In vitro screening for population variability in toxicity of pesticide-containing mixtures

Nour Abdo[1,2], Barbara A. Wetmore[3], Grace A. Chappell[1,4], Damian Shea[5], Fred A. Wright[5,6], Mathew T. Martin[7], Ivan Rusyn[1,4]

[1]Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill, NC, USA; [2]Department of Public Health, Jordan University of Science and Technology, Ibrid, Jordan; [3]The Hamner Institutes for Health Sciences, Research Triangle Park, NC, USA; [4]Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA; [5]Departments of Biological Sciences and [6]Statistics and the Bioinformatics Research Center, North Carolina State University, Raleigh, NC, USA; [7]National Center for Computational Toxicology, US Environmental Protection Agency, Research Triangle Park, NC, USA

Corresponding author: Ivan Rusyn, MD, PhD, 4458 TAMU, Texas A&M University, College Station, TX 77843, USA; phone: +1-979-459-9866; email: irusyn@tamu.edu

Acknowledgements

The authors wish to thank Philippe Marlot and Munir Pirmohamed (Institute for Translational Medicine, University of Liverpool, UK) for technical assistance and useful discussions on study design. This work was supported, in part, by grants from U.S. EPA (STAR RD83516601) and NIH (P42 ES 005948). The views expressed are those of the authors and do not necessarily reflect the statements, opinions, views, conclusions, or policies of the U.S. EPA, NIH, or the United States government. The authors declare they have no competing financial interests.

ABSTRACT

Population-based human *in vitro* models offer exceptional opportunities for evaluating the potential hazard and mode of action of chemicals, as well as variability in responses to toxic insults among individuals . This study was designed to test the hypothesis t hat comparative population genomics with efficient *in vitro* experimental design can be used for evaluation of the potential for hazard, mode of action, and the extent of population variability in responses to chemical mixtures. We selected 146 lymphoblast cell lines from 4 ancestrally and geographically diverse human populations based on the availability of genome sequence and basal RNA -seq data. Cells were exposed to two pesticide mixtures -- an environmental surface water sample comprised primarily of org anochlorine pesticides and a laboratory-prepared mixture of 36 currently used pesticides – in concentration response and evaluated for cytotoxicity. On average, the two mixtures exhibited a similar range of *in vitro* cytotoxicity and showed considerable inter-individual variability across screened cell lines. However, when *in vitro*-to-*in vivo* extrapolation (IVIVE) coupled with reverse dosimetry was employed to convert the in vitro cytotoxic concentrations to oral equivalent dose s and compared to the upper bound of predicted human exposure, we found that a nominally more toxic chlorinated pesticide mixture would pose less risk as compared to the current use pesticide mixture. Multivariate genome -wide association mapping revealed an association between the toxicity of current use pesticide mixture and a polymorphism in rs1947825 in *C17orf54*. A genetic pathway analysis showed a significant association between metabolism pathways and the cytotoxicity of the chlorinated pesticide mixture. We conclude that a combination of *in vitro* human population-based screening followed by dosimetric adjustment and comparative population genomics analyses enables quantitative evaluation of human health hazard from complex environmental mixtures. Additionally, such an approach yields testable hypotheses regarding potential toxicity mechanisms.

INTRODUCTION

Pesticides are chemicals that are used to kill, repel, or control certain forms of plant or animal life that are considered to be pests (Krieger, 2010). Adverse health effects of pesticides can range from mild skin and mucous membrane irritation to more severe outcomes such as neurotoxicity and cancer (Bassil et al., 2007; Rother, 2014; Sanborn et al., 2007). Moreover, potential for adverse effects following exposure may be higher among relatively vulnerable populations, including women, children, the elderly, the immune-compromised and the malnourished (Jurewicz and Hanke, 2008; Perry et al., 2014). There are several challenges in the evaluation of the human health hazard of pesticides. First, pesticides have variable modes of action (MOA) dependent on use and activity, and are meant to be harmful and toxic to pests, but not humans. Second, because they are widely used in agricultural and household settings, people are frequently exposed to pesticide residues. Third, humans are typically exposed to mixtures of pesticides, creating challenges in hazard evaluation (Feron et al., 1998; Manikkam et al., 2012).

While safety testing of the individual pesticides is conducted according to established regulatory guidelines (Babut et al., 2013), evaluation of the toxicity of mixtures is less structured (U.S. EPA, 2002). The cumulative risk assessment is conducted for mixtures of chemicals with common mechanisms of toxicity, even though data are usually available only for individual chemicals. Indeed, current toxicity testing paradigms have been questioned for their failure to consider commonly occurring co-exposures and the magnitude of human population variability in response to chemicals (National Research Council, 2009).

Whole animal testing is difficult to employ for evaluating the hazards of chemical mixtures. In contrast, *in vitro* testing allows greater flexibility, as chemicals can be grouped according to their effects on key biologic pathways or tested over a broad range of concentrations to capture varied exposure scenarios in a rapid and inexpensive manner

(Andersen and Krewski, 2009) . The resulting data could enable an informed and focused approach to the problem of ass essing risk in human populations that are exposed to mixtures. Furthermore, with an experimental *in vitro* design that represents a human population, we are able to explore not only the hazard, but also its intrinsic variability across different concentration ranges (Lock et al., 2012; O'Shea et al., 2011) . Such information would be valuable to inform regulatory decisions that could mor e fully protect public health and sensitive subpopulations (Abdo et al., 2015).

In this study, we addressed the hypothesis that comparative population genomics with efficient *in vitro* experimental design can be used for evaluation of the potential for hazard, mode of action, and the extent of population variability in responses to chemical mixtures. We screened 146 lymphoblast cell lines (LCLs) from four ancestrally and geographically diverse populations with public ly available genotypes and sequencing data from the 1000 Genomes Project (1000 Genomes Project Consortium, 2010). Cells were exposed to two pesticide mixtures (an environmental sample, comprised primarily of a mixture of organochlorines extracted from a passive surface water sampling device, and a mixture of 36 currently used pesticides) at 8 concentrations. Cell viability was evaluated in a 96 -well plate format. Cytotoxic response was assessed using an effective concentration threshold of 10% (EC_{10}), designed to be relevant to the dose-response evaluation commonly used in quantitative risk assessment practice and to meaningfully capture ranges of variation in response across individuals. Genom e-wide association mapping and genetic association pathway analyses were performed to evaluate the genetic determinants of susceptibility. Furthermore, *in vitro* -to-*in-vivo* extrapolation by reverse pharmacokinetics was utilized to translate the *in vitro* concentrations to oral equivalents, which were then compared to predicted human cumulative exposures.

MATERIALS AND METHODS

Experimental Design

Cell lines. A set of 146 immortalized LCLs was acquired from Coriell Cell Repositories (Camden, NJ). The 146 c ell lines represent 4 ancestrally and geographically diverse populations (Table 1) : Utah residents with Northern & European ancestry (CEU); Tuscan in Italy (TSI); Yoruban in Ibadan, Nigeria (YRI); and British from England & Scotland (GBR). Cell lines were chosen based on the availability of dense genotyping information (1000 Genomes Project Consortium et al., 2012) . Screening was conducted in two batches, and cell lines were randomly divided into batches without regard to family structure , but with equal representation of population and gend er. Cells were cultured in RPMI 1640 media (Gibco, Carlsbad, CA) supplemented with 15% fetal bovine serum (HyClone, South Logan, UT) and 1% penicillin - streptomycin (Gibco) and cultured at 37°C with 5% CO_2. Media was changed every 3 days. Cell count and viability were assessed once a day for five days for all cell lines using Cellometer Auto T4 Plus (Nexcelom Bioscience, Lawrence, MA). Cells were grown to a concentration of up to 10^6 cells/ml, volume of at least 100 ml, and viability of >85% before exposures. After centrifugation, the cells were re -suspended in fresh media. Cells (100 µl containing 10^4 cells) were aliquoted to each well in a 96 -well treatment plate (following the addition of the chemicals) and mixed using the Biomek 3000 robo t. Plates were incubated for 24 h after treatment at 37°C and 0.5% CO_2.To increase the robustness of the data and to evaluate reproducibility , each cell line was seeded in at least two plates so that each compound would be screened in each cell line on 2 or more plates.

Chemical Mixtures. Cells were exposed to two environmental chemical mixtures. F irst mixture, referred to as " chlorinated pesticide mixture" throughout the manuscript , is an environmental sample obtained from a universal passive sampling device deployed for 30 days

in surface water next to a chlorinated pesticide storage facility . In this extract, 10 pesticides were present in detectable quantities in the post -collection labo ratory analysis (see T able 2 for a complete list of pesticide chemicals identified by mass spectrometry) . The s econd mixture, referred to as "current use pesticide mixture", was a laboratory-generated mixture of 36 currently used pesticides with relative c oncentrations selected to mimic fractional composition of the pesticide exposures in E astern North Carolina (Table 3). Stock solutions of each mixture were further diluted with dimethyl sulfoxide (DMSO) 8 -fold in ½ -log step -wise manner. Final cumulative concentrations ranged from 0.032 to 370.4 μM for the current use pesticide mixture and from 0.022 to 65.7 μM for the chlorinated pesticide mixture in 0.5% (vol/vol) DMSO. The mixtures were aliquoted to 96 -well plate format using Biomek 3000 robot (Beckman Coulter, Inc., B rea CA). The negative control was DMSO at 0.5%; the positive control was tetra -octyl ammonium bromide at 46 μM.

