DNA repair capacity in multiple pathways predicts chemoresistance in glioblastoma multiforme

Citation: Nagel, Zachary D., Gaspar J. Kitange, Shiv K. Gupta, Brian A. Joughin, Isaac A. Chaim, Patrizia Mazzucato, Douglas A. Lauffenburger, Jann N. Sarkaria, and Leona D. Samson. "DNA Repair Capacity in Multiple Pathways Predicts Chemoresistance in Glioblastoma Multiforme." Cancer Research (October 28, 2016).

As Published: http://dx.doi.org/10.1158/0008-5472.CAN-16-1151

Publisher: American Association for Cancer Research (AACR)

Persistent URL: http://hdl.handle.net/1721.1/105169

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

DNA repair capacity in multiple pathways predicts chemoresistance in glioblastoma multiforme.

Running Title: DNA repair capacity predicts chemoresistance in glioblastoma

Authors:

Zachary D. Nagel[1], Gaspar J. Kitange[2], Shiv K. Gupta[2], Brian A. Joughin[3], Isaac A. Chaim[1], Patrizia Mazzucato[1], Douglas A. Lauffenburger[1], Jann N. Sarkaria[2], Leona D. Samson[1].

[1] Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139. [2] Department of Radiation Oncology, Mayo Clinic, Rochester, MN 55905. [3] David H. Koch Institute for Integrative Cancer Research at MIT

Corresponding Author: Leona Samson, Massachusetts Institute of Technology, Department of Biological Engineering, Building 56-255, 77 Massachusetts Ave, Cambridge, MA 02139

Conflict of Interest: The authors declare no conflicts of interest.

Abbreviations:
MGMT, O^6-methylguanine DNA methyltransferase;
O^6-MeG, O^6-methylguanine
HR, homologous recombination;
NER, nucleotide excision repair;
NHEJ, non-homologous end joining;
MMR, mismatch repair;
DRC, DNA Repair Capacity;
FM-HCR, fluorescence-based multiplexed host cell reactivation;
MLR, multiple linear regression;
LASSO, least absolute shrinkage and selection operator;
ROC, receiver operating characteristic;
AUC, area under the curve;
MNNG, N-methyl-N'-nitro-N-nitrosoguanidine;
TMZ, temozolomide;
MMS, methyl methanesulfonate;
BCNU, 1,3-bis(2-chloroethyl)-1-nitrosourea;
PDX, patient derived xenograft;
GBM, glioblastoma multiforme;

Abstract: Cancer cells can resist the effects of DNA-damaging therapeutic agents via utilization of DNA repair pathways, suggesting that DNA repair capacity (DRC) measurements in cancer cells could be used to identify patients most likely to respond to treatment. However, the limitations of available technologies have so far precluded adoption of this approach in the clinic. We recently developed fluorescence-based multiplexed host cell reactivation (FM-HCR) assays to measure DRC in multiple pathways. Here we apply a mathematical model that uses DRC in multiple pathways to predict cellular resistance to killing by DNA-damaging agents. This model, developed using FM-HCR and drug sensitivity measurements in 24 human lymphoblastoid cell lines, was applied to a panel of 12 patient-derived xenograft (PDX) models of glioblastoma (GBM) to predict GBM response to treatment with the chemotherapeutic DNA damaging agent temozolomide (TMZ). This work showed that, in addition to changes in O6-methylguanine DNA methyltransferase (MGMT) activity, small changes in mismatch repair (MMR), nucleotide excision repair (NER), and homologous recombination (HR) capacity contributed to acquired TMZ resistance in PDX models, and lead to reduced relative survival prolongation following TMZ treatment of orthotopic mouse models in vivo. Our data indicate that measuring the combined status of MMR, HR, NER, and MGMT provided a more robust prediction of TMZ resistance than assessments of MGMT activity alone.

Introduction

Therapy resistance is a universal problem in cancer treatment, and our understanding of its origins remains incomplete. Many cancers can initially be treated effectively with chemotherapy and radiation. However, some cancers are innately resistant to therapy, and most cancers acquire resistance to therapy over the course of treatment. Consequently, patients who do not benefit from treatment are needlessly subjected to the severe and sometimes lasting side effects of chemotherapy and radiotherapy, including an increased risk

of second primary cancers. Moreover, precious time is lost, during which alternative and potentially more effective therapy could have been applied.

Most cancer therapy regimens include chemotherapy and radiotherapy, both of which damage DNA, killing cancer cells by inducing damage that interferes with DNA replication and transcription, ultimately activating cell death pathways. Efficient repair of DNA damage can render cells resistant to killing. Each drug may induce several different types of DNA damage, and at least 6 major DNA repair pathways, as well as numerous sub-pathways, are involved in the repair of damage (1-3). Accordingly, the DNA repair pathways involved in acquired resistance mechanisms vary from drug to drug, and all of the major DNA repair pathways can alter the sensitivity of cancer cells to DNA damaging agents.

GBM is an aggressive cancer with a median survival of 12-15 months under the standard of care treatment of aggressive surgery, radiation and chemotherapy. Almost all patients succumb to this disease, reflecting intrinsic or acquired resistance to frontline treatment with radiotherapy combined with the alkylating agent TMZ (4). TMZ is one of several chemotherapeutic S_N1-type methylating agents, including dacarbazine and procarbazine, that induce a similar spectrum of methylated DNA bases, among which O^6-methylguanine (O^6-MeG) is the predominant cytotoxic DNA lesion (5). Promoter hypermethylation of the *MGMT* gene, encoding a protein that repairs O^6-MeG, has been used to identify patients whose tumors are MGMT-deficient and are thus more likely to benefit from TMZ therapy (6). However, multiple DNA repair pathways affect the sensitivity of cells to killing with S_N1-type methylating agents (1), suggesting that successful personalized therapy based on DRC in GBM will require multiplexed measurements of repair activity. Progress toward this goal has been significantly hindered by a lack of quantitative, high-throughput assays for multiple DNA repair pathways, and the challenges of predicting phenotypes from genetic or other ‘omics approaches (7).

Here we apply recently developed FM-HCR assays to measure DRC for multiple repair pathways in a set of human cell lines (8). We present mathematical models that predict sensitivity to DNA damaging agents from DRC data. We then apply FM-HCR assays to assess global DRC in PDX models of GBM. Our mathematical model predicts the effectiveness of TMZ therapy both for killing GBM cells *in vitro* and for prolonging the survival of GBM tumor-bearing mice upon TMZ treatment *in vivo*. The model integrates DNA repair capacity across four repair pathways, and outperforms predictions made based on tumor MGMT activity alone. This work lays the foundation for using functional assays to make quantitative, personalized predictions concerning the effectiveness of cancer therapy.

Materials and Methods

PDX Models and Cell culture.

Lymphoblastoid cells, from the Coriell Cell Repository (Camden, NJ), were maintained in log phase in RPMI media supplemented with 15% fetal bovine serum plus Penicillin and Streptomycin (8). Cells were obtained in 2001. The Coriell Cell Repository authenticates and tests cell cultures for contamination with mycoplasma, bacteria and fungi; cells have not been authenticated by the authors. For cell cultures derived from GBM PDX lines, flank GBM tumor xenografts were excised, dissociated into single cell suspensions, and cultured in DMEM with 10% FBS as previously described (9).

MMS and TMZ Sensitivity Assays.

Lymphoblastoid cell lines were assayed for sensitivity to BCNU (reported previously in (10)) and TMZ using a cell proliferation assay that measures quenching of DNA-bound Hoechst 33258 fluorescence by BrdU incorporated during DNA synthesis (10,11). Cells were suspended in 300 µL serum free media at 4.5 x 10^5 cells per mL, and treated at a final

concentration 240 μM TMZ for 1 hour at 37 °C. Cells were centrifuged for 5 minutes at 450 g, and washed once with 200 μL PBS. Next, cells were suspended in 280 μL complete media, and subsequently cultured in flat bottom 96-well plates. After two doubling times, BrdU was added to achieve a final concentration of 45 μM. After two additional doubling times, cells were stained with Hoechst 33258 and analyzed as previously reported (10). Sensitivity to MNNG and sensitivity to MMS were measured using a previously described cell proliferation assay (12). Briefly, logarithmically growing cells were treated with0.4 mM MMS or left untreated. After 72 hours, % Control growth was calculated from the ratio of the total number of viable treated cells divided by the total number of viable untreated cells using a Vi-CELL cell viability analyzer with trypan blue staining to exclude dead cells.

