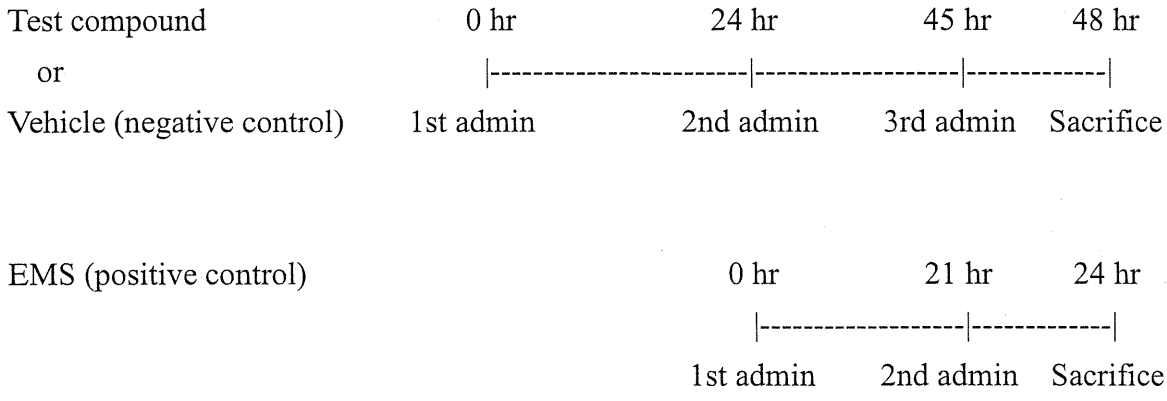
produce mortality, or an unacceptable level of animal distress. Selection of doses will be based on the toxicity of the test substance but will not exceed 2000 mg/kg/day.

1.4.2. Administration to animals

The test substance will be administered three times orally by gavage, 24 and 21 hours apart, i.e. the second administration is 24 hours after the first administration, and the third administration is 21 hours after the second administration (at 3 hours before animal sacrifice). EMS will be administered twice orally by gavage at 24 hours and 3 hours before animal sacrifice. The administration regimes are summarized in a figure below; this protocol enables us to integrate the comet and micronucleated erythrocyte assay into one assay (see note 4). The dosage volume will be 0.1 mL per 10 g body weight in rats on the basis of the animal weight just before administration.



1.4.3. Measurement of body weight and examination of animal conditions

Individual body weights will be measured in accordance with local SOPs and just prior to administration (the weight at this time will be used to determine the volume of each substance administered) and at the time of termination. The clinical signs of the animals will be observed from just after dosing to just before tissue removal with an appropriate interval according to the SOP in each testing facility.

1.4.4. Tissue sampling

Animals will be humanely killed at 3 hours after third administration of a test substance and at 3 hours after second treatment of EMS, consistent with Section C "Animal Welfare and 3Rs". The stomach and the liver will be removed (see note 5). Tissues will be placed into ice-cold mincing buffer, rinsed sufficiently with the cold mincing buffer to remove residual blood (more rinses would likely be needed if exsanguination is not used), and stored on ice until processed. For histopathology, samples will be obtained from the same

liver lobe, and from a minimal possible area of stomach.

1.4.5. Preparation of single cells

Single cell preparation should be done within one hour after animal sacrifice (see note 6). The liver and the stomach will be processed as follows:

Liver: A portion of the left lateral lobe of the liver will be removed and washed in the cold mincing buffer until as much blood as possible has been removed (see note 7). The portion will be minced with a pair of fine scissors to release the cells. The cell suspension will be stored on ice for 15-30 seconds to allow large clumps to settle (or, the cell suspension will be strained through a Cell Strainer to remove lumps and the remaining suspension will be placed on ice), and the supernatant will be used to prepare comet slides.

Stomach: The stomach will be cut open and washed free from food using cold mincing buffer. The forestomach will be removed and discarded. The glandular stomach will be then placed into cold mincing buffer and incubated on ice for from 15 to 30 minutes. After incubation, the surface epithelia will be gently scraped two times using the a scalpel blade or a Teflon scrapper. This layer will be discarded and the gastric mucosa rinsed with the cold mincing buffer. The stomach epithelia will be carefully scraped 4-5 times (or more, if necessary) with a scalpel blade or Teflon scrapper to release the cells. The cell suspension will be stored on ice for 15-30 seconds to allow large clumps to settle (or, the cell suspension will be strained with a Cell Strainer to remove clumps and the remaining suspension will be placed on ice), and samples of the supernatant used to prepare comet slides.

