Table 1
Performance of in silico systems

	Ames result	+	−	Total	Sensitivity (%)	Specificity (%)	Concordance (%)	Applicability (%)
CGX database								
DEREK	+	288	64	352				
	−	69	267	336	81.8	79.5	80.7	97.9
	Total	357	331	688				
MCase	+	235	32	267				
	−	6	249	255	88.0	97.6	92.7	74.3
	Total	241	281	522				
AWorks	+	267	89	356				
	−	149	187	336	75.0	55.7	65.6	98.4
	Total	416	276	692				
ECJ database								
DEREK	+	19	7	26				
	−	21	159	180	73.1	88.3	86.4	100.0
	Total	40	166	206				
MCase	+	13	7	20				
	−	13	133	146	65.0	91.1	88.0	80.6
	Total	26	140	166				
AWorks	+	19	7	26				
	−	54	124	178	73.1	69.7	70.1	99.0
	Total	73	131	204				

MCase: MultiCASE; AWorks: ADMEWorks.

number of chemicals evaluated; and N_{all} is total number of chemicals subjected.

3. Results

Among the set of 703 CGX chemicals with published Ames data, 358 were positive and 345 were negative. The results of the in silico evaluation are summarized in Table 1. The highest sensitivity, specificity, and concordance with Ames assay results was provided by MCase, then followed by DEREK. However, the systems that showed the best applicability were AWorks and (almost the same) DEREK, then followed by MCase. For the database of 206 ECJ chemicals, 26 were positive and 180 were negative. The outcomes of the in silico analyses are summarized in Table 1. The pattern of performance was very similar to that with the 703 chemicals in the CGX database.

Fig. 1 shows the cumulative percent of Ames positive chemicals against molecular weight. It can be seen that 87.1% of those positive chemicals had molecular weights less than 1000, and 96.4% had molecular weights less than 3000; in other words, only 3.6% of the chemicals with a molecular weight >3000 gave a positive response in the Ames assay. Seven of 194 Ames positive chemicals had a molecular weight >3000 and four of these seven polymers had epoxy groups.

When we combined the in silico systems, the performance was different from that when assessed individually (Table 2). If we considered the in silico mutagenicity as positive (or negative) when two or more systems gave positive (or negative) evaluations, 87.8

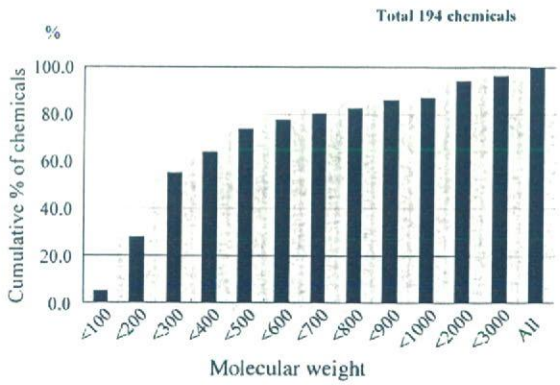

Fig. 1. Cumulative percentage of chemicals based on their molecular weight. 194 Ames positive chemicals were analyzed. 7/194 chemicals were more than 3000 molecular weight and Ames positive and 4/7 contained epoxy groups.

Table 2
Performance of in silico systems after combined

CGX database

In silico / Ames	++ or +++	-- or ---	Total	Sensitivity (%)	Specificity (%)	Concordance (%)	Applicability (%)
+	279	40	319				
-	42	249	291	**87.8**	**85.6**	**86.7**	**86.8**
Total	321	289	610				
	+++	---					
+	166	1	167				
-	3	127	130	**99.4**	**97.7**	**98.7**	**42.2**
Total	168	129	297				

FCJ database

In silico / Ames	++ or +++	-- or ---	Total	Sensitivity (%)	Specificity (%)	Concordance (%)	Applicability (%)
+	19	7	26				
-	23	147	170	**73.1**	**86.5**	**84.7**	**95.1**
Total	42	154	196				
	+++	---					
+	13	2	15				
-	5	94	99	**86.7**	**94.9**	**93.9**	**55.3**
Total	18	96	114				

Table 3
Performances of DEREK and MCase in several published papers.

Target compounds	In silico system	Sensitivity (%)	Specificity (%)	Concordance (%)	Applicability (%)	Reference
394 Drugs	DEREK	52	75	74	94[a]	[11]
	MCase	48	93	90	91[a]	
217 Non-drugs	DEREK	86	50	81	100[a]	[10]
	MCase	91	62	83	100[a]	
520 Drug candidates	DEREK	28	80	72	100	[13]
	MCase	50	86	81	41	
	DEREK + MCase	29	95	88	29	
	DEREK + Mcase + TOPKAT	75	96	95	15	
123 Drug candidates	DEREK	8[b]	31[c]	61	100[d]	[4]
	MCase (A2H)	13[b]	15[c]	72	97[d]	
	Topcat (Ames Mut)	18[b]	15[c]	67	98[d]	
	DEREK + MCase	6[b]	19[c]	75	97[d]	
	DEREK + Mcase + TOPKAT	5[b]	9[c]	86	46[d]	
94 Non-drugs	DEREK	63	81	76	100	[13]
	MCase	40	90	76	75	
	DEREK + MCase	47	100	85	56	
	DEREK + Mcase + TOPKAT	55	100	86	37	
516 Non-drugs	DEREK	6[b]	24[c]	70	100[d]	[4]
	MCase (A2H)	7[b]	12[c]	81	98[d]	
	Topcat (Ames Mut)	25[b]	19[c]	56	97[d]	
	DEREK + MCase	2[b]	16[c]	82	98[d]	
	DEREK + Mcase + TOPKAT	7[b]	10[c]	83	43[d]	

[a] Calculated by us
[b] % False negative.
[c] % False positive.
[d] (1-Indeterminate).

and 73.1% sensitivity, 85.6 and 86.5% specificity, 86.7 and 84.7% concordance, and 86.8 and 95.1% applicability were obtained for the CGX and ECJ databases, respectively. If we considered the in silico mutagenicity as positive (or negative) only when all three systems gave positive (or negative) evaluations, all performance measures (sensitivity, specificity, etc.) increased up to 98.7 and 93.9%. However, applicability decreased to 42.2 and 55.3%, which meant only about half of the chemicals in the CGX and ECJ databases could be evaluated. One chemical, *o*-phenylphenol [90-43-7], was positive in the Ames test but negative by all three in silico systems and three chemicals, carboxymethylnitrosourea [60391-92-6], methidathion [950-37-8], 1-nitroso-3,5-dimethyl-4-benzoylpiperazine [61034-40-0], were negative in the Ames test although all three in silico system gave positive evaluation for mutagenicity in the CGX database. When we used the ECJ database, 2-amino-1-naphthalenesulfonic acid [81-16-3] and 2-vinylpyridine [100-69-6] were positive in the Ames test but negative by all three in silico systems and there was no chemical that was negative in the Ames assay and all positive in in silico system. These exceptional chemicals are listed in Table 3 together with such chemicals taken from literatures.

4. Discussion

It is important to construct a strategy for efficient evaluation of the toxicity of a large number of existing chemicals. Even so-called short-term assays, e.g., Ames assay and in vitro chromosomal aberration assay, can practically assess only 100 chemicals per year according to our experiences in Japan. In this case, it will take 180 years to assess the outstanding 18,000 existing chemicals for genotoxicity, and it will take even longer when repeat dose toxicity tests are also performed, as these are not short-term assays. We therefore need higher-throughput systems to assess chemical safety, or at least to set priorities for those chemicals that should be tested in in vitro and/or in vivo tests. In silico systems have the capability for high throughput but have not yet been well validated for assessment of human hazard, although some regulatory bodies have started to use these methods.

Correlation between the Ames assay result and molecular weight could be explained by the lack of membrane permeability of high molecular weight chemicals, making it more difficult for them to reach target molecules such as DNA and proteins that contribute to the fidelity of cell division. Therefore, only a few chemicals with molecular weight >3000 gave positive responses in the Ames assay. This phenomenon is also true for induction of chromosomal aberrations in vitro (data not shown). The other important issue is the contribution of epoxy group in the polymer. Although of molecular weight >3000, some polymers with an epoxy group gave positive results in both the Ames and chromosomal aberration assays. Epoxy embedding reagents employed in electron microscopy (e.g., epon and araldite) have been reported as mutagenic in the Ames assay [8]. According to these findings, we should include a step to evaluate molecular weight and existence of any epoxy groups in the molecule.

In the present study, we used the CGX database recently published by Kirkland et al. [1] for microbial mutagenicity data on 358 carcinogens and 345 non-carcinogens for validation of three commercially available in silico (Q)SAR systems. When applied individually, MCase gave high sensitivity, specificity, and concordance compared to other two systems. One of the reasons may be because the CGX database contained many results from the U.S. National Toxicology Program (NTP), and the learning dataset of MCase would have used many of the same results. Therefore, some of them were evaluated by direct matching. Moreover, the applicability of MCase was relatively low compared with the other systems in this study (Table 1). MCase judged 119 chemicals as inconclusive and one chemical as marginal, and could not evaluate 67 chemicals. Such selectivity in MCase may contribute to the high concordance. On the other hand, the other systems were not influenced directly by the NTP data. We applied the in silico systems to another dataset, the ECJ database, that does not contain the NTP data and we obtained similar patterns of sensitivity, specificity, etc.

