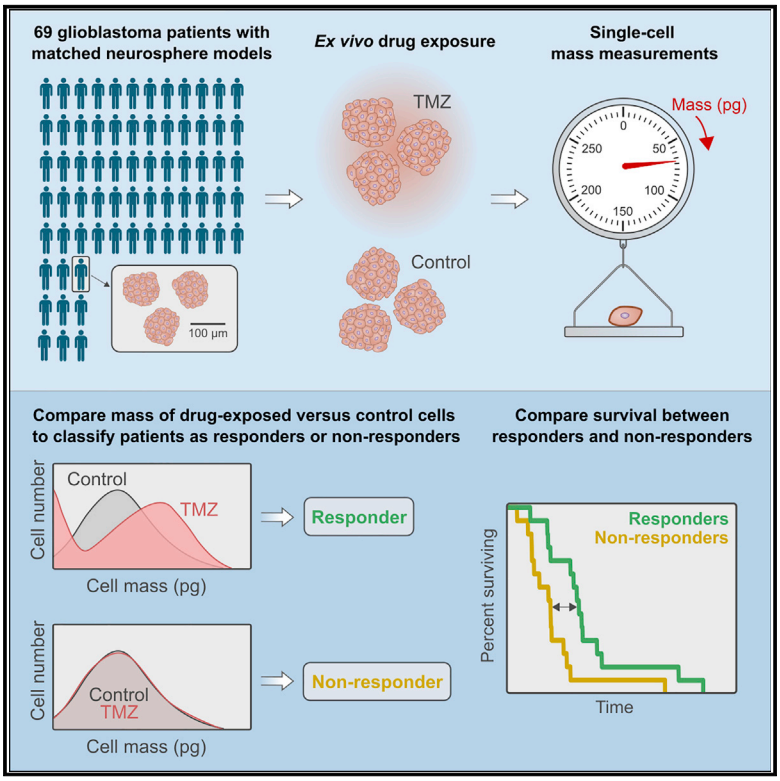
Cell Reports

Article

Functional drug susceptibility testing using single-cell mass predicts treatment outcome in patient-derived cancer neurosphere models

Graphical abstract



Highlights

- In a retrospective study, *ex vivo* drug testing predicts GBM patient survival
- Testing is based on detecting subtle changes in single-cell mass
- Predictive power of functional testing is comparable to MGMT promoter methylation
- Mass biomarker could be used in situations where genomic biomarkers are unavailable

Authors

Max A. Stockslager, Seth Malinowski, Mehdi Touat, ..., Kin-Hoe Chow, Keith L. Ligon, Scott R. Manalis

Correspondence

keith_ligon@dfci.harvard.edu (K.L.L.), srm@mit.edu (S.R.M.)

In brief

Stockslager et al. find in a retrospective study that functional drug susceptibility testing predicts the response of patients with glioblastoma to chemotherapy. By detecting subtle changes in tumor cell mass after *ex vivo* drug exposure, treatment response can be predicted with power comparable to the standard-of-care genomic biomarker.

Stockslager et al., 2021, Cell Reports *37*, 109788
October 5, 2021
https://doi.org/10.1016/j.celrep.2021.109788

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Functional drug susceptibility testing using single-cell mass predicts treatment outcome in patient-derived cancer neurosphere models

Max A. Stockslager,[1,2,11] Seth Malinowski,[3,11] Mehdi Touat,[3,4,5] Jennifer C. Yoon,[2] Jack Geduldig,[3] Mahnoor Mirza,[2] Annette S. Kim,[6] Patrick Y. Wen,[7] Kin-Hoe Chow,[8] Keith L. Ligon,[3,4,6,8,9,12,*] and Scott R. Manalis[1,2,4,10,12,13,*]

[1]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Koch Institute for Integrative Cancer Research, Cambridge, MA, USA
[3]Department of Oncologic Pathology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA
[4]Broad Institute of Harvard and MIT, Cambridge, MA, USA
[5]Sorbonne Université, Inserm, CNRS, UMR S 1127, Institut du Cerveau et de la Moelle épinière, ICM, AP-HP, Hôpitaux Universitaires La Pitié Salpêtrière - Charles Foix, Service de Neurologie 2-Mazarin, Paris, France
[6]Department of Pathology, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA
[7]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA USA
[8]Center for Patient-Derived Models, Dana-Farber Cancer Institute, Boston, MA, USA
[9]Department of Pathology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA
[10]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
[11]These authors contributed equally
[12]Senior author
[13]Lead contact
*Correspondence: keith_ligon@dfci.harvard.edu (K.L.L.), srm@mit.edu (S.R.M.)
https://doi.org/10.1016/j.celrep.2021.109788

SUMMARY

Functional precision medicine aims to match individual cancer patients to optimal treatment through *ex vivo* drug susceptibility testing on patient-derived cells. However, few functional diagnostic assays have been validated against patient outcomes at scale because of limitations of such assays. Here, we describe a high-throughput assay that detects subtle changes in the mass of individual drug-treated cancer cells as a surrogate biomarker for patient treatment response. To validate this approach, we determined *ex vivo* response to temozolomide in a retrospective cohort of 69 glioblastoma patient-derived neurosphere models with matched patient survival and genomics. Temozolomide-induced changes in cell mass distributions predict patient overall survival similarly to O^6-methylguanine-DNA methyltransferase (MGMT) promoter methylation and may aid in predictions in gliomas with mismatch-repair variants of unknown significance, where MGMT is not predictive. Our findings suggest cell mass is a promising functional biomarker for cancers and drugs that lack genomic biomarkers.

INTRODUCTION

Cancer precision medicine seeks to match each individual patient to the most effective available therapy. To date, precision medicine has largely been based on genomic profiling, in which patients with certain pre-defined genetic alterations are identified and matched with drugs targeting those specific abnormalities. Despite early success, progress in identifying additional actionable mutations has been slow, and today fewer than 20% of patients with metastatic cancer are eligible for U.S. Food and Drug Administration (FDA)-approved genome-guided drugs (Marquart et al., 2018). This slow progress has led some to speculate that most of the "low-hanging fruit" of actionable mutations have already been identified and that additional complementary approaches will be needed to continue making progress in improving outcomes for patients with cancer (Friedman et al., 2015).

As a complement to genomic precision medicine, functional precision medicine aims to match patients with cancer to effective therapies by performing *ex vivo* drug susceptibility testing directly on biopsied tumor cells. The vision of functional precision medicine is to predict drug susceptibility at the level of individual patients by sampling the tumor, exposing the tumor cells to candidate drugs *ex vivo*, and then measuring the cellular response using integrative readouts, such as cell growth, proliferation, or apoptotic signaling. Unlike genomic profiling, which typically focuses on identifying small subsets of patients harboring a few specific, well-understood targetable mutations, functional testing could potentially be used to test susceptibility to a broad range of drugs, regardless of the patient's specific disease or genomic background.

Despite continued efforts to develop functional assays for drug susceptibility testing in cancer (Letai, 2017), there are still

no such assays currently used in clinical practice. A key factor that has limited the translation of functional assays to the clinic is a general lack of studies that directly assess whether *ex vivo* drug susceptibility testing is correlated with measures of patient treatment outcome, such as overall survival duration on therapy. This lack of compelling evidence was the basis for a clinical practice guideline published by the American Society of Clinical Oncology, which concluded that, as of 2011, there was not sufficient evidence to support the use the use of functional drug susceptibility testing in clinical oncology practice; the group has not published a revised guideline since then (Burstein et al., 2011). Most studies linking functional testing to patient outcome have only obtained matched clinical outcomes for small numbers of patients (for example, n = 4–9 patients per study; Bhola et al., 2020; Cetin et al., 2017; Tiriac et al., 2018; Vlachogiannis et al., 2018), in part because of the logistical challenges involved in obtaining clinical follow-up.

To address this limitation, we sought to identify a tractable model system in which to conduct a retrospective study comparing functional drug susceptibility testing with patients' clinical outcomes. Glioblastoma (GBM) is one cancer in which such a large-scale retrospective study is possible: using three-dimensional (3D) tumor neurosphere culture techniques, "patient-derived neurosphere models" can be established from primary tumor resections, and stored long-term as viable cells. Patient-derived neurosphere models have been robustly validated to preserve key phenotypic and genotypic features of the patient's tumor *in vitro*. Further, because patients with GBM typically receive only a single drug (temozolomide) for the course of their treatment, overall survival duration provides a meaningful readout of each patient's susceptibility to temozolomide against which to compare *ex vivo* drug susceptibility testing.

In this study, we performed functional drug susceptibility testing on a retrospective cohort of 69 genomically characterized patient-derived neurosphere models (generated by the Dana-Farber Cancer Institute Center for Patient Derived Models) with matched patient treatment history information, including overall survival duration. Although primary tumor samples or more complex patient-derived organoid models (Jacob et al., 2020) may ultimately be the most predictive and convenient tumor model for clinical implementation of functional testing, the immediate availability of banked patient-derived neurosphere models, which are similar to organoids in their 3D growth, with matched patient treatment history information took precedence in this study.

We measured the response of these GBM patient-derived neurosphere models to temozolomide (TMZ), an alkylating chemotherapy agent that serves as the standard-of-care treatment for GBM and, then, asked whether functional TMZ susceptibility testing correlated with the duration that patients survived after standard-of-care treatment that included TMZ. This study serves as a useful intermediate step between initial validation experiments (using immortalized cell lines with little clinical relevance) and a full-scale prospective clinical trial (using fresh patient tissue, but with additional technical and logistical challenges), allowing us to evaluate whether a particular functional assay predicts treatment outcome in a more tractable model system.

RESULTS

Single-cell mass measurements for functional drug susceptibility testing

A variety of functional assays have been used as readouts for *ex vivo* drug susceptibility testing. The ideal assay would work for drugs with a range of cytotoxic or cytostatic mechanisms; would be amenable to acute testing, without requiring extended culture; and would achieve high sample throughput. Previously, our laboratory developed the mass accumulation rate (MAR) assay, which enabled acute testing of both cytotoxic and cytostatic drugs (Cermak et al., 2016; Stevens et al., 2016). In that assay, a microfluidic cell mass sensor, called a suspended microchannel resonator (SMR), measures the instantaneous growth rate of individual drug-treated tumor cells by weighing each cell repeatedly over a period of 10–15 min. By measuring drug-induced changes in single-cell growth rates, the MAR assay has been shown to predict drug susceptibility in a variety of cancer model systems, including clinical response in a small cohort of patients with multiple myeloma who were treated with a variety of standard-of-care agents (Cetin et al., 2017). However, despite promising early results, clinical implementation of the MAR assay has been challenging because of the assay's low throughput; typically, at least 1–2 h are required to measure enough cells for each drug condition, limiting the amount of testing that can be performed in the limited window during which a primary tumor sample remains viable.

