Supplementary Methods

TI-SWR – TI-SWR analysis was performed in Matlab using the *stepwise* function in conjunction with several scripts to automate analysis iterations and post-processing of results.

Dependent variable data vectors (6x1) were generated to capture the change in specific cellular phenotypes from time, t-1, to time, t, following DNA damage for a subset of the cellular phenotypes measured. This metric was calculated as follows:

$$\frac{cell\_phenotype_t - cell\_phenotype_{t-1}}{\Delta t}$$

The subset of cellular phenotypes analyzed includes: 1) ΔcC3+/cPARP+$_{6\text{-}12hr}$, 2) ΔcC3+/cPARP-$_{6\text{-}12hr}$, 3) ΔcC3-/cPARP-$_{6\text{-}12hr}$, 4) ΔG1$_{12\text{-}24hr}$, 5) ΔS$_{12\text{-}24hr}$, 6) ΔG2/M$_{12\text{-}24hr}$. These phenotypes were chosen for analysis based on large dynamic changes in these measurements over the indicated time-periods. For example, the dependent variable vector capturing the change in percentage of cells in S phase between 12 and 24 hr following DNA damage would look like:

$$\begin{pmatrix} \left.\frac{(\%S_{24hr} - \%S_{12hr})}{24-12hr}\right|_{0\,Dox} \\ \left.\frac{(\%S_{24hr} - \%S_{12hr})}{24-12hr}\right|_{0\,Dox+TNF} \\ \left.\frac{(\%S_{24hr} - \%S_{12hr})}{24-12hr}\right|_{2\,Dox} \\ \left.\frac{(\%S_{24hr} - \%S_{12hr})}{24-12hr}\right|_{2\,Dox+TNF} \\ \left.\frac{(\%S_{24hr} - \%S_{12hr})}{24-12hr}\right|_{10\,Dox} \\ \left.\frac{(\%S_{24hr} - \%S_{12hr})}{24-12hr}\right|_{10\,Dox+TNF} \end{pmatrix}$$

Independent variable data matrices (6x14) were generated for each of the 10 time-points at which signaling data was measured. Each of these 10 independent variable matrices capture the activities and/or levels of the 14 signals measured under each of the 6 treatment conditions investigated, at a single one of the ten time-points. For example, a single column of the

independent variable matrix capturing the signal activity/levels at 4 hr following DNA damage would look like:

$$\begin{pmatrix} pAkt_{0\,Dox} \\ pAkt_{0\,Dox+TNF} \\ pAkt_{2\,Dox} \\ pAkt_{2\,Dox+TNF} \\ pAkt_{10\,Dox} \\ pAkt_{10\,Dox+TNF} \end{pmatrix}_{4\,hr}$$

A second column captures the p-ERK1/2 measurements under all 6 treatment conditions, a third column captures the p-JNK measurements, etc. A forward stepwise regression algorithm, implemented as a built-in function in MATLAB, was used to regress a single dependent variable data vector, representing the change in a single cellular phenotype over time, on each of the independent variable matrices constructed for time-points preceding the cellular phenotype under investigation, in turn. For example, if we investigate the potentially causal relationship between time-dependent signaling events and the change in percentage of apoptotic cells between 6 and 12 hr following DNA damage, we regress a dependent variable data vector that looks like this:

$$\begin{pmatrix} \left.\frac{(\%\,apoptosis_{12hr} - \%\,apoptosis_{6hr})}{12-6hr}\right|_{0\,Dox} \\ \left.\frac{(\%\,apoptosis_{12hr} - \%\,apoptosis_{6hr})}{12-6hr}\right|_{0\,Dox+TNF} \\ \left.\frac{(\%\,apoptosis_{12hr} - \%\,apoptosis_{6hr})}{12-6hr}\right|_{2\,Dox} \\ \left.\frac{(\%\,apoptosis_{12hr} - \%\,apoptosis_{6hr})}{12-6hr}\right|_{2\,Dox+TNF} \\ \left.\frac{(\%\,apoptosis_{12hr} - \%\,apoptosis_{6hr})}{12-6hr}\right|_{10\,Dox} \\ \left.\frac{(\%\,apoptosis_{12hr} - \%\,apoptosis_{6hr})}{12-6hr}\right|_{10\,Dox+TNF} \end{pmatrix}$$

on each of the independent variable matrices constructed for time-points prior to or at the 12 hr time-point (e.g. independent variable matrices constructed for the 0.25, 0.5, 1, 1.5, 2, 4, 8 and

12 hr time-point), as only measured signaling that occurs prior to or at the time of the phenotype is potentially causal.

A multiple linear regression analysis posits that all n independent variables contribute to the dependent variable with some weight, β_n, and a final model with n independent variables looks like:

$$Y(x_1, \ldots, x_n) = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n + \varepsilon$$

where Y is the dependent variable, β_0 is a constant term, $\beta_1 \ldots \beta_n$ are regression coefficients for independent variables $x_1 \ldots x_n$ and ε is the residual error. Here we use the stepwise regression method to identify which, if any, of the measured signals at a given time-point following DNA damage correlate with and are potentially causal for a cellular phenotype that follows the signaling event.