Each in silico system showed different outcomes on some chemicals complimentary by some extent. These different evaluation patterns were mainly due to the different evaluation rules. The DEREK is a rule-based system, AWorks is a discriminant-based system mainly depending on physicochemical descriptors, and MCase is a hybrid system based on a database. Therefore, we concluded that in silico evaluation could be optimized by combining the evaluations from the three systems. Sensitivity, specificity and concordance were increased when we combined the three in silico systems to make a final conclusion of mutagenicity (Table 1). Concordance was much higher after combining but the applicability became poor (42.2%). When two of the in silico systems gave the same evaluations, the applicability (86.8%) was good but the concordance was lower (86.7%) than when all three were combined (98.7).

Recently, several in silico studies for prediction of mutagenicity have been conducted on drugs or non-

Table 4
Exceptional chemicals that showed Ames test gave positive but all three in silico systems (DEREK, MCase, TOPKAT/AWorks) gave negative and Ames test gave negative but all three systems gave positive

Compound	CAS	Ames test	DEREK	MCase	TOPKAT/Aworks	Source[a]
Bupropion	34911-55-2	+	–	–	–	1
Citalopram	59729-33-8	+	–	–	–	1
Naloxone	465-65-6	+	–	–	–	1
Oxcarbazepime	28721-07-5	+	–	–	–	1
Quetiapine	111976-69-7	+	–	–	–	1
Rabeprazole	117976-89-3	+	–	–	–	1
Zolmitriptan	139264-17-8	+	–	–	–	1
2-(2-Methylpropyl) thiazole	18640-74-9	+	–	–	–	2
2-Chloropyridine	109-09-1	+	–	–	–	2
Pyrogallol	87-66-1	+	–	–	–	2
o-Phenylphenol	90-43-7	+	–	–	–	3
2-Amino-1-naphthalenesulfonic acid	81-16-3	+	–	–	–	3
2-Vinylpyridine	100-69-6	+	–	–	–	3
Fosfomycin	23155-02-4	–	+	+	+	1
Toremifene	89778-26-7	–	+	+	+	1
Poly (2-hydroxypropyl methacrylate)	25703-79-1	–	+	+	+	2
Carboxymethylnitrosourea	60391-92-6	–	+	+	+	3
Methidathion	950-37-8	–	+	+	+	3
1-Nitroso-3,5-dimethyl-4-benzoylpiperazine	–	+	+	+	3	

[a] 1: Synder et al. [11] (with TOPKAT), 2: White et al. [13] (with TOPKAT), 3: this study (with AWorks).

drug chemicals with commercially available programs, e.g., DEREK, MCase or TOPKAT, or newly developed computational approaches [4,9–12]. The performances of DEREK and MCase in several of these studies are summarized in Table 4. Generally, similar performance in sensitivity, specificity, concordance, and applicability were shown between DEREK and MCase but with some exceptions, e.g., sensitivity in 520 drug candidates [13], specificity in 516 non-drugs [4], and applicability in 520 pharmaceutical drug candidates and 94 non-drugs [13]. These differences might be due to the chemical class of target compounds in each database. However, there was no remarkable difference in performance whether the chemical was intended for use as a pharmaceutical, agricultural, or industrial agent. Our results on performance of in silico systems showed similarity with the published analyses. With respect to the combination of in silico prediction systems, White et al. [13] reported that combination improved the overall accuracy and specificity, but sensitivity was barely above the 50% level (Table 4). On the other hand, their analysis showed quite low applicability in the combination of three prediction systems, DEREK, MCase and TOPKAT. Our analysis of the combination of DEREK, MCase and AWorks showed good improvements in sensitivity, specificity and concordance, but applicability was low, especially in the 3-system combination.

Exceptional chemicals that gave positive Ames results but were negative in all three in silico systems (DEREK, MCase, TOPKAT/AWorks), and those that were negative in the Ames test but gave positive evaluations in all three systems, are summarized in Table 4. This table, which includes data from Synder et al. [11] and White et al. [13] shows there are 19 exceptional chemicals from both drug and non-drug families. Although it would be unrealistic to expect zero exceptions using this approach, further improvement of the prediction systems is needed. We do not have good reasons to explain the discordance, therefore we will verify the results from both sides, i.e., in silico system and Ames test.

Considering these outcomes, we propose a decision tree (Fig. 2), in order to evaluate chemical induction of gene mutation. We may use the decision tree to prioritize chemicals to be assayed by in vitro and/or in vivo tests. A final goal being that eventually, chemical mutagenicity will be evaluated by in silico systems alone for regulatory use. The decision tree consists of three steps; namely to assess the molecular weight, the existence of epoxy groups, and the in silico evaluation for genotoxicity. Based on the purpose of the in silico evaluation, the tree might be altered by the different final call of the in silico evaluation, i.e., regarding as positive (negative) all three systems show positive (negative). The choice of definition for final call applying to the decision tree should be based on the balance between accuracy of eval-

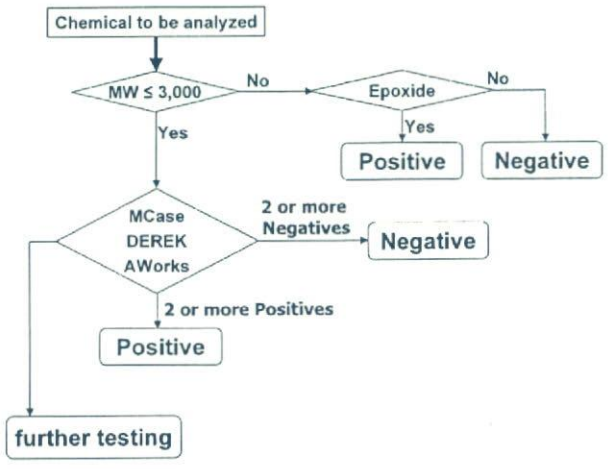

Fig. 2. Decision tree. In in silico evaluation, when two or more give positive then the final call is "positive" and two or more negative then call "negative".

uation and applicability, which are especially important for regulatory purpose. The decision should be made on a case-by-case basis depending upon the purpose of the decisions to be made.

Acknowledgements

The authors thank Dr. D. Kirkland, Covance, for his critical review and kind English edition and Mr. T. Ehara, MHLW, for his invaluable discussion. Authors also want to thank Ms. H. Akiyama (CTC laboratories, Japan), Akamatsu and Naitoh (Charles River Japan, Japan), and Kitajima, Suiroi, and Yuta (Fujitsu Kitakyushu, Japan) for their technical assistance. This work supported by the Health and Labour Sciences Research Grants (H15-Chemistry-003).

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] Law Concerning the Evaluation of Chemical Substances and Regulation of Their Manufacture, etc., Law No. 117, 16 October 1973 as last amended by Law No.49, 28 May 2003.

[3] M.T.D. Cronin, J.S. Jaworska, J.D. Walker, M.H.I. Comber, C.D. Watts, A.P. Worth, Use of QSARs in international decision-making frameworks to predict health effects of chemical substances, Environ. Health Perspect. 111 (2003) 1391–1401.

[4] G.M. Pearl, S. Livingstone-Carr, S.K. Durham, Integration of computational analysis as sentinel tool in toxicological assessments, Curr. Topoics Med. Chem. 1 (2001) 247–255.

[5] A. Hirose, M. Takahashi, M. Kamata, M. Ema, M. Hayashi, Development of genotoxicity predicting QSAR system for registered and existing industrial chemicals in Japan, Toxicol. Appl. Pharmacol. 197 (2004) 358.

[6] N. Greene, P.N. Judson, J.J. Langowski, C.A. Marchant, Knowledge-based expert systems for toxicity and metabolism prediction: DEREK, StAR and METEOR, SAR QSAR Environ. Res. 10 (1999) 299–314.

[7] H.S. Rosenkranz, A.R. Cunningham, Y.P. Zhang, H.G. Claycamp, O.T. Macina, N.B. Sussman, S.G. Grant, G. Klopman, Development, characterization and application of predictive-toxicology models, SAR QSAR Environ. Res. 10 (1999) 277–298.

[8] M.P. Murray, J.E. Cummins, Mutagenic activity of epoxy embedding reagents employed in electron microscopy, Environ. Mutagen. 1 (1979) 307–313.

[9] N.F. Cariello, J.D. Wilson, B.H. Britt, D.J. Wedd, B. Burlinson, V. Gombar, Comparison of the computer programs DEREK and TOPKAT to predict bacterial mutagenicity. Deductive estimate of risk from existing knowledge. Toxicity prediction by computer assisted technology, Mutagenesis 17 (4) (2002) 321–329.

[10] J.R. Votano, M. Parham, L.H. Hall, L.B. Kier, S. Oloff, A. Tropsha, Q. Xie, W. Tong, Three new consensus QSAR models for the prediction of Ames genotoxicity, Mutagenesis 19 (5) (2004) 365–377.

[11] R.D. Snyder, D.E. Ewing, L.B. Hendry, Evaluation of DNA intercalation potential of pharmaceuticals and other chemicals by cell-based and three-dimensional computational approaches, Environ. Mol. Mutagen. 44 (2) (2004) 163–173.

