated the factors that contributed to overall survival using a multivariate Cox proportional hazards model. The analysis revealed that HOTAIR expression, KPS score, and total resection correlated independently with overall survival (hazard ratio [HR] = 2.933, P = .005; HR = 0.508, P = .048; HR = 0.416, P = .034, respectively) when considering gender, Ki-67, EGFR, proliferating cell nuclear antigen (PCNA), topoisomerase II, and glutathione S-transferase (GST)–π expression (P < .3, univariate Cox regression analysis).

**HOTAIR is a Marker for Glioma Molecular Subtype**

The Cancer Genome Atlas (TCGA) network described a robust gene expression–based molecular classification of GBM into classical, mesenchymal, neural, and proneural subtypes.\(^\text{15}\) We applied the TCGA classification system to the CGGA1, REMBRANDT, and GSE4290 data and annotated the samples according to the 4 TCGA subtypes using the prediction analysis of microarrays classifier. One-way ANOVA indicated a markedly significant difference in HOTAIR expression between the 4 glioma subtypes in the CGGA1 datasets. In particular, HOTAIR expression in the classical and mesenchymal subtypes was higher than the expression in the neural and proneural subtypes (Fig. 2A). Recently, we proposed a new classification system containing 2 major subtypes: mesenchymal-like (containing classical and mesenchymal subtypes) and proneural-like (containing neural and proneural subtypes), which is primarily based on the treatment efficacy of temozolomide in the different subtypes.\(^\text{12}\) Further, we found that HOTAIR expression in the mesenchymal-like subtype was significantly higher than in the proneural-like subtype (P < .001; Fig. 2B). Moreover, our recent study has identified 3 major groups of gliomas from 220 CGGA1 samples (referred to as G1, G2, and G3) that have distinctly different clinical prognoses and molecular characteristics.\(^\text{16}\) HOTAIR expression in the G3 subtype was statistically higher than that in the G1 or G2 subtype (Fig. 2C). A similar trend was observed in the REMBRANDT and GSE4290 data (Fig. 2).

HOTAIR regulates Polycomb repressive complex 2 (PRC2)–dependent histone H3 lysine 27 trimethylation and gene silencing in breast and colorectal cancers.\(^\text{14,17}\) Because EZH2, suppressor of zeste 12 homolog (SUZ12), and embryonic ectoderm development (EED) are the core components of PRC2,\(^\text{18}\) we first explored their expression profile in CGGA1. The level of EZH2

---

**Table 1.** Clinical and molecular pathology features of GBM samples in association with HOTAIR expression

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low</th>
<th>High</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, female/male</td>
<td>18/27</td>
<td>19/25</td>
<td>.761</td>
</tr>
<tr>
<td>Age at diagnosis, y</td>
<td>42.3 ± 11.5</td>
<td>49.0 ± 13.2</td>
<td>.012</td>
</tr>
<tr>
<td>Overall survival</td>
<td>498.9 ± 232.7</td>
<td>395.0 ± 192.9</td>
<td>.026</td>
</tr>
<tr>
<td>KPS score (≥80/ &lt;80)</td>
<td>34/11</td>
<td>25/19</td>
<td>.062</td>
</tr>
<tr>
<td>Resection (subtotal/total)</td>
<td>24/21</td>
<td>25/19</td>
<td>.164</td>
</tr>
<tr>
<td>IDH1 mutation (no mutation/mutation)</td>
<td>23/12</td>
<td>39/1</td>
<td>.000</td>
</tr>
<tr>
<td>MGMT promoter methylation (unmethylation/methylation)</td>
<td>15/12</td>
<td>24/5</td>
<td>.027</td>
</tr>
<tr>
<td>MGMT (low/high)</td>
<td>13/32</td>
<td>14/28</td>
<td>.654</td>
</tr>
<tr>
<td>Ki-67 (low/high)</td>
<td>19/26</td>
<td>24/18</td>
<td>.164</td>
</tr>
<tr>
<td>EGFR (low/high)</td>
<td>25/20</td>
<td>11/31</td>
<td>.005</td>
</tr>
<tr>
<td>PCNA (low/high)</td>
<td>37/8</td>
<td>27/15</td>
<td>.058</td>
</tr>
<tr>
<td>PTEN (low/high)</td>
<td>0/45</td>
<td>2/42</td>
<td>.444</td>
</tr>
<tr>
<td>TOPO II (low/high)</td>
<td>17/28</td>
<td>18/24</td>
<td>.629</td>
</tr>
<tr>
<td>GST-π (low/high)</td>
<td>22/23</td>
<td>25/17</td>
<td>.320</td>
</tr>
</tbody>
</table>

Abbreviations: TOPO II, topoisomerase II; PTEN, phosphatase and tensin homolog.