A stepwise regression analysis with forward selection is an iterative variable selection method for unbiased identification of locally optimal regression models that include only a subset of the candidate independent variables. The overarching hypothesis is that not all, and perhaps none, of the independent variables contribute to the dependent variable, such that an initial model looks like:

$$Y = \beta_0 + \varepsilon$$

Next, the effect of adding any single variable of the n independent variables to the model is tested such that all univariate regression models are sampled, e.g.:

$$\begin{aligned} Y(x_1) &= \beta_0 + \beta_1 x_1 + \varepsilon \\ Y(x_2) &= \beta_0 + \beta_2 x_2 + \varepsilon \\ &\vdots \\ Y(x_n) &= \beta_0 + \beta_n x_n + \varepsilon \end{aligned}$$

where x_i is one of the n signaling measurements at a specific time-point and β_i is the associated regression coefficient. If any of the time-dependent signals in the independent variable matrix are correlated with the cellular phenotype, a univariate model including that signal will provide a

better fit of the data than the initial null model. Provided that at least one of the signals in the independent variable matrix is correlated with the cellular phenotype, the current model is updated to the univariate model that includes the single most strongly correlated signal. For inclusion, a variable was required to display an $R^2 \geq 0.65$, and the slope of the subsequent regression line had to be statistically significantly different from 0 (as confirmed by a p-value ≤ 0.05 using a linear regression t-test).

Next, the effect of adding any second variable from the set of n-1 remaining independent variables to the current univariate model is tested such that all bivariate models are sampled holding the first included independent variable constant. For example, if independent variable x_2 provides the best univariate model, then subsequent bivariate models tested would look like:

$$Y(x_2,x_1) = \beta_0 + \beta_2 x_2 + \beta_1 x_1 + \varepsilon$$
$$Y(x_2,x_3) = \beta_0 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$
$$\vdots$$
$$Y(x_5,x_n) = \beta_0 + \beta_2 x_2 + \beta_n x_n + \varepsilon$$

If the addition of any signal improves the model fit, the signal whose addition to the model most significantly improves the correlation with the dependent variable cellular phenotype is added . The addition of signals to the model stops when the addition of no other signal significantly improves the correlation between the signals and the cellular phenotype.

For each cellular phenotype investigated (see above), this process of variable selection and model building is used to build a model relating time-dependent signals to phenotype for each of the signal measurement time-points that precede the phenotype such that a set of time-point specific signal-phenotype models is associated with each phenotype. For example, the change in the percentage of apoptotic cells between 6 and 12 hr following DNA damage can at most be associated with 7 time-dependent signal-phenotype models (for 0.25, 0.5, 1, 2, 4, 8, 12 hr following DNA damage), but might for example be associated with only 5 time-dependent signal-phenotype models if no signals are significantly correlated with the phenotype at 2 of those time-points.

Supplemental Figure Captions

Supplemental Figure 1. G1 cells contains markers of DSBs following doxorubicin exposure. Cells were treated with 2μM doxorubicin as in Figure 1, stained for γH2AX and DNA content (propidium iodide) at 8 hrs after the onset of treatment, and analyzed by flow cytometry. Nearly 50% of cells in G1 (2N DNA content) stained positively for γ-H2AX, compared with ~90% of cells in S or G2/M. Data is the mean of 2 experiments with error bars denoting standard error of the mean.

Supplemental Figure 2. Representative single dataset. Quantification of Western blot and flow cytometry data for one of the six treatment conditions investigated (2 μM doxorubicin without TNF-α). Signal activation data is plotted normalized to the maximum value reached for each signal under this treatment condition, which was assigned a value of 1.0. Cell cycle and cell death responses measured by flow cytometry data are shown as raw percentage of the total population. Mean values and standard error of the mean are shown for 2-3 independent experiments for all signaling measurements except p-ATM, and for ≥ 4 independent experiments for cell response measurements.

Supplemental Figure 3. Quantification of Western blot data for signaling measurements of activity under all six treatment conditions investigated. Measured signals quantified include p-Nbs1 (S343), p-H2AX (S139), NFκB-DNA binding activity (via capture ELISA), p-p38 (T180/Y182), p-JNK (T183/Y185), p-Akt (S473), p-ERK1/2 (T202/Y204), p-p53 (S15), and total p53. Data is the mean of 2-3 independent experiments and is normalized to the signal measured at 0 hr (prior to treatment) to yield a measure of "Signal Intensity". Error bars denote standard error of the mean. When not visible, the error bars are obscured by the plot symbol.