Here, we developed a functional assay that retains the key advantages of the MAR assay (i.e., compatibility with a wide variety of drug mechanisms and not requiring extended culture) and achieves the higher sample throughput needed for clinical implementation (Figure 1A). Our approach, the "SMR mass assay," is based on detecting subtle changes in the mass distributions of drug-exposed versus control cells (Figure 1B). We first used the SMR mass sensor to measure the buoyant mass (referred to hereafter simply as "mass") of single cells in baseline growth conditions by detecting a shift in the resonance frequency of a hollow micro-cantilever beam as cells flow through it. The details of the SMR system have been described thoroughly in previous work from our group (Burg et al., 2007; Cermak et al., 2016; Son et al., 2012; Stockslager et al., 2019) and are highlighted in the Method details section. Using the SMR, we weighed individual cells with precision near 50 fg, which is on the order of 0.1% of the buoyant mass of a typical tumor cell. To achieve higher throughput, we performed these single-cell mass measurements using our recently described parallel SMR array (Figure 1C), a microfluidic device containing 16 SMRs connected fluidically in parallel and operated simultaneously on the same microfluidic chip (Stockslager et al., 2019). This enabled us to weigh a population of cells with maximum throughput of thousands of cells per minute. By sequentially measuring the mass of ~2,000 cells from each sample, then comparing the mass distributions of drug-exposed versus control cells, we can detect changes in mean cell mass as small as 3% with high statistical power and confidence and use as little as 2 min of instrument time per sample. This approach represents an ~30-fold improvement in the rate at which samples can be measured compared with that of the

CellPress
OPEN ACCESS

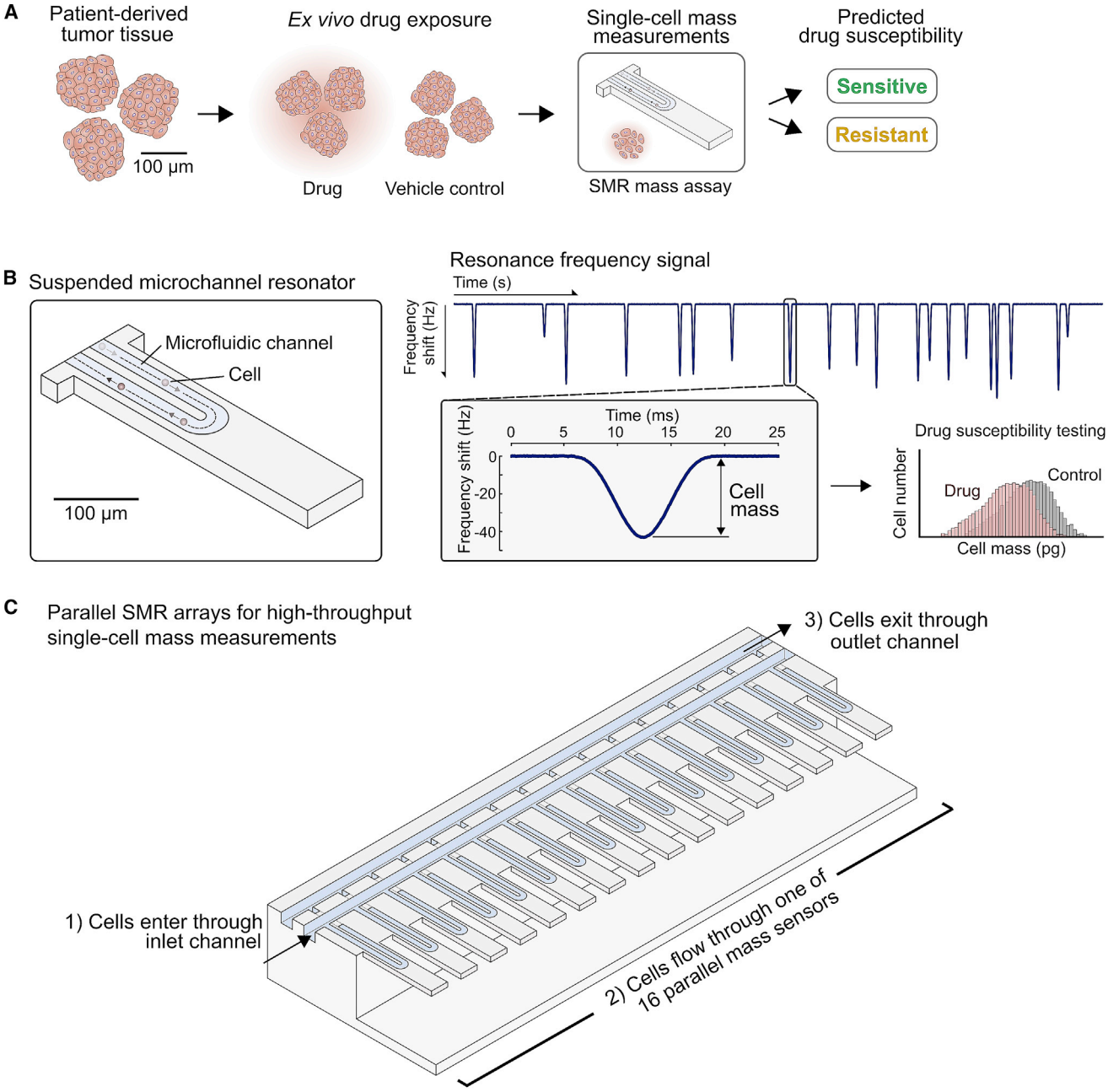


Figure 1. Single-cell mass measurements for functional drug susceptibility testing

(A) General workflow for drug susceptibility testing. Patient-derived tumor tissue is isolated, dissociated to form patient-derived neurosphere models, exposed *ex vivo* to candidate drugs, and then, dissociated again to single cells for mass measurements.

(B) Schematic of a single suspended microchannel resonator (SMR), a microfluidic sensor that weighs single cells as they flow through a resonating micro-cantilever beam. Cell mass is measured by detecting a shift in the cantilever's resonance frequency as the cell passes through. Cell mass can be used as a readout for *ex vivo* drug susceptibility testing, by comparing the mass of drug-exposed tumor cells to untreated controls.

(C) Using the parallel SMR array, single-cell mass distributions can be measured at high throughput because cells are simultaneously weighed by multiple SMR sensors in parallel on the same microfluidic chip.

MAR assay, in which each cell must be weighed repeatedly to measure its growth rate.

Because (unlike the MAR assay) the SMR mass assay has not previously been applied to *ex vivo* drug susceptibility testing, we first used conventional cancer cell lines as a model system to validate that the single-cell mass measurements were consistent with expected patterns of drug susceptibility and resistance. We used two model systems: (1) two BCR-ABL^{+} leukemia cell lines

treated with the BCR-ABL inhibitors imatinib and ponatinib, and (2) two EGFR-mutant lung adenocarcinoma cell lines treated with the EGFR inhibitors gefitinib and osimertinib (Method details). For both model systems, the SMR mass assay was consistent with expected susceptibility versus the resistance of each cell line to each drug (Figure S1).

Single-cell mass measurements for predicting temozolomide susceptibility in GBM

After validating the SMR mass assay for predicting drug susceptibility in conventional cancer cell lines, we next asked whether single-cell mass measurements could predict the clinical response of patients with GBM to TMZ. To benchmark the performance of the SMR mass assay, we also measured TMZ response using the Cell Titer-Glo (CTG) assay, which measures ATP levels as a proxy for numbers of viable cells, because of the assay's frequent use in *ex vivo* drug susceptibility testing (Kulesz-Martin et al., 2013; Tiriac et al., 2018; Vlachogiannis et al., 2018).

We performed functional testing as follows (Figure 2A; Method details): GBM patient-derived neurospheres were seeded in culture, exposed to either 20 μM TMZ or a vehicle control, and then, functional readouts were taken at fixed time points (3, 5, 7, 10, 12, and 14 days of TMZ exposure) with feeding at regular intervals over the course of the drug exposure. Neurospheres remained intact for the duration of drug treatment but were dissociated to single cells immediately before mass measurements using the SMR.

We observed a wide range of biophysical TMZ responses in the patient-derived neurosphere models (Figure 2B). Interestingly, in many TMZ-responsive models, the SMR mass assay measured an increase in mean cell mass after TMZ exposure, consistent with the fact that TMZ arrests susceptible cells at the G2/M checkpoint (Hirose et al., 2001), where individual cells typically have double their initial mass just before division. This accumulation of larger cells was often accompanied by an accumulation of small particles (<25 pg), likely because of the accumulation of dead cells and debris in the culture, especially at later time points. In some models, the accumulation of small particles and debris dominated, and there was not an obvious accumulation of larger cells. Other patient-derived models (e.g., BT330) showed a striking lack of TMZ response (Figure 2B, right).

As expected, the CTG assay typically measured a reduction in total ATP levels relative to controls (Figure 2C, left) in the TMZ-responsive models and less reduction in TMZ-non-responsive models (Figure 2C, right).

To quantify the TMZ responsiveness of each model, we defined a "response score" for each functional assay, a single statistic describing the extent to which each functional readout changes when the cells are exposed to TMZ (Figures 2B and 2C, right; Method details). For both the SMR mass assay and the CTG assay, a greater response score indicated a larger response to the drug exposure. The SMR mass response score is defined as the Hellinger distance (Kitsos and Toulias, 2017) between the vehicle and TMZ-treated cell mass distributions, averaged across three time points of drug exposure (days 5, 7, and 10). The Hellinger distance is a statistic that measures the degree of difference between two mass distributions, regardless of how the mass distributions might differ. This statistic has the advantage of capturing multiple factors that are reflected as changes to the drug-exposed cell mass distribution, such as the accumulation of larger cells due tolate-cell-cycle arrest and the accumulation of small particles due to cytotoxicity.

The CTG response score is defined as the average of the CTG luminescence signal (normalized to the control at each time point) across all time points of drug exposure; larger CTG response scores correspond to greater reduction in ATP levels in drug-exposed samples relative to controls. In future steps toward clinical implementation, it will be important to understand the biological mechanism by which a drug of interest alters cell phenotypes and how those phenotypic changes manifest in the assay results. However, in this study, we chose to use these response scores as a surrogate for the degree of drug susceptibility due to our limited knowledge of the biological mechanism by which temozolomide induces these changes to cell mass (the SMR assay) or ATP levels (the CTG assay).

Of the 69 patient-derived neurosphere models on which any functional testing was performed (Table S1), data were successfully collected for 67/69 models using the SMR mass assay (full dataset shown in Figure 3) and 55/69 models using the CTG assay (full dataset shown in Figure S2).