[12] R.D. Snyder, G.S. Pearl, G. Mandakas, W.N. Choy, F. Goodsaid, I.Y. Rosenblum, Assessment of the sensitivity of the computational programs DEREK, TOPKAT, and MCASE in the prediction of the genotoxicity of pharmaceutical molecules, Environ. Mol. Mutagen. 43 (3) (2004) 143–158.

[13] A.C. White, R.A. Mueller, R.H. Gallavan, S. Aaron, A.G. Wilson, A multiple in silico program approach for the prediction of mutagenicity from chemical structure, Mutat. Res. 539 (1–2) (2003) 77–89.

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Mutation Research 627 (2007) 10–30
Genetic Toxicology and Environmental Mutagenesis
www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

In vivo erythrocyte micronucleus assay III. Validation and regulatory acceptance of automated scoring and the use of rat peripheral blood reticulocytes, with discussion of non-hematopoietic target cells and a single dose-level limit test

Makoto Hayashi [a,*], James T. MacGregor [b], David G. Gatehouse [c], David H. Blakey [d], Stephen D. Dertinger [e], Lilianne Abramsson-Zetterberg [f], Gopala Krishna [g], Takeshi Morita [h], Antonella Russo [i], Norihide Asano [j], Hiroshi Suzuki [k], Wakako Ohyama [l], Dave Gibson [m]

[a] *Division of Genetics & Mutagenesis, National Institute of Health Sciences, 1-18-1 Kamiyoga Setagaya-ku, Tokyo 158-8501, Japan*
[b] *Toxicology Consulting Services, 201 Nomini Drive, Arnold, MD 21012, USA*
[c] *Covance Laboratories Ltd, Otley Rd, Harrogate, North Yorkshire, HG3 1PY, UK*
[d] *Health Canada, Environ Health Center Room 309A, Tunney's Pasture 0803A, Ottawa, Ont., Canada K1A 0L2*
[e] *Litron Laboratories, 200 Canal View Blvd., Rochester, NY 14623, USA*
[f] *National Food Administration, Toxicology Division, Box 622, S-751 26 Uppsala, Sweden*
[g] *Department of Toxicology, MGI Pharma, 6611 Tributary Street, Baltimore, MD 21224, USA*
[h] *Division of Safety Information on Drug, Food and Chemicals, National Institute of Health Sciences, 1-18-1, Kamiyogo, Setagaya-ku, Tokyo 158-8501, Japan*
[i] *Department of Biology, University of Padova, Via U. Bassi 58/b, 35131 Padova, Italy*
[j] *Toxicological Research Center, Nitto Denko Corporation, 1-1-2, Shimohozumi, Ibaraki, Osaka 567-8680, Japan*
[k] *Ina Research Inc., 2148-188 Nishiminowa, Ina-shi, Nagano 399-4501, Japan*
[l] *Yakult Central Institute for Microbiological Research, 1796 Yaho, Kunitachi-shi, Tokyo 186-8650, Japan*
[m] *The Procter & Gamble Company, PO Box 538707, Cincinnati, OH 45253-8707, USA*

Received 28 February 2006; received in revised form 31 July 2006; accepted 14 August 2006
Available online 6 December 2006

Abstract

The *in vivo* micronucleus assay working group of the International Workshop on Genotoxicity Testing (IWGT) discussed new aspects in the *in vivo* micronucleus (MN) test, including the regulatory acceptance of data derived from automated scoring, especially with regard to the use of flow cytometry, the suitability of rat peripheral blood reticulocytes to serve as the principal cell population for analysis, the establishment of *in vivo* MN assays in tissues other than bone marrow and blood (for example liver, skin, colon, germ cells), and the biological relevance of the single-dose-level test.

Our group members agreed that flow cytometric systems to detect induction of micronucleated immature erythrocytes have advantages based on the presented data, e.g., they give good reproducibility compared to manual scoring, are rapid, and require only small quantities of peripheral blood. Flow cytometric analysis of peripheral blood reticulocytes has the potential to allow monitoring of chromosome damage in rodents and also other species as part of routine toxicology studies. It appears that it will be applicable to humans as well, although in this case the possible confounding effects of splenic activity will need to be considered closely. Also, the consensus of the group was that any system that meets the validation criteria recommended by the IWGT (2000)

* Corresponding author. Tel.: +81 3 3700 9872; fax: +81 3 3700 2348.
E-mail address: hayashi@nihs.go.jp (M. Hayashi).

1383-5718/$ – see front matter
doi:10.1016/j.mrgentox.2006.08.010

should be acceptable. A number of different flow cytometric-based micronucleus assays have been developed, but at the present time the validation data are most extensive for the flow cytometric method using anti-CD71 fluorescent staining especially in terms of inter-laboratory collaborative data. Whichever method is chosen, it is desirable that each laboratory should determine the minimum sample size required to ensure that scoring error is maintained below the level of animal-to-animal variation.

In the second IWGT, the potential to use rat peripheral blood reticulocytes as target cells for the micronucleus assay was discussed, but a consensus regarding acceptability for regulatory purposes could not be reached at that time. Subsequent validation efforts, combined with accumulated published data, demonstrate that blood-derived reticulocytes from rats as well as mice are acceptable when young reticulocytes are analyzed under proper assay protocol and sample size.

The working group reviewed the results of micronucleus assays using target cells/tissues other than hematopoietic cells. We also discussed the relevance of the liver micronucleus assay using young rats, and the importance of understanding the maturation of enzyme systems involved in the processes of metabolic activation in the liver of young rats. Although the consensus of the group was that the more information with regard to the metabolic capabilities of young rats would be useful, the published literature shows that young rats have sufficient metabolic capacity for the purposes of this assay. The use of young rats as a model for detecting MN induction in the liver offers a good alternative methodology to the use of partial hepatectomy or mitogenic stimulation. Additional data obtained from colon and skin MN models have been integrated into the data bases, enhancing confidence in the utility of these models.

A fourth topic discussed by the working group was the regulatory acceptance of the single-dose-level assay. There was no consensus regarding the acceptability of a single dose level protocol when dose-limiting toxicity occurs. The use of a single dose level can lead to problems in data interpretation or to the loss of animals due to unexpected toxicity, making it necessary to repeat the study with additional doses. A limit test at a single dose level is currently accepted when toxicity is not dose-limiting.

Keywords: Micronucleus assay; Automation; Flow cytometry; Rat blood micronucleus assay; Validation; Liver micronucleus assay; Skin micronucleus assay; Colon micronucleus assay; Germ cell micronucleus assay

1. Introduction

The International Workshop on Genotoxicity Test Procedures (IWGTP), first held in February 1993 in Melbourne, Australia in conjunction with the 6th International Conference on Environmental Mutagens, has developed consensus on a number of major methodological issues related to regulatory genetic toxicology assays [1]. The conclusions reached at the Melbourne IWGTP meeting by the expert working group on the micronucleus assay have been summarized by Hayashi et al. [2]. The second IWGTP was held in Washington D.C., March 25–26, 1999, and the report of this meeting is complimentary to that of the first report of the Melbourne meeting. In the 2nd IWGTP, the *in vivo* micronucleus test working group discussed a number of issues, including integration of the repeated-dose micronucleus assay into general toxicology studies, the use of automated scoring techniques, micronucleus scoring in other tissues, and methods for differentiating micronuclei derived from acentric chromosome fragments and from centromeric chromosomes [3].

In the 4th International Workshop on Genotoxicity Testing (IWGT), in San Francisco on September 8 and 9, 2005 (an *in vivo* micronucleus assay working group meeting did not take place at the 3rd IWGT, in Plymouth, 2002) a half-day-meeting was held to consider new developments related to the use of the *in vivo* micronucleus test and aimed:

(1) to clarify whether automated methods such as those employing flow cytometry or image analysis are sufficiently well validated to allow the acceptance of such data by the regulatory authorities,
(2) to determine whether the use of rat peripheral blood reticulocytes as the primary endpoint evaluated is sufficiently well validated to allow the acceptance of such data by the regulatory authorities,
(3) to review the current status and any new data related to the *in vivo* micronucleus assay using tissues other than bone marrow and peripheral blood (for example liver, skin, colon, etc.),
(4) to evaluate the acceptability of the single-dose-assay not only at the limit dose (i.e., 2000 mg/kg) but also at other single dose levels when defined by other specific rationales.

1.1. Automated scoring for regulatory use

One of the most important characteristics of the rodent micronucleus assay is the simplicity of scoring its endpoint, namely identification of a small membrane-bound chromatin-containing body in the cytoplasm of the erythrocyte from which the main nucleus has been eliminated during the last stage of hematopoiesis. The characteristics of the micronucleus (a smooth-boundaried body that stains strongly and specifically for chromatin) lend themselves well to automated scoring by image analysis or flow cytometry.