Supplemental Figure 4. The majority of predictive power for apoptosis in PLSR models is captured by 2 principal components, with or without core DNA damage response molecules as independent variables. (**A**) Quantification of the cumulative variance captured by the model (R^2(cum.)), and cross-validated capability of the model to predict the percentage of apoptotic cell-death between 6 and 12 hr following treatment (Q^2(cum)), for a PLSR model retaining 1, 2 or 3 principal components. Cross-validated predictive capability is 0.28, 0.69 and 0.75 with 1, 2, or 3 principal components, respectively, indicating that retention of a third principal component does not contribute much marginal predictive power. (**B**) Quantification of the cumulative variance captured by the 2-component model (R^2(cum.)), and cross-validated capability of the model to predict the percentage of apoptotic cell-death between 6 and 12 hr following treatment (Q^2(cum)), for four separate 2-component PLSR models built with independent variable sets consisting of either (1) all measurements of signal activity, (2) only measurements of signal activity from 'core' DDR signals (p-ATM, p-H2AX, p-Nbs1, p-p53, total p53, p-p38), (3) only measurements of signal activity from 'non-core' DDR signals (p-

ERK, p-JNK, p-Akt, p-p38, NFκB), and (4) only measurements of total protein levels (total Nbs1, H2AX, ERK, Akt, JNK; not including any measurements of signal activity). (p-p38 MAPK was included in both the non-core and core DDR signaling sets, since, while p38 signaling is not generally considered part of the canonical DNA damage response (Zhou & Elledge, 2000), data from our lab and others suggests that it functions downstream of ATM and ATR in this context, and can make important contributions to the phenotypic response (Lafarga et al, 2009; Raman et al, 2007; Reinhardt et al, 2007). Cross-validated predictive capability is high and relatively constant across the first three models built with all activity data, core-DDR signal activity data, or non-core DDR signal activity, indicating that non-core DDR signal activity independently contains information regarding the cellular life-death decision following DNA damage that can be used to predict the life-death decision. A model built with only total protein levels in the independent variable set has dramatically lower predictive power, indicating that measures of activity contain far more predictive information than total protein levels.

Supplemental Figure 5. The majority of the predictive power for cell cycle progression in PLSR models is captured by 2 principal components. Quantification of the cumulative variance captured by the model (R^2(cum.)), and cross-validated predictive capability (Q^2(cum)) of PLSR models retaining 1, 2 or 3 principal components was examined using (**A**) the percentage of cells in G2/M at 24 hr following treatment, or (**B**) the percentage of cells in G1 at 24 hr following treatment as dependent variables. Retention of a third principal component contributes minimally to the predictive power of the model.

Supplemental Figure 6. Late Akt activity after DNA damage is a pro-survival signal. Cells were treated with 10μM doxorubicin as in Figure 1. The AKT inhibitor VIII (CalBiochem) was present either from 0-2 hrs, or from 4 hrs onwards, and apoptotic cells quantified by flow cytometry at 12 hrs following treatment. Inhibition of AKT during the first 2 hrs after damage had little effect on cell death, while inhibition of late AKT activity resulted in a pronounced increase observed following the inhibition of late AKT activity. Data is the mean of 2 experiments with error bars denoting standard error of the mean.

Supplemental Table 1. A summary of signaling measurements used in this study.

Supplemental Table 2. Non-Erk predictors of S-phase progression after DNA damage identified by TI-SWR do not co-correlate with Erk activity. To computationally examine whether Erk activity indirectly affected other signaling molecules that were responsible for controlling S-phase progression, we used TI-SWR to correlate time-dependent signals with S-phase progression from 12 to 24 hrs, this time omitting ERK data from the candidate independent variable sets. Signals at only 4 time points showed statistically significant correlations with S-phase progression: levels of γH2AX + total H2AX at the 2 hr time point, levels of phospho-Nbs1 at the 4 hr time point, and levels of total p53 at the 16 and 24 time points. Regression coefficients (β), p-values, and the marginal increase in the correlation coefficients of the model for each of the independent variables are shown; asterisks indicate the

primary signal from the univariate models. A second round of TI-SWR was then performed to investigate time-dependent co-correlations of other signals with this subset of S-phase correlated signals, focusing on whether ERK activity at any time point co-correlated with, and was potentially causal, for any of these 'second choice' S-phase signal correlates. No such co-correlation between Erk and these other signals was detected.

