添付資料３

Title: Report of the JaCVAM initiative international validation study of the *in vivo* rodent alkaline Comet assay for the detection of genotoxic carcinogens: the 4th (definitive) phase-2nd step

Issued: Yoshifumi Uno, D.V.M., Ph.D., a Validation Management Team (VMT) member
Date: March 21, 2012
Status: Draft Version-1

Notes: this document is prepared to summarize the *in vivo* Comet assay validation process and results in the 4th (definitive) phase-2nd step. The methods are mentioned minimally in this document, because the details are described in the study protocol and the study plan. An article for submission to a scientific journal will be provided separately based on this document, the study protocols and the study plans (and the other documents if necessary and available).

A Table of Contents

1. Introduction

An *in vivo* rodent alkaline Comet assay is a genotoxicity assay to detect DNA damage in almost all tissues theologically, and is practically used worldwide for detecting genotoxic chemicals. The assay is applied to some regulatory aspects for investigation of genotoxic potential of test chemicals, and currently expected as a promising second *in vivo* genotoxicity assay in the ICH-S2(R1) guidance behind *in vivo* micronucleus assay with bone marrow and/or peripheral blood. The methods of Comet assay were often discussed in the meetings of the International Workshop on Genotoxicity Testing (IWGT) and the International Comet Assay Workshop (ICAW), and the consensus articles have been published on the methodology (1-3).

The assay, however, has not been validated formally with a standardized study protocol. In addition, since the reports on predictive capability of *in vivo* rodent Comet assay for carcinogenicity are limited (4-6), the investigation of predictive capability in multiple laboratories using one study protocol validated would be more useful to understand the accurate performance of *in vivo* Comet assay. The Japanese Environmental Mutagen Society/the Mammalian Mutagenicity Study Group (JEMS/MMS) decided to have a (international) collaborative study of *in vivo* Comet assay in 2003, and conducted a preliminary collaborative study on Comet assay procedure, e.g., comparison of assay results between whole cells and isolated nuclei (7). Contemporaneously, the UK Environmental Mutagen Society/the Industrial Genotoxicity Group had a concern about the establishment of OECD guideline on Comet assay (Dr. Brian Burlinson, personal communication). To cover all the concerns about validation efforts of *in vivo* Comet assay, the Japanese Center for the Validation of Alternative Methods (JaCVAM) is organizing the international validation study in and after April, 2006, in cooperation with the U.S. National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), the European Centre for the Validation of Alternative Methods (ECVAM), and JEMS/MMS.

The purpose of this validation study is to evaluate the ability of the *in vivo* Comet assay to identify genotoxic chemicals as a potential predictor of rodent carcinogenicity, and as an alternative follow-up assay to more commonly used *in vivo* rodent Unscheduled DNA Synthesis (UDS) assay. The goal of this validation effort is to establish the OECD guideline on *in vivo* rodent alkaline Comet assay.

2. Background and Purpose

In the 1st step of the 4th phase validation study, the purpose was to examine the extent of reproducibility and variability of *in vivo* Comet assay results among laboratories using coded test chemicals and a positive control ethyl methanesulfonate (EMS), when experiments were conducted in accordance with the Comet assay protocol-version 14. In the review of study data, the Validation Management Team (VMT) confirmed the reproducibility and variability of assay results among laboratories. Thus VMT decided to move on the 2nd step of 4th phase validation study with an expanded set of test chemicals in accordance with the Comet assay protocol-version 14.2 (Appendix 1).

The purpose of the 2nd step is to investigate the predictive capability of the assay against carcinogenicity of test chemicals (see the study plan: Appendix 2).

3. Experimental Period

December, 2009 - February, 2012

4. Participant Laboratories

Fourteen laboratories* participated in the 2nd step of 4th phase validation study, which include four leading laboratories[#] that have a lot of experiences of Comet assay and join the 1st to 3rd phase validation studies, and ten laboratories that passed our recruitment process for this 4th phase validation study.

* Merck Research Laboratories[#] (code: Lab B), BioReliance[#] (Lab C), Huntingdon Life Sciences[#] (Lab D), Food and Drug Safety Center[#] (Lab E), The Institute of Environmental Toxicology (Lab F), Novartis Pharma (Lab G), AstraZeneca (Lab H), Sumitomo Chemical (Lab I), Mitsubishi Chemical Medience (Lab J), Janssen R&D (Lab K), Health Canada (Lab L), Covance (Lab M), Bayer Schering Pharma (Lab N), and Integrated Laboratory System (Lab O).

5. Success Criteria in the Study Plan (Appendix 2)

To obtain the predictive capability (values of positive sensitivity and negative specificity) of the assay against carcinogenicity of test chemicals. VMT discussed at Salt Lake meeting held on March 12, 2010, whether or not expected positive sensitivity and negative specificity should be set as success criteria before starting this step of validation study, and concluded that it was unnecessary because such values would be calculated resultantly after the validation study.

6. Materials and Methods

In this section, outlines of the materials and methods are described. The details are referred to the validation study protocol version 14.2 (Appendix 1), and the study reports written in each laboratory (Appendix 4).