Cytotoxicity profiling . The CellTiter -Glo Luminescent Cell Viability (Promega, Madison, WI) assay was used to assess intracellul ar ATP concentration, a marker for cytotoxicity, 40 h post treatment. Time points were selected based on previous experiments at the National Institutes of Health Chemical Genomics Center (Xia et al., 2008) . A ViewLux plate reader (PerkinElmer, Shelton, CT) was used to detect luminescent intensity.

Data Processing

Cytotoxicity EC_{10} estimation and outlier detection . Cytotoxicity data were normalized relative to positive/negative controls as described elsewhere (Abdo, et al., 2015). We derived an effective concentration 10^{th} percentile (EC_{10}) to provide a single cytotoxicity dose summary per chemical and cell line. The derivation of EC_{10} was based on the logistic model:

with , where is the observed normalized signal representing proportion of surviving cells (which we term the "cytotoxicity value"), is the log(concentration) for each chemical, and is the limiting mean cytotoxicity value for the zero concentration. was set to zero, to avoid difficulties in estimating the minimum cytotoxicity value for chemicals with low cytotoxicity. An exception was made for chemicals in which the cytotoxicity value at the highest concentration was higher than 0.4, as a very few number of plates/chemicals did not reliably reach maximum cytotoxicity. In those instances was set at the observed cytotoxicity at the maximum concentration. Inspection of these data revealed good fits in such instances. Although in principle should have been 1.0, a number of plates exhibited a drift from this value, and thus the parameter was estimated from the data.

Fitting for the parameters proceeded by maximum likelihood using numerical optimization in *R* v2.15. An automatic outlier detection algorithm was devised by considering the impact of dropping each concentration value in succession, and removing those values for which the maximum likelihood improved by a factor of 10 or more and refitting the model using the non-outlying observations.

Normalizing for batch effects. Batch effects were evaluated by running principal component analysis. EC_{10} values were adjusted for batch effect using the ComBat method (Johnson et al., 2007).

Concentration response for populations and individuals. For each pesticide mixture, the three-parameter logistic regression described above in EC_{10} estimation was fit to concentration-response data for each cell line. The variation in the EC_{10} estimates was used as illustrative of population variation in true EC_{10} values, although additional sampling variation underlies each

EC_{10} estimate. An overall logistic concentration -response curve was fit to the aggrega ted data across all individuals (Figure 1).

Reproducibility and correlation between mixtures . Pearson and Spearman correlation coefficients (r) between pairs of replicate plates were used to assess experimental reproducibility and the correlation between the two mixtures. For this analysis, the two replicate plates were selected for each mixture and cell line pair.

Chemical/Mixture Specific Adjustment Factor (CSAF). Variability in response for each mixture across the 146 cell lines was derived as the longest tail of the variability distribution (in our case the ratio of the 50^{th} percentile to the 5^{th} percentile was greater than the ratio of the 95^{th} to the 50^{th} percentile) using the World Health Organization guidance for chemical -specific adjustment factors (World Health Organization, 2005).

Chemical descriptors . Chemical descriptors were calculated using Dragon version 5.5 (Mauri et al., 2006) . Constant and near constant descriptors as well as highly correlated descriptors were excluded and descriptor values were normalized on a scale from 0 to 1.

Differences in cytotoxicity across different populations. *Analysis of Variance (ANOVA)* was performed to assess population differences in cytotoxicity between the four screened populations for each mixture.

Genotypes. The primary source of genotypes was obtain ed as described in Abdo et al. (2015). SNPs with a call rate below 99%, minor allele frequency (MAF)<0.05, or Hardy–Weinberg equilibrium p-value $<1x10^{-3}$ were excluded.

Multivariate Association Analysis (MAGWAS) . The MAGWAS analysis of covariance model (Brown et al., 2012) was used for association mapping. The approach allows for use of the full concentration -response profile, as opposed to a univariate summary (such as EC_{10}) as a single response, with the advantage of robustness and power unde r a wide v ariety of association

patterns. The model used for association for the jth individual and genotype i for the chemical/SNP was:

,

where is the vector of responses (across the eight concentrations) for the j^{th} individual having genotype i, is the design matrix of covariates, including sex, indicator variables for laboratory batch, and the first ten genotype principal components, and is the eight -vector of parameters modeling the effects of genotype i on the response. The model assumes that the error terms are multivariate normally distributed, with mean vector $\mathbf{0}$ and variance-covariance matrix $\mathbf{\Sigma}$, allowing for dependencies in the observations. P-values were obtained using Pillai's trace (Pillai, 1955). Because this method makes use of asymptotic theory, markers with fewer than 20 individuals representing any genotype were removed, leaving 692,013 SNPs for analysis.

Estimation of C_{ss} using in vitro in vivo extrapolation (IVIVE) and Monte Carlo Simulation. Key determinants of steady -state pharmacokinetics were experimentally measured for chemicals and published previously (Wetmore et al., 2012, Wetmore et al., submitted) . Briefly, plasma protein binding was measured using rapid equilibrium dialysis (Wetmore et al., 2012) and t he rate of hepatic metabolism of the parent compound was determined using the substrate depletion approach (Rotroff et al., 2010; Wetmore, et al., 2012) . See flow chart for these analyses in Supplemental Figure 1)

These data were then used to calculate chemical steady -state blood concentrations (C_{ss}) as previously described , with modification (Wetmore et al. 2012 , Wetmore et al., submitted). The base equation used to calculate static C_{ss} is based on constant uptake of a daily oral dose and factors in blood binding, hepatic clearance and non -metabolic renal clearance . The daily oral

dose was set to 1 μg/kg/day to reflect ambient environmental exposures. A correlated Monte Carlo approach was employed (Jamei et al., 2009) using Simcyp (Simcyp v.1.3; Certara, Sheffield, UK) to simulate variability across a population of 10,000 individuals equally comprised of males and females, 20-50 years of age. A coefficient of variation of 30% was used for intrinsic and renal clearance. The median, upper and lower fifth percentiles for the C_{ss} were obtained as output.

Calculation of oral equivalent dose values. In conventional use, pharmacokinetic models are used to relate exposure concentrations to a blood or tissue concentration. This is typically referred to as "forward dosimetry." In contrast, the models can also be reversed to relate blood or tissue concentrations to an exposure concentration, which is referred to as "reverse dosimetry" (Tan et al., 2007). Based on the principal of reverse dosimetry, the median, upper and lower 5th percentiles for the C_{ss} were used as conversion factors to generate oral equivalent doses according to the following formula:

In the equation above, the oral equivalent dose value is linearly related to the *in vitro* EC_{10} and inversely related to C_{ss}. This equation is valid only for first-order metabolism that is expected at ambient exposure levels. An oral equivalent value was generated for each chemical-cell line combination and summed to provide a cumulative oral equivalent value for each cell line.

Predicted exposure limits. Pesticide specific predicted exposures were obtained as previously detailed in (Wambaugh et al., 2013). The pesticide specific exposure limit was available for 35 out of the 36 pesticides in the current-use pesticide mixture and for 6 out of 10

pesticides in the chlorinated pesticide mixture. Missing values were replaced by the highest exposure within each mixture. Then, a cumulative exposure was computed for each mixture from the upper 95th percentile (see flow chart in Supplemental Figure 2).

RESULTS

Cytotoxicity of pesticide mixtures in vitro

Screening was conducted in a 96 -well plate format using a robotic system to facilitate reproducibility and throughput. The 146 cell lines were randomly assigned to two batches with blocking to achieve balancing by sex and population. Each cell line was plated on two plates to evaluate technical reproducibility and pesticide mixtures were added at 8 different concentrations ranging from 0.032 to 370.4 μM for current use pesticide mixt ure, and from 0.022 to 65.7 μM for chlorinated pesticide mixture. Positive and negative controls for cytotoxicity, as assessed by intracellular ATP concentrations, were included on each plate. Normalization to the control for each plate was performed as de scribed in the Materials and Methods section separately for each cell line. EC_{10}s were derived, batch -corrected and averaged across replicate plates for each cell line.

To visualize "individual" vs. "population" response to each pesticide mixture, we fit ted a 3-parametric logistic regression to each cell line's concentration -response, as well a single concentration-response curve for the entire population, as illustrated in Figure 1 . Population variability in cytotoxicity of each mixture is shown as a histogram of EC_{10} values. Both mixtures demonstrated considerable inter -individual variability in cytotoxicity . To evaluate the reproducibility of the EC_{10} values, pair -wise correlations among duplicate plate pairs were calculated for each mixture . Highly significant correlations were observed for both mixtures ($p<0.0001$). For current pesticides mixtures r[Pearson's]=0.62 and ρ[Spearman]=0.55. For

chlorinated pesticides mixture, *r*[Pearson's]=0.65 and ρ[Spearman]=0.56. Overall reproducibility for both mixtures was also significant ($p<0.0001$) with *r*[Pearson's]=0.62 and ρ[Spearman]=0.54.

We found that both mean and median EC_{10} values for *in vitro* cytotoxicity, as well as the range among cell lines tested, were not significantly different between the two mixtures (Figure 2). Using these data, the extent of population variation in *in vitro* cytotoxicity may be derived to serve as a surrogate for cellular variation in the toxicodynamic relationship between systemically available concentrations and toxic responses (Zeise et al., 2013). We calculated a toxicodynamic variability factor for these human cell lines as $10^{(q_{95}-q_{50})}$, analogous to a chemical-specific toxicodynamic uncertainty factor (UFd) for inter-individual variability (World Health Organization, 2005), and found it to be around 3-fold for either mixture (Table 4).