FM-HCR Assays.

FM-HCR assays in lymphoblastoid cell lines were carried out as previously described (8). For FM-HCR assays in PDX lines, GBM cells were seeded at 50,000 cells per well in a 24-well tissue culture plate 24 hours before transfection. 0.5 μg of plasmid mixtures composed of previously described reporter plasmids for the MGMT, MMR, NER, NHEJ, and HR pathways were transfected into cells using 'lipofectamine 2000 LTX with plus^{TM} reagent' (Thermo Fisher). At 24 hours post-transfection, cells were analyzed for fluorescent reporters by flow cytometry and DRC for both lymphoblastoid cells and GBM cells was calculated in terms of % Reporter Expression as previously described (8).

*Mathematical Modeling***.**

The overall modeling approach is summarized in **Fig. S1**. Following log-transformation of FM-HCR data, Z-scores for relative DNA repair capacity among the 24 lymphoblastoid cell lines were generated for each pathway using **Eqn. 1**,

$$Z_{ij} = \frac{x_{ij} - \bar{x}_{ij}}{\sigma_j} \qquad (1)$$

where Z_{ij} is the Z-score for a DNA repair pathway in cell line i in pathway j, x_{ij} is the DNA repair capacity for a given pathway j in cell line i, $\bar{x}_{ij}$ is the mean value of the DNA repair capacity in pathway j over the 24 lymphoblastoid cell lines (i = 1-24), and σ_j is the standard deviation of the DNA repair capacity in pathway j. Z-scores were also generated for relative DNA repair capacity among the 12 PDX models of GBM using the same approach. Because there is an inverse relationship between % Reporter Expression and DRC for the MGMT reporter (larger signal represents lower activity), Z-scored MGMT data were multiplied by -1. Simple linear regression was used to generate the models #1 and #4, denoted as "MGMT" in **Table 1**. These models relate MGMT activity to sensitivity to killing with TMZ and MNNG. A graphical interpretation of the parameters associated with model #1 can be generalized to the multiple linear models, and is detailed in supplemental **Fig. S2**. MATLAB scripts (provided in the supplemental information) were run to generate linear models using DRC in 5 pathways in the lymphoblastoid cell lines as the predictor variables and sensitivity to cell killing (measured as % control growth) with the DNA damaging agents TMZ, MNNG, MMS, or BCNU as the response variable. First, multiple linear regression (MLR) models were fit to all 5 DNA repair pathways; these models are listed as "MLR" models in **Table 1**. Then, LASSO regression was performed to select optimal combinations of DRC predictor variables for incorporation into multiple linear regression models of sensitivity to cell killing with each DNA damaging agent (13). LASSO regression was carried out in MATLAB to return least squares regression coefficients for a set of 100 values of the regularization coefficient λ, which determines the size of the penalty for coefficients remaining in the model. The value of λ is gradually increased until the least squares regression coefficients for all five predictor variables take a value of zero, resulting in a family of combinations of predictor variables. MLR was performed and

evaluated with exhaustive leave-one-out cross validation using the combination of predictor variables selected by LASSO at each value of λ; those for which LASSO returned least squares regression coefficients equal to zero were excluded from MLR analysis. R^2 parameters were calculated for the relationship between predicted sensitivities obtained from cross-validation versus the known sensitivity for each cell line, and the model yielding the largest R^2 values were selected. Plots of sensitivity predicted from leave-one-out cross-validation versus observed sensitivity are presented in supplemental **Fig. S3**. MLR models using the combination of predictor variables selected using LASSO are listed as "MLR_L" in **Table 1**.

Receiver operating characteristic (ROC) analysis was carried out in MATLAB using PDX sensitivities calculated from mathematical models and the actual sensitivity status. DRC data were input into the models in **Table 1** to calculate the predicted sensitivity for each PDX model. To evaluate the performance of the models, ROC curves were generated, and the area under the curve (AUC) computed (**Table 1**). ROC analysis provides a quantitative estimate and graphical representation of the accuracy of predictive assays (14), and is particularly useful when the test result, reported on a continuous scale, is to be used to predict a binary variable, as in the present case. The ROC curve shows the relationship between the true positive rate and the false positive rate as the discrimination threshold is varied. For an ideal assay, AUC = 1.0, indicating that the discrimination threshold can be set in such a way that all true positives can be correctly called without calling any false positives; the distributions of values given by the assay for the positive and negative populations are completely separated. By contrast, for an assay that only performs as well as random chance, the distributions are indistinguishable, and AUC = 0.5, because there are equal chances of calling a positive or a negative at all discrimination thresholds. The true positive rate is equivalent to assay sensitivity, and the false positive rate is calculated from 1 - (assay specificity).

Results

Models Relating DNA Repair Capacity to Alkylation-Induced Cell Killing in Human Lymphoblastoid Cell Lines

24 lymphoblastoid cell lines derived from genetically diverse, apparently-healthy individuals were used as a model system (15). These cell lines have been previously shown to exhibit a broad range of sensitivity to killing by the S_N1-type alkylating agent MNNG (12), and a DNA cross-link forming alkylating agent, BCNU (10). We have now additionally measured the sensitivity of these cell lines to MMS, an S_N2-type alkylating agent, for which the predominant cytotoxic lesions induced are 3-methyladenine and 7-methylguanine, and the sensitivity of a subset of the 24 cell lines to the S_N1-type alkylating chemotherapy agent TMZ. TMZ is regarded as a functional analog of MNNG because it methylates DNA via the same reactive methyldiazonium ion intermediate (16), and TMZ sensitive cells are also sensitive to MNNG (17).

A broad and continuous range of sensitivity to MNNG (12), TMZ, BCNU (10), and MMS, is observed across the 24 lymphoblastoid cell lines (**Fig. 1A, Table S1**). The cell lines were organized in order of decreasing MNNG sensitivity and assigned to color-coded quartiles to facilitate comparisons with respect to TMZ, BCNU and MMS sensitivity. Whereas MNNG and TMZ sensitivity were strongly correlated (R = 0.81) as expected, MNNG sensitivity was less strongly correlated to BCNU sensitivity (R = 0.54) and MMS sensitivity (R = 0.49). Thus, consistent with clinical observations, sensitivity to one agent does not necessarily predict sensitivity to other agents, especially when they induce a different spectrum of DNA damage (18,19).

The relationships among sensitivities to the four DNA damaging agents and previously measured repair capacity in five pathways: NHEJ, HR, MMR, NER and MGMT were evaluated

in the lymphoblastoid cell lines (**Fig. 1B**; **Table S1, S2**) (8). The data in **Fig. 1B** have been Z-scored to emphasize repair capacity variation among the 24 cell lines. No statistically significant linear correlations were observed among repair capacities for the 5 reporters (**Table S3, S4**), and each cell line has a unique constellation of repair capacities. MGMT activity has previously been shown to correlate with MNNG sensitivity in the 24 cell lines (12). The present study, with MGMT activity measured by FM-HCR, confirms this result, and further finds that MGMT activity correlates significantly with sensitivity to killing by TMZ and BCNU (**Fig. S3**). Pairwise linear correlations between DRC in each individual pathway and sensitivity to killing with DNA damaging agents were evaluated based on p-values calculated for linear trends, as well as the value of the correlation coefficient R (**Table S5, S6**). A majority of the correlations were found to be stronger when calculated for sensitivity versus the logarithm of DNA repair capacity. All modeling work was thus carried out using the base 10 logarithm of repair capacities as the predictor variable.

