Supplemental Information for:

Insights into Gemcitabine Resistance and the Potential for Therapeutic Monitoring

Teklab Gebregiworgis,[a,†] Fatema Bhinderwala,[a,b,†] Vinee Purohit,[c] Nina V. Chaika,[c] Pankaj K. Singh,[c,d,e,f,] and Robert Powers[a,b,*]

[a]*Department of Chemistry, University of Nebraska - Lincoln, Nebraska, USA. 68588* [b]*Nebraska Center for Integrated Biomolecular Communication*

[c] *The Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Omaha, Nebraska, USA. 68198*

[d]*Department of Biochemistry and Molecular biology, University of Nebraska Medical Center, Omaha, Nebraska, USA. 68198*

[e]*Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, Nebraska, USA. 68198*

[f]*Department of Genetics, Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, Nebraska, USA. 68198*

†Equal contribution

*To whom correspondence should be addressed:

Robert Powers
University of Nebraska-Lincoln
Department of Chemistry
722 Hamilton Hall
Lincoln, NE 68588-0304
Email: rpowers3@unl.edu
Phone: (402) 472-3039
Fax: (402) 472-9402

Table of Contents:

Methods and materials

Cell culture for developing GemR cell lines

T3M4 and Capan-1 (ATCC) wild type (WT) pancreatic cancer cells were cultured in DMEM with increasing concentrations of gemcitabine over a period of about 6 months. The resistance status of the cells at each gemcitabine dose was determined by calculating an IC_{50} using the MTT cytotoxicity assays. At the end of the 6-month treatment, the cell lines generated (GemR) were approximately 500-1000 fold resistant to gemcitabine as compared to the parental (WT) cells.

Cell cultures for metabolomics

Wild type T3M4 and Capan-1 cells (WT), which are gemcitabine sensitive cells, and the corresponding gemcitabine resistant cells (GemR) were cultured with or without the addition of gemcitabine for the NMR metabolomics studies. A total of four cell culture groups were prepared for both the T3M4 and Capan-1 cell types that consisted of: (1) untreated (WT) cells, (2) WT cells treated with 10 nM gemcitabine for a period of 12 hours (WT+), (3) untreated GemR cells, and (4) GemR cells treated with 10 nm gemcitabine for 12 hours (GemR+). For the 1D ^{1}H NMR dataset, ten replicates of T3M4 and eight replicates of Capan-1 cell cultures per group were grown in DMEM with 0.5 mM ^{12}C-glucose. For the 2D ^{1}H- ^{13}C HSQC NMR dataset, three replicate Capan-1 cell cultures per group were grown in DMEM with 0.5 mM $^{13}C_6$-glucose.

Metabolite extraction

For each of the cell culture replicates, 1 x 10^6 cells were cultured and harvested at 70-80% confluence in a 100 mm petri dish. The media was then removed by aspiration upon harvesting, and the cells were washed twice with PBS at pH 7.2. The cells were then lysed by the addition of

1 mL of an 80% methanol:water mixture at -80^{o}C followed by incubation for 15 minutes in a -80^{o}C freezer. The lysed cells were then transferred to a 1 mL Eppendorf tube using a cell scraper. The cell lysate was centrifuged at 16,200 *g* for 5 minutes and the supernatant was collected and transferred to a clean Eppendorf tube. 250 µL of nanopure water (Nanopure, Dubuque, IA) was added to the cell debris, the sample was mixed by pipetting, and then centrifuged as before. The two supernatants were combined, and the sample was dried by vacuum evaporation (SpeedVac® Plus, Savant, Thermo Scientific, Waltham, MA) followed by freeze-drying (Labconco, Kansas City, MO).