Single-cell mass biomarker is consistent with known molecular biomarkers for TMZ susceptibility

To assess whether functional testing predicts patient response to TMZ, we first asked whether the SMR mass assay and the CTG assay were consistent with known molecular biomarkers for TMZ susceptibility. The best-known molecular biomarker for TMZ susceptibility is methylation of the O^6-methylguanine-DNA methyltransferase (MGMT) promoter (Touat et al., 2020). A second relevant biomarker is mismatch-repair (MMR) deficiency: even in MGMT-methylated tumor cells, which are normally expected to be more responsive to TMZ, MMR deficiency can lead to profound TMZ resistance because those cells continue to cycle and accumulate a large drug-induced mutational burden (Touat et al., 2020). To isolate MGMT promoter methylation as a response predictor, we initially limited our analysis to the subset of 64/69 models that had no MMR alterations (Figure 4A).

We compared functional response scores between MGMT-methylated and MGMT-unmethylated models and found that, for both the SMR and CTG assays, the MGMT-unmethylated models had significantly lower functional TMZ response scores than MGMT-methylated models (Figures 4B and 4C). To compare the degree to which each functional biomarker differed between MGMT-methylated and MGMT-unmethylated models, we computed the receiver operator characteristic area-under-the-curve statistic (ROC AUC); the values were comparable for both functional assays (ROC AUC values: SMR mass assay, 0.75; CTG assay, 0.81). Across all 69 patient-derived models, the SMR mass-response score and the CTG response score were moderately positively correlated (Spearman $\rho = 0.71$; Figure 4D).

For robustness, for 58 of the 68 patient-derived neurosphere models in our cohort with known model MGMT status (mMGMT), we also independently measured or collected from clinical records the MGMT promoter methylation status of the primary

CellPress
OPEN ACCESS

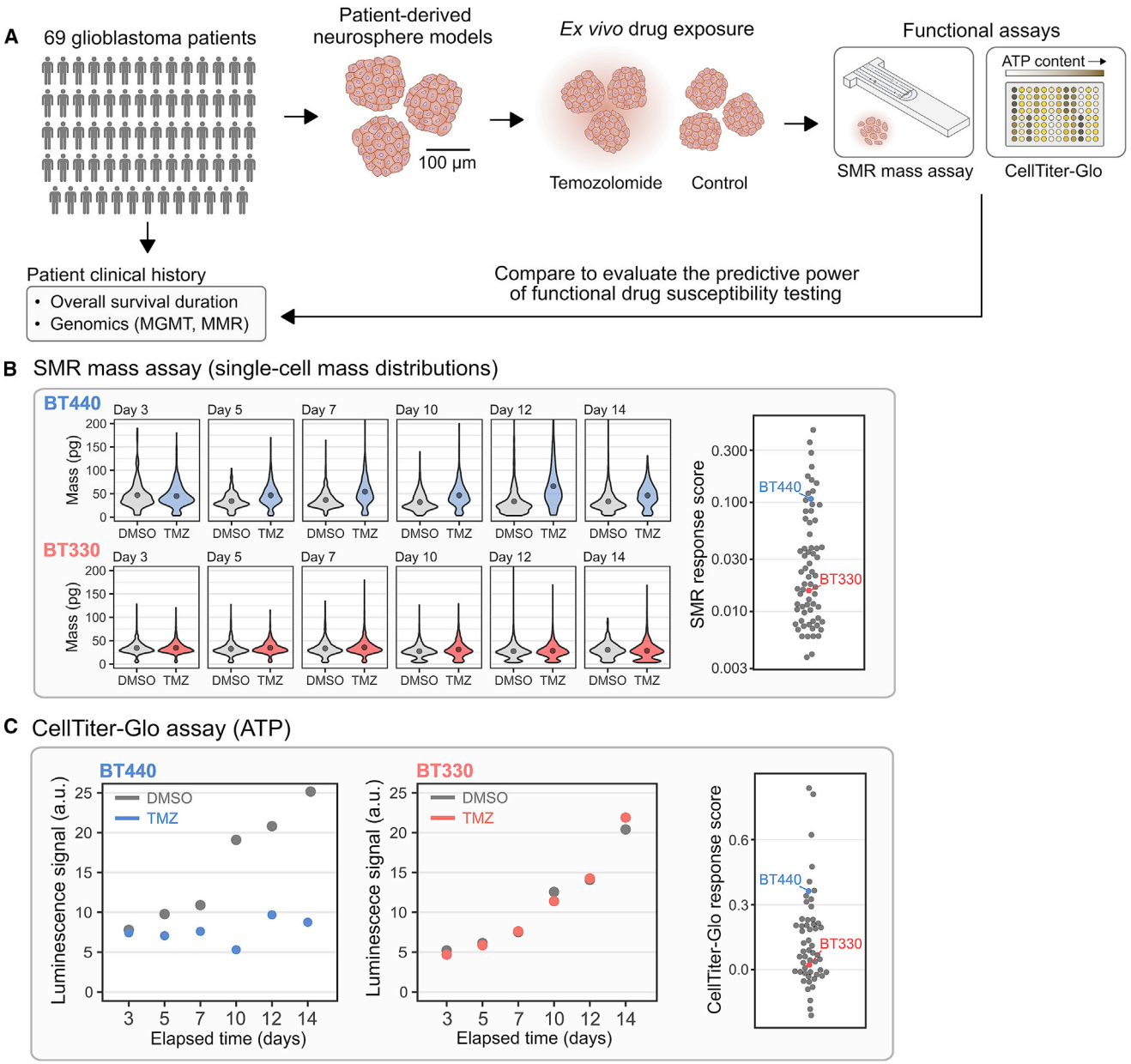


Figure 2. Profiling TMZ responsiveness in glioblastoma patient-derived neurosphere models

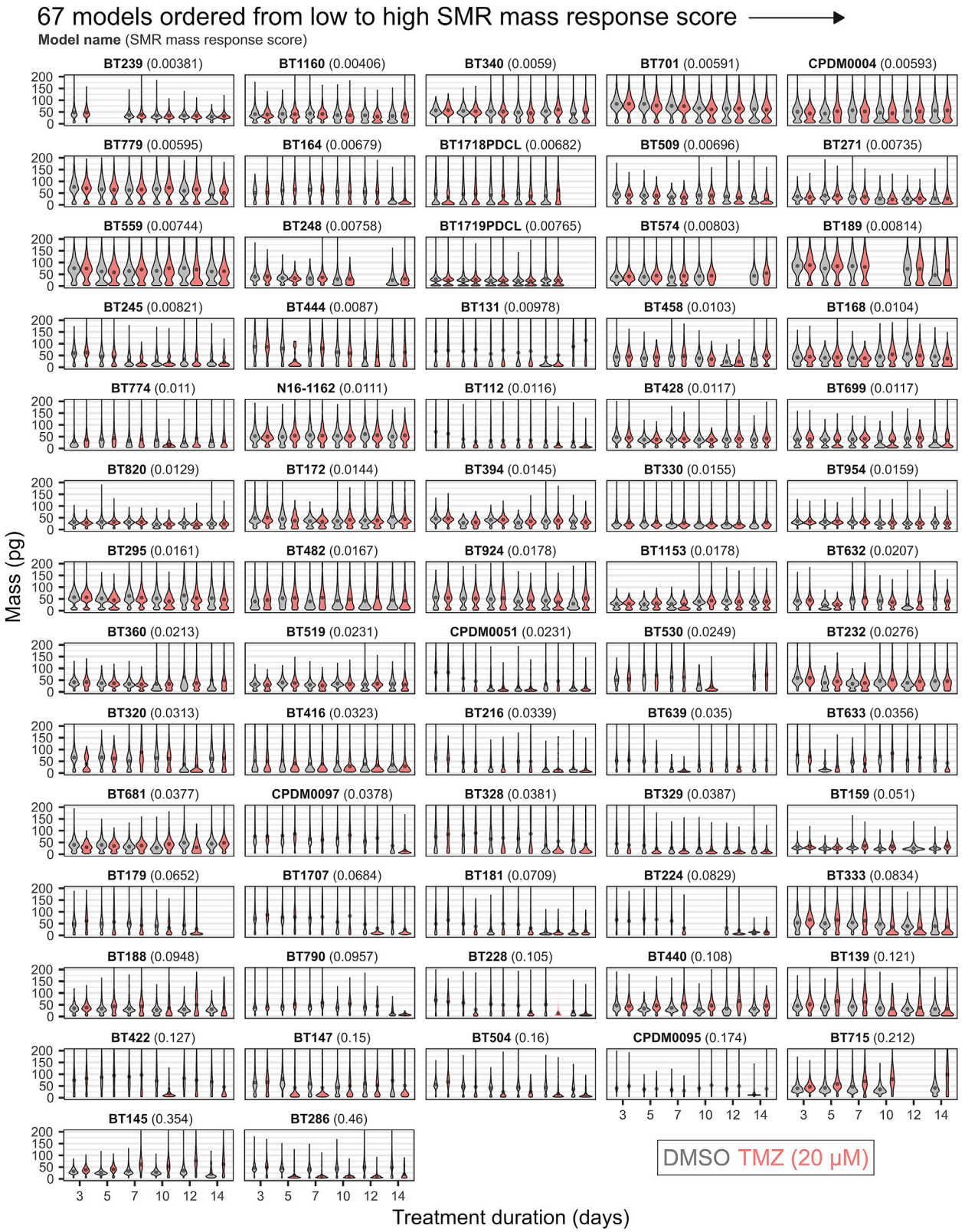
(A) Workflow for performing functional drug susceptibility testing in glioblastoma (GBM). Patient-derived neurosphere models are established from primary tumors, exposed to temozolomide (TMZ) or a vehicle control, then single-cell mass measurements (or for comparison, the CellTiter-Glo metabolic assay) are used to measure the TMZ response at multiple time points of drug exposure. Biophysical assay results are compared against MGMT promoter methylation status and overall survival duration.

(B) For the SMR mass assay, TMZ-responsive models generally increase in mean cell mass over time compared with that of controls.

(C) For the CellTiter-Glo (CTG) assay, TMZ-responsive models have generally reduced CTG luminescence signals compared with controls.

patient tumor sample (pMGMT). Both mMGMT and pMGMT had similar correlations with functional response scores (Method details).

Functional testing would be particularly useful for patients for whom existing genomic biomarkers are non-predictive. In GBM, one such group is the subset of patients with recently described alternative mechanisms of TMZ resistance not mediated by MGMT promoter methylation. For example, MGMT promoter methylation may not reliably predict TMZ susceptibility in MMR-mutated patients, who comprise 5/69 patients in our cohort and a similar fraction of patients with GBM overall (McLendon et al., 2008). Deficient function of the MMR pathway results in TMZ resistance, even in MGMT-methylated patients; therefore, patients with MMR-inactivating mutations would be



(legend on next page)

expected to be resistant to TMZ. Consistently, the two models in our cohort with known MMR-inactivating mutations (N16-1162 and BT1160) were both functionally non-responsive to TMZ, despite one (N16-1162) having the methylated MGMT promoter (Figure S3). However, for other patients—such as those with MMR missense mutations—it remains unclear whether each patient's particular mutation is expected to result in loss of MMR functionality and, therefore, lead to TMZ resistance. Interestingly, we found that, of the three patients in our cohort with MMR missense mutations of unknown clinical significance, two were functionally non-responsive to TMZ (BT559 and BT168), whereas only one was functionally responsive (BT422), despite all three models having the methylated MGMT promoter (Figure S3). The functionally non-responsive patient with known survival duration had shorter-than-average overall survival (1.5 months), whereas the functionally responsive patient had longer-than-average overall survival (20.4 months), although a much larger cohort of patients with mutant MMR would be needed to determine whether that trend is supported. These observations suggest that functional testing has the potential to predict treatment outcome independent of genomics, which would be useful for subsets of patients for whom genomic biomarkers are currently non-predictive.