We have already discussed the application of automated systems to scoring of the rodent micronucleus assay in Washington D.C. in 1999 at the 2nd IWGTP [3]. We summarized the discussion on this topic as follows: "In summary, although early versions of automated scoring techniques suffered from unacceptable levels of artifacts and/or were deficient with respect to distinguishing immature erythrocytes from mature erythrocytes, the modern systems described herein have overcome these limitations. These flow cytometry and image analysis techniques can be expected to replace microscopy-based scoring by providing reliable data in a more efficient manner. However, rigorous validation and adequate quality-control systems are necessary prerequisites to the use of automated scoring." At the 2nd IWGTP, we also discussed the validation criteria and drew the following conclusions:

"Considering the OECD general points, each system should also meet the following validation criteria:

1. Results obtained should be demonstrated to be comparable with those obtained by direct manual microscopic scoring of:
   (a) micronucleated immature erythrocyte frequency
   (b) micronucleated mature erythrocyte frequency, and
   (c) percentage of immature erythrocytes among total erythrocytes.

   This demonstration must be made for each tissue and species analyzed, and should verify that all cell types of interest (e.g., micronucleated immature and mature erythrocyte, immature and mature erythrocyte) are adequately identified and scored.
2. Micronuclei arising from both chromosome fragments and whole chromosomes should be demonstrated to be scored consistently.
3. Scoring should be shown to be consistent within and between experiments.
4. Preferably, laboratories should establish how known potential artifacts (platelets, basophilic granules, Heinz bodies, aggregated RNA, etc.) behave in their system.

At the 4th IWGT, we re-visited this topic and we focused on the flow cytometric method for the micronucleus assay. Image analysis is acknowledged to also be a useful automated methodology, but because image analysis is more closely related to manual microscopy and the image can be retrieved to the microscopic field to confirm the accuracy manually validation is more straightforward than with flow cytometry. With flow cytometry, micronucleated cells are normally not visualized, although it is possible to sort the target cells onto glass slides to confirm microscopically that the target cells are correct.

Flow cytometric analysis is rapid and provides the ability to analyze a large number of cells. A number of papers that demonstrate successful automated scoring methodologies have been published [4–19], although some of the earlier papers were focused on technological development rather than regulatory validation. Several publications have demonstrated that low dose effects of clastogens could be determined by analyzing large numbers of young erythrocytes that were not practical to evaluate by manual microscopy [20–25]. Additionally, recent interlaboratory data has shown that the reproducibility of flow cytometric analysis is much greater than microscopic scoring when a standard is used to calibrate the flow cytometry system [26,27]. Therefore it was the consensus of the group that flow cytometric analysis increases the capacity, and to a certain extent the sensitivity, of the *in vivo* micronucleus assay. As shown by Asano and colleagues [25], flow cytometry is capable of evaluating sufficient numbers of cells per animal such that variation among individuals becomes the limiting factor that defines the minimal micronucleus response that can be detected. Consequently whichever method is used, it is desirable that each laboratory should determine the minimum cell sample size required to ensure that scoring error is maintained below the level of animal-to-animal variation.

The group reconfirmed their view that data acquired by any automated method should be acceptable for regulatory use once it has met the validation criteria which were defined by the IWGT Micronucleus Working Group in Washington D.C., 1999 [3]. An important consensus agreement is that at the present time, sufficient validation has occurred for the flow cytometric CD71 method, as more extensive inter-laboratory collaborative data are available compared to the other procedures [24,26–28]. The outcomes of these inter-laboratory collaborative studies show good agreement between microscopy- and flow cytometry-generated data, as well as good intra- and inter-laboratory reproducibility for this methodology (see Appendices A and B).

During the discussions it was stated that data obtained by the flow cytometric analysis had already been evaluated and accepted by the regulatory authorities in the US [29] and EU (e.g., [30]).

1.2. Rat Blood reticulocytes for regulatory use

Accumulated evidence has established that it is acceptable to rely on the measurement of MN-reticulo-

cyte frequency in peripheral blood as well as bone marrow. This conclusion was restricted to the mouse in the ICH consensus because of presumed splenic elimination of MN-erythrocytes by the rat spleen [31,32]. In the second IWGT at Washington D.C. in 1999, we discussed the use of peripheral blood of rat as target cells to be analyzed as well as the use of mouse peripheral blood based on the Japanese collaborative trial data [33,34] (see Appendix C). However, the 1999 working group was unable to reach a consensus agreement because there was still some concern of rat spleen function. It was the consensus of the 2005 group that accumulated evidence shows that analysis of micronucleated peripheral blood reticulocytes is an acceptable cell population for the micronucleus assay in rats as well as mice [27]. Further, preliminary data suggest that peripheral blood may be an acceptable sampling tissue in other species including human [15,35–40]. In the case of humans as well as rats, the possible confounding effects of splenic activity will need to be considered (see Appendices A–C).

1.3. Micronucleus assays using tissues other than bone marrow

One of the most important issues in the risk assessment of chemicals is the mode of action for carcinogenesis. At the present time it is considered that if the mechanism involves mutagenicity, then a threshold cannot be considered and a linear model would be used to set any potential acceptable daily intake (ADI). The Salmonella/microsome reverse mutation assay (as well as other *in vitro* assays) have played an important role in identifying whether a chemical is mutagenic or not. But, *in vitro* assay data cannot predict whether the chemical expresses its mutagenicity at the target site of carcinogenicity. Based on the weight of evidence philosophy, it is important to assess mutagenicity at the site of carcinogenesis. Theoretically transgenic animal models can assess gene mutations in any organ. On the other hand, chromosomal damage can be assessed only in proliferating organs and the bone marrow (hematopoietic) cells have been frequently used for this purpose. In the 2nd IWGT meeting in Washington D.C., we reviewed and evaluated the micronucleus assay using tissues other than bone marrow [3]. Here, we revisit the same topic again because the assessment of chromosomal damage using different organs has become more important and has been used increasingly for risk assessment purposes. We updated the tables of the micronucleus assay using tissues other than bone marrow including the collaborative validation trials organized by the Collaborative Study Group for the Micronucleus Test (CSGMT)/Japanese Environmental Mutagen Society (JEMS)/Mammalian Mutagenicity Study Group (MMS) [41–44] (see Appendix D). One of the major developments was the introduction of a new staining method using acridine orange together with DAPI that stains the cytoplasm and nucleus clearly and differentially [45].

A liver micronucleus assay with rats and mice was developed by Tates et al. [46,47], who used partial hepatectomy (PH) to stimulate liver cell proliferation. Subsequent investigators used mice [48–50], and induced liver cell proliferation by chemical damage, which induced mitogenesis, e.g., 4-acetylaminofluorene [51] or carbon tetrachloride [52], or used modifications of timing of treatments or sampling intervals [50,53]. The basic assay procedures and protocols have been established and validated [49], and approximately 40 chemicals have been tested so far. Several carcinogens that are negative or weakly positive in the erythrocyte micronucleus assay showed genotoxic effects in regenerating liver.

Recently, a hepatocyte micronucleus assay has been described in which liver of young rats is used without partial hepatectomy or chemical mitogen treatment [54–56]. The CSGMT/MMS engaged collaborative trials on this topic and the first report has been published [41]. The method does not need any additional treatment to detect MN induction by a test chemical in hepatocytes. The group agreed that the use of young rats as a model for detecting MN induction in the liver offers an alternative methodology to the use of partial hepatectomy or mitogenic stimulation based on the following analysis of published literature.

The metabolic capabilities of young rats have been studied and there was no essential difference in the total amount of cytochrome P450 between immature (50–60 g body weight) and adult (150–190 g body weight) in Long Evans rat [57–59]. It is well known that there is no difference among rat strains in main P450 species. In more detail, no difference between immature and adult in CYP2B1/2, immature > adult in CYP2A1 and CYP1A1 were reported [57]. Cytochromes P450 2B1/2 and 1A2 were determined in the growing neonate and fetus of control and 3-methylcholanthrene-pretreated rats. CYP1A2 activity was highest in the 1–2-week-old animals and then decreased with age. The inducibility of this activity by 3-methylcholanthrene was low in the young animals, but increased with age. CYP 2E1 in SD rat showed almost equivalent activity in 4-week-old animals compared to adult animals [60]. CYP2C11, which is the dominant P450 in SD male rat liver, of 4-week-old SD rat showed less activity than adult [58]. Gonzalez et al. [61]

reported that the activity of testosterone 6β-hydroxylase in 4-week-old male rat (strain unknown) was approximately a half of that of adult, but the content was high in the young rat and the total activity thought to be sufficient for the assay [61]. In summary, 4-week-old-rat liver can metabolize important xenobiotic chemicals, e.g., heterocyclic hydrocarbons, heterocyclic amines, nitrosamines, almost as effectively as in adult rats.

Additional experiences with colon and skin MN models has increased the databases enhancing confidence in the utility of these models.

Although not discussed at the meeting, recently germ-cell mutagenicity has come into the spotlight, e.g., WHO initiative project "Globally Harmonized System of Classification and Labeling of Chemicals (GHS)". At present, there are still only limited methods of determining germ cell mutagenesis *in vivo* (e.g., the specific locus test, dominant lethal test, transgenic mutation assays), and these are not frequently performed for regulatory use. The micronucleus assay using spermatids is an additional model that may be valuable for assessing chemical-induced germ cell mutation. The use of a micronucleus assay in mammalian germ cells was proposed in the early 1980s when two different approaches were described by Lähdetie and Parvinen [62] and by Tates et al. [63].