NMR sample preparation and data collection

The dried metabolite extracts were reconstituted in 600 µL of 50 mM phosphate buffer in 99.8% D_2O (Isotec, St. Louis, MO) at pH 7.2 (uncorrected). 50 µM of 3-(tetramethysilane) propionic acid-2,2,3,3-d_4 (TMSP-d_4) was added as a chemical shift reference. The solution was mixed by gentle vortexing, centrifuged at 16,200 *g* for 5 minutes, and then transferred to a 5 mm NMR tube. The NMR data was collected using a Bruker AVANCE DRX 500 MHz spectrometer equipped with 5 mm triple-resonance cryogenic probe (^{1}H, ^{13}C, and ^{15}N) with a Z-axis gradient. The sample collection was automated using a BACS-120 sample changer, ATM (automatic tuning and matching), and Bruker IconNMR™ software. The 1D ^{1}H NMR spectra were collected at 300K with 32K data points, 128 scans, 16 dummy scans, and a spectral width of 5,483 Hz using an excitation sculpting pulse sequence (Nguyen et al. 2007). The 2D ^{1}H-^{13}C HSQC NMR experiments were collected with 2K data points and a spectrum width of 4,735 Hz in the direct dimension and 64 data points and a spectrum width of 17,607 Hz in the indirect dimension. Each 2D ^{1}H-^{13}C HSQC spectrum was collected at 300K with 64 scans, 16 dummy scans and a 1.5 s

relaxation delay. The NMR dataset (accession number in progress) was deposited in the Metabolomics Workbench (http://www.metabolomicsworkbench.org/) database.

NMR data analysis

Our MVAPACK software (http://bionmr.unl.edu/mvapack.php) was used to analyze the NMR metabolomics data (Worley and Powers 2014). A principal component analysis (PCA) model was generated from the 1D ^{1}H NMR dataset. The 1D ^{1}H NMR spectra were Fourier transformed, automatically phased, and binned using adaptive intelligent binning (De Meyer et al. 2008; Worley and Powers 2015). The dataset was normalized using probabilistic quotient normalization (Dieterle et al. 2006) the noise was removed as previously described (Halouska and Powers 2006), and the dataset was scaled using Pareto scaling. The normalized and scaled NMR bins were then used to generate receiver operator characteristics (ROC) models with linear-support vector machine (SVM) as the classification method and SVM as the feature ranking method built into MetaboAnalyst (https://www.metaboanalyst.ca/) (Chong et al. 2018). The tree diagrams were generated from the PCA scores plot using our PCA/PLS-DA Utilities as previously described (Worley and Powers 2013; Werth et al. 2010). Five 1D ^{1}H NMR spectra from T3M4 cell lysates were identified as outliers and excluded from the statistical analysis. As apparent from the PCA scores plot using the entire T3M4 data set (Figure S4), one spectrum is outside the 95% confidence ellipse for the entire PCA model, one spectrum does not cluster with any group, and three spectra have apparently swapped labels. The orthogonal projections to latent structures discriminant analysis (OPLS-DA) scores plot and back-scaled loadings were generated using the entire NMR spectra following alignment with the icoshift algorithm (Savorani et al. 2010) and removal of solvent peaks (3.30-3.38 ppm for methanol and 4.75-4.85

ppm for water). The 2D ^{1}H-^{13}C HSQC NMR spectra were processed using NMRPipe and analyzed using NMRViewJ (Delaglio et al. 1995; Johnson and Blevins 1994). The Human Metabolomics Database (Wishart et al. 2013), Platform for RIKEN Metabolomics (Akiyama et al. 2008), Madison Metabolomics Consortium (Cui et al. 2008) and Chenomx NMR Suite 7.6 (http://www.chenomx.com/) were used for metabolite identification. Metabolites were assigned using 0.08 ppm and 0.25 ppm chemical shift error for ^{1}H and ^{13}C resonances, respectively. Accordingly, metabolites were identified to MSI level 2. Peak intensities were sum normalized using the total intensity of the spectrum.

References

Akiyama, K., Chikayama, E., Yuasa, H., Shimada, Y., Tohge, T., Shinozaki, K., et al. (2008). PRIMe: A Web site that assembles tools for metabolomics and transcriptomics. *In Silico Biol.*, *8*(3,4), 339-345.

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., et al. (2018). MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Research, 46*(W1), W486-W494, doi:10.1093/nar/gky310.