Next, we evaluated the extent of the similarity of cytotoxic responses to the mixtures across cell lines. Strong (significant even after removal of the three outlier cell lines) correlation (r[Pearson]=0.53, $p<0.0001$; ρ[Spearman]=0.25, $p<0.01$) was observed between the mixtures, illustrating appreciable degree of concordance in individual cell line responses (Figure 3a). There were no suggestive patterns of population clustering in the correlation between the mixtures and neither mixture exhibited significant differences among the populations tested (Supplemental Figure 3). It is of note, however, that GBR cell lines were the most sensitive, while YRI cell lines the least sensitive to *in vitro* cytotoxicity of these mixtures. Moreover, within-population variability was greater for the current use pesticide mixture as compared to the chlorinated pesticide mixture, especially when considering the range of the upper quartile to the lower quartiles.

The finding of the significant concordance in responses to both mixtures is of interest because there is no individual chemical overlap (Tables 2 and 3). These results may suggest

potential shared mechanisms for cytotoxicity. To further explore chemical similarity among compounds in each mixture, we performed principal components analysis using chemical descriptors. We found that two mixtures overlap in their chemical descr iptor space (Figure 3b), which may partially explain the correlation between two mixtures. While some of the individual components in both mixtures are closely related isomers, no clustering of compounds based on the known pesticidal mode of action (http://www.irac-online.org/documents/moa-classification/?ext=pdf) was observed.

We also compared the strength of the correlation between two mixtures to that of a pair-wise comparison between any pair of compounds in another study that evaluated cytotoxicity of 179 diverse environmental compounds and drugs in a population of lymphoblast cell lines (Abdo, et al., 2015). The correlation between two mixtures tested in this study (ρ[Spearman]=0.25) was comparable to the median correlation of a randomly chosen pair from 15931 possible combinations in the previous cytotoxicity experiment (Figure 3c).

In vitro *-to-* in vivo *extrapolation of cytotoxicity of pesticide mixtures in a population-based model to oral human equivalents and predicted human exposure levels*

To conduct a comparative analysis of *in vitro* cytotoxicity measures of pesticide mixtures with potential human exposures, we computed oral equivalent doses for both mixtures using the reverse dosimetry approach (Wetmore, 2015). *In vitro* pharmacokinetic data (Wetmore et al. 2012, Wetmore et al., submitted) were available for 31 of the 36 chemicals present in the current use pesticide mixture, and for 4 of the 10 chemicals in the chlorinated pesticide mixture. In comparison to 180 ToxCast Phase II chemicals similarly assessed for *in vitro* pharmacokinetics, the C_{ss} values for the 31 current use pesticides had a similar distribution (a median C_{ss} <1 μM and 95th percentile ≈200 μM). Only two of the chemicals in the current use pesticide mixture had very high C_{ss} values, ethalfluralin (350 μM) and flumetralin (277 μM), the rest were below 8

μM. The distribution of the values for 4 compounds in the chlorinated pesticide mixture was different, the maximum C_{ss} value was 58.5 μM.

Because there is no standard approach for evaluation of pharmacokinetics of mixtures, for the purposes of pharmacokinetic modeling in this study we assumed that pharmacokinetics of each chemical will not be significantly affected by the presence of other chemicals in the mixture. Given that cytotoxicity was measured across 146 individual cell lines, separate oral equivalents were calculated for each individual based on the percentage of a given chemical in the mixture (Tables 2 and 3). Furthermore, because some chemicals in each mixture were without *in vitro* pharmacokinetic parameters, oral equivalent doses were computed based on four different scenarios (see Supplemental Figure 1 for the workflow). We substituted missing Css values with either median or largest (based on the most conservative simulation assuming no hepatic clearance, high blood binding and only renal clearance, referred to as a "worst-case-scenario") C_{ss} value of other chemicals in the mixture. In addition, oral equivalents were calculated with and without weighting of the EC_{10} by the percentage of chemical in the mixture.

The C_{ss} values were derived using the Simcyp software with Monte Carlo simulations to account for the population variability in pharmacokinetics in healthy individuals (Northern European, 20-50 years of age, equally mixed sex). To be reasonably conservative, the upper 95th percentile values from a series 10 simulation (1000 individuals each for estimating pharmacokinetics variability) per trial were used to determine the oral equivalents. This analysis showed that even though *in vitro* cytotoxicity of the mixtures, *i.e.*, EC_{10} values, was quantitatively (with respect to the mean, median or distribution) indistinguishable (Figure 2), highly significant differences ($p<0.01$ or greater) arise when oral equivalents are computed from *in vitro* EC_{10} values (Figure 4). The chlorinated pesticide mixture was predicted as about an order of magnitude more toxic ($p<0.001$) than the current use pesticide mixture. Oral equivalent

doses for both mixtures were not significantly different (<0.5 -fold difference) when the median C_{ss} value was used instead of the "worst case scenario" to replace missing values. However, oral equivalent dose was remarkably shifted for both mixtures (>1.2 -fold change) when EC_{10} values were weighed according to the relative proportion of the chemical in the mixture. Still, the relationship based on the oral equivalent dose was the same in all scenarios; the chlorinated pesticide mixture was more toxic than the current use pesticide mixture.

To further interpret the outcome of these experiments in the context of human health risk, we examined the relationship of the calculated oral dose equivalent with estimated human exposures to these mixtures. First, we computed a cumulative exposure value for each mixture based on the exposure estimates for each individual chemical obtained from ExpoCast (Wambaugh, et al., 2013), a framework that estimated human exposure potential for 1936 chemicals. Predicted estimates of exposure were available for 35 of the 36 chemicals present in the current use pesticides mixture, and 6 of the 10 chemicals in the chlorinated pesticide mixture. To remain conservative, missing values were substituted with the highest predicted exposure from Expocast data for a chemical in the respective mixture (see Supplemental Figure 2). Next, cumulative exposure for each mixture was computed as the upper 95th percentile and compared to oral equivalent doses for *in vitro* cytotoxicity (Figure 4). While human exposure estimates were lower than oral dose equivalent *in vitro* cytotoxic doses for both mixtures, a much greater margin of safety is evident for the chlorinated pesticide mixture than for the current use pesticide (5-fold or greater *vs* less than 2-fold, respectively). This indicates a wider margin of safety for the chlorinated pesticide mixture than the current-use pesticide mixture.

Relationships between cytotoxicity of pesticide mixtures and genotype

Because the cell lines used in this study are densely genotyped (1000 Genomes Project Consortium, et al., 2012), association analysis was performed between the quantitative estimates

of cytotoxicity and genetic variability among 146 individuals included in screening. Genotyping data was processed as detailed in Methods. Sex, experimental batch and date, population, and the first ten genotype principal components were included as covariates in multivariate ANCOVA genome-wide association analysis (Brown, et al., 2012), a sensitive method designed for evaluating a pattern of variation of cytotoxicity measurements due to genotype. Despite a relatively small population of 146 cell lines, a highly suggestive ($p<6.5e^{-08}$) association was observed between cytotoxicity of the current use pesticide mixture and a locus on Chr17 (Figure 5a) [near the Bonferroni threshold, and genome-wide significant by the criterion of (Dudbridge and Gusnanto, 2008)]. The most highly associated SNP (rs1947825) is located in an open reading frame C17orf54 (Figure 5b). When the cytotoxicity concentration-response patterns for cells with each of three genotypes for rs1947825 were examined (Figure 5c), we found that the major allele (AA) confers greater sensitivity, with the heterozygous genotype (AT) falling consistently in the middle across all concentrations.

Next, "pathway" association analysis of gene sets/ontologies (Schaid et al., 2012) was performed using EC_{10} phenotypes and the same compendium of SNPs using the *gene set scan* approach, which performs resampling to compute significance of SNPs, genes, and ontologies (*i.e.*, KEGG) in a hierarchical manner. For each mixture and ontology, we applied family-wise error rate (FWER) control using 10,000 resamples, and report in Table 5 all of the ontology findings with FWER<0.2 in order to be inclusive. Interestingly, several metabolism pathways were significantly associated with cytotoxicity response to the chlorinated pesticide mixture. The top contributing genes within each of those pathways were mainly from the uridine diphosphate glucuronosyltransferases (UGT) family. UGT genes are highly polymorphic in humans and play an important role in metabolism of various drugs and xenobiotics, including pesticides (Burchell, 2003).

DISCUSSION

Following the advice from the U.S. National Academies on developing a long-range strategic plan to update and advance the way environmental agents are tested for toxicity (Krewski et al., 2011), substantial advancements in high-throughput approaches to characterize the biological activity that may be indicative of potential human health hazard of environmental chemicals *in vitro* have been implemented (Collins et al., 2008). Nonetheless, difficulties are many in conducting human health risk assessments from *in vitro* endpoints (Crump et al., 2010; Judson et al., 2011). A major challenge in human health assessments is developing a comprehensive understanding of population variability in susceptibility to chemical toxicity (Zeise, et al., 2013). Regulatory risk assessment incorporates multiple uncertainty factors that are based on default assumptions and only recently experimental approaches have become available to provide scientific data to replace defaults in inter-individual variability in toxicokinetics (Wetmore et al., 2014) and toxicodynamics (Abdo, et al., 2015).