An individual study protocol was prepared in each laboratory in accordance with the validation study protocol v.14.2. The experiments proceeded in each facility based on their own study protocol and SOP.

6-1. Animal species, strain, and sex

Rats were selected in this validation effort because of popular species in toxicology studies. Crl:CD(SD) male rats were used.

6-2. Test chemical, vehicle, and dose level

Forty coded test chemicals were used in this study (Table 1). The reasons for selection of those test chemicals are mentioned in the chemical selection report written by Dr. Takeshi Morita (Appendix 3). Briefly, test chemicals were selected from four categories based on their genotoxicity and carcinogenicity properties, i.e., genotoxic carcinogen, genotoxic non-carcinogen, non-genotoxic carcinogen, and non-genotoxic non-carcinogen. Genotoxicity is defined as a positive result in Ames test or standard *in vivo* genotoxicity test such as bone-marrow micronucleus assay. Carcinogenicity is defined as positive results in rodent (rat and/or mouse) carcinogenicity study, or known human carcinogens. Test chemicals include organic and inorganic chemicals, and have many types of mode of action for genotoxicity and/or carcinogenicity, e.g., base-alkylation, aneugenic effects, bulky adduct formation, cross-link formation, epoxide formation, nucleoside analog, cytotoxicity, and peroxisome proliferation.

Test chemicals were randomly coded in JaCVAM (Table 1), and sent to an assigned chemical master of each testing facility, who was independent of the validation study. Each coded test chemical was basically evaluated in one laboratory (Table 1), because VMT considered that the reproducibility and variability of assay results had been already confirmed among laboratories using the four coded test chemicals and the positive control EMS in the 1st step of 4th phase validation study. The vehicle for a coded test chemical was appropriately selected in each testing facility (Table 1). The dose levels were also decided in each facility based on the results of preliminary dose-finding study designed in each laboratory (Appendix 4), although VMT provided some toxicological information such as the LD_{50}, if available from

published literatures, in order to assist the dose selection process in each laboratory. Exceptionally, the dose levels and vehicles of three chemicals, A4204, A4206 and A4219 were directed by VMT because of the retests in another laboratory (Table 1).

Table 1 Test chemical code, test chemical name, category, laboratory tested, vehicle, and dose levels

Test chemical code	Test chemical name (CASRN)	Category of genotoxicity and carcinogenicity	Lab tested (coded lab name)	Vehicle	Dose level (mg/kg/day)
A4114	2-Acetylaminofluorene (53-96-3)	Genotoxic carcinogen	Lab O	Corn oil	250, 500, 1000
A4201	1,3-Dichloropropene (542-75-6)	Genotoxic carcinogen	Lab C	Corn oil	50, 100, 200
A4202	Ethionamide (536-33-4)	Non-genotoxic non-carcinogen	Lab C	Corn oil	125, 250, 500
A4203	Buslfan (55-98-1)	Genotoxic carcinogen	Lab C	Corn oil	10, 20, 40
A4204	*N*-Nitrosodimethylamine (62-75-9)	Genotoxic carcinogen	Lab L	Saline	2.5, 5, 10
			Lab O *	Saline *	0.63, 1.25, 2.5 *
A4205	Ampicillin trihydrate (7177-48-2)	Non-genotoxic non-carcinogen	Lab L	Saline	25, 50, 100
				Corn oil	500, 1000, 2000
A4206	1,2-Dimethylhydrazine dihydrochloride (306-37-6)	Genotoxic carcinogen	Lab L	Saline	6.25, 12.5, 25
			Lab O *	Saline *	1.56, 3.13, 6.25 *
A4207	Isobutyraldehyde (78-84-2)	Non-genotoxic non-carcinogen	Lab B	Corn oil	500, 1000, 2000
A4208	Cisplatin (15663-27-1)	Genotoxic carcinogen	Lab B	0.5% CMC	6, 12.5, 25
A4209	Azidothymidine (30516-87-1)	Genotoxic carcinogen	Lab B	0.5% CMC	500, 1000, 2000
A4210	*p*-Dichloroaniline (106-47-8)	Genotoxic carcinogen	Lab D	Corn oil	37.5, 75, 150
A4211	*t*-Butylhydroquinone (1948-33-0)	Non-genotoxic non-carcinogen	Lab D	Corn oil	131.3, 262.5, 525
A4212	Methyl carbamate (598-55-0)	Non-genotoxic carcinogen	Lab D	Saline	500, 1000, 2000

