Selective Detection of the D-enantiomer of 2-Hydroxyglutarate in the CSF of Glioma Patients with Mutated Isocitrate Dehydrogenase

Juliya Kalinina1, Jun Ahn2, Narra S. Devi1, Liya Wang3, Yuancheng Li3, Jeffrey J. Olson1,9, Michael Glantz4, Thomas Smith5, Ella L. Kim6, Alf Giese6, Randy L. Jensen7, Clark C. Chen4, Bob S. Carter8, Hui Mao3,9, Miao He2, and Erwin G. Van Meir1,9,10

Abstract

Purpose: Elevation in D-2-Hydroxyglutarate (D-2HG) has recently emerged as a mandatory byproduct of mutated Isocitrate Dehydrogenase (IDH) genes 1 and 2 in glioma patients. The goal of the present study was to demonstrate the feasibility of detection of elevated levels of D-2HG in the cerebrospinal fluid (CSF) of glioma patients that carry point substitutions in the IDH gene.

Experimental Design: We developed a mass spectrometry (MS)–based platform to detect and quantify the D- and L-forms of 2HG in the CSF of glioma patients. Three independent cohorts of patients were analyzed, comprising a total of 176 samples derived from 84 patients. The levels of D- and L-2HG were used to stratify patients into IDH wild-type or IDH-mutated groups using an empirically obtained threshold of 0.69 μmol/L.

Results: Using this platform, a greater than 17-fold mean increase in D-2HG was observed in the CSF of patients with IDH mutant versus wild-type gliomas. The means for the D-2HG levels in CSF were 0.427 μmol/L in wild-type and 7.439 μmol/L in mutant groups. The C statistic for the receiver operator curve was 0.938, with 84% sensitivity, 90% specificity, and 89% accuracy to detect D-2HG. The levels of D- and L-2HG in CSF from wild-type patients varied by location of CSF draw (cisternal–ventricular–lumbar).

Conclusions: Our findings demonstrate that the CSF of patients harboring IDH mutant gliomas contain increased levels of D-2HG, which can be reliably detected with a MS-based platform.

Introduction

Somatic heterozygous mutations are present in metabolic enzyme genes Isocitrate Dehydrogenase (IDH) 1 and 2 in patients with gliomas and acute myeloid leukemia (AML), as well as a variety of other neoplasms (1–4). The mutations occur in the active site of the enzymes and alter mutant IDH1/2 affinities for their natural substrate, isocitrate, and expected product, α-keto-gluutarate (α-KG). The mutant IDH1/2 enzymes acquire a neomorphic activity, which leads to the reduction of α-KG to the D-enantiomer of 2-Hydroxyglutarate (2HG) (5). The D-2HG metabolic byproduct can accumulate in greater than 100-fold excess in IDH-mutated glioma and AML patients’ tumor cells, with intracellular concentration of D-2HG in the 5 to 35 micromolar range per gram of tumor (5–10). D-2HG behaves as an oncometabolite, promoting oncogenesis by reprogramming the cellular epigenome, collagen synthesis, and cell signaling by inhibiting a class of α-KG–dependent enzymes (11–13). Furthermore, glioma patients bearing IDH1/2 mutations represent a distinct clinical entity, with a unique biology, and display improved overall survival, as compared with their wild-type (WT) counterparts (1, 3, 14). IDH1/2 mutations are found in most of World Health Organization (WHO)-defined grade II/III gliomas, and represent about 10% of grade IV glioblastomas (14).

Although the exact role of excess D-2HG amid altered metabolism and malignant transformation is still being actively investigated, the diagnostic and prognostic value of mutated IDH1/2 and D-2HG is already well recognized. Antibodies directed at specific mutants of IDH1 have been successfully developed (15), and are already utilized in clinical laboratories to distinguish gliomas carrying IDH mutations from other lesions. In addition, the feasibility of exploiting D-2HG diagnostically as a noninvasive imaging marker for tumors harboring IDH mutations using magnetic resonance spectroscopy (MRS) has been shown (8, 16–18).

We have developed a mass spectrometry (MS)–based platform for the detection and accurate quantification of D- and L-forms of 2HG in the cerebrospinal fluid (CSF) of brain tumor patients. This method is less invasive than neurosurgical procedures, more

Note: Supplementary data for this article are available to Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

J. Kalinina and J. Ahn are co-first authors.

Note: Juliya Kalinina is now a full-time employee of Merck & Co., Inc. Miao He is now a faculty at the Department of Pathology and Laboratory Medicine at the University of Pennsylvania.