Furthermore, no clear framework has been set to evaluate potential toxicity of chemical mixtures in non-animal alternative models. Few environmental chemical mixtures have been evaluated, especially at environmentally relevant concentrations (Carvalho et al., 2014), with regulatory decisions primarily based on a single compound evaluation. However, potentiation and synergistic interactions of chemicals in mixtures is of great concern (Cedergreen, 2014). It has been shown that exposure to chemical mixtures, including pesticides, often occurs with each chemical in the mixture present at respective safety limit concentrations (Carvalho, et al., 2014). Moreover, evaluation of chemical mixtures with similar modes of action, without consideration of realistic exposure in the environment, might underestimate the toxicological risk associated with their exposure (Hadrup, 2014).

To address the challenges of assessment of potential hazard of complex mixtures while accounting for potential inter-individual variability, we aimed to provide quantitative measures for population-based *in vitro* toxicity of pesticide mixtures. We also used a reverse dosimetry approach to translate *in vitro* cytotoxicity estimates to oral equivalent doses and compared those to estimates of human exposure (Wetmore et al., 2012). Although investigation of population variability in toxicity of hundreds of individual chemicals is ongoing (Abdo, et al., 2015; Lock, et al., 2012), to our knowledge this study is first to examine inter-individual variability in response to mixtures.

This screening approach showed that both pesticide mixtures that were tested exhibited appreciable inter-individual variation in cytotoxicity. Interestingly, the toxicodynamic uncertainty factor for both pesticide mixtures (3.0 and 3.05) derived from the population variability in our present study was similar to the median inter-individual variability for the 179 individual chemicals previously tested. This finding is consistent with the default uncertainty factor for toxicodynamic difference among humans ($10^{0.5}$) that is used in risk assessments when no chemical-specific data are available (World Health Organization, 2005). On average, there was no significant difference between the *in vitro* cytotoxicity concentrations (*i.e.*, EC_{10}) of the current use pesticide mixture and the chlorinated pesticide mixture. However, incorporation of dosimetry with the *in vitro* data and conversion to an oral equivalent dose for each mixture revealed that a significantly lower dose of a chlorinated pesticide mixture would lead to an internal concentration equal to the cyto toxicity-eliciting EC_{10}. Conversion of the *in vitro* data in this manner allows a risk-relevant ranking of the mixtures that considers chemical pharmacokinetic behavior along with additional exposure data to adjust the potencies. Incorporation of human dosimetry and predicted human exposure is necessary for greater

confidence in the "presumed hazard" from *in vitro* high throughput screening alone (Gangwal et al., 2012).

It is not surprising that the cumulative human predicted exposure limit is much higher for the current-use pesticide mixture compared to the chlorinated pesticide mixture, which mostly consisted of pesticides withdrawn from the market. The current-use pesticide mixture included 36 currently used pesticides and mimicked real exposure levels in Eastern North Carolina, with atrazine pesticides being the most abundant. Atrazine is among the highest used (64-80 million pounds annually in the United States) agricultural pesticides (Barr et al., 2007). Therefore, the predicted exposure limit for the current-use pesticide mixture was expected to be high, and in our case it was very close to the calculated cytotoxic oral equivalent dose.

In addition to demonstrating how an *in vitro* human population-based model system may be used to evaluate potential hazard of complex mixtures, we also took advantage of the availability of genetic information on the cells to evaluate genotype-phenotype associations. Recognizing the genetic underpinning of cytotoxicity may offer valuable insights into the underlying casual physiological variation and biologically-associated pathways. The significant locus (*C17orf54*) identified in this study is in a presumably non-coding genomic region, consistent with 90% of the significant findings from human GWAS studies to date (Fraser, 2013). The long intergenic non-protein coding RNA 469 resides in the region. A critical role of non-coding RNAs in response to carcinogen and toxicant exposure is an emerging area of investigation in toxicology (Marrone et al., 2014), and a potential relationship between cytotoxicity of the current use pesticide mixture and the long non-coding RNA remains to be explored.

The pathway analyses of genetic variability revealed that polymorphisms in UGT metabolizing enzymes are significantly associated with inter-individual variability in cytotoxicity

for the chlorinated pesticide mixture. UGTs metabolize pesticides and other xenobiotics to less-toxic water-soluble glucuronides and facilitate their excretion in bile and urine (Burchell, 2003; Meech et al., 2012). This finding is also noteworthy because UGT enzymes are genetically polymorphic, with more than 200 human alleles identified to date (Stingl et al., 2014). Polymorphisms in UGT1 and UGT2 families can alter enzymatic role, cellular processes, or gene expression, thereby possibly affecting individual cell's cytotoxic response. The majority of xenobiotics are metabolized mainly by UGT 1A1, 1A3, 1A4, 1A9 and 2B7 (Stingl, et al., 2014), which were the top ten significant genes associated with the cytotoxicity of the chlorinated pesticide mixture. This finding suggests that variation in the genes coding for these enzymes may be particularly relevant in metabolizing chlorinated pesticides.

There are a number of limitations to extrapolating from *in vitro* toxicity profiling using lymphoblasts to humans, including severe limitations in metabolic capacity of these cells, acute nature of exposure, questions about target organ adverse effects, and no opportunity to consider other important variables such as age, lifestyle factors and diet. It is also yet to be established how chemicals may interact with one another in mixtures, both in terms of pharmacokinetics and in terms of toxicity; the assumptions made in our work with regards to reverse dosimetry and treatment of missing values may constrain the interpretation of the data presented in this work. There remains a pressing need to screen individual pesticides, in addition to their mixtures, in order to test these assumptions. These limitations notwithstanding, our work highlights the value of a population-based *in vitro* survey combined with assessment of oral equivalents and human exposures for pesticides and other chemicals. These experiments advance our understanding of the genetic underpinnings of susceptibility-related regulatory networks in response to toxicants.

References

1000 Genomes Project Consortium (2010). A map of human genome variation from population-scale sequencing. *Nature* **467**(7319), 1061-1073.

1000 Genomes Project Consortium, Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., Handsaker, R. E., Kang, H. M., Marth, G. T., and McVean, G. A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**(7422), 56-65.

Abdo, N., Xia, M., Brown, C. C., Kosyk, O., Huang, R., Sakamuru, S., Zhou, Y. H., Jack, J. R., Gallins, P., Xia, K., Li, Y., Chiu, W. A., Motsinger-Reif, A. A., Austin, C. P., Tice, R. R., Rusyn, I., and Wright, F. A. (2015). Population-based in vitro hazard and concentration-response assessment of chemicals: the 1000 genomes high-throughput screening study. *Environ Health Perspect* **123**(5), 458-466.

Andersen, M. E., and Krewski, D. (2009). Toxicity testing in the 21st century: bringing the vision to life. *Toxicol Sci* **107**(2), 324-330.

Babut, M., Arts, G. H., Barra Caracciolo, A., Carluer, N., Domange, N., Friberg, N., Gouy, V., Grung, M., Lagadic, L., Martin-Laurent, F., Mazzella, N., Pesce, S., Real, B., Reichenberger, S., Roex, E. W., Romijn, K., Rottele, M., Stenrod, M., Tournebize, J., Vernier, F., and Vindimian, E. (2013). Pesticide risk assessment and management in a globally changing world--report from a European interdisciplinary workshop. *Environ Sci Pollut Res Int* **20**(11), 8298-8312.

Barr, D. B., Panuwet, P., Nguyen, J. V., Udunka, S., and Needham, L. L. (2007). Assessing exposure to atrazine and its metabolites using biomonitoring. *Environ Health Perspect* **115**(10), 1474-1478.

Bassil, K. L., Vakil, C., Sanborn, M., Cole, D. C., Kaur, J. S., and Kerr, K. J. (2007). Cancer health effects of pesticides: systematic review. *Can Fam Physician* **53**(10), 1704-1711.

Brown, C. C., Havener, T. M., Medina, M. W., Krauss, R. M., McLeod, H. L., and Motsinger-Reif, A. A. (2012). Multivariate methods and software for association mapping in dose-response genome-wide association studies. *BioData Min* **5**(1), 21.

Burchell, B. (2003). Genetic variation of human UDP-glucuronosyltransferase: implications in disease and drug glucuronidation. *Am J Pharmacogenomics* **3**(1), 37-52.

Carvalho, R. N., Arukwe, A., Ait-Aissa, S., Bado-Nilles, A., Balzamo, S., Baun, A., Belkin, S., Blaha, L., Brion, F., Conti, D., Creusot, N., Essig, Y., Ferrero, V. E. V., Flander-Putrle, V., Furhacker, M., Grillari-Voglauer, R., Hogstrand, C., Jona , A., Kharlyngdoh, J. B., Loos, R., Lundebye, A. K., Modig, C., Olsson, P. E., Pillai, S., Polak, N., Potalivo, M., Sanchez, W., Schifferli, A., Schirmer, K., Sforzini, S., Sturzenbaum, S. R., Softeland, L., Turk, V., Viarengo, A., Werner, I., Yagur-Kroll, S., Zounkova, R., and Lettieri, T. (2014). Mixtures of Chemical Pollutants at European Legislation Safety Concentrations: How Safe Are They? *Toxicological Sciences* **141**(1), 218-233.

Cedergreen, N. (2014). Quantifying synergy: a systematic review of mixture toxicity studies within environmental toxicology. *PLoS One* **9**(5), e96580.