We reviewed and summarized assay results during the 2nd IWGT, Washington, 1999 [3]. We searched references to update the summary table but not many articles have been published after the last survey (Appendix D). One of the possible reasons might be that identification of aneugenesis detected by the FISH method or the use of simpler methods that determine DNA damage may be being used more frequently than cytogenetic analysis or measurement of mutations *per se*. Although we did not prepare a summary table, for readers' information we cite the references [64–70].

The table for the micronucleus assay results using fetal/neonatal animals by transplacental treatment was revised and is presented in Appendix D.

1.4. Relevance of a single-dose-level assay

Under the current guidelines, e.g., OECD TG474 [71], the single-dose-level micronucleus assay is acceptable under limited conditions. The consensus during the IWGT meeting at Melbourne in 1993 was that negative data was acceptable from a study using a single dose level only when animals were treated at the limit dose of 2000 mg/kg and there was no depression of bone marrow proliferation and no sign of any adverse effects clinically. The single dose limit test is thus currently used just to demonstrate a clear negative response with non-toxic agents. As a basic guidance, for agents that are toxic below the 2000 mg/kg dose limit at least three dose levels separated by a factor between 2 and SQR (10) should be used [72].

The major concern with regard to the use of the single dose-level approach is the possibility of a 'downturn phenomenon'. Theoretically, the 'downturn phenomenon' should not occur when immature erythrocytes were targeted for analysis and the monotonic dose–response may be obtained when the sample time was optimized for each dose-level. The 'downturn phenomenon', however, has been shown to occur on some occasions although it is not clear whether this is due to excessive cytotoxicity within the bone marrow or to some other, as yet undefined, mechanism. In these cases, the group agreed, positive effects could still be obtained at the second highest dose. The use of a single dose level could sometimes lead to problems in data interpretation or to the loss of animals due to unexpected toxicity. In such cases it would be necessary to repeat the study with additional doses and this would lead to the use of an increased number of animals. After discussion of these issues, there was no consensus for the wider use of a single dose level protocol. Although it was agreed that it is important to reduce the number of experimental animals whenever possible, very few participants were in favor of such an approach in this case. There was an equal division among the participants concerning the need to use two or three doses-levels in this instance.

Acknowledgements

The authors would like to thank all participants at the *in vivo* working group of IWGT, San Francisco, September 9, 2005 for their extensive discussion, especially those who provided specific data, proposals, and suggestions. We also thank Dr. S. Ozawa for his literature review and evaluation of the expression of metabolic capacity in young rat liver.

Appendix A

Inter-laboratory variation and reproducibility of micronucleus induction using model chemicals (cyclophosphamide, cisplatin, and vinblastine) evaluated in peripheral blood by flow cytometry, by methanol-fixed acridine orange staining, and acridine orange supravital staining, and also methanol-fixed acridine orange staining in bone marrow cells are summarized from an extensive validation study [27].

Treatment	Intact/Splx	n	Lab	Avg. %MN-RET ± S.E.M. (fold increase)			
				BM-AO	PB-AO	PB-SVAO	PB-FCM
0 CP	Intact	5	L2				0.13 ± 0.02
		5	L3				0.16 ± 0.01
		5	L1	0.09 ± 0.03	0.15 ± 0.03		0.12 ± 0.01
		5	L9	0.04 ± 0.01	0.10 ± 0.04		
		5	L10	0.09 ± 0.04	0.05 ± 0.03		
		5	L11	0.22 ± 0.08	0.18 ± 0.06		
		5	L5			0.21 ± 0.07	
		5	L6			0.19 ± 0.04	
		5	L7			0.18 ± 0.02	
10 CP	Intact	5	L2				1.39 ± 0.24 (10.7)
		5	L3				1.39 ± 0.22 (8.7)
		5	L1	3.42 ± 0.28 (38.0)	1.62 ± 0.33 (10.8)		1.15 ± 0.15 (9.6)
		5	L9	1.52 ± 0.12 (38.0)	0.88 ± 0.25 (8.8)		
		5	L10	1.92 ± 0.36 (21.3)	0.41 ± 0.12 (8.2)		
		5	L11	2.74 ± 0.40 (12.5)	1.72 ± 0.37 (9.6)		
		5	L5			2.11 ± 0.38 (10.0)	
		5	L6			2.33 ± 0.31 (12.3)	
		5	L7			1.23 ± 0.34 (6.8)	
0 CP	Splx	5	L2				0.17 ± 0.02
		5	L3				0.19 ± 0.02
		5	L1	0.15 ± 0.02	0.16 ± 0.03		0.15 ± 0.02
		5	L9	0.12 ± 0.03	0.13 ± 0.03		
		5	L10	0.20 ± 0.04	0.09 ± 0.03		
		5	L11	0.32 ± 0.08	0.14 ± 0.04		
		5	L5			0.39 ± 0.07	
		5	L6			0.22 ± 0.03	
		5	L7			0.31 ± 0.05	
10 CP	Splx	5	L2				2.19 ± 0.26 (12.9)
		5	L3				2.43 ± 0.25 (12.8)
		5	L1	3.34 ± 0.29 (22.3)	3.31 ± 0.40 (20.7)		2.03 ± 0.18 (13.5)
		5	L9	1.39 ± 0.15 (11.6)	1.37 ± 0.17 (10.5)		
		5	L10	1.95 ± 0.33 (9.8)	0.61 ± 0.13 (6.8)		
		5	L11	3.06 ± 0.60 (9.6)	2.56 ± 0.42 (18.3)		
		5	L5			3.74 ± 0.50 (9.6)	
		5	L6			3.35 ± 0.24 (15.2)	
		5	L7			1.84 ± 0.37 (5.9)	
CP Aggregate Correlation With Ref. Lab (R^2)		20	L2				NA
		20	L3				0.9797
		20	L1	NA	NA		0.9742
		20	L9	0.9273	0.8842		
		20	L10	0.8551	0.8834		
		20	L11	0.8512	0.9396		
		20	L5			NA	
		20	L6			0.8901	
		20	L7			0.8173	
0 CisPl	Intact	5	L2				0.08 ± 0.02
		5	L3				0.09 ± 0.01
		5	L1	0.13 ± 0.03	0.08 ± 0.01		0.09 ± 0.01
		5	L9	0.08 ± 0.03	0.09 ± 0.02		
		5	L10	0.07 ± 0.04	0.05 ± 0.04		
		5	L11	0.17 ± 0.03	0.20 ± 0.02		
		5	L5			0.17 ± 0.08	
1.0 CisPl	Intact	5	L2				0.53 ± 0.14 (6.6)
		5	L3				0.48 ± 0.12 (5.3)
		5	L1	1.11 ± 0.24 (8.5)	0.65 ± 0.14 (8.1)		0.48 ± 0.09 (5.3)
		5	L9	0.32 ± 0.08 (4.0)	0.41 ± 0.12 (4.6)		
		5	L10	0.33 ± 0.10 (4.7)	0.28 ± 0.11 (5.6)		
		5	L11	0.62 ± 0.17 (3.6)	0.88 ± 0.38 (4.4)		
		5	L5			0.85 ± 0.26 (5.0)	
0 CisPl	Splx	5	L2				0.16 ± 0.02
		5	L3				0.16 ± 0.02
		5	L1	0.14 ± 0.03	0.17 ± 0.03		0.15 ± 0.02

Treatment	Intact/Splx	n	Lab	Avg. %MN-RET ± S.E.M. (fold increase)			
				BM-AO	PB-AO	PB-SVAO	PB-FCM
		5	L9	0.11 ± 0.01	0.12 ± 0.01		
		5	L10	0.06 ± 0.02	0.22 ± 0.04		
		5	L11	0.17 ± 0.07	0.29 ± 0.07		
		5	L5			0.15 ± 0.02	
1.0 CisPl	Splx	5	L2				0.89 ± 0.18 (5.6)
		5	L3				0.91 ± 0.19 (5.7)
		5	L1	1.06 ± 0.25 (7.6)	1.12 ± 0.17 (6.6)		0.90 ± 0.14 (6.0)
		5	L9	0.38 ± 0.13 (3.5)	0.54 ± 0.11 (4.5)		
		5	L10	0.52 ± 0.12 (8.7)	1.12 ± 0.20 (5.1)		
		5	L11	1.05 ± 0.30 (6.2)	0.86 ± 0.05 (3.0)		
		5	L5			1.11 ± 0.23 (7.4)	
CisPl Aggregate Correlation With Ref. Lab (R^2)		20	L2				NA
		20	L3				0.9846
		20	L1	NA	NA		0.9555
		20	L9	0.5165	0.8336		
		20	L10	0.7554	0.7481		
		20	L11	0.769	0.5104		
0 VB	Intact	5	L2				0.10 ± 0.01
		5	L3				0.11 ± 0.02
		5	L1	0.10 ± 0.02	0.13 ± 0.03		0.16 ± 0.01
		5	L9	0.06 ± 0.02	0.07 ± 0.02		
		5	L10	0.18 ± 0.05	0.08 ± 0.01		
		5	L11	0.15 ± 0.03	0.24 ± 0.03		
		3	L5			0.34 ± 0.07	
0.125 VB	Intact	3	L2				0.43 ± 0.08 (4.3)
		3	L3				0.41 ± 0.05 (3.7)
		5,3,3	L1	1.46 ± 0.18 (14.6)	0.57 ± 0.09 (4.4)		0.47 ± 0.07 (2.9)
		5,3	L9	0.65 ± 0.13 (10.8)	0.43 ± 0.15 (6.1)		
		5,5	L10	0.51 ± 0.11 (2.8)	0.11 ± 0.05 (1.4)		
		5,3	L11	0.70 ± 0.13 (4.7)	0.32 ± 0.06 (1.3)		
		2	L5			0.65 ± 0.20 (1.9)	
0 VB	Splx	4	L2				0.17 ± 0.03
		4	L3				0.20 ± 0.02
		4	L1	0.12 ± 0.01	0.21 ± 0.03		0.20 ± 0.02
		4	L9	0.04 ± 0.01	0.04 ± 0.02		
		4	L10	0.11 ± 0.03	0.09 ± 0.02		
		4	L11	0.26 ± 0.08	0.30 ± 0.12		
		4	L5			0.35 ± 0.10	
0.125 VB	Splx	5	L2				0.57 ± 0.19 (3.4)
		5	L3				0.72 ± 0.18 (3.6)
		5,3,5	L1	0.87 ± 0.24 (7.3)	0.60 ± 0.20 (2.9)		0.78 ± 0.22 (3.9)
		5	L9	0.34 ± 0.16 (8.5)	0.23 ± 0.10 (5.8)		
		5,5	L10	0.51 ± 0.15 (4.6)	0.13 ± 0.03 (1.4)		
		5	L11	0.75 ± 0.25 (2.9)	0.36 ± 0.08 (1.2)		
		3	L5			0.95 ± 0.53 (2.7)	
VB Aggregate Correlation With Ref. Lab (R^2)		17	L2				NA
		17	L3				0.9321
		17	L1	NA	NA		0.9584
		19, 15	L9	0.8134	0.3666		
		19, 15	L10	0.5644	0.0278		
		19, 15	L11	0.5436	0.0992		