A4213	Methyl methanesulfonate (66-27-3)	Genotoxic carcinogen	Lab G	Saline	20, 40, 80
A4214	2,6-Diaminotoluene (823-40-5)	Genotoxic non-carcinogen	Lab G	Corn oil	150, 300, 600
A4215	5-Fluorouracil (51-21-8)	Genotoxic non-carcinogen	Lab G	Saline	25, 50, 100
A4216	8-Hydroxyquinoline (148-24-3)	Genotoxic non-carcinogen	Lab N	Corn oil	125, 250, 500
A4217	Hydroquinone (123-31-9)	Genotoxic carcinogen	Lab N	Saline	125, 250, 500 [1)]
A4218	Saccharin (81-07-2)	Non-genotoxic carcinogen	Lab N	Corn oil	500, 1000, 2000
A4219	Sodium arsenite (7784-46-5)	Genotoxic carcinogen	Lab M	Saline	7.5, 15, 30
			Lab O *	Saline *	7.5, 15, 30 *
A4220	Thioacetamide (62-55-5)	Non-genotoxic carcinogen	Lab M	Saline	19, 38. 75
A4221	Diethanolamine (111-42-2)	Non-genotoxic carcinogen	Lab M	Saline	175, 350, 700
A4222	*p*-Phenylenediamine dihydrochloride (624-18-0)	Genotixic non-carcinogen	Lab K	Saline	25, 50, 100
A4223	*o*-Phenylphenol sodium salt (132-27-4)	Non-genotixic carcinogen	Lab K	Corn oil	250, 500, 1000
A4224	2,4-Diaminotoluene (95-80-7)	Genotoxic carcinogen	Lab K	Saline	100, 150, 200
					37.5, 75, 150
A4225	4,4'-Oxydianiline (101-80-4)	Genotoxic carcinogen	Lab H	0.5% CMC	50, 100, 200
A4226	*o*-Anisidine (90-04-0)	Genotoxic carcinogen	Lab O	Corn oil	150, 300, 600
A4227	Sodium chloride (7647-14-5)	Non-genotoxic non-carcinogen	Lab O	Water	500, 1000, 2000
A4228	Acrylonitrile (107-13-1)	Genotoxic carcinogen	Lab E	Corn oil	15.7, 31.3, 62.5
A4229	9-Aminoacridine hydrochloride monohydrate (52417-22-8)	Genotoxic non-carcinogen	Lab E	Corn oil	15.7, 31.3, 62.5
A4230	Ethanol	Non-genotoxic	Lab E	Saline	500, 1000, 2000

	(64-17-5)	carcinogen			
A4231	1,2-Dibromomethane (106-93-4)	Genotoxic carcinogen	Lab J	Corn oil	25, 50, 100
A4232	*p*-Anisidine (104-94-9)	Genotoxic non-carcinogen	Lab J	0.5% CMC	125, 250, 500
A4233	*o*-Anthranilic acid (118-92-3)	Non-genotoxic non-carcinogen	Lab J	0.5% CMC	500, 1000, 2000
A4234	Benzene (71-43-2)	Genotoxic carcinogen	Lab I	Corn oil	500, 1000, 2000
A4235	Di(2-ethylhexyl)phthalate (117-81-7)	Non-genotoxic carcinogen	Lab I	Corn oil	500, 1000, 2000
A4236	Trisodium EDTA monohydrate (10378-22-0)	Non-genotoxic non-carcinogen	Lab I	Saline	500, 1000, 2000
A4237	Cadmium chloride (10108-64-2)	Genotoxic carcinogen	Lab F	Saline	20, 40, 80
A4238	Chloroform (67-66-3)	Non-genotoxic carcinogen	Lab F	Corn oil	125, 250, 500
A4239	D,L-Menthol (15356-70-4)	Non-genotoxic non-carcinogen	Lab F	Corn oil	125, 250, 500

* The vehicle and dose levels were directed by VMT, because those chemicals were retested in another laboratory due to the reasons described in the section 7-2.

6-3. Positive control

EMS has been used as a positive control in a sequence of this validation effort, because it is a well-known genotoxic chemical for multiple organs. EMS was dissolved in physiological saline, and administered to rats at the dose level of 200 mg/kg twice (21 hr interval) with oral administration.

6-4. VMT consensus about expected assay results for test chemicals

It would be necessary to have expected assay results for each category of four classes of test chemicals (i.e., genotoxic carcinogen, genotoxic non-carcinogen, non-genotoxic carcinogen, and non-genotoxic non-carcinogen) before review of assay results in order to avoid any bias for the evaluation of assay results. VMT discussed and decided expected assay results at Salt Lake meeting held on March 12, 2010, as follows.

6-4-1. Genotoxic carcinogen: positive results will be expected in the liver and/or

the stomach. Target organ specificity of carcinogenicity may be considered for the interpretation of negative results, but it should be minimized because our validation study protocol is designed for screening purpose against carcinogenicity without consideration of the target organ specificity.

6-4-2. Genotoxic non-carcinogen: negative results will be preferred in both the liver and the stomach, but positive results will be acceptable because this category of chemicals is considered to have genotoxic activity essentially.

6-4-3. Non-genotoxic carcinogen: negative results will be preferred in both the liver and the stomach, but positive results will be acceptable because this category of chemicals may have some genotoxic mode of actions for carcinogenicity. But, in case of positive response of known cytotoxic agents, more careful consideration will be required for the interpretation of positive results.

6-4-4. Non-genotoxic non-carcinogen: negative results will be expected.

The above were internal consensus of VMT, and not disclosed to testing facilities, because it was necessary to keep secret about what categories of test chemicals were selected and used in this validation study.