Collins, F. S., Gray, G. M., and Bucher, J. R. (2008). Toxicology. Transforming environmental health protection. *Science* **319**(5865), 906-907.

Crump, K. S., Chen, C., and Louis, T. A. (2010). The future use of in vitro data in risk assessment to set human exposure standards: challenging problems and familiar solutions. *Environ Health Perspect.* **118**(10), 1350-1354.

Dudbridge, F., and Gusnanto, A. (2008). Estimation of significance thresholds for genomewide association scans. *Genet Epidemiol* **32**(3), 227-234.

Feron, V. J., Cassee, F. R., and Groten, J. P. (1998). Toxicology of chemical mixtures: international perspective. *Environ Health Perspect* **106 Suppl 6**, 1281-1289.

Fraser, H. B. (2013). Gene expression drives local adaptation in humans. *Genome Res* **23**(7), 1089-1096.

Gangwal, S., Reif, D. M., Mosher, S., Egeghy, P. P., Wambaugh, J. F., Judson, R. S., and Hubal, E. A. (2012). Incorporating exposure information into the toxicological prioritization index decision support framework. *Sci Total Environ* **435-436**, 316-325.

Hadrup, N. (2014). Evidence from pharmacology and pathophysiology suggests that chemicals with dissimilar mechanisms of action could be of bigger concern in the toxicological risk assessment of chemical mixtures than chemicals with a similar mechanism of action. *Regul Toxicol Pharmacol* **69**(3), 281-283.

Jamei, M., Marciniak, S., Feng, K., Barnett, A., Tucker, G., and Rostami-Hodjegan, A. (2009). The Simcyp population-based ADME simulator. *Expert Opin Drug Metab Toxicol* **5**(2), 211-223.

Johnson, W. E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* **8**(1), 118-127.

Judson, R. S., Kavlock, R. J., Setzer, R. W., Cohen Hubal, E. A., Martin, M. T., Knudsen, T. B., Houck, K. A., Thomas, R. S., Wetmore, B. A., and Dix, D. J. (2011). Estimating Toxicity-Related Biological Pathway Altering Doses for High-Throughput Chemical Risk Assessment. *Chem Res Toxicol* **24**(4), 451-462.

Jurewicz, J., and Hanke, W. (2008). Prenatal and childhood exposure to pesticides and neurobehavioral development: review of epidemiological studies. *Int J Occup Med Environ Health* **21**(2), 121-132.

Krewski, D., Westphal, M., Al Zoughool, M., Croteau, M. C., and Andersen, M. E. (2011). New directions in toxicity testing. *Annu Rev Public Health* **32**, 161-178.

Krieger, R. (2010). *Haye's Handbook of Pesticide Toxicology*. 3rd ed. Academic Press, Waltham, MA.

Lock, E. F., Abdo, N., Huang, R., Xia, M., Kosyk, O., O'Shea, S. H., Zhou, Y. H., Sedykh, A., Tropsha, A., Austin, C. P., Tice, R. R., Wright, F. A., and Rusyn, I. (2012). Quantitative high-throughput screening for chemical toxicity in a population-based in vitro model. *Toxicol Sci* **126**(2), 578-588.

Manikkam, M., Tracey, R., Guerrero-Bosagna, C., and Skinner, M. K. (2012). Pesticide and insect repellent mixture (permethrin and DEET) induces epigenetic transgenerational inheritance of disease and sperm epimutations. *Reprod Toxicol* **34**(4), 708-719.

Marrone, A. K., Beland, F. A., and Pogribny, I. P. (2014). Noncoding RNA response to xenobiotic exposure: an indicator of toxicity and carcinogenicity. *Expert Opin Drug Metab Toxicol* **10**(10), 1409-1422.

Mauri, A., Consonni, V., Pavan, M., and Todeschini, R. (2006). DRAGON Software: An easy approach to molecular descriptor calculations. *MATCH Commun Math Comput Chem* **56**, 237-248.

Meech, R., Miners, J. O., Lewis, B. C., and Mackenzie, P. I. (2012). The glycosidation of xenobiotics and endogenous compounds: versatility and redundancy in the UDP glycosyltransferase superfamily. *Pharmacol Ther* **134**(2), 200-218.

National Research Council (2009). *Science and Decisions: Advancing Risk Assessment*. National Academies Press, Washington, DC.

O'Shea, S. H., Schwarz, J., Kosyk, O., Ross, P. K., Ha, M. J., Wright, F. A., and Rusyn, I. (2011). In vitro screening for population variability in chemical toxicity. *Toxicol Sci* **119**(2), 398-407.

Perry, L., Adams, R. D., Bennett, A. R., Lupton, D. J., Jackson, G., Good, A. M., Thomas, S. H., Vale, J. A., Thompson, J. P., Bateman, D. N., and Eddleston, M. (2014). National toxicovigilance for pesticide exposures resulting in health care contact - An example from the UK's National Poisons Information Service. *Clin Toxicol (Phila)* **52**(5), 549-555.

Pillai, K. (1955). Some new test criteria in multivariate analysis. *Ann Math Stat* **26**, 117-121.

Rother, H. A. (2014). Communicating pesticide neurotoxicity research findings and risks to decision-makers and the public. *Neurotoxicology* **45**, 327-337.

Rotroff, D. M., Wetmore, B. A., Dix, D. J., Ferguson, S. S., Clewell, H. J., Houck, K. A., LeCluyse, E. L., Andersen, M. E., Judson, R. S., Smith, C. M., Sochaski, M. A., Kavlock, R. J., Boellmann, F., Martin, M. T., Reif, D. M., Wambaugh, J. F., and Thomas, R. S. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol Sci* **117**(2), 348-358.

Sanborn, M., Kerr, K. J., Sanin, L. H., Cole, D. C., Bassil, K. L., and Vakil, C. (2007). Non-cancer health effects of pesticides: systematic review and implications for family doctors. *Can Fam Physician* **53**(10), 1712-1720.

Schaid, D. J., Sinnwell, J. P., Jenkins, G. D., McDonnell, S. K., Ingle, J. N., Kubo, M., Goss, P. E., Costantino, J. P., Wickerham, D. L., and Weinshilboum, R. M. (2012). Using the gene ontology to scan multilevel gene sets for associations in genome wide association studies. *Genet Epidemiol* **36**(1), 3-16.

Stingl, J. C., Bartels, H., Viviani, R., Lehmann, M. L., and Brockmoller, J. (2014). Relevance of UDP-glucuronosyltransferase polymorphisms for drug dosing: A quantitative systematic review. *Pharmacol Ther* **141**(1), 92-116.

Tan, Y. M., Liao, K. H., and Clewell, H. J., 3rd (2007). Reverse dosimetry: interpreting trihalomethanes biomonitoring data using physiologically based pharmacokinetic modeling. *J Expo Sci Environ Epidemiol* **17**(7), 591-603.

U.S. EPA (2002). Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity. In (Office of Pesticide Programs, Washington, DC.

Wambaugh, J. F., Setzer, R. W., Reif, D. M., Gangwal, S., Mitchell-Blackwood, J., Arnot, J. A., Joliet, O., Frame, A., Rabinowitz, J., Knudsen, T. B., Judson, R. S., Egeghy, P., Vallero, D., and Cohen Hubal, E. A. (2013). High-throughput models for exposure-based chemical prioritization in the ExpoCast project. *Environ Sci Technol* **47**(15), 8479-8488.

Wetmore, B. A. (2015). Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment. *Toxicology*, in press.

Wetmore, B. A., Allen, B., Clewell, H. J., 3rd, Parker, T., Wambaugh, J. F., Almond, L. M., Sochaski, M. A., and Thomas, R. S. (2014). Incorporating population variability and susceptible subpopulations into dosimetry for high-throughput toxicity testing. *Toxicol Sci* **142**(1), 210-224.

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Sochaski, M. A., Rotroff, D. M., Freeman, K., Clewell, H. J., 3rd, Dix, D. J., Andersen, M. E., Houck, K. A., Allen, B., Judson, R. S., Singh, R., Kavlock, R. J., Richard, A. M., and Thomas, R. S. (2012). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol Sci* **125**(1), 157-174.

World Health Organization (2005). Chemical-specific adjustment factors for interspecies differences in human variability: Guidance document for use of data in dose/concentration-response assessment. http://www.inchem.org/documents/harmproj/harmproj/harmproj2.pdf

Xia, M., Huang, R., Witt, K. L., Southall, N., Fostel, J., Cho, M. H., Jadhav, A., Smith, C. S., Inglese, J., Portier, C. J., Tice, R. R., and Austin, C. P. (2008). Compound cytotoxicity profiling using quantitative high-throughput screening. *Environ Health Perspect* **116**(3), 284-291.

Zeise, L., Bois, F. Y., Chiu, W. A., Hattis, D., Rusyn, I., and Guyton, K. Z. (2013). Addressing human variability in next-generation human health risk assessments of environmental chemicals. *Environ Health Perspect* **121**(1), 23-31.

Table 1. Human populations from which lymphoblast cell lines were selected for this study.

Population	# of Cell lines screened	% of Total	N males	N females
CEU: Utah residents with Northern & Western European ancestry	47	32.2	24	23
YRI: Yoruban in Ibadan, Nigeria	40	27.4	19	21
TSI: Tuscan in Italy	32	21.9	16	16
GBR: British from England & Scotland	27	18.5	14	13
Total	146	100	73	73

Table 2. Chemicals contained in the chlorinated pesticide mixture.

