Mutagenesis vol. 21 no. 1 pp. 11–13, 2006
Advance Access Publication 27 September 2005
doi:10.1093/mutage/gei053

Performance of flow cytometric analysis for the micronucleus assay—a reconstruction model using serial dilutions of malaria-infected cells with normal mouse peripheral blood

Dorothea Torous*, Norihide Asano[1], Carol Tometsko, Siva Sugunan, Stephen Dertinger, Takeshi Morita[2] and Makoto Hayashi[3]

Litron Laboratories, 200 Canal View Boulevard, Rochester, NY 14623, USA. [1]Toxicological Research Center, Nitto Denko Corporation, 1-1-2 Shimohozumi, Ibaraki, Osaka 567-8680, Japan, [2]Division of Safety Information on Drug, Food and Chemicals and [3]Division of Genetics and Mutagenesis, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

To confirm the performance and statistical power of a flow cytometric method for scoring micronucleated erythrocytes, reconstruction experiments were performed. For these investigations, peripheral blood erythrocytes from untreated mice, with a micronucleated erythrocyte frequency of ~0.1% were combined with known quantities of *Plasmodium berghei* (malaria) infected mouse erythrocytes. These cells had an infected erythrocyte frequency of ~0.7%, and mimic the DNA content of micronuclei (MN). For an initial experiment, samples with a range of MN/malaria (Mal) content were constructed and analysed in triplicate by flow cytometry until 2000, 20 000 and 200 000 total erythrocytes were acquired. In a second experiment, each specimen was analysed in triplicate until 2000, 20 000, 200 000 and 1 000 000 erythrocytes were acquired. As expected, the sensitivity of the assay to detect small changes in rare erythrocyte sub-population frequencies was directly related to the number of cells analysed. For example, when 2000 cells were scored, increases in MN/Mal frequencies of 3.9- or 2.7-fold were detected as statistically significant. When 200 000 cells were analysed, a 1.2-fold increase was detected. These data have implications for the experimental design and interpretation of micronucleus assays that are based on automated scoring procedures, since previously unattainable numbers of cells can now be readily scored.

Introduction

From a statistical point of view, in order to achieve a higher power of detection, sample size should be increased. For many experimental situations, it is not always feasible to increase the number of subjects studied. When the event under consideration is rare as to cause appreciable scoring error, then an alternative would be to enhance the precision of each measurement. For example, in the rodent erythrocyte micronucleus assay, the evaluation of 2000 immature erythrocytes per animal and 5 animals per dose group represents commonly cited minimum values. Owing to the rarity of micronucleated cells, even this minimal assay design results in tedious and time-consuming efforts. The use of flow cytometry (1–3) realizes the ability to evaluate high numbers of erythrocytes, something that is impossible to achieve by manual microscopy. By reducing scoring error in this manner, flow cytometry has the potential to increase statistical power.

In the present study, we evaluated the relationship between statistical power to detect a rare erythrocyte sub-population, i.e. micronucleated or malaria-infected erythrocytes (MN/Mal), and the total number of erythrocytes analysed. These experiments were accomplished using a reconstruction model whereby known quantities of malaria-infected erythrocytes were added to blood from an untreated mouse. Malaria is a known model for micronucleated erythrocytes, as they endow the target cells of interest with a micronucleus-like DNA content (4,5). The samples were analysed by flow cytometry to measure the MN/Mal frequency through the interrogation of 2000 (2k), 20 000 (20k), 200 000 (200k) and 1 000 000 (1m) erythrocytes. The results presented here show the capability of flow cytometric technology to reduce scoring error, and also the extent to which this affects the ability to detect small changes to baseline micronucleus frequencies.

Materials and methods

Staining of blood specimens

Methanol-fixed blood from untreated and malaria-infected mice used in this study were two 'biological standards' which accompany the Mouse MicroFlow®PLUS kits (Litron Laboratories, NY). MicroFlow PLUS kits were the source of these specimens.

Before analysis, malaria-infected specimens and untreated mouse specimens were washed out of fixative with ~12 ml Hank's Balanced Salt Solution. Procedures for the 3-colour labelling technique which appear in the MicroFlow®PLUS instruction manual (version 031230) were scaled up ~7-fold in order to provide at least 10 ml each of control and malaria blood in a cell density range that is recommended for this assay (between ~2000 and 6000 events/s). Anti-CD71-FITC, anti-CD61-PE and all other flow cytometry reagents were also supplied in the kits. After the labelling procedures were accomplished, the cell density of the malaria-infected sample was adjusted so that it was equal to that of the control blood sample. Initial cell densities were measured with a Coulter Counter, model ZM. After adjustment with additional propidium iodide staining solution, equal cell densities were confirmed by Coulter Counter measurements. Normalization of cell densities was an important experimental design consideration, as this allowed us to calculate the expected MN/Mal frequencies in the diluted samples once the frequencies of the original control (0.10 and 0.09% for Experiments 1 and 2, respectively) and malaria-infected (0.67 and 0.70% for Experiments 1 and 2, respectively) samples were determined with high precision (i.e. control and malaria-infected %MN/Mal frequencies are the mean value of triplicate analyses with 1m erythrocytes per analysis).

Dilution of malaria blood specimen

Malaria-infected blood (Sample H) was diluted with control blood (Sample A) in the following ratios (v/v): 1:1 (Sample G), 1:3 (Sample F), 1:7 (Sample E), 1:15 (Sample D), 1:31 (Sample C) and 1:63 (Sample B). These blood specimens were stored at 4°C until flow cytometric analysis, which occurred on the same day. Each sample was analysed three times to evaluate reproducibility.

Flow cytometric analysis

All samples were analysed according to the MicroFlow® PLUS 3-colour technique. One deviation to the kit-supplied data acquisition and analysis template was that the frequency of erythrocytes with malaria or micronuclei was determined without restriction to CD71-expression level. That is, the

*To whom correspondence should be addressed. Email: dtorous@litronlabs.com

© The Author 2005. Published by Oxford University Press on behalf of the UK Environmental Mutagen Society.
All rights reserved. For permissions, please email: journals.permissions@oupjournals.org

Mal and MN frequencies measured and reported here are based on total peripheral blood erythrocytes. A second deviation from standard practices is that the default stop mode of 20 000 reticulocytes was not utilized. Rather, each specimen was analysed until the following number of erythrocytes were acquired: 2k, 20k and 200k erythrocytes in the first experiment and 2k, 20k, 200k and 1m erythrocytes in the second experiment.

Statistical analysis

The average of triplicate MN/Mal measurements associated with the control blood sample were compared with those associated with each of the other seven specimens by the Fisher's exact method. A P-value of 0.05 divided by 7 (number of sample groups) was considered evidence of a statistically significant difference. Expected versus observed MN/Mal frequencies were graphed for each measurement performed in the second experiment. Microsoft Excel (Microsoft Corp., Seattle, Washington) was used to determine a best-fit line. The associated equations and r^2 values were determined.

Results

Data from Experiments 1 and 2 are summarized in Table I and include the expected and observed MN/Mal frequencies. The MN/Mal frequencies shown are the average of triplicate analyses. As shown in Table I, for measurements based on 2k erythrocytes, samples with expected MN/Mal frequencies of 0.39 and 0.24% were found to be significantly different from control samples, in Experiments 1 and 2, respectively. These values correspond to fold increases of 3.9 and 2.7 for the first and second experiment, respectively. As more erythrocytes were analysed per sample, the detection limit was improved. For instance, measurements based on the evaluation of 200k erythrocytes per analysis show statistical significance for expected MN/Mal samples of 0.12 and 0.11%. These values correspond to an increase of ~1.2-fold. In fact for the second experiment, when a stop mode of 1m erythrocytes was investigated, statistical significance was observed between the control blood sample (0.09% MN/Mal) and the specimen with the lowest frequency of malaria (0.10% MN/Mal; $P = 0.00005$).

As an aid for visualizing the performance characteristics associated with the various number of cells analysed, scattergrams showing %MN/Mal measurement are presented (Fig. 1). Best-fit lines and equations are included with these graphs, and illustrate the degree to which the experimentally derived data agree with the linear relationship that is known to exist among MN/Mal frequencies for these specimens.

Discussion

To evaluate the performance and statistical power of a flow cytometric approach to score micronucleated erythrocytes, we performed a reconstruction model experiment by the serial dilution of malaria-infected mouse blood with normal mouse

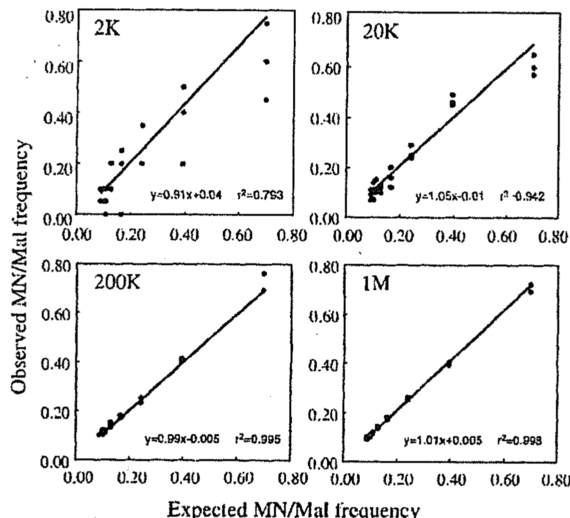

Fig. 1. Scattergram of expected versus observed MN/Mal frequencies. Each of three replicate measurements is plotted for these specimens. Best-fit linear lines are graphed, with associated equations. r^2 values document the degree of reproducibility observed.

Table I. MN/Mal frequencies (%) and P-values for comparisons with sample A

Sample	Expected (%)	Number of cells analysed/sample							
		2k		20k		200k		1m	
		(%)	P-value	(%)	P-value	(%)	P-value	(%)	P-value
Experiment 1									
A	0.10	0.07		0.09		0.10			
B	0.11	0.08	*0.50000*	0.11	*0.20764*	0.11	*0.13403*		
C	0.12	0.05	*0.77349*	0.10	*0.46272*	0.12	*0.00007*		
D	0.14	0.08	*0.50000*	0.13	*0.03463*	0.14	*0.00000*		
E	0.18	0.22	*0.02452*	0.21	*0.00000*	0.18	*0.00000*		
F	0.25	0.18	*0.05924*	0.30	*0.00000*	0.25	*0.00000*		
G	0.39	0.57	*0.00000*	0.36	*0.00000*	0.41	*0.00000*		
H	0.67	0.57	*0.00000*	0.73	*0.00000*	0.67	*0.00000*		
Experiment 2									
A	0.09	0.08		0.09		0.10		0.09	
B	0.10	0.08	*0.62305*	0.11	*0.20374*	0.11	*0.04822*	0.10	*0.00005*
C	0.11	0.05	*0.85547*	0.12	*0.06479*	0.12	*0.00001*	0.11	*0.00000*
D	0.13	0.13	*0.29053*	0.12	*0.08887*	0.14	*0.00000*	0.15	*0.00000*
E	0.17	0.15	*0.21198*	0.16	*0.00038*	0.17	*0.00000*	0.17	*0.00000*
F	0.24	0.30	*0.00531*	0.26	*0.00000*	0.24	*0.00000*	0.25	*0.00000*
G	0.40	0.37	*0.00076*	0.47	*0.00000*	0.41	*0.00000*	0.40	*0.00000*
H	0.70	0.60	*0.00000*	0.61	*0.00000*	0.71	*0.00000*	0.70	*0.00000*

Shading indicates those samples that are significantly different from respective control samples.

blood. As expected, the present results illustrate that the power of rare event detection is directly related to the number of cells analysed per specimen. By analysing 3m (triplicate of 1m) cells per group, 0.10% is significantly different ($P = 0.00005$) when compared with 0.09%. Even so, it must be appreciated that the biological significance of minute changes must be considered in addition to statistical significance.