6-5. Administration of test chemical to animals

Each coded test chemical was administered to rats at the three dose levels in three times (24 and 21 hours interval) by oral gavage. This administration regimen was designed to allow the combination genotoxicity assay of micronucleus and Comet, in consideration of the 3R's principle for animal use, although investigation into micronucleus induction was optional and data on micronucleus are not included in this validation study (micronucleus data may be included in study reports written by testing facilities: Appendix 4).

6-6. Organs analyzed

The liver and the stomach (glandular stomach) are selected in this validation effort, because the former is the primary organ for the metabolism of absorbed chemicals, and the latter is a site of first contact of chemicals in orally administration. These organ analyses are recommended for screening purpose of genotoxic chemicals in the previous discussion in ICAW (2).

6-7. Data-acceptance criteria

Data-acceptance criteria were determined based on the 1st to 3rd phase validation study results.

6-7-1. Negative (vehicle) control

Means of %DNA in tail are 1-8% in the liver and 1-30% (preferably 1-20%) in the stomach.

6-7-2. Positive control (EMS)

Effect (difference of means of % DNA in tail between groups of EMS and vehicle control) is statistically significant increases (see the section 6-8.) and is 5% or higher in the liver and the stomach. Those are primary criteria for data-acceptance.

6-8. Data analysis of % DNA in tail

% DNA in tail was used as the primary endpoint of this validation study, because it is considered linearly related to the DNA break frequency over a wide range of damaged DNA levels (2). The other parameter such as tail moment and tail length may be calculated and reported in the study reports written in testing facilities (Appendix 4), but no statistical analysis was applied to such parameter in this validation study report.

Three conceptual key terms, "Endpoint", "Estimate" and "Effect" were defined and used in the data analysis of this validation study. Briefly, Endpoint is defined as individual observed values for a parameter such as % DNA in tail. Estimate is defined as a mean calculated with values of Endpoint in each animal. Effect is defined as difference (hereafter designated as Effect (diff.)) or ratio (hereafter designated as Effect (ratio)) of a mean of Estimate between a negative control group and a treatment group. A general purpose of data analysis in validation studies is to investigate how large variation exists among data from several testing facilities, and Effect is considered as a good indicator to understand the variation of Comet assay parameters among testing facilities. VMT noticed through the 1st to the 3rd phase validation studies that Effect (diff.) was more effective for the comparison of variation than Effect (ratio), because Effect (ratio) depended on the magnitude of negative control values (i.e. lower negative control values easily produced higher Effect (ratio)) and would be often misread in the evaluation of responses induced with a test chemical administration. Therefore Effect (diff.) was used to evaluate the assay results.

Dunnett's test (two-sided, $P<0.05$) and linear Trend test (two-sided, $P<0.05$) were applied to Effect (diff.) in the groups of coded test chemicals. The two-sided analysis was used because both increases and decreases in the Comet parameter

could be detected. A decreased % DNA in tail was expected to be a good index to detect cross-linking agents. For the positive control group, Student's t-test (one-sided, $P<0.025$) was applied to the Effect (diff.).

6-9. Histopathology

Regarding the evaluation of cytotoxicity, the "Gold Standard" for assessing levels of necrosis and apoptosis when an *in vivo* Comet assay gave positive results is concluded to be histopathology, which is a consensus of IWGT meeting in 2005 (3), and it is also pointed out that there is a need to standardize ways to present histopathological findings in the meeting (3). In this validation study, when increased % DNA in tail was observed in the liver and/or stomach, the organ(s) was histopathologically examined by pathologists of testing laboratory, except for retests of A4204, A4206 and A4219 which were examined by pathologists in Biosafety Research Center (BSRC).

Based on the IWGT consensus and suggestion (3), VMT decided that histopathology results should be primarily considered to interpret a relationship between positive findings in Comet assay and cytotoxicity of test chemicals. Increased % hedgehog was another candidate for cytotoxicity evaluation, but it was not adopted in this validation study although the data were collected, because increased % hedgehog would indicate not only cellular toxicity but also DNA damage.

In histopathological examination, VMT considered that necrosis and/or degeneration (or findings indicating degenerative changes) would be main indicators of cellular toxicity. In contrast, apoptosis would indicate both cellular toxicity and DNA damage. Thus necrosis and/or degeneration were mainly focused for interpretation of positive findings in Comet assay.

To decide standardized ways to present histopathological findings, VMT firstly reviewed histopathological data of test chemicals showing negative results in Comet assay. In most of negative compounds, there seemed no histopathological finding indicating cytotoxicity related to chemical treatment, except for the following cases.

1) A4215: single cell necrosis in glandular stomach mucosa (grade: minimal; incidence: two of five (2/5) rats at 25 mg/kg, and 5/5 rats at 50 and 100 mg/kg)
2) A4217: single cell necrosis in hepatocytes (minimal to slight; 5/5 rats at 500 mg/kg)
3) A4235: eosinophilic degeneration/necrosis in hepatocytes (slight; 5/5 rats at 2000 mg/kg)

4) Histopathology of stomach in many chemicals (e.g., A4223, A4232, A4233, ···) showed ulcer/erosion in the glandular stomach or forestomach, and the grade of lesions was minimal to slight (or mild). Since these findings are often observed in rats incidentally by stress, they may not always indicate cytotoxicity.

