A novel Cas9-targeted long-read assay for simultaneous detection of IDH1/2 mutations and clinically relevant MGMT methylation in fresh biopsies of diffuse glioma

Thidathip Wongsurawat1*, Piroon Jenjaroenpun1, Annick De Loose2, Duah Alkam1, David W. Ussery1, Intawat Nookaew1, Yuet-Kin Leung3, Shuk-Mei Ho3, John D. Day2 and Analiz Rodriguez2*

Abstract

Molecular biomarkers provide both diagnostic and prognostic results for patients with diffuse glioma, the most common primary brain tumor in adults. Here, we used a long-read nanopore-based sequencing technique to simultaneously assess IDH mutation status and MGMT methylation level in 4 human cell lines and 8 fresh human brain tumor biopsies. Currently, these biomarkers are assayed separately, and results can take days to weeks. We demonstrated the use of nanopore Cas9-targeted sequencing (nCATS) to identify IDH1 and IDH2 mutations within 36 h and compared this approach against currently used clinical methods. nCATS was also able to simultaneously provide high-resolution evaluation of MGMT methylation levels not only at the promoter region, as with currently used methods, but also at CpGs across the proximal promoter region, the entirety of exon 1, and a portion of intron 1. We compared the methylation levels of all CpGs to MGMT expression in all cell lines and tumors and observed a positive correlation between intron 1 methylation and MGMT expression. Finally, we identified single nucleotide variants in 3 target loci. This pilot study demonstrates the feasibility of using nCATS as a clinical tool for cancer precision medicine.

Keywords: Glioblastoma, MGMT, IDH, Targeted sequencing, CRISPR/Cas9, Long-read sequencing, Nanopore, Molecular marker, Methylation

* Correspondence: twongsurawat@uams.edu; arodriguez@uams.edu
1Department of Biomedical Informatics, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA
2Department of Neurosurgery, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA
Full list of author information is available at the end of the article

© The Author(s). 2020 Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
**Introduction**

Diffuse gliomas (DG) comprise 80% of primary malignant central nervous system tumors in adults and traditionally were diagnosed with pathological criteria to define histological type (e.g., astrocytoma, oligodendroglioma, or oligoastrocytoma) and malignancy grade (e.g., grades I-IV) [1–3]. In 2016, the World Health Organization (WHO) diagnostic guidelines incorporated molecular markers into the classification of DGs [4, 5]. Many of these diagnostic biomarkers also serve as prognostic indicators, and the neuro-oncology community has supported this integration of molecular markers into clinical practice [6]. However, to date, there is wide variability in biomarker assessment because molecular techniques and test validity are inconsistent throughout the world and even within geographic regions [7, 8]. Therefore, the use of novel sequencing techniques that can assess multiple biomarkers simultaneously is an attractive option to overcome current clinical practice limitations. In this pilot study, we explore the use of nanopore Cas9-targeted sequencing (nCATS) to accomplish these goals.

To diagnose DG, the presence of isocitrate dehydrogenase 1 and 2 (IDH1/2) gene mutation is required for subtype identification and is also a prognostic molecular marker [4, 9]. The methylation status of the O6-methylguanine-DNA methyltransferase (MGMT) promoter is used routinely to guide chemotherapeutic treatment decisions, especially in glioblastoma (GBM) (e.g., grade IV astrocytoma), which is the most common type of DG. Thus, IDH and MGMT are the most commonly assayed molecular markers in patients with DG [10].

Various methods can be used to screen for IDH1/2 mutation and MGMT promoter methylation. Typically, IDH1/2 mutation screening is performed with an immunohistochemistry (IHC) assay specific for the most common mutation at IDH1 arginine 132 (arginine to histidine, R132H). However, IHC cannot detect other less common mutations, including IDH1 R132S, R132C, R132G, and R132L substitutions or IDH2 R172K. Polymerase chain reaction (PCR) or Sanger sequencing is thus recommended as a second-step test for IHC-negative tumors [4, 11].

Assaying MGMT methylation requires identifying the modification of cytosine residues on CpG islands in the promoter, which includes 98 CpG dinucleotides surrounding the transcription start site. These assays vary in the methodology used and the promoter region assessed. However, most interrogate only a fraction of the CpG sites to predict the transcriptional activity of the MGMT gene and in turn to predict potential therapeutic response to temozolomide (TMZ), an oral chemotherapy drug. Two differentially methylated regions (DMRs) cover CpGs 25–50 (DMR1) and CpGs 73–90 (DMR2) and have been demonstrated to correlate with transcriptional silencing [12]. DMR2 has some cis-acting sites that control the transcription of MGMT in a cell-based reporter study [13]. The presence of MGMT promoter methylation portends responsiveness to TMZ treatment [14, 15], but the degree of methylation corresponding to TMZ treatment response is a subject of debate, and there is no consensus on which assay method is optimal. Commonly used methods such as methylation-specific PCR, pyrosequencing, and mass spectrometry (MassARRAY®) introduce PCR bias and are restricted to study limited sequence length due to bisulfite treatment [16].

