January 29, 2020

Title: 3Rs-Friendly Approach to Exogenous Metabolic Activation that Supports High-Throughput Genetic Toxicology Testing

Authors: Shuchang Tian[1]*, Aiyana Cyr[1,2]*, Karen Zeise[1,3]*, Steven M. Bryce[1], Nikki Hall[1], Jeffrey C. Bemis[1], Stephen D. Dertinger[1]**

Affiliation:

[1]Litron Laboratories, Rochester, New York, USA

[2]Current affiliation: University of Connecticut, Connecticut, USA

[3]Current affiliation: University of Michigan, Ann Arbor, Michigan, USA

Notes:

*These authors contributed equally to this work.

**Corresponding author: SDD, Litron Laboratories, 3500 Winton Place, Rochester, NY, 14623; sdertinger@litronlabs.com

Key words: DNA damage, metabolic activation, γH2AX, p53, flow cytometry

Running Title: 3Rs-Friendly Approach to Exogenous Metabolic Activation

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/em.22361

Abstract

MultiFlow® DNA Damage — p53, γH2AX, Phospho-Histone H3 is a miniaturized, flow cytometry-based assay that provides genotoxic mode of action information by distinguishing clastogens, aneugens, and non-genotoxicants. Work to date has focused on the p53-competent human cell line TK6. While mammalian cell genotoxicity assays typically supply exogenous metabolic activation in the form of concentrated rat liver S9, this is a less-than-ideal approach for several reasons, including 3Rs considerations. Here, we describe our experiences with low concentration S9 and saturating co-factors which were allowed to remain in contact with cells and test chemicals for 24 continuous hours. We exposed TK6 cells in 96-well plates to each of 15 reference chemicals over a range of concentrations, both in the presence and absence of 0.25% v/v phenobarbital/β-naphthoflavone-induced rat liver S9. After 4 and 24 hr of treatment cell aliquots were added to wells of a microtiter plate containing the working detergent/stain/antibody cocktail. After a brief incubation robotic sampling was employed for walk-away flow cytometric data acquisition. PROAST benchmark dose (BMD) modelling was used to characterize the resulting dose-response curves. For each of the 8 reference pro-genotoxicants studied, relative nuclei count, γH2AX, and/or p53 biomarker BMD values were order(s) of magnitude lower for 0.25% S9 conditions compared to 0% S9. Conversely, several of the direct-acting reference chemicals exhibited appreciably lower cytotoxicity and/or genotoxicity BMD values in the presence of S9 (e.g., resorcinol). These results prove the efficacy of the low concentration S9 system, and indicate that an efficient and highly scalable multiplexed assay can effectively identify chemicals that require bioactivation to exert their genotoxic effects.

Introduction

Among the significant challenges associated with screening large numbers of chemicals for genotoxic potential is the low throughput capacity of conventional hazard identification-type assays. Progress is being made in this area, and includes miniaturization and automated scoring of: *in vitro* micronuclei [Bryce et al., 2010]; the alkaline comet assay [Ge et al., 2015]; γH2AX and other indicators of DNA damage [Audebert *et al.,* 2010; Smart *et al.,* 2011; Tsamou *et al.,* 2012; Garcia-Canton *et al.,* 2013; Nikolova *et al.,* 2014; Cheung *et al.,* 2015; Khoury *et al.,* 2016]; a panel of five complementary high throughput screening assays described by Hsieh and colleagues [2019]; a toxicogenomic signature of genotoxicity [Li *et al.,* 2015, 2017]; and fluorescent reporters of DNA damage and other cellular stress indicators using one or more engineered mammalian cell lines [Hastwell et al., 2006; Hendriks et al., 2012].

Our work in this area includes development of the MultiFlow® DNA Damage Kit, an *in vitro* assay formatted as an add-and-read test that efficiently prepares mammalian cells in microtiter plates for flow cytometric analysis. The multiplexed biomarkers measured are: i) phosphorylation of H2AX at serine 139 (γH2AX) to detect DNA double strand breaks, ii) phosphorylation of histone H3 at serine 10 (p-H3) to identify mitotic cells, iii) nuclear p53 content as an indicator of p53 activation in response to DNA damage, iv) frequency of 8n+ cells to monitor polyploidization, and v) relative nuclei counts (RNC) to provide information about treatment-related cytotoxicity [Bryce et al., 2016].

The MultiFlow assay has demonstrated good sensitivity and specificity for detecting diverse genotoxicants, and furthermore provides information about the predominant mode of genotoxic action—clastogenicity versus aneugenicity [Bryce et al., 2014, 2016, 2017, 2018;

Bernacki et al., 2016; Dertinger et al., 2019]. The bulk of the work to date has been accomplished with the p53-proficient human cell line TK6. Since TK6 cells do not have appreciable cytochrome P450 enzyme activities, they cannot reliably detect pro-genotoxicants unless an exogenous source of metabolic activation is added to the test system. Historically, *in vitro* genotoxicity assays are therefore tested in the presence and absence of rat liver microsomes (i.e., the so-called S9 fraction) with necessary cofactors [Ames et al., 1973].

Instead of using rat liver microsomes, some laboratories provide metabolic capacity by working with metabolically competent cells, typically of hepatocyte origin. However, many of these cell lines express low levels of important P450 isoforms, while others have good expression profiles but their proprietary nature adds costs that cannot always be accommodated in high volume- early screening-type environments [Westerink and Schoonen, 2007; Le Hégart et al., 2014].

For various reasons then, the use of high concentration rat liver S9 continues to be widely employed. However, there are several aspects that are less than ideal. First, the historic use of Aroclor-induced rat liver S9 has become problematic. The supply of Aroclor is nearly exhausted and it will not be replenished (personally communication, R. Cammeron, Molecular Toxicology, Inc). This means laboratories accustomed to using Aroclor-induced S9 will need to switch to an alternate induction scheme, for example phenobarbital/β-naphthoflavone, and ensure it is effective for their application(s). Second, the traditional use of high concentration S9 (typically 2-4% v/v final) provides opportunities to reduce animal numbers, and thereby advance 3Rs [Russell and Burch, 1959]. Third, high concentration S9 is inherently cytotoxic to cultured mammalian cells, and this explains why most protocols call for washing cells free of test

chemical and S9 enzymes after several hours of exposure. This is especially suboptimal in high throughput testing environments, since it is a subtraction step that is not as easily automated relative to processing steps that call for simple transfer(s) or addition(s). Finally, the requirement for centrifugation and aspiration has the potential to impact cell health and numbers which may be mistakenly attributed to test article toxicity as opposed to processing inconsistencies.

To address these issues with exogenous metabolic activation systems we investigated an alternate strategy based on phenobarbital/β-naphthoflavone-induced rat liver S9. Importantly, the final concentration of S9 was 0.25% v/v—the maximal non-cytotoxic concentration. Owing to the low cytotoxicity of the system, it was possible to maintain enzymes/co-factor mix with cells and test article for the complete exposure period (in the case of TK6-based MultiFlow assays, 24 hr). This is similar to the approach used by the ToxTracker® system [Hendriks et al., 2012], where mouse stem cells are exposed to test article in the presence and absence of 0.25% Aroclor-induced rat S9 for 24 hr. For the initial proof-of-principle experiments described herein, we focused on clastogens, especially those that require enzymatic activation to most efficiently form DNA-reactive metabolites. Our encouraging experiences with the MultiFlow assay are described herein, along with a discussion about advantages of low concentration S9 approaches.

Materials and Methods

Chemicals

The identity of the 15 test chemicals are provided in Table I, along with information about metabolic activation, predominant mode of action, etcetera. Most are specified in the European

Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) publication that lists chemicals that are useful for validating *in vitro* genotoxicity assays [Kirkland *et al.,* 2016]. Importantly, the non-genotoxicants are cytotoxic to mammalian cells through a variety of mechanisms, and the pro-genotoxicants collectively require a diverse set of CYP450 enzymes to form DNA-reactive electrophiles.

Cell Culture and Treatments

TK6 cells were purchased from ATCC® (cat. no. CRL-8015). Cells were grown in a humidified atmosphere at 37°C with 5% CO_2, and were maintained at or below 1 x 10^6 cells/mL. The culture medium consisted of RPMI 1640 with 200 µg/mL sodium pyruvate (both from Sigma-Aldrich, St. Louis, MO), 200 µM L-glutamine, 50 units/mL penicillin and 50 µg/mL streptomycin (from Mediatech Inc., Manassas, VA), and 10% v/v heat-inactivated horse serum (Gibco®, a Thermo Fisher Scientific Company, Waltham, MA).

The low S9 concentration, continuous exposure activation approach utilized a NADPH regeneration system and frozen rat liver S9 from Molecular Toxicology Inc. (Boone, NC). Specifically, Regensys™ 'A' and Regensys™ 'B' reagents were combined, and this solution was used to prepare phenobarbital/β-naphthoflavone-induced rat liver S9 at a concentration of 2.5% v/v. This 10x solution was maintained on ice until it was added at a 1:9 ratio to TK6 cells adjusted to 2 x 10^5/mL in culture medium for a final S9 concentration of 0.25%. The 0% S9 cultures were TK6 cells at 2 x 10^5/mL in culture medium culture medium without cofactors or S9.

Dose range-finding experiments were performed to generate 24 hr relative nuclei count (RNC) data for each chemical. Chemical treatments occurred in U-bottom 96 well plates, with

198 µL TK6 cell suspensions with and without S9/cofactor mix as described above. Test chemicals prepared in DMSO were added at 2 µL/well for final DMSO concentrations of 1% v/v. The highest test chemical concentration was 10 mM unless solubility or previous experience with a chemical indicated this would be overly cytotoxic. Testing occurred at 20 concentrations in single wells, and each concentration differed from the one above by a factor of 70.71%. Solvent was tested in at least 10 replicate wells spread throughout the plate.

Definitive experiments with the 15 test chemicals were also conducted in U-bottom 96 well plates, with 198 µL TK6 cells with and without S9/cofactor mix as described above. As with the dose range-finding experiments, test chemicals prepared in DMSO were added at 2 µL/well. The highest concentration was derived from the dose range-finding experiment. In the case of non-cytotoxic freely soluble chemicals, the top concentration was 10 mM. When precipitate was evident, the lowest precipitating concentration was evaluated. Otherwise, our goal for the highest concentration tested was 60-80% reduction to RNC at 24 hr, but only two concentrations within the range 70-80% reduction were permitted [Dertinger et al., 2019]. The additional ten lower concentrations were tested using the 70.71% dilution scheme described above. Each of the 11 concentrations was tested in triplicate wells, whereas DMSO controls were evaluated in 10 replicate wells spread throughout the plate. Upon addition of test chemicals the plates were immediately incubated in a humidified atmosphere at 37°C with 5% CO_2 for 24 hr.

Other experiments were performed with cyclophosphamide- and dibenzo[a,l]pyrene-exposed TK6 cells in order to directly compare results from low, continuous S9 treatment versus a traditional approach to metabolic activation (i.e., short-term exposure with a high S9

concentration). The low, continuous S9 exposure occurred with 0.25% S9, and took place as described above. The short-term exposure with high concentration S9 was performed with the same rat liver S9 and cofactor mix, but in this case the final S9 concentration was 2% v/v, delivered from a 20% v/v stock solution. After 4 hr of treatment, cells were washed two times (*via* centrifugation, aspiration of 150 µL supernatants, and resuspension with phosphate buffered saline). After final resuspension with growth medium, the cells were reincubated for an additional 20 hr.

MultiFlow Assay

For dose range-finding experiments, cells exposed to test chemicals for 4 and 24 hr were resuspended with pipetting, then 25 µL were removed from each well and added to a new 96-well plate containing 50 µL/well of pre-aliquoted working MultiFlow reagent solution. This solution was prepared with reagents in the MultiFlow® DNA Damage Kit — p53, γH2AX, Phospho-Histone H3 (Litron Laboratories, Rochester, NY). Note that for these dose range-finding experiments, antibodies were omitted from the reaction mix. In this manner, we simultaneously measured cytotoxicity (i.e., RNC values) as well as fluorescence in the p53 and γH2AX channels for evidence of test article-associated fluorescence ("background fluorescence") that could impact the definitive assay. After incubation at room temperature for 30 min, samples were analyzed *via* flow cytometry.

For the definitive experiments, TK6 cells were prepared for analysis using reagents and instructions included in the MultiFlow® DNA Damage Kit — p53, γH2AX, Phospho-Histone H3. As with the dose range-finding experiments, at the 4 and 24 hr time points, an aliquot of 25

µL/well was added to wells containing 50 µL of pre-aliquoted working MultiFlow reagent solution without antibodies present. This was done for one of three replicate wells per concentration. Additionally, for each of the three replicate wells per concentration, an aliquot of 25 µL/well was added to wells containing 50 µL of pre-aliquoted working MultiFlow reagent solution with antibodies present. After incubation at room temperature for 30 min, samples were analyzed *via* flow cytometry.

Flow cytometric analysis was carried out using a Miltenyi Biotec MACSQuant® Analyzer 10 flow cytometer with integrated 96-well MiniSampler device. Stock photomultiplier tube detectors and associated optical filter sets were used to detect fluorescence emissions associated with the fluorochromes: FITC (detected in the B1 channel, PE (B2 channel), propidium iodide (B3 channel), and Alexa Fluor® 647 (R1 channel).

Representative bivariate graphs, gating logic, and position of regions were described in detail in a previous report [Bryce et al., 2016]. Briefly, two biomarker measurements, γH2AX and p53, were based on the shift in median channel fluorescence intensity relative to same-plate solvent controls. Polyploidy and p-H3 biomarker measurements were based on their frequency among other nuclei. Nuclei to counting bead ratios were calculated for each sample, and these ratios were used to determine absolute nuclei counts (those with 2n and greater DNA-associated propidium iodide fluorescence). Nuclei counts were used to derive RNC, and %cytotoxicity was calculated as 100% minus %RNC at 24 hr for each concentration tested.

MultiFlow Data Analysis: Pre-Processing

Definitive study MultiFlow data were corrected for background fluorescence in those

rare instances when test chemical alone, in the absence of fluorescent antibodies, shifted the median channel fluorescence of the p53 and/or γH2AX biomarkers relative to solvent control wells. This was accomplished in a test chemical- and concentration-specific manner by subtracting the mean solvent control median fluorescence value (no antibodies present) from the median fluorescent value (antibodies present) observed at each test article concentration. This value corresponds to fluorescence that cannot be attributed to antibody reagents, and was subtracted from the corresponding median fluorescence channel observed at the same concentration in the presence of the complete labeling solution, that is, the reagent mix that included antibodies.

After baseline fluorescence corrections as necessary, MultiFlow data were prepared for analysis by converting 4 and 24 hr γH2AX and p53 median fluorescence values, and p-H3 and polyploidy frequency measurements, into fold-change values. This was accomplished on an endpoint-, well-, and time point-specific basis by dividing the biomarker measurement by the mean solvent control value (Microsoft Excel, v16.16.14). This was performed for every test article concentration that was not excluded due to excessive cytotoxicity or other limits described above.

MultiFlow Data Analysis: Machine Learning Ensemble

The use of three ML models, multinomial logistic regression (LR), artificial neural network (ANN), and random forest (RF), has been described in detail previously [Bryce et al., 2018]. These various models utilize 4 and 24 hr MultiFlow data fold-change values and predict whether a chemical exhibits clastogenic, aneugenic, or clastogenic and aneugenic activity. Each

model's output was synthesized into MoA calls as follows. A clastogenic MoA required two successive concentrations to exhibit clastogen probability scores ≥ 80%, or one concentration to exhibit a clastogen probability score ≥ 90%. An aneugen MoA required two successive concentrations to exhibit aneugen probability scores ≥ 80%, or one concentration to exhibit an aneugen probability score ≥ 90%. Non-genotoxic was defined as the absence of two successive concentrations exhibiting clastogen or aneugen probability scores ≥ 80%, and no single concentration exhibiting a clastogen or aneugen probability score ≥ 90%.

Results from the three clastogen models were synthesized into a final prediction based on a majority vote ensemble. That is, a simple majority (≥ 2/3 clastogen-positive models) was necessary for a final clastogen call. (Note that since the experiments described herein focused on metabolic activation and clastogenicity, and because no aneugen probability scores ≥ 80% were observed, data for the aneugen-responsive biomarkers (p-H3 and polyploidy) and aneugen machine learning probabilities are not presented.)

MultiFlow Data Analysis: BMD Analyses

Benchmark dose (BMD) analyses [Wills et al., 2015] were performed for several MultiFlow biomarker responses using PROAST v67.0, which was accessed through the European Food Safety Authority online tool, see: https://efsa.onlinelibrary.wiley.com/doi/pdf/10.2903/sp.efsa.2019.EN-1489. Critical Effect Size (CES) values were -0.3 for the %RNC endpoint (i.e., 30% reduction), and 0.3 for fold-change p53 and γH2AX endpoints (i.e., 30% increase). One chemical at a time was analyzed, with the S9 condition as the covariate. With very few exceptions, an exponential model was used to fit

the dose response curves. To facilitate potency comparisons across S9 conditions, BMD values are reported for all chemicals, biomarkers, and time points. In those cases where PROAST could not reliably fit a model (i.e., because a response did not reach the specified CES), or where the BMD estimate was beyond the top concentration tested, we cited the BMD as the top concentration tested with a greater than symbol. BMD values for with and without S9 were used to derive a "S9 Potentiation Ratio", which is the BMD value in the absence of S9 divided by the BMD value in the presence of S9. For example, a value of 10 corresponds to chemical/biomarker/time point combination exhibiting 10-fold higher potency in the presence of S9, whereas a value of 0.1 corresponds 10-fold reduced potency in the presence of S9.

Results and Discussion

Preliminary Experiments

Early work focused on determining the maximal amount of rat liver S9 that could be tolerated by TK6 cells for 24 hr without causing significant cytotoxicity. Where S9 concentrations ≥ 0.5% showed evidence of cytotoxicity as expressed by reduced %RNC and increased frequencies of cells with impaired membrane integrity, 0.25% S9 in combination with 10% v/v cofactor mix had little to no effect (data not shown). This is consistent with the ToxTracker protocol that utilizes 0.25% S9 for mouse stem cell experiments, and explains our use of 0.25% for the experiments that followed.

Other experiments were performed to directly compare results generated by the low (0.25%) S9 approach to those based on traditional short-term treatment with a high S9 concentration. As shown by Figure 1, CP caused robust γH2AX and p53 responses that

correctly classified the agent as clastogenic. The 24 hr γH2AX and p53 responses were remarkably similar across the S9 systems. On the other hand, several significant differences were evident for other biomarkers/time points. For instance, appreciably greater cytotoxicity (reduced RNC) was evident at lower concentrations when 2% S9 was used. Also, the RNC values for replicate 2% S9 wells were more variable compared to the low S9 approach, likely due to the additional cell processing steps that were necessary. We quantified the extent of RNC variation by considering the ratio between the BMD upper confidence limit (BMDU) to the BMD lower confidence limit (BMDL) [Wills et al., 2015]. Whereas this ratio was low for the continuous S9 cultures (8.09 / 5.89 or 1.4), it was considerably higher for the 2% S9 cultures that required washout (1.67 / 0.0774 or 22). Finally, whereas the magnitude of 4 hr γH2AX responses were similar between low and high concentration S9, the latter's dose response curve was shifted to the left.

Figure 2 shows results for TK6 cells exposed to DB[a,l]P with low, continuous S9 versus the standard short-term 2% approach. Again, both systems produced clear γH2AX and p53 responses that correctly classified DB[a,l]P as clastogenic. The RNC dose response curves were quite similar, although replicate wells were more variable for high concentration S9. (BMDU to BMDL ratios = 1.8 and 10.7 for continuous S9 versus washout S9, respectively.) Whereas the 4 hr γH2AX dose response curve was shifted slightly to the left for the high S9 condition, the reverse was true for 24 hr γH2AX and p53 curves. Furthermore, 24 hr γH2AX and p53 responses exhibited markedly higher magnitudes in the low, continuous S9 system.

Pilot studies with B[a]P are noteworthy, as they alerted us to the possibility that some test chemicals can impart fluorescence to detergent-liberated nuclei that may affect one or both

biomarkers that are based on median fluorescence, i.e., p53 and γH2AX. Figure 3 shows 24 hr p53 results over a range of B[a]P concentrations. Since nuclei that were not brought into contact with antibodies exhibited increasing fluorescence with increasing B[a]P concentration, this effect must be attributed to background fluorescence, not the p53 biomarker. Once background fluorescence is subtracted from the values obtained with antibodies present, a more accurate assessment of the biomarker is made. For the experiments described herein, background fluorescence was only appreciable for three chemicals (2AAN, B[a]P, and brefeldin A) and the one biomarker read on the FITC channel, i.e., p53.

Given the promising preliminary results described above, we proceeded to test 15 reference chemicals with and without the low concentration S9 system.