References

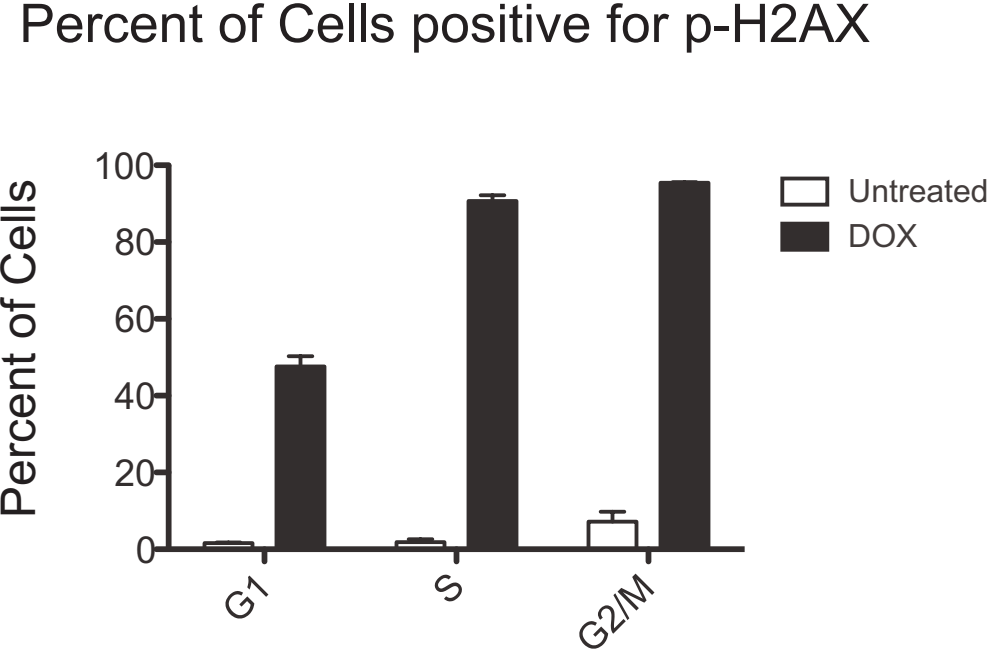
Lafarga V, Cuadrado A, Lopez de Silanes I, Bengoechea R, Fernandez-Capetillo O, Nebreda AR (2009) p38 Mitogen-activated protein kinase- and HuR-dependent stabilization of p21(Cip1) mRNA mediates the G(1)/S checkpoint. *Mol Cell Biol* **29:** 4341-4351

Raman M, Earnest S, Zhang K, Zhao Y, Cobb MH (2007) TAO kinases mediate activation of p38 in response to DNA damage. *Embo J* **26:** 2005-2014

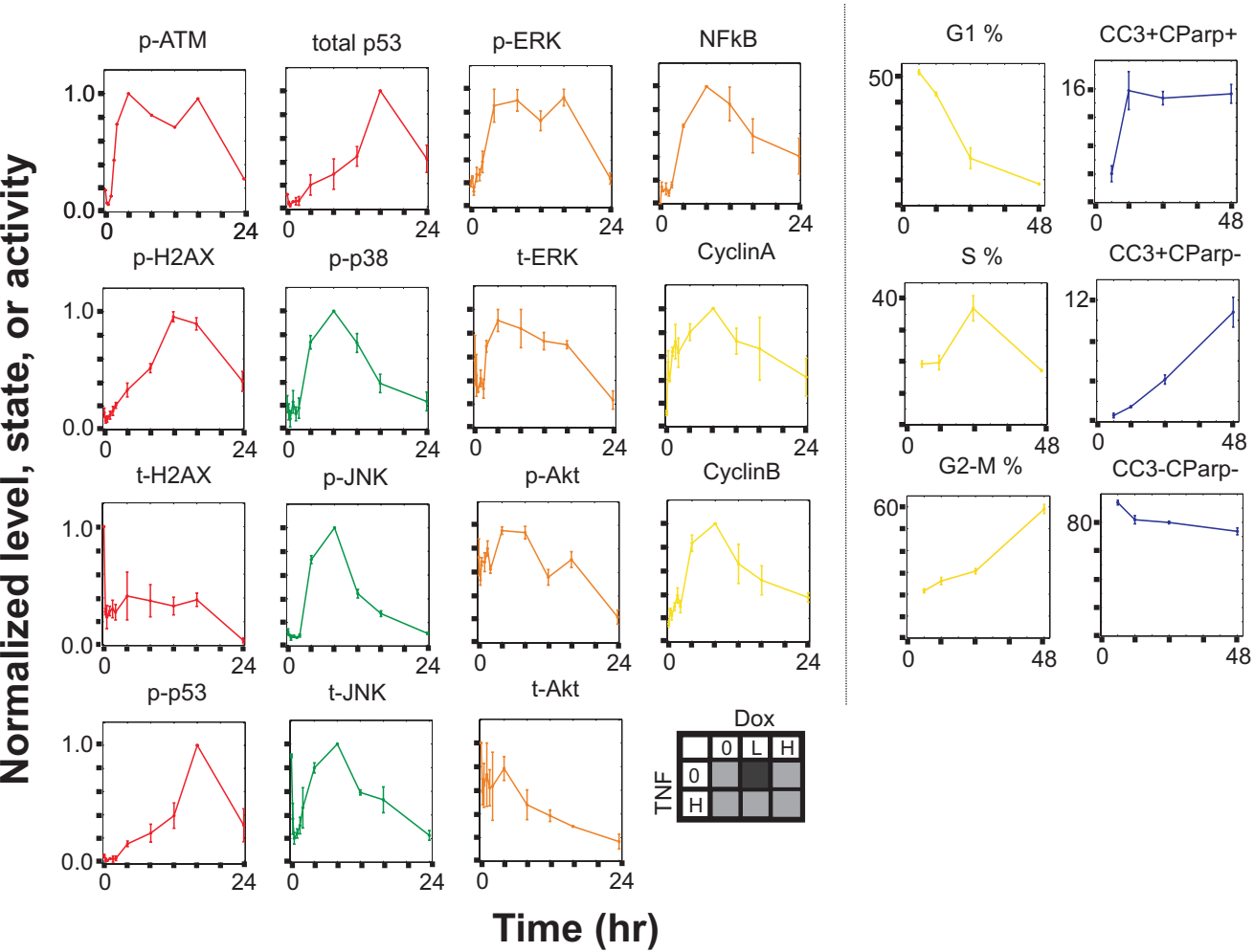
Reinhardt HC, Aslanian AS, Lees JA, Yaffe MB (2007) p53-deficient cells rely on ATM- and ATR-mediated checkpoint signaling through the p38MAPK/MK2 pathway for survival after DNA damage. *Cancer Cell* **11:** 175-189