Mathematical models were developed to test whether DRC in multiple pathways is correlated to sensitivity to DNA damaging agents (**Fig. S1**). A simple linear model ($Y = mx + b$) relates MNNG sensitivity to MGMT activity (model #1 in **Table 1**). The positive slope in this model indicates that a 1 standard deviation increase in MGMT activity leads to a 13 point increase in % control growth (**Fig. S2**). This relationship was expected based on the inverse correlation between MGMT activity and MNNG sensitivity that has been observed previously (12), and presently in **Tables S4** and **S5** and **Fig. S3**. A linear model relating MGMT activity to TMZ sensitivity similarly indicated that an increase in MGMT activity leads to an increase in TMZ resistance (model #4 in **Table 1**). A MATLAB script, available in the supplemental information, was used next to generate a multiple linear regression (MLR) model relating Z-scored DRC data in all 5 DNA repair pathways to cytotoxicity induced by each of the four agents, MNNG, TMZ, BCNU and MMS. In these models (#2, #5, #7, and #9 in **Table 1**), DRC

values for each pathway in the 24 cell lines in **Fig. 1B** served as the 5 (independent) predictor variables, and the sensitivity of the same 24 cell lines (reported as % control growth in **Fig. 1A)** was the (dependent) response variable. The resulting MLR models take the form given below in **Eqn. 2**,

$$Y = m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 + m_5x_5 + b \qquad (2)$$

where *Y* represents the response variable (MNNG sensitivity) reported as % Control Growth; x_i, are the DNA repair capacity predictor variables for the five substrates reported as z-scored log % reporter expression, m_i are the slopes along the corresponding dimensions, and *b* is a constant that represents the *Y* intercept. These models (#2, #5, #7, and #9) are denoted as "MLR" in **Table 1**. The slopes and y-intercepts of these models are interpreted in the same manner as described above for model #1. For all models, the goodness of fit is reported as the R^2 parameter, and predictive accuracy is reported as the R^2_{LOO} parameter for predicted sensitivity versus observed sensitivity, calculated from exhaustive leave-one-out cross-validation (**Table 1, Fig. S4**). An additional parameter in **Table 1**, area under the curve (AUC), reports how accurately each model predicts TMZ induced cell killing from DRC for a panel of PDX models of GBM. The application of the models to predict TMZ sensitivity in GBM is discussed in more detail below.

The models were further refined using a second MATLAB script (available in the supplemental information) running the LASSO algorithm (13) to select combinations of DRC predictor variables to be used for building multiple linear regression (MLR) models of the form given in **Eqn. 3**.

$$Y = m_1x_1 + m_2x_2 \ldots + m_nx_n + b \qquad (3)$$

The number and identity of predictor variables (n (max 5)) used to build the model is established from the results of the LASSO algorithm, which minimizes over-fitting by excluding variables that do not improve model performance under cross-validation (13). Models obtained using this approach (#3, #6, #8, and #10) for each DNA damaging agent are denoted as "MLR_L" in **Table 1**. The TMZ MLR_L retained 4 pathways (MGMT, HR, MMR, and NER); the MNNG MLR_L model retained 3 pathways (MGMT, HR, and MMR). The MLR_L models for BCNU and MMS retained only a single pathway (MGMT and HR, respectively). As expected, for each alkylating agent, the predictive accuracy of the MLR_L models was improved relative to the predictive accuracy of the MLR models; this is reflected in higher R^2_{LOO} values (**Table 1**), which correspond to better fits in plots of predicted sensitivity versus observed sensitivity for the 24 lymphoblastoid cell lines under leave-one-out cross-validation (**Fig. S4**).

The use of Z-scored DRC data for generating the mathematical models allows for a straightforward interpretation of the slopes (m_i) that are associated with each model: the steeper the slope, the more strongly the model responds to a given number of standard deviations about the mean for that particular pathway. A slope of 0 would indicate a model in which drug sensitivity is entirely *independent* of the DNA repair pathway in question.

Sensitivity Models for the S_N1-type alkylating agents TMZ and MNNG:

The MLR models for the S_N1 alkylating agents TMZ and MNNG (#2 and #5) are similar in several respects (**Table 1**). First, the coefficients for MGMT and HR are large and positive in both models, indicating that the models are relatively sensitive to small changes in repair capacity in these pathways, and higher MGMT or HR repair capacity leads to a lower sensitivity to killing with either agent. Second, MMR takes a negative coefficient in both models, indicating that higher MMR capacity leads to higher sensitivity to killing. This aspect of

the modeling is borne out in the raw data; higher MMR activity in cell line #6 versus cell line #10 without major differences in other pathways is associated with higher MNNG sensitivity in cell line #6 (**Fig. 1**). Third, the NHEJ coefficient is the smallest in both models, suggesting that this pathway plays the smallest role in sensitivity to S_N1 type alkylating agents. Finally, the NER coefficient is positive in both models, indicating that higher NER capacity is associated with lower sensitivity to S_N1 alkylating agents. Both the MNNG and TMZ MLR_L models (#3 and #6) exclude the NHEJ pathway, further indicating that this pathway does not contribute significantly to sensitivity to either agent for this panel of cells. The MLR_L models also both retain MGMT, MMR and HR, but differ in that the TMZ model also includes NER. To test whether the differences in MLR_L models (#3 and #6) were due to the inclusion of 5 additional cell lines in model #6, an MNNG sensitivity model was generated for the same subset of 19 cell lines for which the TMZ models were generated (not shown). Like model #3, this MLR_L model also did not include the NER pathway. We infer that the differences in models #3 and #6 are not due to the use of a subset of cell lines as the training population for the TMZ model. As detailed in the supplemental information, these models are consistent with previously established roles for the MGMT, MMR, HR and NER pathways affecting the sensitivity of cells to killing with S_N1-type alkylating agents (**Fig. 2**), although the quantitative contributions of each pathway have not previously been determined.

Sensitivity Models for BCNU and MMS

The BCNU and MMS sensitivity models are distinguished in several respects from the S_N1 alkylating agent sensitivity models. For both the BCNU and MMS MLR models (#7 and #9), the MMR coefficient was small (**Table 1**), and was not retained in the MLR_L models (#8 and #10), indicating that MMR capacity does not play a major role in determining cellular sensitivity in the models. Both MLR_L models retain only a single DNA repair pathway, and the

MMS MLR_L model is the only model that excludes the MGMT pathway. The MLR_L models indicate that HR plays a significant role in determining MMS sensitivity, whereas MGMT activity is the major contributor to BCNU sensitivity. As with the TMZ and MNNG sensitivity models, models #7-10 are consistent with previously established roles for DRC affecting the cellular MMS and BCNU sensitivity (see supplemental information).

Measuring DRC in GBM

The TMZ and MNNG models suggest that relatively small changes in MGMT, MMR, HR and NER could be responsible for changes in sensitivity to killing with S_N1-type alkylating agents in cancer cells. To test whether relationships between DRC and TMZ sensitivity in lymphoblastoid cell lines also apply to GBM, we measured changes in DRC that accompany TMZ resistance acquired *in vivo* for several paired sensitive and resistant PDX models of GBM (**Fig. 3, Supplemental Table S7**), as well as one PDX model (GBM46) innately TMZ resistant model (9). The parental PDX models (GBM12, GBM14, GBM22, GBM39, GBM59, and GBM46) were previously classified as “resistant” or “sensitive” to TMZ based on whether TMZ treatment produces a statistically significant survival benefit *in vivo* (20), and statistically significant increases in TMZ resistance in the PDX models (GBM12TMZ_3080, GBM12TMZ_5476, GBM14TMZ, GBM39TMZ, and GBM59_TMZ) have been measured *in vitro* either by neurosphere formation or by cell proliferation assays (21).

Strikingly, we observe statistically significant changes in DRC that are consistent with the relationship between DRC and S_N1-type alkylating agent sensitivity suggested by the MNNG and TMZ sensitivity models (**Table 1**). In cell lines that acquired TMZ resistance, we observed increased MGMT activity (GBM12/GBM12TMZ_3080), increased HR capacity (GBM14/GBM14TMZ and GBM22/GBM22TMZ), or decreased MMR capacity (GBM22/GBM22TMZ and GBM39/GBM39TMZ) relative to the TMZ sensitive parental cell lines

(**Fig. 3A**). Statistically significant changes in NHEJ (GBM39/GBM39TMZ) were also observed (**Fig. 3A**). Although the differences were not statistically significant, NER activity trended higher in all of the acquired resistance models relative to their corresponding parental lines, except for GBM14. A trend toward reduced MMR and elevated HR capacity was also observed in GBM59TMZ relative to GBM59, however these apparent changes likewise did not reach statistical significance (Fig 3B). Overall, the data indicate that the relationship between DRC and sensitivity to TMZ and MNNG in lymphoblastoid cell lines (**Table 1**) are applicable to TMZ sensitivity in GBM.

Applying MLR to predict sensitivity to TMZ in GBM

The ability of mathematical models to classify PDX models of GBM as TMZ sensitive or TMZ resistant was tested. ROC plots for a subset of the models are presented in **Fig. 4A**. The MLR_L model trained using TMZ sensitivity in lymphoblastoid cell lines (#6) is the best predictor of TMZ sensitivity in GBM (AUC = 0.86). The MLR_L model generated from DRC for sensitivity to MNNG (#3) yielded an AUC of 0.74. Predictions using a model for MMS sensitivity yielded an AUC = 0.71 (model #10), and a model generated for BCNU sensitivity (model #8) barely outperformed random chance for this panel of PDX models (AUC = 0.54).

