Author's response to reviews

Title: The effect of oxythioquinox exposure on normal human mammary epithelial cell gene expression: A microarray analysis study

Authors:
Maureen R Gwinn (mwg9@cdc.gov)
Diana L Whipkey (drw9@cdc.gov)
Ainsley Weston (agw8@cdc.gov)

Version: 4 Date: 13 September 2004

Author's response to reviews:

Title : Gene Expression Profile of Normal Human Mammary Epithelial Cells Following Exposure to Oxythioquinox as Monitored by DNA Microarrays
Authors: Maureen R Gwinn, Diana L Whipkey and Ainsley Weston

Dear Dr. Gwinn,

Thank you for revising your interesting manuscript (above) and for providing detailed responses to the reviewer comments. We have now received the reports from the four reviewers. Two of them have no further comments, although Dr Yang notes in regard to his first comment that free software is available, such as TreeView, that may allow most of the clustering analysis and heatmaps, and that such global gene expression profiles will provide very useful reference information for future gene expression studies among human populations. Further comments from two reviewers are copied below.

We realize that this report may not be a definitive study, but we feel that pilot data of the kind that you have obtained are useful when accompanied with proper caveats. We therefore believe that your work is potentially suitable for publication in Environmental Health, and we would like you to address the reviewer comments in the form of a revised manuscript, together with a point by point response to the comments. We will then be happy to consider the paper further for publication, which may necessitate sending it back to the reviewers.

Some formatting changes are also necessary. The revised title does not include study type. Please see http://www.ehjournal.net/info/instructions/ for the correct format. The abstract cannot exceed 350 words (the strict limit used by PubMed), and you have 356. Please remove all hyphens at line breaks (you can move the word down a line by using a soft-return (control+enter)). All references need to be properly formatted according to our instructions. For both tables, columns of data should be made visibly distinct by ensuring the borders of each cell display as black lines. RT-PCR is missing in the list of abbreviations. Not all instances of "select genes" were changed to "selected genes" as requested by a reviewer. Please crop figures so that they include a minimal amount of white space, probably best by converting to another format, crop, and reupload (make sure they are at least 300dpi or greater and ok when changed to a width of 600 pixels).

Your revised manuscript should be in strict accordance with our instructions for authors, cf. the pre-acceptance checklist at http://www.biomedcentral.com/info/edgr-preacceptcheck.asp
You may also want to look at one of the previously published papers for guidance.

When you have revised your manuscript in the light of the reviewers' comments and made the required changes to the format of the paper, please re-upload the revised files, and make any necessary changes to your manuscript submission forms on the Environmental Health web site, by clicking 'My Environmental Health', logging in with your user name and password and then clicking the manuscript title to go through to your submission screen. You should provide your point-by-point response to the reviewers' comments, along with a summary of the changes you have made, by uploading them at the end of the submission process, using the 'Cover letter' box.

We would expect you to return the revised paper within three weeks (i.e. by 17 September, 2004), but if you
Kristian Almstrup:
Many of my questions have been answered sufficiently, however I still have two major concerns with this manuscript:

**Major points**

The link is not working: http://www.cdc.gov/niosh/extsupp-mat/weston\AlteredGenes.htm

Without access to the data, I cannot make a hierarchical clustering and get an impression of how an unselected clustering comes out. I once again encourage the authors to acquire software, which are capable of doing hierarchical clustering. Such software is freely available i.e. at The Institute for Genomic Research (www.tigr.org/software/tm4).

A hyphen was removed from the link while formatting this paper for EHJournal. This has been repaired, so data should be accessible. This accessibility will allow other investigators, if interested, to reanalyze the data in a variety of software packages. As stated previously, our analysis was performed with dedicated software for Affymetrix analysis which does not at this time have the capabilities to do hierarchical clustering. To reanalyze the data at this point with new software as mentioned above may distort the results presented here due to differences in algorithms for both software packages. The software suggested by the reviewer will be considered for future work in microarray analysis.

I am a bit confused as to what the purpose of this manuscript is. In the end of the Background section on page 2 it is stated "This study focused on the effect of interindividual variation in gene expression signatures." While in the beginning of the second paragraph on page 5 it is stated "The primary goal of this study was to look for consistent changes common to all donors that could potentially be used as biomarkers of exposure, creating a gene expression profile following OTQ exposure in normal human cells." This is stated again in the beginning of the conclusion, "The overall goal of this project was to create a gene expression profile for OTQ or related pesticide analogues with the hopes of finding genes to be used as potential biomarkers of exposure."