Previously, we have shown that individual differences were negligible in the mouse micronucleus assay when 1000 cells per animal and 5 or 6 mice per dose group were analysed (6–8) and the statistical unit for the evaluation can be assigned to a cell but not to an animal. According to the present results and also results by Asano *et al.* (9), the variability of the data was high when 2k cells were analysed. Under these circumstances, the difference among animals is not apparent, as they are likely to be smaller than the scoring error. While, in the case of the present malaria dilution experiments, when 200k or 1m cells per sample were analysed, the scoring error decreased and converged to a value. This, however, is not true in the case of the actual micronucleus assays using model chemicals (9). When 200k or 1m immature erythrocytes were analysed, differences between individual animals became apparent and there was data variability within each dose group. Therefore, even if the experimental size in the animal experiments is increased, we cannot expect the same increment of detecting power. This finding suggests that optimizing the statistical procedure also includes evaluating individual differences.

Based on the present results, we confirm the accuracy and high performance of the micronucleus assay system using flow cytometry and we propose that the number of reticulocytes analysed for the micronucleus assay using flow cytometry be a minimum of 20k. We suggest that the analysis of 20k reticulocytes is approximately equivalent to the manual microscopic analysis according to test guideline OECD 474 (9,10). We anticipate that the experimental size of the MN assay will be recommended and set by expert committees based on the evaluated data. In addition to statistical sensitivity, biological variability between animals and as a consequence of treatment should also be considered. There appears to be diminishing value to analyses based on 200k or even 1m per animal. These may be useful in certain special circumstances, for instance when looking for evidence of threshold or practical threshold effects (9).

References

1. Grawé,J., Zetterberg,G. and Amnéus,H. (1992) Flow-cytometric enumeration of micronucleated polychromatic erythrocytes in mouse peripheral blood. *Cytometry*, **13**, 750–758.
2. Dertinger,S.D., Torous,D.K. and Tometsko,K.R. (1996) Simple and reliable enumeration of micronucleated reticulocytes with a single-laser flow cytometer. *Mutat. Res.*, **371**, 283–292.
3. Torous,D.K., Hall,N.E., Illi-Love,A.H. *et al.* (2005) Interlaboratory validation of a CD71-based flow cytometric method (Microflow) for the scoring of micronucleated reticulocytes in mouse peripheral blood. *Environ. Mol. Mutagen.*, **45**, 44–55.
4. Tometsko,A.M., Torous,D.K. and Dertinger,S.D. (1993) Analysis of micronucleated cells by flow cytometry. 1. Achieving high resolution with a malaria model. *Mutat. Res.*, **292**, 129–135.
5. Dertinger,S.D., Torous,D.K., Hall,N.E., Tometsko,C.R. and Gasiewicz,T.A. (2000) Malaria-infected erythrocytes serve as biological standards to ensure reliable and consistent scoring of micronucleated erythrocytes by flow cytometry. *Mutat. Res.*, **464**, 195–200.
6. Hayashi,M., Yoshimura,I. Sofuni,T. and Ishidate,M. (1989) A procedure for data analysis of the rodent micronucleus test involving a historical control. *Environ. Mol. Mutagen.*, **13**, 347–356.
7. Hayashi,M., Hashimoto,S., Sakamoto,Y., Hamada,C., Sofuni,T. and Yoshimura,I. (1994) Statistical analysis of data in mutagenicity assays: rodent micronucleus assay. *Environ. Health Perspect.*, **102** (Suppl. 1), 49–52.
8. Adler,I.-D., Bootman,J., Favor,J., Hook,G., Schriever-Schwemmer,G., Welzl,G., Whorton,E., Yoshimura,I. and Hayashi,M. (1998) Recommendations for statistical designs of in vivo mutagenicity test with regard to subsequent statistical analysis. *Mutat. Res.*, **417**, 19–30.
9. Asano,N., Torous,D.K., Tometsko,C.R., Dertinger,S.D., Morita,T. and Hayashi,M. (2005) Practical threshold for micronucleated reticulocyte induction observed for low doses of mitomycin C, Ara-C, and colchicine. *Mutagenesis*, **21**, 15–20.
10. OECD (1997) *Guideline for the Testing of Chemicals: Mammalian Erythrocyte Micronucleus Test. Guideline 474.* Organisation for Economic Cooperation and Development, Paris, France.

Received on May 23, 2005; revised on July 21, 2005; accepted on July 22, 2005

Available online at www.sciencedirect.com

ScienceDirect

Mutation Research 608 (2006) 29–42

MR Genetic Toxicology and Environmental Mutagenesis

www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

Evaluation of the ability of a battery of three *in vitro* genotoxicity tests to discriminate rodent carcinogens and non-carcinogens II. Further analysis of mammalian cell results, relative predictivity and tumour profiles

David Kirkland[a,*], Marilyn Aardema[b], Lutz Müller[c], Makoto Hayashi[d]

[a] *Covance Laboratories Limited, Otley Road, Harrogate HG3 1PY, United Kingdom*
[b] *The Procter and Gamble Company, Miami Valley Laboratories, P.O. Box 538707, Cincinnati, OH 45253-8707, USA*
[c] *Non-Clinical Drug Safety, F Hoffmann-La Roche Ltd., PRBN-S, Bldg 73311b, CH-4070 Basel, Switzerland*
[d] *Division of Genetics and Mutagenesis, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan*

Received 2 February 2006; received in revised form 10 April 2006; accepted 30 April 2006
Available online 12 June 2006

Abstract

One of the consequences of the low specificity of the *in vitro* mammalian cell genotoxicity assays reported in our previous paper [D. Kirkland, M. Aardema, L. Henderson, L. Müller, Evaluation of the ability of a battery of three *in vitro* genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. I. Sensitivity, specificity and relative predictivity, Mutat. Res. 584 (2005) 1–256] is industry and regulatory agencies dealing with a large number of false-positive results during the safety assessment of new chemicals and drugs. Addressing positive results from *in vitro* genotoxicity assays to determine which are "false" requires extensive resources, including the conduct of additional animal studies. In order to reduce animal usage, and to conserve industry and regulatory agency resources, we thought it was important to raise the question as to whether the protocol requirements for a valid *in vitro* assay or the criteria for a positive result could be changed in order to increase specificity without a significant loss in sensitivity of these tests. We therefore analysed some results of the mouse lymphoma assay (MLA) and the chromosomal aberration (CA) test obtained for rodent carcinogens and non-carcinogens in more detail. For a number of chemicals that are positive only in either of these mammalian cell tests (i.e. negative in the Ames test) there was no correlation between rodent carcinogenicity and level of toxicity (we could not analyse this for the CA test as insufficient data were available in publications), magnitude of response or lowest effective positive concentration. On the basis of very limited *in vitro* and *in vivo* data, we could also find no correlation between the above parameters and formation of DNA adducts. Therefore, a change to the current criteria for required level of toxicity in the MLA, to limit positive calls to certain magnitudes of response, or to certain concentration ranges would not improve the specificity of the tests without significantly reducing the sensitivity.

We also investigated a possible correlation between tumour profile (trans-species, trans-sex and multi-site *versus* single-species, single-sex and single-site) and pattern of genotoxicity results. Carcinogens showing the combination of trans-species, trans-sex and multi-site tumour profile were much more prevalent (70% more) in the group of chemicals giving positive results in all three *in vitro* assays than amongst those giving all negative results. However, single-species, single-sex, single-site carcinogens were not very prevalent even amongst those chemicals giving three negative results *in vitro*. Surprisingly, when mixed positive and negative results were compared, multi-site carcinogens were highly prevalent amongst chemicals giving only a single positive result in the battery of three *in vitro* tests.

* Corresponding author. Tel.: +44 1423 848401; fax: +44 1423 848983.
E-mail address: david.kirkland@covance.com (D. Kirkland).

1383-5718/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrgentox.2006.04.017

Finally we extended our relative predictivity (RP) calculations to combinations of positive and negative results in the genotoxicity battery. For two out of three tests positive, the RP for carcinogenicity was no higher than 1.0 and for 2/3 tests negative the RP for non-carcinogenicity was either zero (for Ames + MLA + MN) or 1.7 (for Ames + MLA + CA). Thus, all values were less than a meaningful RP of two, and indicate that it is not possible to predict outcome of the rodent carcinogenicity study when only 2/3 genotoxicity results are in agreement.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Carcinogens; Mouse lymphoma assay; Chromosomal aberration; Genotoxicity; Tumour profiles

1. Introduction

Current regulatory guidelines for genotoxicity testing require chemicals to be evaluated at high concentrations/molarity and/or cellular toxicity *in vitro* for a valid test. The low specificity in the mammalian cell assays reported in our previous paper [1] raises the question as to whether such protocol requirements for a valid assay or the criteria for a positive result (according to ICH [2] and OECD [3] guidelines) could be changed in order to increase specificity, thereby reducing the number of false-positive results obtained. Since the standard approach to determine the biological relevance of positive *in vitro* genotoxicity assays involves conduct of additional *in vivo* genotoxicity assays, efforts to reduce this high false-positive rate could lead to substantial reductions in animal testing, as well as the cost and time to run these tests. Therefore, we examined such questions as whether the sensitivity of a particular mammalian cell assay could be maintained, but the specificity improved, if lower levels of toxicity were required for a valid assay? Did the carcinogens induce genotoxic responses of greater magnitude than the non-carcinogens, such that by requiring a higher level of response for a positive result, specificity might be improved? Also, were the non-carcinogens positive at much higher concentrations than the carcinogens, such that by requiring a lower top concentration than the current 10 mM, specificity might be improved? In this paper we have analysed some of the mouse lymphoma assay (MLA) and chromosomal aberration (CA) test results with rodent carcinogens and non-carcinogens (as defined in our previous paper [1]) in light of these questions.