Nanopore technology (Oxford Nanopore Technologies® or ONT) could overcome the limitations of the aforementioned assays to assess both methylation and mutations. Quantitative methylation assessment without bisulfite conversion is possible with nanopore sequencing, as electrolytic current signals are sensitive to methylation of carbon 5 in cytosine (5mC) [17]. In addition, with the capacity for long-read single-molecule sequencing, multiple CpGs in the promoter region and additional surrounding regions can be captured. Here, we applied nanopore Cas9-targeted sequencing (nCATS) [18] and used the low-cost nanopore MinION device (ONT) to simultaneously assay IDH mutations and MGMT methylation. We also compared our results against currently used clinical tests. We observed a positive correlation between the methylation of all captured CpGs and gene expression levels and showed that both nCATS and existing deep sequencing methods detected the same single nucleotide variants in clinical DG samples.

**Materials and methods**

**Informed consent**

This study included 8 patients diagnosed with glioma. Case records were reviewed, and brain tissue samples were obtained under the approval of the institutional review board at the University of Arkansas for Medical Sciences (IRB protocol #228443). All patients provided written informed consent. Four samples with IDH mutations and 4 with IDH wild type were selected by A.R. However, all samples were processed and analyzed in a single-blind fashion before mutational status was disclosed to the analytical group (T.W. and P.J.).

**DNA samples and DNA extraction for nCATS**

**Control DNA**

IDH1/2 wild type gDNA standards (Horizon Discovery, USA) were used as the negative control for genotyping by PCR and nanopore sequencing (ONT, USA). For positive controls, IDH1 codon 132 mutant DNA (CGT $\rightarrow$ GGT) was obtained from a patient in this study; IDH2 codon 172 mutant DNA (AGG $\rightarrow$ AAG)
was purchased from Horizon Discovery. Exon 4 of IDH1/2 of each standard was amplified using specific primers (Integrated DNA Technologies, USA). PCR conditions for IDH1/2 amplifications were identical, using 100 ng gDNA, 20 mM primers, and 25 μl LongAmp Taq 2x Master Mix (NEB, USA) with the following program: 95 °C 2 min, 25 cycles of [95 °C 15 s, 60 °C 30 s, 65 °C 40 s], 65 °C 10 min, 4 °C hold. PCR reactions were purified with AMPure XP beads (Beckman Coulter, USA) and eluted in 20 μl nuclease-free water (NEB). The purified PCR products were used for library preparation using 1D Native barcoding genomic DNA with EXP-NBD103 and SQK-LSK108 protocols (ONT) and nanopore sequencing with the R9.4.1/FLO-MIN106 flow cell (ONT).

The CpGenome™ DNA Standard Set (MilliporeSigma, USA) containing 5-mC and unmodified cytosines was used for quantitative analysis. The standard DNAs consist of linear, double-stranded DNA (897 bp) with 52 CpG sites; each standard contains either 100% 5-mCs or unmodified cytosines.

The CpGenome™ Human Methylated & Non-Methylated DNA Standard Set (MilliporeSigma) was used as the positive and negative control for nCATS and methylation status assessment. The Methylated DNA Standard is methylated enzymatically at all CpG dinucleotides (>95%). The Non-Methylated DNA Standard contains less than 5% methylated DNA.

**Cell line gDNA**

Four GBM cell lines were used in this study: U87, U251, T98G, and LN18 (Sigma, USA). The cells were grown to 85–90% confluence in 10-cm dishes in DMEM (U87) with 10% fetal bovine serum (FBS); EMEM (U251 and T98G) with 2 mM glutamine, 1% NEAA, 1 mM sodium pyruvate, and 10% FBS; and in DMEM (LN18) with 5% FBS utilizing standard techniques. The cells were washed with PBS before DNA extraction with the AllPrep DNA/RNA Mini Kit (Qiagen, USA). Eluted gDNA was purified and concentrated using AMPure XP beads and eluted in 20–40 μl nuclease-free water and stored at −20 °C.

**Clinical samples**

The study included 8 brain tissue samples graded according to the 2016 WHO classification for diffuse glioma by a board-certified neuropathologist, Murat Gokden M.D. (Table 1). Following surgical resection, tissue samples were immediately frozen on dry ice and stored at −80 °C until DNA extraction. DNA extraction was carried out with the AllPrep DNA/RNA Mini Kit (Qiagen) as described above.