Abbreviations: MN-RET, micronucleated reticulocyte; Splx, splenectomized; BM-AO, bone marrow scored by microscopy using standard acridine orange staining of methanol-fixed slides; PB-AO, peripheral blood scored by microscopy using standard AO staining of methanol-fixed slides; PB-SVAO, peripheral blood scored by microscopy using supravital acridine orange staining; PB-FCM, peripheral blood scored by flow cytometry using the propidium iodide/anti-CD71-based labeling method; S.E.M., standard error of the mean; *by "aggregate correlation" it is meant that the MN-RET value obtained for every animal in a particular study obtained at a reference laboratory (L1, L2 or L5) was plotted against the corresponding value obtained at each other laboratory. The R^2 values shown here were calculated in Microsoft Excel. Experimental design: Groups of intact and splenectomized Sprague Dawley male rats were administered vehicle, cyclophosphamide (10 mg/kg/day), cisplatin (1 mg/kg/day) or vinblastine (0.125 mg/kg/day) for 5 days at 24 h intervals (five animals per group). Twenty-four hours after the last administration, blood and bone marrow were prepared as coded slides or fixed blood suspensions and supplied to several laboratories that have extensive experience performing micronucleus measurements. MN-RET averages, S.E.M., fold-increases, and correspondence to reference laboratories are presented. In some instances, the number of measurements (n) performed per chemical does not equal 20. Except for Vinblastine Splx Control rats where $n = 4$, instances of $n < 5$ were due to excessive toxicity. More detailed information regarding these experiments can be found in MacGregor et al. [27].

Appendix B

Variation and reproducibility of intra- and inter-laboratory scoring of blinded triplicate samples are summarized from an extensive validation study [26].

Laboratory	Method	Compartment	Treatment	%MN-RET		
				Avg. ± S.E.M.	%CV	Fold difference
L1	MeOH-AO	BM	Vehicle	0.15 ± 0.03	33.3	
			CP	3.35 ± 0.10	5.4	22.3
L9	MeOH-AO	BM	Vehicle	0.05 ± 0.05	173.2	
			CP	1.63 ± 0.27	28.4	32.6
L10	MeOH-AO	BM	Vehicle	0.03 ± 0.02	86.6	
			CP	2.33 ± 0.23	17.3	77.7
L11	MeOH-AO	BM	Vehicle	0.18 ± 0.03	31.5	
			CP	2.44 ± 0.20	13.9	13.6
Pooled L1, 9, 10, 11*			*Vehicle*	*0.10 ± 0.02*	*80.5*	
			CP	*2.44 ± 0.21*	*29.1*	*24.4*
L1	MeOH-AO	PB	Vehicle	0.05 ± 0.03	100.0	
			CP	1.77 ± 0.17	16.6	35.4
L9	MeOH-AO	PB	Vehicle	0.05 ± 0.00	0.0	
			CP	0.50 ± 0.03	10.0	10.0
L11	MeOH-AO	PB	Vehicle	0.18 ± 0.04	41.7	
			CP	1.42 ± 0.10	12.3	7.9
Pooled L1, 9, 11			*Vehicle*	*0.09 ± 0.03*	*85.6*	
			CP	*1.23 ± 0.20*	*48.2*	*13.7*
L5	SV-AO	PB	Vehicle	0.13 ± 0.03	43.3	
			CP	1.83 ± 0.15	13.7	14.1
L6	SV-AO	PB	Vehicle	0.12 ± 0.07	99.0	
			CP	1.77 ± 0.32	31.2	14.8
L7	SV-AO	PB	Vehicle	0.22 ± 0.14	113.8	
			CP	1.47 ± 0.27	31.7	6.7
Pooled L5, 6, 7			*Vehicle*	*0.16 ± 0.05*	*94.3*	
			CP	*1.69 ± 0.14*	*24.7*	*10.6*
L1	FCM	PB	Vehicle	0.12 ± 0.02	24.8	
			CP	0.99 ± 0.04	6.5	8.3
L2	FCM	PB	Vehicle	0.11 ± 0.02	31.5	
			CP	1.04 ± 0.04	6.7	9.5
L3	FCM	PB	Vehicle	0.11 ± 0.02	32.9	
			CP	1.11 ± 0.04	6.8	10.1
Pooled L1, 2, 3			*Vehicle*	*0.11 ± 0.01*	*26.5*	
			CP	*1.05 ± 0.03*	*7.6*	*9.5*

Abbreviations: MN-RET, micronucleated reticulocyte; MeOH-AO, acridine orange staining of methanol-fixed smears; SV-AO, supravital staining using acridine orange-coated slides; FCM, flow cytometry; BM, bone marrow; PB, peripheral blood; CP, cyclophosphamide; S.E.M., standard error of the mean; %CV, percent coefficient of variance.

Experimental design: One female Sprague Dawley rat was administered vehicle or cyclophosphamide (10 mg/kg/day) for 6 days at 24 h intervals. Twenty-four hours after the last administration, blood and bone marrow were prepared as coded slides or fixed blood suspensions and supplied to several laboratories that have extensive experience performing micronucleus measurements. The vehicle and the cyclophosphamide specimens were each supplied as three uniquely coded slides or tubes for analysis. MN-RET averages, S.E.M., and mean fold increase for each laboratory are presented. *Like-method data from three or four labs were combined for "pooled" calculations. More detailed information regarding these experiments can be found in Dertinger et al. [26].

Appendix C

Accumulated evidence has established that it is acceptable to rely on the measurement of MN-reticulocyte frequency in peripheral blood as well as bone marrow. In the second IWGT at Washington D.C. in 1999, we have discussed the use of peripheral blood of rat as target cells to be analyzed as well as mouse based on the Japanese collaborative trial data [33,34]. As an appendix, we summarize here again the outcomes of Japanese collaborative trial to evaluate the relevancy of using rat peripheral blood reticulocytes as target cells to be analyzed. The negative control values were $0.14 \pm 0.12\%$ in bone marrow (252 rats) and $0.07 \pm 0.08\%$ in peripheral blood (287 rats). The positive control (cyclophosphamide, 20 mg/kg, oral gavage, sampled 24 h after the last treatment) were $1.8 \pm 0.8\%$ in bone marrow (56 rats) and $0.80 \pm 0.48\%$ in peripheral blood (77 rats). The Table shows the summary table of the collaborative trial on 16 model chemicals that induce micronuclei with different modes of action. Although the absolute values of micronucleated reticulocyte frequencies were different between observations of bone marrow cells and peripheral reticulocytes in negative and positive controls and also model chemicals, the statistically significant induction of micronucleated reticulocytes were observed on the positive control chemical (cyclophosphamide) and tested model chemicals.