**Table 2.** Cox proportional hazards regression analyses of HOTAIR expression and other characteristics in relation to overall survival in GBM

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariable Regression</th>
<th>Multivariable Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>P</td>
</tr>
<tr>
<td>Gender, female/male</td>
<td>1.349</td>
<td>.263</td>
</tr>
<tr>
<td>Increasing age</td>
<td>1.007</td>
<td>.498</td>
</tr>
<tr>
<td>KPS score (≥80)</td>
<td>0.318</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Total resection</td>
<td>0.550</td>
<td>.022</td>
</tr>
<tr>
<td>IDH1 mutation</td>
<td>0.734</td>
<td>.422</td>
</tr>
<tr>
<td>MGMT promoter methylation</td>
<td>1.136</td>
<td>.713</td>
</tr>
<tr>
<td>High HOTAIR</td>
<td>2.021</td>
<td>.009</td>
</tr>
<tr>
<td>High Ki-67</td>
<td>1.508</td>
<td>.127</td>
</tr>
<tr>
<td>High EGFR</td>
<td>1.430</td>
<td>.201</td>
</tr>
<tr>
<td>High MGMT</td>
<td>0.905</td>
<td>.966</td>
</tr>
<tr>
<td>High PCNA</td>
<td>1.571</td>
<td>.111</td>
</tr>
<tr>
<td>High PTEN</td>
<td>1.036</td>
<td>.972</td>
</tr>
<tr>
<td>High TOPO II</td>
<td>1.618</td>
<td>.087</td>
</tr>
<tr>
<td>High GST-π</td>
<td>1.393</td>
<td>.217</td>
</tr>
</tbody>
</table>

Abbreviations: TOPO II, topoisomerase II; PTEN, phosphatase and tensin homolog.
expression was the most significantly associated with glioma grade (Fig. 3A). Then, we tested whether HOTAIR expression levels in 89 GBM samples from CGGA1 were highly correlated with the previously identified gene expression signatures of EZH2. Indeed, gene signatures with HOTAIR-induced EZH2 occupancy \( (P = .134 \text{ and FDR} = 0.141) \) and H3K27me3 occupancy \( (P = .112 \text{ and FDR} = 0.117) \) were not significantly enriched in GBM (Fig. 3B), which is incompatible with previous data that demonstrated that HOTAIR expression induced genome-wide retargeting of PRC2 in breast cancers and colorectal cancers.\(^{14,17}\) Inconsistent results were also obtained from the GBM data in GSE4290 (Supplementary Fig. S2). These results indicate that HOTAIR may regulate gene expression patterns in an EZH2-independent manner in glioma.

**HOTAIR Is a Cell Cycle–Associated Long Noncoding RNA**

To identify the mechanism of HOTAIR involvement in glioma, we first screened differentially expressed genes in HGG compared with LGG and found 1928 upregulated genes and 815 downregulated genes in HGG. Further coexpression analysis of these 2743 genes
revealed that 1400 genes and 637 genes were positively and negatively, respectively, correlated with HOTAIR expression ($P < .05$; Fig. 4A). These genes will be called “HOTAIR-associated genes.” GSEA was used to evaluate the pathways that were differentially expressed between patients with high levels of HOTAIR expression and those with low levels of HOTAIR expression. GSEA analysis revealed that HOTAIR regulates genes primarily associated with cell cycle progression (Fig. 4B).