We also examined whether a consistent pattern of genotoxicity results e.g. positive in Ames, MLA and either *in vitro* micronucleus (MN) or CA tests correlated better with trans-species, trans-sex, multi-site carcinogens than with single-species, single-sex, single-site carcinogens.

Finally, the new concept of relative predictivity (RP) that we introduced in our previous paper [1] and that had been calculated and reported only for those chemicals giving all three results positive or all three results negative in the *in vitro* battery, was examined for other combinations of test results. Previously we reported that a reasonable correlation (RP of two or more) with rodent carcinogenicity was obtained when all three *in vitro* tests gave positive results (>3× more likely that a chemical would be a rodent carcinogen if all three tests were positive), and a reasonable correlation with non-carcinogenicity was obtained when all three tests were negative (>2× more likely that a chemical would be a non-carcinogen if all three tests were negative). As most compounds tested by industry give mixtures of positive and negative results across a battery of three tests (2 positive + 1 negative, or 1 positive + 2 negative), and not all regulatory guidelines request the conduct of an extensive *in vitro* battery, we calculated the RP for rodent carcinogenesis and non-carcinogenesis from these combinations of results.

2. Quantitative analysis of various aspects of the mammalian cell tests

An important source of poor specificity in our previous analysis [1] was for non-carcinogens where a single positive result was obtained in one of the mammalian cell tests in the battery. We therefore decided to examine whether changing the criteria for a valid assay or the criteria for a positive result would eliminate some of the isolated positives with non-carcinogens, whilst preserving the single positive results with carcinogens, thereby improving the specificity without decreasing the sensitivity. We decided to look at three parameters:

- Level of toxicity
- Magnitude of positive response
- Lowest clearly positive concentration

As indicated above, we focussed this analysis on those chemicals where the only positive response was in either the MLA or CA test, accompanied by negative results in the Ames test (we included one chemical with no Ames result), and either negative or no results in the other mammalian cell test.

Analysing these three parameters was relatively easy to do for the MLA because most published data (and many results came from NTP tests [4]) contained sufficient detail. However, very few CA publications, including the many NTP tests, provided detail on toxicity. We did have access to the laboratory notebooks of Dr. Ishidate, but were able to obtain toxicity information on too few clastogens to allow a meaningful analysis to be undertaken. For the MN test, where there are many fewer publications, there were so few situations where the MN was the only positive in the battery that this was not worthy of further analysis at this time.

As is the case in a retrospective analysis of this sort, it is important to recognise that the data from some studies may not meet current guideline requirements. Whilst it would be ideal to review each published result in terms of compliance with current guidelines, and focus our analysis only on those that complied, this goes beyond the scope of the project. As indicated previously [1], we did re-evaluate a significant number of the MLA studies from Mitchell et al. [6], and a number of published studies that we considered to be "technically compromised" were excluded from the analysis. However, there was insufficient information to do this for other study types, in particular the CA test. Nonetheless, we believe the trends observed here would largely still hold true if the analysis was limited to studies compliant with current guidelines.

2.1. Mouse lymphoma assay (MLA) results

The criteria for conclusions of positive, negative, equivocal or technically compromised have been described in our previous paper [1], as have the reasons why some previously reported MLA results were reassessed in order to fit our categorisation. The carcinogens and non-carcinogens where only the MLA gave a positive result (i.e. either both Ames and MN/CA were negative or results were not available) were identified from our Carcinogenicity Genotoxicity eXperience (CGX) database (as detailed in [1]; it can be viewed at http://www.lhasalimited.org/cgx). This amounted to only 45 chemicals, which is a small number, and analysis of other categories of chemicals in our database may provide more useful information. We re-examined the original reports and recorded the following:

- The highest relative total growth (RTG—i.e. lowest level of toxicity) at which a significant increase in mutant frequency occurred. Increases in MF above concurrent control levels of 90×10^{-6} for the agar method and 126×10^{-6} for the microwell method are considered significant according to the latest recommendations [5]. Most of the published studies in our database used the agar method.
- The maximum increase in MF at any concentration giving acceptable RTG (i.e. RTG >0.1).
- The lowest effective concentration at which a significant increase in MF (as defined above) was observed.

For the MLA-positive chemicals shown in Tables 1 and 2, all except one were negative in the Ames test—for pyrilamine maleate there was no Ames test result. Most were also negative for CA and/or MN but for a small number there was no result for either CA or MN, as noted in the Table. For pyrilamine maleate there was also no result in either CA or MN, so the MLA was the only result. For two carcinogens (*o*-benzyl-*p*-chlorophenol and C.I. Direct Blue 218) the MLA data were referred to by Mitchell et al. [6] in their revised EPA Gene-Tox review, and therefore included in our CGX database, but the original data were not available for review.

Tables 1 and 2 show the highest RTG levels (i.e. lowest levels of toxicity) at which significant positive responses were seen for carcinogens and non-carcinogens, respectively. As can be seen, there were no differences between carcinogens and non-carcinogens in the levels of toxicity at which the minimum positive response was observed. Carcinogens and non-carcinogens are distributed across the whole range of toxicity for the point at which the first positive response is seen. The lowest level of toxicity for a biologically relevant positive response ranged from 23 to 89% (RTG from 0.77 to 0.11) for carcinogens, and from 16 to 91% (RTG 0.84–0.09) for non-carcinogens.

Thus, if the upper limit of toxicity required for a valid assay were lowered, several carcinogens would be missed. Also, sufficient non-carcinogens were positive at low levels of toxicity that the specificity would not be improved. For instance, if the toxicity limit were lowered to 50% (consistent with the chromosomal aberration test), seven non-carcinogens would no longer be positive (see Table 2A) but 15 carcinogens would also no longer be positive (see Table 1A).

Tables 1 and 2 also show the maximum induced MF (IMF) at any concentration that produced acceptable (non-excessive) toxicity (i.e. RTG not lower than 0.1). It should be noted that, at the time most of the MLA studies were performed, there was limited appreciation of the importance of small-colony mutants, and therefore some tests may have achieved conditions that favoured growth of small-colony mutants, whereas others would not. Clearly this could have an impact on

Table 1
Toxicity, maximum mutagenic response and lowest mutagenic concentrations in the MLA for Ames-negative carcinogens

Chemical name	CAS number	RTG for minimum positive response[a]	Maximum induced MF ($\times 10^{-6}$) at any treatment with RTG of 0.1 or higher	Lowest effective concentration (μmol/l) for minimum positive or highest non-significant response[a]	Induction of DNA adducts [reference]
A: Carcinogens positive only in MLA (i.e. negative or equivocal in Ames and CA/MN)					
11-Aminoundecanoic acid	2432-99-7	0.51	149	99.4	No data
Benzofuran	271-89-6	0.22	172	1270	No data
o-Benzyl-*p*-chlorophenol	120-32-1	Data not publicly available			
Butylated hydroxytoluene	128-37-0	0.26	145	45	No data
Chlorobenzilate	510-15-6	0.30	244	277	No data
Chloroform	67-66-3	0.34	92	0.4	+ Calf thymus DNA[b] [9] and hepatocytes [10] *in vitro*
4-Chloro-*o*-toluidine	3165-93-3	0.17	71[c]	848	+ Liver *in vivo* [11]
C.I. Direct blue 218	28407-37-6	Data not publicly available			
Cinnamyl anthranilate	87-29-6	0.24	205	30.8	No data
Dicofol	115-32-2	0.30	36[c]	135	No data
N,N'-Diethyl-2-thiourea	105-55-5	0.42	2307	11344	No data
Ethylbenzene	100-41-4	0.34	529	754	No data
Kepone (AKA chlordecone)	143-50-0	0.38	62[c]	71.3	No data
Malonaldehyde sodium salt	24382-04-5	0.67	239	5553	+ Liver *in vivo* [12]
Methyl *tert*-butyl ether	1634-04-4	0.61	181	22.7	+ Hepatocytes *in vitro*[d] [13]
Piperonyl butoxide	51-03-6	0.17	443	148	No data
Piperonyl sulphoxide	120-62-7	0.70	85[e]	7.7	No data
Toluene	108-88-3	0.21	216	2442	+ HL-60 cells *in vitro* [14]
2,4,6-Trichlorophenol	88-06-2	0.11	144	608	No data
Trimethylthiourea	2489-77-2	0.21	377	35533	No data
4-Vinylcyclohexene	100-40-3	0.40	79[c]	740	+ Skin *in vivo* [15]
B: Carcinogens positive in MLA, negative Ames but no test (or TC) in CA/MN					
1,4-Benzoquinone	106-51-4	0.34	200	3.15	+ Plasmid DNA *in vitro* [16]
Dimethyl methylphosphonate	756-79-6	0.73	54[c]	40.3	No data
FD&C Red 1 (Ponceau 3R)	3564-09-8	0.36	100	3440	No data
Procarbazine HCl (Natulan)	366-70-1	0.77	356	49.7	+ Mammary tissue [17], liver and bone marrow [18] *in vivo*
Thioacetamide	62-55-5	0.69	177	10240	+ Liver *in vivo* [19]
Trichloroacetic acid	76-03-9	0.28	121	12241	+ Liver *in vivo*[b] [20]
Vinylidene chloride	75-35-4	0.70	212	64990	No data
C: Carcinogens positive in MLA but no test (or TC) in Ames and CA/MN					
Pyrilamine maleate	59-33-6	0.55	155	1619	No data

NB: There were no chemicals positive in MLA, negative in CA/MN but no test (or TC) in Ames.

[a] Minimum clear positive response considered where mutant frequency (MF) increased at least 90×10^{-6} (all studies used agar method) over concurrent control as recommended by Moore et al. [5]. In some cases, due to spacing of doses, the lowest significant increase in MF was much greater than the minimum, which would therefore have occurred at a higher RTG and lower concentration.

[b] Oxidative damage to DNA.

[c] Minimum increase in MF not achieved. These chemicals should be reclassified as equivocal rather than positive.

[d] DNA-protein crosslinks.

[e] Although induced MF was $<90 \times 10^{-6}$, the dose response was very steep and it is expected that a significant response would have been achieved at higher concentrations.