Is the purpose to describe inter-individual variation or to find biomarkers? I think it is important to make clear as inter-individual variation is demonstrated in the manuscript with the huge variation that is observed between strains in the gene expression analysis. It is however still possible that the observed variation is caused by technical variation when real time-PCR and microarray data differ that much (For DDH gene Strain 1 time point 15 min SLR with RT-PCR gives 1.23 and from the array analysis -0.4. Strain 3 time point 120 min RT-PCR -3.4 and array 0). However, if inter-individual variation is well described, then biomarkers cannot be very well defined. This is illustrated above, as I believe the DDH genes should be a biomarker. It is a difficult task to show both robust biomarkers and their variation since very reliable tools are needed to generate reproducible results across many samples. The authors are thus encouraged to include further results from the follow-up analysis as stated in the second paragraph in the discussion, "Future analysis with a larger number of cell strains will be used to follow-up this analysis on specific genes of interest."

This paper focused on two main goals: acquiring a gene expression profile in NHMECs following exposure to OTQ in order to look for 1) biomarkers of exposure to this pesticide and 2) inter-individual variation in response to this pesticide. We again tried to clarify our goals and strengthen our argument (p ?). Future work on this particular pesticide is not ongoing, for a variety of reasons. Primarily, due to the discontinuation of use in the United States of this pesticide, it is not a priority for analysis at this time.

**Minor points:**

It could be nice if the authors could state the OTQ production per year.

It was stated that OTQ has been discontinued in the US.

Mike Wang:
In this study, the authors conducted a cDNA microarray analysis of global gene expression profiles of cellular responses to a quinoxaline pesticide, oxythioquinox (OTQ). One of the major revisions of the manuscript is the change of its titles, which now emphasizes the effect of inter-individual variation in gene expression profiling revealed by microarray analysis using only four cell strains established from four different human subjects. However, it is a common knowledge that there are large inter-personal variations in gene expressions when microarray technology is applied to population-based human samples. In such kinds of studies, major efforts are focused on pulling out true signals from the random fluctuations of gene expressions in the studied population by any means. A large size of samples is often required to have enough power to achieve this purpose. Thus, with only four individuals included, there are very few scientific values to study the effect of inter-individual variation in response to pesticide exposure using microarray analysis.

The choice of mammary tissue for a model system has already been discussed. These tissues are easily accessible from normal human donors, allowing us to obtain a bank of cells showing a normal variation, as would be seen in the population at large. The use of four cell lines from these tissues has as much value if not more than the use of any one cell line (eg, MCF-7) as is often described in the literature. The use of these primary cultures is to determine potential biomarkers of exposure in humans, and therefore is a better model than an aberrant human cancer cell line or a mouse cell line. Primary cell cultures are closer to the normal cells found in humans than immortalized cell lines, or cells obtained from other animals. That being said, each type of model system has benefits and drawbacks. The primary cultures do show inter-individual variability, which we chose to discuss as a second goal to this study. This can be seen as a benefit or drawback, depending on your viewpoint.

In addition, the results in the manuscript are not adequate to support the conclusions. First, in the Results section of Abstract, the authors indicated, " Cluster analysis examined the effects of OTQ on the cells with specific p53 polymorphisms major and intermediate. The two strains expressing the major variant of p53 had 83 common genes altered (35 increased, 48 decreased) at one or more time point ...". However, except the Abstract, the relevant data were not given in the context of manuscript. Second, the authors used loose criteria of expression altered in three or more of four cell strains at any treatment time point to select genes responded to pesticide exposure. Without statistical justification, the genes identified by these criteria are likely to be random noise given the background that thousands of genes were tested simultaneously.

One of the drawbacks in using primary cultures is the lack of reproducibility between donors as a result of the inter-individual variation. Looking at overall expression in these four cell strains, we were examining for similar trends in expression levels as well as temporal patterns. Genes that were increased at any time point in these strains are the discussed, and should be considered in any further studies on OTQ. Statistical validation of these results is from the Affymetrix analysis software. The algorithm used has been described in detail on their website, and briefly in our text. This algorithm takes into account the nature of microarray analysis to give accurate results regardless of background noise from the technology. Microarrays are still a new enough field that in the future there may be a more consistent method of analysis for all gene expression analysis, but until that time, the Affymetrix analysis software is accepted as means of statistical analysis.