### Table 1: Demographic characteristics of 8 patients

| Patient ID | P553 | P690 | P701 | P568 | P785 | P712 | P816 | P722 |
|------------|------|------|------|------|------|------|------|------|
| **Age**    | 29   | 24   | 37   | 42   | 72   | 37   | 48   | 73   |
| **Gender** | Male | Male | Male | Male | Female | Female | Female | Female |
| **Race**   | White | White | White | White | White | White | White | White |
| **Pathology Diagnosis** | Secondary GBM, WHO Grade 4 | Secondary GBM, WHO Grade 4 | Diffuse astrocytoma, WHO Grade 2 | Diffuse astrocytoma, WHO Grade 2 | Anaplastic astrocytoma, WHO Grade 3 | Anaplastic astrocytoma, WHO Grade 3 | GBM, WHO Grade 4 | GBM, WHO Grade 4 |
| **MGMT Status** | Low level detected | Detected | Detected | Detected | Not detected | Detected | Detected | Detected |
| **IDH**    | Mutant | Mutant | Mutant | Mutant | Not detected | Not detected | Not detected | Not detected |
| **Previous Chemo** | Yes | Yes | No | No | No | No | No | No |
| **Chemo Agent** | TMZ | TMZ | NA | NA | NA | NA | NA | NA |
| **Previous Radiation** | Yes | Yes | No | No | No | No | No | No |
| **Previous Radiation Dose** | 50.4 Gy | 60 Gy | NA | NA | NA | NA | NA | NA |
| **Previous Diagnosis** | Diffuse astrocytoma, WHO Grade 2 | Oligoastrocytoma, WHO Grade 2 | NA | NA | NA | NA | NA | NA |
| **Progression Interval (months)** | 30 | 55 | NA | NA | NA | NA | NA | NA |
| **Vital Status** | Alive | Alive | Alive | Alive | Alive | Alive | Deceased | Deceased |
RNA extraction
For all cell lines and tissue samples, RNA and DNA were extracted from the same samples. The AllPrep DNA/RNA Mini Kit (Qiagen) allows the simultaneous purification of gDNA and total RNA from the same sample.

Purity, quantity, and integrity of DNA and RNA
DNA and RNA purity was assessed in all samples with a NanoDrop-2000 spectrophotometer (Thermo Scientific, USA). DNA concentration was measured using a Qubit3.0 quantification assay (Thermo Scientific). The integrity of DNA and RNA was determined using a TapeStation 2200 (Agilent, USA).

Single guide (sg)RNA design
To design the crRNAs, we used CHOPCHOP as described in the ONT protocol [19]. The specificity of the crRNA was tested with the UCSC In-Silico PCR tool to search against the human genome (hg19). The designed crRNAs, tracrRNA, and HiFi Cas9 were purchased from IDT. The following crRNAs were used: MGMT_promoter_left: ATGAGG GGCCCCACTAATTGA; MGMT_promoter_right: ACCTGA GTATAAGCTCCGTAC; IDH1_left: ACAGTCCATGAATCT AACCCTG; IDH1_right: GGCACCATCGAAATATTCT; IDH2_left: GCTAGGCCAGGAGCTCCAGT; IDH2_right: GCTGTGTTGGGGCCGCTCTCGA.

nCATS library preparation for targeted sequencing by ONT
For each sample, 3.5 μg to 5.5 μg gDNA was used as input for preparing the nCATS library. The library preparation protocol was provided by ONT via the Enrichment Channel, Nanopore Community (protocol version: ENR_9084_v109_revA_04Dec2018). Briefly, gDNA ends were treated with calf intestinal phosphatase (NEB) to reduce the ligation of sequencing adapters to non-target strands. Then, Cas9 ribonucleoprotein complexes (Cas9 RNPs) were freshly prepared and used for generating double-strand breaks at targeted regions of blocked DNA. An adenine (A)-tail was immediately added to the 3′ ends of cut DNA fragments using Taq polymerase and dATP (NEB). The A overhang can pair with the T overhang of nanopore sequencing adapters. Both adapter-ligated DNA and blocked DNA were added to the flow cell for sequencing. The excess unligated adapters were removed with AMPure XP beads (Beckman Coulter). The library (molecules ligated to the adapters) were sequenced with the MinION Mk1B. Each library was sequenced for 36 h on an R9.4.1/FLO-MIN106 flow cell (ONT).

Bioinformatics and statistical analysis
Data processing and mapping reads
The ONT raw signal data (FAST5 files) generated with MinKNOW software (version 1.7.14) were converted to DNA sequence data (FASTQ files) using the GUPPY algorithm (version 3.0.3). Quality control for ONT reads was performed to filter FASTQ files based on a mean quality threshold higher than Phred score 8 and read lengths longer than 200 bases using the NanoFilt program [20]. We aligned the filtered reads to the human reference genome (hg19) using Minimap2 and sorted them with SAMtools (version 1.6).

Nanopore methylation calling
CpG methylation (5mC) calling was performed with Nanopolish v 0.11.0 (17) using the reads (FASTQ files), aligned reads (BAM files), and raw signals (FAST5 files) for each sample. We then calculated the methylation frequency and log-likelihood ratios of methylation at each position using “calculate_methylation_frequency.py” from the Nanopolish package. We filtered out any position with < 10 reads and log-likelihood ratios of < 2.5 in each sample.