Corresponding Author: Erwin G. Van Meir, Winship Cancer Institute, Emory University, 1365C Clifton Rd., NE, Rm CS078, Atlanta, GA 30322. Phone: 404-778-5563; Fax: 404-778-5550; E-mail: evanmei@emory.edu
©2016 American Association for Cancer Research.
Translational Relevance

The diagnosis of neoplasia in the central nervous system is based on clinical symptoms, radiologic imaging, and invasive biopsy. There is an unmet need for the quantitative detection of biomarkers in biofluids of brain tumor patients to complement clinical examination and neuroradiographic images, and obviate the need for invasive neurosurgical procedures. Mutations in the Isocitrate Dehydrogenase (IDH) genes 1 and 2 are present in most of human World Health Organization grade II and III gliomas. The mutated enzymes have a neomorphic activity that results in elevated production of metabolite D-2-Hydroxyglutarate (D-2HG) in the tumors. Here, we demonstrate that D-2HG can be detected through mass spectrometry in the cerebrospinal fluid of glioma patients with mutated IDH1/2, supporting its further development as a clinical test for diagnosis and follow-up.

Methods

CSF and tumor tissue samples collection and processing

A total of 176 CSF samples (N = 84 baseline, plus N = 92 longitudinal) with matching tumor tissue samples (N = 82) or IHC slides (N = 2) from cancer patients were selected for this study (Supplementary Table S1). They include gliomas with astrocytic, oligoastrocytic, and oligodendrocytic features (N = 77), meningiomas (N = 2), ependymomas (N = 2), melanoma (N = 1), adenocarcinoma (N = 1), and mammary gland carcinoma (N = 1), as classified by the World Health Organization (2007 edition). Samples were collected retrospectively from the following institutions: Penn State University College of Medicine (Hershey) and University of Massachusetts Medical School (Worcester, MA) for cohort 1 (N = 29 baseline; N = 92 longitudinal; N total = 121), University of Mainz (Germany) for cohort 2 (N = 39), and University of California (San Diego) and Huntsman Cancer Center (Salt Lake City) for cohort 3 (N = 16).

The CSF originated from the ventricular system (N = 113), the cavernous or subarachnoid space of the convexities (cisternal; N = 44) or from the spinal canal (lumbar; N = 15), with four samples of unknown origin. Most CSF samples were collected just before or during the surgery (baseline CSF samples). In select patients, an intrathecal access device (Ommaya reservoir) was placed for installation of intrathecal chemotherapy after standard of care treatments were exhausted. This allowed for serial CSF withdrawal and analysis (longitudinal samples). The Institutional Review Board (IRB) committees at the respective institutions approved the CSF collections, and patients gave informed consent. Due to IRB restrictions, we received de-identified samples accompanied with only limited clinical information, e.g., clinical diagnosis, age at the time of CSF draw, sex, and the site of procurement of the CSF (Supplementary Table S1). Furthermore, due to the retrospective nature of our study, potential variables such as duration of storage, total volume drawn, time of day for drawing, or position of lumbar puncture were not standardized or recorded.

The CSF were spun after collection at 3,000 rpm at 4°C to remove cells and debris, and were stored in polypropylene tubes at −80°C. Matching brain tissue was supplied either frozen or formalin-fixed and paraffin-embedded.

Standards: Preparation of D- and L-2HG

D-2HG standard stock was prepared by mixing 25 mg of D-2HG (Santa Cruz Biotech) in 2.604 mL of 0.02 N hydrochloric acid. L-2HG standard stock was prepared by mixing 25 mg of L-2-Hydroxyglutarate (Santa Cruz Biotech) in 3.376 mL of 0.02 N hydrochloric acid.

Derivatization of D- and L-2HG

The derivatization step was adapted from a prior study (20). The CSF samples were thawed on ice, filtered using 10 kDa membrane filters, and centrifuged at 10,000 rpm for 10 minutes. 100 μL of neat samples were then transferred to screw cap glass tubes, and 100 μL of 10 μmol/L internal standards were added to each. The solutions were then dried under N₂ at 50°C, after which 100 μL of 50 g/L of diacetyl-L-tartaric anhydride (DATAN) was added to each tube (Fig. 1A). The reaction was allowed to proceed for 30 minutes at 75°C, and then dried under N₂ and lyophilized overnight. Once completely dry, the mixture was reconstituted in 200 μL of ultrapure water and aliquoted for subsequent MS analysis. The reaction produced D- or L-diastereoisomers capable of separation on an achiral column (Waters X Terra MS C18 5 μm, 3.9 × 150 mm Column).

Quality control sample preparation and analysis

Quality control (QC) samples were included and analyzed to assess reproducibility of every run as well as the validity of the method. Clinically “normal” pooled CSF was used to prepare the QC samples. QC-High was formulated by spiking pooled CSF with 50 μmol/L of D- or L-2HG, whereas QC-Low by spiking with 0.5 μmol/L of D- or L-2HG. Twenty runs of “Low” and “High” D- and L-2HG levels were collected, and variations of QC measurements between runs recorded. Calibration curves of standards were constructed with 0, 0.1 μmol/L, 0.5 μmol/L, 5 μmol/L, 10 μmol/L, and 50 μmol/L of D- or L-2HG. Acceptance criteria for QC were set within two standard deviations from the historic mean, and the R² or the Pearson Coefficient of calibration curves were greater than 0.9. Data from a sample run were considered invalid if concurrent QC or calibration curves did not meet the acceptance norms.