In this study, RT-PCR was used to further examine the genes selected by the loose criteria. Although the authors claimed that the RT-PCR results confirmed data found by microarray analysis in the majority of samples, I found that the RT-PCR results were not consistent with microarray data. For two selected genes, CYP2A13 and DDH, there were clear different patterns of expression over treatment times existed among four cell strains (Figure 3 and suggesting large inter-individual variations. For gene CYP2A13, the expression by microarray analysis in three cell strains (except strain 3) was increased at 15 min OTQ treatment and unchanged at 120 min treatment (Figure 3). Whereas, the strain 3 had a reverse pattern, unchanged at 15 min treatment and increased at 120 min treatment. However, even using the same cDNA samples, RT-PCR demonstrated that all four strains were increased at both 15 min and 120 min treatment (Table 2). Furthermore, the microarray data and RT-PCR data for gene DDH were also not agree with each other.

The reasoning behind the variation between RT-PCR and microarray has been discussed (p?). This early rendition of the Affymetrix microarray has been discontinued, and new generations of this array have been found to have increased reproducibility with RT-PCR data. Differences in specificity of primers for RT-PCR and probes used on the array are one possibility for the variability.

Third, the data of the expression levels of the selected genes in Table 1 are peak value of one cell strain of one treatment time point. It is not the proper way to present data of gene expression of four cell strains with three OTQ treatment time points. For gene CYP2A13, the peak expression value in Table 1 is correspondent to the expression value of cell strain 2 at 15 min treatment (Figure 3). And for dihydrodiodide dehydrogenase, the peak expression value is correspondent to the expression value of cell strain 1 at 60 min treatment (Figure 3).
Table 1 is used as a summary table. In order to keep this table more readable and less bulky while still getting the point across as to which genes were of interest, select expression levels were shown. With all of the data made available, this seemed the cleanest way to show the trends in expression patterns. Lastly, there are still some type errors in the context of manuscript:

1. In page 12, the old figure citations were used, "Figures 3a and b", should be Figure 3 and 4.
This has been corrected.
2. Page 13, first paragraph indicates: "Genes altered in three of four cell strain in italics in Table 1." But there are not italics in table 1.
This has been removed.
3. Page 13, second paragraph indicates: "Results are shown in Table 2, with genes altered in three of four cell strain in italics." But there are not italics in table 2.
This has been removed.
4. Page 16 and 17, RTPCR should be RT-PCR.
This has been corrected.

**ADDITIONAL COMMENTS ON AUTHORS RESPONSES**

Original comments: I have several major concerns about this study. First, I am concerned about the study design. Why did the authors decide to use mammary epithelial cells? Whereas several previous studies demonstrated that the OTQ had demonstrated hepatotoxicity in laboratory animals. Hepatocytes and mammary epithelial cells are two different and well-differentiated cell types, and it is possible for them to have different expression profiles in response to OTQ exposure. Therefore, the authors should justify the use of human mammary epithelial cells in the study of OTQ-induced gene expression profiles.

Authors' response: The use of mammary epithelial cells was in order to look at inter-individual variation in response to various occupational chemicals. These cells are more easily obtained from normal human donors, and allowed the creation of a large tissue/DNA bank. Further, although OTQ exposure has not been related to breast cancer (and to my knowledge never studied in relation to this), pesticide use has been related to breast cancer. While the use of hepatocytes has been the main area of testing in the past, it would not have been possible to look at inter-individual variation in normal cell strains with hepatocytes, which was a main focus of the study. Most hepatic cell lines are immortalized, and would have limited the analysis of inter-individual variation. Comments can be found on p 5, 14.

New comments: To look at inter-individual variation in response to various occupational chemicals cannot justify using human mammary epithelial in OTQ exposure study. Please see first paragraph of the comments.
Please see comments above.