Single nucleotide variant calling
SNVs were called over the targeted regions with Nanopolish using FASTQ files, BAM files, and FAST5 files. Nanopolish was used to reanalyze the raw signals after alignment and to calculate SNV allele frequencies from the ONT data at the signal level. The “nanopolish variants” subprogram was used to simultaneously call SNVs with a modified parameter setting: -min-candidate-frequency = 0.15, -min-candidate-depth = 10, --methylation-aware = cpg, --snps, and --ploidy = 2. We reviewed the variant quality of SNVs and visualized them with the Integrative Genomics Viewer and trackViewer [21, 22].

MGMT gene expression analysis with quantitative reverse transcriptase (qRT)-PCR
A total of 1 μg extracted RNA was reverse transcribed to cDNA using Superscript IV reverse transcriptase (Invitrogen, USA). qRT-PCR analysis was performed using iTaq Universal SYBR Green Supermix (BioRad, USA) and the StepOnePlus Real-Time PCR System (Applied Biosystems, USA). Real-time PCR was carried out in technical triplicates; it was run at 95 °C for 10 min, at 40 cycles of 95 °C for 15 s, and at 60 °C for 60 s. A published primer set was used for MGMT and the β-actin gene (ACTB) [23–25]. For data analysis, the average result in each triplicate was used.

Illumina sequencing of patient tumor samples
DNA and RNA sequencing was performed on clinical tumor specimens and saliva samples (from the same patients as the tumor specimens) for 6 of the 8 patients using the Tempus xT assay [26]. Briefly, nucleic acid was extracted from tumor tissue sections with tumor
cellularity greater than 20% using a Chemagic360 instrument and a source-specific magnetic bead protocol. Total nucleic acid was used for DNA library construction, while RNA was further purified by DNase I digestion and magnetic bead purification. The nucleic acid was quantified with a Quant-iT PicoGreen dsDNA Kit or Quant-iT Ribogreen RNA Kit (Life Technologies), and quality was confirmed with a LabChip GX Touch HT Genomic DNA Reagent Kit or LabChip RNA High HT Pico Sensitivity Reagent Kit (PerkinElmer).

For DNA library construction, 100 ng DNA from tumor or normal samples was mechanically sheared to an average size of 200 bp using a Covaris ultrasonicator. The libraries were prepared using the KAPA Hyper Prep Kit. Briefly, DNA underwent enzymatic end repair and A-tailing, followed by adapter ligation, bead-based size selection, and PCR. The captured DNA targets were amplified using the KAPA HiFi HotStart ReadyMix. The amplified target-captured libraries were sequenced on an Illumina HiSeq 4000 System with patterned flow cell technology.

Results
Nanopore sequencing accurately assesses mutational status and methylation levels
The error rate of raw nanopore sequencing reads continues to decrease, allowing the technology to be used for genotyping and methylation assays [17]. Nanopore sequencing errors are largely random and use of a consensus sequence from sufficient read depth can eliminate almost all of the sequencing error. To confirm the ability of nanopore sequencing to accurately genotype the IDH mutations, we sequenced PCR amplicons that were IDH1/2 wild type or IDH1/2 mutant using a nanopore MinION device. This test showed that heterozygous mutations in these 2 genes could be accurately detected, although artificial errors are inevitable (Fig. 1a).

To determine the limit of detection for CpG methylation, we sequenced 2 synthetic DNA standards with that were either 100% methylated or 0% methylated on CpGs and then used Nanopolish for methylation calling [17]. We generated data for 10, 25, 50%, or 75% methylated CpGs by randomly sampling the reads from the 0 and 100% methylated standards. We found that at a low sequencing coverage of ~10 reads (10X), methylation could be measured, but with high variation. Decreasing of coefficient of variation when increasing of sequencing depth was observed (Supplementary Table 1). At higher depth, ≥20X, the standard deviation was lower (Fig. 1b), and methylation levels of 0, 25, 50, 75, and 100% could be distinguished. Thus, 20X was used as the theoretical limit of detection in this study.

nCATS MGMT methylation assay is comparable to pyrosequencing assays
Based on these preliminary data, we then designed guide RNA for the nanopore Cas9-targeted sequencing (nCATS) workflow (18) to test on 4 human GBM cell lines (2 TMZ-sensitive [U87 and U251]) and 2 TMZ-resistant [T98G and LN18] and 8 clinical DG samples (4 IDH mutant and 4 IDH wild type) (Fig. 1c and Table 1). Sequencing depth coverage was an average of 184, 664, and 939 for MGMT, IDH1, and IDH2, respectively (Fig. 1d).

We then used nCATS to perform targeted sequencing of the MGMT gene; this approach captured 98 CpGs (located in promoter and exon 1) and 121 CpGs (in a 5’end of intron 1). The genomic coordinates of CpG loci are shown in Supplementary Table 2. The first 98 CpGs have been studied by others, and a subset of CpGs in this region has been used clinically to assess methylation [16]. Thus, we first focused on the 98 CpGs and used them to compare the methylation levels obtained by nCATS to levels obtained by pyrosequencing assays. Using a methylated and unmethylated DNA standard with >95% vs <5% methylation, respectively, nCATS provided a clear methylation pattern in both samples (Fig. 2a) that was comparable to the results of bisulfite modification-PCR-pyrosequencing for CpGs 1–25 and 70–84 (detail in Supplementary Tables 2 and 3).