2AAF

Results for 2AAF are provided in Figure 4. The presence of S9 did not remarkably affect cytotoxicity as expressed as %RNC, and no substantial effects on the p53 biomarker were observed. On the other hand, γH2AX clearly responded at the 4 and 24 hr time points when S9 was present, and these were of sufficient magnitude for the machine learning ensemble to characterize 2AAF +S9 as clastogenic. Although a modest γH2AX effect was apparent at 4 hr without S9, it was not sufficient for a clastogen call. Based on BMD estimates, S9 potentiated early and late γH2AX responses by approximately 2.6- and 8.9-fold, respectively (Table II).

2AAN

2AAN results are shown in Figure 5. The presence of S9 potentiated 24 hr %RNC, as well

as p53 and γH2AX induction at both time points studied. The machine learning ensemble characterized 2AAN +S9 as clastogenic. The γH2AX responses at the 4 and 24 hr time points without S9 were of sufficient magnitudes for the machine learning ensemble to make a clastogenic call for this condition as well. That being said, Table II highlights the fact the presence of S9 resulted in 1 - 2 orders of magnitude higher potency compared to no S9, for example 167-fold more potent in the case of 4 hr γH2AX.

Anisomycin

Anisomycin caused concentration-dependent reductions to RNC, and as shown by Figure 6, the presence of S9 muted the effect. In fact, as shown by Table II, the largest effect S9 had on Anisomycin was this S9-related protection effect on %RNC. The genotoxicity biomarkers did not respond to any appreciable level, with the exception of a slight increase in 24 hr γH2AX at the highest concentrations tested in cultures without S9. As shown by the machine learning graph, this weak effect did not result in a clastogen prediction, with or without S9.

B[a]P

Results for B[a]P are provided in Figure 7. In the presence of S9, B[a]P caused concentration-dependent reductions to RNC, whereas cultures without S9 showed no signs of cytotoxicity up to the lowest precipitating concentration. Whereas 4hr p53 responses in the presence of S9 were minimal, robust effects were observed for 24 hr p53 as well as the γH2AX biomarker at both time points. The machine learning ensemble characterized B[a]P +S9 as clastogenic. Cultures without S9 showed modestly elevated γH2AX fluorescence at 24 hr, just

enough for the machine learning algorithms to characterize B[a]P without S9 as clastogenic. Table II puts this into perspective—several orders of magnitude higher concentrations of B[a]P were required in the absence of S9 to exert this effect.

Brefeldin A

Brefeldin A caused concentration-dependent reductions to RNC, and as shown by Figure 8, S9 attenuated the effect. In fact, as shown by Table II, the presence of S9 shifted the BMD by approximately 100-fold. The genotoxicity biomarkers were unresponsive despite the considerable cytotoxicity that was induced, and as shown by the machine learning graph, neither the 0 or 0.25% S9 treatment scenario resulted in a clastogen prediction.

CCCP

Results for CCCP are provided in Figure 9. CCCP caused concentration-dependent reductions to RNC, and the absence of S9 shifted the curve to the left, approximately halving the BMD value (Table II). The genotoxicity biomarkers did not respond to any appreciable level, with the exception of a modest increase in 24 hr p53 in cultures with S9. Even so, as shown by the machine learning graph, this effect did not lead to a clastogen prediction.

CP

CP caused concentration-dependent reductions to RNC both in the presence and absence of S9 (Figure 10). However, as shown by Table II, it required on the order of 800x lower CP concentration to achieve the CES when 0.25% S9 was provided. In the presence of S9, CP

induced large increases to γH2AX at both time points, and p53 at 24 hr. In the absence of S9, CP exhibited weaker effects, both in terms of BMD values and biomarker response magnitudes. Indeed, Table II shows that the genotoxicity biomarkers exhibited between 2- and 3-orders of magnitude lower BMD values when low concentration S9 was included for 24 hr. Even so, the machine learning ensemble characterized CP, both with and without S9, as clastogenic.

Cycloheximide

Results for cycloheximide are provided in Figure 11. Cycloheximide caused concentration-dependent reductions to RNC, and S9 only slightly affected the BMD value (Table II). The genotoxicity biomarkers did not respond to any appreciable level, with the exception of a modest increase in 24 hr γH2AX in cultures lacking S9. However, as shown by the machine learning graph, cycloheximide was not predicted to be clastogenic, either with or without S9.

DB[a,l]P

Results for DB[a,l]P are provided in Figure 12. The presence of S9 caused concentration-dependent reductions to RNC, while cultures without S9 showed no signs of cytotoxicity up to the lowest precipitating concentration. Whereas 4 hr p53 exhibited a slight response in the presence of S9, late p53 and γH2AX at both time points showed large increases. Similar patters were observed in cultures exposed to DB[a,l]P without S9, but in each case the magnitude of the response was greatly reduced, and the concentrations required to reach the CES were much higher. Indeed, Table II shows that for the most responsive biomarker (γH2AX at 24 hr), 0.25% S9 increased DB[a,l]P's potency by 4 orders of magnitude. The machine learning

predictions characterized DB[a,l]P, with and without S9, as clastogenic.

DEN

DEN results are shown in Figure 13. The presence of 0.25% S9 increased cytotoxicity, and furthermore potentiated γH2AX responses by more than 3.5- and 10-fold at the 4 and 24 hr time points, respectively (Table II). The machine learning ensemble characterized DEN +S9 as clastogenic, but not until high mM concentrations were reached. On the other hand, even up to the 10 mM limit concentration, in the absence of S9 DEN did not trigger a genotoxic prediction by any of the three machine learning models.

DMBA

Results for DMBA are provided in Figure 14. Concentration-dependent reductions to RNC were observed both with and without S9, however, the presence of S9 made DMBA a much more potent cytotoxicant (21.8-fold difference in potency, Table II). DMBA +S9 exhibited robust p53 responses at both time points, as well as γH2AX at 4 hr. Whereas DMBA -S9 caused an early γH2AX effect and late p53 induction, the magnitude of the responses were reduced, and the concentrations required to reach the CES were much higher. For instance, as shown in Table II, the 4 hr γH2AX biomarker exhibited 89-fold difference in potency between the 0 and 0.25% S9 conditions. The machine learning predictions characterized DMBA clastogenic, both in the presence and absence of S9.

MMC

MMC caused concentration-dependent reductions to RNC both in the presence and absence of S9 (Figure 15), with 0.25% S9 slightly attenuating the cytotoxic effect. The p53 and γH2AX biomarkers markedly responded in a concentration-dependent manner at both time points, with and without S9. As shown by Table II, the ability of MMC to induce p53 and γH2AX tended to be slightly but consistently lower in the presence of S9. The machine learning ensemble characterized MMC, with and without S9, as clastogenic.

PhIP

Results for PhIP are provided in Figure 16. The presence of S9 caused concentration-dependent reductions to RNC, whereas cultures without S9 showed no signs of cytotoxicity up to the lowest precipitating concentration. Whereas the p53 and γH2AX biomarkers were highly responsive in the presence of 0.25% S9, no appreciable effects were observed without S9. It is therefore not surprising that Table II characterizes the potency of PhIP with S9 to be at least 3 orders of magnitude greater in the presence as opposed to the absence of S9, at least for the 4 hr γH2AX biomarker. The machine learning ensemble predicted PhIP +S9 to be clastogenic, whereas -S9 did not trigger a clastogen call.

Resorcinol

Resorcinol caused concentration-dependent reductions to RNC both in the presence and absence of S9 (Figure 17), with a considerable attenuating effect in the presence of 0.25% S9. Whereas the p53 responses were modest at 4 hr, strong induction of p53 with and without S9 occurred at 24 hr. The magnitudes of the γH2AX biomarker responses were pronounced at both

time points, and similar across S9 conditions. That being said, S9 markedly shifted the dose response curves to the right. Table II highlights this effect on BMD values, which were reduced by the presence of S9 by an order of magnitude in several instances. The machine learning ensemble characterized resorcinol, with and without S9, as clastogenic.

Thapsigargin

Results for thapsigargin are provided in Figure 18. Thapsigargin caused concentration-dependent reductions to RNC, with similar profiles occurring with and without S9. The genotoxicity biomarkers did not respond to any appreciable level. As shown by the machine learning graph, thapsigargin was not predicted to be clastogenic, despite the significant levels of cytotoxicity that were induced.

Conclusions

The results of the experiments described herein clearly support the use of 0.25% phenobarbital/β-naphthoflavone-induced rat liver S9 in combination with a NADPH regeneration system as a means to detect chemicals that require enzymatic activation to exert their genotoxic effects. Each of the 8 genotoxicants that are known to be enzymatically metabolized to more reactive electrophiles were found to be clastogenic in this system. In the instances these chemicals were positive with and without S9, the genotoxic potencies were increased by order(s) of magnitude in the presence of 0.25% S9.

Importantly, the 0.25% concentration utilized by the ToxTracker test and evaluated herein with the MultiFlow assay is non-cytotoxic to TK6 cells over an extended time frame (at least 24

hr). This concentration is on the order of 8- to 10-times lower than those that have been traditionally used, and therefore clearly address 3Rs goals. The low concentration also facilitated the continuous exposure scenario that appears to be a large contributor to the effectiveness of the system, as most of the observed genotoxic effects were markedly higher at the later (24 hr) time point. Furthermore, the ability to conduct the assay without the need for centrifugation and aspiration steps benefitted the system in at least two other significant ways. First, the data were collected in a more efficient manner, with a clear potential for fully automated cell processing. Second, the cytotoxicity measurements for replicate wells were less variable, likely attributable to the fewer processing steps required. This suggests the data were not simply being acquired more efficiently, but also in a more reliable and reproducible manner.

Ideally, the encouraging results presented here will stimulate additional work that evaluates the transferability of the low S9 strategy to other laboratories, and investigates the generalizability of the method to other cell lines and endpoints.

AUTHOR CONTRIBUTIONS

ST, AC, and KZ performed benchtop work. SMB, NH, JCB, and SDD designed experiments and chose chemicals. ST, AC, KZ, and NH created Excel files that supported efficient data capture and processing. SMB suggested BMD analyses, and SDD accomplished them. All authors contributed to the writing of the manuscript.

ACKNOWLEDGMENTS

This work was funded by a grant from the National Institute of Health/National Institute of Environmental Health Sciences (NIEHS; grant no. R44ES029014). The contents are solely the responsibility of the authors, and do not necessarily represent the official views of the NIEHS. The authors would like to thank Carol Swartz who provided advice about working Moltox's NADPH regeneration system, and Maik Schuler, Maria Engel, and Richard Spellmen who encouraged direct comparisons between 0.25% S9 and traditionally used concentrations.

CONFLICT OF INTEREST STATEMENT

The authors are/were employed by Litron Laboratories. Litron has a patent covering the flow cytometry-based assay described in this manuscript and sells a commercial kit based on these procedures: MultiFlow® DNA Damage Kit — p53, γH2AX, Phospho-Histone H3.

Figure Legends

Figure 1. MultiFlow results are shown for TK6 cells exposed to cyclophosphamide in the presence of 2% S9 that was washed out after 4 hours (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 2. MultiFlow results are shown for TK6 cells exposed to dibenzo[a,l]pyrene in the presence of 2% S9 that was washed out after 4 hours (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 3. MultiFlow p53 results are shown for TK6 cells exposed to benzo[a]pyrene in the presence of 0.25% S9 for 24 continuous hours. Whereas red triangles correspond to fluorescence obtained with fluorescent antibodies against p53, black circles are results observed when no antibodies were included in the reaction mix. The increased fluorescence observed with increasing concentrations of benzo[a]pyrene in the 'no antibodies' samples therefore corresponds to test article-associated background fluorescence, and these data were used to correct for this phenomenon (blue triangles).

Figure 4. MultiFlow results are shown for TK6 cells exposed to 2-acetylaminofluorene in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled

with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 5. MultiFlow results are shown for TK6 cells exposed to 2-amminoanthracene in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 6. MultiFlow results are shown for TK6 cells exposed to anisomycin in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 7. MultiFlow results are shown for TK6 cells exposed to benzo[a]pyrene in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note

that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 8. MultiFlow results are shown for TK6 cells exposed to brefeldin A in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 9. MultiFlow results are shown for TK6 cells exposed to carbonyl cyanide *m*-chlorophenyl hydrazine in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 10. MultiFlow results are shown for TK6 cells exposed to cyclophosphamide in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 11. MultiFlow results are shown for TK6 cells exposed to cycloheximide in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 12. MultiFlow results are shown for TK6 cells exposed to dibenzo[a,l]pryene in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -

0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 13. MultiFlow results are shown for TK6 cells exposed to diethylnitrosamine in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 14. MultiFlow results are shown for TK6 cells exposed to 7,12-dimethylbenz[a]anthracene in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 15. MultiFlow results are shown for TK6 cells exposed to mitomycin C in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that

large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 16. MultiFlow results are shown for TK6 cells exposed to 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 17. MultiFlow results are shown for TK6 cells exposed to resorcinol in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

Figure 18. MultiFlow results are shown for TK6 cells exposed to thapsigargin in the absence of S9 (red Xs), or in the presence of 0.25% S9 for 24 continuous hours (black triangles). Note that large symbols denote averages, and small symbols represent data for individual replicate wells. Relative Nuclei Counts (RNC), p53, and γH2AX data were modeled with PROAST software to provide Benchmark Dose estimates based on Critical Effect Sizes of -0.3, 0.3 and 0.3, respectively. The upper-right graph shows machine learning probabilities based on Random Forest (RF), Logistic Regression (LR), and Neural Network (NN) models.

REFERENCES

Ames BN, Durston WE, Yamasaki E, Lee FD. 1973. Carcinogens are mutagens: A simple test system combining liver homogenates for activation and bacteria for detection. PNAS 70:2281-2285.

Arif JM, Gupta RC. 1997. Microsome-mediated bioactivation of dibenzo[a,l]pyrene and identification of DNA adducts by 32P-postlabeling. Carcinogenesis 18:1999-2007.

Audebert M, Riu A, Jacques C, Hillenweck A, Jamin EL, Zalko D, Cravedi JP. 2010. Use of the γH2AX assay for assessing the genotoxicity of polycyclic aromatic hydrocarbons in human cell lines. Toxicol Lett 199:182-192.

Bernacki DT, Bryce SM, Bemis JC, Kirkland D, Dertinger SD. 2016. γH2AX and p53 in TK6 cells discriminate promutagens and nongenotoxicants in the presence of rat liver S9. Environ Mol Mutagen 57:546-558.

Bryce SM, Avlasevich SL, Bemis JC, Phonethepswath S, Dertinger SD. 2010. Miniaturized flow

cytometric *in vitro* micronucleus assay represents an efficient tool for comprehensively characterizing genotoxicity dose-response relationship. Mutat Res 703:191-199.

Bryce SM, Bemis JC, Mereness JA, Spellman RA, Moss J, Dickinson D, Schuler MJ, Dertinger SD. 2014. Interpreting in vitro micronucleus positive results: simple biomarker matrix discriminates clastogens, aneugens, and misleading positive agents. Environ Mol Mutagen 55:542-555.

Bryce SM, Bernacki DT, Bemis JC, Dertinger SD. 2016. Genotoxic mode of action predictions from a multiplexed flow cytometric assay and a machine learning approach. Environ Mol Mutagen 57:171-189.

Bryce SM, Bernacki DT, Bemis JC, Spellman RA, Engel ME, Schuler M, Lorge E, Heikkinen PT, Hemmann U, Thybaud V, Wilde S, Queisser N, Sutter A, Zeller A, Guérard M, Kirkland D, Dertinger SD. 2017. Interlaboratory evaluation of a multiplexed high information content in vitro genotoxicity assay. Environ Mol Mutagen 58:146-161.

Bryce SM, Bernacki DT, Smith-Roe SL, Witt KL, Bemis JC, Dertinger SD. 2018. Investigating the generalizability of the MultiFlow® DNA Damage Assay and several companion machine learning models with a set of 103 diverse test chemicals. Toxicol Sci 162:146-166.

Carriére V, de Waziers I, Courtois YA, Leroux JP, Beaune PH. 1992. Cytochrome P450 induction and mutagenicity of 2-aminoanthracene (2AA) in rat liver and gut. Mutat Res 268:11-20.

Cheung JR, Dickinson DA, Moss J, Schuler MJ, Spellman RA, Heard PL. Histone markers identify the mode of action for compounds positive in the TK6 micronucleus assay. Mutat Res 777:7-16.

Dertinger SD, Kraynak AR, Wheeldon RP, Bernacki DT, Bryce SM, Hall N, Bemis JC, Galloway SM, Escobar PA, Johnson GE. 2019. Predictions of genotoxic potential, mode of action, molecular targets, and potency via a tiered MultiFlow® assay data analysis strategy. Environ Mol Mutagen 60:513-533.

de Graaf AO, van den Heuvel LP, Dijkman HB, de Abreu RA, Birkenkamp KU, de White T, van der Reijden BA, Smeitink JA, Jansen JH. 2004. Bcl-2 prevents loss of mitochondria in CCCP-induced apoptosis. Exp Cell Res 299:533-540.

European Food Safety Authority (EFSA). 2010. Scientific opinion on the use of resorcinol as a food additive. EFSA Journal 8(1):1411.

Futami T, Miyagishi M, Taira K. 2005. Identification of a network involved in thapsigargin-induced apoptosis using a library of small interfering RNA expression vectors. J Biol Chem 280:826-831.

Garcia-Canton C, Anadon A, Meredith C. 2013. Assessment of the in vitro γH2AX assay by high content screening as a novel genotoxicity test. Mutat Res 757:158-166.

Ge J, Chow DN, Fessler JL, Weingeist DM, Wood DK, Engelward BP. Micropatterned comet assay enables high throughput and sensitive DNA damage quantification. Mutagenesis 30:11-19.

Hastwell PW, Chai L, Roberts KJ, Webster TW, Harvey JS, Rees RW, Walmsley RW. 2006. High-specificity and high-sensitivity genotoxicity assessment in a human cell line: validation of the GreenScreen HC *GADD45a-GFP* genotoxicity screening assay. Mutat Res 607:160-175.

Hendriks G, Atallah M, Morolli B, Calléja F, Ras-Verloop N, Huijskens I, Raamsman M, van de

Water B, Vrieling H. 2012. The ToxTracker assay: novel GFP reporter systems that provide mechanistic insight into the genotoxic properties of chemicals. Toxicol Sci 125:285-298.

Hsieh JH, Smith-Roe SL, Huang R, Sedykh A, Shockley KR, Aurerbach SS, Merrick BA, Xia M, Tice RR, Witt KL. 2019. Identifying compounds with genotoxicity potential using Tox21 high-throughput screening assays. Chem Res Toxicol 32:1384-1401.

Khoury L, Zalko D, Audebert M. 2016. Complementarity of phosphorylated histones H2AX and H3 quantification in different cell lines for genotoxicity screening. Arch Toxicol 90:1983-1995.

Kirkland D, Kasper P, Martus H-J, Müller L, van Benthem J, Madia F, Corvi R. 2016. Updated recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests. Mutat Res 795:7-30.

Krais AM, Speksnijder EN, Melis JPM, Singh R, Caldwell A, de Costa CG, Luijten M, Phillips DH, Arlt VM. 2016. Metabolic activation of 2-amino-1-methyl-6-phenylimidazo [4,5-*b*]pyridine and DNA adduct formation depends on p53: Studies in *Trp53(+/+)*, *Trpp53(+/-)* and *Trp53(-/-)* mice. Molecular Cancer Biology 138:976-982.

Le Hégarat L, Mourot A, Huet S, Vasseur L, Camus S, Chesné C, Fessard V. 2014. Performance of comet and micronucleus assays in metabolically competent HepaRG cells to predict in vivo genotoxicity. Toxicol Sci 138:300-309.

Li HH, Hyduke DR, Chen R, Heard P, Yauk CL, Aubrecht J, Fornace AJ Jr. 2015. Development of a toxicogenomics signature for genotoxicity using a dose-optimization and informatics strategy in human cells. Environ Mol Mutagen 56:505-519.

Li HH, Chen R, Hyduke DR, Williams A, Frötschl R, Ellinger-Ziegelbauer H, O'Lone R, Yauk CL,

Aubrecht J, Forance AJ Jr. 2017. Development and validation of a high-throughput transcriptomic biomarker to address 21st century genetic toxicology needs. PNAS 114:E10881-E10889.

Moon JL, Kim SY, Shin SW, Park J-W. 2012. Regulation of brefeldin A-induced ER stress and apoptosis by mitochondrial NADP+ -dependent isocitrate dehydrogenase. *Biochem* Biophys Res Commun 417:760-764.

Nikolova T, Dvorak M, Jung F, Adam I, Krämer E, Gerhold-Ay A, Kaina B. 2014. γH2AX assay for genotoxic and nongenotoxic agents: Comparison of H2AX phosphorylation with cell death response. Toxicol Sci 140:103-117.