Table 2
Toxicity, maximum mutagenic response and lowest effective mutagenic concentrations in the MLA Ames-negative non-carcinogens

Chemical name	CAS number	RTG for minimum positive response[a]	Maximum induced MF ($\times 10^{-6}$) at any treatment with RTG of 0.1 or higher	Lowest effective concentration (μmol/l) for minimum positive or highest non-significant response[a]	Induction of DNA adducts [reference]
A: Non-carcinogens positive only in MLA (i.e. Ames and CA/MN negative or equivocal)					
Aldicarb	116-06-3	0.84	111	21024	No data
Anilazine	101-05-3	0.72	298	10.5	No data
Barium chloride dihydrate	10326-27-9	0.40	27[b]	4094	No data
1,2-Dichlorobenzene	95-50-1	0.81	357	68	+ Calf thymus DNA *in vitro* [21]
FD&C yellow number 6	2783-94-0	0.79	62[b]	8842	No data
4-Hexylresorcinol	136-77-6	0.29	110	154	No data
Hydrochlorothiazide	58-93-5	0.51	264	2519	No data
Isopropyl-*N*-(3-chlorophenyl)carbamate	101-21-3	0.13	68[b]	147	No data
Malaoxon	1634-78-2	0.22	75[b]	477	No data
Oxytetracycline HCl	2058-46-0	0.30	4329	616	No data
Phenyl-β-naphthylamine	135-88-6	0.36	140	59.3	No data
Sodium diethylcarbamate trihydrate	148-18-5	0.55	326	0.39	No data
Sulfisoxazole	127-69-5	0.39	96	3741	No data
2,3,5,6-Tetrachloro-4-nitroanisole	2438-88-2	0.55	151	24.1	No data
B: Non-carcinogens positive in MLA, negative in Ames but no test (or TC) in CA/MN					
Alpha-methyldopa sesquihydrate	41372-08-1	0.27	164	168	No data
Rhodamine 6G	989-38-8	0.09	94	16.7	No data

NB: There were no chemicals positive in MLA, negative in CA/MN but no test (or TC) in Ames.
[a] Minimum clear positive response considered where mutant frequency (MF) increased at least 90×10^{-6} (all studies used agar method) over concurrent control as recommended by Moore et al. [5]. In some cases, due to spacing of doses, the lowest significant increase in MF was much greater than the minimum, which would therefore have occurred at a higher RTG and lower concentration.
[b] Minimum increase in MF not achieved. Chemical should be reclassified as equivocal rather than positive.

the IMF, and highlights some of the difficulties of trying to apply current standards to old studies. Non-carcinogens (Table 2) were just as likely to produce very high IMF as were carcinogens (Table 1), and some carcinogens produced very low IMF. In fact, several carcinogens (as well as four non-carcinogens) failed to achieve the minimum IMF (90×10^{-6} for the agar method) currently recommended for a biologically relevant response [5] and should be re-classified as equivocal. These changes will be made in the web version of the CGX database (http://www.lhasalimited.org/cgx). Thus, if we were to consider increasing the required IMF for a positive response, several carcinogens would be missed.

The lowest concentrations of chemicals at which the minimum positive responses were seen (or in the case of those where minimum IMF was not reached, the highest non-significant concentration) are also shown in Tables 1 and 2. The lowest-effective concentrations clearly vary widely from <1 to >10,000 μM for both carcinogens (Table 1) and non-carcinogens (Table 2). It will be noted that some chemicals are listed where the concentration for a minimum positive response was >10 mM. If tested according to current standards such high concentrations would not be used. Thus five carcinogens and two non-carcinogens would drop out of the analysis if positive results were only included when obtained below 10 mM. However, the range of lowest effective concentrations would still be large, both for carcinogens and non-carcinogens. Thus, changing the maximum concentration requirement for a valid test might improve the specificity of the MLA but it would lower the sensitivity.

We investigated whether formation (or not) of DNA adducts might discriminate between MLA positives of different potency. The *in vitro* and *in vivo* DNA-adduct data that could be found in the literature are also included in Tables 1 and 2. The following comments can be made on these limited data:

- There were nine carcinogens positive only in the MLA that had adduct data, and all formed DNA adducts. Within this sub-group, the level of toxicity varied widely (from 23 to 83% toxicity for a minimum pos-

itive response), the maximum IMF varied from 71 to 356×10^{-6}, and the lowest effective concentration varied over six orders of magnitude from 0.4 to 64,990 μM.

- There was only one non-carcinogen giving an isolated positive result in the MLA for which adduct data could be found, and this compound formed DNA adducts. This was positive at low toxicity (19% toxicity for minimum positive response).

Since all the MLA-positive carcinogens and the one MLA-positive non-carcinogen induced DNA adducts, it is not possible to use these data to draw conclusions regarding DNA reactivity and the other parameters we studied (level of toxicity required, minimum effective concentration or magnitude of response). However, it does not appear that any changes to currently recommended toxicity levels or maximum concentration of test chemical, or requiring a higher induced mutant frequency for a positive "call" would improve the specificity of the MLA without causing serious deterioration in the sensitivity.

2.2. Chromosomal aberration (CA) test results

As mentioned earlier, there were very few published examples where CA results were accompanied by toxicity data. Although one of us (MH) was able to look at the laboratory notebooks of Dr. Ishidate, who published many of the CA results in our previous analysis [1], we

Table 3
Analysis of CA test responses for Ames-negative carcinogens

Chemical name	CAS number	Survival (cell count, confluence or mitotic index as a % of control) for minimum positive response[a]	Maximum % of aberrant cells (excluding gaps) at any treatment with survival 50% or higher	Lowest effective concentration (μmol/l) for minimum positive or highest non-significant response[a]	Induction of DNA adducts [reference]
A: Carcinogens positive only in CA[b] (i.e. negative in Ames and MLA)[c]					
Pentachloronitrobenzene	82-68-8	Not given	19	8.1	No data
Naphthalene	91-20-3	Not given	16	234	No data
Zearalenone	17924-92-4	Not given	13	47.1	+ Kidneys, liver and ovaries *in vivo* [22]
B: Carcinogens positive in CA[b], negative in Ames but no test (or TC) in MLA[d]					
Atrazine	1912-24-9	64	28.7	5.0	− Liver *in vivo* [23]
Heptachlor	76-44-8	Not given	14	67.0	+ Hop plants *in vivo* [24]
Hexanamide	628-02-4	Not given	10	39071	No data
4-Methoxyphenol	150-76-5	Not given	6.8	250	No data
N-Methyloacrylamide	924-42-5	Not given	26	2473	No data
Methylphenidate HCl	298-59-9	Not given	9	3707	No data
Nafenopin	3771-19-5	100	Cannot be established	30.0	No data
Nitrobenzene	98-95-3	Not given	33.2	61.3	+ Liver *in vivo* [25]
C: Carcinogens positive in CA[b] but no test (or TC) in Ames and MLA[e]					
Haloperidol	52-86-8	Not given	Only one concentration used	Only one concentration used	No data
Retinol acetate[f]	127-47-9	Not given	13	200	No data

NB: There were no chemicals positive in CA, negative in MLA but no test (or TC) in Ames.

[a] Minimum positive response was considered as >5% cells with aberrations (excluding gaps) as this would usually be considered a positive response for CHO, CHL cells and human lymphocytes.

[b] Chemicals that were also positive in MN were excluded from these tables.

[c] Another chemical in this category, but for which there were insufficient data in the publications to complete any of the columns, was sodium saccharin.

[d] Other chemicals in this category, but for which there were insufficient data in the publications to complete any of the columns, were aldrin, aniline HCl, asbestos, carboxymethylnitrosourea, clofibrate, lead acetate, methimazole, styrene and 12-*O*-tetradecanoylphorbol 13-acetate.

[e] Other chemicals in this category, but for which there were insufficient data in the publications to complete any of the columns, were 1-amyl-1-nitrosourea, sodium barbital, manganese ethylenebisthiocarbamate and petasitenine.

[f] Equivocal in MN.

were only able to find toxicity data on very few additional clastogens. Therefore, it was not possible to analyse (as we did above for the MLA) whether any change to the required level of toxicity (currently at least 50% according to OECD [3]) would provide improved specificity whilst retaining sensitivity.

For a small number of chemicals that were positive only in this test, we could identify the maximum induced CA response (although we could not know whether this occurred at levels of toxicity much higher than 50%) and the lowest effective concentration required for a minimum positive response. For the latter, since most published data were from Chinese hamster cells, we accepted that CA frequencies (excluding gaps) of >5% would probably be statistically significant and exceed historical control ranges. These data are tabulated in Tables 3 and 4 for carcinogens and non-carcinogens, respectively. Unfortunately there were a number of chemicals for which the published reports contained insufficient details to establish the maximum frequency of aberrant cells or the minimum concentration for a positive effect.

From Tables 3 and 4 it can be seen that the maximum percentage of aberrant cells ranged from 9 to 33% for carcinogens and from 7 to 23% for non-carcinogens. The minimum positive concentration ranged from 5 to 39,000 μM for carcinogens and from 24 to 5000 μM for non-carcinogens. In other words, for this small sample of Ames-negative clastogens there were no clear differences in maximum response or minimum positive concentration between carcinogens and non-carcinogens. It will be noted that one chemical (hexanamide, Table 3) is listed where the concentration for a minimum positive response was >10 mM. If tested according to current standards such high concentrations would not be used. Thus this one carcinogen would drop out of the analysis if positive results were only included when obtained below 10 mM. However, the range of lowest effective concentrations would still be large, both for carcinogens and non-carcinogens. Thus it does not appear that any change to the maximum concentration, or the introduction of a requirement for the induced frequency of aberrant cells to reach a certain level would improve the specificity of the CA test without causing serious deterioration in the sensitivity.

As for the analysis of MLA data, we investigated whether DNA reactivity may correlate better with potency. Any *in vitro* or *in vivo* DNA-adduct data that could be found in the literature for these chemicals positive only in the CA test are shown in Tables 3 and 4.

Table 4
Analysis of CA test responses for Ames-negative non-carcinogens

Chemical name	CAS number	Survival (cell count, confluence or mitotic index as a % of control) for minimum positive response[a]	Maximum % of aberrant cells (excluding gaps) at any treatment with survival 50% or higher	Lowest effective concentration (μmol/l) for minimum positive or highest non-significant response[a]	Induction of DNA adducts [reference]
A: Non-carcinogens positive only in CA[b] (i.e. Ames and MLA negative)[c]					
Diphenhydramine HCl	147-24-0	Not given	17	343	No data
4,4′-Isopropylidenediphenol	80-05-7	45	23	399	+ Liver *in vivo* [26] and Syrian hamster cells *in vitro* [27]
Tin (II) chloride	7772-99-8	Not given	22	132	No data
B: Non-carcinogens positive in CA, negative in Ames but no test (or TC) in MLA[d]					
Carbromal	77-65-6	Not given	7	5061	No data
Chlorpheniramine maleate	113-92-8	Not given	18	1279	No data
Fenvalerate	51630-58-1	Not given	20	23.8	No data
Hexachlorocyclopentadiene	77-47-4	Not given	9.5	27.5	No data

NB: There were no chemicals positive in CA, negative in MLA but no test (or TC) in Ames.