1.4.6. Slide preparation

Slide preparation should be done within one hour after single cell preparation (see note 6). Comet slides will be prepared using laboratory specific procedures. The volume of the cell suspension added to 0.50% low melting agarose to make the slides will not decrease the percentage of low melting agarose by more than 10% (i.e., not below 0.45%) .

1.4.7. Lysis

Once prepared, the slides will be immersed in chilled lysing solution overnight in a refrigerator under a light proof condition (see note 6). After this incubation period, the slides will be rinsed in purified water or neutralization solution to remove residual detergent and salts prior to the alkali unwinding step.

1.4.8. Unwinding and electrophoresis

Slides will be randomly placed onto a platform of submarine-type electrophoresis unit

and the electrophoresis solution added. A balanced design will be used (see note 8). The electrophoresis solution will be poured until the surfaces of the slides are completely covered with the solution. The slides will be left to be unwind for 20 minutes. Next, the slides will be electrophoresed at 0.7 V/cm for at least 20 minutes, with a constant voltage at approximately 300 mA (see note 9). The current at the start and end of the electrophoresis period should be recorded. The temperature of the electrophoresis solution through unwinding and electrophoresis should be maintained at a constant temperature <10°C . The temperature of the electrophoresis solution at the start of unwinding, the start of electrophoresis, and the end of electrophoresis should be recorded.

1.4.9. Neutralization and dehydration of slides

After completion of electrophoresis, the slides will be immersed in the neutralization buffer for at least 5 minutes. All slides will be dehydrated by immersion into absolute ethanol (≥99.6%) for at least 5 minutes if slides will not be scored soon, allowed to air dry, and then stored until scored at room temperature, protected from humidity > 60 %. Once scored, slides should be retained and stored under low humidity conditions (e.g., in a desiccator) for potential rescoring.

1.4.10. DNA staining, comet visualization and analysis

Coded slides will be blind scored according to laboratory specific SOPs. The slides will be stained with SYBR Gold according to manufacturer's specifications. The comets will be measured via a digital (e.g. CCD) camera linked to an image analyzer system using a fluorescence microscope at magnification of 200X. For each sample (animal/tissue), fifty comets per slide will be analyzed, with 2 slides scored per sample (see note 10). Approximately 10 areas/slide should be observed at 5 cells or less/field (see note 11), taking care to avoid any selection bias, overlap counting of cells, and edge areas of slides. Heavily damaged cells exhibiting a microscopic image (commonly referred to as hedgehogs) consisting of small or non-existent head and large, diffuse tails will be excluded from data collection if the image analysis system can not properly score them (see note 12). However, the frequency of such comets should be determined per sample, based on the visual scoring of 100 cells per sample. The comet endpoints collected will be % tail DNA, tail length in microns measured from the estimated edge of the head region closest to the anode (see note 13), and, if possible for a particular image analysis system, Olive tail moment [= a measure of tail length (a distance between a center of head mass and a center of tail mass; microns) X a measure of DNA in tail (% tail DNA/100): Olive et al., 1990]. (see note 14)

1.4.11. Histopathology

When a positive Comet assay response is obtained for a tissue, a sample histopathological assessment will be conducted to evaluate for the presence of apoptotic and/or necrotic cells according to the SOP in each testing facility.

2. STATISTICS

Different approaches for data analysis have been proposed for comet data generated across a range of test substance dose levels (Lovell et al. 1999; Hartmann et al. 2003; Wiklund and Agurell 2003). The primary endpoint of interest for DNA migration is the % tail DNA. In addition, the distribution of migration patterns among cells within an animal will be considered. The percentage of "hedgehogs" will also be evaluated as a function of treatment. The unit of analysis for a specific tissue is the individual animal.

In data analysis process of this validation study, three conceptual key terms, i.e. "Endpoint", "Estimate", and "Effect" are defined and used. 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 a particular "Endpoint" in each animal. "Effect" is defined as difference of an average of "Estimate" between a negative control group and a treatment group (see note 15). Dunnett's test (two-sided, $P<0.05$) and linear Trend test (two-sided, $P<0.05$) will be applied to "Effect" to judge positive or negative as assay results. For the positive control group, Student's t-test (one-sided, $P<0.025$) will be applied to the "Effect".