Single-cell mass measurements predict overall survival in patients newly diagnosed with GBM

Next, we asked whether single-cell mass measurements could retrospectively predict the overall survival of patients with GBM who were treated with standard-of-care therapy that included TMZ. To limit potentially confounding effects, we limited our analysis to the 30 patient-derived models derived from patients who, at the time, were newly diagnosed, had wild-type isocitrate dehydrogenase (IDH) status, were subsequently treated with TMZ, and had known overall survival. The median overall survival among this group was 14.6 months.

We used the ROC to assess the sensitivity and specificity with which the SMR and CTG functional biomarkers predicted overall survival of 15 months or greater (Figure 5A). For each functional biomarker, we computed the ROC AUC as a measure of the biomarker's predictive power. The SMR assay was moderately predictive of 15-month survival (ROC AUC statistic, 0.78), whereas the CTG assay was somewhat less predictive of 15-month survival (ROC AUC statistic, 0.66). For robustness, we confirmed that the predictive power was similar when using other binary survival outcomes, such as 12-month, 18-month, 21-month, and 24-month survival (Figure S4A).

Next, we compared overall survival distributions between TMZ-responsive and non-responsive patients. Ideally, the predictive power would be evaluated by dividing the cohort into a training set (to set a response threshold to classify patients as responders versus non-responders) and a test set (to compare survival between responders and non-responders). However, because of the limited size of our cohort, we instead chose to classify the 50% of patients with the highest functional-response scores as TMZ responders (Figure 5B). This approach is reasonable because previously published imaging results suggest that approximately one-half of patients with glioma respond to TMZ (Lee, 2016; Ollier et al., 2017). Although the exact fraction of TMZ-responsive patients in GBM is debated, we noted that our results were robust when different fractions of patients were classified as TMZ responders (in the range of 25%–75% responders, which generally includes the clinically estimated range of responses to TMZ; Figure S4B).

Overall survival was significantly longer for SMR mass assay responders than it was for SMR mass assay non-responders (Figure 5C; median overall survival, 20.1 versus 12.2 months, respectively; log-rank $p = 0.01$). Although CTG responders had longer median survival than CTG non-responders (18.8 versus 10.3 months, respectively), there was not a statistically significant difference in overall survival distributions between those two groups (Figure 5C; log-rank $p = 0.11$). The predictive power of the SMR mass assay was comparable to that of the patient-derived mMGMT status (Figure 5C; log-rank $p = 0.001$; median survival, 18.8 months versus 8.2 months for patients with methylated versus unmethylated mMGMT, respectively). A detailed statistical comparison is included in Figures S5A and S5B.

To confirm that the SMR and mMGMT biomarkers did not reach statistical significance due only to their increased sample size relative to the CTG (28 eligible patients had SMR data and 29 patients had known mMGMT status, whereas only 25 patients had CTG data), we repeated this analysis limited to the subset of 25 patients with CTG data and found similar results (Method details).

We asked whether the SMR mass assay could be combined with the MGMT promoter methylation biomarker to make more accurate predictions of patient survival outcomes. By combining functional biomarkers with MGMT, one can achieve either greater sensitivity or specificity for identifying TMZ-resistant patients (Figures S5C–S5E).

DISCUSSION

This retrospective study demonstrates that single-cell mass measurements can be used as a functional readout for *ex vivo* drug susceptibility testing and that, in particular, the SMR mass assay can be used in individual patients with GBM to predict overall survival duration after TMZ treatment, with predictive power comparable to the existing standard molecular biomarker. Although the SMR mass assay and MGMT appear to have similar predictive capability in GBM, the SMR mass assay has the potential to serve as a complementary and orthogonal biomarker for providing an integrative readout of cell

Figure 3. Single-cell mass measurements for profiling the TMZ responsiveness of 67 GBM patient-derived neurosphere models

Models are ranked in order of low to high SMR mass-response score, which is based on the Hellinger distance between the mass distributions of DMSO and TMZ-treated cell populations at each time point. Points overlaid on the histograms indicate the mean mass for each drug condition and time point. The horizontal coordinate represents density (dimensions pg^{-1}), with the same horizontal scale for all distributions for each model. Some histograms appear narrow because at least one sample for that model accumulated many particles of the same size, compressing the horizontal (density) scale; this is common for TMZ-responsive models (bottom) because of the accumulation of low-mass particles (<25 pg) in the culture as a result of drug-induced cell death.

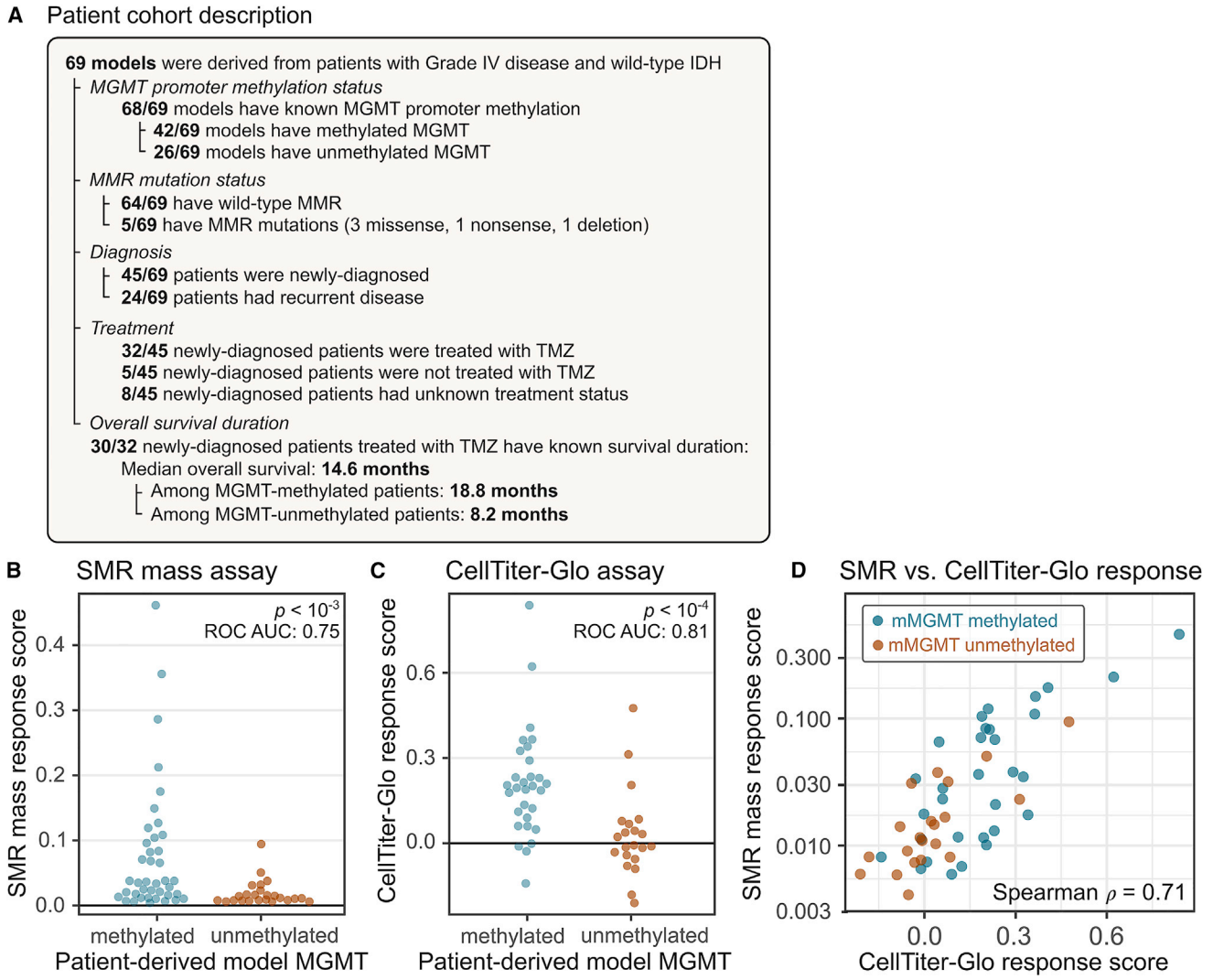


Figure 4. Functional drug susceptibility testing is consistent with MGMT methylation for predicting TMZ susceptibility

(A) Clinicopathologic characteristics of the patient cohort from which the patient-derived models were established.

(B and C) MGMT-methylated patient-derived models have significantly higher functional-response scores compared with that of the MGMT-unmethylated models, for both (B) the SMR mass assay (Wilcoxon rank-sum, $p < 0.001$; ROC AUC, 0.75) and (C) the CellTiter-Glo assay (Wilcoxon rank-sum, $p < 0.0001$; ROC AUC, 0.81).

(D) Correlation between SMR mass response score and CellTiter-Glo response score (shown for models with known MGMT status).

response in a rapid and cost-effective manner. In the future, such a biomarker could be useful to predict drug susceptibility in diseases and patient subsets for which predictive genomic or clinical biomarkers are not available. However, to establish that an assay is predictive, one needs to first compare its predictions against a validated indicator of a patient's drug susceptibility or resistance, and GBM is a useful model system for such a study because of the presence of MGMT as well-validated molecular biomarker and the availability of banked patient-derived neurosphere models with matched clinical data. The close agreement among functional testing, MGMT, and overall survival provides a foundation of evidence that functional testing predicts patient outcome, which will need further development through prospective studies in other cancers.