[a] Minimum positive response was considered as >5% cells with aberrations (excluding gaps) as this would usually be considered a positive response for CHO, CHL cells and human lymphocytes.

[b] Chemicals that were also positive in MN were excluded from these tables.

[c] Other chemicals in this category, but for which there were insufficient data in the publications to complete any of the columns, were acetohexamide, benzoin, FD&C Red number 3 and tetracycline HCl.

[d] Other chemicals in this category, but for which there were insufficient data in the publications to complete any of the columns, were sodium benzoate, caffeine, chlorpropamide, FD&C Yellow number 5 and vinyl toluene.

Adduct data were found for only a small number of chemicals. However, the following comments can be made:

- Of the carcinogens that were positive only for CA, there were only four with adduct data, and one of these (atrazine) did not form DNA adducts. Whereas we might have expected DNA-reactive chemicals to be positive at lower concentrations, the non-DNA-reactive atrazine was positive at the lowest concentration of all the 13 carcinogens studied in this set.
- Of the carcinogens with adduct data the maximum CA response was quite variable (13–33%) as was the lowest effective concentration (5–67 μM).
- Only one non-carcinogen had adduct data (which was positive).

Since all except one CA-positive carcinogen and the one CA-positive non-carcinogen induced DNA adducts, it is not possible to draw conclusions regarding DNA reactivity and the other parameters we studied (magnitude of response, or lowest effective concentration). Thus it does not appear that the specificity of the CA test can be improved, without serious deterioration in sensitivity, by changing the maximum required test concentration or setting a limit for magnitude of response.

3. Tumour profiles

We reviewed the tumour profiles as detailed in the Carcinogenic Potency DataBase (CPDB) of Gold [7] for those carcinogens giving positive results across all three *in vitro* tests (Table 5) or giving negative results across all three *in vitro* tests (Table 6). The prevalence of the trans-species, trans-sex and multi-site carcinogens amongst the two categories of results we examined (positive or negative across all three *in vitro* tests) is shown in Table 7A. The prevalence of chemicals showing "all positive" or "all negative" amongst the single-species, single-sex, single-site carcinogens is shown in Table 7B. The following conclusions can be drawn from these analyses:

- There are no differences between "all positive for genotoxicity" and "all negative for genotoxicity" amongst the carcinogens as far as inducing tumours in both males and females. In other words, trans-sex carcinogens are equally likely to be genotoxic in all *in vitro* tests as to be non-genotoxic in all *in vitro* tests.
- Trans-species carcinogens are 39% more prevalent in the "all positive for genotoxicity" group than in the "all negative for genotoxicity" group.
- Multi-site carcinogens are 31% more prevalent in the "all positive for genotoxicity" group than in the "all negative for genotoxicity" group.
- Carcinogens showing the combination of trans-species, trans-sex and multi-site tumour profile are much more prevalent (70% more) in the "all positive for genotoxicity" group than in the "all negative for genotoxicity" group.
- Single-species, single-sex, single-site carcinogens are not very prevalent even amongst those chemicals giving three negative results *in vitro*.

We also reviewed the tumour profiles for carcinogens giving positive results in only one of the *in vitro* genotoxicity tests whilst being negative in the other two tests in the battery. The results are shown in Table 8A for Ames-positive only, Table 8B for MLA-positive only, and Table 8C for MN- or CA-positive only. The prevalence of trans-species, trans-sex or multi-site carcinogens amongst these compounds positive in a single genotoxicity test is shown in Table 9. The following conclusions can be drawn:

- Multi-site carcinogens are surprisingly prevalent amongst those chemicals inducing only single positive genotoxicity results.
- Multi-site carcinogens are more prevalent amongst Ames-positive genotoxins, but the database is very small.
- The prevalence of trans-species or trans-sex or multi-site carcinogens amongst single genotoxicity test positive compounds is intermediate between those in "all positive" and "all negative" groups.

4. Relative predictivity for batteries with mixed positive and negative results

Because we did not observe any obvious changes to protocols for the mammalian cell tests that would improve specificity without impairing sensitivity, we revisited the topic of relative predictivity described in our previous paper [1] and we calculated relative predictivity (RP) for rodent carcinogenicity from different combinations of test results. As described previously, positive RP for predicting carcinogenicity was acceptable (>2) only for the Ames test of the single tests, but was most informative when all three tests were positive. We also calculated the RP for non-carcinogenicity from one, two or three *in vitro* tests all giving negative results. Only a few combinations of all negative results gave acceptable RP for non-carcinogenicity.

Table 5
Tumour profiles for carcinogens (tested in rats and mice, males and females) positive in Ames plus MLA plus either MN or CA

Carcinogen	CAS number	Tumour profile		
		Rats and mice	Males and females	Multi-site
N-Acetoxy-2-acetylaminofluorene	6098-44-8	Yes	Yes	Yes
2-Acetylaminofluorene	53-96-3	Yes	Yes	Yes
2-Aminoanthracene[a]	613-13-8	Yes	Yes	Yes
2-Aminofluorene[a]	153-78-6	Yes	Yes	Yes
2-Amino-4-nitrophenol	99-57-0	No	No	No
4-Amino-2-nitrophenol	119-34-6	No	No	No
2-Amino-5-nitrothiazole	121-66-4	No	No	Yes
5-Azacytidine	320-67-2	Yes	Yes	Yes
Benzo[a]pyrene	50-32-8	Yes	Yes	No[b]
Benzyl chloride	100-44-7	Yes	Yes	Yes
bis(2-Chloro-1-methylethyl)ether, technical grade	108-60-1	No	Yes	Yes
Calcium chromate	13765-19-0	Yes	Yes	Yes
Captan	133-06-2	Yes	Yes	Yes
Chlorodibromomethane	124-48-1	No	Yes	No
3-(Chloromethyl)pyridine HCl	6959-48-4	Yes	Yes	No
C.I. Disperse blue 1	2475-45-8	No	Yes	No
C.I. Disperse orange 2 (1-amino-2-methyl-anthrquinone)	82-28-0	Yes	Yes	Yes
Cyclophosphamide monohydrate	6055-19-2	Yes	Yes	Yes
Cytembena	21739-91-3	No	Yes	Yes
2,4-Diaminoanisole sulphate	39156-41-7	Yes	Yes	Yes
2,4-Diaminotoluene	95-80-7	Yes	Yes	Yes
1,2-Dibromo-3-chloropropane	96-12-8	Yes	Yes	Yes
1,2-Dibromoethane	106-93-4	Yes	Yes	Yes
Dichloroacetic acid	79-43-6	Yes	Yes	No
Dichloromethane	75-09-2	Yes	Yes	Yes
2,6-Dichloro-*p*-phenylenediamine	609-20-1	No	Yes	No
1,2-Dichloropropane	78-87-5	No	Yes	No
Dichlorvos	62-73-7	Yes	Yes	Yes
Diglycidyl resorcinol ether, technical grade	101-90-6	Yes	Yes	No
3,3′-Dimethylbenzidine	119-93-7	Yes	Yes	Yes
Dimethyl hydrogen phosphite	868-85-9	No	No	Yes
Epichlorohydrin	106-89-8	No	Yes	No
1,2-Epoxybutane	106-88-7	No	No	Yes
Ethyl methanesulphonate	62-50-0	Yes	Yes	Yes
Formaldehyde	50-00-0	Yes	Yes	Yes
Furylfuramide (AF-2)	3688-53-7	Yes	Yes	No[b]
Glycidol	556-52-5	Yes	Yes	Yes
Hydrazine sulphate	10034-93-2	Yes	Yes	Yes
Methylazoxymethanol acetate	592-62-1	No	Yes	Yes
4,4′-Methylenedianiline 2HCl	13552-44-8	Yes	Yes	Yes
N-Methyl-*N*′-nitro-*N*-nitrosoguanidine	70-25-7	Yes	Yes	Yes
2-Naphthylamine	91-59-8	Yes	Yes	No[b]
o-Nitroanisole	91-23-6	Yes	Yes	Yes
5-Nitro-2-furaldehyde semicarbazone (AKA Nitrofurazone)	59-87-0	Yes	No	Yes
1-[(5-Nitrofurfurylidene)amino]hydantoin (AKA Nitrofurantoin)	67-20-9	Yes	No	No[b]
2-Nitro-*p*-phenylenediamine	5307-14-2	No	No	No
4-Nitroquinoline-*N*-oxide	56-57-5	Yes	Yes	Yes
N-Nitrosodimethylamine (dimethylnitrosamine)	62-75-9	Yes	Yes	Yes
p-Nitrosodiphenylamine	156-10-5	Yes	No	No
4,4′-Oxydianiline	101-80-4	Yes	Yes	Yes
Phenobarbital	50-06-6	No	Yes	No
Beta-Propiolactone	57-57-8	Yes	Yes	No
1,2-Propylene oxide	75-56-9	Yes	Yes	Yes
Quercetin	117-39-5	No	Yes	Yes
p-Quinone dioxime	105-11-3	No	No	No
Selenium sulphide	7446-34-6	Yes	Yes	Yes
Styrene oxide	96-09-3	Yes	Yes	No
1,2,3-Trichloropropane	96-18-4	Yes	Yes	Yes
Zinc dimethyldithiocarbamate (Ziram)	137-30-4	No	Yes	Yes

NT: not tested in that species or sex.
[a] Carcinogens positive in Ames + MLA + MN but not tested or equivocal in CA. All others positive in Ames + MLA + CA.
[b] Single-site tumours only within a species. However, different tumours may have arisen in rats and mice.

Table 6
Tumour profiles for carcinogens negative in Ames plus MLA plus either MN or CA

Carcinogen	CAS number	Tumour profile		
		Rats and mice	Males and females	Multi-site
3-Amino-1,2,4-triazole (Amitrole)	61-82-5	Yes	Yes	Yes
tert-Butyl alcohol	75-65-0	Yes	No[a]	No[b]
5-Chloro-*o*-toluidine	95-79-4	No	Yes	Yes
Decabromodiphenyl oxide	1163-19-5	No	Yes	No
Diethanolamine	111-42-2	No	Yes	Yes
Di(2-ethylhexyl)phthalate	117-81-7	Yes	Yes	No
1,4-Dioxane	123-91-1	Yes	Yes	Yes
DL-Ethionine	67-21-0	Yes	Yes	No
Melamine	108-78-1	No	No	No
Methyl carbamate	598-55-0	No	Yes	No
Nitrilotriacetic acid, trisodium salt, monohydrate	18662-53-8	No	Yes	Yes
N-Nitrosodiphenylamine	86-30-6	No	Yes	No
Progesterone	57-83-0	Yes	Yes	Yes
Pyridine	110-86-1	Yes	Yes	No[b]
Reserpine	50-55-5	Yes	Yes	Yes
2,3,7,8-Tetrachlorodibenzo-*p*-dioxin	1746-01-6	Yes	Yes	Yes
1,1,2,2-Tetrachloroethane	79-34-5	No	Yes	No
Tris(2-ethylhexyl)phosphate	78-42-2	No	No	No

By inhalation.
[a] Tumours in male rats but female mice.
[b] Single-site tumours only within a species. However, different tumours may have arisen in rats and mice.