Otteneder M, Lutz WK. Correlation of DNA adduct level with tumor incidence: carcinogenic potency of DNA adducts. Mutat Res 424:237-247.

Rodriguez-Antona C, Ingelman-Sundberg M. 2006. Cytochrome P450 pharmacogenetics and cancer. Oncogene 25:1679-1691.

Russell WMS, Burch RL. 1959. The principles of humane experimental technique. London: Methuen and Co.

Smart DJ, Ahmedi KP, Harvey JS, Lynch AM. 2011. Genotoxicity screening via the γH2AX by flow assay. Mutat Res 715:25-31.

Tsamou M, Jennen DG, Claessen SM, Magkoufopoulou C, Kleinjans JC, van Delft JH. 2012. Performance of in vitro γH2AX assay in HepG2 cells to predict in vivo genotoxicity. Mutagenesis 27:645-652.

Westerink WM, Schoonen WG. 2007. Cytochrome P450 enzyme levels in HepG2 cells and cryopreserved primary human hepatocytes and their induction in HepG2 cells. Toxicol In

Vitro 21:1581-1591

Wills JW, Johnson GE, Doak SH, Soeteman-Hernández LG, Slob W, White PA. 2015. Empirical analysis of BMD metrics in genetic toxicology part I: in vitro analyses to provide robust potency rankings and support MOA determinations. Mutagenesis 31:255-263.

Yamazaki H, Oda Y, Funae Y, Imaoka S, Inui Y, Guengerich FP, Shimada T. 1992. Participation of rat liver cytochrome P450 2E1 in the activation of n-nitrosodimethylamine and n-nitrosodiethylamine to products genotoxic in an acetyltransferase-overexpressing *Salmonella typhimurium* strain (NM2009). Carcinogenesis 13:979-985.

Youngblom JH, Wiencke JK, Wolff S. 1989. Inhibition of the adaptive response of human lymphocytes to very low doses of ionizing radiation by the protein synthesis inhibitor cycloheximide. Mutat Res 227:257-261.

Table II. Benchmark Dose Estimates.

	BMD (μM)				
Treatment	**4hr p53**	**24hr p53**	**4hr γH2AX**	**24hr γH2AX**	**24hr RNC**
2AAF without S9	>311	>311	>311	>311	98
2AAF with S9	>311	>311	119	35	184
S9 Potentiation Ratio	NC	NC	>2.6	>8.9	0.53
2AAN without S9	>55	>55	40	10.9	>55
2AAN with S9	7.9	2.5	0.24	0.67	2.1
S9 Potentiation Ratio	7	>22	167	16.3	>26.2
Anisomycin without S9	>2.5	>2.5	>2.5	2.1	0.08
Anisomycin with S9	>14.1	>14.1	>14.1	>14.1	0.48
S9 Potentiation Ratio	NC	NC	NC	<0.15	0.17
B[a]P without S9	>88	>88	>88	36	>88
B[a]P with S9	>31	1.1	0.19	2x10^-5	2.8
S9 Potentiation Ratio	NC	>80	>463	1.8x10^6	>31.4
Brefeldin A without S9	>0.5	>0.5	>0.5	>0.5	0.13
Brefeldin A with S9	>356	>356	>356	>356	12.8
S9 Potentiation Ratio	NC	NC	NC	NC	0.01
CCCP without S9	>35	>35	>35	>35	2.0
CCCP with S9	>35	26	>35	>35	4.4
S9 Potentiation Ratio	NC	>1.3	NC	NC	0.45
CP without S9	>10,000	7,570	>10,000	5,523	5,981
CP with S9	>156	8.7	99	1.7	7.5
S9 Potentiation Ratio	NC	870	>101	3,249	797
Cycloheximide without S9	>15	>15	>15	12	0.25
Cycloheximide with S9	>15	>15	>15	>15	0.31

S9 Potentiation Ratio	NC	NC	NC	<0.8	0.8
DB(a,l)P without S9	>50	>50	25	>50	>50
DB(a,l)P with S9	0.9	0.052	0.0011	0.0013	0.041
S9 Potentiation Ratio	>56	>962	22,727	>38,462	>1,220
DEN without S9	>10,000	>10,000	>10,000	>10,000	>10,000
DEN with S9	>10,000	>10,000	2,867	1,004	>10,000
S9 Potentiation Ratio	NC	NC	>3.5	>10	NC
DMBA without S9	>194	116	13.3	>194	98
DMBA with S9	5.9	4.9	0.15	>194	4.5
S9 Potentiation Ratio	>32.9	23.7	88.7'	NC	21.8
MMC without S9	0.802	0.0988	0.05031	0.00862	0.07407
MMC with S9	1.576	0.136	0.06122	0.03391	0.2363
S9 Potentiation Ratio	0.51	0.73	0.82	0.25	0.31
PhIP without S9	>178	>178	>178	>178	>178
PhIP with S9	56.81	5.142	0.1093	0.2281	9.764
S9 Potentiation Ratio	>3.1	>34.6	>1,629	>780	>18.2
Resorcinol without S9	>12.5	2.6	6.6	0.79	0.59
Resorcinol with S9	>71	24.6	25.8	14.4	12.6
S9 Potentiation Ratio	NC	0.11	0.26	0.05	0.047
Thapsigargin without S9	>21	>21	>21	>21	0.2
Thapsigargin with S9	>15	>15	>15	>15	0.14
S9 Potentiation Ratio	NC	NC	NC	NC	1.4

Abbreviations: Chemical abbreviations = same as Table I; BMD = benchmark dose; RNC = relative nuclei count; NC = not calculated

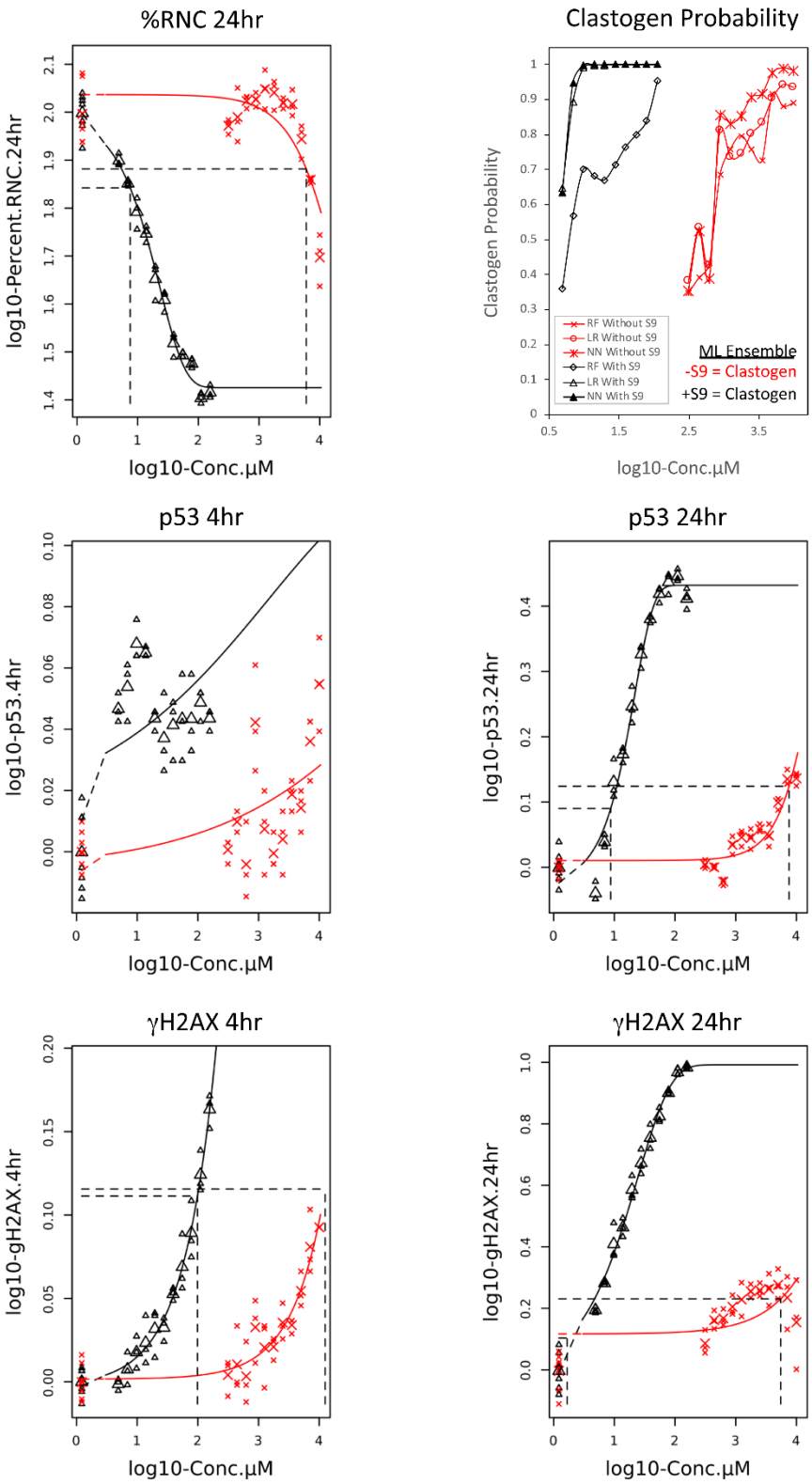
Table I. Test chemicals, source, biological effects.

Chemical (Abbreviation)	**CAS No., Source**	**Notes about Biotransformation, Miscellaneous Info**	**References**
2-Acetylaminofluorene (2AAF)	53-96-3, Sigma-Aldrich	Clastogen, requires metabolic activation (CYP1A2), forms C8 adduct on guanine	Otteneder and Lutz, 1999; Kirkland *et al.,* 2016
2-Aminoanthracene (2AAN)	613-13-8, Sigma-Aldrich	Clastogen, aromatic amine, requires metabolic activation (CYP1B1, 2A family)	Carriére *et al.,* 1992
Anisomycin	22862-76-6, Sigma-Aldrich	Cytotoxicant, protein synthesis inhibitor; *in vitro* MN neg. with high levels of apoptosis	Personal communication, Maik Schuler, Richard Spellman, Maria Engel
Benzo[a]pyrene (B[a]P)	50-32-8, Sigma-Aldrich	Clastogen, polycyclic aromatic hyrocarbon, requires metabolic activation (CYP1A1, 1B1, epoxide hydrolase), forms bulky adducts	Kirkland *et al.,* 2016
Brefeldin A	20350-15-6, Sigma-Aldrich	Cytotoxicant, ER-golgi transporter inhibitor, ER stress-induced apoptosis	Moon *et al.,* 2012
Carbonyl cyanide m-chlorophenyl hydrazone (CCCP)	555-60-2, Sigma-Aldrich	Cytotoxicant, uncouples oxidative phosphorylation	de Graaf *et al.,* 2004
Cyclophosphamide monohydrate (CP)	6055-19-2, Sigma-Aldrich	Clastogen, nitrogen mustard, requires metabolic activation (CYP2B6, CYP2C19, CYP2C9 and CYP3A4/5)	Kirkland *et al.,* 2016; Rodriguez-Antona and Inglelman-Sundberg, 2006
Cycloheximide	66-81-9, Sigma-Aldrich	Cytotoxicant, protein synthesis inhibitor	Youngblom *et al.,* 1989

Dibenzo[a,l]pyrene (DB[a,l]P)	191-30-0, Sigma-Aldrich	Clastogen, polycyclic aromatic hydrocarbon, requires metabolic activation (thought to be primarily activated by CYP1A1)	Arif and Gupta, 1997
Diethylnitrosamine (DEN)	55-18-5, Sigma-Aldrich	Clastogen, requires metabolic activation to form alkylating agent (likely involves CYP2E1 which is not highly expressed in rat liver S9); often only positive at high concentrations (≥10 mM)	Yamazaki *et al.,* 1992;
7,12-Dimethylbenzanthracene (DMBA)	57-97-6, Sigma-Aldrich	Clastogen, requires metabolic activation (CYP1B1), forms bulky adducts	Kirkland *et al.,* 2016
Mitomycin C (MMC)	50-07-7, Sigma-Aldrich	DNA-DNA crosslinks, also alkylating activity and oxidative damage	Kirkland et al., 2016
2-amino-1-methyl-6-phenylimidazo [4,5-b]pyridine (PhIP)	105650-23-5, Toronto Research Chemicals	Clastogen, heterocyclic amine, requires metabolic activation (CYP1A family)	Kirkland *et al.,* 2016; Krais *et al.,* 2016
Resorcinol	108-46-3, Sigma-Aldrich	*In vitro* mammalian cell pos. (MLA assay with and without metabolic activation pos., *in vitro* human lymphocyte MN pos. in absence of metabolic activation); *in vitro* findings not confirmed *in vivo* (mouse MN neg.)	EFSA Journal, 2010
Thapsigargin	67526-95-8, Sigma-Aldrich	Cytotoxicant, ER stress-induced apoptosis	Futami *et al.,* 2005