Cui, Q., Lewis, I. A., Hegeman, A. D., Anderson, M. E., Li, J., Schulte, C. F., et al. (2008). Metabolite identification via the Madison Metabolomics Consortium Database. *Nat. Biotechnol.*, *26*(2), 162-164.

De Meyer, T., Sinnaeve, D., Van Gasse, B., Tsiporkova, E., Rietzschel, E. R., De Buyzere, M. L., et al. (2008). NMR-Based Characterization of Metabolic Alterations in Hypertension Using an Adaptive, Intelligent Binning Algorithm. *Anal. Chem.*, *80*(10), 3783-3790, doi:10.1021/ac7025964.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR, 6*, 277-293, doi:10.1007/BF00197809.

Dieterle, F., Ross, A., Schlotterbeck, G., & Senn, H. (2006). Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in 1H NMR Metabonomics. *Analytical Chemistry, 78*(13), 4281-4290, doi:10.1021/ac051632c.

Halouska, S., & Powers, R. (2006). Negative impact of noise on the principal component analysis of NMR data. *J Magn Reson, 178*(1), 88-95, doi:10.1016/j.jmr.2005.08.016.

Johnson, B. A., & Blevins, R. A. (1994). NMRView, A Computer Program for the Visualization and Analysis of NMR Data. *J. Biomol. NMR, 4*, 603-614.

Nguyen, B. D., Meng, X., Donovan, K. J., & Shaka, A. J. (2007). SOGGY: Solvent-optimized double gradient spectroscopy for water suppression. A comparison with some existing techniques. *J. Magn. Reson., 184*(2), 263-274, doi:10.1016/j.jmr.2006.10.014.

Savorani, F., Tomasi, G., & Engelsen, S. B. (2010). icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. [10.1016/j.jmr.2009.11.012]. *J. Magn. Reson., 202*(2), 190-202, doi:10.1016/j.jmr.2009.11.012.

Werth, M. T., Halouska, S., Shortridge, M. D., Zhang, B., & Powers, R. (2010). Analysis of Metabolomic PCA Data using Tree Diagrams. *Analytical Biochemistry, 399*(1), 56-63.

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., et al. (2013). HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res, 41*(Database issue), D801-807, doi:10.1093/nar/gks1065.

Worley, B., & Powers, R. (2013). Multivariate Analysis in Metabolomics. *Curr Metabolomics, 1*(1), 92-107, doi:10.2174/2213235X11301010092.

Worley, B., & Powers, R. (2014). MVAPACK: a complete data handling package for NMR metabolomics. *ACS Chem Biol, 9*(5), 1138-1144, doi:10.1021/cb4008937.

Worley, B., & Powers, R. (2015). Generalized adaptive intelligent binning of multiway data. *Chemometr Intell Lab Syst, 146*, 42-46, doi:10.1016/j.chemolab.2015.05.005.