Constituent name	MW	Constituent CAS#	µg in 1 mL	µM in 1 mL	% in 1mL
α-Benzene hexachloride (BHC)	290.8	319-84-6	107	0.368	5.60
β-Benzene hexachloride (BHC)	290.8	319-85-7	55	0.189	2.88
γ-Benzene hexachloride (Lindane)	290.8	58-899/55963-79-6	151	0.519	7.90
δ-Benzene hexachloride (BHC)	290.8	319-86-8	41	0.141	2.15
cis-Chlordane	409.8	5103-71-9	18	0.044	0.67
trans-Chlordane	409.8	5103-74-2	15	0.037	0.56
4,4'-DDD (Dichlorodiphenyldichloro ethane)	320.1	72-54-8	293	0.915	13.94
4,4'-DDE (Dichlorodiphenyldichloro ethylene)	318.0	72-55-9	1,193	3.75	57.11
4,4'-DDT (dichlorodiphenyltrichloro ethane)	354.5	50-29-3	176	0.496	7.56
Dieldrin	380.9	60-57-1	41	0.108	1.64
		Cumulative concentration	2090	6.57	**100**

Table 3. Chemicals contained in the current use pesticide mixture.

Constituent name	MW	Constituent CAS#	µg in 1 mL	µM in 1 mL	% in 1mL
Metolachlor	283.8	94449-58-8/51218-45-2	115	0.405	22.77
2,6-Diethylaniline	149.2	579-66-8	1,259	8.44	19.76
Molinate	187.3	2212-67-1	139	0.742	19.45
Tebuthiuron	228.3	34014-18-1	65	0.285	14.74
Trifluralin	335.5	1582-09-8/75635-23-3	78	0.232	2.40
Chlorothalonil	265.9	1897-45-6	29	0.109	2.00
Prometon	225.3	1610-18-0	74	0.328	1.42
Butylate	217.4	2008-41-5	193	0.888	1.35
Benfluralin	335.3	1861-40-1	76	0.227	1.31
Alachlor	269.8	15972-60-8	37	0.137	1.23
Ethoprop	242.3	13194-48-4	45	0.186	1.09
Desisopropyl atrazine	173.6	1007-28-9	1,271	7.32	1.04
Metribuzin	214.3	21087-64-9	98	0.457	0.89
Diazinon	304.4	333-41-5	89	0.292	0.79
Disulfoton	274.4	298-04-4	26	0.095	0.77
Aldicarb	190.3	116-06-3	92	0.484	0.74
Methyl parathion	263.2	298-00-0	36	0.137	0.66
Ethalfluralin	333.3	55283-68-6	82	0.246	0.63
Pebulate (Tilliam)	203.4	1114-71-2	56	0.275	0.61
Cyanazine	240.7	21725-46-2/11096-88-1	31	0.129	0.55
Permethrin	391.3	52645-53-1	39	0.1	0.52
Carbofuran	221.3	1563-66-2	85	0.384	0.50
Chlorpyrifos (Dursban)	350.6	2921-88-2	71	0.202	0.48
Prometryne	241.4	7287-19-6/83653-07-0	42	0.174	0.47
Carbaryl	201.2	63-25-2	106	0.527	0.39
Desethyl atrazine	187.6	6190-65-4	1,352	7.21	0.37
Flumetralin	421.7	62924-70-3	81	0.192	0.37
Dacthal	332	65862-98-8/1861-32-1	15	0.045	0.37
Atrazine	215.7	1912-24-9	1,178	5.46	0.35
Simazine	201.7	122-34-9	101	0.501	0.35
Terbufos	288.4	13071-79-9	42	0.146	0.35
Fonofos (Dyfonate)	246.3	944-22-9	32	0.13	0.32
Pendimethalin	281.3	40487-42-1	33	0.117	0.29
Fenamiphos	303.4	22224-92-6	54	0.178	0.27
Tribufos (DEF 6)	314.5	78-48-8	41	0.13	0.26
Napropamide	271.4	15299-99-7	37	0.136	0.12
		Cumulative concentration	7200	37.0	**100**

Table 4. Summary statistics for the range in EC_{10} values for each mixture.

Pesticide mixture	Mean*	STD^{+}	Range	Median	Q_{05}*	Q_{95}^{φ}	UFd$^{\lambda}$
Chlorinated pesticides	11.6	1.96	(0.180-40.6)	13.1	4.36	21.7	3.00
Current use pesticides	11.1	1.85	(0.649-39.9)	11.9	3.89	24.7	3.05

*, All values (except for UFd column) are in μM.

$^{+}$ The standard deviation of the mean EC_{10}

* The value corresponding to the 5th percentile of EC_{10} across 146 averaged values for each mixture.

$^{\varphi}$ The value corresponding the 95th percentile of EC_{10} across 146 averaged values for each mixture.

$^{\lambda}$ The population toxicodynamic uncertainty factor $[10^{(q95-q50)}]$ for each mixture.

Table 5. Significant and suggestive (FWER P-value <0.2) EC_{10} –gene set associations.

Mixture	Gene set	Gene set name	N of genes	*Adjusted p*-value (FWER)	Top genes in the gene set
Chlorinated pesticides	KEGG	Ascorbate and aldarate metabolism	22	0.009	UGT2B11, UGT2B7, UGT1A3, UGT1A7, UGT1A4, UGT1A5, UGT1A6
Chlorinated pesticides	KEGG	Starch and sucrose metabolism	48	0.034	UGT2B11, UGT2B7, UGT1A3, UGT1A7, UGT1A4, UGT1A5, UGT1A6
Chlorinated pesticides	KEGG	Porphyrin and chlorophyll metabolism	39	0.06	EARS2, UGT2B11, UGT2B7, BLVRA, UGT1A3, UGT1A7, UGT1A4
Chlorinated pesticides	KEGG	Pentose and glucuronate interconversions	28	0.08	UGT2B11, UGT2B7, UGT1A3, UGT1A7, UGT1A4, UGT1A5, UGT1A6
Chlorinated pesticides	KEGG	Nitrogen metabolism	23	0.08	CA6, GLUL, CA2, CA4, HALCTH, CA5A
Chlorinated pesticides	KEGG	p53 signaling pathway	68	0.185	DDB2, CCNE2, CHEK1, TP73, CD82, SFN, SERPINE1

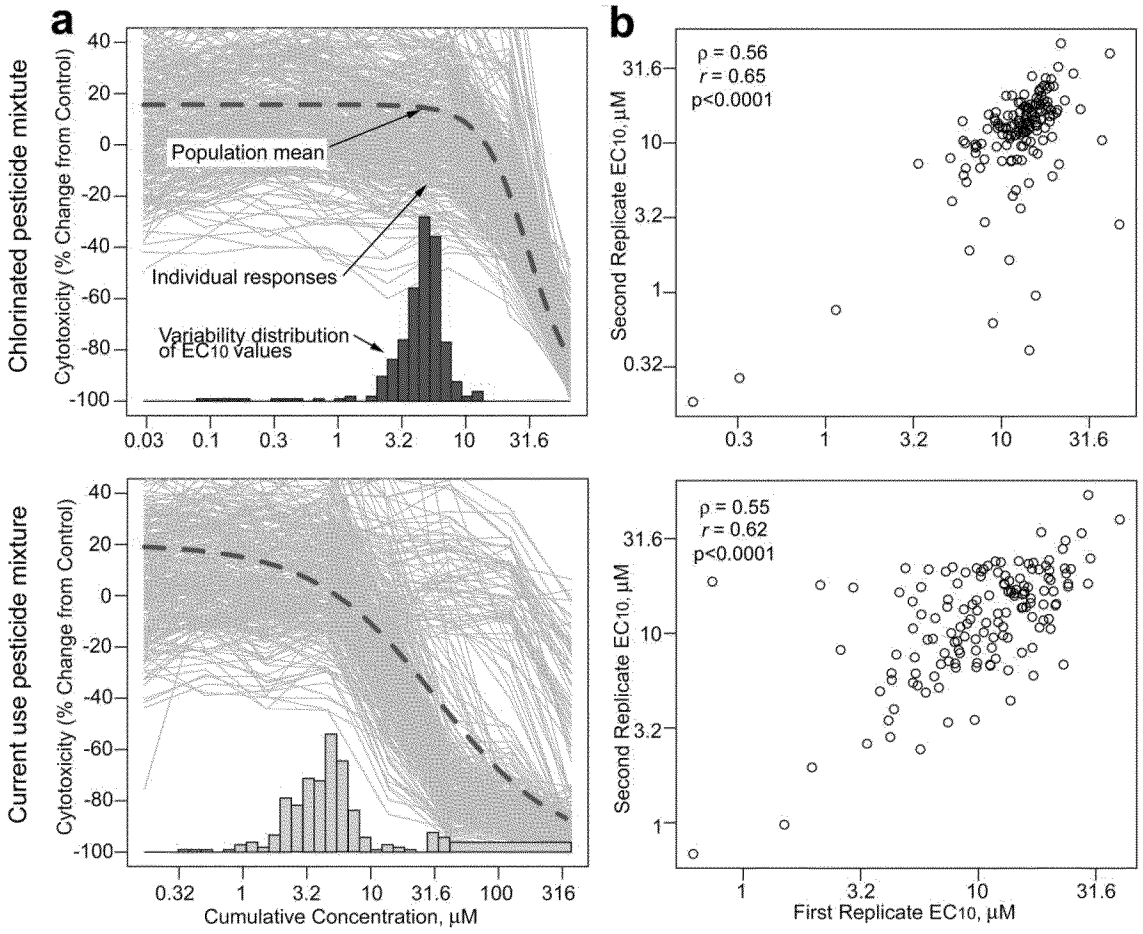


Figure 1. Inter-individual and population variability and reproducibility of the cytotoxicity of pesticide -containing mixtures in human lymphoblast cell lines . (A) A population concentration response was modeled using *in vitro* cytotoxicity of the chlorinated pesticide mixture (top) and the current use pesticide mixture (bottom). Logistic dose–response modeling was applied to each individual cell line , with individual data shown by thin gray lines. Bars represent a histogram of the individual EC_{10} values, and the dashed curve represents the fit of the logistic model to the pooled data. **(B)** Intra-experimental reproducibility of EC_{10} values for within-batch replicate plates for cell lines for the chlorinated pesticide mixture (top) and the current use pesticide mixture (bottom). Spearman and Pearson's correlation coefficients are shown.