We next applied nCATS to 4 well-characterized GBM cell lines (described above). The percent methylation of these 4 cell lines assayed by nCATS also correlated positively (r = 0.73, P = 6.9 × 10^-8 to r = 0.94, P = 2.2× 10^-16) with the percent methylation returned by pyrosequencing (Fig. 2b). At this point, we concluded that methylation data derived from nCATS is comparable to data derived from pyrosequencing assays when applied to a homogeneous sample (e.g. an immortalized glioma cell line).

Simultaneous evaluation of methylation and mutation biomarkers in patients with diffuse glioma
We next confirmed that nCATS can be used in clinical samples that have heterogenous cell populations opposed to the glioma cell lines. To test the accuracy of nCATS to assay MGMT methylation and IDH1/2 mutations in clinical samples. For MGMT methylation, we compared the nCATS data to data generated with bisulfite modification-PCR-pyrosequencing or the MassARRAY® System performed by 2 independent Clinical Laboratory Improvement Amendments (CLIA)-certified labs. There was a statistically significant positive correlation (r = 0.64, P = 1.04 × 10^-5 to r = 0.80, P = 4.39 × 10^-10) between nCATS quantitative methylation and pyrosequencing (Fig. 2c). MassARRAY® results were semiquantitative and only denoted methylation levels in 3 categories (not detected: <10%; low methylation: 10–
Fig. 1 (See legend on next page.)
Though, we observed high expression of MGMT found an unexpected result in the T98G cell line. Although, we observed high expression of MGMT as previous studies [27] but observed methylation level and gene expression were not opposed (Fig. 3a and b). This unexpected result led us to investigate the methylation of additional CpGs with nCATS (CpG 99–219). CpGs that had strong correlation (> 0.7) between MGMT expression and methylation were selected for by clustering analysis including 12 CpGs in the exon 1 and 34 CpGs in the intron 1. Hierarchical clustering according to CpG sites showed 2 clear position-dependent clusters: CpGs in exon 1 were clustered together and separated from CpGs in intron 1 (Fig. 3d). Hierarchical clustering of the 8 samples (4 cell lines and 4 tumors) demonstrated 2 distinct clusters: 2 TMZ-sensitive cell lines with similar methylation profiles were clustered together, while 2 TMZ-resistant cell lines and the 4 clinical samples were clustered together (Fig. 3d). Moreover, we found that intronic CpG methylation positively correlated with MGMT expression (r = 0.78, P = 0.024); whereas, exonic CpG methylation remained negatively correlated with MGMT expression (r = −0.77, P = 0.026) (Fig. 3e).

To test additional tumor grades, 4 tumor samples classified as primary WHO grade III or IV (high-grade gliomas) were assayed with qRT-PCR for MGMT expression and nCATS for methylation. These 4 samples differed from the previous clinical samples not only in tumor classification, but they came from IDH wild type patients. MGMT expression (Fig. 4a) and MGMT methylation pattern (Fig. 4b) varied between samples. The data for these 4 samples were combined with data for the 8 previous samples (including cell lines) for correlation analysis. With 12 samples, a negative correlation between MGMT expression and methylation in exon 1 was present (r = −0.51) but not statistically significant (P = 0.093). However, there was a statistically significant positive correlation for MGMT expression and methylation in intron 1 (r = 0.67, P = 0.016) (Fig. 4c). For IDH genotyping in these last four clinical samples, nCATS detected IDH1 and IDH2 as wild type, consistent with Illumina and Sanger sequencing results.

**nCATS identified single nucleotide variants**

Finally, we showed that nCATS could be used to identify single nucleotide variants (SNVs) in MGMT and IDH1/2
We compared nanopore sequencing with Illumina sequencing and also verified the absence of the pathogenic SNVs in germ-cell DNA using Illumina-sequenced saliva samples from 6 of the patients (no Illumina data available for P785 and P816). nCATS and Illumina returned similar genotypes for *MGMT* loci 1 and 2.
For locus 2, both methods detected heterozygous alleles (C/A) in both tumor and saliva from Patient 712. For locus 3, nCATS detected heterozygous alleles in all samples, while Illumina showed heterozygous alleles in only 1 sample. For loci 4, 5 (IDH1), and 6 (IDH2), nCATS and Illumina consistently detected somatic variants (the variants were not identified in saliva samples).