Zhou BB, Elledge SJ (2000) The DNA damage response: putting checkpoints in perspective. *Nature* **408:** 433-439

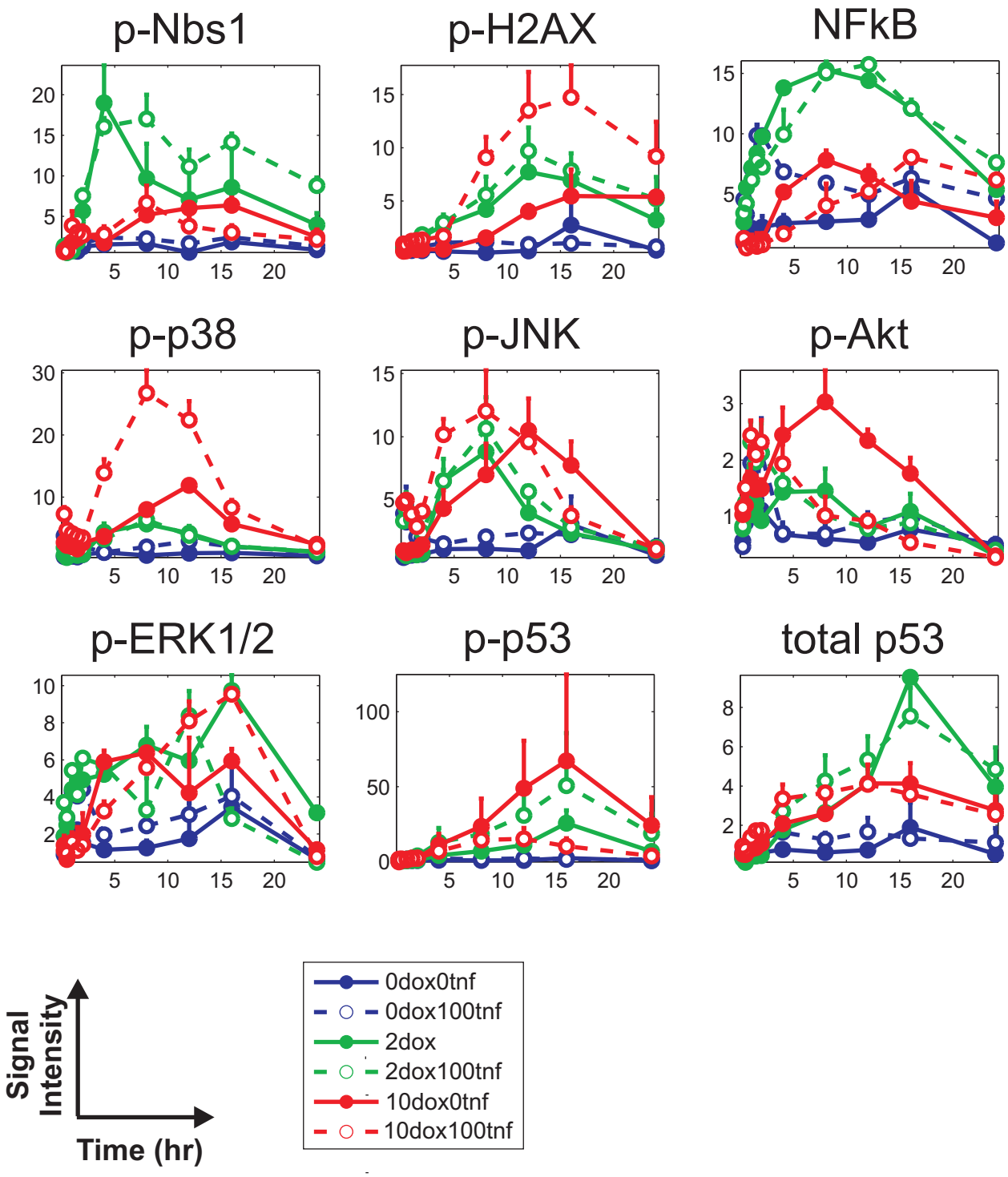
Supplementary Figure 1, Tentner et al.



Supplementary Figure 2, Tentner et al.