LC/MS analysis

LC/MS was used to detect D- or L-2HG in CSF with stable-isotope-labeled internal standards (D, L-[3,3,4,4-²H₄]-2-hydroxyglutaric acid). AB-Sciex QTRAP 5500 was used in negative ion mode.
mode. The analytes were separated using LC (Shimadzu) with a C18 HPLC column in-line with a C18 HPLC guard column (Waters). IDH1/2 genes’ mutation status was unknown at the time of MS analysis. The MS analysis was done using a Quantitative Gradient Method using 7.0% Acetonitrile in Ammonium Formate buffer (pH = 3.6; flow rate: 0.4 mL/min). Quattro Micro Tune Parameters using negative electro-spray ion mode at 750 V and 4,500 V with –7.5 V collision energy were used. Scan parameters of 200 m/z to 550 m/z were employed.

NMR analyses

Nuclear Magnetic Resonance (NMR) analyses were performed on frozen tumor samples as previously described (16).

Data processing/statistical analysis

Calculations were performed on the Analyst software suite, ABIX. The concentrations of D- and L-2HG in each sample were calculated using an isotope dilution method. Final results were corrected for dilution factors and reported as μmoles/L (μmol/L).

T-test, one-way ANOVA comparisons (Tukey correction), and receiver operator curve (ROC) analyses were performed using GraphPad Prism software (v6.04). The data are expressed as geometric mean with 95% confidence intervals (CI).

Identification of patient tumors with mutated IDH1/2

Sequencing and IHC procedures were previously described (16).

Results

MS-based assay development for D- and L-2HG detection in human CSF

Methodology development for the identification of D- and L-2HG in human CSF using MS. To develop a robust and accurate method for the measurement of D- and L-2HG in human CSF using MS, we adapted a protocol for measurement of urinary D- and L-2HG by LC-MS (20). Using a derivatization step of a racemic 2HG mixture in human CSF with DATAN (Fig. 1A), we were able to achieve robust separation of the D- and L-2HG isomers on the achiral column coupled to the tandem MS/MS instrumentation with retention times of 3.5 minutes for L-2HG and 4.3 minutes for D-2HG (Fig. 1B).

Analytical measurement/sensitivity range determination. Assay sensitivity and analytical range determination were evaluated in triplicate using a 10-point, 2-fold serial dilution of standard D- or L-2HG in MilliQ grade water from 0.0005 to 10,000 μmol/L. The linearity was obtained between 0.05 and 1,000 μmol/L, i.e., low and high limits of detection (LOD/HOD), respectively. The limit of quantification was validated at 0.1 μmol/L with a limit of detection at 0.05 μmol/L. Samples with D- or L-2HG concentration >1 mmol/L were re-measured with dilutions. The coefficient of variation (CV) of 20 measurements of both high and low levels of D- and L-2HG was <20% (Table 1).

Stability of samples. CSF sample stability was assessed in two studies using QCs and two representative CSF samples: (1) a
Table 1. Summary of assay properties of the D- and L-2HG in human glioma CSF

<table>
<thead>
<tr>
<th>Assay parameter</th>
<th>D-2HG added</th>
<th>Endogenous D-2HG species in CSF</th>
<th>L-2HG added</th>
<th>Endogenous L-2HG species in CSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical reference range (µmol/L)</td>
<td>0.05–1,000</td>
<td>0.05–1,000</td>
<td>0.05–1,000</td>
<td>0.05–1,000</td>
</tr>
<tr>
<td>LLOQ (µmol/L)</td>
<td>0.1</td>
<td></td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Inter-assay precision (%CV)</td>
<td>3.7</td>
<td>13.8</td>
<td>4.4</td>
<td>17.9</td>
</tr>
<tr>
<td>Intra-assay precision (%CV)</td>
<td>0.8</td>
<td>4.2</td>
<td>0.7</td>
<td>2.4</td>
</tr>
<tr>
<td>CSF spike recovery (%)</td>
<td>95.6–107.0</td>
<td></td>
<td>93.2–109.0</td>
<td></td>
</tr>
<tr>
<td>Freeze-thaw stability</td>
<td>6X thaw cycles</td>
<td>6X thaw cycles</td>
<td>6X thaw cycles</td>
<td>6X thaw cycles</td>
</tr>
<tr>
<td>Stability at RT for 24 hours</td>
<td>Stable</td>
<td>Stable</td>
<td>Stable</td>
<td>Stable</td>
</tr>
<tr>
<td>Interference with L-2HG</td>
<td>None (up to 50 µmol/L)</td>
<td>None (up to 50 µmol/L)</td>
<td>None (up to 50 µmol/L)</td>
<td>None (up to 50 µmol/L)</td>
</tr>
</tbody>
</table>