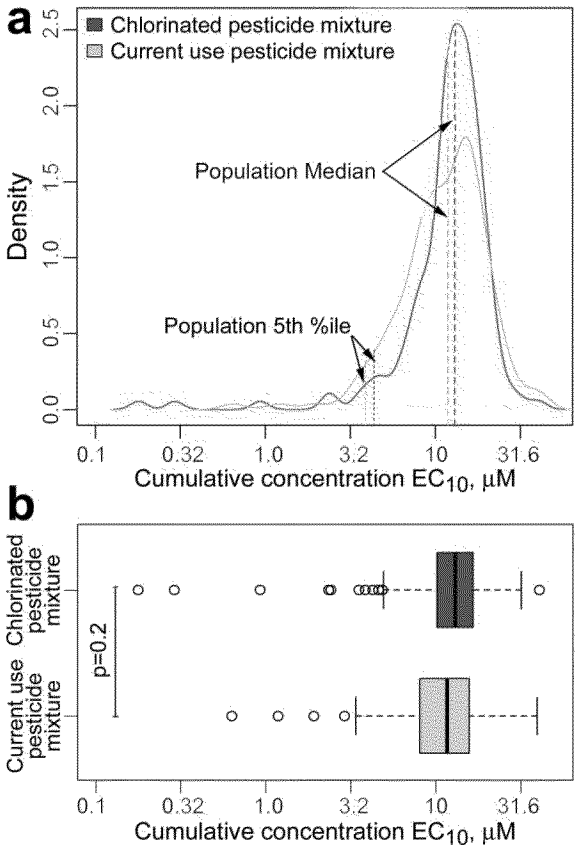


Figure 2. Distribution of EC_{10}s across 146 cell lines for each mixture. **(A)** A density plot for the distribution and mean of EC_{10} of each pesticide mixture (red: chlorinated pesticide mixture, blue: current use pesticide mixture) across 146 cell lines. **(B)** Box plots (box represents first and third quartiles; vertical line inside the box, the median; whiskers are the 1.5 inter-quantile range; circles are outliers with >1.5 IQR above minimum or maximum).

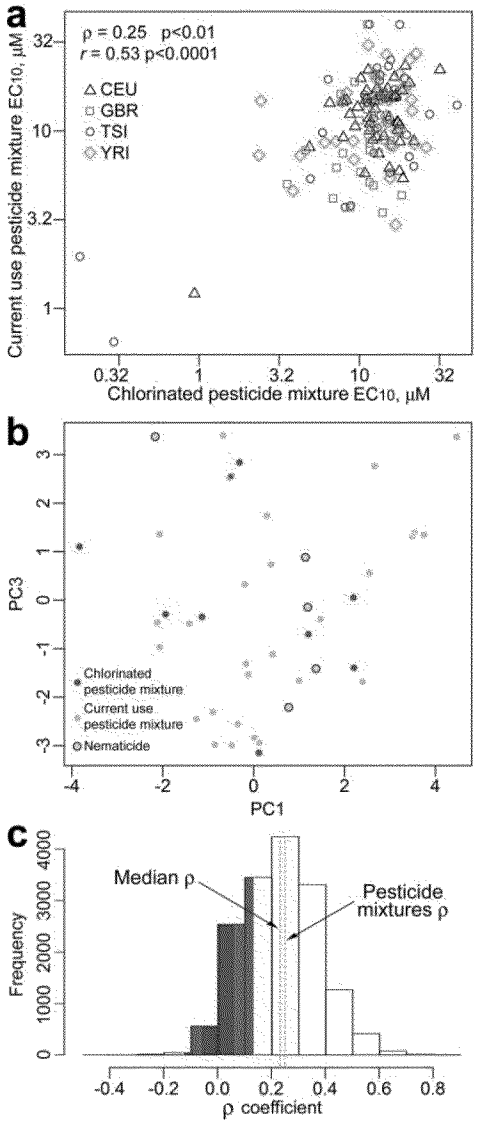


Figure 3. Comparative analysis of the mixtures. **(A)** Scatter plot comparison of EC_{10} values of each cell line between pesticide mixtures. Symbols represent populations as shown in the inset. Pearson and Spearman correlations are also shown. **(B)** Scatter plot of 1 st and 3 rd principal components of the molecular descriptors of the individual chemicals in each pesticide mixture. **(C)** Frequency histogram of 15,931 pair-wise correlation values (Spearman) among 179 chemicals screened in (Abdo, et al., 2015). The green dashed line represents a median ρ value for all correlations, and the red dashed line represents pairwise correlation of pesticide mixtures. Blue shading represents non-significant correlations after correction for false discoveries.

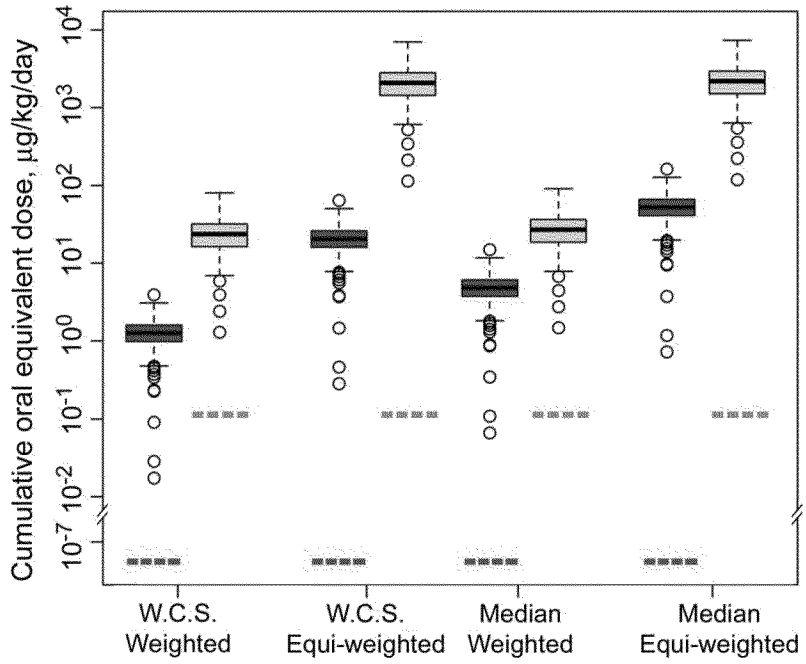


Figure 4. ***In vitro* -to-*in vivo* extrapolation of cytotoxicity EC_{10} values.** Box plots (box represents first and third quartiles; horizontal line inside the box is the median; whiskers are the 1.5 inter-quantile range; circles are outliers with >1.5 IQR above minimum or maximum) of the cumulative oral doses for each pesticide mixture (red: chlorinated pesticide mixture, blue: current use p esticide mixture) across 146 cell lines in four different scenarios for handling missing data , w eighted by chemical percentage in the mixture or not ("equi-weighted"), and assuming the "worst case scenario " (WCS) vs median for missing values. Red and blue dotted horizontal lines indicate the estimated cumulative human oral exposure levels to each pesticide mixture.