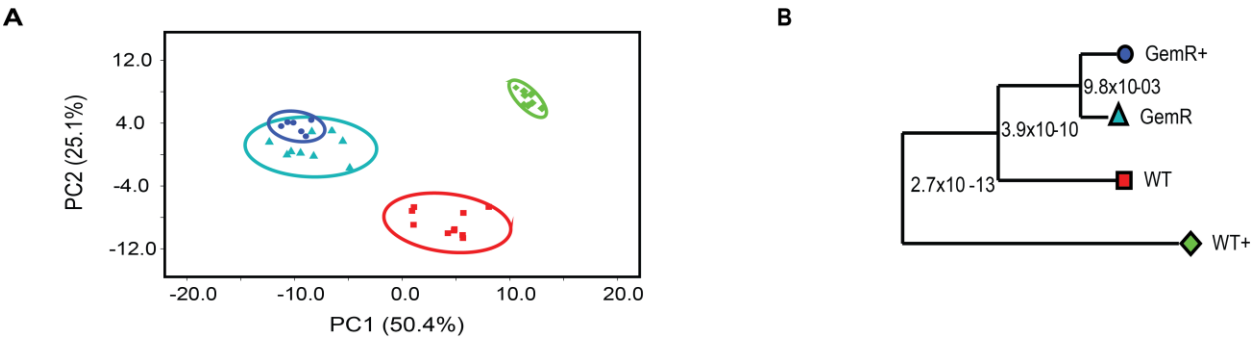

Figure S1: (**A**) PCA scores plot generated from 1D ^{1}H NMR spectra from cell lysates of wild type Capan-1 cells (WT, ■, n=8), WT cells treated with 10 nM of gemcitabine (WT+, ♦, n=8), gemcitabine-resistant (GemR, ●, n=8) cells, and GemR cells treated with 10 nM of gemcitabine (GemR+, ▲, n=8). The ellipses correspond to 95% confidence intervals for a normal distribution. (**B**) Metabolic tree diagram generated from the PCA score plots. The number at each node is the *p*-value calculated from the Mahalanobis distance between each group. The coloring is identical to the PCA scores plot.