Table 7A
Prevalence of trans-species, trans-sex or multi-site carcinogens amongst consistent positive or negative genotoxicity responses

	Trans-species carcinogens (%)	Trans-sex carcinogens (%)	Multi-site carcinogens (%)[a]	Trans-species, multiple site and trans-sex carcinogens (%)
Positive in Ames plus MLA plus either MN or CA	69.5 (41/59)	83.1 (49/59)	72.9 (43/59)	56 (33/59)
Negative in Ames plus MLA plus either MN or CA	50.0 (9/18)	83.3 (15/18)	55.6 (10/18)	33 (6/18)

[a] Including those carcinogens producing single-site tumours within one species but tumours at different sites in different species.

Table 7B
Prevalence of single-species, single-sex and single-site carcinogens amongst consistent positive or negative genotoxicity responses

	Single-species, single-sex, single-site carcinogens (%)
Positive in Ames plus MLA plus either MN or CA	6.8 (4/59)
Negative in Ames plus MLA plus either MN or CA	11.1 (2/18)

Table 8A
Tumour profile for carcinogens positive in Ames but negative in MLA, MN or CA

Carcinogen	CAS number	Tumour profile		
		Rats and mice	Males and females	Multi-site
D&C red 9	5160-02-1	No	No	Yes
Trifluralin, technical grade	1582-09-8	No	No	Yes
Urethane[a]	51-79-6	Yes	Yes	Yes

[a] There are a number of reports that urethane is negative in the Ames test. As noted in our previous paper [1] we decided to accept the deliberations of Zeiger [28] for this test with urethane.

Table 8B
Tumour profile for carcinogens (with results in rats and mice, males and females) positive in MLA but negative in Ames, MN or CA

Carcinogen	CAS number	Tumour profile		
		Rats and mice	Males and females	Multi-site
11-Aminoundecanoic acid	2432-99-7	No	No	Yes
Benzofuran	271-89-6	Yes	Yes	Yes
o-Benzyl-*p*-chlorophenol	120-32-1	No	No	No
Butylated hydroxytoluene	128-37-0	No	No	Yes
Chlorobenzilate	510-15-6	No	Yes	No
Chloroform	67-66-3	Yes	Yes	Yes
C.I. Direct blue 218	28407-37-6	Yes	Yes	No[a]
Cinnamyl anthranilate	87-29-6	Yes	Yes	Yes
Dicofol	115-32-2	No	No	No
N,N'-Diethyl-2-thiourea	105-55-5	No	Yes	No
Ethylbenzene	100-41-4	Yes	Yes	Yes
Kepone (AKA Chlordecone)	143-50-0	Yes	Yes	No
Malonaldehyde sodium salt	24382-04-5	Yes	Yes	Yes
Methyl *tert*-butyl ether	1634-04-4	Yes	Yes	Yes
Piperonyl butoxide	51-03-6	Yes	Yes	No
Piperonyl sulphoxide	120-62-7	No	No	No
Toluene	108-88-3	No	Yes	Yes
Trichloroethylene (with and without epichlorhydrin)	79-01-6	Yes	Yes	Yes
2,4,6-Trichlorophenol;	88-06-2	Yes	Yes	No[a]
Trimethylthiourea	2489-77-2	No	No	No

[a] Single-site tumours only within a species. However, different tumours may have arisen in rats and mice.

Table 8C
Tumour profile for carcinogens (with results in rats and mice, males and females) positive in MN or CA but negative in Ames and MLA

Carcinogen	CAS number	Tumour profile		
		Rats and mice	Males and females	Multi-site
DDT	50-29-3	Yes	Yes	Yes
17-β-Estradiol	50-28-2	Yes	Yes	Yes
Nitrilotriacetic acid	139-13-9	Yes	Yes	Yes
Oxazepam	604-75-1	No	Yes	Yes
Pentachloronitrobenzene	82-68-8	No	No	No
Saccharin, sodium[a]	128-44-9	No	No	No
Zeralenone	17924-92-4	No	Yes	Yes

[a] Sodium saccharin induces bladder tumours via chronic irritation from the formation of precipitates at high doses, and the CA positive response was only obtained at concentrations >10 mM, so it is unlikely these effects are connected.

Table 9
Prevalence of trans-species, trans-sex or multi-site carcinogens amongst single-positive genotoxicity responses

	Trans-species carcinogens (%)	Trans-sex carcinogens (%)	Multi-site carcinogens (%)[a]	Trans-species, multiple site and trans-sex carcinogens (%)
Positive Ames only	33.3 (1/3)	33.3 (1/3)	100 (3/3)	33.3 (1/3)
Positive MLA only	55.0 (11/20)	70.0 (14/20)	60.0 (12/20)	45.0 (9/20)
Positive MN or CA only	42.9 (3/7)	71.4 (5/7)	71.4 (5/7)	42.9 (3/7)
Positive in any single test but Negative in the other two	50.0 (15/30)	66.7 (20/30)	66.7 (20/30)	43.3 (13/30)

[a] Including those carcinogens inducing single-site tumours within one species, but inducing tumours at different sites in different species.

Table 10A
Numbers of carcinogens and non-carcinogens with 2/3 results positive or negative[a]

	Ames plus MLA plus MN (%)	Ames plus MLA plus CA (%)
Carcinogens		
2/3 Tests positive	16/50 (32.0)	51/176 (29.0)
2/3 Tests negative	6/50 (12.0)	30/176 (17.0)
Non-carcinogens		
2/3 Tests positive	8/12 (66.7)	22/76 (28.9)
2/3 Tests negative	0/12 (0)	22/76 (28.9)

[a] E and TC results excluded from analysis (counted as blank = no result).

Table 10B
Relative predictivity of carcinogen/non-carcinogen status when 2/3 results indicative (data from Table 10A)

Combination of three tests	Chemicals positive in 2/3 tests (%)		Relative predictivity that 2/3 positive results indicate a carcinogen A/B	Chemicals negative in 2/3 tests (%)		Relative predictivity that 2/3 negative results indicate a non-carcinogen C/D
	Carcinogens (A)	Non-carcinogens (B)		Non-carcinogens (C)	Carcinogens (D)	
Ames plus MLA plus MN	32.0	66.7	0.48	0	12.0	0
Ames plus MLA plus CA	29.0	28.9	1.00	28.9	17.0	1.70

As a battery of three *in vitro* tests most often gives a mixture of positive and negative results, we decided to calculate RP for predicting carcinogenicity when two tests were positive but one negative, and to calculate RP for predicting non-carcinogenicity when two tests were negative but one positive. The numbers of results falling into each category are shown in Table 10A and the RP values in Table 10B. It can be seen that in all cases the RP values are less than the meaningful value of two, and in some cases are <1, which means there would be a better chance of predicting the outcome of the carcinogenicity studies by flipping a coin.

As all of the RP values were worse for 2/3 tests than for 3/3 tests, we decided not to analyse further the RP from 1/3 tests.

5. Discussion and conclusions

Although limited data were available, our analyses of chemicals giving isolated positive genotoxicity results in one of the mammalian cell tests, suggest that revising the requirements for toxicity limits or highest test concentration, or to apply some "threshold" response that must be exceeded, would improve the poor specificity of the MLA and CA tests, but would at the same time significantly impair sensitivity. An examination of DNA reactivity within these small groups did not suggest any associations between DNA reactivity and level of toxicity or concentration needed for a minimum positive response, or the magnitude of the genotoxic response. It is possible that if we examine all MLA and CA positives, and view separately those carcinogens known to interact with DNA, and/or known to have structural alerts, we may find that the "profile" of the genotoxicity results (concentration, level of toxicity, magnitude of response) would be more useful in predicting carcinogenic outcome. This is currently being explored and will be the subject of a future manuscript.

An analysis of tumour profiles indicated that trans-species, trans-sex and multi-site carcinogens are much more prevalent amongst those chemicals giving positive results in all three *in vitro* genotoxicity tests than amongst those giving negative results in all three tests. This is consistent with the relative predictivity (RP) analysis we reported previously [1] where chemicals giving positive results in all three *in vitro* tests would be at least 3× more likely to be rodent carcinogens than non-carcinogens. However, unexpectedly, single-species, single-sex and single-site carcinogens were not prevalent amongst those chemicals giving negative results in all three *in vitro* tests. This suggests that non-genotoxic carcinogens are just as likely to produce single-site tumours in one sex of one species as to produce multi-site tumours in both sexes of rats and mice. It was also surprising that multi-site carcinogens were highly prevalent amongst those chemicals positive in one but negative in the other two genotoxicity tests. Multi-site carcinogens were particular prevalent amongst Ames-positive chemicals, but the

database was small and further examples are needed for a rigorous evaluation. Thus, it appears that the profile of genotoxicity results cannot be used with any confidence to predict the tumour profile in subsequent carcinogenicity studies. This is consistent with a more modern view that mutational events and chromosomal rearrangements are, although required for tumourigenesis, not a major driving factor for prevalence and incidence of tumours in lifetime rodent carcinogenicity tests with chemical compounds [8]. In such long-term experiments, hormonal disturbances, genetic strain background, general tumour promotional effects, etc. are very important. Such effects may equally be exerted by many of the investigated chemicals that also possess genotoxic properties.

Although our previous RP analysis [1] indicated that chemicals giving positive results in all three *in vitro* tests were >3× more likely to be rodent carcinogens than non-carcinogens, and that chemicals giving all negative results were >2× more likely to be non-carcinogens than carcinogens, in industry mixed positive and negative results are most common. As might be expected, RP analysis of "mixed" genotoxicity results revealed that these cannot be used to predict carcinogenic outcome. It should be noted that predictivity analyses are dependent on the choice of compounds analysed and influenced by the prevalence of carcinogens and non-carcinogens in the data set. Different predictivities from those presented here could be obtained with different numbers of carcinogens and non-carcinogens in the data set. Also, only some of the mechanisms (e.g. genotoxic) giving rise to cancer are the same as those leading to mutations and chromosomal aberrations, and non-genotoxic mechanisms would not be expected to be predicted by genotoxicity assays. Therefore, the predictivity of genotoxicity tests for cancer will depend on the mechanism.