Discussion

In this study, we used nanopore Cas9-targeted long-read sequencing (nCATS) to simultaneously assess 2 prognostic molecular markers in diffuse glioma clinical samples and cell lines—MGMT methylation and IDH1/2 mutations. nCATS enables enrichment of genomic regions without amplification [18, 28], quantitative analysis of methylation on native DNA, and identification of single nucleotide variants. Gilpatrick et al. assessed clinical cancer biomarkers (e.g., TP53, KRAS, and BRAF) with nCATS in breast cancer cell lines and 1 patient tumor sample, demonstrating its feasibility [18]. Here, we demonstrated the feasibility of using nCATS on several clinical solid tumor samples to assess both genetic and epigenetic prognostic biomarkers that are clinically relevant.

nCATS allowed for simultaneous evaluation of IDH1/2 mutational status and MGMT methylation level in a streamlined workflow, resulting in biomarker assessment within 36 h (Fig. 1c). The ability of nanopore sequencing to evaluate methylation from native DNA sequences obviated the need for bisulfite modification [17], and we were able to achieve adequate depth coverage without amplification even in clinical samples. Our assessment of IDH mutational status correlated with clinically used Sanger methods and was further compared with Illumina sequencing (Fig. 4d).

MGMT methylation assessment is currently highly variable, as both the methodology used and the gene region evaluated are not consistent between clinicians. Further, no cutoff value in MGMT methylation level has been verified to correlate with MGMT expression; thus, no clinical consensus exists [16, 29]. Many institutions evaluate 2 differentially methylated regions (DMRs)
**Fig. 4** nCATS can simultaneously quantify MGMT CpG methylation and detect single nucleotide variants (SNVs) in glioma clinical samples. 

**a**, MGMT gene expression in 4 IDH wild type samples by qRT-PCR. Data are the mean ± SD (3 technical replicates). **b**, Methylation pattern by nCATS and MassARRAY. **c**, Correlation between MGMT expression and exon 1 methylation and between MGMT expression and intron 1 methylation. **d**, SNVs in MGMT and IDH1/2 were assayed with nCATS and Illumina sequencing in tumor and saliva samples from 6 patients. Data were plotted with trackViewer. No data were available for P785 and P816.
within the MGMT promoter and exon 1 that have been shown to correlate with MGMT expression in cell lines and patient cohorts [13, 30]; MGMT methylation is then used to predict responsiveness to temozolomide (TMZ) therapy. Our institution uses MassARRAY® and stratifies patients into 3 groups: no methylation (< 10%), low methylation (10–30%), and high methylation (> 30%). In this study, nCATS data from both cell lines and patient samples correlated with both MassARRAY® data and pyrosequencing (Fig. 2c and 4b). However, some patients who are below this arbitrary cutoff value (e.g., 10%) do respond to TMZ therapy [31–33], placing them in a “gray zone” and producing a clinical quandary. With this in mind, Chai et al. developed a novel CpG averaging model for pyrosequencing data that defines the MGMT promoter as being methylated when at least 3 CpGs exceed their respective cutoff values; this allows clinicians to better stratify patients with very low levels of methylation (e.g., < 10%) [34]. We demonstrate that nCATS can be used to quantify CpG methylation in multiple regions of the MGMT gene and may provide further insight into the variability of treatment responses. In the future, this long-read sequencing method could provide a reliable and thorough quantitative assessment of MGMT to develop a cutoff methylation value, but a large validation cohort will be needed.

Given the long-read sequencing capacity of nCATS, we were also able to quantify CpG methylation along the entire MGMT promoter, exon 1, and a portion of intron 1. One of the TMZ-resistant cell lines (T98G) did not have the expected inverse correlation between MGMT promoter methylation level and MGMT expression. There was a positive correlation between methylation of intronic CpG sites and MGMT expression for all GBM cell lines, the IDH mutant sample, and wild type DG samples (Fig. 3e and 4c). The result was in agreement with recent studies demonstrating the role of CpG methylation in the gene body (outside the promoter) in regulating levels of MGMT gene expression, with higher levels of gene body cytosine modification correlating with higher MGMT expression [27, 35]. This finding suggests a potential benefit of assaying gene body methylation, as the intron could be important for determining MGMT expression; however, a larger sample size is needed.

Finally, we identified 2 SNVs in the promoter region of MGMT, and one of them (rs1625649) had prognostic impact on patients with MGMT methylated glioblastoma [36, 37]. In MGMT, inconsistency between nCATS and Illumina result was also observed. In locus no.3 (Fig. 4d), nCATS detected 2 alleles in all patients while Illumina showed 2 alleles in only P568. We then considered the DNA sequence in this region and found 6 consecutive guanines (homopolymer) in this locus. For the current version of nanopore, homopolymer rich regions are the major source of errors. Therefore, for this locus, nCATS could not deliver accurate genotyping when using this version of nanopore (R9.4.1). An updated version of nanopore is being developed that incorporates a longer sensor to overcome errors in homopolymer rich regions.

Our nCATS technique also identified mutation variants (locus no.4–5 (Fig. 4)) in IDH1 and IDH2. The variants in IDH1 are associated with survival in patients with acute myeloid leukemia [38], but their prognostic value in GBM is not known [39]. However, with the advent of new IDH-directed therapies, variants in IDH1/2 may be of significance in the future [40]. These insights could lead to the incorporation of SNVs as an additional factor in therapeutic decision making, which can be done contemporaneously along with biomarker identification with nCATS.