			Bone marrow			Peripheral blood		
			MNPCE			MNPCE		
Chemicals	Species	Route	Dose (mg/kg)	(%)	Fold Inc.	Dose (mg/kg)	(%)	Fold Inc.
p-Aminoazobenzene	Rat	ip	80	0.41	3.40	80	0.38	3.50
	Mouse	ip	150	0.70	5.00	150	1.14	8.10
o-Aminoazotoluene	Rat	ip	600	0.11	0.90	600	0.16	8.00
	Mouse	ip	400	0.60	3.00			
Benzene	Rat	po	2000	2.34	16.7	2000	2.16	18.0
	Mouse	po	2000	2.75	34.4	2000	2.75	6.10
Diepoxybutane	Rat	po	144	4.00	22.2	144	0.78	26.0
	Mouse	po	72	8.68	36.2			
Ethyl methanesulfonate	Rat	ip	300	1.36	45.3	300	0.75	12.5
	Mouse	ip	200	1.48	8.20			
N-Ethyl-*N*-nitrosourea	Rat	ip	75	2.78	27.8	75	1.36	12.4
	Mouse	ip	100	7.85	60.4	50	3.78	11.8
5-Fluorouracil	Rat	ip	40	1.39	10.7	80	0.69	34.5
	Mouse	ip	50	3.13	24.1	50	2.54	18.1
Hydrazine	Rat	ip	75	0.58	2.80	12.5	0.15	2.50
	Mouse	ip				50	1.18	5.90
6-Mercaptopurine	Rat	ip	50	0.92	6.10	25	0.53	5.30
	Mouse	ip	50	6.98	27.9	25	2.9	7.30
MNNG	Rat	ip	100	1.20	4.00	100	0.75	3.75
	Mouse	ip	150	2.55	10.2	100	0.8	3.10
Nitrogen mustard	Rat	ip	2	3.24	23.5	2	0.66	7.50
	Mouse	ip	2	7.17	29.6	2	4.98	21.4
1,3-Propane sultone	Rat	ip	60	2.15	13.4	30	0.66	7.30
	Mouse	ip				72	1.92	11.3
Propylene oxide	Rat	ip	200	0.33	2.20	300	0.4	3.1
	Mouse	ip	300	0.53	8.80	300	0.4	6.7
Triethylenemelamine	Rat	ip	0.5	5.58	11.8	0.13	1.08	6.80
	Mouse	ip	1	7.52	32.7	1	6.64	66.4
Urethane	Rat	ip	1000	1.29	11.7	1000	1.6	40.0
	Mouse	ip				1000	3.28	16.4
Vincristine	Rat	ip	0.2	0.99	7.60	0.08	0.26	4.3
	Mouse	ip	0.125	11.68	89.8	0.3	4.34	54.3

Light shadow: frequency in mice was more than twice that in rats. Dark shadow: frequency in rats was more than twice that in mice. *MS/Ae strain mice.

Appendix D

Micronucleus assay results using tissues other than haematopoietic tissue have become increasingly important for risk assessment to assess chemical clastogenicity at the target site of carcinogenesis or other adverse effects. We have summarized the data on micronucleus induction in liver, colon, and skin in Table D.1 including new data published after the 2nd IWGTP report. In Table D.2, we add new data on the micronucleus assay results on germ cells (mainly spermatids). These data using germ cells are more weighted in "Globally Harmonized System of Classification and Labeling of Chemicals (GHS)" for labeling of chemicals. Table D.3 shows micronucleus assay results using fetal/neonatal animals by transplacental treatment.

Table D.1
Micronucleus assay results using liver, colon, and skin

[Specific tissue/organ] Chemical	MN in erythrocyte		MN in the specific tissue/organ	
	Result	Reference	Result	Reference
[Liver]				
2-Acetylaminofluorene	+	[76]	+/−	[50,51,54,77]
4-Acetylaminofluorene	−	[78]	−	[51]
Acrylamide	I	[79]	+	[49]
4-Aminobiphenyl	+	[79]	+	[48]
Amsacrine (*m*-AMSA)	+	[50]	+	[50]
o-Anthranilic acid	−	[78]	−	[77]
L-Ascorbic acid	−	[81]	−	[77]
Auramine O	−	[79]	+	[49]
Benzene	+	[76]	−	[50,80]
e-Caprolactam	−	[81]	−	[77]
Carbon tetrachloride	−	[79]	−	[80]
Clofibrate			+	[49]
4-Chloro-*o*-phenylenediamine (CPDA)	+	[33]	−	[41]
Cyclophosphamide	+	[76]	+/−	[50,54,77]
2,4-Diaminotoluene	I	[79]	+	[77]
Di(2-ethylhexyl)phthalate (DEHP)	−	[79]	−	[41]
Diethylnitrosamine	−	[79]	+	[46–48,50,54,56]
p-Dimethylamino-azobenzene (DAB)	−	[79]	+	[41]
4-Dimethylamino-3′-methylazobenzene	−	[50]	+	[50]
6-Dimethylaminophenylazo-benzthiazole	−		+	[51]
7,12-Dimethylbenz[*a*]anthracene	+	[76]	−	[54]
Dimethylnitrosamine	+	[76]	+	[41,46–49,51]
	−/+	[73,74]		
1,1-Dimethylhydrazine	+	[79]	+	[48,77]
1,2-Dimethylhydrazine	+	[79]	+	[51]
2,4-Dinitrotoluene	−	[75]	+	[77]
1,4-Dioxane	I/−	[79,82]	+	[82]
Ethylmethanesulfonate	+	[76]	−	[53]
Ethylnitrosourea	+	[76]	+	[47,50]
5-Fluorouracil	+	[76]	+	[50]
Kojic acid	+	[41]	−	[41]
D-Mannitol	−	[81]	−	[77]
Methoxychlor	−	[77]	−	[77]
4,4′-Methylenedianiline (MDA)	−	[41]	−	[41]
Methylmethanesulfonate	+	[76]	−	[41,47]
Mitomycin C	+	[76]	+	[50,54,77]
2-Nitrofluorene	−	[54]	+	[54]
N-Nitrosomorpholine	+	[79]	+	[83]
Phenazopyridine hydrochloride	+	[79]	+	[77]
Potassium chromate (IV) (K_2CrO_4)	+	[76]	+	[50]
β-Propiolactone	−	[79]	+	[48]
4-*N*-Pyrrolindinylazobenzene			−	[51]
Quinoline	I	[41]	+	[41]
Styrene oxide	−	[79]	+	[49]

Table D.1 (*Continued*)

[Specific tissue/organ] Chemical	MN in erythrocyte		MN in the specific tissue/organ	
	Result	Reference	Result	Reference
Selenious acid (H_2SeO_3)	+	[50]	+	[50]
o-Toluidine	+	[41]	–	[41]
[Colon epithelium]				
Carbendazim	+	[84–86]	+	[87]
	–	[87]		
Colchicine	+	[76,87–91]	+	[87]
Cyclophosphamide	+	[76,89,92,93]	+	[93]
1,2-Dimethylhydrazine	+	[79]	+	[42,87,93–96]
	–	[87,92,93,96]		
Griseofulvin	–	[79,87]	+	[87]
Methylnitrosourea	+	[79]	+	[42]
Mitomycin C	+	[76,89]	+	[42]
Tubulazole	+	[87,89,91]	+	[87]
[Skin]				
Acetone	–	[97]	–[a]	[98,99]
			–	[43]
2-acetylaminofluorene (2-AAF)	+	[100]	+	[101]
Anthracene			–	[102]
Benz[*a*]anthracene	I	[79]	+	[102]
Benzene (BEN)	+	[76]	–	[101]
Benzo[*a*]pyrene	+	[76]	+[a]	[44,98,102,103]
Benzo[*e*]pyrene			–	[102]
Catechol	+	[109]	–	[101]
Chrysene			+[a] –	[98,102]
Clinafloxacin(CLFX) + UVA			+ (oral)	[104]
Colchicine (COL)	+	[76]	+	[101]
Cyclophosphamide	+	[76]	–	[101]
			+	[105]
Dibenz[*a,c*]anthracene			+	[102]
Dibenz[*a,h*]anthracene	+	[79]	+	[102]
Dichlorvos	–	[79]	+[a]	[106]
Diethylstilbesterol (DES)	–	[76]	–	[101]
15,16-Dihydrocyclopenta[*a*]-phenanthrene-17-one			–	[103]
15,16-Dihydro-11-methylcyclo-penta[*a*]phenanthrene-17-one			+	[103]
7,12-Dimethylbenz[*a*]anthracene	+	[76]	+[a]	[98,99,107]
			+	[44,102]
1,2-Dimethyl hydradine dihydrocloride	+	[79]	+	[101]
17-β-Estradiol	–	[79]	–	[101]
2-Ethyl-1,3-hexanediol			–	[105]
N-Ethyl-*N'*-nitro-*N*-nitrosoguanidine (ENNG)	+	[108]	+	[44,105]
5-Fluorouracil (5-FU)	+	[76]	+	[101]
Hydrazine HCl (HDZ)	+	[79]	–	[101]
Hydroquinone	+	[109]	–	[101]
o-Hydroxybiphenyl			–	[101]
Levofloxacin (LVFX) + UVA			– (oral)	[104]
Lomefloxacin (LFLX) + UVA			+ (oral)	[104]
6-Mercaptopurine	+	[76]	+	[101]
3-Methylcholanthrene			+	[102]
Methyl methanesulfonate	+	[76]	+[a]	[99]
			+	[43,101]
N-Methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG)	+	[76]	+	[44]
Mitomycin C	+	[76]	+	[43,105]

Table D.1 (*Continued*)

[Specific tissue/organ] Chemical	MN in erythrocyte		MN in the specific tissue/organ	
	Result	Reference	Result	Reference
Nickel chloride ($NiCl_2$)	–	[79]	I	[101]
p-Nitrophenol			–	[105]
4-Nitroquinoline 1-oxide (4NQO)	+	[110]	+	[44]
N-Nitrosodiethylamine (DEN)	–	[79]	–	[101]
	+ (Liver)	[56]		
Pyrene			–	[98,102]
TPA			–	[99]
TPA + mezerein			+	[111]
Trichloroethylene			–	[105]
Triethylenemelamine	+	[76]	+	[112]
Urethane	+	[76]	+	[98]
Vinblastine	+	[113]	+	[101,105]

Abbreviations: TPA: Phorbol-12-myristate-13-acetate; +: positive, –: negative, I: inconclusive (i.e., there were contradictory results from different laboratories).