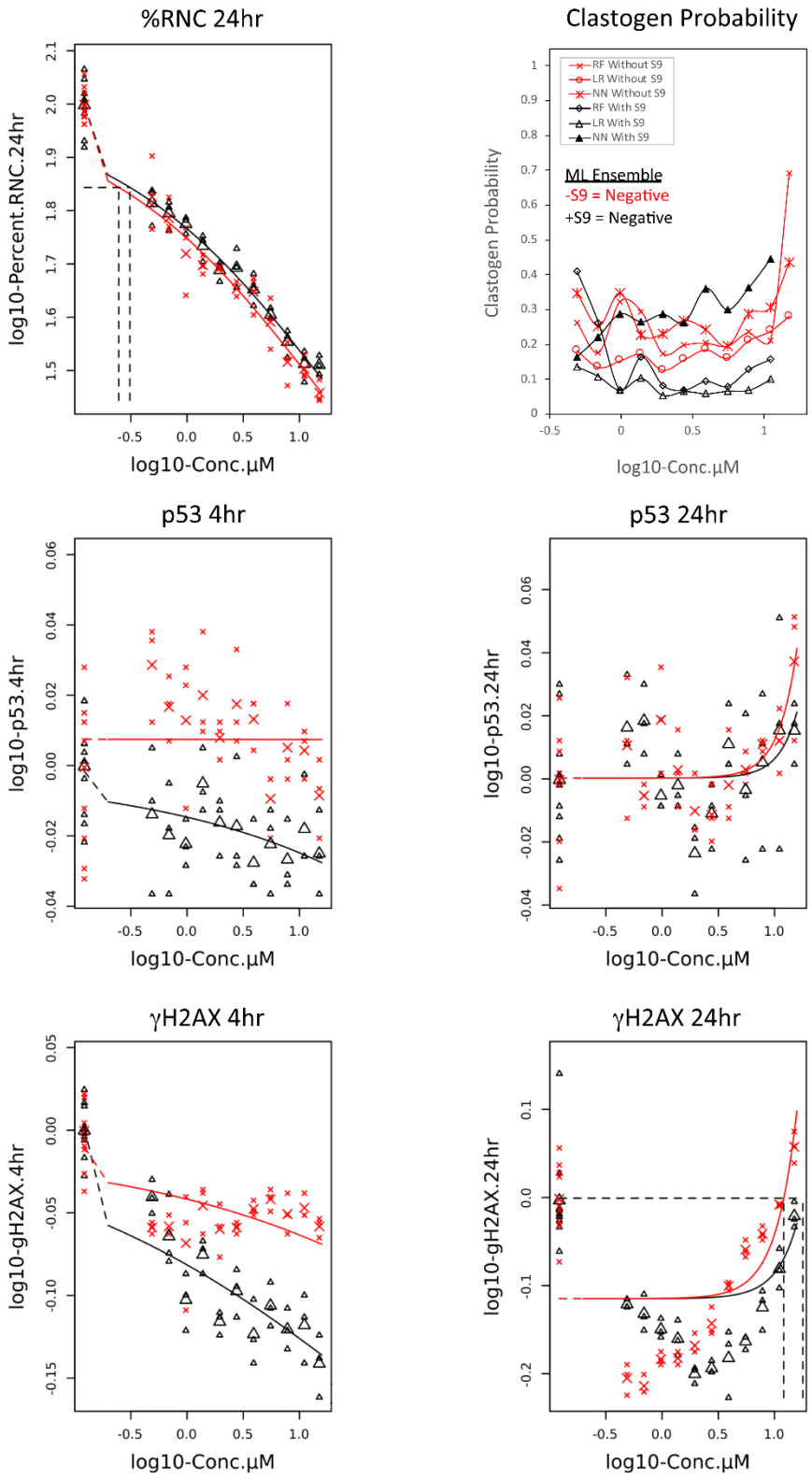
Cyclophosphamide

X = without S9 Δ = with S9



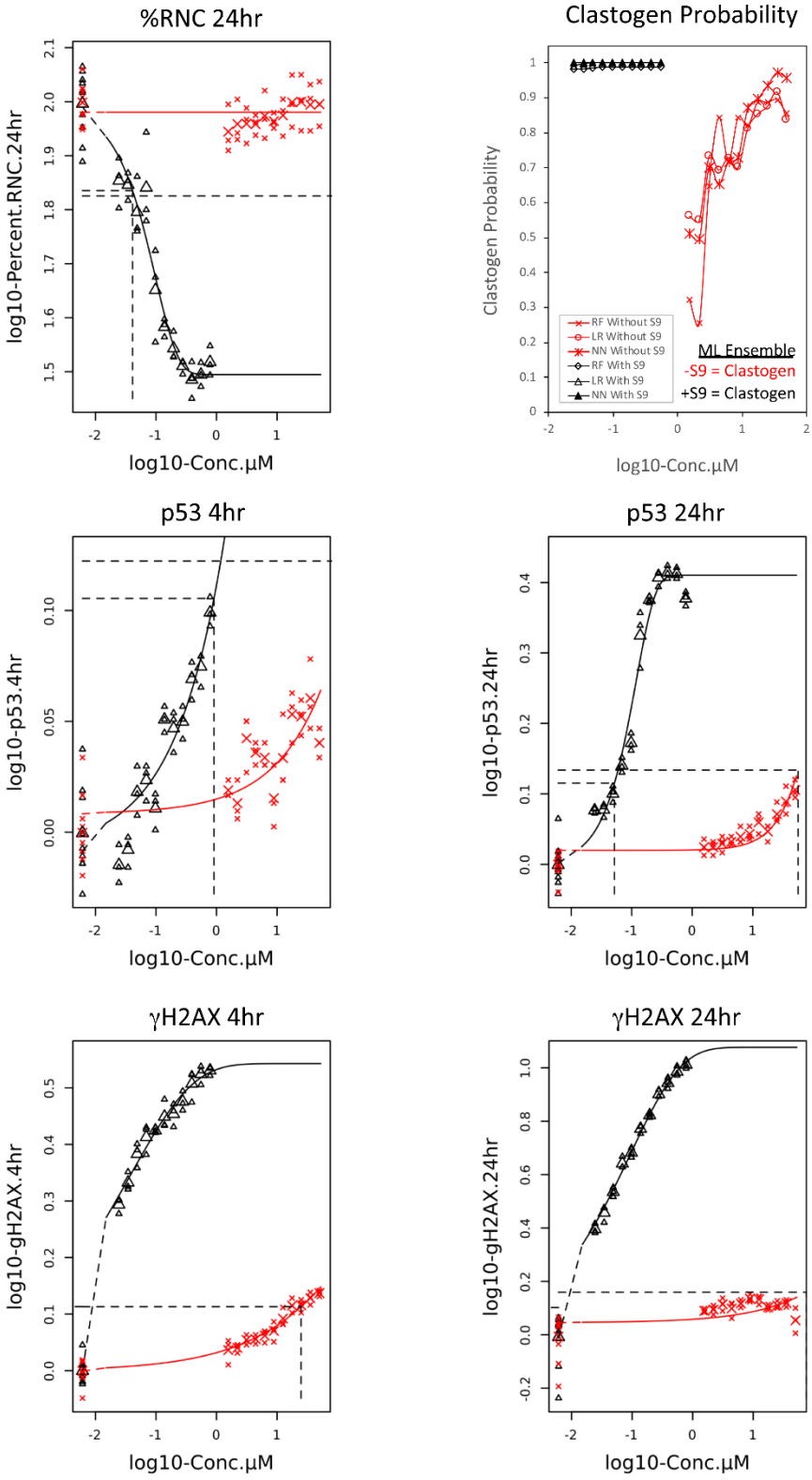
Cycloheximide

X = without S9 Δ = with S9



Dibenzo[a,l]pyrene

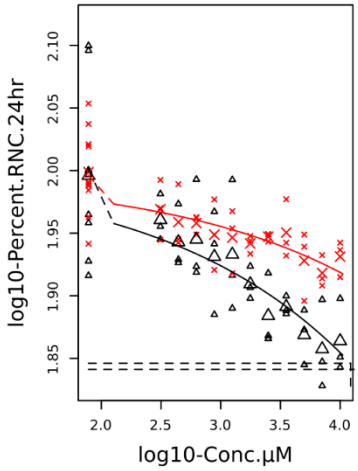
X = without S9 Δ = with S9



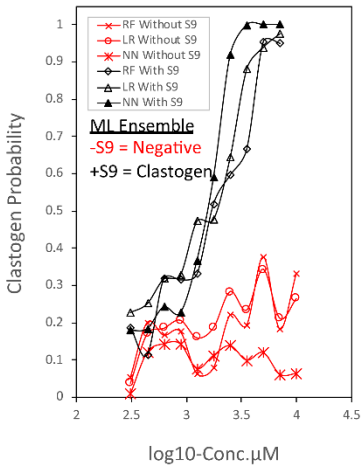
Diethylnitrosamine

X = without S9 Δ = with S9

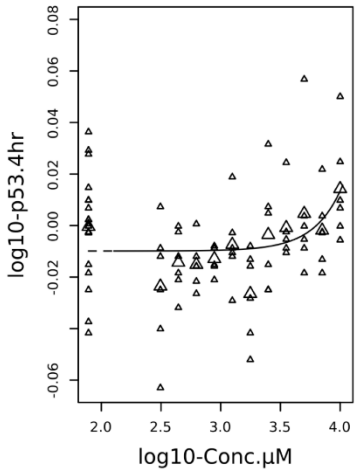
%RNC 24hr



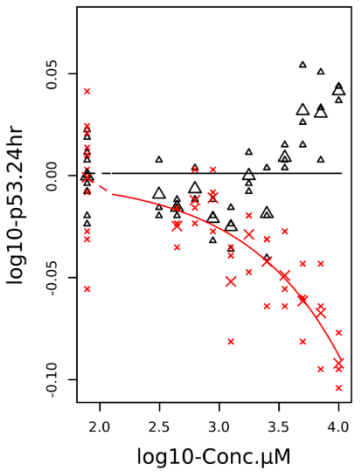
Clastogen Probability



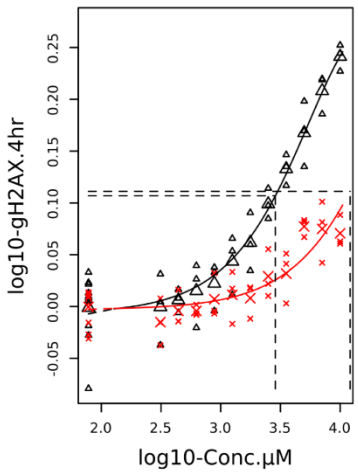
p53 4hr



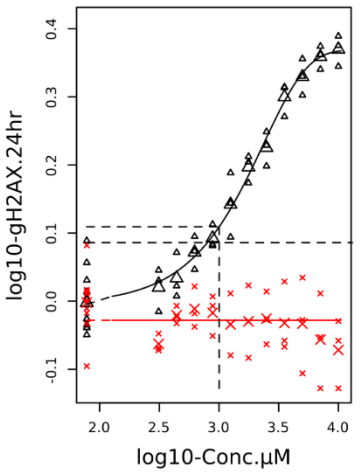
p53 24hr



γH2AX 4hr

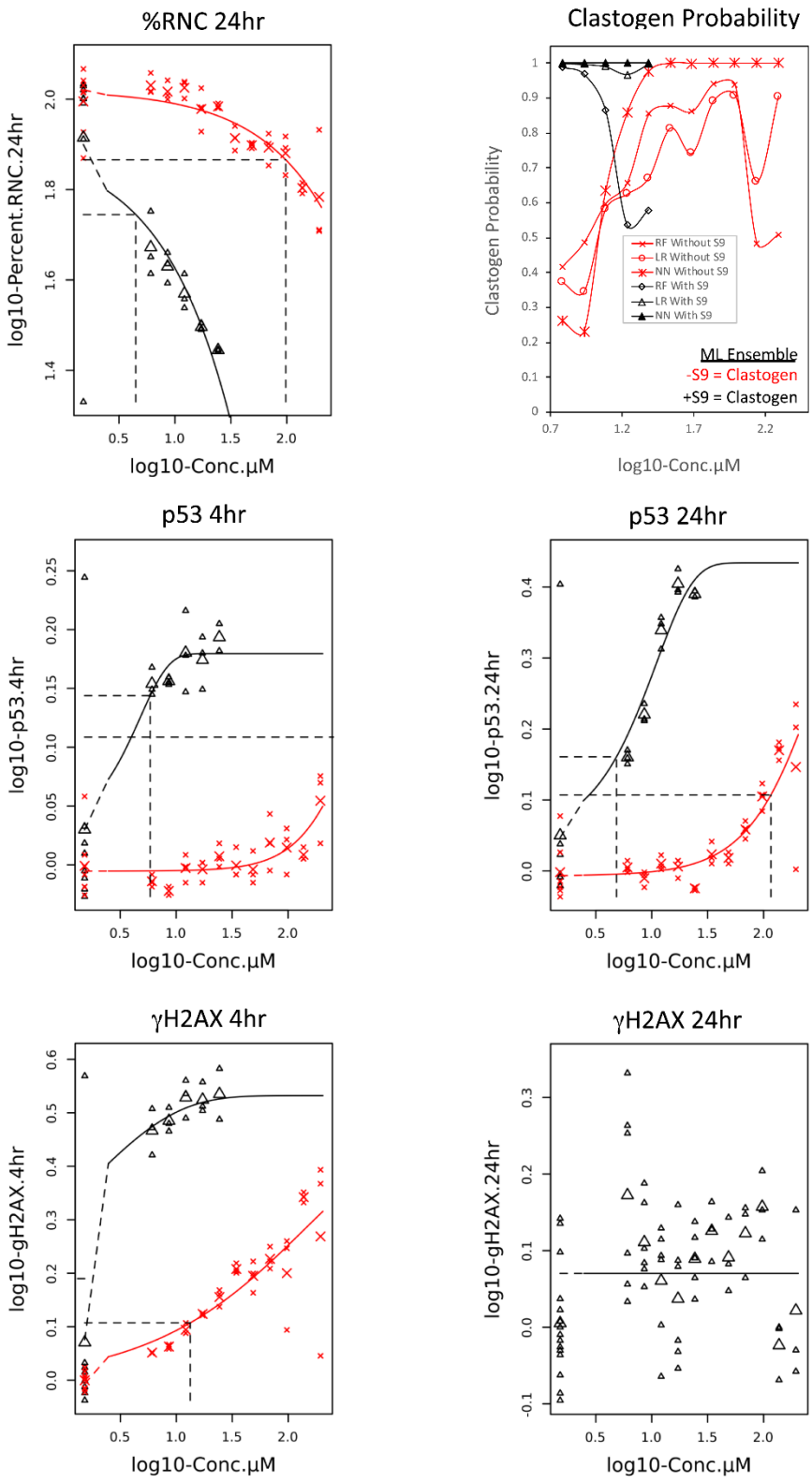


γH2AX 24hr



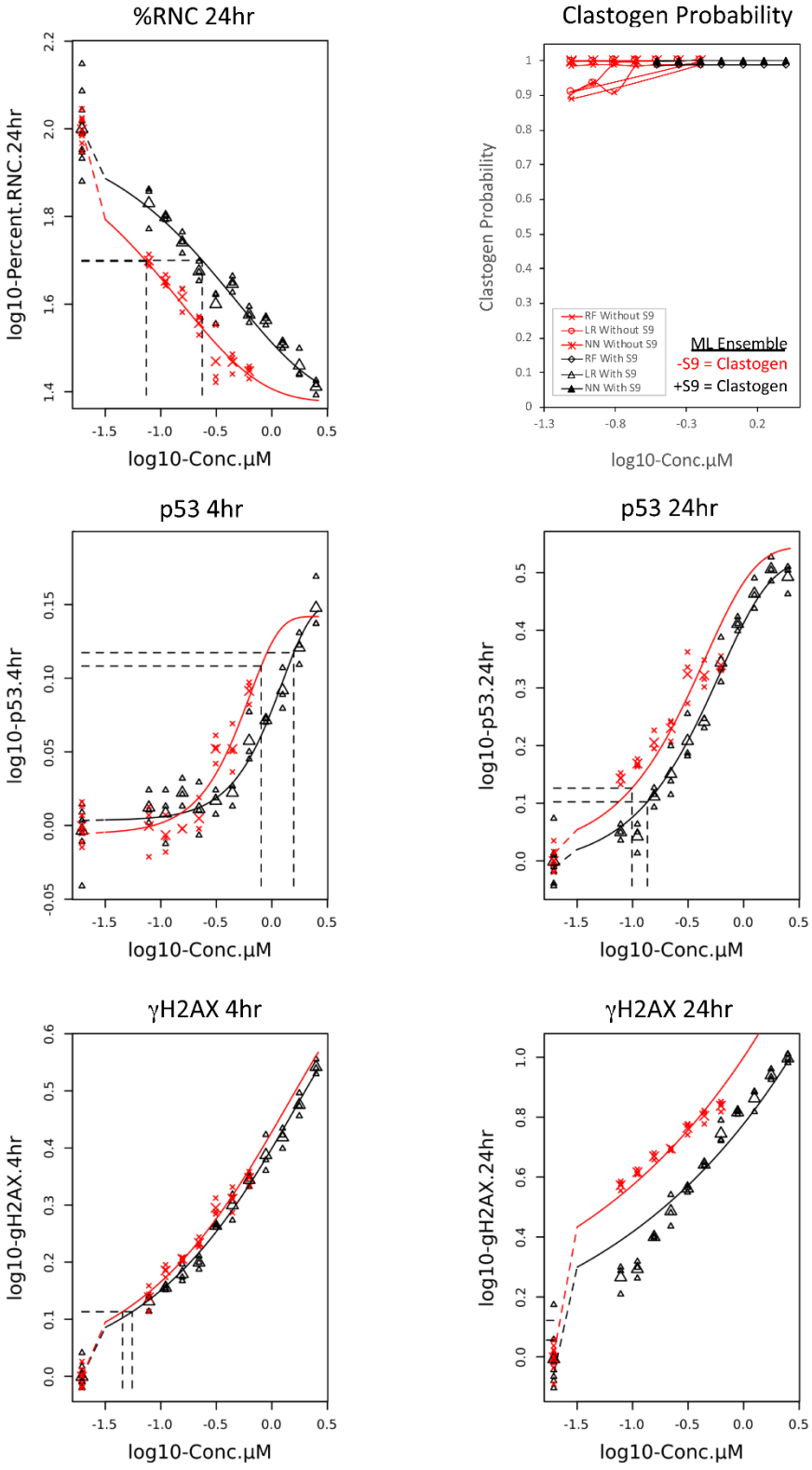
7,12-Dimethylbenz[a]anthracene

X = without S9 Δ = with S9



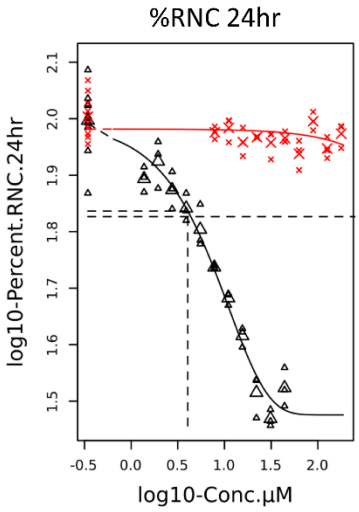
Mitomycin C

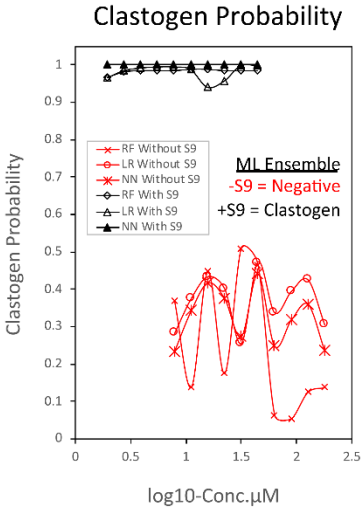
X = without S9 Δ = with S9

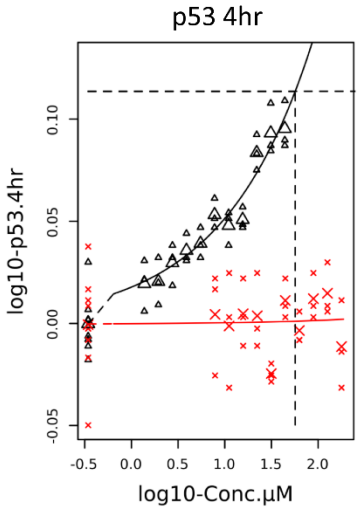


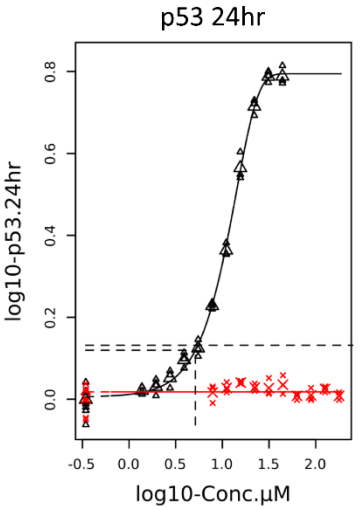
2-Amino-1-methyl-6-phenylimidazo[4,5-b]pyridine