In conclusion, the nCATS technique provides results within 2 days of surgical resection, potentially at lower capital cost than traditional methods. We demonstrated feasibility in clinical solid tumor samples and used DG as a model given that both genetic and epigenetic biomarkers are used clinically. The nCATS method also provided assessment of MGMT methylation throughout a larger gene region in comparison to currently used methods. There is great potential to use nCATS clinically to standardize molecular marker testing in DG and provide insights into patient variability to treatment response. Furthermore, nanopore platforms can be cost-effective and high-throughput, making them accessible in countries with limited resources [41]. In the future, we plan to design a comprehensive nCATS based DG diagnostic panel and testing on larger prospective cohorts. nCATs requires > 3 μg of high-quality DNA as starting material, making testing formalin-fixed specimens impractical. Obtaining tissue from fresh samples requires consideration of choosing a region with low necrosis and high tumor content in order to optimize DNA extraction. Nevertheless, the nCATS method provides a promising tool for enhancing cancer precision medicine with the potential for simultaneously assessing multiple molecular targets.

**Supplementary information**

Supplementary information accompanies this paper at https://doi.org/10.1186/s40478-020-00963-0.

**Additional file 1.**

**Acknowledgements**

The UAMS Tissue Biorepository and Procurement Service aided in obtaining patient samples.
Authors’ contributions  
T. W. and A.R. developed the concept and design. T.W. and P.J. developed the methodology. Data was acquired by T.W., D.A., A.D., J.D.D., and A.R. Data analysis and interpretation was done by T.W., P.J., Y.L., S.H., and A.R. Writing and revision of manuscript was completed by T.W., P.J., Y.L., S.H., and A.R. All authors reviewed the article. The author(s) read and approved the final manuscript.

Funding  
This work was funded in part by the University of Arkansas for Medical Sciences (UAMS) Department of Neurosurgery, College of Medicine (AR); the Arkansas Biosciences Institute (AR); a UAMS Winthrop P. Rockefeller Cancer Institute Seeds of Science Grant (AR); the Helen Adams & Arkansas Research Alliance Endowed Chair (T.W.); and the National Institute of General Medical Sciences of the National Institutes of Health award no. P20GM112503 (T.W.).

Availability of data and materials  
All raw data are available in BioProject (PRJNA549038).

Ethics approval and consent to participate  
This study conforms to the ethical guidelines for human research and the Health Insurance Portability and Accountability Act. This was a single institution, retrospective study which was approved by our Institutional Review Board (IRB # 228443) at the University of Arkansas for Medical Sciences. Tissue donation was performed after consent from the patient or a person with legal authority to grant permission. Samples used for research under BioProject ID: PRINAS49038 are blinded to personal identifying information.

Consent for publication  
Not applicable. All patient data is not identifiable.

Competing interests  
The authors declare that they have no conflicts of interest with respect to the contents of this manuscript.

Author details  
1. Department of Biomedical Informatics, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA. 2. Department of Neurosurgery, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA. 3. Department of Pharmacology and Toxicology, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA.