As functional precision medicine approaches move toward clinical implementation, key practical considerations are emerging. First, is the throughput of the assay. In this study, it took 4 months for two researchers to perform functional testing on 69 patient-derived neurosphere models. The main rate-limiting factor in our study was the time required to culture the models. However, the next rate-limiting factor would have been the throughput of profiling the models using the SMR mass assay. The on-device data collection time was approximately 20 min per SMR sample; therefore, measuring a total of 12 conditions per patient model (two drug conditions and six time points), required a total of 3 h of SMR instrument time per patient model. This level of throughput is likely compatible with current clinical pathology workflows. Further, we estimate that

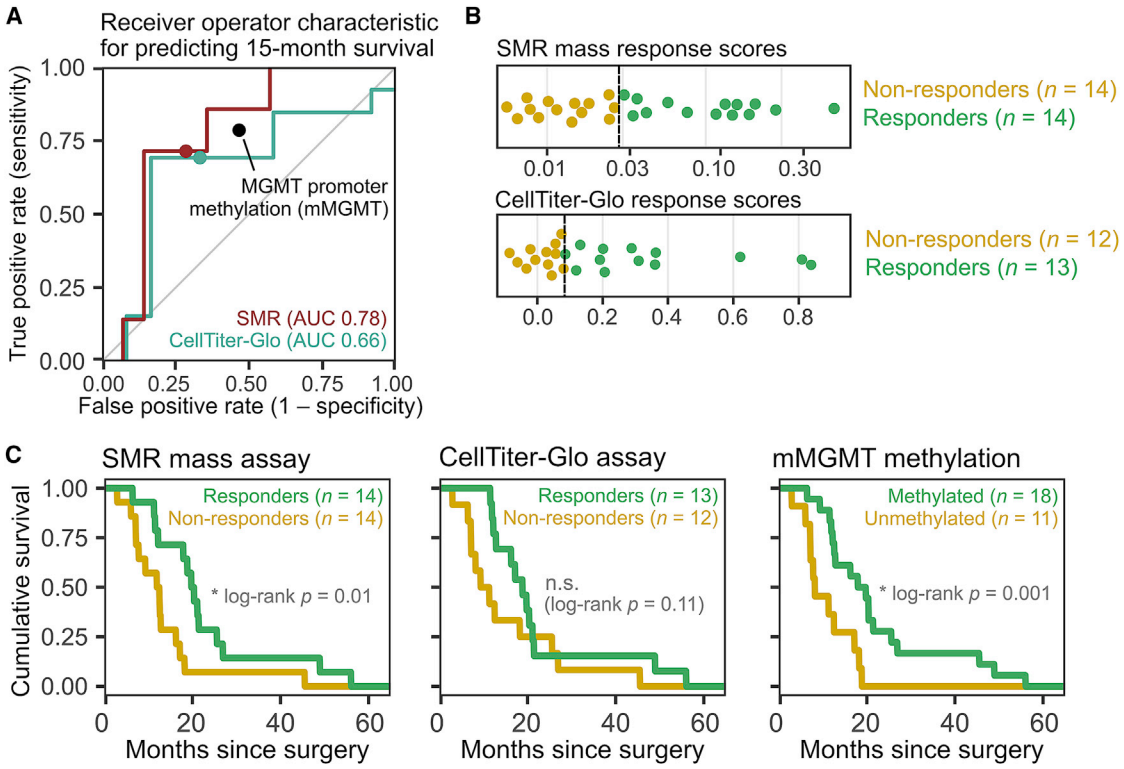


Figure 5. Single-cell mass measurements predict overall survival duration

(A) Receiver operator characteristic (ROC) between continuous functional biomarkers (the SMR mass response score and CellTiter-Glo response score) and a binary survival outcome (15-month survival). Points indicate the performance of the thresholds used for the classification in (B) and (C).

(B) For each assay, patients within the top 50% of response scores were labeled as functional "responders," and the bottom 50% were labeled as "non-responders."

(C) Overall survival distributions of functional responders versus non-responders. For the SMR mass assay, TMZ responders survived significantly longer on therapy than TMZ non-responders did (median overall survival, 20.1 versus 12.2 months, respectively; log-rank p = 0.01). Median survival was longer for CellTiter-Glo assay responders than it was for non-responders (median overall survival, 18.8 versus 10.3 months, respectively), but overall survival distributions were not significantly different between these groups (log-rank p = 0.11). The predictive power of the SMR mass assay was comparable to the predictive power of mMGMT promoter methylation (median overall survival, 18.8 versus 8.2 months for mMGMT-methylated versus mMGMT-unmethylated patients, respectively; log-rank p = 0.001).

the required time per sample could be significantly reduced: as we have shown previously (Stockslager et al., 2019), using more concentrated cell samples, the SMR can achieve mass-measurement throughput reaching thousands of cells per minute, potentially reducing the required SMR instrument time per sample by several fold.

A second practical consideration is sample consumption, i.e., the number of tumor cells required to perform functional testing. Depending on the cancer, a typical solid-tumor biopsy or resection might contain several hundred thousand tumor cells, setting an upper bound on the number of drug conditions or replicates on which testing can be performed. In our study, because the patient-derived neurosphere models could be propagated *ex vivo* to obtain more sample material, tumor cells were abundant, and so we did not fully optimize the study design to use cells efficiently. Despite that, across all time points, our assays consumed a total of only 120,000 cells per patient for each drug tested. These sample requirements are comparable to next-generation sequencing, which is already a standard component of many clinical workflows and requires as much as 1 μg of DNA from typically 100,000 tumor cells (Bäumer et al., 2018). However, the sample input requirements of the SMR mass assay could be reduced by several fold because sufficient statistical power can be obtained by weighing as few as 2,000 cells per sample, which is considerably less than our current protocol (Method details).

A third practical consideration is the duration for which cells must be exposed to drugs *ex vivo* before a response can be detected. Although evaluating this metric was not a focus when designing this study, the SMR mass assay can detect *ex vivo* drug response faster than traditional indicators of cell viability. For example, in this work, we showed that the SMR mass assay can detect the response of the leukemia cell line Ba/F3 BCR-ABL to imatinib in as little as 8 h (Figure S1), whereas, in previous work, we found that the traditional cell-viability markers DAPI and annexin V did not detect the same response even after 24 h of drug exposure (Stevens et al., 2016). Although this difference in response speed was not important for performing drug susceptibility testing on the GBM patient-derived neurosphere models, which can propagate indefinitely *ex vivo* without

loss of viability, it will be a more important consideration when moving to primary patient samples with a limited window of viability and when drugs that act more quickly than temozolomide are tested.

A frequent concern with functional drug susceptibility testing approaches is that *ex vivo* testing only measures tumor-cell-intrinsic drug susceptibility and does not capture other factors that determine a patient's response to therapy and overall survival duration. In particular, our study did not assess the effects of radiation therapy, which most patients with GBM receive concurrently with TMZ (Martínez-Garcia et al., 2018) and which is known to provide additional survival benefit. In this work, we chose to focus specifically on assessing tumor-cell-intrinsic drug susceptibility to limit the complexity of the study. However, in future work, it will be of great interest to evaluate whether functional testing could also be used to predict the degree of response to radiation or combination therapies in a more quantitative manner than existing predictors.

We found that the SMR mass assay was consistent with the MGMT biomarker and that SMR responders survived for significantly longer on therapy than non-responders did. It is interesting to note that, although the CTG assay was also consistent with the MGMT biomarker and CTG responders had longer median survival than non-responders, in our hands, there was not a statistically significant difference in overall survival distributions between CTG responders versus non-responders. However, this finding should not be viewed as an indictment of the CTG assay as a whole and should be viewed in light of the large body of existing work in which the assay has successfully been used for drug susceptibility testing. As others have described previously (Niepel et al., 2019), even simple experiments, such as measuring the response of cells to *in vitro* drug exposure, can be affected by factors such as cell-counting protocols, micro-titer plate selection, and cell-seeding density; all of which result in significant inter-center variability. We ran the CTG assay using one particular assay format (i.e., plate type, cell-seeding density, feeding schedule, drug dose, and time-point selection), and although our selection was a reasonable choice, it is entirely possible that, under different conditions, the assay could be more predictive. Further studies might reveal biological underpinnings that explain the differences in drug responsiveness using biophysical versus metabolic assay readouts.

An interesting finding of this work is that both functional assays identified a continuous spectrum of TMZ responsiveness, rather than a bimodal distribution with clearly distinct TMZ-responsive and TMZ-resistant models. This may suggest that patients with GBM also exhibit a continuous spectrum of cell-intrinsic TMZ susceptibility versus resistance, rather than the binary classification typically represented by the presence or absence of MGMT promoter methylation. The notion of a continuous spectrum of TMZ responsiveness is consistent with previous work showing that MGMT promoter methylation status is also continuously distributed across patients and that there is not a clear, unambiguous level of methylation separating "methylated" from "unmethylated" patients (Hegi et al., 2019). Furthermore, recent studies suggest that quantitative levels of MGMT methylation in gliomas may also correlate with the response to TMZ (Mathur et al., 2020). It is possible that a continuous distribution of functional response scores could be used to make more quantitative predictions of survival for individual patients, by classifying patients into multiple levels of TMZ susceptibility depending on the extent of their functional TMZ response. Assessing whether these approaches are predictive will require larger patient cohorts and greater statistical power. However, if validated, more quantitative, graded predictions of drug susceptibility could be a valuable addition to existing clinical workflows.

A key distinction between this work and most other applications of functional drug susceptibility testing is that this study focuses on stratifying patients by predicted susceptibility to one specific drug, as opposed to selecting between several candidate drugs. The most common application of functional drug susceptibility testing is, given a patient and several candidate drugs, which drug is likely to be the most effective? Previous studies in leukemia (Swords et al., 2018) and multiple myeloma (Cetin et al., 2017) have focused on this application, in which the goal is to test many candidate drugs against a patient's tumor and to select the most effective from among them. However, our study focused on a simpler question: given a single standard-of-care drug and a population of patients, which patients are most likely to benefit from the drug? Although more limited in scope compared with studies that test dozens of novel drugs and combinations, such studies focused on single clinical scenarios of need are more likely to be translated into clinical practice and FDA approval as diagnostics.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Description of patient cohort
  - GBM patient-derived model initiation and culture
  - Ba/F3 BCR-ABL and Ba/F3 BCR-ABL T315I cell lines
  - PC9 and PC9-GR4 cell lines
- METHOD DETAILS
  - Functional testing of PC9 and PC9-GR4 cell lines
  - Functional testing of Ba/F3 BCR-ABL and Ba/F3 BCR-ABL T315I cell lines
  - Functional testing of GBM patient-derived models
  - Functional testing of GBM patient-derived neurosphere models
  - SMR operation
  - Measurement of MGMT promoter methylation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - SMR mass response score
  - CellTiter-Glo response score
  - Receiver operator characteristic analysis
  - Comparing throughput of the MAR and SMR mass assays

- Validation of the SMR mass assay using conventional cancer cell lines
- Comparing patient-derived model MGMT status to patient MGMT status
- Comparing overall survival in a reduced patient subset

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109788.

ACKNOWLEDGMENTS

The authors thank David Weinstock for providing the Ba/F3 BCR-ABL T315I cell line, Pasi Jänne for providing the PC9 and PC9-GR4 cell lines, Ahmed Idbaih for providing N16-1162, and Jesse Boehm and members of the Cancer Cell Line Factory for providing other patient-derived models. This work was supported by the MIT Center for Precision Cancer Medicine and the Cancer Systems Biology Consortium U54 CA217377 (S.R.M.), R33 CA191143 (S.R.M.), and Cancer Center support (core) grant P30-CA14051 from the National Cancer Institute, as well as P50 CA165962 (K.L.L.) and R01CA219943 (K.L.L.).