Original comments: Second, the sample size of only four cell strains is not large enough to conduct microarray analysis using primary cell lines established from human tissues. Genetic heterogeneity in human population results in tremendous variations of gene expressions. In microarray analysis, the scales of interpersonal variations are comparable or larger than the true signals, which is usually a major problem of successfully applying this technology on human samples. All of the data of this study, including immunoassay of p53 protein levels (Figure 1), microarray analysis (Figure 3) and RT-PCR (Table 2), clearly demonstrate that there are large interpersonal variations among four cell strains at baseline and each time point of OTQ treatment. Without proper adjustments of the interpersonal variations, it is very difficult to distinguish the true OTQ-induced responses from the random variations of expressions within the human populations. In addition, the batch effect of microarray analysis is another major source of artificial variation that can disturb the interpretation of data mining results. In this study, there were fewer details on how the microarray analyses were conducted, such as whether the RNA samples of vehicle control and each time points of each cell strain were analyzed in one batch experiment or in different batches. There is also a lack the details of data scaling and normalization. Thus, with the consideration of larger interpersonal variations of gene expressions and batch effects of microarray technology, the size of only four primary cell strains from human tissues cannot give enough power to allow the authors to assess the OTQ-induced responses.

Authors' response: Microarray analysis for each cell strain was performed in one day. All controls and treatments were performed at the same time for each individual cell strain. Normalization of the data was performed with the Affymetrix software based on a scaling factor (1500) looking at the results of all genes on the array. This information will be added to the paper. These are all good points, but suggest to me that the paper doesn't stress part of what we wanted to examine was the interindividual variation. Comments will be revised to stress the importance of this aspect of the study (p 5, 9, 14).

New comments: The way of authors' response to small sample size in this study is to change the goal and the title of study to look at inter-individual variation in response to various occupational chemicals. I don't think this issue was addressed properly. Please see first paragraph of the comments.

The title of this manuscript was changed per the editor's suggestion, in order to better fit the format of this journal. The goal has not been changed, but was reworded to clarify. Please see comments above regarding small sample size, as we feel this has been addressed sufficiently.

Original comments: Third, the data analysis is not conducted properly. The authors initially used SOM
method to identify a number of patterns of gene expressions at four time points in each cell strain, and then searched for common patterns across cell strains. According to SOM patterns, genes with a SLR to vehicle control >0.6 (~ larger than 1.5 fold change) in at least two of four cell strains were considered to be significantly changed their expressions in response to OTQ exposure. However, the microarray analyses using human samples are not like those using laboratory animals, where the individual animal can be regarded as replicate samples because of homogenous genetic background. Considering genetic heterogeneity and large interpersonal variations, an increase or decrease of expression in two or more of only four subjects cannot give statistically significant conclusions. With the available vehicle controls, the approach of paired t-tests of comparing each time point to baseline should be used in selecting genes with altered expressions on a larger size of human samples. Moreover, all RNA samples from each cell strain should be analyzed in the same batch of microarray analysis to control the systematic variations.

Authors’ response: All RNA from each cell strain was analyzed in the same batch of microarrays at the same time. All microarrays used were from the same lot as well. There was still array variability, but this was done to help control for that. Further, we did not use these samples as biological replicates, but to examine the variation between them and agree that a larger sample size is needed to confirm any results found here. We currently have over 40 tissues banked and ready for analysis if further studies are to be performed for follow-up. This study was originally undertaken to determine if there were any clear-cut differences between the individual cell strains selected, especially related to p53 haplotypes. This was not found to be the case, so further study is not warranted at this time (p 14).

New comments: the authors still not address how to adjust the inter-personal variations in their analysis. Changes were made in the last revision to state the use of housekeeping genes on p 9.

Original comments: Lastly, there is a concern that the authors conducted RT-PCR using the same cDNA synthesized from microarray analysis, and did not included housekeeping controls. Since there are large variations among reverse transcriptions, RT-PCR using cDNA synthesized from microarray analysis cannot be considered as an independent analysis to confirm the results of microarray analysis. To confirm the results of microarray analysis, RT-PCR should start from RNA samples.

Authors’ response: This is a valid point, but in general the RT-PCR is performed on the same cDNA to assess the results of the two methods. Housekeeping controls used were both 18S and GAPDH to normalize for discrepancies between samples. This information will be added. Since the analysis was performed in duplicate, any fluctuations in reverse transcription should have shown up as a bad replicate, and would have been re-analyzed (p 9).

New comments: Several reference of using RT-PCR to confirm microarray data were all started from RNA samples, not using the same cDNA produced in microarray analysis. The use of cDNA can also be found to validate microarray data, looking for confirmation of array results. These can also be found in various references, as well as many references that do not clearly state whether they used fresh RNA or cDNA from the array experiment. There can be a case made for both independent RNA sources and cDNA from the array experiment.