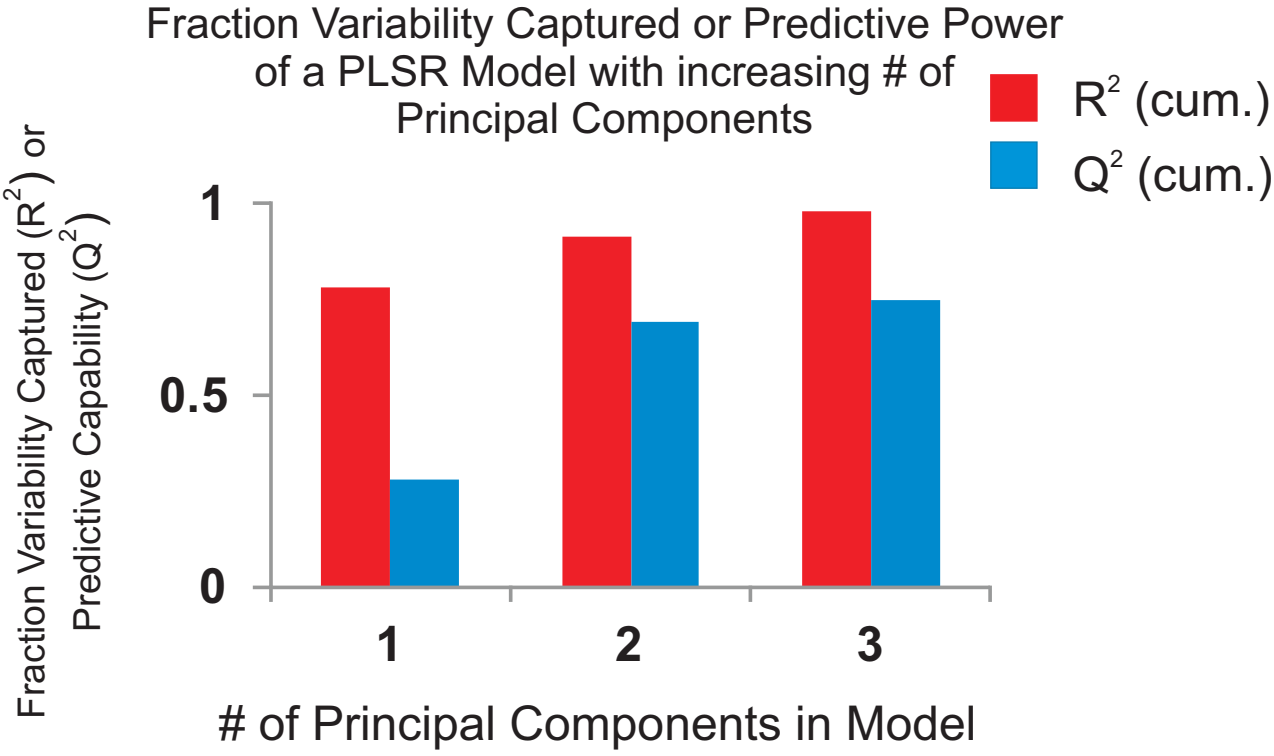


Supplementary Figure 3, Tentner et al.