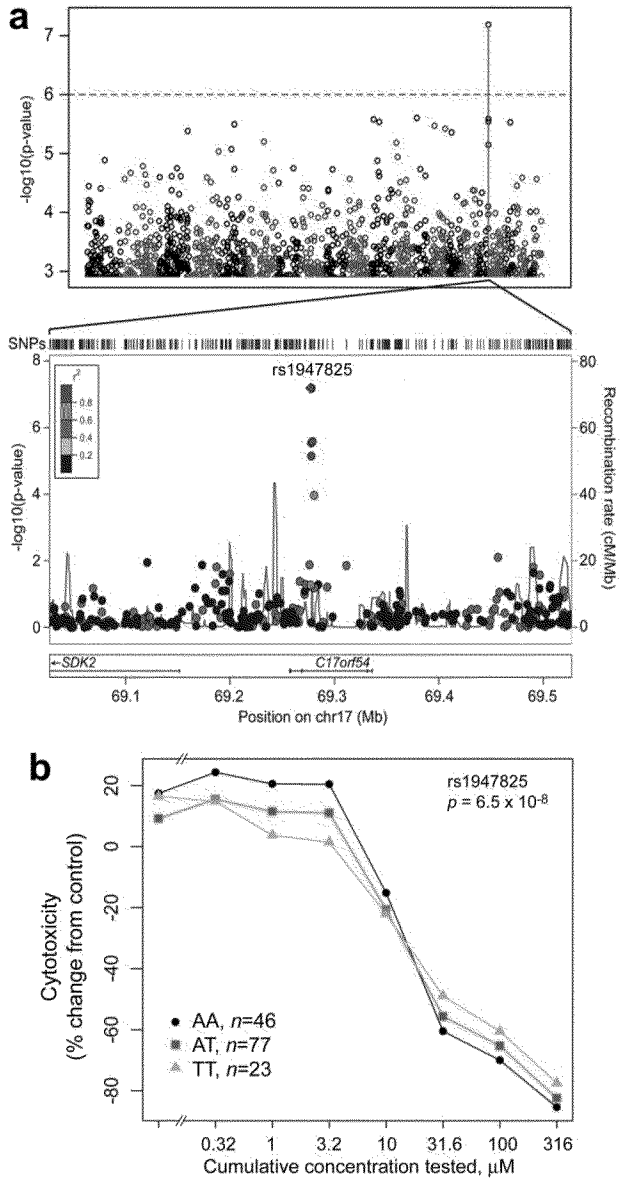
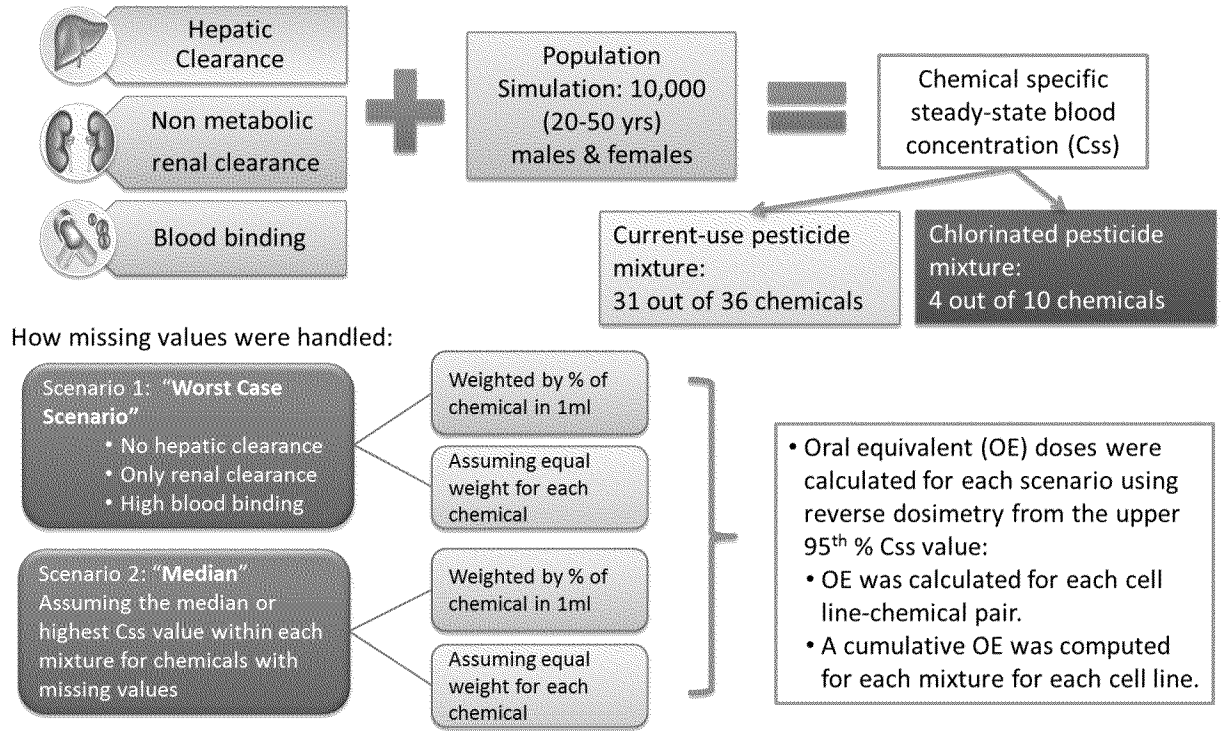
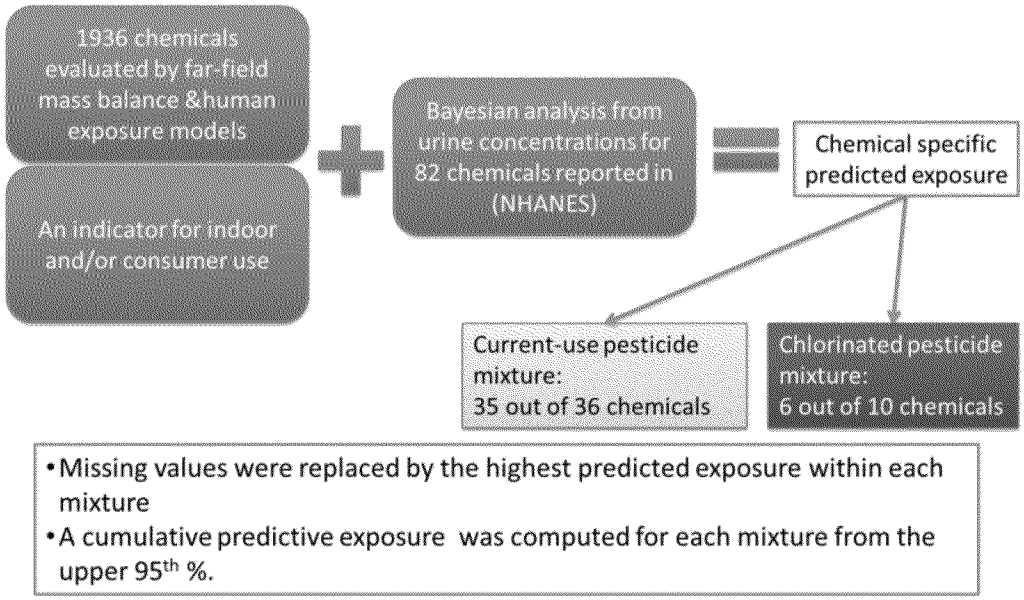


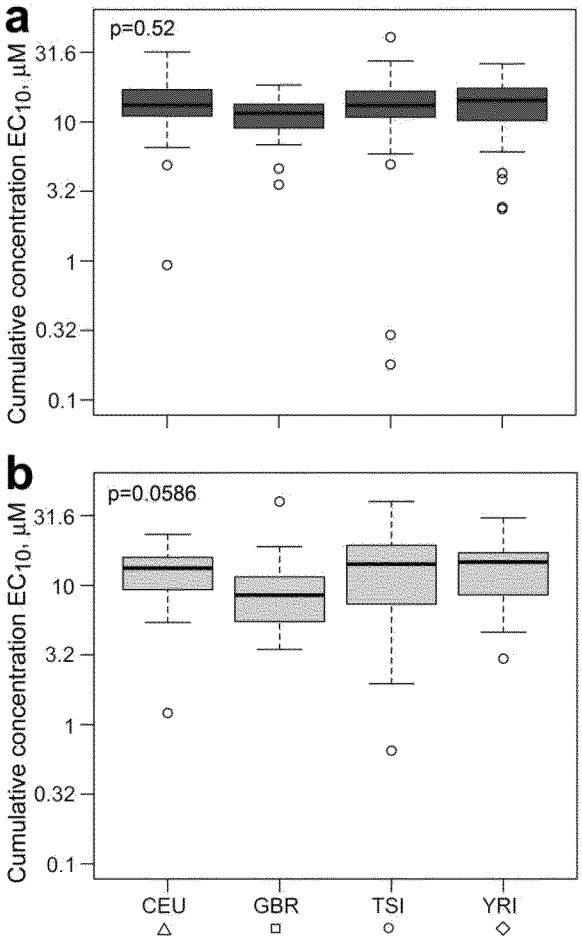
Figure 5. Genome-wide association analysis of population variability in cytotoxicity of th e current u se pesticide mixture. (A) Manhattan plot of MAGWAS -log10(p) vs. genomic position for association of genotype and cytotoxicity to current use pesticide mixture. The dashed blue line indicates suggestive association (expected once per genome scan). A LocusZoom plot of the most significant ($P=6.5\times10^{-8}$) region at SNP rs1947825. **(B)** Average concentration-response profiles of cytotoxicity of current use pesticide mixture plotted separately for each genotype at rs1947825.



Supplemental Figure 1. Workflow diagram for calculations of o ral equivalent doses for pesticide mixtures used in this study . Chemical specific steady -state values were obtained as previously described in (Wetmore, et al., 2012).



Supplemental Figure 2. Workflow diagram for calculations of predicted exposure levels for pesticide mixtures used in this study . Chemical specific predicted exposure was obtained as previously described in (Wambaugh, et al., 2013).



Supplemental Figure 3. Boxplots of inre-population differences in cytotoxicity of pesticide mixtures. Box plots (box represents first and third quartiles; horizontal line inside the box is the median; whiskers are the 1.5 inter -quantile range; circles are outliers with >1.5 IQR above minimum or maximum) of the EC_{10} values (red: chlorinated pesticide mixture, blue: current use pesticide mixture) across cell lines separated by population. ANOVA *p*-values for population differences are shown in the inset.