3. DATA AND REPORTING

3.1.1. Treatment of results

Individual animal data and group summaries will be presented in a fixed tabular form that will be provided from the VMT.

3.1.2. Evaluation and interpretation of results

A positive response is defined as a statistically significant change in the % tail DNA in at least one dose group in comparison with the vehicle control value using Dunnett's test (two-sided, $P<0.05$) as well as a statistically significant linear Trend test (two-sided, $P<0.05$). A negative response is defined as the statistically nonsignificant change in both Dunnett's test and the linear Trend test, and an equivocal response is defined as the statistically significant change in either of Dunnett's test or the linear Trend test. The

positive control should produce a statistically significant increase in Student's t-test (one-sided, $P<0.025$), and if not, the study data will not be acceptable. Where a positive response is obtained in a test substance group, the investigator(s) will assess the possibility that a cytotoxic rather than a genotoxic effect is responsible based on the percentage of "hedgehogs" and histopathology (see note 16). Positive results indicate that the test substance induce DNA damage in the target tissue(s) investigated. Negative results indicate that, under the test conditions used, the test substance does not induce DNA damage *in vivo* in the tissue(s) evaluated.

3.1.3. Study report

The study report from each testing facility will at least include the following information:

3.1.3.1. Test substance and positive/negative controls

Identification; Chemical Abstracts Service Registry number (when available); supplier, lot number and purity (when available); physiochemical properties relevant to the conduct of the study, if known; justification for choice of vehicle; and solubility and stability of the substances in the solvent/vehicle, if known.

3.1.3.2. Test animals

Species/strain used; number, age and sex of animals; source, housing conditions, quarantine and acclimation procedure, and animal identification and group assignment procedure; individual weight of the animals on the day of receipt, at the end of the acclimation period, and before administration (at the time of grouping), including body weight range, mean and standard deviation for each group; and choice of tissue(s) and justification.

3.1.3.3. Reagents to prepare reagent solutions

Identification; supplier; lot number; and time limit for usage if known.

3.1.3.4. Test conditions

Data from range-finding study, if conducted; rationale for dose level selection; details of test substance preparation; details of the administration of the test substance; methods for verifying that the test substance reached the general circulation or target tissue, if applicable; details of food and water quality; detailed description of treatment and sampling schedules; method of measurement of toxicity, including histopathology; detailed methods of single cell preparation; method of slide preparation, including duration between tissue sampling and slide preparation, agarose concentration, lysis conditions (duration for lysis, etc.), alkali conditions and pH, alkali unwinding time and

temperature, electrophoresis conditions (pH, V/cm, mA, and temperature at the start of unwinding and the start and the end of electrophoresis) and staining procedure; criteria for scoring comets and number of comets analyzed per slide, per tissue and per animal; evaluation criteria; criteria for considering studies as positive, negative or equivocal.

3.1.3.5. Results

Signs of toxicity, including histopathology in the appropriate tissue(s) if applicable; individual and mean values for DNA migration (and ranges) and % hedgehogs in individual tissue, animal, and group; concurrent positive and negative control data; and statistical evaluation.

3.1.3.6. Discussion of the results and/or conclusion, as appropriate.

4. ARCHIVES AND REVIEW

The study report and all raw data (including slide samples and image data) from this study will be retained according to the SOP in each testing facility. All raw data will be submitted to the management team for review if required.

5. NOTES

1) We evaluated the data of the 3rd phase validation studies as to whether or not fewer (two, three or four) animals were sufficient in the positive control group to show a statistically significant increase in the Effect (difference) with a one-tailed student's t-test ($P<0.025$). The analysis results were presented and discussed at the Florence meeting held on August 25-26, 2009, and the participants felt that the reduction of animal number would be possible but the slight decrease in the statistic power might require additional experiments and result in the increase in animal usage. Thus the VMT decided to continue using five animals as the positive control in this validation effort. We may need to further investigate the appropriate number of animals/group afterwards based upon power calculation.
2) We will likely need to specify shelf life for some solutions as we reconcile lab-specific protocols.
3) The VMT extensively discussed at the Osaka meeting held on Feb. 4-6, 2009 how a preliminary dose-finding study should be done to choose an appropriate high dose level, because selection of a suitable high dose would be closely related to the sensitivity/specificity of genotoxicity assays in general. The VMT decided to request each facility to submit its own protocol for dose-selection, and the VMT will review them and then direct each facility to use its own protocol as it is or to follow a

dose-finding study protocol recommended by the VMT.