As indicated previously [1], the disappointing findings from these analyses suggest it may be time for a complete rethink with regard to *in vitro* genotoxicity testing. It does not appear that a redefinition of the conditions (cytotoxicity, solubility, etc.) for *in vitro* testing will improve the low specificity for prediction of rodent tumourigenesis, and therefore new, more robust assays may be needed. We should also consider, however, that the rodent bioassay is not relevant for predicting human carcinogenicity in many cases, and therefore new, more robust carcinogenicity assays may also be needed.

Regarding the availability of complete data sets, or of critical information (e.g. cytotoxicity) related to single-assay results, the authors are astonished that only small subsets of chemicals have full data sets. Hence, despite more than 30 years of genotoxicity testing, the published literature still contains large data gaps. These gaps certainly do not facilitate far-reaching and important conclusions on correlations between important parameters. Thus in many cases our only recourse is to some kind of patchwork analysis when working with published data. As expected, most adduct data have been reported for the carcinogens and there are very few studies of adducts formed by non-carcinogens.

References

[1] D. Kirkland, M. Aardema, L. Henderson, L. Müller, Evaluation of the ability of a battery of three *in vitro* genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. I. Sensitivity, specificity and relative predictivity, Mutat. Res. 584 (2005) 1–256.

[2] ICH Topic S2A, Genotoxicity: guidance on specific aspects of regulatory genotoxicity tests for pharmaceuticals. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, Harmonised tripartite guideline CPMP/ICH/141/95, approved September 1995.

[3] OECD, Guidelines for the Testing of Chemicals: Genotoxicity. Organisation for Economic Co-operation and Development, Paris, Revised and new guidelines adopted, 1997.

[4] NTP website at http://ntp-server.niehs.nih.gov.

[5] M.M. Moore, M. Honma, J. Clements, G. Bolcsfoldi, B. Burlinson, M. Cifone, J. Clarke, R. Delongchamp, R. Durward, M. Fellows, B. Gollapudi, S. Hou, P. Jenkinson, M. Lloyd, J. Majeska, B. Myhr, M. O'Donovan, T. Omori, C. Riach, R. San, L.F. Stankowski Jr., A.K. Thakur, F. Van Goethem, S. Wakuri, I. Yoshimura, Mouse lymphoma thymidine kinase gene mutation assay: follow-up meeting of the international workshop on genotoxicity tests—Aberdeen, Scotland, 2003, Environ. Mol. Mutagen. 47 (2006) 1–5.

[6] A.D. Mitchell, A.E. Auletta, D. Clive, P.E. Kirby, M.M. Moore, B.C. Myhr, The L5178Y/$tk^{+/-}$ mouse lymphoma specific gene and chromosomal mutation assay. A phase III report of the U.S. Environmental protection Agency Gene-Tox Program, Mutat. Res. 394 (1997) 177–303.

[7] L.S. Gold, The carcinogenic potency project. http://potency.berkeley.edu/cpdb.html (2004).

[8] R.A. Beckman, L.A. Loeb, Genetic instability in cancer: theory and experiment, Semin. Cancer Biol. 15 (2005) 423–435.

[9] A.B. Direnzo, A.J. Gandolfi, I.G. Sipes, Microsomal bioactivation and covalent binding of aliphatic halides to DNA, Toxicol. Lett. 11 (1982) 243–252.

[10] E.J. Beddowes, S.P. Faux, J.K. Chipman, Chloroform, carbon tetrachloride and glutathione depletion induce secondary genotoxicity in liver cells via oxidative stress, Toxicology 187 (2003) 101–115.

[11] P. Bentley, F. Waechter, F. Bieri, W. Staubli, W. Muecke, Species differences in the covalent binding of *p*-chloro-*o*-toluidine to DNA, Arch. Toxicol. Suppl. 9 (1986) 163–166.

[12] A. Kautiainen, C.E. Vaca, F. Granath, Studies on the relationship between hemoglobin and DNA adducts of malonaldehyde and their stability *in vivo*, Carcinogenesis 14 (1993) 705–708.

[13] M. Casanova, H.A. Heck, Lack of evidence for the involvement of formaldehyde in the hepatocarcinogenicity of methyl *tertiary*-butyl ether in CD-1 mice, Chem. Biol. Interact. 105 (1997) 131–143.

[14] N.W. Gaikwad, W.J. Bodell, Formation of DNA adducts in HL-60 cells treated with the toluene metabolite *p*-cresol: a potential biomarker for toluene exposure, Chem. Biol. Interact. 145 (2003) 149–158.

[15] N. Mabon, K. Randerath, ^{32}P-postlabeling of 1,3-butadiene and 4-vinyl-1-cyclohexene metabolite-DNA adducts: *in vitro* and *in vivo* applications, Toxicology 113 (1996) 341–344.

[16] M. Gaskell, K.I. McLuckie, P.B. Farmer, Genotoxicity of benzene metabolites *para*-benzoquinone and hydroquinone, Chem. Biol. Interact. 153/154 (2005) 267–270.

[17] L.Y. Fong, D.E. Jensen, P.N. Magee, DNA methyl-adduct dosimetry and O^6-alkylguanine-DNA alkyl transferase activity determinations in rat mammary carcinogenesis by procarbazine and *N*-methylnitrosourea, Carcinogenesis 11 (1990) 411–417.

[18] V. Plesta, C. Valavanis, J.H. van Delft, M.J. Steenwinkel, S.A. Krytopoulos, DNA damage and mutagenesis induced by procarbazine in lambda lacZ transgenic mice: evidence that bone marrow mutations do not arise primarily through miscoding by O^6-methylguanine, Carcinogenesis 18 (1997) 2191–2196.

[19] A.E. Pegg, M.E. Dolan, D. Scicchitano, K. Morimoto, Studies of the repair of O^6-alkylguanine and O^4-alkylthymine in DNA by alkyltransferases from mammalian cells and bacteria, Environ. Health Perspect. 62 (1985) 109–114.

[20] E.W. Austin, J.M. Parrish, D.H. Kindler, R.J. Bull, Lipid peroxidation and formation of 8-hydroxydeoxyguanosine from acute doses of halogenated acetic acids, Fund. Appl. Toxicol. 31 (1996) 77–82.

[21] M. Paolini, L. Pozzetti, P. Silingardi, C. Della Croce, G. Bronzetti, G. Cantelli-Forti, Isolation of a novel metabolising system enriched in phase-II enzymes for short-term genotoxicity assays, Mutat. Res. 413 (1998) 205–217.

[22] A. Pfohl-Leszkowicz, L. Chekir-Ghedira, H. Bacha, Genotoxicity of zearalenone, an estrogenic mycotoxin: DNA adduct formation in female mouse tissues, Carcinogenesis 16 (1995) 2315–2320.

[23] S.E. George, R.W. Chadwick, M.J. Kohan, J.C. Allison, S.H. Warren, R.W. Williams, Atrazine treatment potentiates excretion of mutagenic urine in 2,6-dinitrotoluene-treated Fischer 344 rats, Environ. Mol. Mutagen. 26 (1995) 178–184.

[24] J.A. Laouedj, C. Schenk, A. Pfohl-Leszkowicz, G. Keith, D. Schontz, P. Guillemaut, B. Rether, Detection of DNA adducts in declining hop plants grown on fields formerly treated with heptachlor, a persistent insecticide, Environ. Pollut. 90 (1995) 409–414.

[25] H. Li, H. Wang, H. Sun, Y. Liu, K. Liu, S. Peng, Binding of nitrobenzene to hepatic DNA and hemoglobin at low doses in mice, Toxicol. Lett. 139 (2003) 25–32.

[26] A. Atkinson, D. Roy, *In vivo* DNA adduct formation by Bisphenol A, Environ. Mol. Mutagen. 26 (1995) 60–66.

[27] T. Tsutsui, Y. Tamura, E. Yagi, K. Hasegawa, M. Takahashi, N. Maizumi, F. Yamaguchi, J.C. Barrett, Bisphenol-A induces cellular transformation, aneuploidy and DNA adduct formation in cultured Syrian hamster embryo cells, Int. J. Cancer 75 (1998) 290–294.

[28] E. Zeiger, Genotoxicity database, in: L.S. Gold, E. Zeiger (Eds.), Handbook of Carcinogenic Potency and Genotoxicity Databases, CRC Press Inc., Boca Raton, 1997, pp. 687–729.

Genes and Environment, Vol. 28, No. 4 pp. 141-152 (2006)

Commentary

Globally Harmonized System on Hazard Classification and Labeling of Chemicals and Other Existing Classification Systems for Germ Cell Mutagens

Takeshi Morita[1,3], Makoto Hayashi[2] and Kaoru Morikawa[1]

[1]Division of Safety Information on Drug, Food and Chemicals,
[2]Division of Genetics and Mutagenesis, National Institute of Health Sciences, Tokyo, Japan

(Received August 8, 2006; Revised August 26, 2006; Accepted August 28, 2006)

The Globally Harmonized System (GHS) on hazard classification and labeling of chemicals will be implemented globally by 2008. The GHS includes (a) harmonized criteria for classifying chemicals and chemical mixtures according to their health, environmental and physical hazards, and (b) harmonized hazard communication elements, including requirements for labeling and safety data sheets. Germ cell mutagenicity is included in the GHS health hazard classes in addition to carcinogenicity. This means increased significance for then results of genetic toxicology testing for the classification of chemicals. GHS requires the classification of chemicals if they are germ cell mutagens (categories 1A, 1B and 2) or not. Several classification systems for germ cell mutagens have been proposed in the EU, Germany, US, Canada, in advance of the adoption of the GHS. In this paper, these classification systems including GHS are introduced and summarized to provide the basis of the hazard classification of germ cell mutagens. Though the objectives, target audiences and criteria of these classification systems are different, the GHS will become standard for hazard classification. Hazard classification is a significant first step in risk communication. Further development of risk evaluation criteria and communication on germ cell mutagens is expected.

Key words: GHS, hazard classification, germ cell mutagenicity, germ cell mutagens

Introduction

The Globally Harmonized System (GHS) of classification and labeling of chemicals is a single, globally harmonized system to address classification of chemicals, labelling, and safety data sheets, which has been developed by the United Nations (UN). The GHS document has been prepared and published by the secretariat of the United Nations Economic Commission for Europe (1,2). The GHS covers all hazardous chemicals except for pharmaceuticals, food additives, cosmetics, and pesticide residues in food in terms of labeling at the point of intentional intake. The GHS is based on currently available data and thus compliance with these criteria will not require retesting of chemicals for which acceptable test data already exists.