The above findings indicate that minimal to slight (or mild) histopathological changes related to cytotoxicity would not affect increased % DNA in tail in the liver and the stomach. Therefore, VMT decided that the grade and incidence of the histopathological findings should be considered for the interpretation. For example, when one of five animals showed minimal to slight (or mild) necrosis and/or degenerative findings in a target organ, such weak changes would not affect Comet assay results. In addition, study directors and/or pathologists in most of participant laboratories did not consider that single cell necrosis or single cell death with grade of minimal to slight (or mild) would affect Comet assay results. VMT has agreed to the interpretation considered in those laboratories.

The above criteria were presented to participants in Kyoto meeting held on September 12-13, 2011, and accepted. Based on those criteria, VMT and the meeting-participants interpreted the significance of positive findings in Comet assay under the blind-condition of chemical names (Appendix 6).

6-10. % hedgehog

The frequency of hedgehogs, % hedgehog was determined per sample, based on the visual scoring of 100 cells per sample. No statistical analysis was applied to % hedgehog.

7. Results

In this section, outlines of the study results are described. The details are referred to the study reports written in each laboratory (Appendix 4) and the statistical analysis results by Dr. Takashi Omori (Appendix 5).

7-1. Control groups

Figs. 1 and 2 summarize the lab-orderly means of % DNA in tail (Estimate) in the vehicle and positive control EMS groups in the liver and the stomach, respectively.

7-1-1. Vehicle control groups

Figs. 3 and 4 are enlarged-figures of the lab-orderly Estimate in the vehicle control

groups in the liver and the stomach, respectively. All the values in the stomach satisfied the preferable data-acceptance criteria 1-20%, and there seems no tendency within/between-laboratory variability. Those in the liver also met the data-acceptance criteria 1-8% except for one experiment with a coded test chemical A4236 in Lab I (the actual mean value in the liver was 0.7%). In the liver, there seems very slight between-laboratory variability, but no within-laboratory variability, indicating that experiments were well-controlled in each laboratory. In the stomach, there seems no clear within/between-laboratory variability.

Figs. 5 and 6 show the vehicle-orderly Estimate in the vehicle control groups in the liver and the stomach, respectively. In the liver, the Estimate seems no different between four vehicles, water, saline, 0.5%CMC, and corn oil. In the stomach, in contrast, the values seem slightly higher in corn oil compared with the other watery vehicles. <This comparison will be re-discussed with results of statistics later.>

As a deviation from the validation study protocol v.14.2, DNA was not stained with SYBR Gold but with ethidium bromide (EtBr) in Lab M due to careless mistake, i.e., Lab M routinely uses EtBr and then used it in this validation study, too. Thus the control data were separately shown in Fig. 7. Since EtBr is generally used for staining DNA in Comet assay, VMT decided to accept the data in Lab M, although this deviation should be taken into consideration for the data review. The negative control values in both organs satisfied the data acceptance criteria.

7-1-2. Positive control groups

Figs. 7 and 8 show Effect (diff.) of mean %DNA in tail between the vehicle control group and the positive control group in the liver and the stomach, respectively. All of Effect (diff.) show statistically significant increases with Student's t-test (one-sided, $p<0.025$) both in the liver and the stomach, and also showed 5% or higher values. Therefore, it was judged that all the positive control values satisfied the primary data-acceptance criteria. Regarding the magnitude of increased % tail DNA in both organs, there seems very slight between-laboratory variability, but no within-laboratory variability, indicating that experiments were well controlled in each laboratory.

As mentioned in the section 7-1-1., DNA was stained with EtBr, not SYBR Gold in Lab M, and thus the control data were separately shown in Fig. 9. The positive control values in both organs satisfied the data acceptance criteria.

7-2. Justification of test chemical evaluation

From the viewpoint of data-acceptance criteria, VMT considered that all the data on coded test chemicals would be acceptable for the evaluation of predictive capability of *in vivo* Comet assay. Some deviations were noted as mentioned in the section 7-1., but they were considered noncritical and not to affect the evaluation of coded test chemicals with Comet assay.

However, at Huntingdon meeting held on February 7-9, 2011, three deviations from the validation study protocol were pointed out as shown in the following items 1) to 3). In addition, at Kyoto meeting held on September 12-13, 2011, an issue on histopathological analysis was noted as the following item 4). After those meeting, all the issues were solved as mentioned in each item.