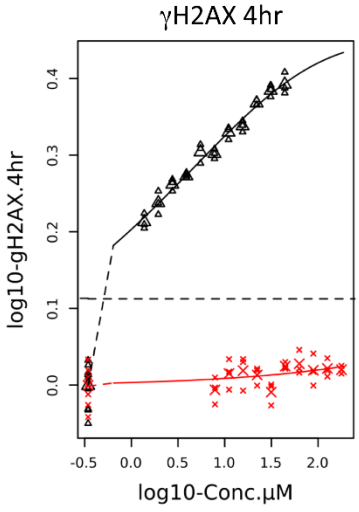
X = without S9 Δ = with S9

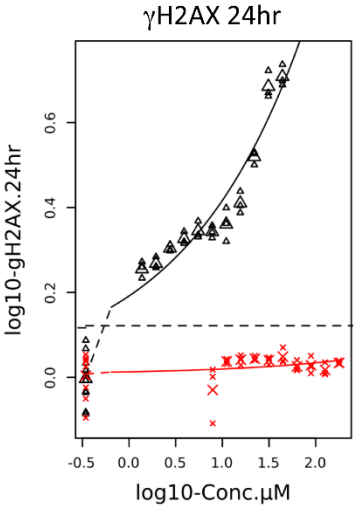








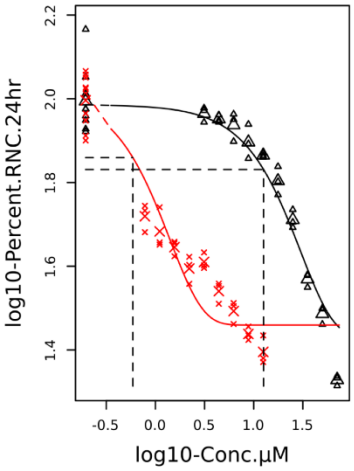




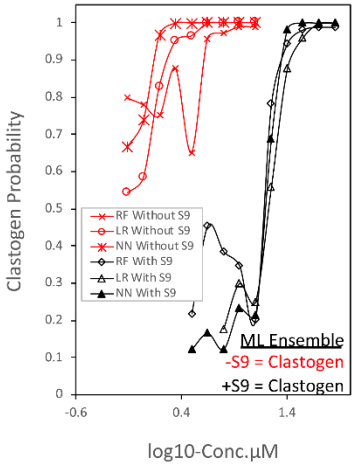
Resorcinol

X = without S9 Δ = with S9

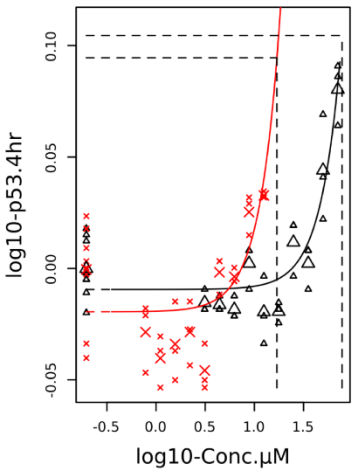
%RNC 24hr



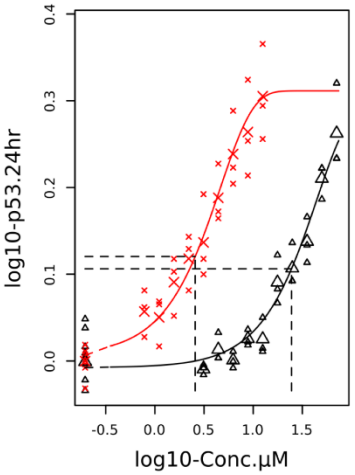
Clastogen Probability



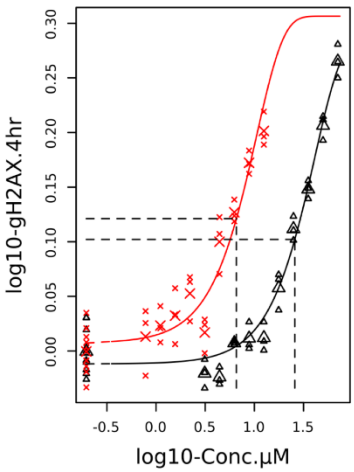
p53 4hr



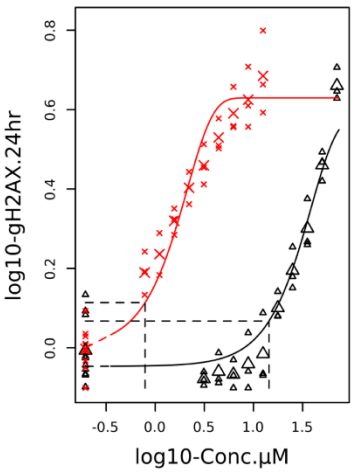
p53 24hr



γH2AX 4hr

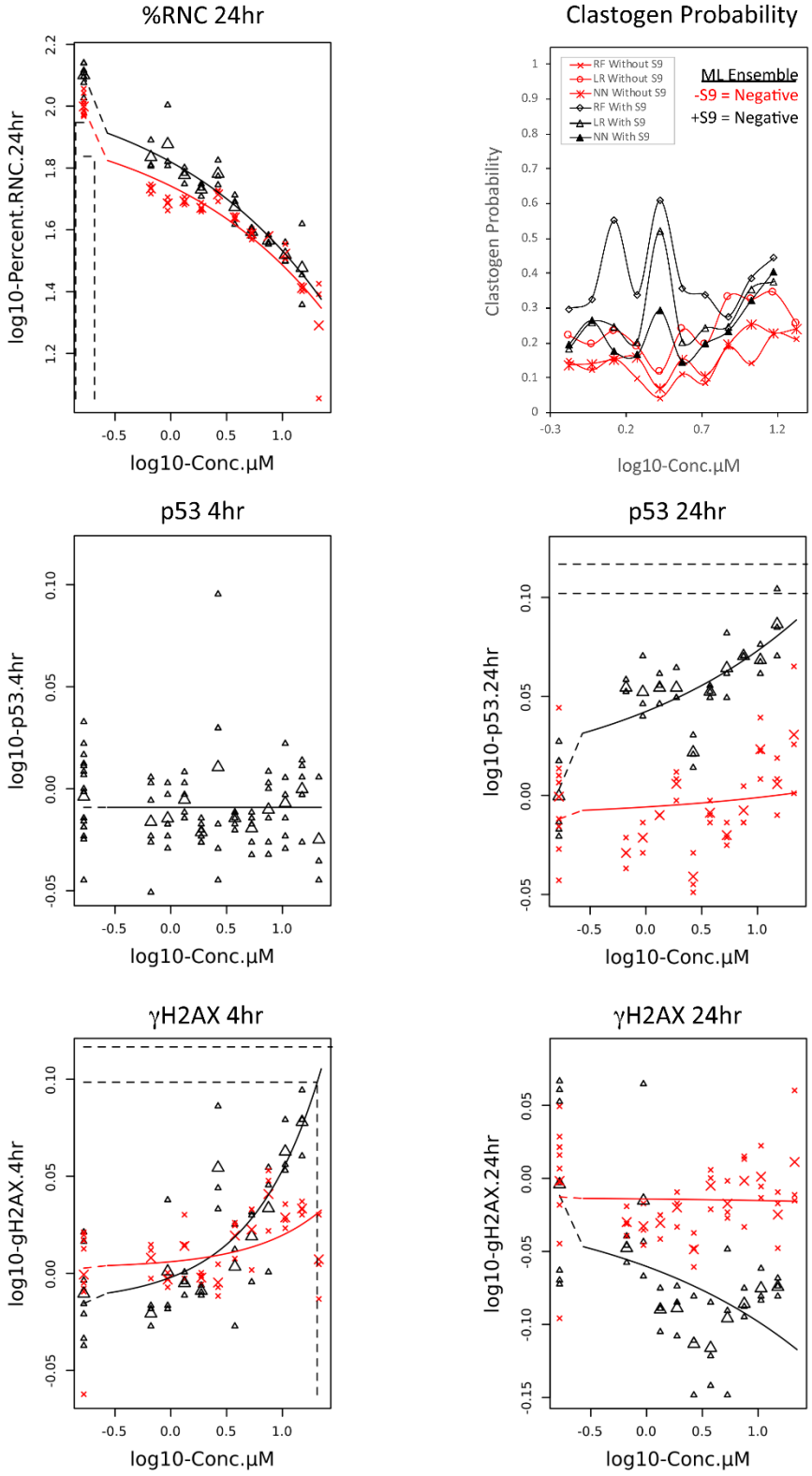


γH2AX 24hr



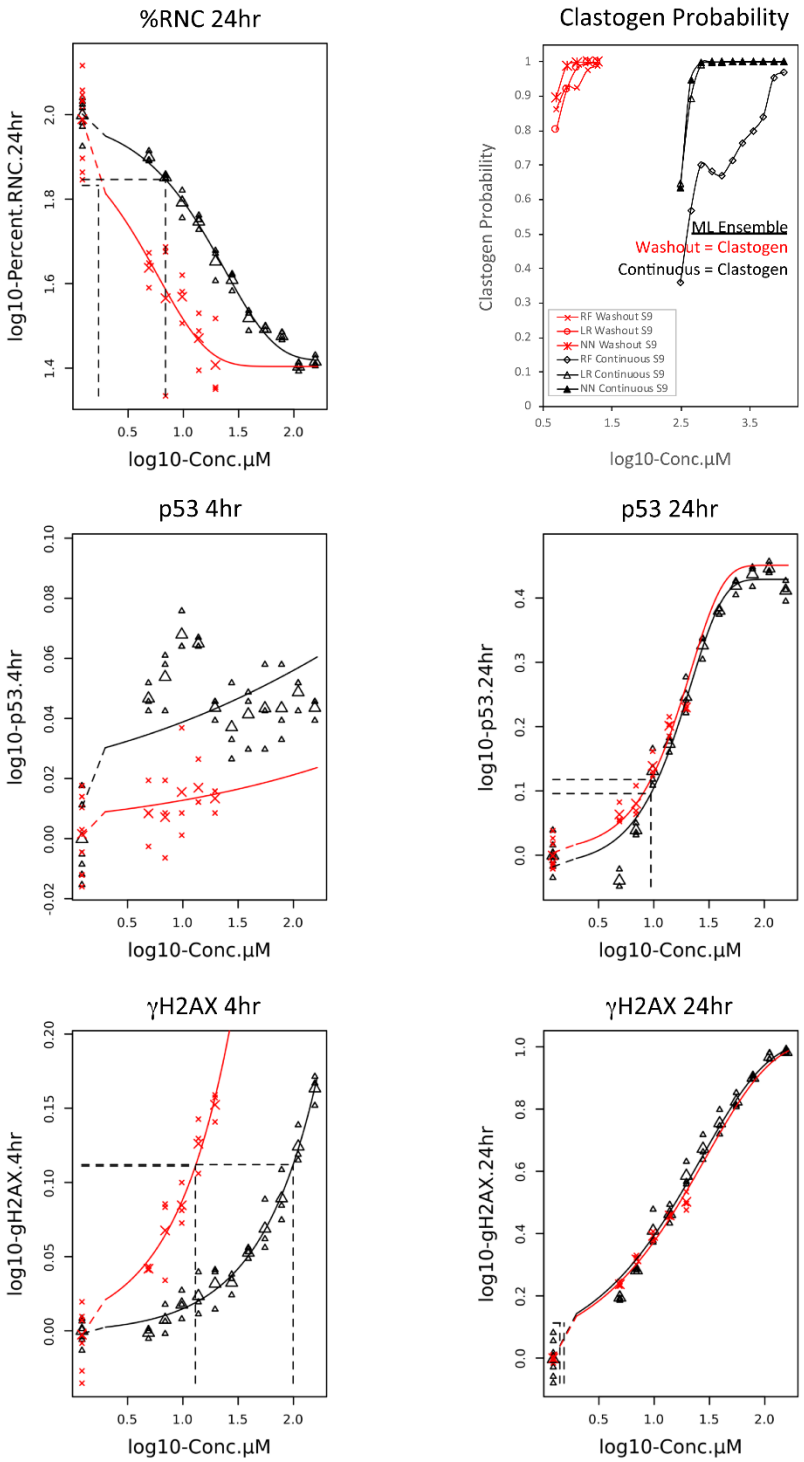
Thapsigargin

X = without S9 Δ = with S9



Cyclophosphamide

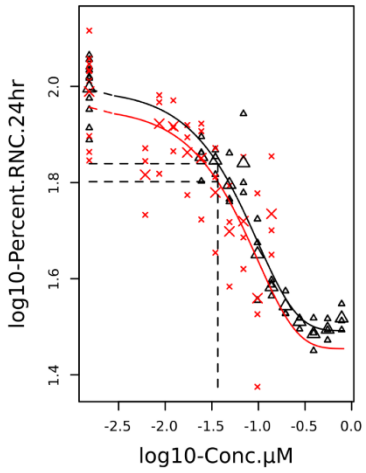
X = washout S9 Δ = continuous S9



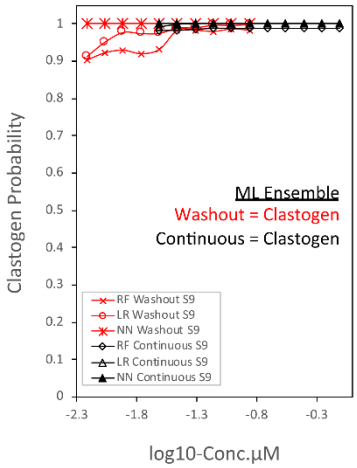
Dibenzo[a,l]pyrene

X = washout S9 Δ = continuous S9

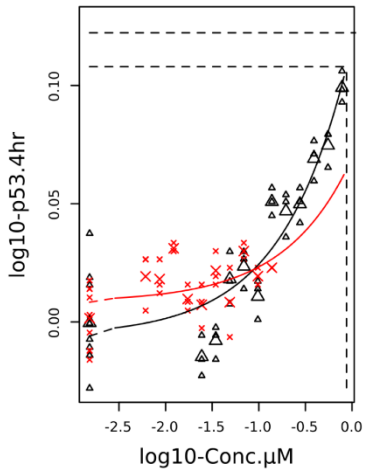
%RNC 24hr



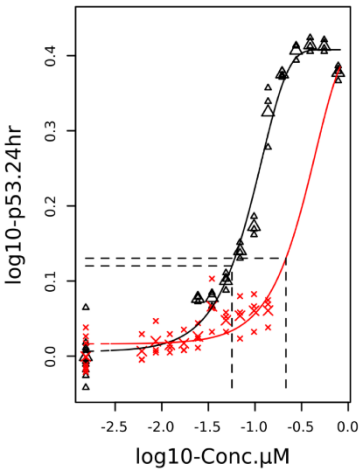
Clastogen Probability



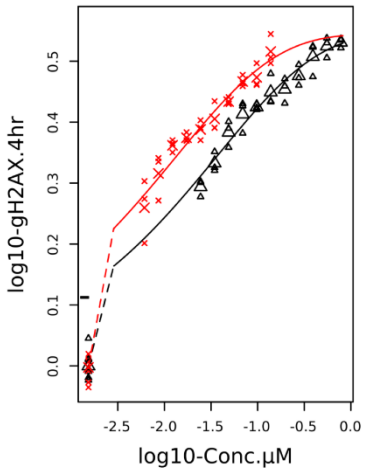
p53 4hr



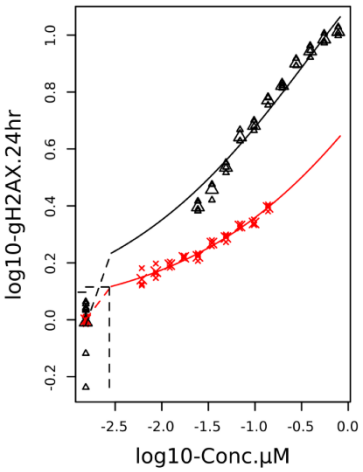
p53 24hr

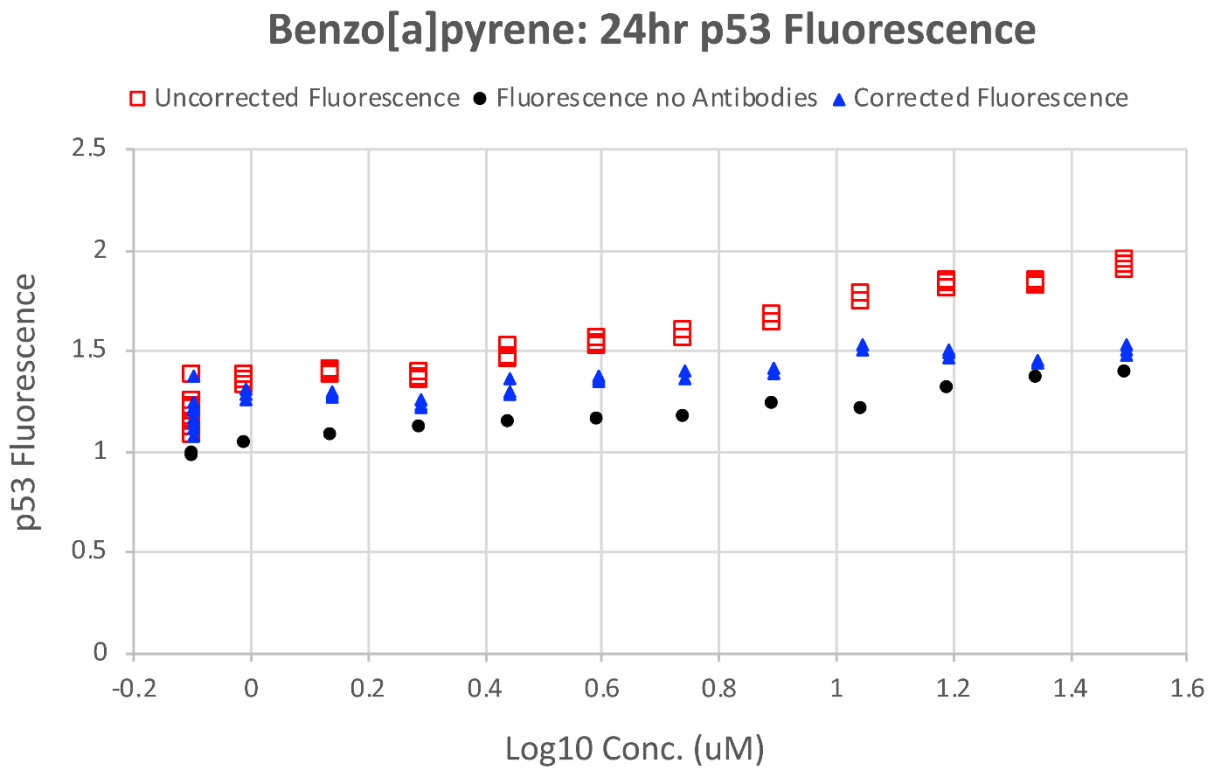


γH2AX 4hr



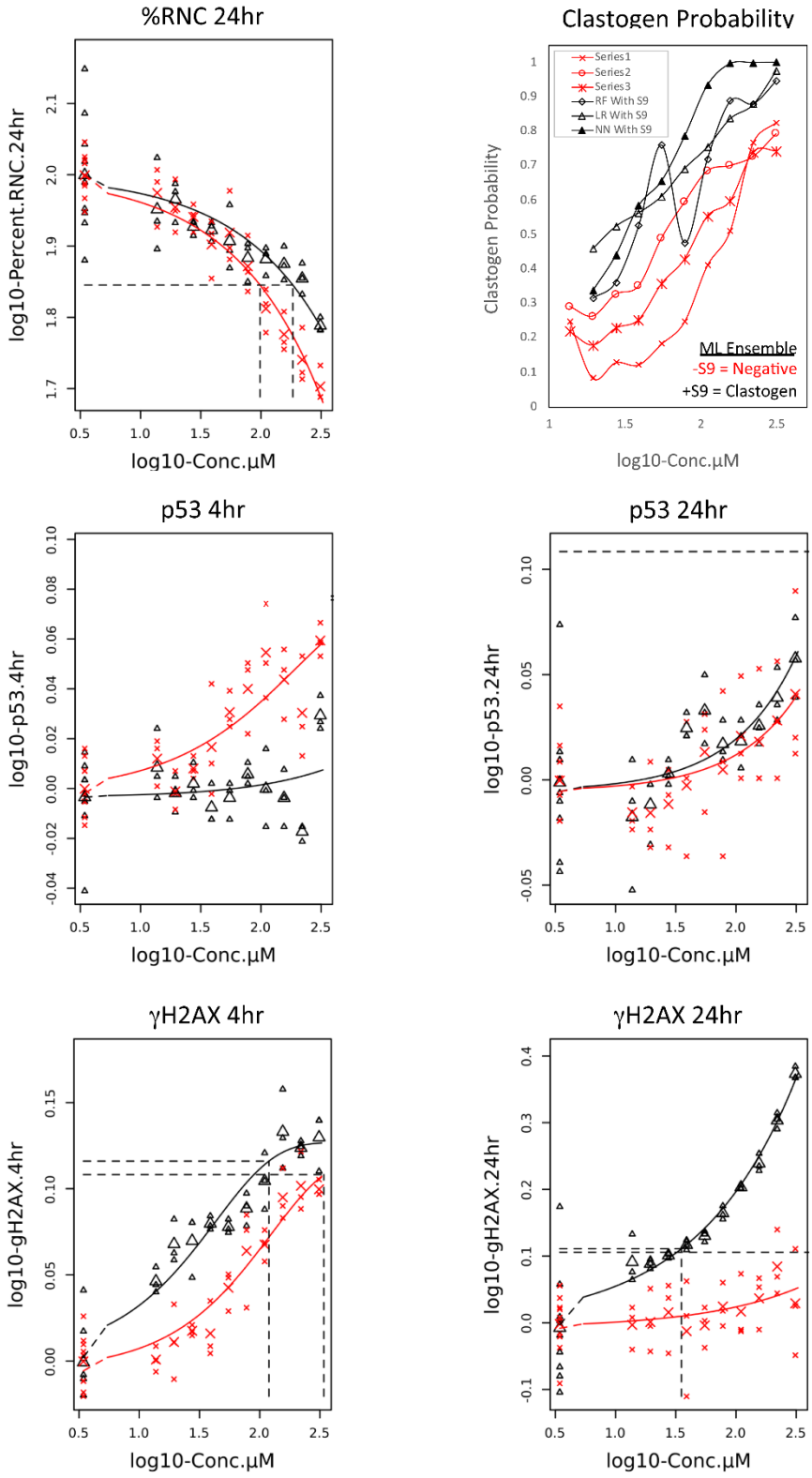
γH2AX 24hr





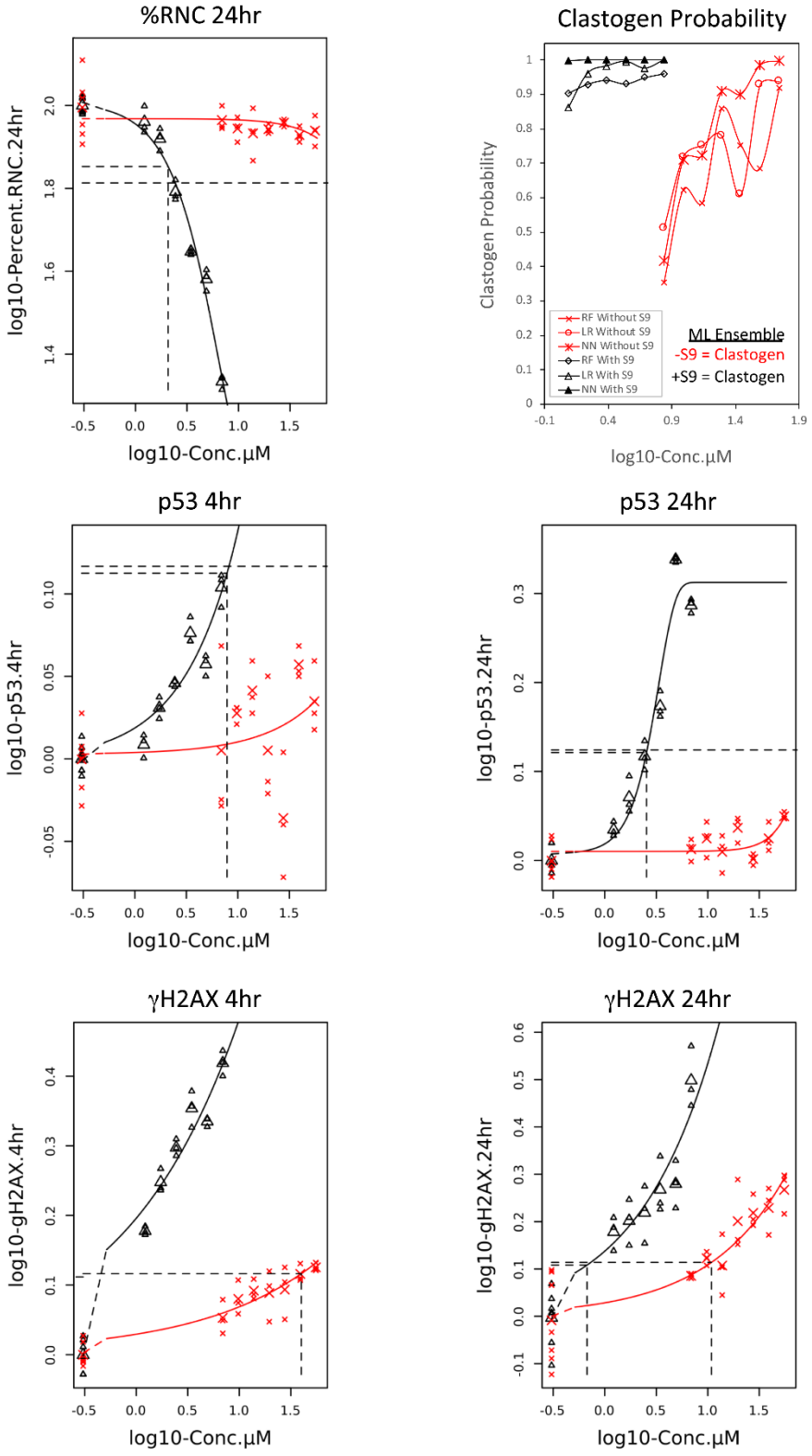
2-Acetylaminofluorene

X = without S9 Δ = with S9



2-Aminoanthracene

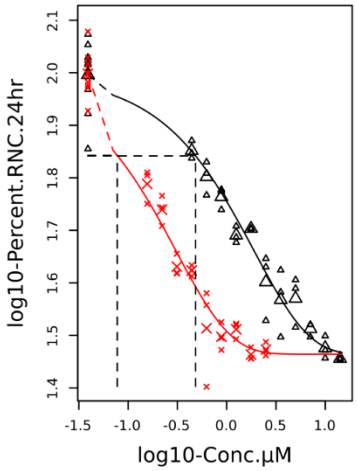
X = without S9 Δ = with S9



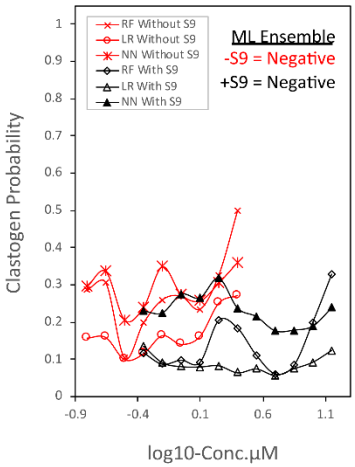