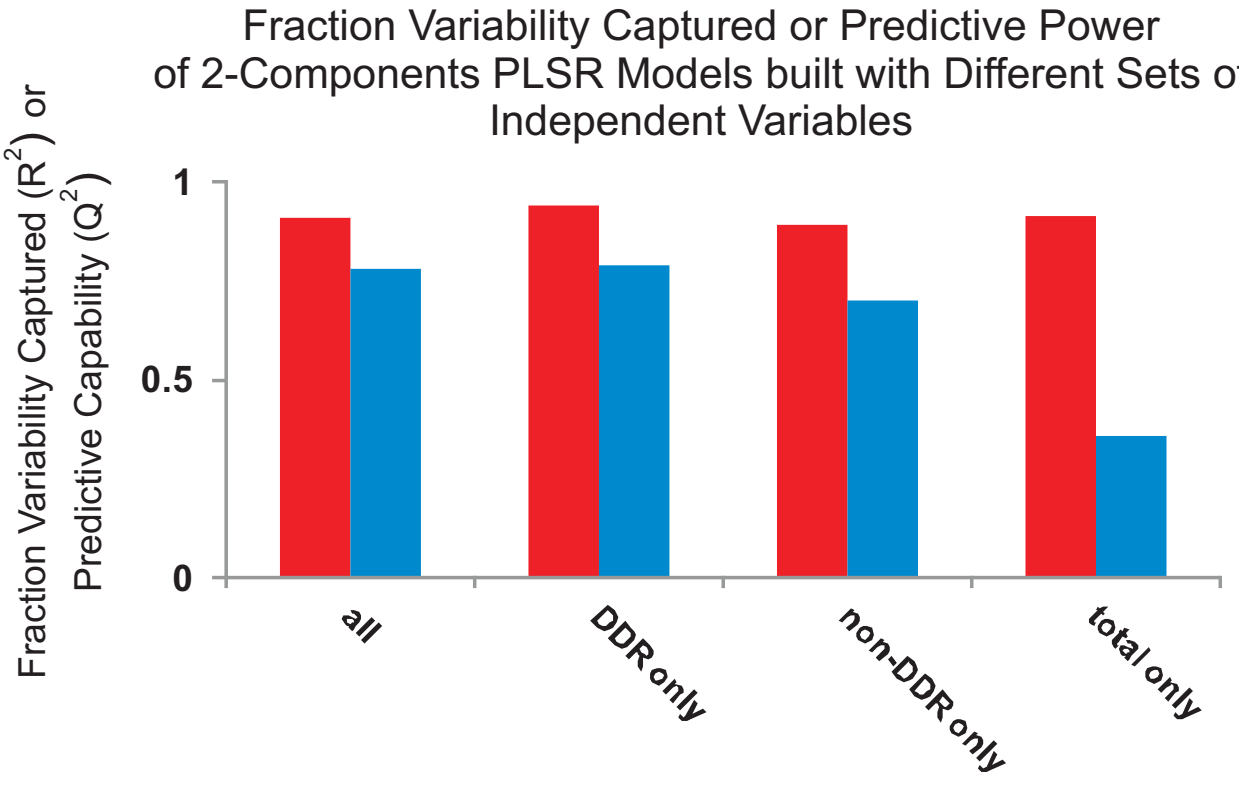


Supplementary Figure 4, Tentner et al.