1) Test chemical code A4205 should be retested, because no toxicity was observed in the highest dose level of 100 mg/kg. This issue happened due to a misunderstanding about the validation study protocol, i.e., A4205 was soluble up to 10 mg/mL in saline which is the first choice of vehicles in the validation study protocol, and Lab L misunderstood that the solubility limit was one of the reasons to select the highest dose level (since dosing volume was designated as 10 mL/kg, the highest dose level was set at 100 mg/kg). Lab L accepted the request of retest, and reexamined A4205 up to 2000 mg/kg (upper limit dose of the validation study protocol) with corn oil as a vehicle.
2) Test chemical code A4217 should be retested, because no toxicity was observed in the highest dose level of 500 mg/kg. In accordance with the request, Lab N retested A4217 with the dose level of 750 mg/kg. However, animal death was observed at the dose level, indicating that A4217 showed extremely severe toxicity with slight increase in the dose levels. Therefore, 500 mg/kg of A4217 was considered to be the maximum tolerated dose, and the Comet assay data up to 500 mg/kg were justified for the evaluation.
3) Test chemical code A4219 should be retested due to equivocal judgment in the statistical analysis conducted in VMT, although Lab M judged A4219 as positive based on their overall analysis. Since Lab M could not accept the request of retest, Lab O reexamined A4219.
4) Test chemical codes A4204 and A4206 showed increased % DNA in tail in the liver in all dose levels, but histopathology for the liver was only examined at the highest dose levels, and cytotoxic changes were observed. In those cases, histopathology data were needed for the treatment groups of mid and low dose levels, too. VMT asked Lab L the possibility of retests, but Lab L could not accept it. Lab O on behalf of Lab L retested A4204 and A4206.

Finally, from all the points of view to justify this validation study, VMT concluded that all the data on coded test chemicals including results of retests would be acceptable for the evaluation of predictive capability of *in vivo* Comet assay.

7-3. Coded test chemical groups

Table 2 show the summary of results on all test chemicals. The table includes three types of judgment on Comet assay results, i.e., lab judgment, VMT judgment, and final judgment. Lab judgment is that each testing facility calls the Comet assay results as "Positive" or "Negative" based on their own statistical analysis results and histopathological consideration, and some laboratories also consider their own historical control range. VMT judgment is simply conducted based on the statistical analysis results without consideration of histopathology, and it is shown as "Increase", "Decrease", "Equivocal", or "No change" in % DNA in tail. "Increase" means a statistically significant increase in both Dunnett's and linear Trend tests, "Decrease" means a statistically significant decrease in both statistical tests, "Equivocal" means a statistically increase or decrease in either statistical test, and "No change" means no statistical significance in both statistics. Final judgment was decided through the discussion with all participants in Kyoto meeting under a blind-condition with coded test chemicals, based on the consideration of both statistics and histopathology results, and it is shown as "Positive" or "Negative" (or "Equivocal"). When a statistically significant increase in % DNA in tail is noted in either the liver or the stomach, it is simply judged to be "Positive". Appendix 6 is a modified material originally used in Kyoto meeting, which includes the summary data on test chemicals showing increased % tail DNA.

7-3-1. Genotoxic carcinogens

12 of 19 genotoxic carcinogens tested induced a statistically significant increase % tail DNA in the liver and/or the stomach, and VMT judgment for the 12 chemicals was "Increase". In 10 of the 12 chemicals, the final judgments were "Positive", which were coincident with the lab judgment. However, two chemicals, acrylonitrile and thioacetamide were inconsistent between the lab judgment and the final one.

Regarding acrylonitrile, in the first experiment of laboratory, it showed an equivocal response in the liver (a statistically significant increase in linear Trend test, but not in Dunnett's test) and a decrease in the stomach (a statistically significant decrease in both statistics). In the retest for the liver (not done for the stomach), it showed a

statistically significant increase in both statistics in the liver. Based on the statistics, VMT judged the response "Increase" in the liver, and "Decrease" in the stomach. Acrylonitrile, however, was judged "Negative" in the testing facility, because the increased % tail DNA in the liver was within the historical control range of testing facility (note: the decreased % tail DNA in the stomach was not discussed). No histopathological change related to the chemical treatment was found in the liver. In the glandular stomach, edema in lamina propria/submucosal layer was found at 31.3 and 62.5 mg/kg/day (grade: very slight or slight). In forestomach examined at 62.5 mg/kg/day, degeneration/necrosis in squamous cells (grade: moderate), edema in lamina propria/submucosal layer (grade: moderate), degeneration (vacuolar) in muscle fiber of muscular layer (grade: slight), and ulcer (non-graded) were found As the final judgment, acrylonitrile is judged to be "Positive" in the liver in consideration of the reproducible results in two independent experiments and no cytotoxic finding in the liver. Lab judgment in consideration of their own historical control range was considered inappropriate, because data from experiments conducted with the lab-original study protocol were included to establish the historical control rage.

Thiacetamide induced a statistically significant increase % tail DNA in the liver and stomach, and the VMT judgment was "Increase" in both organs. However, it was judged "Negative" in the testing facility, because cytotoxicity was noted in both organs in the histopathological examination and/or increased % hedgehogs (liver). In the histopathology for the liver, hepatocellular degeneration (grade: minimal to moderate), central vein phlebitis (minimal to moderately severe), increased mitosis (minimal or slight), and bile duct hyperplasia (minimal) were found in all treatment groups. In the stomach, no abnormality was noted at 19, 38 mg/kg/day, and erosion/ulcer (one of 6 rats, no-graded), and forestomach gastritis (one of 5 rats, no-graded) were found at 75 mg/kg/day. VMT considered that clear histopathological changes indicating cytotoxicity were observed in the liver, but not in the stomach. Therefore, this chemical is considered to be "Positive" in the stomach as the final judgment.