The best model (#6) also correctly predicts TMZ resistance in the acquired resistance models relative to the corresponding parental cell lines in all of the paired PDX models (**Fig. S5**). Importantly, using DRC data for multiple pathways in GBM cell lines improves the predictive accuracy for TMZ sensitivity models versus those that include only the MGMT pathway. MGMT activity measured by FM-HCR was correlated to methylation status (R=0.61, **Fig S6**), but the extent of promoter methylation varied over a ~10-fold range among GBM lines that all had very low MGMT activity, consistent with previous observations indicating that widely used MGMT tumor promoter methylation assays may not accurately predict MGMT

transcript levels and activity (22). When cells are ranked according to TMZ sensitivity as predicted by model #4, which includes only MGMT activity (**Fig. 4B**), two TMZ resistant cell lines, GBM22TMZ and GBM39TMZ, are predicted to be the most sensitive, with predicted % control growth ~ 10%, and AUC = 0.54. By contrast, when PDX models are ranked according to TMZ sensitivity predicted by model #6, which takes MGMT, MMR, HR and NER into account, the predicted response across the GBM cell lines is much more accurate relative to previously measured sensitivity (20,21); the improved accuracy can be appreciated both qualitatively (**Fig. 4C**) and quantitatively (AUC = 0.86).

Using DRC-based modeling to Predict Effectiveness of Therapy with TMZ in vivo

The ultimate goal of predicting the sensitivity of GBM cells to S_N1 alkylating agents is to generate a metric that could potentially guide cancer therapy decisions. This goal requires that repair capacity *in vitro* predicts therapeutic response *in vivo*, where the sensitivity of cancer cells to DNA damaging therapies can be strongly influenced by tumor microenvironment (23). The PDX models for which we have measured DRC profiles have previously been shown to exhibit a broad range of responsiveness to *in vivo* TMZ therapy, measured by relative survival prolongation of tumor bearing mice treated with TMZ versus tumor bearing mice receiving placebo (21). Relative survival prolongation is calculated by dividing the median survival time for treated mice by the median survival time for untreated mice; larger values reflect longer survival in treated mice relative to untreated mice.

To test whether the DRC profiles measured using FM-HCR in cultured cells predict TMZ therapy effectiveness *in vivo*, relative survival prolongation in TMZ treated PDX mouse models was compared against TMZ sensitivities predicted using the models in **Table 1**. Sensitivities predicted using a model that includes MGMT activity alone (model #4) are only weakly correlated to relative survival prolongation, and PDX models with low MGMT expression (black

dots in **Fig. 5A**) range from highly responsive to completely refractory to treatment with TMZ. a A stronger correlation is observed between survival prolongation and predicted sensitivity calculated for the PDX models using a TMZ sensitivity model that includes MGMT, MMR, HR and NER (model #6), (**Fig. 5B**). These results indicate that relatively small changes in HR, MMR and NER capacity collectively contribute to changes in GBM TMZ sensitivity, and that the resistance mechanisms detailed in **Fig. 2** operate *in vivo*. When the analysis is restricted to MGMT-deficient PDX models, the correlation between relative survival prolongation and the combined DRC score is even stronger (**Fig. 5C**). The strong correlation in MGMT deficient cells underscores the importance of DNA repair pathways other than MGMT in determining the effectiveness of TMZ therapy. Importantly, these results demonstrate that DRC measurements carried out in cells *in vitro* can at least partially predict therapeutic response *in vivo*.

Discussion

The ability to choose the optimal therapy for each patient would have tremendous ramifications for cancer treatment. Advance knowledge of whether a particular therapy will be effective would increase the likelihood of achieving the primary goal of killing cancer cells, and avoid subjecting patients to unnecessary treatments that cause serious side effects and increase the risk of additional cancers (24-26). DRC continues to emerge as an important determinant for the response of cancer cells to therapy (27), making accurate DRC assessments a key goal for precision medicine (7).

The potential to predict TMZ sensitivity from DRC in multiple pathways underscores need to measure MGMT, HR, NER and MMR activity in order to make the most robust DNA repair-based predictions regarding the effectiveness of TMZ therapy. However, direct measurements of DNA repair activity in clinical samples has been limited because previous

methods were cumbersome, time consuming, incompatible with multiplexing, and in most cases did not report on the functional status of repair (7). Tumor MGMT status is considered a strong predictor of therapeutic response in GBM (28), and clinicians use MGMT promoter methylation assays or in some cases immunohistochemical analysis of excised tumor tissue to infer MGMT activity in GBM (29). Nevertheless, concerns remain about standard clinical MGMT assays that unlike FM-HCR, do not directly measure MGMT activity; immunohistochemical analysis is subject to individual bias (30), and although MGMT activity was, as expected, inversely correlated to MGMT promoter methylation in the GBM lines in this study (**Fig. S6**), the high degree of variability in the extent of promoter methylation among MGMT-deficient GBM lines adds to existing evidence that promoter methylation status is not always a reliable measure of tumor MGMT activity (9,22). Similarly, microsatellite instability assays are used to assess MMR status in clinical samples, but this approach is not quantitative and may not report minor changes in MMR capacity that are sufficient for TMZ resistance (31,32). Several approaches have been taken to estimate HR capacity in cancer cells (33), but they rely on transcriptional profiling and mutational signatures that may not predict the functional status of the HR pathway, and are unlikely to detect small differences in functional activity. Finally, unscheduled DNA synthesis has long been used as a method of estimating NER capacity in patient cells (34), but is relatively laborious. By contrast, FM-HCR assays provide a rapid functional assessment of DNA repair capacity in key pathways that we have shown are useful predictors of TMZ sensitivity in lymphoblastoid and PDX lines.

The present work has been focused on modeling TMZ sensitivity in GBM, but the fundamental principles governing resistance to DNA damaging agents, particularly those driven by relative changes in DRC, are expected to be applicable to multiple agents and multiple cell types. Strikingly, a model trained using DNA repair capacity and TMZ sensitivity data for lymphoblastoid cells in culture accurately predicted TMZ sensitivity of GBM cells

treated in mouse PDX models. Additional cell-autonomous therapy resistance mechanisms, including changes in DNA damage signaling (35), damage detoxification mechanisms (36), and drug efflux pumps are expected to complement DNA repair-based resistance mechanisms (37). Single cell-based assays for processes that contribute to resistance, such as FM-HCR, hold the major advantage of being amenable to identifying small, therapy resistant subpopulations within a tumor, which can expand and emerge as a resistant tumor following treatment (38). Therefore, models integrating FM-HCR assays for DNA repair-based resistance mechanisms with cell-based functional assessments of other therapy resistance mechanisms hold promise for further enhancing predictions of therapy response and advancing personalized cancer treatment.

In summary, our results indicate that TMZ resistance in cancer cells is associated with differences in repair capacity in at least four DNA repair pathways (MMR, HR, NER and MGMT), and that a combined DNA repair score derived from FM-HCR assays provides a more robust prognostic biomarker than MGMT activity alone for TMZ therapy.

Acknowledgements

This work was supported by the National Institutes of Health (DP1-ES022576 and U54-CA112967). The authors thank Rebecca Fry, Brad Hogan, and Meriem Sefta (MIT) for help with measuring MMS sensitivity and TMZ sensitivity in the Coriell Cell lines. The authors further acknowledge Mark Schroeder and Brett Carlson (Mayo Clinic) for propagating xenograft lines.

References

1. Fu D, Calvo JA, Samson LD. Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat Rev Cancer* **2012**;12:104-20.
2. Lindahl T, Wood RD. Quality control by DNA repair. *Science* **1999**;286:1897-905.