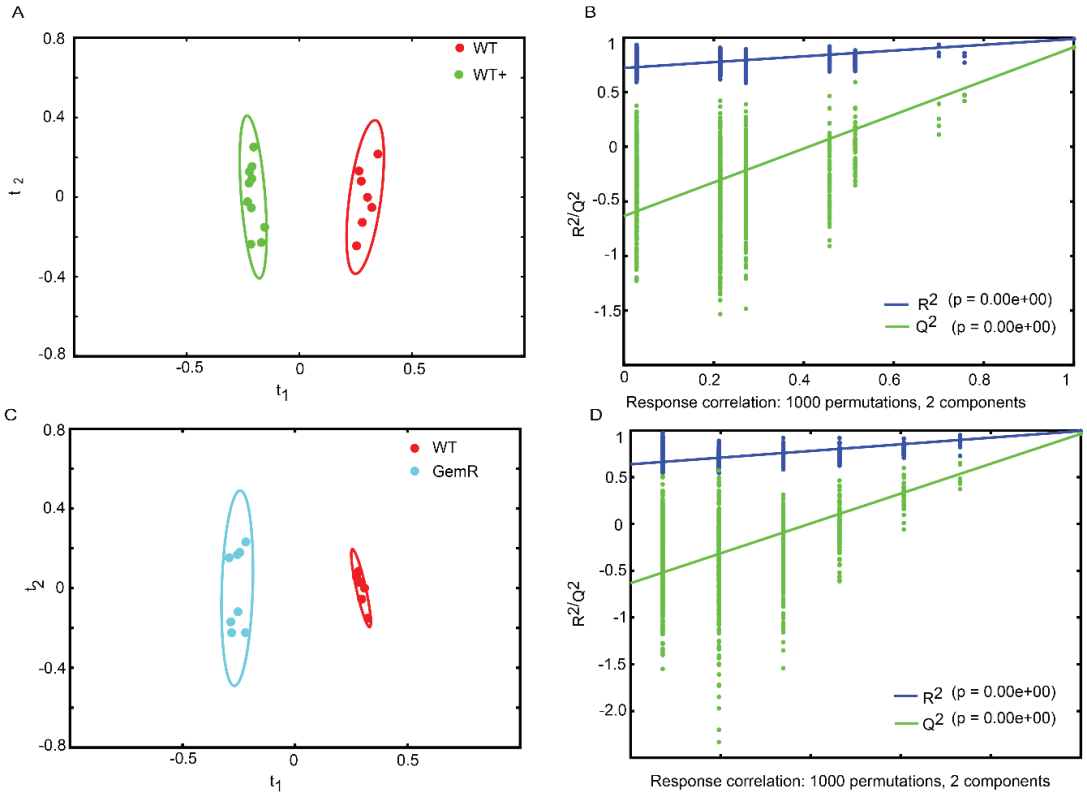

Figure S2: OPLS-DA models generated from the 1D ^{1}H NMR spectra of T3M4 cell lysates. (**A**) OPLS-DA scores plot comparing T3M4 wild-type cells with (WT+) and without (WT) the addition of 10 nM of gemcitabine. The validated OPLS-DA model yielded an R^2 of 0.99, Q^2 of 0.90 and a CV-ANOVA p-value of 7.94 x 10^{-7}. (**B**) Plot of R^2 and Q^2 values from permutation test for OPLS-DA models comparing WT and WT+ using 1000 random permutations of group membership. The observed R^2 and Q^2 *p*-values were both 0. (**C**) OPLS-DA scores plot comparing gemcitabine sensitive (WT) and gemcitabine-resistant (GemR) cells. The validated OPLS-DA model yielded an R^2 of 0.99, Q^2 of 0.0.96 and a CV-ANOVA p-value of 2.51 X 10^{-7}. (**D**) Plot of R^2 and Q^2 values from permutation test for OPLS-DA models comparing WT and GemR using 1000 random permutations of group membership. The observed R^2 and Q^2 *p*-values were both 0. The coloring of the symbols in the scores plot corresponds to: WT (red), WT+ (green): GemR (teal).

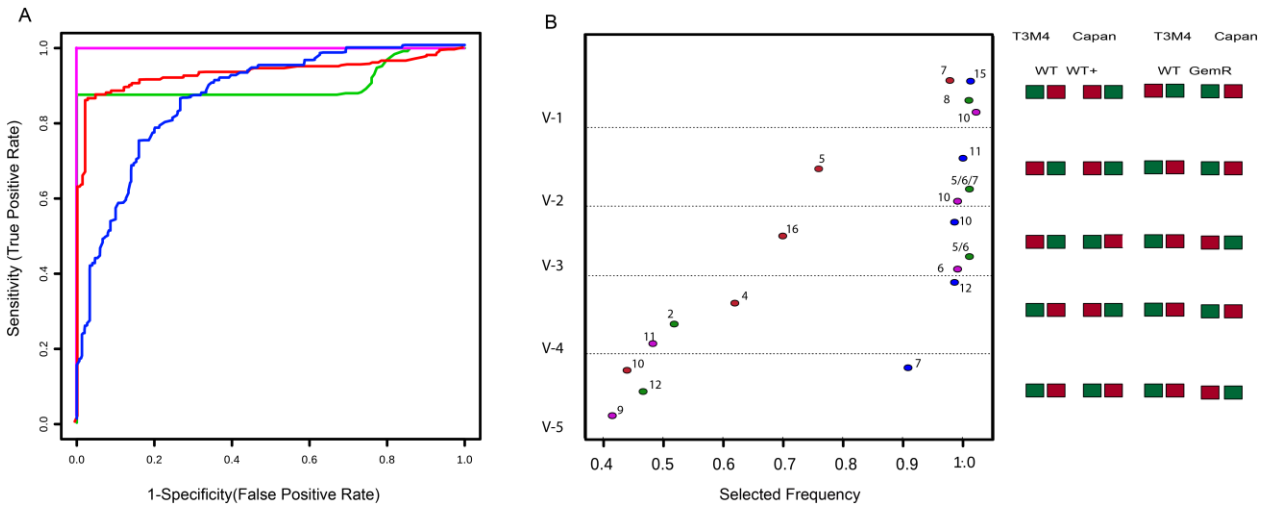

Figure S3. ROC curves generated from 1D ^{1}H NMR spectra of T3M4 and Capan-1 cell lysates. (**A**) Overlay of four ROC curves comparing NMR spectral features between T3M4 WT and WT+ cells (red, AUC 0.948), Capan-1 WT and WT+ cells (blue, AUC 0.853), T3M4 WT and GemR cells (green, AUC 0.909), and Capan-1 WT and GemR cells (pink, AUC 1.0) (**B**) The top five significant spectral bin values (V1 to V5) used to predict group membership in each ROC curve. Each spectral feature is colored according to the ROC curve it belongs to and is annotated using the same metabolite numbering scheme as in Figure 1C, E: 1, branched chain amino acids (leucine, isoleucine, valine); 2, lactate; 3, alanine; 4, acetate; 5, glutamate; 6, glutamine; 7,glutathione; 8,malate; 9,aspartate; 10, creatinine; 11, creatinine phosphate; 12, glycerophosphocholine; 13, taurine, 14, glycine, 15, myo-inositol 16, AXP (AMP, ADP, and ATP). 1, branched chain amino acids (leucine, isoleucine, valine); 2, lactate; 3, alanine; 4, acetate; 5, glutamate; 6, glutamine; 7,glutathione; 8,malate; 9,aspartate; 10, creatinine; 11, creatinine phosphate; 12, glycerophosphocholine; 13, taurine, 14, glycine, 15, myo-inositol 16, AXP (AMP, ADP, and ATP). Bins colored red indicate a relative increase in intensity, while bins colored green indicate a relative decrease in intensity. The bin order from left to right is:

T3M4 WT, T3M4 WT+, Capan-1 WT, Capan-1 WT+, T3M4 WT, T3M4 GemR, Capan-1 WT, Capan-1 GemR.

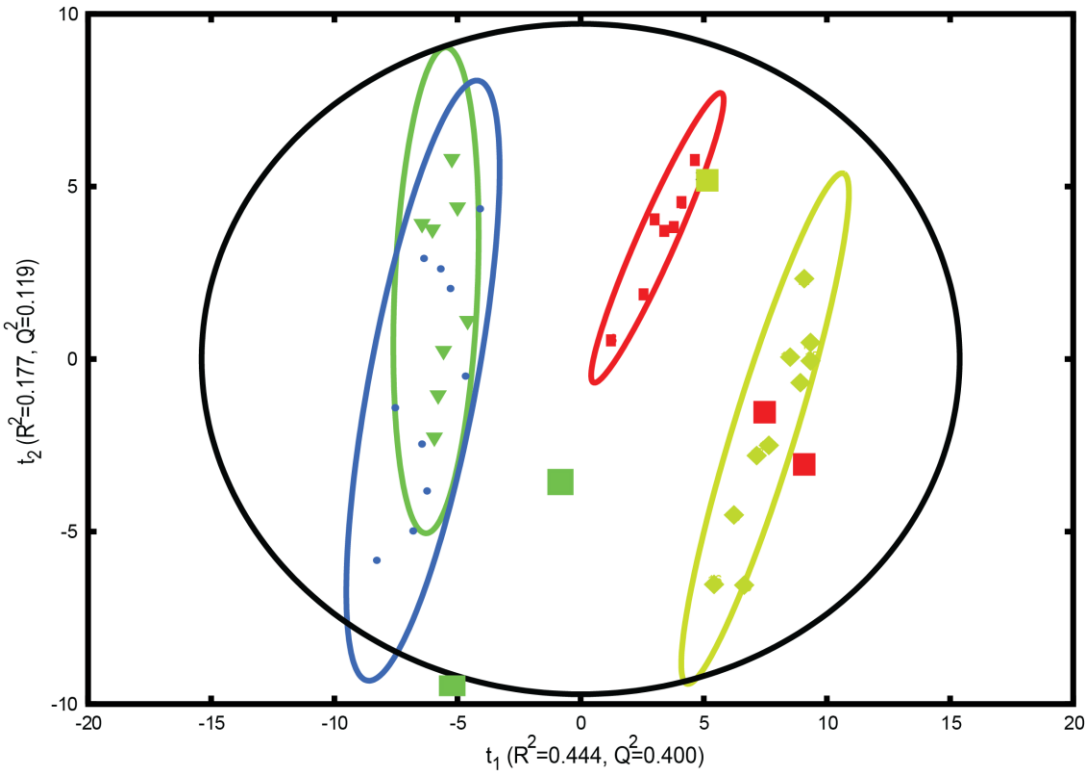

Figure S4. PCA scores plot generated from forty 1D ^{1}H NMR spectra collected from cell lysates of wild type T3M4 cells (WT, ■), WT cells treated with 10 nM of gemcitabine (WT+, ♦), gemcitabine-resistant (GemR, ●) cells, and GemR cells treated with 10 nM of gemcitabine (GemR+, ▲). The ellipses correspond to 95% confidence intervals for a normal distribution excluding the outliers. Each cluster contains ten biological replicates. The large squares (various colors) identify data points removed from further analysis.