Anisomycin

X = without S9 Δ = with S9

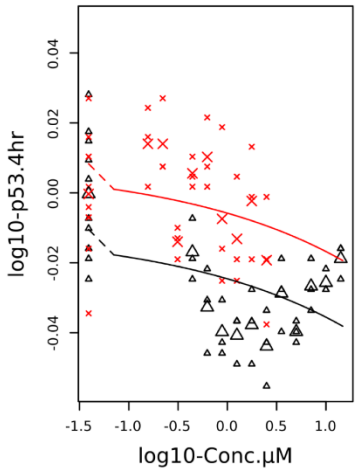
%RNC 24hr



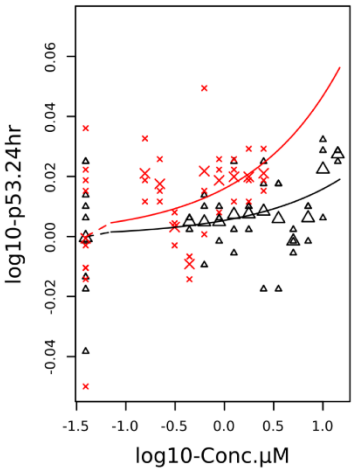
Clastogen Probability



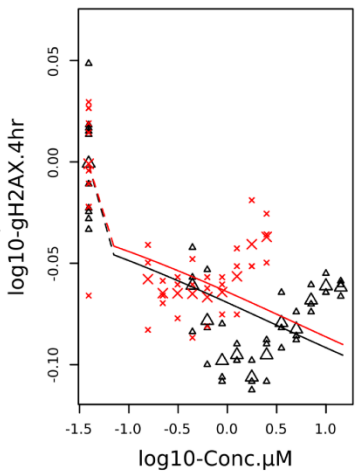
p53 4hr



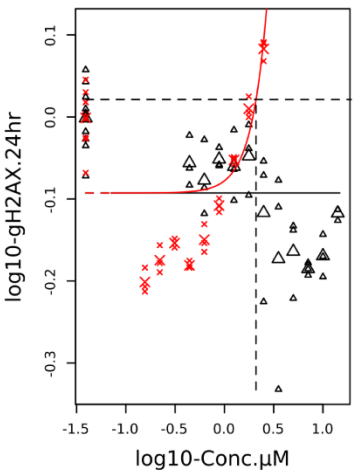
p53 24hr



γH2AX 4hr

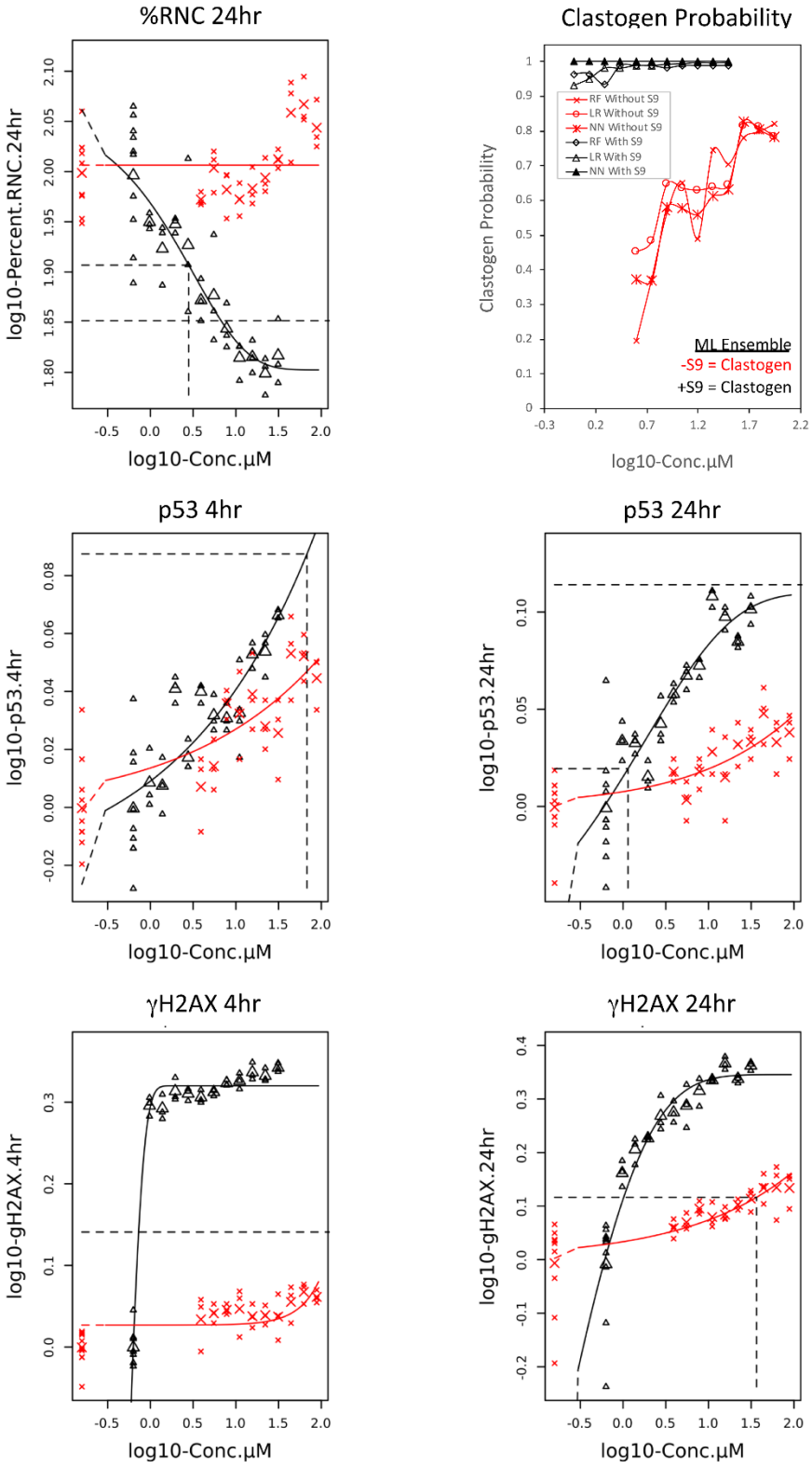


γH2AX 24hr



Benzo[a]pyrene

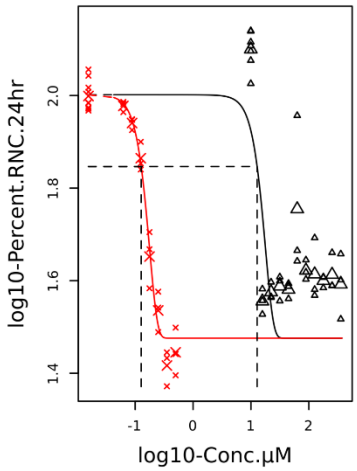
X = without S9 Δ = with S9



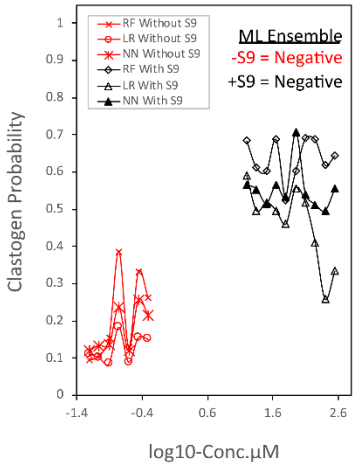
Brefeldin A

X = without S9 Δ = with S9

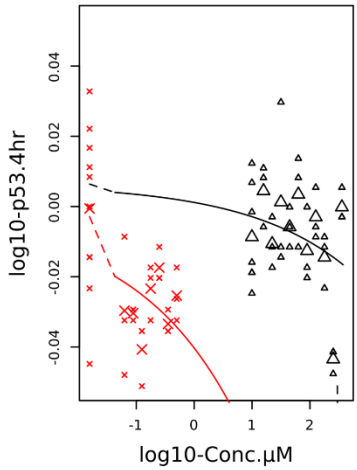
%RNC 24hr



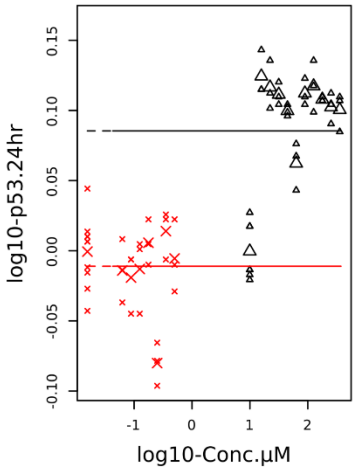
Clastogen Probability



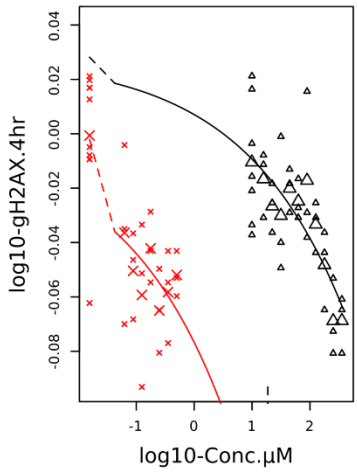
p53 4hr



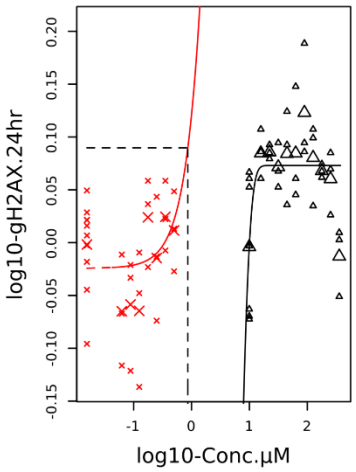
p53 24hr



γH2AX 4hr



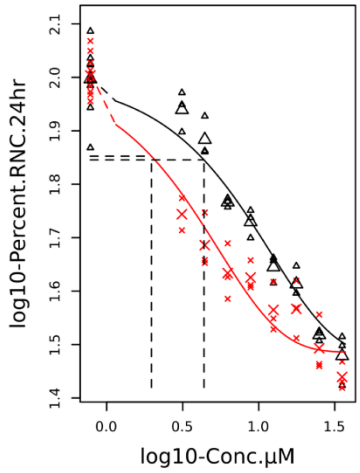
γH2AX 24hr



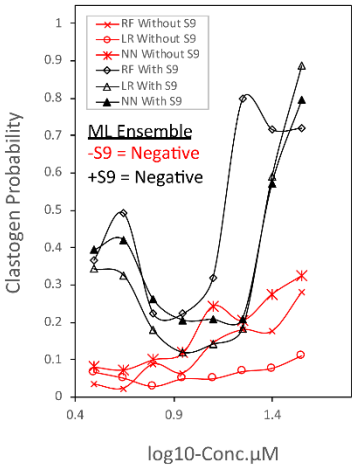
Carbonyl cyanide *m*-chlorophenyl hydrazine

X = without S9 Δ = with S9

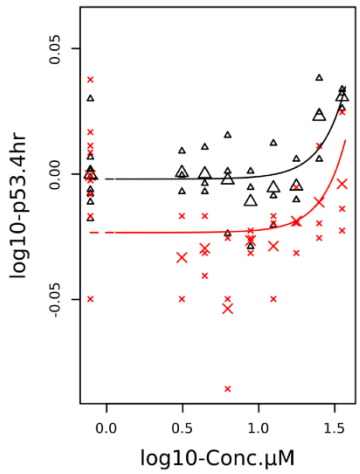
%RNC 24hr



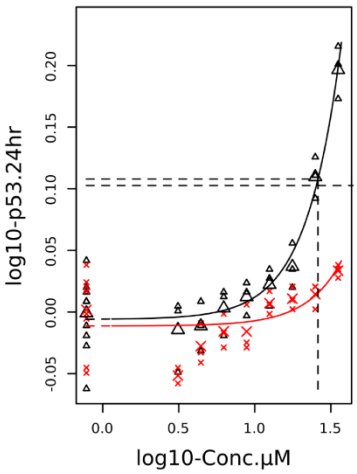
Clastogen Probability



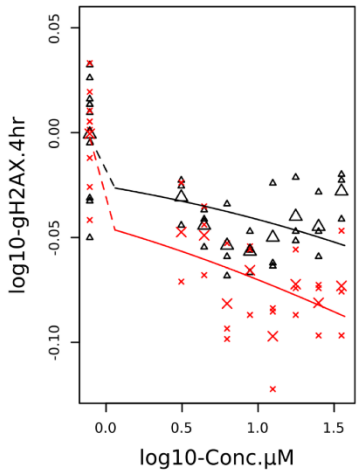
p53 4hr



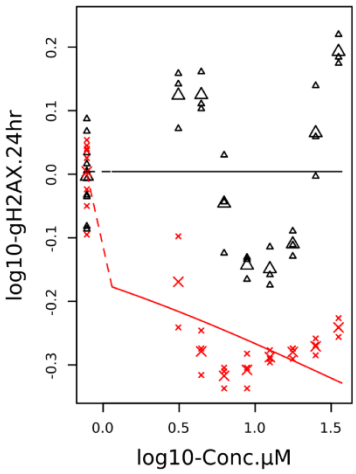
p53 24hr

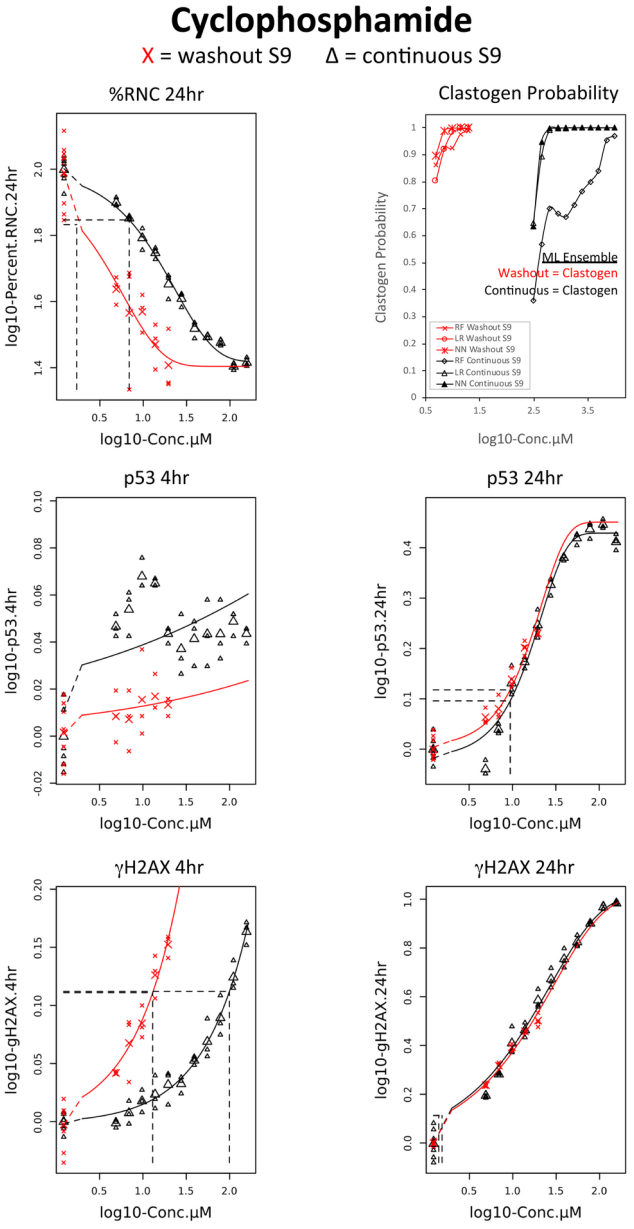


γH2AX 4hr

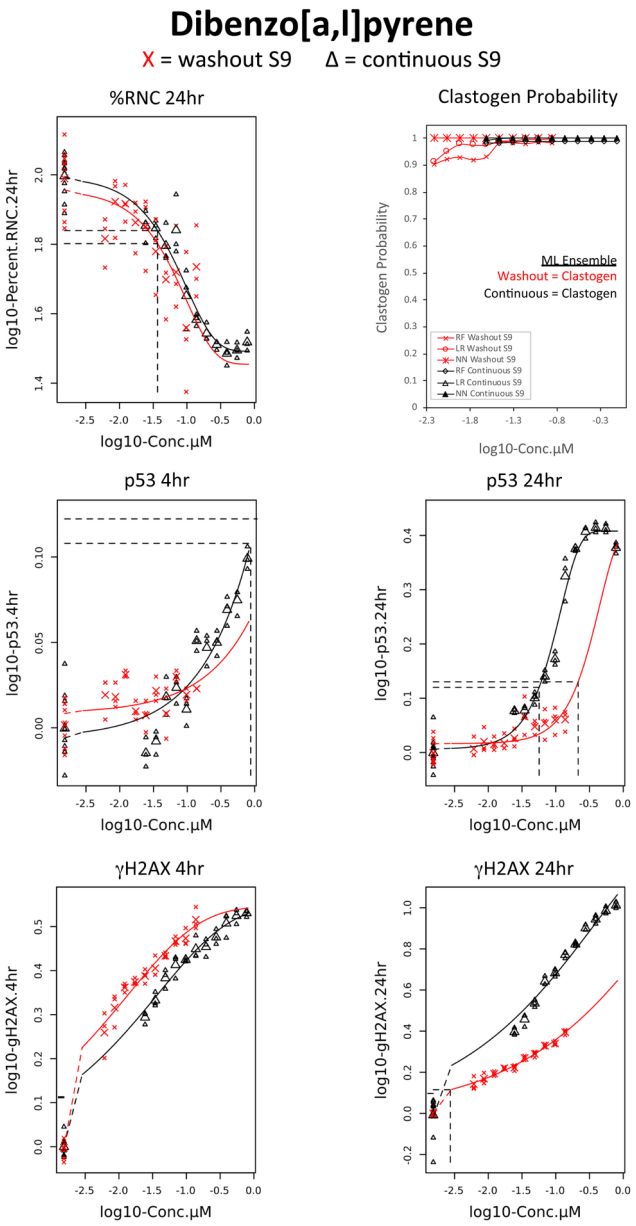


γH2AX 24hr

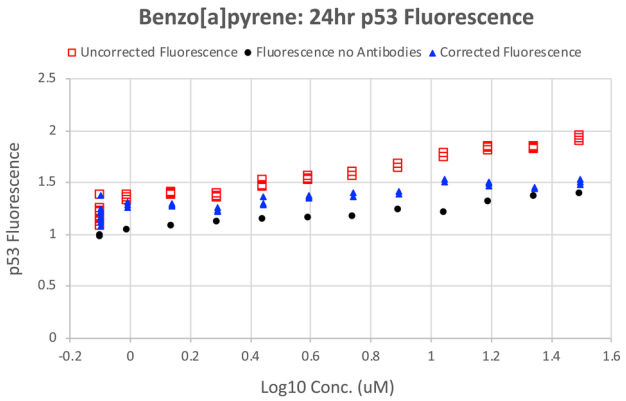




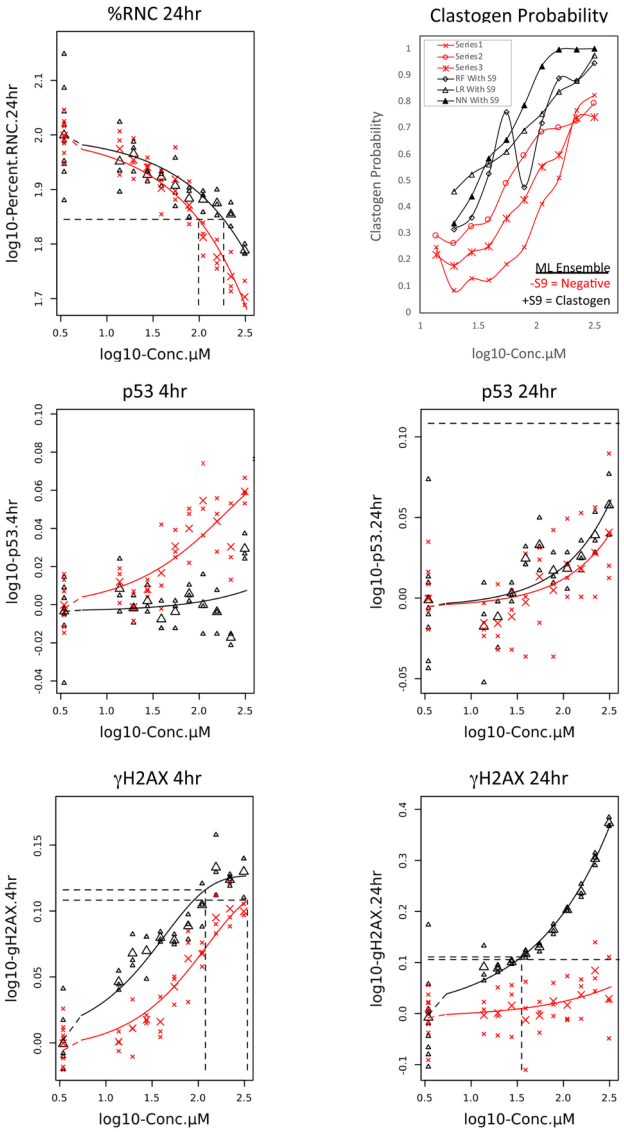
EM_22361_Figure 1_CP low vs highv191213.tif