AUTHOR CONTRIBUTIONS

M.A.S., S.M, S.R.M., and K.L.L. designed the experiments. M.A.S., S.M., J.C.Y., and M.M. performed the experiments. S.M., P.Y.W., A.S.K., and J.G. collected and interpreted patient clinical and molecular data. S.M., M.T., and K.-H.C. established patient-derived neurosphere models. M.A.S. and S.M. analyzed the data. M.A.S. wrote the manuscript with contributions from S.M., K.L.L., and S.R.M. All authors reviewed and approved the manuscript.

DECLARATION OF INTERESTS

S.R.M. and K.L.L. are founders of Travera. S.R.M. is a founder of Affinity Biosensors and an inventor of a relevant patent (US8087284B2). K.L.L. receives consulting fees from BMS, Rarecyte, and Integragen, research funding to DFCI from BMS, Lilly, and Amgen. The other authors declare no competing interests.

Received: January 13, 2021
Revised: August 17, 2021
Accepted: September 10, 2021
Published: October 5, 2021

REFERENCES

Bäumer, C., Fisch, E., Wedler, H., Reinecke, F., and Korfhage, C. (2018). Exploring DNA quality of single cells for genome analysis with simultaneous whole-genome amplification. Sci. Rep. *8*, 7476.

Bhola, P.D., Ahmed, E., Guerriero, J.L., Sicinska, E., Su, E., Lavrova, E., Ni, J., Chipashvili, O., Hagan, T., Pioso, M.S., et al. (2020). High-throughput dynamic BH3 profiling may quickly and accurately predict effective therapies in solid tumors. Sci. Signal. *13*, 1–12.

Burg, T.P., Godin, M., Knudsen, S.M., Shen, W., Carlson, G., Foster, J.S., Babcock, K., and Manalis, S.R. (2007). Weighing of biomolecules, single cells and single nanoparticles in fluid. Nature *446*, 1066–1069.

Burstein, H.J., Mangu, P.B., Somerfield, M.R., Schrag, D., Samson, D., Holt, L., Zelman, D., and Ajani, J.A.; American Society of Clinical Oncology (2011). American Society of Clinical Oncology clinical practice guideline update on the use of chemotherapy sensitivity and resistance assays. J. Clin. Oncol. *29*, 3328–3330.

Calistri, N.L., Kimmerling, R.J., Malinowski, S.W., Touat, M., Stevens, M.M., Olcum, S., Ligon, K.L., and Manalis, S.R. (2018). Microfluidic active loading of single cells enables analysis of complex clinical specimens. Nat. Commun. *9*, 4784.

Cermak, N., Olcum, S., Delgado, F.F., Wasserman, S.C., Payer, K.R., A Murakami, M., Knudsen, S.M., Kimmerling, R.J., Stevens, M.M., Kikuchi, Y., et al. (2016). High-throughput measurement of single-cell growth rates using serial microfluidic mass sensor arrays. Nat. Biotechnol. *34*, 1052–1059.

Cetin, A.E., Stevens, M.M., Calistri, N.L., Fulciniti, M., Olcum, S., Kimmerling, R.J., Munshi, N.C., and Manalis, S.R. (2017). Determining therapeutic susceptibility in multiple myeloma by single-cell mass accumulation. Nat. Commun. *8*, 1613.

Esteller, M., Hamilton, S.R., Burger, P.C., Baylin, S.B., and Herman, J.G. (1999). Inactivation of the DNA repair gene O^6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res. *59*, 793–797.

Friedman, A.A., Letai, A., Fisher, D.E., and Flaherty, K.T. (2015). Precision medicine for cancer with next-generation functional diagnostics. Nat. Rev. Cancer *15*, 747–756.

Hegi, M.E., Genbrugge, E., Gorlia, T., Stupp, R., Gilbert, M.R., Chinot, O.L., Nabors, L.B., Jones, G., Van Criekinge, W., Straub, J., and Weller, M. (2019). MGMT promoter methylation cutoff with safety margin for selecting glioblastoma patients into trials omitting temozolomide: A pooled analysis of four clinical trials. Clin. Cancer Res. *25*, 1809–1816.

Hirose, Y., Berger, M.S., and Pieper, R.O. (2001). p53 effects both the duration of G2/M arrest and the fate of temozolomide-treated human glioblastoma cells. Cancer Res. *61*, 1957–1963.

Hochmair, M.J., Buder, A., Schwab, S., Burghuber, O.C., Prosch, H., Hilbe, W., Cseh, A., Fritz, R., and Filipits, M. (2019). Liquid-biopsy-based identification of EGFR T790M mutation-mediated resistance to afatinib treatment in patients with advanced EGFR mutation-positive NSCLC, and subsequent response to osimertinib. Target. Oncol. *14*, 75–83.

Jacob, F., Salinas, R.D., Zhang, D.Y., Nguyen, P.T.T., Schnoll, J.G., Wong, S.Z.H., Thokala, R., Sheikh, S., Saxena, D., Prokop, S., et al. (2020). A patient-derived glioblastoma organoid model and biobank recapitulates inter- and intra-tumoral heterogeneity. Cell *180*, 188–204.e22.

Kitsos, C., and Toulias, T. (2017). Hellinger distance between generalized normal distributions. Br. J. Math. Comput. Sci. *21*, 1–16.

Kulesz-Martin, M.F., Lagowski, J., Olson, S., Wortham, A., West, T., Thomas, G., Ryan, C., and Tyner, J.W. (2013). A molecular case report: functional assay of tyrosine kinase inhibitors in cells from a patient's primary renal cell carcinoma. Cancer Biol. Ther. *14*, 95–99.

Lee, S.Y. (2016). Temozolomide resistance in glioblastoma multiforme. Genes Dis. *3*, 198–210.

Letai, A. (2017). Functional precision cancer medicine-moving beyond pure genomics. Nat. Med. *23*, 1028–1035.

Marquart, J., Chen, E.Y., and Prasad, V. (2018). Estimation of the percentage of US patients with cancer who benefit from genome-driven oncology. JAMA Oncol. *4*, 1093–1098.

Martínez-Garcia, M., Álvarez-Linera, J., Carrato, C., Ley, L., Luque, R., Maldonado, X., Martínez-Aguillo, M., Navarro, L.M., Vaz-Salgado, M.A., and Gil-Gil, M. (2018). SEOM clinical guidelines for diagnosis and treatment of glioblastoma (2017). Clin. Transl. Oncol. *20*, 22–28.

Mathur, R., Zhang, Y., Grimmer, M.R., Hong, C., Zhang, M., Bollam, S., Petrecca, K., Clarke, J., Berger, M.S., Phillips, J.J., et al. (2020). MGMT promoter methylation level in newly diagnosed low-grade glioma is a predictor of hypermutation at recurrence. Neuro-oncol. *22*, 1580–1590.

McLendon, R., Friedman, A., Bigner, D., Van Meir, E.G., Brat, D.J., Mastrogianakis, G.M., Olson, J.J., Mikkelsen, T., Lehman, N., Aldape, K., et al.; Cancer Genome Atlas Research Network (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature *455*, 1061–1068.

Niepel, M., Hafner, M., Mills, C.E., Subramanian, K., Williams, E.H., Chung, M., Gaudio, B., Barrette, A.M., Stern, A.D., Hu, B., et al.; LINCS Consortium (2019).

A multi-center study on the reproducibility of drug-response assays in mammalian cell lines. Cell Syst. *9*, 35–48.e5.

Olcum, S., Cermak, N., Wasserman, S.C., and Manalis, S.R. (2015). High-speed multiple-mode mass-sensing resolves dynamic nanoscale mass distributions. Nat. Commun. *6*, 7070.

Ollier, E., Mazzocco, P., and Ricard, D. (2017). Analysis of temozolomide resistance in low-grade gliomas using a mechanistic mathematical model. Fundam. Clin. Pharmacol. *31*, 347–358.

Son, S., Tzur, A., Weng, Y., Jorgensen, P., Kim, J., Kirschner, M.W., and Manalis, S.R. (2012). Direct observation of mammalian cell growth and size regulation. Nat. Methods *9*, 910–912.

Stevens, M.M., Maire, C.L., Chou, N., Murakami, M.A., Knoff, D.S., Kikuchi, Y., Kimmerling, R.J., Liu, H., Haidar, S., Calistri, N.L., et al. (2016). Drug sensitivity of single cancer cells is predicted by changes in mass accumulation rate. Nat. Biotechnol. *34*, 1161–1167.

Stockslager, M.A., Olcum, S., Knudsen, S.M., Kimmerling, R.J., Cermak, N., Payer, K.R., Agache, V., and Manalis, S.R. (2019). Rapid and high-precision sizing of single particles using parallel suspended microchannel resonator arrays and deconvolution. Rev. Sci. Instrum. *90*, 085004.

Swords, R.T., Azzam, D., Al-Ali, H., Lohse, I., Volmar, C.H., Watts, J.M., Perez, A., Rodriguez, A., Vargas, F., Elias, R., et al. (2018). Ex-vivo sensitivity profiling to guide clinical decision making in acute myeloid leukemia: A pilot study. Leuk. Res. *64*, 34–41.

Tiriac, H., Belleau, P., Engle, D.D., Plenker, D., Deschênes, A., Somerville, T.D.D., Froeling, F.E.M., Burkhart, R.A., Denroche, R.E., Jang, G.H., et al. (2018). Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. Cancer Discov. *8*, 1112–1129.

Touat, M., Li, Y.Y., Boynton, A.N., Spurr, L.F., Iorgulescu, J.B., Bohrson, C.L., Cortes-Ciriano, I., Birzu, C., Geduldig, J.E., Pelton, K., et al. (2020). Mechanisms and therapeutic implications of hypermutation in gliomas. Nature *580*, 517–523.

Vlachogiannis, G., Hedayat, S., Vatsiou, A., Jamin, Y., Fernández-mateos, J., Khan, K., Lampis, A., Eason, K., Huntingford, I., Burke, R., et al. (2018). Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science *926*, 920–926.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Glioblastoma patient-derived neurosphere models	DFCI Center for Patient-Derived Models	https://www.dana-farber.org/cpdm
Chemicals, peptides, and recombinant proteins		
Temozolomide	Selleck Chemicals	CCRG81045
Imatinib	Selleck Chemicals	STI571
Ponatinib	Selleck Chemicals	AP24534
Gefitinib	Selleck Chemicals	ZD1839
Osimertinib	Selleck Chemicals	AZD9291
Critical commercial assays		
CellTiter-Glo	Promega	G7570
Deposited data		
Functional testing and genomics raw data	Mendeley Data	https://doi.org/10.17632/9fyc3mk76x
Experimental models: Cell lines		
Ba/F3 BCR-ABL	Weinstock lab, DFCI	CVCL_UE63
Ba/F3 BCR-ABL T315I	Weinstock lab, DFI	CVCL_UE64
PC9	Jänne lab, DFCI	CVCL_B260
PC9-GR4	Jänne lab, DFCI	CVCL_DH34
Software and algorithms		
Raw data analysis code	Mendeley Data	https://doi.org/10.17632/8fp65yrn34

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Scott Manalis (srm@mit.edu).