NOTE: Lowest limit of quantification (LLOQ) is the lowest back-interpolated standard that provides signal 2X over background with percent coefficient of variance <20%. For other measures, mean or range %CV is shown. Spike recovery ranges are reported only for the exogenous D- or L-2HG added. Freeze-thaw stability refers to the number of freeze-thaw cycles that allow <10% change in the signal from the original controls. Stability at room temperature (RT) for 24 hours of a sample was confirmed as “stable” if it met the acceptance criteria of <10% change from the original controls. Samples were run in duplicates.

freeze-thaw study at −80°C and (2) a stability study at room temperature for 24 hours. To assess the response of 2HG to freeze-thaw stability, CSF samples stored at −80°C were thawed completely at room temperature, then placed back at −80°C until 14X freeze-thaw cycles (over the course of 4 months) were achieved. Our results indicate that D- and L-2HG in CSF can withstand up to 6 freeze-thaw cycles (CV = 9.24%; Table 1). In addition, to define the stability of D- and L-2HG, the samples were assayed on day 1 and then reassayed on day 2 after keeping them at room temperature for 24 hours (CV = 4.81%; Table 1). The results indicate that the samples are stable at room temperature at least over the course of 24 hours.

Recovery of D- and L-2HG in human CSF. Next, we analyzed the recovery of exogenous D- or L-2HG in 4 representative CSF samples ± spiking with 50 µmol/L of D- or L-2HG standards. The percentage of recovery ranged between 95.6%–107.0% and 93.2%–109.0% for D- and L-2HG, respectively (Table 1).

Inter- and intra-assay variability testing. To test intra-assay precision, Low and High QCs and 2 representative samples of 100 µL human CSF were used to measure the concentrations of D- or L-2HG reference compounds. The results of 10 replicates of the intra-day study showed the %CV is between 0.7% and 4.2% (Table 1).

The inter-assay precision was assessed on 10 separate days/runs using duplicate measurements. Low and High QCs and 2 representative CSF samples were utilized. The inter-assay %CV ranged between 3.7 and 17.9 (Table 1).

Both intra- and inter-assay precision thus fell well within the acceptance criteria of CV < 20%.

Effect of L-2HG interference. To test interference of L-2HG in the measurement of D-2HG, we also performed interference experiments in six representative CSF samples spiked with five different levels (0.1 µmol/L, 0.5 µmol/L, 5 µmol/L, 10 µmol/L, 50 µmol/L) of standard L-2HG. Spiked samples were then run together with unspiked (endogenous) CSF samples. Our results demonstrate that the presence of L-2HG in the samples (up to 50 µmol/L) does not interfere with the accurate measurement of D-2HG in human CSF using our assay (Table 1).

Testing of the assay in three cohorts of CSF from glioma patients with or without IDH1 mutations

To test our developed assay with the overall goal of discriminating WT from IDH1/2-mutated gliomas based on 2HG levels in the patient CSF, we obtained three individual human CSF cohorts (total N = 176; baseline N = 84; longitudinal N = 92; Supplementary Tables S1 and S2) in a blind fashion. Cohorts 1, 2, and 3 included N = 29, N = 39, and N = 16 baseline CSF samples, respectively. In addition, 20 of 29 patients from Cohort 1 had repeated CSF draws, adding 92 longitudinal CSF samples. All of the samples were representative of glioma patients at various stages of disease progression (WHO grades II to IV) or control malignancies with unknown IDH1/2 mutation status at the time of 2HG analysis (Table 2). The cohorts were tested independently with Cohort 1 being resampled twice.

Our results indicate that all of the D- and L-2HG values measured in the CSF samples through MS analysis fall within the determined analytical reference range for the developed assay: 0.05–1,000 µmol/L. The absolute levels of D-2HG ranged from 0.11–8.38 µmol/L in Cohort 1; 0.26–22.10 µmol/L in Cohort 2; and 0.10–109 µmol/L in Cohort 3. L-2HG metabolite levels ranged between 0.14–6.14 µmol/L (Cohort 1); 0.4–4.08 (Cohort 2); and 0.246–2.14 µmol/L (Cohort 3).

Confirmation of the mutational status in the matching glioma tissues

To test whether elevated D-2HG levels in the CSF could discriminate between patients with WT versus mutant IDH1/2 gliomas, we next obtained matching tumor tissues from the same patients (N = 82 tissues) and identified 16 IDH1-mutated gliomas using DNA sequencing and immunohistochemistry (Supplementary Tables S1 and S2). A single false positive was obtained using IHC, as well as two ambiguous samples, and one sample that tested positive but could not be verified using DNA sequencing.

Next, based on the sequencing and immunohistochemical data, we segregated baseline or baseline + longitudinal CSF samples into two groups, based on the WT or mutant IDH status of the corresponding patient's tumors. We plotted their D- or L-2HG values as a function of identified mutation, either as individual or combined cohorts (Fig. 2). We found significant agreement between the CSF data and the tumor data.