Received: 23 May 2020 Accepted: 11 June 2020

Published online: 20 June 2020

References  
1. Ostrom QT, Gittleman H, Truitt G, Boscia A, Kruchko C, Barnholtz-Sloan JS. (2018) CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015. Neuro Oncol 20:v1–v86 Available from: https://academic.oup.com/neo-oncology/article/20/suppl_4/v1/5090960. [cited 2018 Dec 27]
2. Li K, Lu D, Guo Y, Wang C, Liu X, Liu Y et al Trends and patterns of incidence of diffuse glioma in adults in the United States, 1973–2014. Cancer Med 7:5281–5290 Available from: http://www.ncbi.nlm.nih.gov/pubmed/30175510. [cited 2020 Feb 16]
3. Perry A, Wesseling P. (2016) Histologic classification of gliomas. Handb Clin Neurol 71:95–95 Available from: http://www.ncbi.nlm.nih.gov/pubmed/26948349. [cited 2020 Feb 16]
4. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol. 2016; 131:803–820. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27157931. [cited 2019 Sep 30]
5. Patterson J, Wongsurawat T, Rodriguez A. A Glioblastoma genomics primer for clinicians. Med Res Arch Knowledge Enterprise Journals; 2020;8:2034
6. Aldape K, Nejad R, Louis DN, Zadab G (2017) Integrating molecular markers into the World Health Organization classification of CNS tumors: a survey of the neuro-oncology community. Neuro Oncol 23:336–344 Available from: http://www.ncbi.nlm.nih.gov/pubmed/27688263. [cited 2020 Feb 16]
7. Santosh V, Sravaya P, Gupta T, Muzumdar D, Chacko G, Suri V et al (2019) ISNO consensus guidelines for practical adaptation of the WHO 2016 classification of adult diffuse gliomas. Neurol India 67:173–182 Available from: http://www.ncbi.nlm.nih.gov/pubmed/30860119. [cited 2020 Feb 16]
8. Wohrer A, Krestensson BW, Vital A, Hainfeld JR. (2017) Patterns of diagnostic marker assessment in adult diffuse glioma: a survey of the European Confederation of Neuropathological Societies (Euro-CNS). Clin Neuropathol 36:5–14 Available from: http://www.ncbi.nlm.nih.gov/pubmed/27966427. [cited 2020 Feb 16]
9. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W et al (2009) IDH1 and IDH2 Mutations in Gliomas. N Engl J Med 360:765–773 Available from: http://www.ncbi.nlm.nih.gov/pubmed/19228619. [cited 2019 Sep 29]
10. Weller M, van den Bent M, Tonn JC, Stupp R, Preusser M, Cohen-Jonathan-Moyal E et al (2017) European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendrogliomas. Lancet Oncol 18:e315–e329
11. Capper D, Weisert S, Salis J, Habel A, Meyer J, Jager D et al (2010) Characterization of R132H mutation-specific IDH1 antibody binding in brain tumors. Brain Pathol 20:245–254. https://doi.org/10.1111/j.1750-3639.2009.00352.x [cited 2019 Oct 12]
12. Bienkowski M, Berghoff AS, Marosi C, Wohrer A, Heinzl H, Hainfeld JR et al (2015) Clinical neuropathology practice guide 5-2015: MGMT methylation pyrosequencing in glioblastomas: unresolved issues and open questions. Clin Neuropathol 34:25–257
13. Malley DS, Hamoudi RA, Kociatkowski S, Pearson DM, Collins VP, Ichimura K (2011) A distinct region of the MGMT CpG island critical for transcriptional regulation is preferentially methylated in glioblastoma cells and xenografts. Acta Neuropathol 121:651–661 Available from: http://link.springer.com/10.1007/s00401-011-0803-5. [cited 2019 May 15]
14. Nalmanström A, Granberg BH, Marosi C, Stupp R, Frappaz D, Schultz H et al (2012) Temozolomide versus standard 6-week radiotherapy versus hypofractionated radiotherapy in patients older than 60 years with glioblastoma: the Nordic randomised, phase 3 trial. Lancet Oncol 13:916–926 Available from: https://linkinghub.elsevier.com/retrieve/pii/S1470204512702656. [cited 2018 Dec 26]
15. Wick W, Platten M, Meisner C, Felsberg J, Tabatabai G, Simon M et al (2012) Temozolomide chemotherapy alone versus radiotherapy alone for malignant astrocytoma in the elderly: the NOA-08 randomised, phase 3 trial. Lancet Oncol 13:707–715 Available from: https://linkinghub.elsevier.com/retrieve/pii/S147020451270164X. [cited 2018 Dec 26]
16. Mansouri A, Hachem LD, Mansouri S, Nassiri F, Lapereiere NJ, Xia D et al (2018) MGMT promoter methylation status testing to guide therapy for glioblastoma: refining the approach based on emerging evidence and current challenges. Neuro Oncol Available from: https://academic.oup.com/neo-oncology/article/doi/10.1093/neuonc/nyy309. [cited 2018 Dec 22]
17. Simpson JT, Workman RE, Zuzarte PC, David M, Duni LJ, Timp W (2017) Detecting DNA cytosine methylation using nanopore sequencing. Nat Methods 14:407
18. Gilpatrick T, Lee I, Graham JE, Raimondu E, Bowen R, Heron A et al (2020) Targeted nanopore sequencing with Cas9-guided adapter ligation. Nat Biotechnol 1–6 Available from: http://www.ncbi.nlm.nih.gov/pubmed/32042167. [cited 2020 Feb 16]
19. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E (2019) CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res 47:W171–W174 Available from: https://academic.oup.com/nar/article/47/W1/W171/5491735. [cited 2020 Feb 16]
20. De Coster W, D’Heret S, Schulz DT, Cruts M, Van Broeckhoven C (2018) NanoPack: visualizing and processing long-read sequencing data. Berger B, editor. Bioinformatics 34:2666–2669 Available from: https://academic.oup.com/bioinformatics/article/34/15/2666/4934095. [cited 2020 Feb 16]
21. Ou J, Zhu LI (2019) trackViewer: a Bioconductor package for interactive and integrative visualization of multi-omics data. Nat Methods 16:421
22. Robinson JT, Thovaldsdóttir H, Winckler W, Gutmann M, Lander ES, Getz G et al (2011) Integrative genomics viewer. Nat Biotechnol 29(1):24–26
23. Coutarel L, Kliz T, Cohen L, Shen SS, Costa M (2016) Molecular Mechanisms of Malignant Transformation by Low Dose Cadmium in Normal Human Bronchial Epithelial Cells. Shi X, editor. PLoS One 11:e0155002 Available from: http://dx.doi.org/10.1371/journal.pone.0155002. [cited 2020 Feb 16]