3. Hoeijmakers JHJ. Genome maintenance mechanisms for preventing cancer. *Nature* **2001**;411:366-74.
4. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New England Journal of Medicine* **2005**;352:987-96.
5. Colvin M. Alkylating Agents and Platinum Antitumor Compounds. In: Kufe DW, Pollock RE, Weichselbaum RR, Bast RC, Gansler TS, Holland JF, et al., editors. Holland-Frei Cancer Medicine. 6 ed. Volume 2. Hamilton (ON): BC Decker; 2003.
6. Pegg AE. Multifaceted Roles of Alkyltransferase and Related Proteins in DNA Repair, DNA Damage, Resistance to Chemotherapy, and Research Tools. *Chem Res Toxicol* **2011**;24:618-39.
7. Nagel ZD, Chaim IA, Samson LD. Inter-individual variation in DNA repair capacity: A need for multi-pathway functional assays to promote translational DNA repair research. *DNA Repair* **2014**;19:199-213.
8. Nagel ZD, Margulies CM, Chaim IA, McRee SK, Mazzucato P, Ahmad A, et al. Multiplexed DNA repair assays for multiple lesions and multiple doses via transcription inhibition and transcriptional mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America* **2014**;111:E1823-32.
9. Kitange GJ, Carlson BL, Schroeder MA, Grogan PT, Lamont JD, Decker PA, et al. Induction of MGMT expression is associated with temozolomide resistance in glioblastoma xenografts. *Neuro-Oncology* **2009**;11:281-91.
10. Valiathan C, McFaline JL, Samson LD. A rapid survival assay to measure drug-induced cytotoxicity and cell cycle effects. *DNA Repair* **2012**;11:92-98.
11. Sefta M. A study of inter-individual differences in the DNA damage response. Cambridge, MA: Massachusetts Institute of Technology; 2012.
12. Fry RC, Svensson JP, Valiathan C, Wang E, Hogan BJ, Bhattacharya S, et al. Genomic predictors of interindividual differences in response to DNA damaging agents. *Genes Dev* **2008**;22:2621-26.
13. Tibshirani R. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society Series B-Methodological* **1996**;58:267-88.
14. Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian journal of internal medicine* **2013**;4:627-35.
15. Collins FS, Brooks LD, Chakravarti A. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* **1998**;8:1229-31.
16. Newlands ES, Stevens MF, Wedge SR, Wheelhouse RT, Brock C. Temozolomide: a review of its discovery, chemical properties, pre-clinical development and clinical trials. *Cancer treatment reviews* **1997**;23:35-61.
17. Briegert M, Kaina B. Human monocytes, but not dendritic cells derived from them, are defective in base excision repair and hypersensitive to methylating agents. *Cancer Res* **2007**;67:26-31.
18. Fischer DS, Knobf MT, Durivage HJ. The Cancer Therapy Handbook. St. Louis, MO: Mosby; 1997.
19. Wen PY, Kesari S. Malignant gliomas in adults. *New England Journal of Medicine* **2008**;359:492-507.
20. Carlson BL, Grogan PT, Mladek AC, Schroeder MA, Kitange GJ, Decker PA, et al. Radiosensitizing effects of temozolomide observed in vivo only in a subset of O6-methylguanine-DNA methyltransferase methylated glioblastoma multiforme xenografts. *International journal of radiation oncology, biology, physics* **2009**;75:212-9.

21. Kitange GJ, Mladek AC, Carlson BL, Schroeder MA, Pokorny JL, Cen L, et al. Inhibition of histone deacetylation potentiates the evolution of acquired temozolomide resistance linked to MGMT upregulation in glioblastoma xenografts. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2012**;18:4070-9.
22. Kreth S, Thon N, Eigenbrod S, Lutz J, Ledderose C, Egensperger R, et al. O-methylguanine-DNA methyltransferase (MGMT) mRNA expression predicts outcome in malignant glioma independent of MGMT promoter methylation. *PLoS One* **2011**;6:e17156.
23. Tredan O, Galmarini CM, Patel K, Tannock IF. Drug resistance and the solid tumor microenvironment. *J Natl Cancer Inst* **2007**;99:1441-54.
24. Leone G, Voso MT, Sica S, Morosetti R, Pagano L. Therapy related leukemias: Susceptibility, prevention and treatment. *Leukemia & Lymphoma* **2001**;41:255-76.
25. Wood ME, Vogel V, Ng A, Foxhall L, Goodwin P, Travis LB. Second malignant neoplasms: assessment and strategies for risk reduction. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **2012**;30:3734-45.
26. Oeffinger KC, Bhatia S. Second primary cancers in survivors of childhood cancer. *Lancet (London, England)* **2009**;374:1484-5.
27. Helleday T, Petermann E, Lundin C, Hodgson B, Sharma RA. DNA repair pathways as targets for cancer therapy. *Nat Rev Cancer* **2008**;8:193-204.
28. Hegi ME, Diserens A, Gorlia T, Hamou M, de Tribolet N, Weller M, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *New England Journal of Medicine* **2005**;352:997-1003.
29. Stupp R, Brada M, van den Bent MJ, Tonn JC, Pentheroudakis G. High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **2014**;25 Suppl 3:iii93-101.
30. Preusser M, Janzer RC, Felsberg J, Reifenberger G, Hamou MF, Diserens AC, et al. Anti-O6-methylguanine-methyltransferase (MGMT) immunohistochemistry in glioblastoma multiforme: Observer variability and lack of association with patient survival impede its use as clinical biomarker. *Brain Pathology* **2008**;18:520-32.
31. McFaline-Figueroa JL, Braun CJ, Stanciu M, Nagel ZD, Mazzucato P, Sangaraju D, et al. Minor changes in expression of the mismatch repair protein MSH2 exert a major impact on glioblastoma response to temozolomide. *Cancer Res* **2015**.
32. Shin KH, Park JG. Microsatellite instability is associated with genetic alteration but not with low levels of expression of the human mismatch repair proteins hMSH2 and hMLH1. *Eur J Cancer* **2000**;36:925-31.
33. Timms KM, Abkevich V, Hughes E, Neff C, Reid J, Morris B, et al. Association of BRCA1/2 defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. *Breast cancer research : BCR* **2014**;16:475.
34. Kelly CM, Latimer JJ. Unscheduled DNA synthesis: a functional assay for global genomic nucleotide excision repair. *Methods Mol Biol* **2005**;291:303-20.
35. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* **2011**;144:646-74.
36. Rocha CR, Garcia CC, Vieira DB, Quinet A, de Andrade-Lima LC, Munford V, et al. Glutathione depletion sensitizes cisplatin- and temozolomide-resistant glioma cells in vitro and in vivo. *Cell death & disease* **2014**;5:e1505.
37. Noguchi K, Katayama K, Sugimoto Y. Human ABC transporter ABCG2/BCRP expression in chemoresistance: basic and clinical perspectives for molecular cancer therapeutics. *Pharmacogenomics and personalized medicine* **2014**;7:53-64.

38. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **2012**;481:506-10.

Figure Legends

Fig. 1. DRC and Sensitivity data for Coriell lymphoblastoid cell line panel. a) 24 lymphoblastoid cell lines organized (from left to right) in descending order of sensitivity to 4 different DNA damaging agents. The colors assigned to bars for each cell line in the top panel (MNNG) by quartile are maintained in the TMZ, BCNU, and MMS panels. Sensitivity to MNNG, and MMS were measured at 72 hours. Sensitivity to BCNU and TMZ were measured at 4 doubling times post-exposure. b) Heat map representing Z-scores for relative DNA repair capacity among the 24 lymphoblastoid cell lines in the five pathways listed in the column headings. Purple indicates cell lines with repair capacity below the mean; yellow indicates cell lines with repair capacity above the mean, and white represents cell lines with repair capacity at or near the mean. Z-units are reported at the right side of the figure. Cell lines are ordered by sensitivity to MNNG.

Fig. 2. Model for DNA TMZ resistance mechanisms mediated by DNA repair capacity. Double stranded DNA is represented by parallel black lines. A normal G:C base pair is indicated at the top of the figure. Strand breaks are indicated as gaps in the black lines. Gray arrows and text indicate DNA repair processes that mediate sensitivity and resistance to DNA methylation damage.

Fig. 3. DNA repair capacity measured using FM-HCR in GBM cells. a) Statistically significant changes in DRC in PDX models of acquired TMZ resistance. Note that for the MGMT assay,

lower % Reporter Expression corresponds to higher MGMT activity. Error bars represent the standard deviation of at least 3 measurements. b) A heat map representing Z-scored DNA repair capacity in 5 pathways for all 12 GBM models. Purple indicates cell lines with repair capacity below the mean; yellow indicates cell lines with repair capacity above the mean, and white represents cell lines with repair capacity at or near the mean.