Table 2. Identified mutations from the tissues and quantified D- and L-2HG levels from matched CSF in a combined cohort of 1–3 based on WHO Grade
Figure 2.
Quantification of D- and L-2HG in CSF samples. A, Levels of D-2HG found in the CSF of Cohorts 1 (baseline; \( N = 29 \); left), 2 (\( N = 39 \); middle), and 3 (\( N = 16 \); right) are plotted as a function of \( \text{IDH1/2} \) mutation status. In Cohort 1, patients underwent CSF draws both before (baseline samples) and at different times after surgery (longitudinal samples). B, D-2HG levels of all CSF samples of Cohort 1 (baseline plus longitudinal; \( N = 121 \); left), of combined Cohorts 1, 2, and 3 at baseline (\( N = 84 \); middle), or at baseline plus longitudinal (\( N = 176 \); right) are shown as a function of \( \text{IDH1/2} \) mutation status. C, L-2HG levels in same CSF samples as A. D, L-2HG levels in same CSF samples as B. Scatter plots with geometric mean and 95% CIs are indicated. “Mutant” and “WT” represent samples with a mutated or WT \( \text{IDH1/2} \) gene in the matching tumor, respectively.
differences between D-2HG levels in the two groups in all cohorts (sampled individually or in combination; Fig. 2A and B; Table 2; Supplementary Tables S1 and S3). In a combined cohort, 12 of 16 baseline (or 37/44 baseline + longitudinal CSF samples) positive for IDH1/2 mutation demonstrated D-2HG levels of >0.69 μmol/L, whereas 55 of 68 baseline (or 118/132 baseline + longitudinal) WT group CSF demonstrated <0.69 μmol/L levels. These results suggest that the two groups can be separated based on their D-2HG levels using an empirically calculated threshold of 0.69 μmol/L.

The D-2HG levels in the WT group (N = 132) ranged between 0.10 and 2.57 μmol/L, and in the mutant group (N = 44) from 0.31 to 109 μmol/L. The means for the D-2HG levels were 0.427 μmol/L in the WT group and 7.439 μmol/L in the mutant group, comprising a 17-fold increase (mean difference −7.012; 95% CI, −10.11 to −3.916; P < 0.0001; Table 2; Supplementary Table S3). In addition, we consider the 0.6 to 0.7 μmol/L range as a "gray zone," where the D-2HG level cannot accurately predict whether the matching patient tumors contain IDH1/2 mutations. Note that 7 of 176 samples fell in this region.

With regard to L-2HG levels, the differences between the two groups sampled in a combined fashion were not significant, with ranges for the WT group of 0.14–6.14 μmol/L and mutant group between 0.40 and 4.08 μmol/L. The mean values were 0.87 and 0.99 for WT and mutant groups, respectively (Fig. 2C and D; Table 2; Supplementary Table S3).

Receiver operator characteristic analysis was performed on the D-2HG data from the combined cohorts pool. It demonstrated an area under the curve of 0.938 (0.882–0.983; P < 0.0001) with a sensitivity of 81% and specificity of 89%, respectively, with an area under the curve of 0.938 (0.882–0.983; 95% CI, 0.896–0.980; P < 0.0001; Supplementary Fig. S1B). For comparison, the ROC curves for L-2HG were not significant (Supplementary Fig. S1C and S1D).

D-2HG CSF levels related to tumor grade

A trend toward higher D-2HG levels in tumors of higher WHO grade was observed, even though it was not significant due to the limited number of samples per group (Supplementary Fig. S2; Table 2). Of note, 2 of 3 mutated CSF from grade II patients had D-2HG levels below our empirical threshold of 0.69 μmol/L, which prevented the discrimination between WT and mutant grade II gliomas.

D-2HG CSF levels related to patient gender and age

There was no correlation between patient gender and D-2HG levels (Supplementary Fig. S3). A significant negative correlation between D-2HG levels and age was observed (R = −0.575; P = 0.0033), likely reflecting the younger age of patients with IDH1/2 mutated tumors (40) as no correlation with age was present in mutated-only samples.

Comparison of D-2HG levels in CSF from different anatomical sources

Comparison of CSF obtained through ventricular (VT), cisternal (CT), or lumbar (LP) draws showed that the D-2HG levels were elevated in all samples derived from patients with IDH mutant gliomas, regardless of anatomical origin (Fig. 3; Table 3). The data were significant for ventricular (P = 0.0002; N (WT:Mut) = 80:33), cisternal (P = 0.0022; N (WT:Mut) = 36:8), and lumbar (P = 0.0283; N (WT:Mut) = 12:3) draws with P values proportional to number of samples. L-2HG values showed no differences between WT and Mut groups regardless of site of CSF draw.