Five of 19 genotoxic carcinogens tested failed to induce a statistically significant increase or decrease in % tail DNA in the liver and the stomach. One of 19, 4,4'-oxydianiline showed a statistically significant decrease in % DNA in tail in the stomach, but no change in the liver.

Another one, sodium arsenaite induced increases in % tail DNA in the liver, but the statistical significance was only noted with the linear Trend test in Lab M, and with

Dunnett's test in Lab O. This chemical was considered to produce marginal increase in % tail DNA in the liver in two independent experiments, and VMT finally judged this chemical "Equivocal".

In % hedgehog analysis, only two "Positive" chemicals in the liver, 1,2-dimethylhydrazine dihydrochloride and *N*-nitrosodimethylamine, and one "Positive" chemical in the stomach, methyl methanesulfonate clearly increased hedgehogs in the positive organs with dose-dependency, but the other test chemicals including "Positive" chemicals were considered not to clearly affect % hedgehog.

7-3-2. Genotoxic non-carcinogens

One of six genotoxic non-carcinogens tested, 2,6-diaminotoluene induced a statistically significant increase in % tail DNA in the liver, and the VMT judgment was "Increase" in the liver. In histopathology for the liver, no abnormality was noted at 150 mg/kg/day, and mitoses in hepatocyes and hemopoiesis were less frequently or absent at 300 and 600 mg/kg/day. Because of no clear cytotoxicity in the liver, the final judgment was "Positive", which was coincident with the lab judgment. No changes of % hedgehog were noted in all test chemicals in this category.

7-3-3. Non-genotoxic carcinogens

One of seven non-genotoxic carcinogens tested, chloroform induced a significant % tail DNA response in the liver (significant in Dunnett's test in the 250 and 500 mg/kg/day groups) with dose-dependency. In histopathology for the liver, the following chamges were observed: centrilobular hepatocellular single cell necrosis (grade: +), centrilobular hepatocellular necrosis (one of 5 rats, grade: +), centrilobular hepatocellular vacuolation (grade: + or ++), centrilobular hepatocellular enlargement/granular change (grade: + or ++), centrilobular inflammatory cell infiltration (grade: + or ++), and centrilobular hemorrhage (one of 5 rats, grade: ++) at 250 mg/kg/day; and centrilobular hepatocellular single cell necrosis (grade: +), centrilobular hepatocellular necrosis (one of 5 rats, grade: ++), centrilobular hepatocellular vacuolation (grade: + or ++), centrilobular hepatocellular enlargement/granular change (grade: + or ++), centrilobular inflammatory cell infiltration (grade: +), and centrilobular hemorrhage (one of 5 rats, grade: ++) at 500 mg/kg/day. Since serious cytotoxic findings such as moderate grade (++) of hepatocellular necrosis and vacuolation followed by inflammatory cell infiltration and hemorrhage were observed with dose-dependency, the increased % tail DNA in the liver was considered to be related to cytotoxicity, and thus the final

judgment was "Negative", which was coincident with the lab judgment. As a result, all test chemicals in this category were evaluated as "Negative" as the final judgment. No clear changes of % hedgehog were noted in all test chemicals in this category.

7-3-4. Non-genotoxic non-carcinogens

One of eight non-genotoxic non-carcinogens tested, *t*-butylhydroquinone (*t*-BHQ) induced a significant % tail DNA response in the liver, and the VMT judgment was "Increase" in the liver. However, the lab judgment was "Negative", because the increased % DNA in tail was within their historical control range which was established in testing conditions based on the validation study protocol (Brian Burlinson, personal communication). Finally, *t*-BHQ was judged to be "Positive" in the liver based on the statistical analysis results, but VMT decided to note in this validation study report that the chemical was judged to be "Negative" in the testing facility. No changes of % hedgehog were noted in all test chemicals in this category.

8. Discussion

The purpose of the 2nd step of 4th phase international validation study was to investigate the predictive capability of *in vivo* rat alkaline Comet assay against carcinogenicity of test chemicals. In general, the predictive capability of a screening genotoxicity assay for carcinogens is simply described as "positive sensitivity" for carcinogens, and "negative specificity" for non-carcinogens. Such simple approach is, of course, possible, but may not fit the evaluation of predictive capability in this validation study, because we already know that test chemicals have two different aspects of biological/toxicological properties, i.e., genotoxic properties and carcinogenic properties, although it is still possible to calculate the positive sensitivity and the negative specificity only for the carcinogenic properties. The most appropriate approach to evaluate the predictive capability using this validation study results would be to focus the consistency with expected assay results determined before the validation study (see the section 6-4.). Therefore, we discuss the predictive capability of *in vivo* rat alkaline Comet assay from the aspect of four categories of test chemicals, i.e. genotoxic carcinogens, genotoxic non-carcinogens, non-genotoxic carcinogens, and non-genotoxic non-carcinogens.