The goal of the GHS is to identify the intrinsic hazards found in chemicals and chemical mixtures and to convey the information about these hazards to the target audiences including consumers, workers, transport workers, and emergency responders. The World Summit on Sustainable Development in Johannesburg on 4th September, 2002 encouraged countries to implement the GHS as soon as possible with a view to having the system fully operational by 2008. In Japan, an inter-ministerial committee was organized in 2001 to share information about the GHS among ministries and to play a pivotal role in the UN-Subcommittee. Seven government offices, *i.e.*, Ministry of Health, Labour and Welfare (MHLW), Ministry of Economy, Trade and Industry (METI), Ministry of the Environment (MOE), Ministry of Internal Affairs and Communications (MIC), Ministry of Agriculture, Forestry and Fisheries (MAFF), Ministries of Land Infrastructure and Transport (MLIT) and Ministry of Foreign Affairs (MOFA), and experts from national laboratories and industries participated in the committee. The committee's activities include: (i) translation of the GHS into Japanese (3); (ii) information sharing among ministries with respect to the relevant domestic laws; (iii) classification of chemicals under each relevant domestic law (4); and (iv) deliberation on the agenda items and documents of the UN Sub-Committee meetings and decision-making about the Japanese position.

The GHS includes the following two elements: harmonized criteria for classifying substances and mixtures according to their health (10 hazard classes),

[3]Correspondence to: Takeshi Morita, Division of Safety Information on Drug, Food and Chemicals, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. Tel: +81-3-3700-1141, Fax: +81-3-3700-1483, E-mail: morita-tk@nihs.go.jp

© The Japanese Environmental Mutagen Society

environmental (1 hazard class) and physical hazards (16 hazard classes); and harmonized hazard communication elements, including requirements for labeling and safety data sheets. The GHS requests classification of the following 10 hazard classes in health hazard, *i.e.*, (i) acute toxicity, (ii) skin corrosion/irritation, (iii) serious eye damage/eye irritation, (iv) respiratory or skin sensitization, (v) germ cell mutagenicity, (vi) carcinogenicity, (vii) reproductive toxicity, (viii) specific target organ systemic toxicity—single exposure, (ix) specific target organ systemic toxicity—repeated exposure, and (x) aspiration hazard.

It is striking that not mutagenicity per se but germ cell mutagenicity specifically is included in GHS health hazard class in addition to carcinogenicity. The GHS focuses on heritable effects by mutagens. Several classification systems of mutagens or germ cell mutagens have been proposed from European Union (EU), Germany, United States of America (US), Canada, etc. The criteria for germ cell mutagens of GHS and other systems are reviewed and summarized to aid understanding of the control of chemicals by these regulations.

Classification Systems for Germ Cell Mutagens

Classification systems for mutagens or germ cell mutagens in GHS, EU, Germany, USA, Canada, and Japan are described below.

GHS: In the GHS (2), the term "mutation" applies both to heritable genetic changes that may be manifested at the phenotypic level and to the underlying DNA modifications when known (including, for example, specific base pair changes and chromosomal translocations). The term "mutagenic" and "mutagen" will be used for chemicals giving rise to an increased occurrence of mutations in populations of cells and/or organisms. The more general terms "genotoxic" and "genotoxicity" apply to chemicals or processes which alter the structure, information content, or segregation of DNA, including those which cause DNA damage by interfering with normal replication processes, or which in a non-physiological manner (temporarily) alter its replication.

In order to achieve classification, GHS states that 'Test results are considered from experiments determining mutagenic and/or genotoxic effects in germ and/or somatic cells of exposed animals. The system is hazard based, classifying chemicals on the basis of their intrinsic ability to induce mutations in germ cells. The scheme is, therefore, not meant for the (quantitative) risk assessment of chemical substances. Classification for heritable effects in human germ cells is made on the basis of well conducted, sufficiently validated tests, preferably as described in OECD Test Guidelines. Evaluation of the test results should be done using expert judgment and all the available evidence should be weighed for classification. The classification of individual substances should be based on the total weight of evidence available, using expert judgment. In those instances where a single well-conducted test is used for classification, it should provide clear and unambiguously positive results. If new, well validated, tests arise these may also be used in the total weight of evidence to be considered. The relevance of the route of exposure used in the study of the chemical compared to the route of human exposure should also be taken into account.'

The criteria for classification of germ cell mutagens places chemicals in one of three categories, category 1 being used for chemicals known to induce heritable mutations (category 1A) or known to be regarded as if they induce heritable mutations in germ cells of humans (category 1B); category 2 for chemicals which cause concern for humans owing to the possibility that they may induce heritable mutations in the germ cells of humans (2). The criterion for category 1A is positive evidence from human epidemiological studies. The criteria for category 1B are as follows: (i) positive result(s) from *in vivo* heritable germ cell mutagenicity tests in mammals; or (ii) positive result(s) from *in vivo* somatic cell mutagenicity tests in mammals, in combination with some evidence that the substance has the potential to cause mutations to germ cells. This supporting evidence may, for example, be derived from mutagenicity/genotoxic tests in germ cells *in vivo*, or by demonstrating the ability of the substance or its metabolite(s) to interact with the genetic material of germ cells; or (iii) positive results from tests showing mutagenic effects in the germ cells of humans, without demonstration of transmission to progeny; for example, an increase in the frequency of aneuploidy in the sperm cells of exposed people. The criteria for category 2 are positive evidence obtained from experiments in mammals and/or in some cases from *in vitro* experiments, obtained from: (i) somatic cell mutagenicity tests *in vivo*, in mammals; or (ii) other *in vivo* somatic cell genotoxicity tests which are supported by positive results from *in vitro* mutagenicity assays. In addition, following criteria are included as *Note*: Chemicals which are positive in *in vitro* mammalian mutagenicity assays, and which also show chemical structure activity relationship to known germ cell mutagens, should be considered for classification as Category 2 mutagens.

If there are not enough data for the evaluation of mutagenicity of the chemical, it regards as "classification not possible" (Fig. 1). If there is no concern of induction of heritable mutations in the germ cells of humans or insufficient evidence for inclusion in category 1 or 2, the chemicals are regarded as "not classified". Hazard categories and their criteria for germ cell mutagens in GHS are summarized in Table 1. For classification of chemical mixtures, the mixture will be

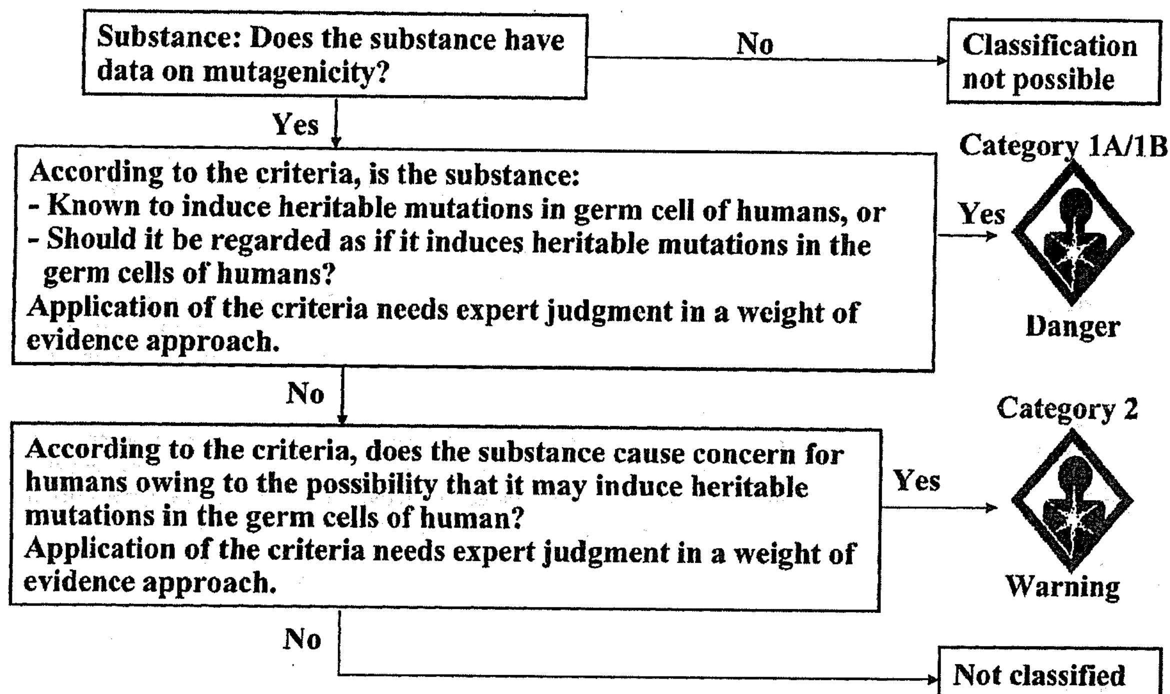

Fig. 1. Decision logic for the classification of germ cell mutagens in GHS (2)

classified as a mutagen when at least one ingredient has been classified as a category 1 or category 2 mutagen and is present at or above the cut-off value/concentration limits below for category 1 ($\geq 0.1\%$) or category 2 ($\geq 1.0\%$), respectively.

European Union (EU): The criteria for classification of mutagens in EU are described in Commission directive (5, 6). There are three categories: substances known to be mutagenic to human (category 1) for which there is sufficient evidence to establish a causal association between human exposure to the chemical and heritable genetic damage; substances which should be regarded as if they were mutagenic to human (category 2) for which there is sufficient evidence to provide a strong presumption that human exposure to the chemical may result in the development of heritable genetic damage, generally on the basis of appropriate animal studies and other relevant information; substances which cause concern for humans owing to possible mutagenic effect (category 3) for which there is evidence from appropriate mutagenicity studies, but it is insufficient to place the substance in category 2. EU criteria for classification of chemicals are summarized in Table 2.

This system is primarily based on intrinsic hazard, despite the statement in the Annex (7) that 'the object of classification is to identify all the physicochemical, toxicological and ecotoxicological properties of substances and preparations which may constitute a risk during normal handling or use' (4).

Germany: Maximale Arbeitsplatz-Konzentration (MAK) Commission in Germany proposed 5 categories for classification of germ cell mutagens at the workplace (8, 9). These are germ cell mutagens which have been shown to increase the mutant frequency in the progeny of exposed humans (category 1); germ cell mutagens which have been shown to increase the mutant frequency in the progeny of exposed mammals (category 2); chemicals which have been shown to induce genetic damage in germ cells of humans and/or animals, or which produce mutagenic effects in somatic cells of mammals *in vivo* and the chemicals have been shown to reach the germ cells in an active form (category 3A); chemicals which are suspected of being germ cell mutagens because of their genotoxic effects in mammalian somatic cells *in vivo*; in exceptional cases, chemicals for which there are no *in vivo* data but which are clearly mutagenic *in vitro* and structurally related to known *in vivo* mutagens (category 3B); and germ cell mutagens, the potency of which is considered to be so low that, provided the MAK value (Maximum Concentration at the Workplace) is observed, their contribution to genetic risk for man is expected not to be significant (category 5). Category 4 is not applicable in germ cell mutagenicity because this classification system has been