Fig. 4. Classification of GBM PDX models as TMZ sensitive or TMZ resistant. a) Plots showing ROC curves (black lines) for sensitivity models for BCNU (model #8); MMS (#10); MNNG (#3); TMZ (#6). The numbers refer to models detailed in **Table 1**. AUC is calculated from the area under the ROC curves, and can adopt values between 0 (all samples are classified incorrectly) and 1 (all samples are classified correctly). The dotted red line represents a theoretical ROC curve for a model that classifies samples randomly (AUC = 0.5). b) PDX models organized in ascending order of predicted percent control growth as calculated from model #4. Lower predicted percent control growth represents higher predicted sensitivity to TMZ. c) PDX models organized in ascending order of predicted percent control growth as calculated from model #6. TMZ sensitivity has been determined previously for the PDX models; black bars indicate PDX models that TMZ resistant (TMZ^R), and red bars indicate PDX models that are TMZ sensitive (TMZ^S). A Mann-Whitney U test indicated that the two populations (TMZ^R and TMZ^S) were significantly different in panel c ($p = 0.03$), but not in panel b.

Fig. 5. Correlation between relative survival prolongation for PDX mouse models treated with TMZ and predicted percent control growth for PDX models treated with TMZ calculated from a) model #4 or b,c) model #6. In panel c, PDX models are restricted to those expressing low levels of MGMT (black dots). Cells with low MGMT levels were defined as those for which

average % reporter expression for the MGMT assay was at least 50%; this value of % reporter expression was observed previously for a cell line that is known to be MGMT deficient (TK6). PDX models with higher levels of MGMT reporter expression are represented with open squares.

Table 1. Linear models relating sensitivity to DNA damaging agents to DNA repair capacity in multiple pathways. The models are numbered under column heading "#". R^2_{FIT} = the goodness of fit for each model; R^2_{FIT} can adopt values between 0 (no fit at all) and 1 (perfect fit). R^2_{L00} is a measure of the predictive accuracy for each model, and is calculated from exhaustive leave-one-out cross-validation presented graphically in **Fig. S2**. Once again, R^2_{L00} can adopt values between 0 (no fit at all) and 1 (perfect fit). AUC is a statistic that measures the performance of the models when they are used to classify the PDX models as sensitive versus resistant. AUC can adopt values between 0 (all samples are classified incorrectly) and 1 (all samples are classified correctly). AUC is calculated from receiver operating characteristic curves (**Fig. 5**) generated using the sensitivity of PDX models calculated from the respective models and the actual status of each GBM model (sensitive or resistant).

Agent	#	Model	Equation	R^2_{FIT}	R^2_{LOO}	AUC
MNNG	1	MGMT	$\mathbf{13.3} * MGMT + \mathbf{59}$	0.41	0.33	0.54
	2	MLR	$\mathbf{13.8} * MGMT - \mathbf{9.3} * MMR + \mathbf{6.8} * HR + \mathbf{4.4} * NER - \mathbf{2.3} * NHEJ + \mathbf{59}$	0.68	0.49	0.89
	3	MLR_L	$\mathbf{13.3} * MGMT + \mathbf{8.1} * HR - \mathbf{7.0} * MMR + \mathbf{59}$	0.63	0.50	0.74
TMZ	4	MGMT	$\mathbf{19.0} * MGMT + \mathbf{71}$	0.61	0.50	0.54
	5	MLR	$\mathbf{17.5} * MGMT + \mathbf{8.5} * NER + \mathbf{8.3} * HR - \mathbf{3.6} * MMR - \mathbf{0.2} * NHEJ + \mathbf{71}$	0.90	0.81	0.86
	6	MLR_L	$\mathbf{17.4} * MGMT + \mathbf{8.5} * NER + \mathbf{8.2} * HR - \mathbf{3.5} * MMR + \mathbf{71}$	0.90	0.84	0.86
BCNU	7	MLR	$\mathbf{10.4} * MGMT + \mathbf{2.5} * NER + \mathbf{2.2} * NHEJ - \mathbf{1.3} * MMR + \mathbf{0.3} * HR + \mathbf{33}$	0.54	0.19	0.80
	8	MLR_L	$\mathbf{10.6} * MGMT + \mathbf{33}$	0.49	0.40	0.54
MMS	9	MLR	$\mathbf{6.9} * HR + \mathbf{5.8} * NHEJ + \mathbf{5.6} * NER + \mathbf{5.2} * MGMT - \mathbf{1.5} * MMR + \mathbf{44}$	0.38	0.06	0.71
	10	MLR_L	$\mathbf{8.9} * HR + \mathbf{44}$	0.17	0.07	0.71

Figure 1

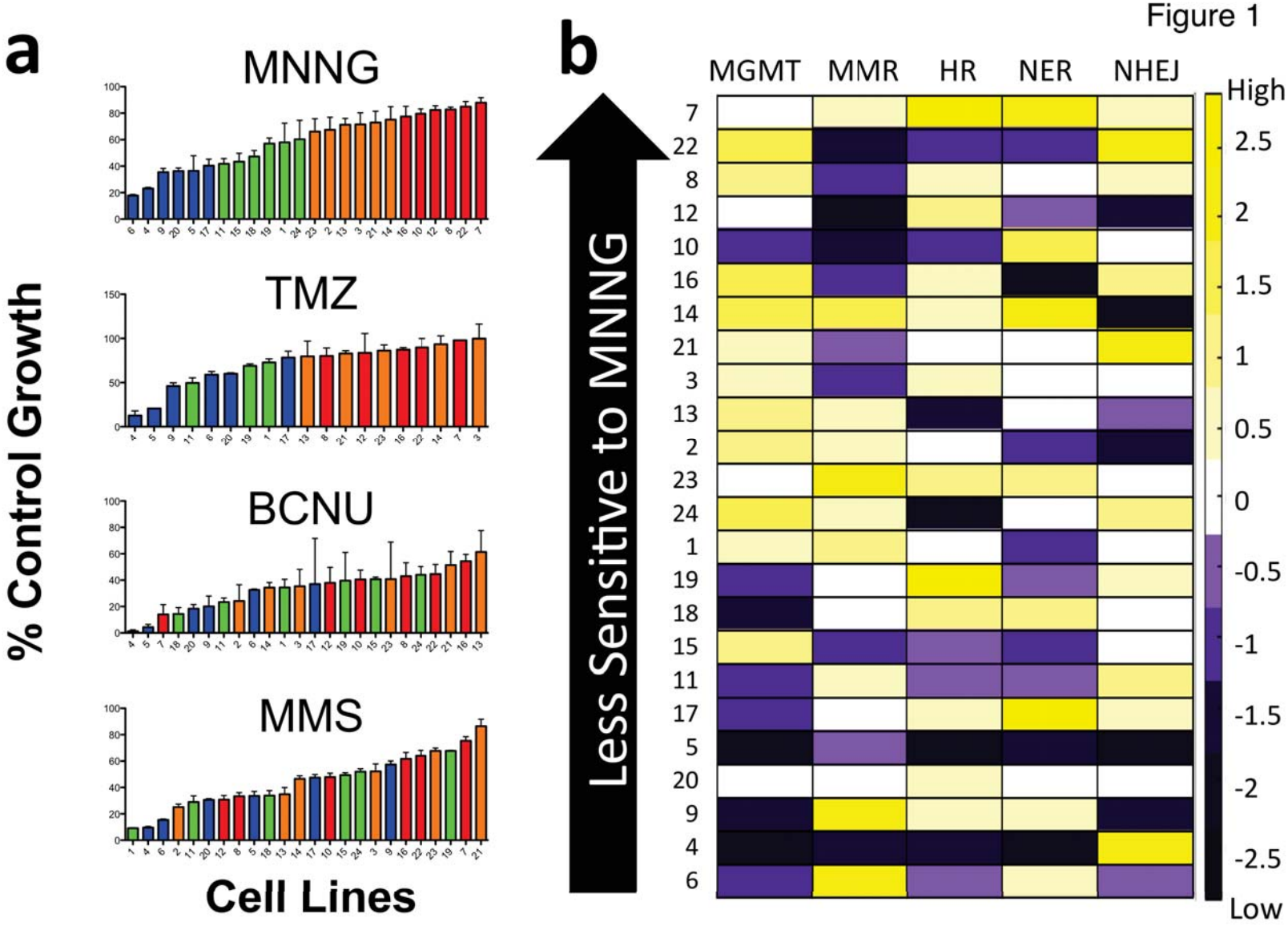
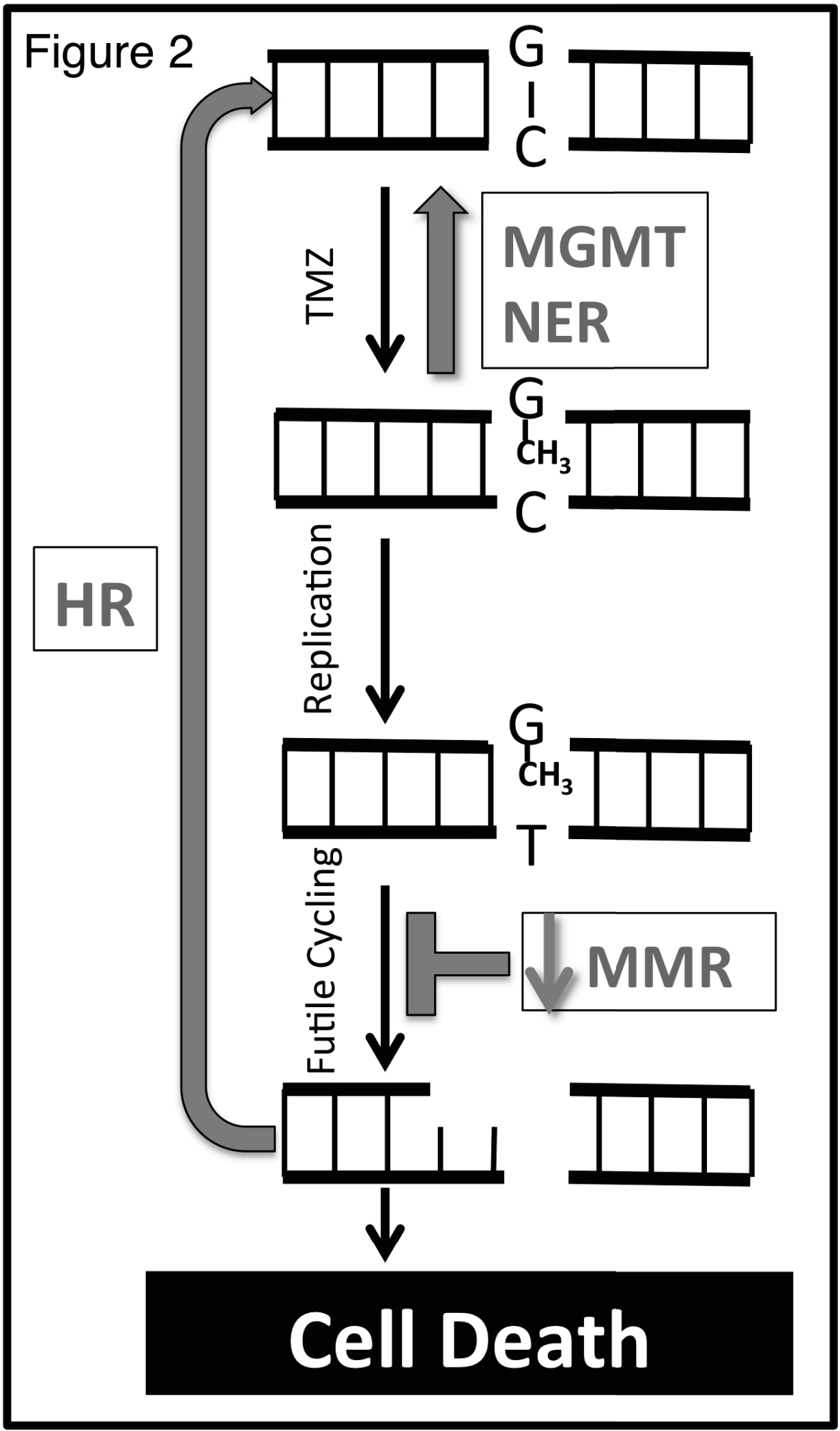