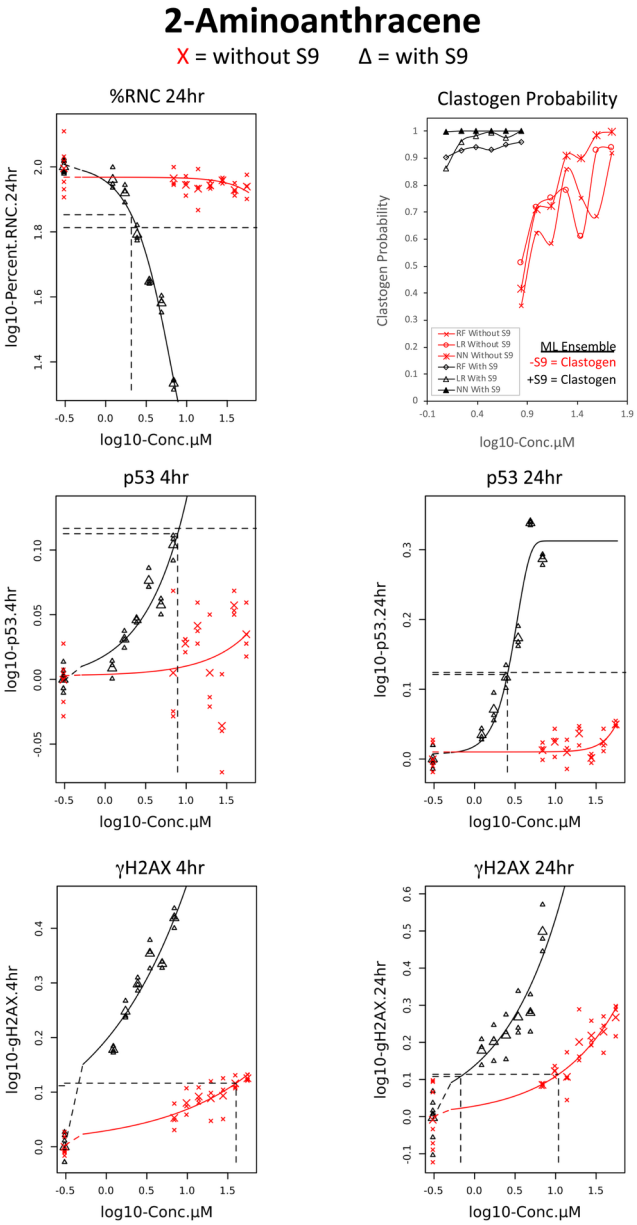
EM_22361_Figure 2_DBP low vs high_v191213.tif



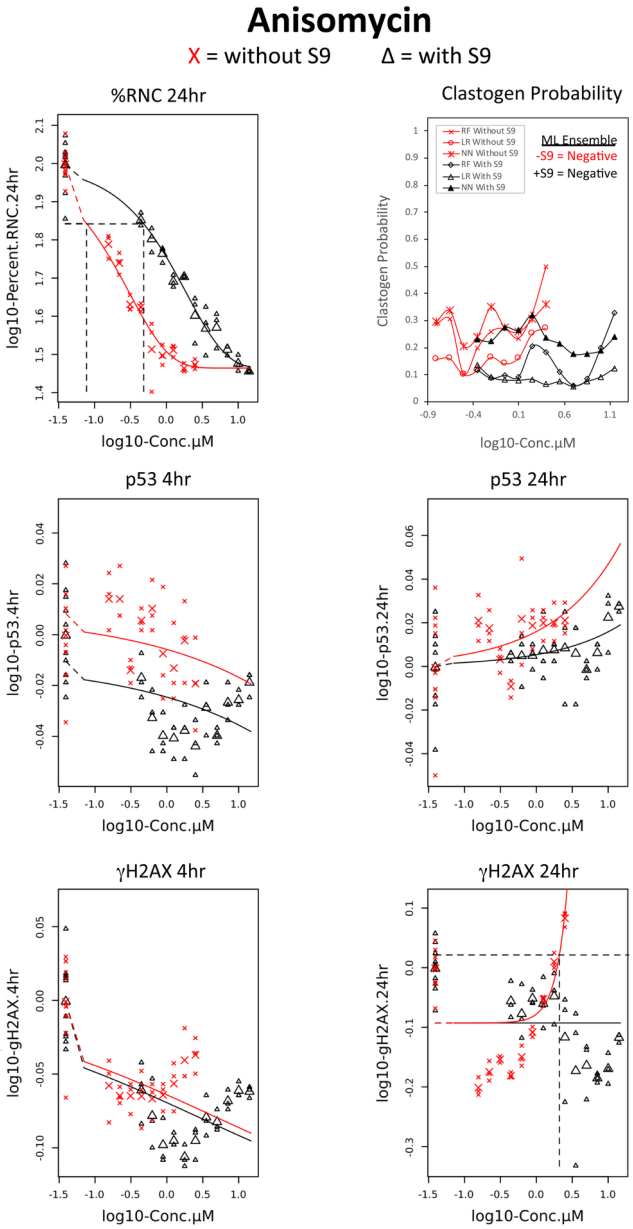
EM_22361_Figure 3_background fluorescence_v191209.tif



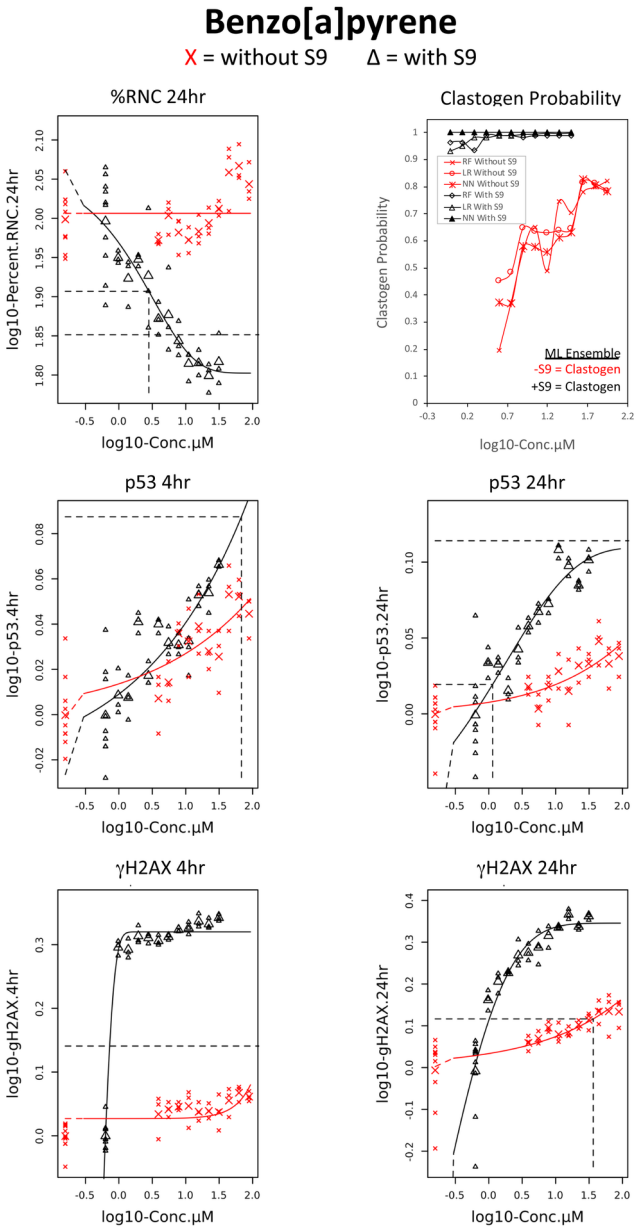
EM_22361_Figure 4_2AAF_v200108.tif



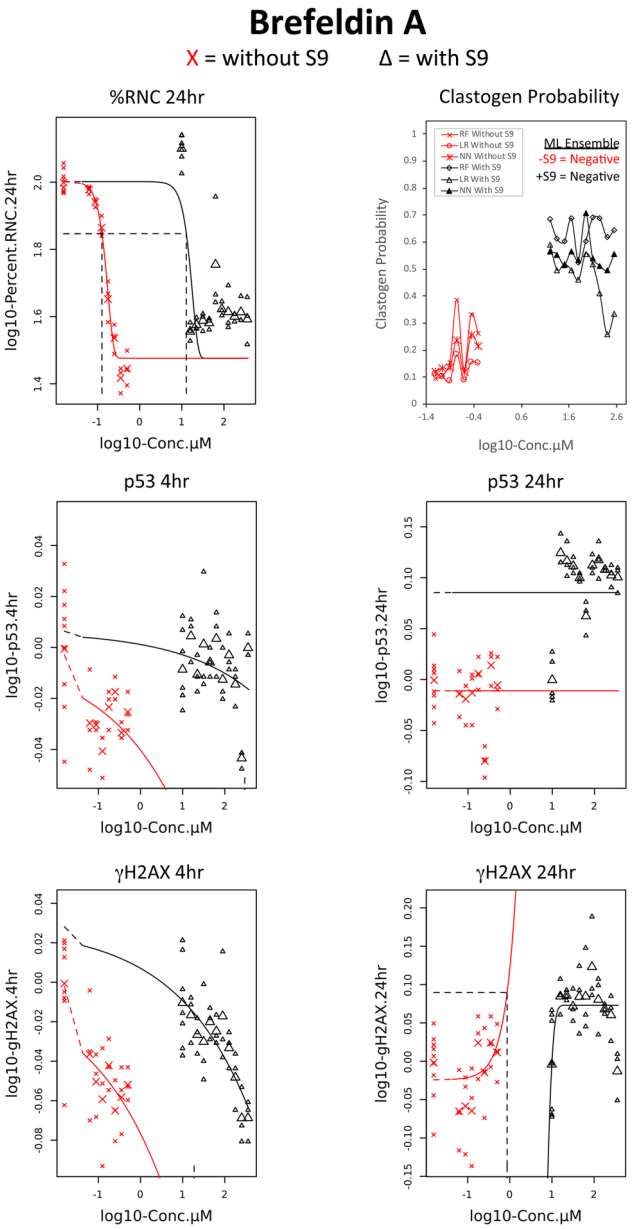
EM_22361_Figure 5_2AAN_v191211.tif



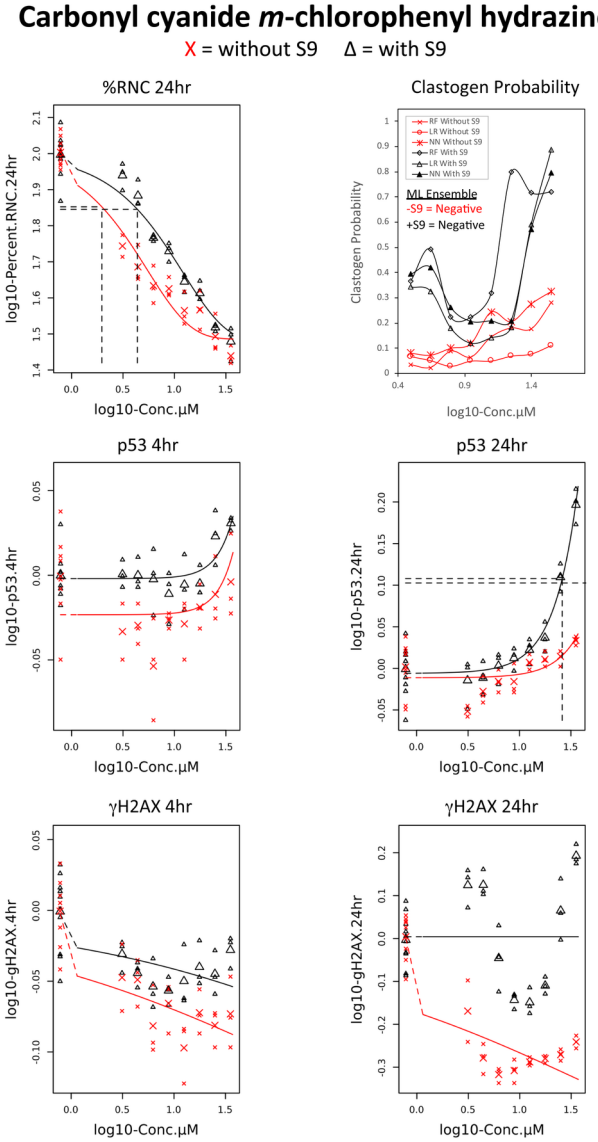
EM_22361_Figure 6_Anisomycin_v191211.tif



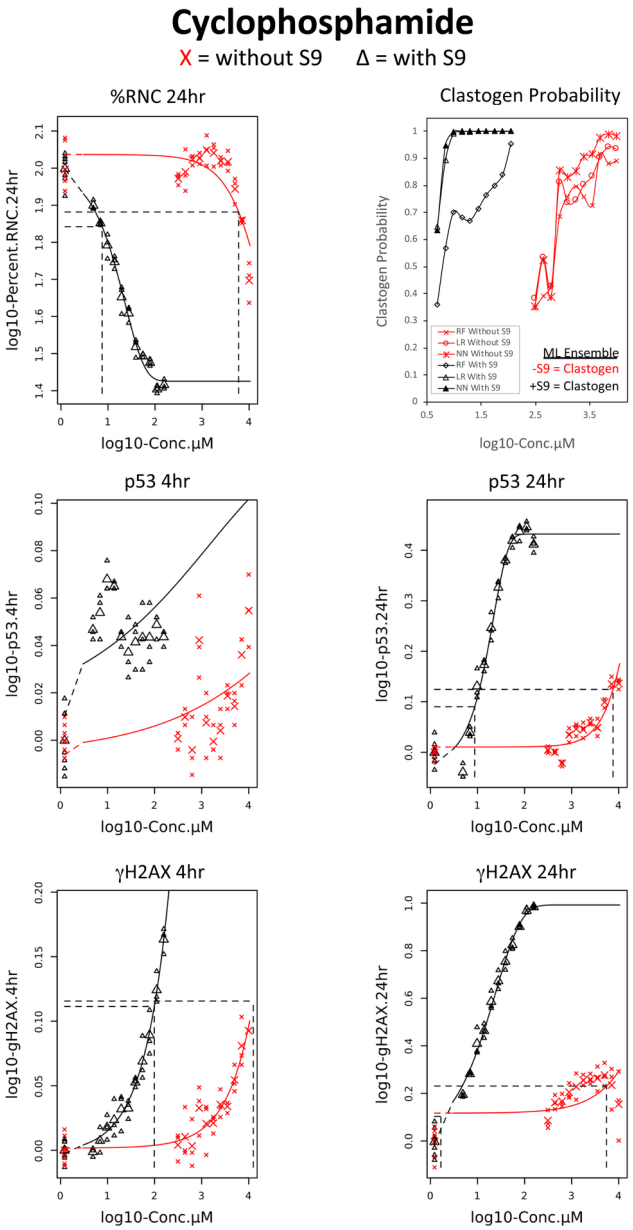
EM_22361_Figure 7_BaP_v191209.tif



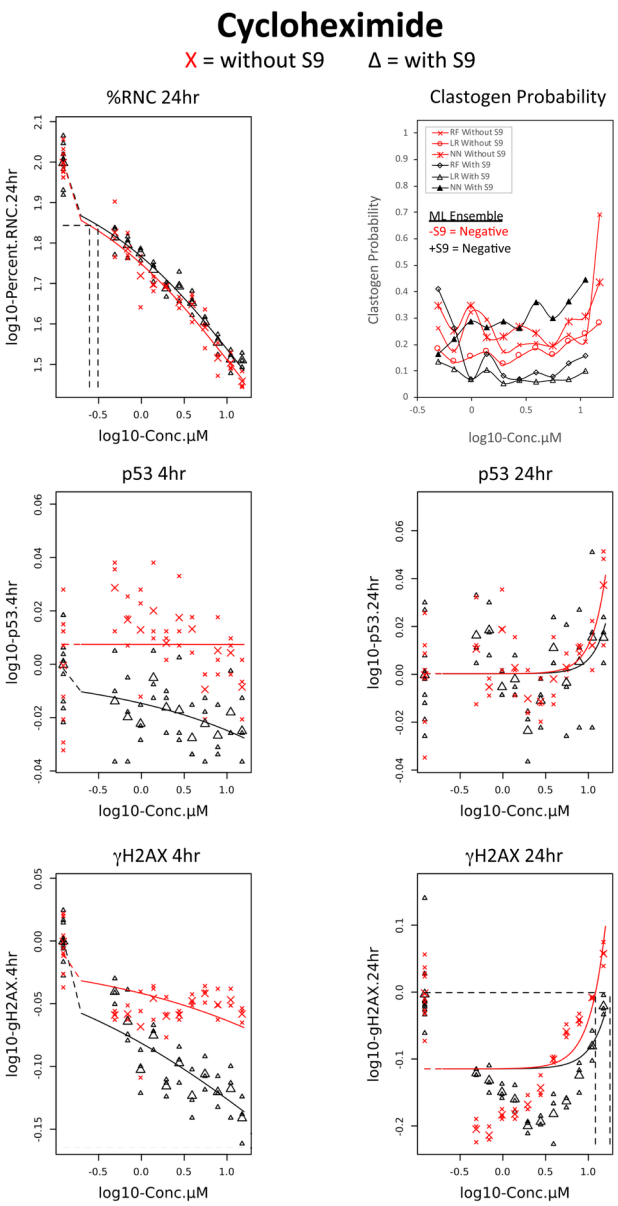
EM_22361_Figure 8_Brefeldin A_v191210.tif