Interestingly, we observed that the levels of both D- and L-2HG were site-dependent (Supplementary Fig. S4), suggesting heterogeneous distribution of metabolites among different CSF compartments. The highest levels were observed in cisternal, and the lowest in lumbar-derived CSF (CT>VT>LP). This finding prompted us to re-stratify 2HG thresholds based on the anatomical origin of CSF yielding the following new cut-offs for D-2HG: CT~1.46 μmol/L; VT~0.95 μmol/L; LP~0.33 μmol/L (Supplementary Table S4). These new and better-defined cut-offs enabled us to calculate site-specific values for sensitivity, specificity, and accuracy (Supplementary Table S5). The overall accuracy of distinguishing WT from mutant IDH glioma patients on the basis of D-2HG concentration in lumbar CSF improved to 100% because it “rescued” an originally false-negative sample. These measurements provide a foundation for a prospective study in a larger cohort of site-specific samples.
Comparison of MS-based to HRMAS NMR-based methodologies to detect D-2HG in human brain tumor tissues

Finally, we attempted to see how the MS-based methodology compares to that of NMR (16). To this purpose, we measured D-2HG levels in parallel by both techniques in an independent set of 12 glioma tissues with IDH mutations (Supplementary Fig. S5; Supplementary Table S6). We found a significant correlation between both methods ($r = 0.6639; P = 0.0185$), with the sensitivity of the MS-based methodology in the order of 200-fold higher than NMR.

Together, our findings demonstrate the feasibility of a MS-based platform for the specific quantification of the D-enantiomer of 2HG in CSF and provide a complementary biomarker methodology in biofluid (by MS) to that in tissues (via NMR).

### Discussion

There is a pressing need for the quantitative detection of biomarkers in patients with neoplastic diseases of the central nervous system. Currently, brain lesions are initially identified through neurological examination and imaging procedures (CT and/or MRI scans), followed by pathologic examination of tissues obtained by invasive stereotactic biopsy or tumor resection. There is no laboratory test that predicts tumor recurrence, and imaging studies are often confounded by postsurgical and posttreatment changes, including radiation, corticosteroids, pseudoprogression, antiangiogenic agents, immunotherapies, and targeted therapies (21–25). A biomarker-based and easy to administer method to monitor tumor growth, response to therapy, and recurrence would be of great complementary benefit.

To determine the potential utility of the D-2HG metabolite as such a marker for the stratification and follow-up of patients with gliomas, we developed a robust and quantitative MS-based assay that can differentially detect and quantify D- and L-2HG metabolites in the CSF of patients with gliomas with an analytical range of 0.05 to 1,000 μmol/L. The assay was subsequently extensively characterized, including inter- and intra-assay variability, recovery, linearity, stability, and interference experiments, with the results suggesting that D-2HG and L-2HG calculated using this assay represent robust measurements.

Applying the MS-based D-2HG detection assay to three independent glioma CSF cohorts collected at several collaborating institutions, we were able to discriminate glioma patients with IDH1/2 mutations with an overall empirical D-2HG threshold of 0.69 μmol/L. The CSF of patients with WT IDH1 tumors had D-2HG levels below 0.69 μmol/L, while that of patients with IDH1/2-mutated tumors consistently displayed higher levels. These values are consistent with the published range for “normal” levels for D-2HG in human CSF of 0.07 to 0.3 μmol/L (26). The sensitivity, specificity, and accuracy of this discrimination using this threshold were 84%, 90%, and 89%, respectively. It is important to note that there was a "gray" zone (0.6–0.7 μmol/L), and one IDH1-mutated and six WT group CSF samples had D-2HG levels in this range, suggesting that the accurate cut-off values in discriminating WT from IDH1/2-mutated CSF may need to be further refined with testing larger annotated collections of CSF.

A prior report has measured physiologic levels of L-2HG between 0.3 and 2.3 μmol/L in CSF (26), which agrees with our measurements. Most of our samples contained L-2HG levels of <2.3 μmol/L (N = 169). Six samples were found to have L-2HG between 2.3 and 3.1 μmol/L and three outliers contained L-2HG at 3.8, 4.08, and 6.14 μmol/L.

Our study assumes the presence of mutated IDH1/2 enzymes is the major driver of elevated D-2HG/L-2HG ratios in tumors; however, cancer-related alterations in other genes may also lead to moderately increased D- or L-2HG. For example, in tumors with elevated Muc, mitochondrial glutaminase is activated, increasing the flux from glutamine to glutamate to ketoglutarate, and then to 2HG. Likewise, it is conceivable that other mutations occurring in the background of IDH mutant tumors may modify the flux.