a
MNNG
TMZ
BCNU
MMS
% Control Growth
Cell Lines
b
MGMT
MMR
HR
NER
NHEJ
Less Sensitive to MNNG
High
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
Low

Figure 2


G
C
TMZ
MGMT
NER
G
CH3
C
HR
Replication
G
CH3
T
Futile Cycling
MMR
Cell Death

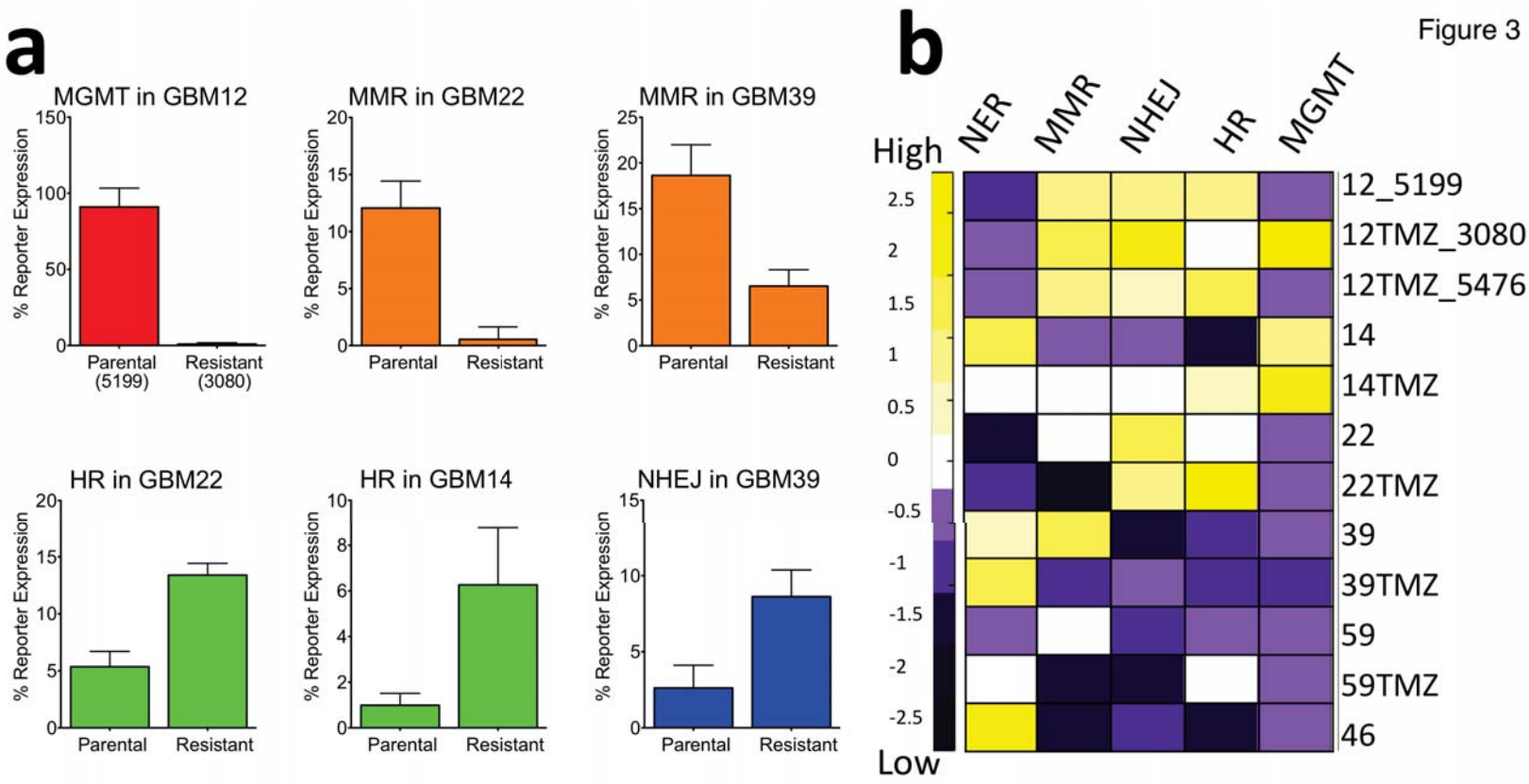

a
MGMT in GBM12
% Reporter Expression
150
100
50
0
Parental (5199)
Resistant (3080)
MMR in GBM22
% Reporter Expression
20
15
10
5
0
Parental
Resistant
MMR in GBM39
% Reporter Expression
25
20
15
10
5
0
Parental
Resistant
HR in GBM22
% Reporter Expression
20
15
10
5
0
Parental
Resistant
HR in GBM14
% Reporter Expression
10
8
6
4
2
0
Parental
Resistant
NHEJ in GBM39
% Reporter Expression
15
10
5
0
Parental
Resistant
b
NER
MMR
NHEJ
HR
MGMT
High
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
Low
12_5199
12TMZ_3080
12TMZ_5476
14
14TMZ
22
22TMZ
39
39TMZ
59
59TMZ
46