[a] In vivo/in vitro method.

Table D.2
Micronucleus assay results on spermatids

Compound	Method[a]	Species tested[b]	No. of doses	No. of animals	No. of SPD/animal	Stages tested	Results	Reference
Acrylamide	S	M	3	4–5	1000–2000	Diakinesis-MI/MII	–	[114]
						Diff. SPG-preleptotene	+	
Acrylamide	D	R	3	5	2000	Diplotene diakinesis	–	[115]
						Late pachytene	–	
						Preleptotene	+	
Acrylamide	S	R	3	5	1000	DIAKINESIS	–	[116]
						Diplotene	–	
						Diff. SPG/preleptotene	+	
Butadiene	S	M	3	5	2000	Late meiotic prophase	–	[117]
						Diff. SPG/prelept./early prophase	+	
Butadiene	S	M	3	3–5	1000	Late meiotic prophase	–	[118]
						Prelept./early prophase	–	
						Diff. SPG/preleptotene	+	
Butadiene dioloepoxide	D	R	2	4–5	2000	Diplotene diakinesis	+	[119]
						Late pachytene	+	
						Preleptotene	+	
Carbendazim	D	R	3	6	1500	Stage I	+	[120]
						Stage V	+	
						Stage VII	–	
Chloral hydrate	S	M	3	5	1000	Diakinesis-MI	–	[121]
						Leptotene-zygotene	–	
						Preleptotene	+	
						Stem cell SPG	+	
	S	M	1	2–3	2000	Pachytene-diakinesis	–	[122]
						Preleptoten	–	
						Stem cell SPG	+	
1,2,3,4-Diepoxybutane	S	M	2	5	2000	Diakinesis/meiotic divisions	+	[123]
						Diff. SPG/preleptotene	+	
1,2,3,4-Diepoxybutane	S	M	2	4–5	1000	Diakinesis	–	[118]
						Diplotene	+	
						Zygotene	–	
						Diff. SPG-preleptotene	–	

Table D.2 (*Continued*)

Compound	Method[a]	Species tested[b]	No. of doses	No. of animals	No. of SPD/animal	Stages tested	Results	Reference
1,2,3,4-Diepoxybutane	D	R	2	4–5	2000	Diplotene diakinesis	+	[119]
						Late pachytene	–	
						Preleptotene	+	
						Stem cell SPG	+	
1,2,3,4-Diepoxybutane	S	R	4	4–5	1000	Diakinesis	–	[118]
						Diplotene	+	
						Zygotene	+	
						Diff. SPG-preleptotene	+	
3,4-Epoxybutene	S	M	3	5	2000	Diakinesis/MI/MII	–	[123]
						Diff. SPG/preleptotene	+	
3,4-Epoxybutene	S	M	2	5	1000	Diakinesis	–	[118]
						Diplotene	–	
						Diff. SPG-preleptotene	+	
3,4-Epoxybutene	D	R	5	4–5	2000	Diplotene diakinesis	+	[119]
						Late pachytene	+	
						Preleptotene	+	
						Stem cell SPG	–	
3,4-Epoxybutene	S	R	2	4–5	1000	Diakinesis	+	[118]
						Diplotene	+	
						Diff. SPG-preleptotene	+	
Etoposide	D	R	2	2–5	1000	Diplotene diakinesis	+	[115]
						Late pachytene	+	
						Preleptotene	+	
	D	M	1	3	1000	Diplotene-diakinesis	+	[124]
Merbarone	D	M	2	3	1000	Diplotene-diakinesis (CREST)	+	[124]
						Prometaphase-MI	–	
						Prometaphase-MII	–	
Potassium bromate	S	M	3	10	2000	Spermatogonial stem cells–spermatogomia/ spermatocytes	–	[125]
Trichloroethylene	S	M	3	6	1000	Preleptotene-early pachytene	+	[121]
Trophosphamide	S	M	2	4–5	1000	Diakinesis/MI/MII	–	[126]
						Diff. SPG-preleptotene	+	
Trophosphamide	D	R	2	5	2000	Dipl.diak.	+	[127]
						Late pachytene	+	
						Preleptotene	+	
Merbarone	D	M	2	3	1000	Diplotene-diakinesis	+	CREST
						Prometaphase-MI	–	[124]
						Prometaphase-MII	–	
VP-16	D	M	1	3	1000	Diplotene-diakinesis	+	CREST [124]
Carbendazim	D	R	3	6	1500	Stage I	+	CREST
						Stage V	+	[120]
						Stage VII	–	
Potassium bromate	S	M	3	10	2000	Spermatogonial stem cells–spermatogomia/ spermatocytes	–	In drinking water for 8 weeks, antikinetochore [125]
Chloral hydrate	S	M	1	2–3	2000	Pachytene-diakinesis	–	CREST, FISH
						Preleptoten	–	[128]
						Stem cell SPG	+	

[a] S: suspension method, D: dissection method.
[b] M: mouse, R: rat.

Table D.3
Micronucleus assay results using feta/neonatal animals by transplacental treatment

Reference	Species/tissue	Test agent	Treatment regimen	Results/comments
[129]	Mouse; maternal bone marrow, fetal liver, and fetal blood	γ-Radiation, MMS, procarbazine, MMC, B[*a*]P	i.p. injection during 15–16th day of pregnancy	Increased MN in maternal bone marrow, fetal liver and fetal blood by all test agents
[130]	Mouse; maternal bone marrow, fetal liver, neonatal blood	DEN	i.p. injection during 15–16th day of pregnancy	Increased MN in fetal liver and neonatal blood; no increase in MN in maternal bone marrow
[131]	Mouse; maternal bone marrow, and fetal liver	γ-Radiation, MMS, procarbazine, MMC, B[*a*]P, cyclophosphamide, DEN	i.p. injection on 16th day of pregnancy	Increased MN in maternal bone marrow, and fetal liver by all test agents, except DEN which was positive only in fetal liver; fetal liver was relatively more sensitive
[132]	Mouse; maternal bone marrow, fetal liver, and fetal blood	Cyclophosphamide, triethylenemelamine, methotrexate	Various gestation days	Fetal liver MN test is useful to detect mutagens and embryotoxicity
[133]	Mouse; maternal bone marrow and fetal liver	MNNG, cyclophosphamide, methotrexate, 5-fluorouracil, 6-mercaptopurine, triethylene-melamine	Various routes on 13th day of pregnancy and/or 5 times daily	Increased MN in maternal bone marrow, and fetal liver by all test agents, except MNNG, which was positive only in fetal liver; fetal liver was relatively more sensitive
[134]	Rat; maternal bone marrow, embryonic blastocytes	Chlorambucil	i.p. injection on gestation days 1–3	Increased MN in maternal bone marrow and embryonic blastocytes; the latter being more sensitive
[135]	Mouse; fetal liver	Cigarette smoke	Inhalation on 15th and 16th day of gestation, long-term exposure on premating through mating and/or 16th day of gestation	No increased MN in fetal liver cells; but a sister chromatid exchange assay was positive under these conditions
[136]	Mouse and rat; maternal bone marrow, fetal liver, fetal blood	MMC, DMBA	MMC: ip on 12th to 17th day for mouse, DMBA: oral on 11th to 16th day for mouse and 13th to 18th day for rat	Increased MN in all tissues studied, with fetal blood cells being more sensitive
[137]	Mouse; maternal bone marrow and fetal liver	B[*a*]P	Various gestation days	Increased MN in both tissues studied, with fetal liver cells being more sensitive
[138]	Mouse; maternal bone marrow and fetal liver	4-Nitroquinoline 1-oxide	s.c. dosing on 12th, 14th, or 16th day of gestation	Increased MN in both tissues studied, with fetal liver cells being more sensitive based on equal adduct levels
[139]	Mouse; maternal bone marrow, fetal liver, and fetal blood	Benzene	i.p. injections on 14th day of gestation	Increased MN in all tissues studied, with fetal liver cells being more sensitive
[140]	Mouse; maternal bone marrow and fetal liver	Benzidine, heliotrine, monocrotaline, urethane	i.p. injection on 17th–19th days of gestation	Increased MN in both tissues studied, with fetal liver cells being more sensitive
[141]	Mouse; fetal blood	MMS, heliotrine, lasiocarpine, monocrotaline, B[*a*]P	i.p. or p.o. on day 15 or 16 of gestation	Increased MN in fetal blood erythrocytes, both young (uniformly-stained RNA) and older (stippled) RNA-positive erythrocytes
[142]	Mouse; maternal bone marrow, fetal liver, and fetal blood	Cyclophosphamide, DEN, *N*-nitroso-*N*-ethylurea, MMC	i.p. injections on 15th or 16th day of gestation	Increased MN in maternal bone marrow, fetal liver, and fetal blood by all test agents, except DEN, which was positive only in fetal liver and fetal blood