Due to limited sample numbers for each grade, we were unable to determine whether there is a statistically significant relationship between D-2HG levels in CSF and WHO grade of IDH-mutant tumors. The data appear to suggest that IDH-mutant Grade II gliomas release less D-2HG in CSF, which may impede discrimination between WT and mutant groups, but this needs further study due to low sample number. Density of tumor cells in tumor tissue, amount of enzyme synthesized per tumor cell, and activity of mutant enzyme in different cellular contexts will all affect D-2HG production and could vary between grades. Other confounders most likely influence the transport of metabolites from the bulk tumor to CSF: for example, edema and bulk flow from solid tumor to adjacent brain, distance from the lesions to the brain surface, and brain anatomy including a sulcus, and the inter-hemispheric and sylvian fissures. This all may affect how fast and how many tumor-produced substances enter in the CSF. Absence of clinical, imaging, and treatment information also precluded us from analyzing possible correlations of D-2HG levels with tumor volume, proximity to the ventricles, and response to therapy. The present study, nonetheless, provides impetus to analyze larger well-characterized longitudinal collections in an effort to answer these important questions.

To determine whether site of CSF sampling affects 2HG concentration, we stratified our overall threshold, and found that L- and D-2HG levels are, in fact, site-dependent, with the highest

### Table 3. D/L-2HG concentrations segregated by the site of CSF draw (t-test statistics)

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Site of CSF draw</th>
<th>Geometric mean ± SEM (μmol/L)</th>
<th>95% CI of geometric mean</th>
<th>P value difference between WT and Mut</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-2HG</td>
<td>Lumbar</td>
<td>0.173</td>
<td>2.084</td>
<td>0.35–0.221</td>
</tr>
<tr>
<td></td>
<td>Cisternal</td>
<td>0.568</td>
<td>4.824</td>
<td>0.483–0.668</td>
</tr>
<tr>
<td></td>
<td>Ventricular</td>
<td>0.314</td>
<td>2.317</td>
<td>0.279–0.353</td>
</tr>
<tr>
<td>L-2HG</td>
<td>Lumbar</td>
<td>0.303</td>
<td>0.626</td>
<td>0.202–0.455</td>
</tr>
<tr>
<td></td>
<td>Cisternal</td>
<td>1.259</td>
<td>1.491</td>
<td>1.065–1.489</td>
</tr>
<tr>
<td></td>
<td>Ventricular</td>
<td>0.622</td>
<td>0.888</td>
<td>0.540–0.716</td>
</tr>
</tbody>
</table>

*NS, not significant.
levels observed in central (cisternal or ventricular), and the lowest in distal (lumbar) CSF (CT>VT>LP). Measurements in a single patient with an IDH1-mutant tumor where both lumbar and ventricular CSF were available confirmed the existence of a gradient. The lumbar CSF sample had 3- to 10-fold lower D-2HG levels than repeated ventricular draws (Sample # 7; Cohort 1). Yet, all measurements were above our site-specific thresholds of 0.33 μmol/L and 0.95 μmol/L for LP and VT, respectively, highlighting the importance of stratifying the samples by the site of origin. Despite these differences, the amounts of D-2HG found in the CSF were significantly different between IDH WT and mutant groups, regardless of CSF site of origin.

While the CSF circulates from ventricles to the spinal canal and then comes back to the brain cisternae before draining into the circulation at the arachnoid villi, it is important to realize there is a connection between the ventricles and the cisterna magna, so the anatomy influences CSF circulation and the distribution of tumor-derived products in the CSF. While studies comparing lumbar and centrally collected CSF are limited, the available literature suggests that gradients of protein and metabolite concentration exist between these samples (27–34). For instance, IgG contents decrease progressively as the CSF moves from the site of intracranial inflammation to the lumbar sac (29). Similar gradients have been noted for other proteins and metabolites in patients suffering from bacterial meningitis (32) or cysticercosis (30). The presence of such gradients is relevant to molecular diagnostics, given results presented here, and a previous study showing that mutated IDH1 mRNA can be more readily detected in cisternal CSF samples relative to lumbar samples (35). The origin of such gradients may stem from the complexity of CSF physiology (36), and they warrant careful consideration in the design of future CSF sample-based molecular diagnostics. The circadian gradient, patient position, as well as the total volume of CSF collected may also affect measurement of D-2HG in CSF by lumbar puncture. It is also possible that patients with partially impaired CSF flow through the aqueduct of sylvius or with arachnoid adhesions over the convexities of the brain due to tumor effects would have abnormal CSF flow resulting in asymmetry in 2HG concentrations. It will be important to further address these issues, as lumbar collection is conceivable as a moderately invasive diagnostic procedure, whereas ventricular or cisternal collection involves placing intrathecal access devices that are more invasive. The latter may be unpractical for routine screening, but could be helpful to guide the treatment such as distinguishing between pseudo- and true tumor progression.

We also assessed 2HG levels in plasma and urine from glioma patients (data not shown), as did others in serum (37, 38), but no correlation between D-2HG levels and the presence of IDH1/2 mutations was found. A recent study also suggested that CSF is better than plasma for the detection of cell-free circulating tumor DNA (39).

We did not observe any correlation between patient gender and D-2HG levels. Consistent with the well-documented association between younger age and IDH1/2 mutated tumors (40), we did find a significant negative correlation between D-2HG levels and age (R = −0.575; P = 0.0033).