Materials availability

Additional samples from 65/69 patient-derived neurosphere models are publicly available for licensing from the Dana-Farber Cancer Institute Center for Patient Derived Models (https://www.dana-farber.org/cpdm); see also the Key resources table and Figure S1.

Data and code availability

- Genomics and functional testing raw data has been deposited at Mendeley Data and is publicly available as of the date of publication (Mendeley Data: https://doi.org/10.17632/9fyc3mk76x)
- The code used to analyze the data has also been deposited at Mendeley Data and is publicly available as of the date of publication (Mendeley Data: https://doi.org/10.17632/8fp65yrn34)
- Any additional information required to reanalyze the data reported in this paper is available from the Lead Contact (Scott Manalis; srm@mit.edu) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Description of patient cohort

All patients, clinical data and models were studied following consent to research (DFCI IRB#10-417) or waiver of consent (DFCI IRB#10-043) per institutional review board procedures at the Dana-Farber Cancer Institute and Brigham and Women's Hospital. We performed functional testing on 69 patient-derived models established from GBM patients, all of whom had wild-type IDH. Of these 69 models, functional response data was successfully collected for 67/69 models using the SMR assay and for 55/69 models

using the CellTiter-Glo assay. Of the 69 models, 64/69 have a wild-type mismatch repair (MMR) genotype, while 5/69 have MMR mutations.

For comparing functional response data to MGMT methylation status, we restricted our analysis to the subset of patients with wild-type MMR genotypes. Of these 64 patients, 63/64 had models with known MGMT methylation status (38/63 methylated, 25/63 unmethylated). Further, of these 64 patients, 62/64 had SMR mass assay data and 52/64 had CellTiter-Glo assay data.

For comparing functional response data to patient survival outcomes, we further restricted our analysis to patients who were newly-diagnosed and were treated with TMZ. Of the 69 total models, 45/769 were derived from newly-diagnosed patients, and of these 45 patients, 32/45 were treated with TMZ.

Of these 32 eligible models (those established from newly-diagnosed patients who were treated with TMZ), overall survival duration was known for 30/32 patients. Of these 30 patients, 28/30 have SMR mass assay data, 25/30 have CellTiter-Glo assay data, and 29/30 have known patient-derived model MGMT status (18/29 methylated, 11/29 unmethylated).

GBM patient-derived model initiation and culture

GBM patient-derived models were established and maintained as described previously (Stevens et al., 2016). Briefly, GBM tumor resections were subjected to enzymatic and mechanical dissociation, then seeded in neurosphere culture conditions. The models were propagated in a proprietary neurosphere culture medium (NeuroCult; STEMCELL Technologies) supplemented with additional growth factors (20 ng/mL epidermal growth factor, 10 ng/mL fibroblast growth factor) and 2 μg/mL heparin. Models were propagated in ultra-low attachment flasks (Corning 3814), and passaged by dissociating to single cells (5 minute treatment with 1X Accutase at 37 °C; STEMCELL Technologies) and resuspending at a concentration of 100k-300k cells/mL. Models and data are available via the Dana Farber Cancer Institute Center for Patient Derived Models (models@dfci.harvard.edu) and were created in laboratories at DFCI (Ligon) and the Broad Institute (K. Ligon and J. Boehm, Cancer Cell Line Factory). Where available from clinical records, the sex of each cell line is reported in Table S1.

Ba/F3 BCR-ABL and Ba/F3 BCR-ABL T315I cell lines

Ba/F3 BCR-ABL and Ba/F3 BCR-ABL T315I cell lines were a gift from the Weinstock laboratory at the Dana-Farber Cancer Institute. Both the Ba/F3 BCR-ABL and Ba/F3 BCR-ABL T315I cell lines were maintained in RPMI-1640 medium (ThermoFisher) supplemented with 10% FBS (Sigma-Aldrich), 25 mM HEPES (GIBCO), and 1X antibiotic/antimycotic (GIBCO). For maintenance, cells were passaged every 2-3 days to a minimum concentration of 75,000 cells/mL.

PC9 and PC9-GR4 cell lines

PC9 and PC9-GR4 cell lines were a gift from the Jänne laboratory at the Dana Farber Cancer Institute. Both cell lines were maintained in RPMI-1640 medium (ThermoFisher) supplemented with 10% FBS (Sigma-Aldrich), 25 mM HEPES (GIBCO), and 1X antibiotic/antimycotic (GIBCO) in 75 cm^2 culture flasks (VWR 10062-860). For maintenance, cells were passaged every 3 days to a concentration 50k cells/mL in a volume of 15 mL (i.e., 10k cells/cm^2). At each passage, medium was aspirated, cells were detached from the culture surface by incubating with 0.25% trypsin-EDTA (ThermoFisher) for 10 minutes at 37°C, then washing with medium and resuspending at the desired concentration.

METHOD DETAILS

Functional testing of PC9 and PC9-GR4 cell lines

For functional testing, cells were trypsinized and seeded in 12-well plates (Argos P1012) at a concentration of 50k cells/mL in a volume of 1 mL. After 24 hours, cells were exposed to gefitinib (Selleck Chemicals), osimertinib (Selleck Chemicals), or a vehicle control (0.1% DMSO). At time points of 8, 16, and 24 hours of drug exposure, cells were trypsinized then resuspended in 150 μL medium for SMR measurement, with a measurement duration of 20 minutes per sample.

Functional testing of Ba/F3 BCR-ABL and Ba/F3 BCR-ABL T315I cell lines

For functional testing, cells were seeded at 75,000 cells/mL in 12-well plates (Argos P1012), with a volume of 1 mL/well. After 24 hours, cells were dosed with imatinib (Santa Cruz Biotechnology), ponatinib (Santa Cruz Biotechnology), or a vehicle control (0.2% DMSO). At 8-10 hours of drug exposure, cells were collected and immediately sampled by a SMR system for up to 20 minutes per sample.

Functional testing of GBM patient-derived models

For the SMR assay, patient-derived models were dissociated to single cells (5-minute treatment with 1X Accutase at 37°C), then seeded at a density of 15,000 cells/mL in 24-well ultra-low attachment plates (Costar 3473), with a volume of 1 mL/well. Because Accutase dissociation is the same technique used in the biweekly neurosphere passage protocol, we do not expect this process to interfere with cell viability. After allowing 24 hours for neurosphere formation, the cells were exposed to 20 μM temozolomide or a vehicle control (0.1% DMSO). Medium was replenished (100 μL/well) at days 3, 6, 10, and 13 after drug exposure, while cells were sampled for SMR measurement on days 3, 5, 7, 10, 12, and 14. For SMR mass measurement, neurospheres were dissociated

to single cells by treatment with 1X Accutase for 10 minutes at 37 °C, then resuspended in 50 uL medium and sampled by the SMR for 20 minutes. Samples with fewer than 300 cells detected in this time period were excluded.

Functional testing of GBM patient-derived neurosphere models

For the CellTiter-Glo assay, patient-derived models were dissociated to single cells (5-minute treatment with 1X Accutase at 37°C), then seeded at a density of 1500 cells/mL in 96-well ultra-low attachment flat-bottom plates (Corning 7007), with a volume of 100 μL/well, with three biological replicates per condition. After allowing 24 hours for neurosphere formation, the cells were exposed to 20 μM temozolomide or a vehicle control (0.1% DMSO). Medium was replenished (10 μL/well) at days 3, 6, 10, and 13 after drug exposure. At days 3, 5, 7, 10, 12, and 14 after drug exposure, the CellTiter-Glo assay was performed following the manufacturer's protocol. For the CellTiter-Glo assay, we excluded time points for which there was unusually high variation between replicates (specifically, we excluded time points for which the coefficient of variation of the measured luminescence signal between replicates was greater than 30%, excluding a total of 7% of the measured time points).

SMR operation

The design and operation of the suspended microchannel resonator (SMR) has been described thoroughly in previous work from our group (Burg et al., 2007; Cermak et al., 2016; Olcum et al., 2015; Son et al., 2012). Briefly, single cells in suspension flow through a micromechanical resonator with an embedded fluidic channel, generating a shift in resonance frequency proportional to the cell's mass. The sensors are calibrated by measuring monodisperse polystyrene beads of known mass. The measurements described here were performed using a parallel SMR array, in which twelve sensors are connected fluidically in parallel and operated simultaneously for increased throughput (Stockslager et al., 2019). Individual SMR chips can be washed between samples and are reusable, typically for several months of routine use. Before measurement, the fluidic channels were passivated with poly-L-lysine-grafted poly(ethylene glycol). Cell samples remained at room temperature throughout the (typical) 20-minute measurement duration.

Measurement of MGMT promoter methylation

Methylation of the CpG island of the MGMT gene was measured using standard methylation-specific PCR at the Brigham and Women's Hospital Center for Advanced Molecular Diagnostics. Specifically, bisulfite treatment converted unmethylated (but not methylated) cytosines to uracil, prior to PCR using primers specific for either the methylated or modified unmethylated DNA (Esteller et al., 1999). PCR products were analyzed using capillary gel electrophoresis in duplicate parallel runs. Partially methylated or ambiguous calls were verified by a molecular pathologist (A.K.) and binned into one of the categories based on this review.

QUANTIFICATION AND STATISTICAL ANALYSIS

Details of each statistical analysis (including statistical tests and exact sample sizes) are provided in the corresponding figure legends. The size of our cohort was determined by the availability of sample material rather than an *a priori* power analysis. When comparing functional response scores between MGMT methylated versus unmethylated patients, we initially limited our analysis to the subset of 64/69 models which had no MMR alterations. Then, when comparing survival between functional responders and non-responders, we limited our analysis to the 30 patient-derived models derived from patients who at the time were newly diagnosed, had wild-type IDH status, were subsequently treated with TMZ, and had known overall survival.

SMR mass response score

For each functional assay, a "response score" was calculated that summarizes the extent to which the sample responds to the drug treatment compared to a vehicle control across all measured time points.