For genotoxic carcinogens, it was expected to call positive judgment. 12 of 19 genotoxic carcinogens tested were judged "Positive", and six of 19 were "Negative",

and one was "Equivocal". Since 13 of 19 genotoxic carcinogens are considered to show at least alert of positive call, the positive sensitivity for genotoxic carcinogens would be calculated to be 68%. In contrast, for non-genotoxic non-carcinogens, it was expected to give negative judgment, and the negative specificity for non-genotoxic non-carcinogens is calculated to be 88%, since 7 of 8 chemicals were judged to be "Negative". For genotoxic non-carcinogens and non-genotoxic carcinogens, it was preferably expected to call negative. The negative specificity for genotoxic non-carcinogens is calculated to be 83%, since 5 of 6 chemicals were judged to be "Negative". That for non-genotoxic carcinogens is calculated to be 100%, since all (seven) chemicals were judged to be "Negative". Therefore, the concordance for expected assay results is calculated to be 80%, since 32 of 40 test chemicals gave the expected assay results, indicating that *in vivo* rat alkaline Comet assay would have a good capability to identify genotoxic chemicals as a potential predictor of rodent carcinogenicity.

Since this step of validation study is the final stage of a series of validation efforts on *in vivo* rat alkaline Comet assay, it should be referred to another purpose through all phases of the validation studies, i.e., to evaluate the ability of the *in vivo* Comet assay as an alternative follow-up assay to more commonly used *in vivo* rodent Unscheduled DNA Synthesis (UDS) assay. The genotoxic carcinogens used in this study included five UDS-positive chemicals, 2-acetylaminofluorene (2-AAF), 2,4-diaminotoluene, 1,2-dimethylhydrazine dihydrochloride, methyl methanesulfonate and *N*-nitrosodimethylamine, and four UDS-negative chemicals, acrylonitrile, *o*-anisidine, 1,3-dichloropropene and 4,4'-oxydianiline. Four of the five UDS positives were judged to be "Positive" in this study, but 2-AAF was "Negative". Two of the four UDS negatives, acrylonitrile and 1,3-dichloropropene were judged to be "Positive" in this study. Although the number of chemicals tested is limited, VMT concluded that Comet assay could be at least equal or maybe more sensitive to detect genotoxic carcinoges which could be detected by UDS assay.

The above discussions are simply focused on the consistency between Comet assay results in this validation study and genotoxicity/carcinogenicity of test chemicals. To evaluate the predictive capability of Comet assay more accurately, it is clearly needed to consider more detailed biological/toxicological properties of test chemicals, i.e., genotoxic and/or carcinogenic mode of actions (MoA) of test chemicals, especially for genotoxic carcinogens judged to be "Negative" in this

study. For such discussion, it should be taken into account that the liver and the stomach (glandular stomach) were only analyzed in this validation study. The reasons why those organs were selected in this validation study was that our validation study protocol was designed for screening purpose against genotoxicity and/or carcinogenicity without consideration of the target organ specificity. The liver is the primary organ for the metabolism of absorbed chemicals, and the stomach is a site of first contact of chemicals in orally administration, which are recommended for the screening purpose of *in vivo* Comet assay in the previous discussion in ICAW (2). Actually, in this validation study, the following chemicals were judged to be "Positive" in the liver and/or the stomach, although both organs were not targets in the rodent carcinogenicity studies: acrylonitrile, azidothymidine, cadmium chloride, *p*-chloroaniline, and thioacetamide (Table 2). In addition, we had decided a rule before starting this step of validation study that target organ specificity of carcinogenicity might be considered for the interpretation of negative results about genotoxic carcinogens, but it should be minimized (see the section 6-4.). Therefore, the following discussion for negative chemicals of genotoxic carcinogens is mainly focused on whether or not the negative results are rational when considering known genotoxic and/or carcinogenic MoA.

1) 2-Acetylaminofluorene (CASRN: 53-96-3; 2-AAF)

2-AAF is a well-known genotoxic carcinogen. The positive results are reported in almost all *in vivo* genotoxicity assays, i.e., rat liver UDS test, rat bone marrow micronucleus (MN) test, mouse Comet assay in the colon, liver, kidney and lung, and gene mutation assay in BigBlue mouse liver (note: reference papers on genotoxicity assay results are described in Appendix 3, as well as the following chemicals discussed). It is reported that the oral administration to animals induced tumor for many organs such as the liver, urinary bladder, and kidney in many animal species (8). To induce genotoxicity and carcinogenicity, metabolic activation of 2-AAF is required, and the critical activation pathway is know to convert 2-AAF to the *N*-hydroxy derivative followed by mainly *N*-SO_4 derivative (9).

In this validation study, this chemical was judged to be negative in both liver and stomach. 2-AAF was also examined in the 1st step of 4th phase validation study, and a positive result (i.e., statistically significant increases in both Dunnett's and Trend tests) in the liver was reported in only one of three laboratories testing this chemical, and the overall judgment of 2-AAF would be