Next, we used a colony formation assay and evaluated cell cycle distribution to determine the role of HOTAIR in glioma cell proliferation. As shown in Fig. 5A, LN229 and U87 cells exhibited a significant reduction in colony formation after 2 weeks of siHOTAIR treatment compared with the control group. In addition, a reduction in HOTAIR expression resulted in a significant increase in cells in the G0/G1 phase in both LN229 and U87 cell lines (Fig. 5B). Next, we explored expression changes in cell cycle proteins. Western blot assays revealed that HOTAIR knockdown triggered a reduction in cyclin D1, cyclin E, cyclin-dependent kinase (CDK)4, CDK2, and E2F1 expression and an induction of p21 and p16 expression (Fig. 5C).

Further, we employed a U87 orthotopic glioma model to detect the in vivo function of HOTAIR. Statistically significant difference in tumor volume indicated by bio-luminescence imaging appeared between the control and siHOTAIR treated group on day 10 (Fig. 5E). Representative images of mice implanted with intracranial tumors are shown in Fig. 5D.

**Discussion**

HOTAIR has been characterized as a negative prognostic factor in liver, colon, and laryngeal squamous cancer patients. In our study, HOTAIR was identified as a critical marker not only for tumor grade and outcome but also for molecular subtype in glioma. HOTAIR expression was low in LGG but high in HGG samples. Glioma patients with high HOTAIR expression had a poorer prognosis for overall survival than did those with low HOTAIR expression. Multivariate Cox regression analysis revealed that HOTAIR was an independent prognostic factor in GBM patients. Our further data indicated that HOTAIR expression in the classical or mesenchymal subtype was higher than in the neural or proneural subtype, which was consistent with HOTAIR expression status in the mesenchymal-like and proneural-like subtypes. This suggests that HOTAIR plays an important role in glioma molecular classification and may serve as a novel therapeutic target for classical and mesenchymal glioma subtypes.

Gupta and colleagues reported that HOTAIR-induced genome-wide retargeting of PRC2, which is composed of EZH2, SUZ12, and EED, led to H3K27me3 modification and promoted metastasis of breast cancer by silencing multiple metastasis-suppressing genes. Additional data that used GSEA also showed that HOTAIR expression induced genome-wide retargeting of PRC2 in colon cancer. However, GSEA in our study suggested that HOTAIR may regulate gene
expression in an EZH2-independent manner in glioma. Recent data suggest that HOTAIR preferentially occupies a GA-rich DNA motif to nucleate broad domains of Polycomb occupancy and H3K27me3 and that HOTAIR occupancy occurs independently of PRC2. These data suggest that HOTAIR-dependent gene regulation in glioma cells is complex and differs significantly from the reports of its activity in other cancer cells.

HOTAIR has been well documented to play a role in the invasion of breast, colon, and liver cancer cells. Another recent study showed that HOTAIR knockdown not only inhibited cell invasion but also decreased cell proliferation, altered cell cycle progression, and induced apoptosis in pancreatic cancer cells. In our study, GSEA demonstrated gene set differences between HOTAIR high- versus low-expressing patients, which indicated that HOTAIR regulates genes involved in cell cycle progression. Additionally, a reduction in HOTAIR expression induced cell cycle G0/G1 arrest and was accompanied by changes in the expression of cell cycle–associated proteins.

In summary, we show that HOTAIR is a favorable factor for malignant progression and poor prognosis in glioma patients and exhibits pro-oncogenic activity by modulating the cell cycle. To our knowledge, this is the first study that used clinical glioma samples to demonstrate the clinical signature and biological function of HOTAIR in glioma. Therefore, understanding the
regulation of HOTAIR expression in glioma could lead to the identification of new therapeutic targets for treating glioma and warrants further investigation.

**Supplementary Material**

Supplementary material is available online at *Neuro-Oncology* (http://neuro-oncology.oxfordjournals.org/).

**Funding**

This work was supported financially by grants from the National High Technology Research and Development Program 863 (2012AA02A508), the International Science and Technology Cooperation Program of China (2012DFA30470), the National Natural Science Foundation of China (91229121, 81272792, 81172406, 81101901), the Jiangsu Province’s Key Discipline of Medicine (XK201117), the Key Foundation of Jiangsu Province Science and Technology Commission (BL2012028), Provincial Initiative Program for Excellency Disciplines, Jiangsu Province, and the Jiangsu Province’s Medical Major Talent program (RC2011051), Natural Science Foundation of Tianjin Municipal Science and Technology Commission (12JCZDJC24300).

**Conflict of interest statement.** None declared.
References