Figure 3

Figure 4

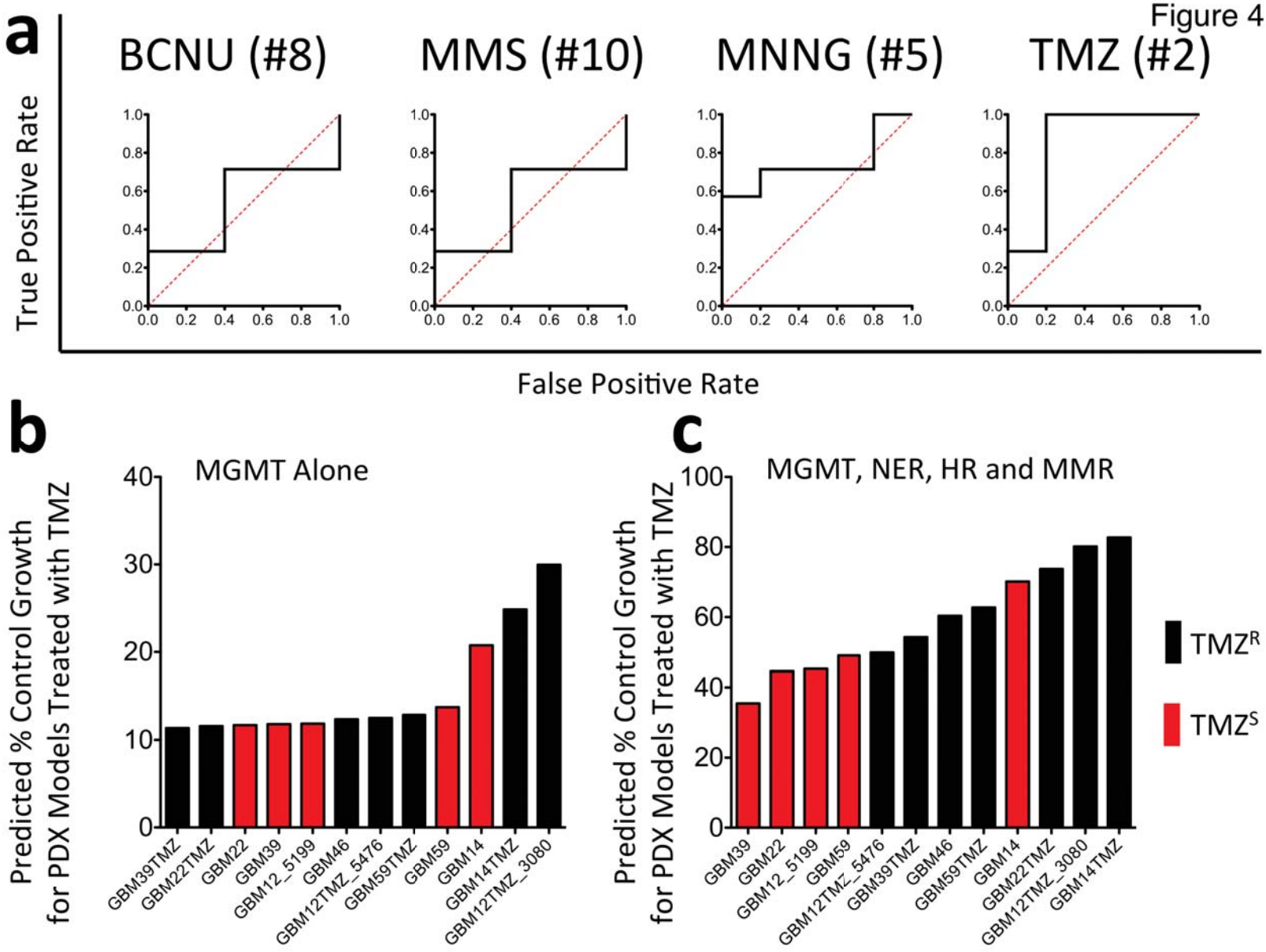
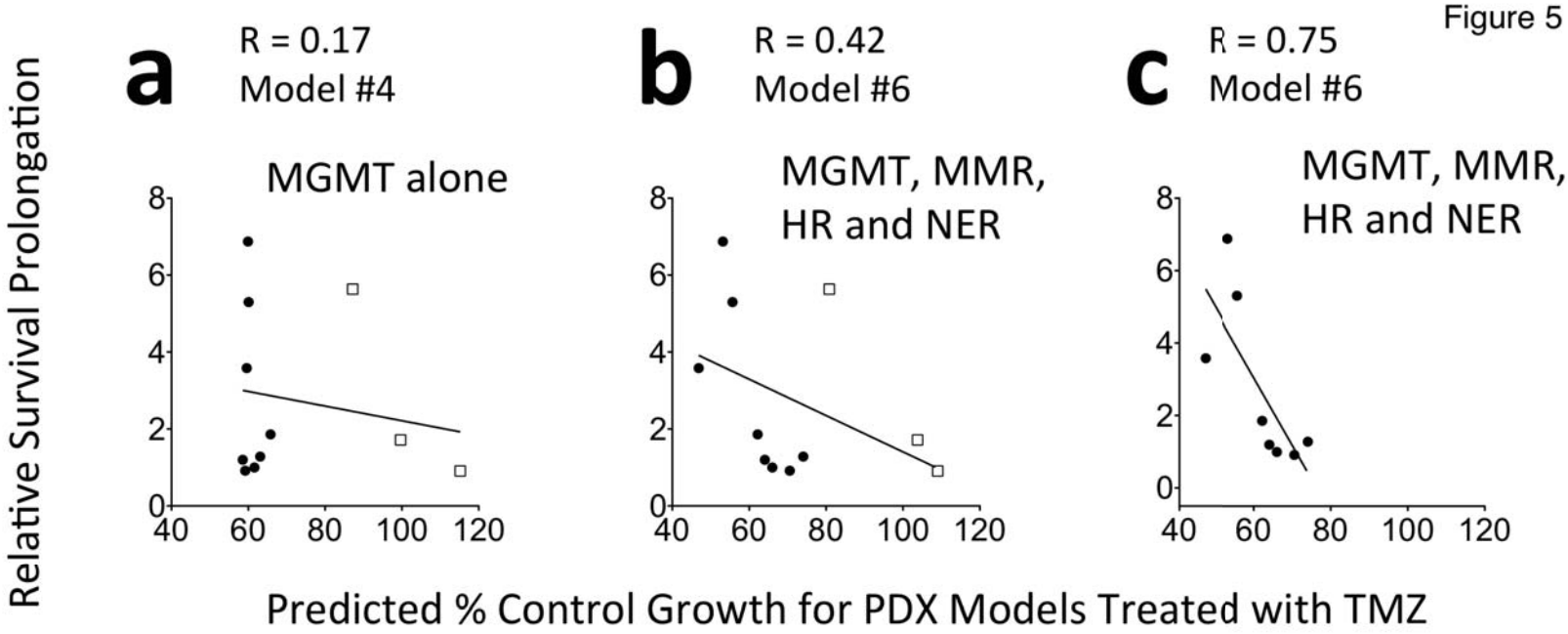

a
BCNU (#8)
MMS (#10)
MNNG (#5)
TMZ (#2)
True Positive Rate
False Positive Rate
b
MGMT Alone
Predicted % Control Growth for PDX Models Treated with TMZ
GBM39TMZ
GBM22TMZ
GBM22
GBM39
GBM12_5199
GBM46
GBM12TMZ_5476
GBM59TMZ
GBM59
GBM14
GBM14TMZ
GBM12TMZ_3080
c
MGMT, NER, HR and MMR
Predicted % Control Growth for PDX Models Treated with TMZ
GBM39
GBM22
GBM12_5199
GBM59
GBM12TMZ_5476
GBM39TMZ
GBM46
GBM59TMZ
GBM14
GBM22TMZ
GBM12TMZ_3080
GBM14TMZ
TMZR
TMZS

Figure 5


a
R = 0.17
Model #4
MGMT alone
b
R = 0.42
Model #6
MGMT, MMR, HR and NER
c
R = 0.75
Model #6
MGMT, MMR, HR and NER
Relative Survival Prolongation
Predicted % Control Growth for PDX Models Treated with TMZ
0
2
4
6
8
40
60
80
100
120