Another important question we attempted to answer is the correlation between repeated CSF draws in the same patients and D-2HG levels. We were not able to establish any correlation in the 20 patients where longitudinal samples were available, possibly related to the short intervals between individual sample draws or other nonstandardized confounding factors.

Importantly for clinical practice, we have demonstrated that the levels of D-2HG in CSF are stable, even with multiple freeze/thaws; as such it makes it possible to send frozen samples to a central clinical laboratory for analysis. Taken together, our findings demonstrate that D-2HG can be detected in the CSF drawn from various sites, and its potential use in the clinic can be further considered. The quantification of D-2HG in lumbar CSF could be used for diagnostic and prognostic purposes by defining the genetic background of the patient’s brain tumor, without the need for invasive biopsy or surgical resection of the tumor. Elevated D-2HG can stratify the lesion as a glioma of the proneural subtype as they carry IDH1/2 mutations (14, 41). While collection of CSF though lumbar puncture is not entirely noninvasive, and is currently not part of the standard procedures in the clinical care of glioma patients, it could be implemented if justified, as it is safe in the majority of brain tumor patients even in the setting of increased intracranial pressure (42–45). Exceptions are patients with large midline shifts, or large posterior fossa masses, in which needle puncture of the dura can lead to continued escape of CSF into the epidural space, and cause brain herniation. Patients with impaired CSF flow due to ventricular obstruction or with marked epidural spinal cord compression should also be excluded. Coagulopathy or thrombocytopenia may be further contraindications, but these are not specific or even particularly common in patients with primary brain tumors. Recent studies have shown that MRS methods specific to detect and measure D-2HG can now be incorporated in the standard clinical MRI scanner, which is safer and noninvasive compared with lumbar punctures (8, 16–18). However, the sensitivity of D-2HG detection using MRS is 25- to 200-fold lower than by the current MS-based methodology. While MS is more sensitive than NMR, this advantage is offset by the significant drop in concentration of D-2HG from tumor tissue to what is released and diluted in the CSF. Although this will need further study, we believe that both MRS data on tumor and MS data on CSF will be complementary, depending on the clinical question and availability of the samples and methods.

Detection of D-2HG in the CSF using the highly sensitive MS method described here holds perhaps the most promise for the longitudinal monitoring of glioma patients, especially in postoperative patients where the intra-cranial tumor pressure has subsided after tumor debulking. Periodical CSF draws will provide a window into the biology of the tumor, and recent studies have shown this liquid can be informative not only for tumor metabolite production, but also for the presence of metastatic tumor cells, tumor-derived exosomes, nucleic acids, and protein (19, 39, 46–49). Ultimately, diagnosis and treatment will benefit from integrating information from both MRS and lumbar CSF depending on the clinical situation.

In summary, our findings establish a clinical association between D-2HG levels in the CSF and the presence of IDH1/2 mutation in the gliomas of three independent patient cohorts. Our study demonstrates the feasibility of an MS-based platform for the specific quantification of the D-enantiomer of 2HG in CSF, without interference from the L-isofrom. A future prospective longitudinal study with well-defined inclusion/exclusion criteria and whose samples and data are collected in a coherent fashion to validate this technology is warranted. With that information in
Disclosure of Potential Conflicts of Interest

J. Olson reports receiving commercial research grants from Genentech and Millennium; other commercial research support from Merck, and is a consultant/advisory board member for the American Cancer Society. R.L. Jensen is a consultant/advisory board member for Medtronic, Pharmacokinesis, and Varian. Juliya Kalinina is now a full-time employee of Merck & Co., Inc. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions

Conception and design: J. Kalinina, J. Ahn, H. Mao, E.G. Van Meir
Development of methodology: J. Kalinina, J. Ahn, Y. Li, H. Mao, M. He
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J. Kalinina, J. Ahn, Y. Li, M. Glantz, T. Smith, A. Giese, R.L. Jensen, C.C. Chen, B.S. Carter, H. Mao

Acknowledgments

We thank Nara Sarojini Devi, Milton Merchant, Lawrence Recht and Zhaobin Zhang for help with patient sample banking.

Grant Support

This work was supported in part by grants from the NIH (CA086335, CA116804, NS096236 to E.G. Van Meir, CA169937 to H. Mao, CA138292 to Walter Curran, Winship Cancer Institute), NINDS training grant T32NS007480-11 (to J. Kalinina and Allan I. Levey), the Brain Tumor Funders Collaborative (to E.G. Van Meir), and the Georgia Cancer Coalition (to J. Kalinina and E.G. Van Meir).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 6, 2015; revised May 18, 2016; accepted May 31, 2016. Published OnlineFirst June 23, 2016.

References


Selecting Detection of the D-enantiomer of 2-Hydroxyglutarate in the CSF of Glioma Patients with Mutated Isocitrate Dehydrogenase