The SMR mass response score is based on the Hellinger distance between the control and drug-treated cell mass distributions. The Hellinger distance is a measure of statistical distance between mass distributions $P(m)$ and $Q(m)$, and is defined as:

$$H(P,Q) = \left[1 - \int_0^\infty \sqrt{P(m)Q(m)}\, dm\right]^{1/2}$$

A Hellinger distance of 0 corresponds to no difference between the control and drug-treated mass distributions (i.e., no drug response), and a Hellinger distance of 1 would correspond to completely non-overlapping mass distributions. This summary statistic has the advantage that is agnostic to whether drug treatment caused mass to increase or decrease, and simply identifies the degree of change in the mass distribution. We evaluated the Hellinger distance between each control-drug pair by computing kernel density estimates of P and Q and then numerically integrating. Further, we obtained bootstrap standard errors and confidence intervals for the Hellinger distance by repeatedly resampling the measured cell mass distributions and re-computing the kernel density estimates and Hellinger distance (200 iterations per sample).

For the GBM patient-derived models, the SMR mass response score is defined as the average of the Hellinger distance between the TMZ and vehicle control samples at days 5, 7, and 10.

CellTiter-Glo response score

For the GBM patient-derived models, the "CellTiter-Glo response score" was defined as $1 - (\textit{average viability across timepoints})$, where "viability" is defined as the ratio of the mean CTG luminescence signal for the TMZ sample to the mean CTG luminescence signal for the vehicle control sample. This metric integrates the difference between the treatment and control conditions across all measured time points, with larger CTG response scores corresponding to a greater reduction in CTG luminescence signal in the drug-treated samples compared to matched controls.

Receiver operator characteristic analysis

To evaluate whether functional biomarkers predicted patient outcome, we computed the receiver operator characteristic (ROC) between the functional biomarker value (i.e., the SMR mass response score or CellTiter-Glo response score) and the binary survival outcome of interest (e.g., 15-month survival). We computed the ROC area-under-the-curve statistic (ROC AUC) by integrating the receiver operator characteristic true positive rate with respect to false positive rate.

Comparing throughput of the MAR and SMR mass assays

For both the MAR assay and the SMR mass assay, throughput (in samples measured per hour of instrument time) is determined from the rate at which individual cells can be measured and the number of cells required to detect a significant MAR or cell mass changes with sufficient statistical power.

Previously, it has been reported that serial SMR array devices measure MAR at a maximum rate of approximately 200 cells/hr (Calistri et al., 2018). While required sample sizes depend strongly on the specific cell type and the expected effect size of the drug, previous studies using MAR have typically reported sample sizes on the order of 100-200 cells (Calistri et al., 2018; Cetin et al., 2017; Stevens et al., 2016). Therefore, we estimated that the MAR assay has average throughput of 1 sample per hour of instrument time.

We used a similar approach to estimate the sample throughput of the parallel SMR array. Based on a previous study, these devices can measure cell mass at a maximum throughput of 6800 particles/minute when optimized for speed (Stockslager et al., 2019). For our order-of-magnitude throughput calculations we used a more conservative estimate of 1000 particles/minute, since most tumor cell samples are sparse and throughput is lower when measuring these generally lower-concentration samples. Next, to estimate required sample sizes, we computed the number of measurements required to detect a significant reduction in mean mass with specified power $(1 - \beta)$, and confidence $(1 - \alpha)$, obtaining the following expression for a drug treatment that reduces mean cell mass from m to $m - \Delta m$:

$$n > \frac{2\left(CV_{population}^2 + CV_{error}^2\right)\left(z_{1-\alpha} + z_{1-\beta}\right)^2}{\left(\Delta m/m\right)^2}$$

where n is the required sample size, $CV_{population}$ is the coefficient of variation of the untreated cells' masses, and CV_{error} is the fractional error of the mass measurement, and z is the inverse normal cumulative distribution function. Although we generally compared drug-treated versus control samples using the Hellinger distance rather than directly comparing the mean mass, this calculation provides a rough estimate of required sample sizes. Using this expression, a sample size of n = 2000 cells allows us to detect a reduction in mean cell mass as small as 3% for a typical sample ($CV_{population}$ = 30%, CV_{error} = 1%) with 80% power and 99% confidence. Therefore, we estimated that the parallel SMR array has average throughput of 1 sample per 2 minutes of instrument time.

Validation of the SMR mass assay using conventional cancer cell lines

Unlike the MAR assay, the SMR mass assay has not been previously applied to *ex vivo* drug susceptibility testing. Therefore, as a first step, we used conventional cancer cell lines as a model system to validate that single-cell mass measurements were consistent with expected patterns of drug susceptibility and resistance. We used two model systems: (1) BCR-ABL-positive leukemia cell lines treated with BCR-ABL inhibitors, and (2) EGFR mutant lung adenocarcinoma cell lines treated with EGFR inhibitors.

First, we replicated a previous experiment in which leukemia cell lines expressing the oncogenic BCR-ABL fusion protein were exposed to the BCR-ABL inhibitors imatinib and ponatinib (Figures S1A–S1D). We have observed previously that in the cell line Ba/F3 BCR-ABL, exposure to BCR-ABL inhibitors arrests cells in G1 and drastically alters their growth rate distributions (Stevens et al., 2016). Consistent with this finding, the SMR detected a significant reduction in mean cell mass within 8 hours of exposure to 1.4 μM imatinib and 100 nM ponatinib (Figure S1A). For Ba/F3 BCR-ABL T315I, a cell line engineered to express an imatinib-resistant mutant BCR-ABL, cell mass was not significantly reduced in response to 1.4 μM imatinib, but as expected, was reduced in response to 100 nM ponatinib (Figure S1B).

To exploit the increased throughput of the SMR mass assay, we next measured dose-response curves for the Ba/F3 and Ba/F3 BCR-ABL T315I cell lines after 8 hours exposure to imatinib (Figures S1C and S1D). Instead of comparing the mean mass between drug-exposed cells and untreated controls, we computed an alternative summary statistic, the Hellinger distance, to evaluate to what extent the cell mass distributions were altered by drug exposure. The Hellinger distance is a statistic that measures the degree of difference between the mass distributions of drug-treated and untreated cells, where larger Hellinger distances reflect greater differences between the treated and untreated mass distributions (Method details). A drug treatment with no effect on the mass distribution would have a Hellinger distance of zero, while drug treatments causing larger shifts in the cell mass distribution would be

assigned larger Hellinger distances, up to a maximum of one. The Hellinger distance statistic has the advantage of capturing effects other than changes in cell mass, such as the accumulation of small debris in the culture due to drug-induced cell death, or the emergence of cell subpopulations of different sizes in response to drug exposure. As expected, Hellinger distance increased with imatinib dose for Ba/F3 BCR-ABL but not for Ba/F3 BCR-ABL T315I (Figures S1C and S1D).

Next, we asked whether the SMR mass assay could also detect expected patterns of drug sensitivity and resistance in a solid tumor cell line model. We used PC9, an EGFR mutant human lung adenocarcinoma cell line (EGFR Del-E745-A750) known to be sensitive to the EGFR inhibitor gefitinib. Additionally, we used PC9-GR4, a resistance model containing EGFR T790M, a secondary mutation which is known to confer resistance to gefitinib (Hochmair et al., 2019). As expected, 1 μM gefitinib induced a large reduction in population cell mass in PC9 after 24 hours of drug exposure (Figure S1E; 19% reduction), but only a small change in PC9-GR4 (Figure S1F; 6% reduction). Also consistently, for both cell lines we observed a large reduction in population cell mass after 24 hour exposure to 100 nM osimertinib, an EGFR inhibitor that also inhibits EGFR T790M (Hochmair et al., 2019) (16% and 34% mean mass reduction respectively for PC9 and PC9-GR4).

In addition to measuring the cell mass response after 24 hours of drug exposure, we also captured the transient gefitinib and osimertinib responses by sampling the cell mass distributions at 8, 16, and 24 hours of exposure (Figures S1G and S1H). As expected, the Hellinger distance increased over time for PC9 but not PC9-GR4 when exposed to gefitinib, but increased over time for both cell lines when exposed to osimertinib. Therefore, cell mass measurements reliably and sensitively identified expected patterns of drug sensitivity and resistance in both liquid and solid tumor cell line model systems with vastly increased throughput compared to previous SMR-based approaches.

Comparing patient-derived model MGMT status to patient MGMT status

We measured the MGMT methylation status of the patient-derived neurosphere models (“mMGMT”) for 68/69 models in our cohort. For 58 of these 68 patients, we also independently measured MGMT methylation status of the primary tumor sample at the time of collection (“pMGMT”). Here we compare in detail for which patients these two MGMT status results (mMGMT and pMGMT) agree and disagree.

Of the 58 models where both pMGMT and mMGMT are known, 34/58 (59%) of models had methylated mMGMT and the remaining 24/58 (41%) had unmethylated mMGMT. Of the 34 models with methylated mMGMT, 18/34 (53%) also had methylated pMGMT, while 16/34 (47%) had unmethylated pMGMT. Of the 24 models with unmethylated mMGMT, 17/24 (71%) also had unmethylated pMGMT, while 7/24 (29%) had methylated pMGMT.

In other words, 18/58 patients (31%) had both methylated pMGMT and methylated mMGMT; 17/58 patients (29%) had both unmethylated pMGMT and unmethylated mMGMT; 7/58 patients (12%) had methylated pMGMT but unmethylated mMGMT; and 16/58 patients (28%) had unmethylated pMGMT but methylated mMGMT.

It remains unclear whether these shifts in MGMT methylation between the patient sample and patient-derived model occur due to bias in the MGMT assay (either an abundance of false negative unmethylated calls from primary samples, or false positive methylation calls from PDM samples), or whether this corresponds to a real phenotypic shift between the primary sample and patient-derived model.

While the pMGMT and mMGMT biomarkers were both correlated with functional testing results (i.e., MGMT methylated samples tended to have more significant TMZ responses), the pMGMT biomarker was slightly more strongly correlated than was the mMGMT biomarker. We computed the ROC AUC statistic to quantify to what extent functional response scores are segregated by both the pMGMT and mMGMT biomarkers. For both assays, the ROC AUC statistic was similar for pMGMT and mMGMT (SMR, ROC AUC 0.79 versus 0.75; CellTiter-Glo, ROC AUC 0.83 versus 0.81).

Comparing overall survival in a reduced patient subset

To confirm that the SMR and mMGMT biomarkers did not reach statistical significance due only to their increased sample size relative to the CellTiter-Glo (28 eligible patients had SMR data and 29 patients had known mMGMT status, while only 25 patients had CellTiter-Glo data), we repeated this analysis limited to the subset of 25 patients with CellTiter-Glo data and found similar results. Specifically, among this subset of patients, overall survival was still significantly longer for SMR assay responders versus non-responders (median overall survival 20.4 versus 12.5 months, log-rank $p = 0.046$), and for patients with MGMT methylated versus unmethylated models (median overall survival 19.7 versus 11.3 months, log-rank $p = 0.01$), but not for CellTiter-Glo responders versus non-responders (median overall survival 18.8 versus 10.3 months, log-rank $p = 0.15$).