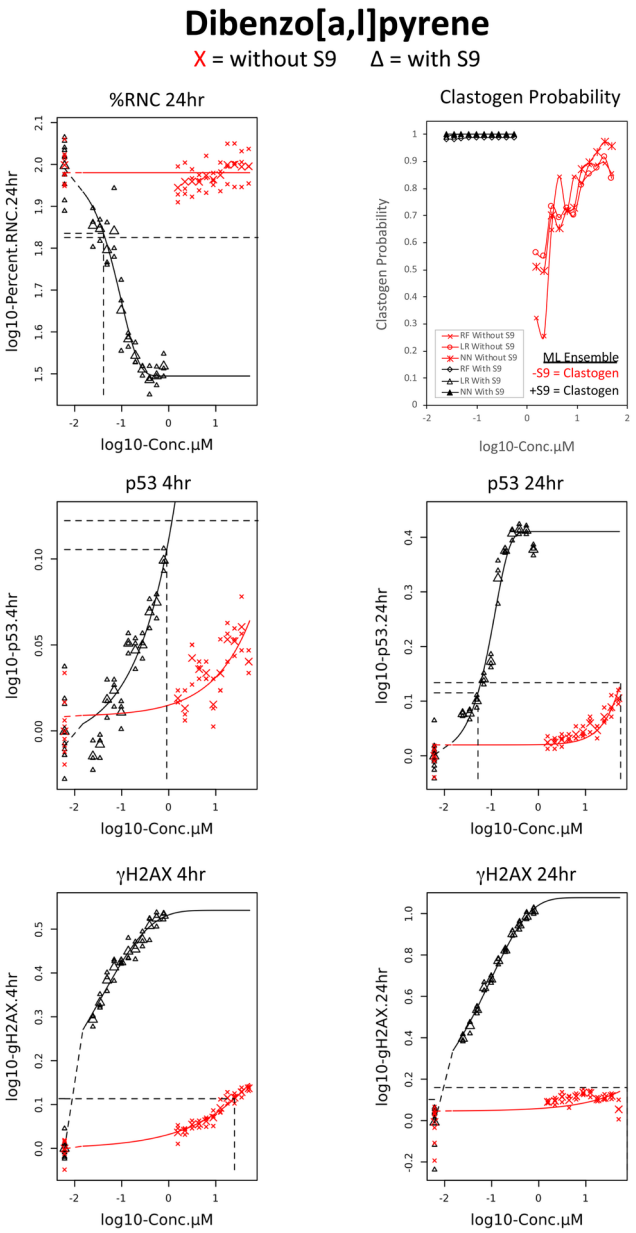
EM_22361_Figure 9_CCCP_v191209.tif



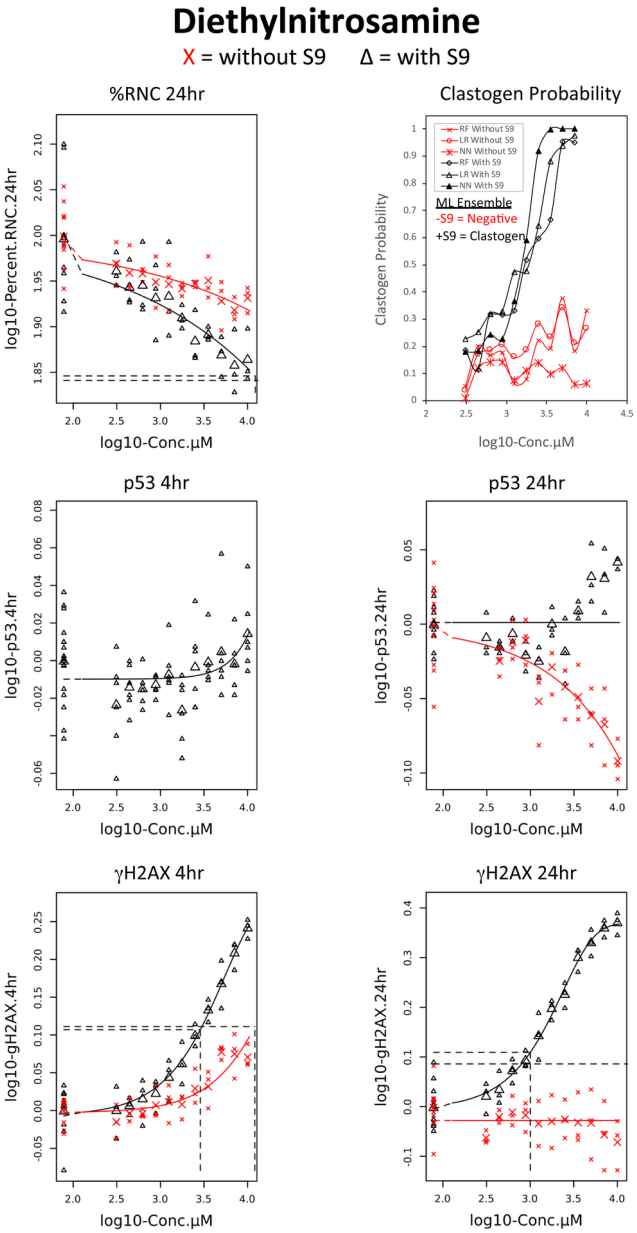
EM_22361_Figure 10_CP_v191209.tif



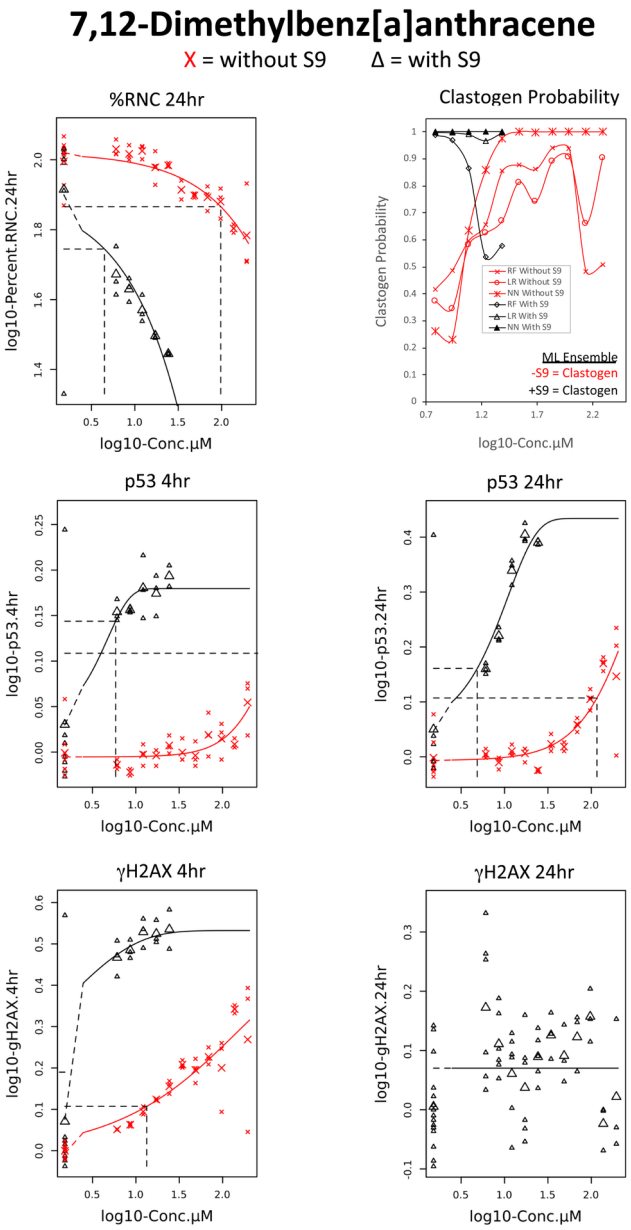
EM_22361_Figure 11_Cyclohex_v191209.tif



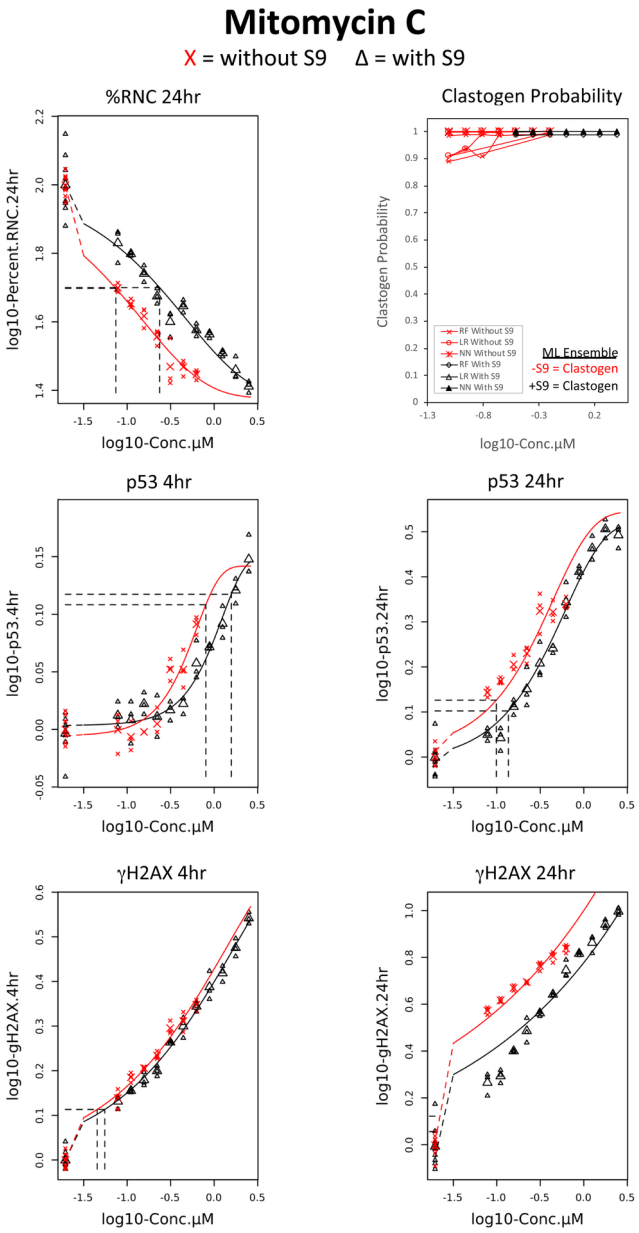
EM_22361_Figure 12_DB_a_l_P_v191209.tif



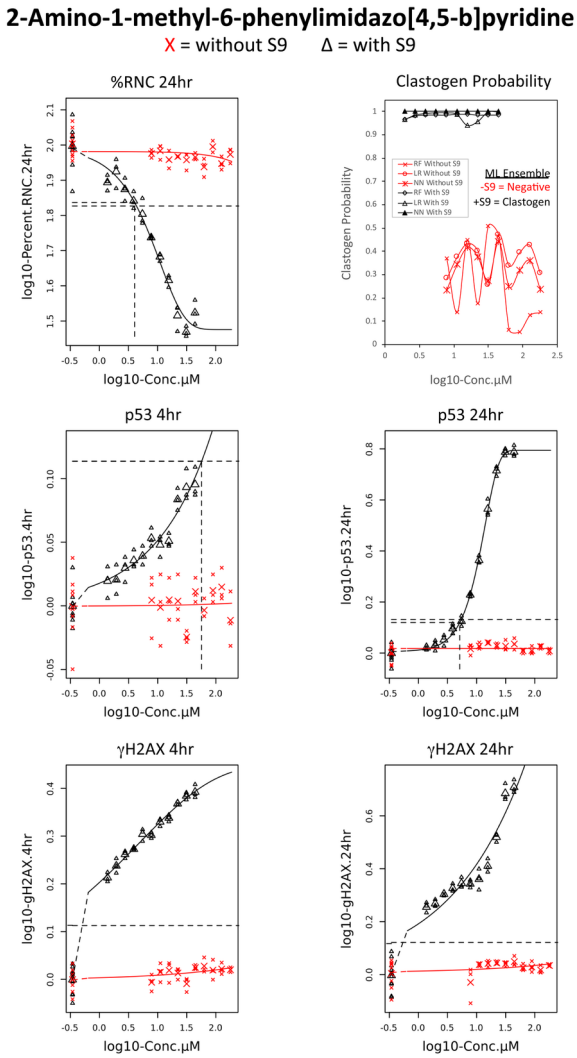
EM_22361_Figure 13_DEN_v191209.tif



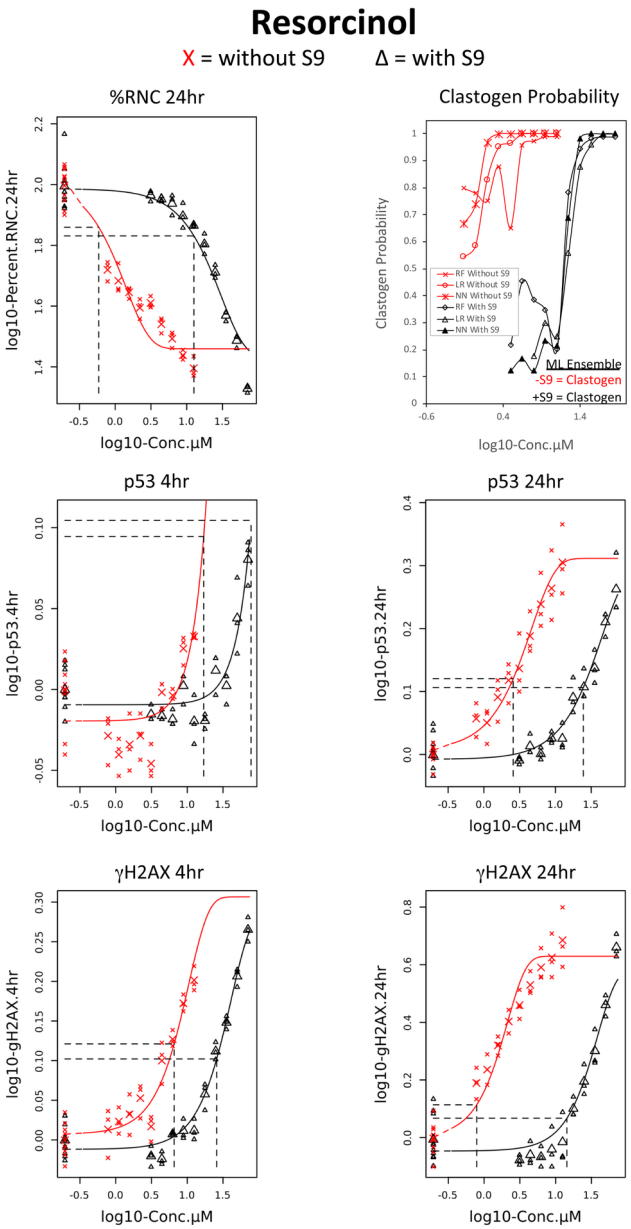
EM_22361_Figure 14_DMBA_v191209.tif



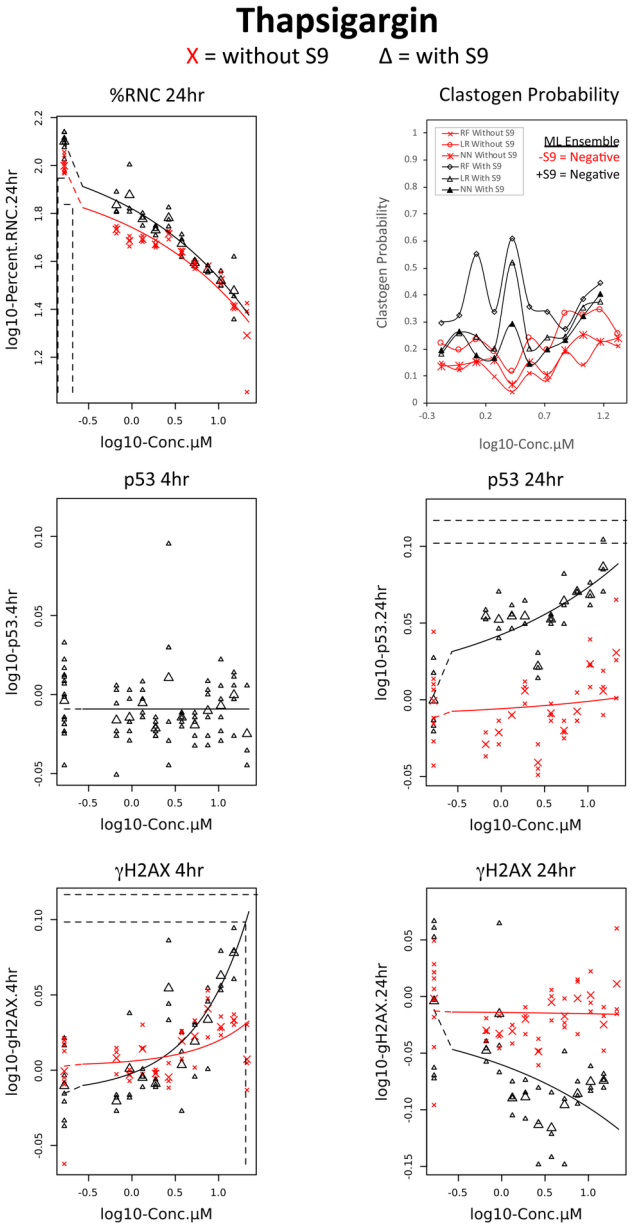
EM_22361_Figure 15_MMC_v191209.tif



EM_22361_Figure 16_PhIP_v191209.tif



EM_22361_Figure 17_Resorc_v191209.tif



EM_22361_Figure 18_Thapsigargin_v191210.tif

Table I. Test chemicals, source, biological effects.

Chemical (Abbreviation)	CAS No., Source	Notes about Biotransformation, Miscellaneous Info	References
2-Acetylaminofluorene (2AAF)	53-96-3, Sigma-Aldrich	Clastogen, requires metabolic activation (CYP1A2), forms C8 adduct on guanine	Otteneder and Lutz, 1999; Kirkland *et al.,* 2016
2-Aminoanthracene (2AAN)	613-13-8, Sigma-Aldrich	Clastogen, aromatic amine, requires metabolic activation (CYP1B1, 2A family)	Carriére *et al.,* 1992
Anisomycin	22862-76-6, Sigma-Aldrich	Cytotoxicant, protein synthesis inhibitor; *in vitro* MN neg. with high levels of apoptosis	Personal communication, Maik Schuler, Richard Spellman, Maria Engel
Benzo[a]pyrene (B[a]P)	50-32-8, Sigma-Aldrich	Clastogen, polycyclic aromatic hyrocarbon, requires metabolic activation (CYP1A1, 1B1, epoxide hydrolase), forms bulky adducts	Kirkland *et al.,* 2016
Brefeldin A	20350-15-6, Sigma-Aldrich	Cytotoxicant, ER-golgi transporter inhibitor, ER stress-induced apoptosis	Moon *et al.,* 2012
Carbonyl cyanide m-chlorophenyl hydrazone (CCCP)	555-60-2, Sigma-Aldrich	Cytotoxicant, uncouples oxidative phosphorylation	de Graaf *et al.,* 2004
Cyclophosphamide monohydrate (CP)	6055-19-2, Sigma-Aldrich	Clastogen, nitrogen mustard, requires metabolic activation (CYP2B6, CYP2C19, CYP2C9 and CYP3A4/5)	Kirkland *et al.,* 2016; Rodriguez-Antona and Inglelman-Sundberg, 2006
Cycloheximide	66-81-9, Sigma-Aldrich	Cytotoxicant, protein synthesis inhibitor	Youngblom *et al.,* 1989
Dibenzo[a,l]pyrene (DB[a,l]P)	191-30-0, Sigma-Aldrich	Clastogen, polycyclic aromatic hydrocarbon, requires metabolic activation (thought to be primarily activated by CYP1A1)	Arif and Gupta, 1997
Diethylnitrosamine (DEN)	55-18-5, Sigma-Aldrich	Clastogen, requires metabolic activation to form alkylating agent (likely involves CYP2E1 which is not highly expressed in rat liver S9); often only positive	Yamazaki *et al.,* 1992;

		at high concentrations (≥10 mM)	
7,12-Dimethylbenzanthracene (DMBA)	57-97-6, Sigma-Aldrich	Clastogen, requires metabolic activation (CYP1B1), forms bulky adducts	Kirkland *et al.,* 2016
Mitomycin C (MMC)	50-07-7, Sigma-Aldrich	DNA-DNA crosslinks, also alkylating activity and oxidative damage	Kirkland et al., 2016
2-amino-1-methyl-6-phenylimidazo [4,5-b]pyridine (PhIP)	105650-23-5, Toronto Research Chemicals	Clastogen, heterocyclic amine, requires metabolic activation (CYP1A family)	Kirkland *et al.,* 2016; Krais *et al.,* 2016
Resorcinol	108-46-3, Sigma-Aldrich	*In vitro* mammalian cell pos. (MLA assay with and without metabolic activation pos., *in vitro* human lymphocyte MN pos. in absence of metabolic activation); *in vitro* findings not confirmed *in vivo* (mouse MN neg.)	EFSA Journal, 2010
Thapsigargin	67526-95-8, Sigma-Aldrich	Cytotoxicant, ER stress-induced apoptosis	Futami *et al.,* 2005

Table II. Benchmark Dose Estimates.

	BMD (μM)				
Treatment	**4hr p53**	**24hr p53**	**4hr γH2AX**	**24hr γH2AX**	**24hr RNC**
2AAF without S9	>311	>311	>311	>311	98
2AAF with S9	>311	>311	119	35	184
S9 Potentiation Ratio	NC	NC	>2.6	>8.9	0.53
2AAN without S9	>55	>55	40	10.9	>55
2AAN with S9	7.9	2.5	0.24	0.67	2.1
S9 Potentiation Ratio	7	>22	167	16.3	>26.2
Anisomycin without S9	>2.5	>2.5	>2.5	2.1	0.08
Anisomycin with S9	>14.1	>14.1	>14.1	>14.1	0.48
S9 Potentiation Ratio	NC	NC	NC	<0.15	0.17
B[a]P without S9	>88	>88	>88	36	>88
B[a]P with S9	>31	1.1	0.19	2x10^-5	2.8
S9 Potentiation Ratio	NC	>80	>463	1.8x10^6	>31.4
Brefeldin A without S9	>0.5	>0.5	>0.5	>0.5	0.13
Brefeldin A with S9	>356	>356	>356	>356	12.8
S9 Potentiation Ratio	NC	NC	NC	NC	0.01
CCCP without S9	>35	>35	>35	>35	2.0
CCCP with S9	>35	26	>35	>35	4.4
S9 Potentiation Ratio	NC	>1.3	NC	NC	0.45
CP without S9	>10,000	7,570	>10,000	5,523	5,981
CP with S9	>156	8.7	99	1.7	7.5
S9 Potentiation Ratio	NC	870	>101	3,249	797
Cycloheximide without S9	>15	>15	>15	12	0.25
Cycloheximide with S9	>15	>15	>15	>15	0.31
S9 Potentiation Ratio	NC	NC	NC	<0.8	0.8
DB(a,l)P without S9	>50	>50	25	>50	>50
DB(a,l)P with S9	0.9	0.052	0.0011	0.0013	0.041
S9 Potentiation Ratio	>56	>962	22,727	>38,462	>1,220
DEN without S9	>10,000	>10,000	>10,000	>10,000	>10,000
DEN with S9	>10,000	>10,000	2,867	1,004	>10,000
S9 Potentiation Ratio	NC	NC	>3.5	>10	NC

DMBA without S9	>194	116	13.3	>194	98
DMBA with S9	5.9	4.9	0.15	>194	4.5
S9 Potentiation Ratio	>32.9	23.7	88.7'	NC	21.8
MMC without S9	0.802	0.0988	0.05031	0.00862	0.07407
MMC with S9	1.576	0.136	0.06122	0.03391	0.2363
S9 Potentiation Ratio	0.51	0.73	0.82	0.25	0.31
PhIP without S9	>178	>178	>178	>178	>178
PhIP with S9	56.81	5.142	0.1093	0.2281	9.764
S9 Potentiation Ratio	>3.1	>34.6	>1,629	>780	>18.2
Resorcinol without S9	>12.5	2.6	6.6	0.79	0.59
Resorcinol with S9	>71	24.6	25.8	14.4	12.6
S9 Potentiation Ratio	NC	0.11	0.26	0.05	0.047
Thapsigargin without S9	>21	>21	>21	>21	0.2
Thapsigargin with S9	>15	>15	>15	>15	0.14
S9 Potentiation Ratio	NC	NC	NC	NC	1.4

Abbreviations: Chemical abbreviations = same as Table I; BMD = benchmark dose; RNC = relative nuclei count; NC = not calculated