A

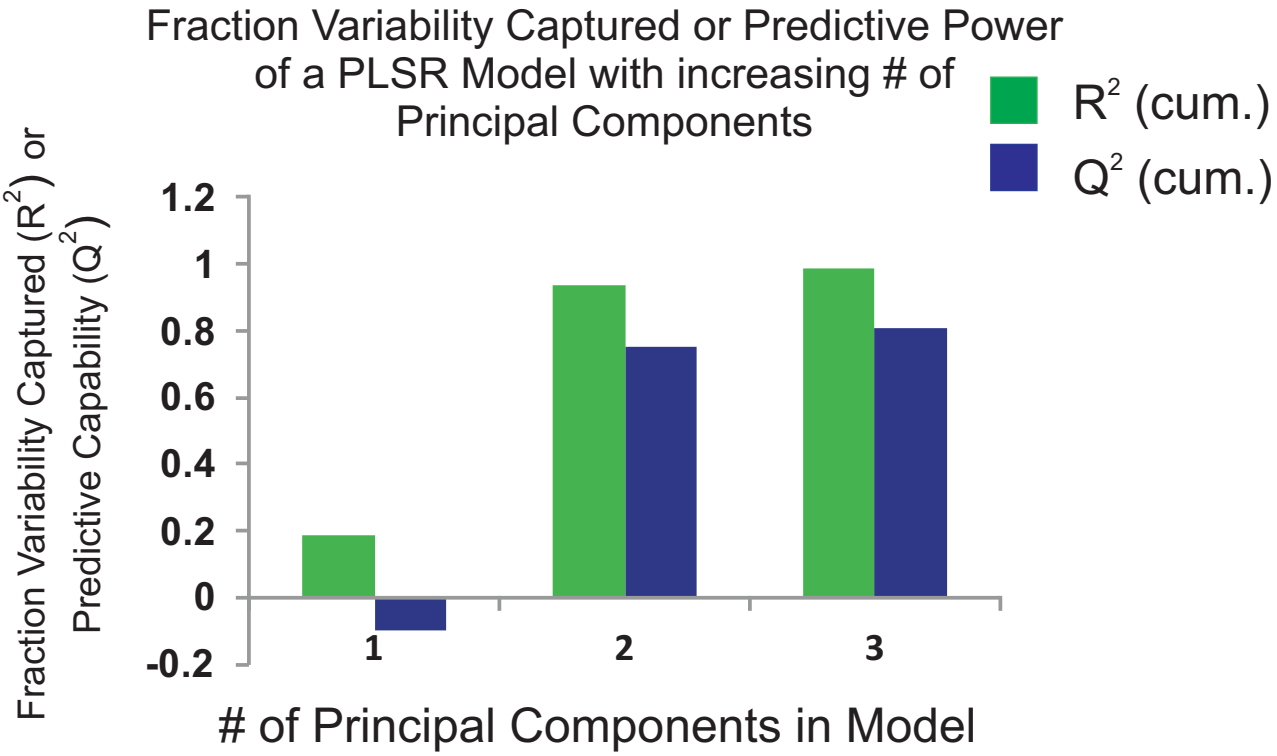


B

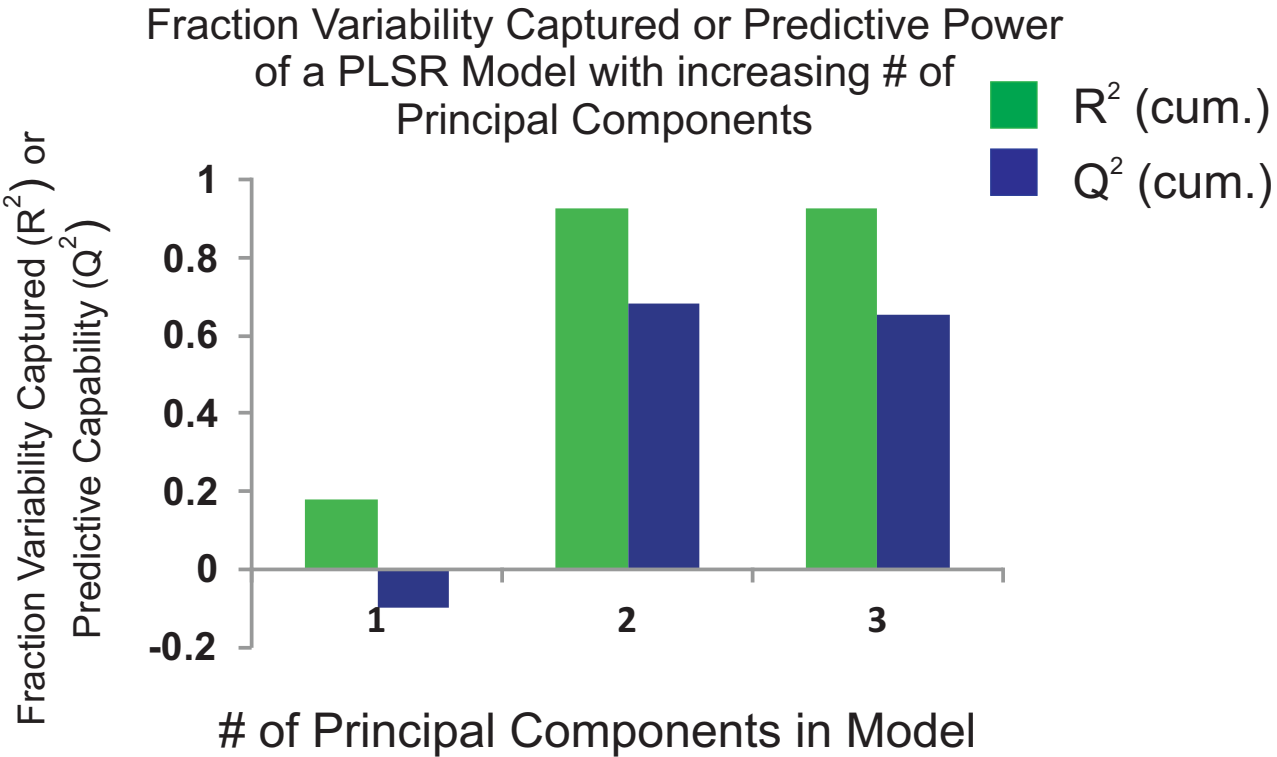


Supplementary Figure 5, Tenter et al.

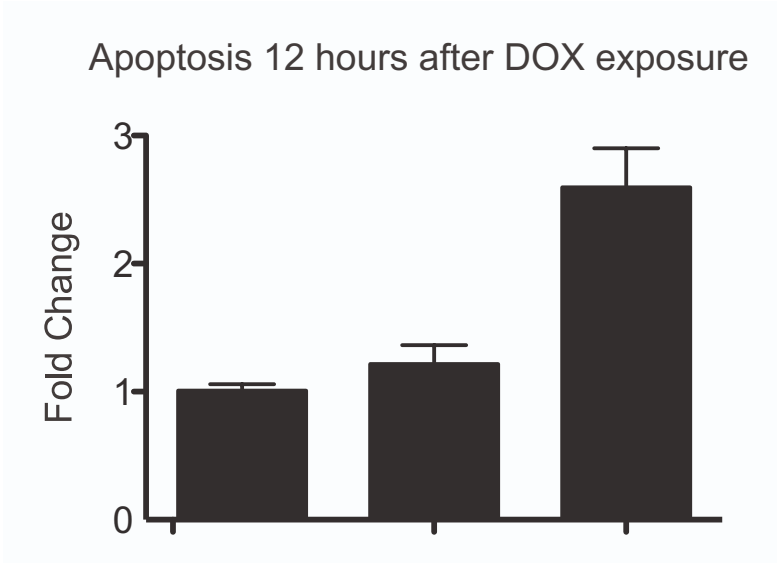
A



B



Supplementary Figure 6, Tentner et al.



Supplementary Table 1, Tentner et al.

Signaling molecule	Measurement	Measurement assay	Measurement time-points
pATM (S1981)	activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
pH2AX (S139)	activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
pNbs1 (S343)	activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
pp53 (S15)	activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
pp38 (T180/Y182)	activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
pJNK (T183/Y185)	activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
pERK1/2 (T202/Y204)	activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
pAkt (S437)	activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
total H2AX	level	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
total Nbs1	level	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
total p53	level/activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
total JNK	level	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
total ERK	level	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
total Akt	level	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
total Cyclin A	level/activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
total Cyclin B	level/activity	quantitative western	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr
NFkB	DNA-binding activity	capture-ELISA	0, 0.25, 0.5, 1, 1.5, 2, 4, 8, 12, 16, 24hr

Supplementary Table 2, Tentner et al.

time-point	signal	ß	p-value	R^2
2h	**p-H2AX***	0.86	0.002	0.66
2h	**total H2AX**	-1.8	0.006	0.32
4h	**p-Nbs1***	0.07	0.027	0.74
16h	**total p53***	0.19	0.002	0.92
24h	**total p53***	0.34	0.027	0.74