4) When following the regimen for EMS as a positive control, micronucleus (MN) induction will be detected in bone marrow but not in peripheral blood. To also detect MN induction in peripheral blood, it would be necessary to administer EMS as well as the other test chemicals three times. It was also pointed out at the Florence meeting (August, 2009) that four times administration of test chemicals excluding the positive control, EMS, would be needed if we expect to detect micronuclei in the peripheral blood.
5) In this validation study, Comet analysis for the liver and the stomach will be conducted. Comet analysis along with MN for the bone marrow and/or the peripheral blood are optional in this validation study.
6) At the Florence meeting, it was pointed out that the duration of tissue sampling should be kept to a set time (e.g. within 10 min) and the duration for lysis should be controlled, in order to obtain more stable negative control values. The VMT considers that such action would be preferred and recommended but not required of participant laboratories because the feasibility would depend on the performance of each laboratory. To further address this issue, the duration of tissue sampling and the duration for lysis should be recorded in the study report of each facility.
7) The size of the liver portions will be at the discretion of the laboratory, because there is no recommendation for standardizing this step.
8) In each electrophoresis run, there should be the same number of slides from each animal in the study; see Attachment 1, an example of how to keep track of each slide during each electrophoresis run. Each laboratory will need to provide its own electrophoresis box chart, as different boxes can accommodate different numbers of slides.
9) Under those electrophoresis conditions, it is expected that an average DNA migration obtained in the negative control group will be 1-8% tail DNA for the liver, and 1-20% tail DNA for the stomach. These ranges were set based on the analysis with negative control data from the 2nd and 3rd phase validation studies, i.e. the average +/- 3XS.D. values were as follows in the 2nd and 3rd phase validation studies, respectively: 3.8+/-4.8 (n=15 from 5 labs) and 3.1+/-3.9 (n=12 from 4 labs) in the liver, and 12.5+/-6.9 (n=12 from 4 labs) and 8.8+/-9 (n=10 from 4 labs) in the stomach. The reason why the lowest value is set at 1 is to be able to detect a significant decrease in % DNA in the tail. The decrease in DNA migration is expected for cross-linkers, and if such agents are intended to be detected using the Comet assay then a decrease

in migration would be easier to detect when the negative control value is at the higher end of the acceptable range. If the negative control average deviates from the range, the duration of electrophoresis will be adjusted to achieve this range.

10) An investigation was conducted to compare with two slides/animal and three slides/animal about some data of the 3rd phase validation study, and the result was presented and discussed at the Florence meeting. As there was no difference between them as far as the present analysis method was used, the VMT decided to use two slides/animal.

11) In order to obtain suitable areas for observation, dilution of cell suspension may be required during the single cell preparation process.

12) This instruction indicates that if a comet is analyzable by the software program then it should be analyzed. However the following cases will be excluded from the analysis: a) analyzable but the recognition by software is considered incorrect (e.g. the automatic recognition of nucleus center is shifted); and b) the staining of nucleus and/or migration is considered poor. At the Florence meeting, more detailed analysis methods were discussed and agreed to, i.e. cells should be classified into three categories, scorable, non-scorable and hedgehog, and also scorable cells with a 90% or more DNA in the tail should not be adopted as part of the data for analysis. The VMT will prepare a color atlas to instruct how to distinguish comet and hedgehog.

13) 'Tail length' is defined as 'Tail migration' in some image analyzers such as Comet IV.

14) At the Atagawa meeting held on March 13-14, 2008, there was discussion about the need to collect data on tail length and Olive tail moment in this validation study. Again, there was brief discussion about this point at the Osaka meeting. The consensus was that % DNA in tail seems to be a sufficient endpoint for validation and therefore these parameters would no longer be analyzed statistically. However, data on tail length and tail moment will continue to be collected in this validation study in case there is a reason to analyze these data in the future.

15) Effect (difference) seems to be more suitable for revealing variation between labs than Effect (ratio), which was pointed out at Osaka meeting in the discussion of the data of 3rd phase validation study.

16) At the present moment, there is no evident data on the consistency between the percentage of "hedgehogs" and histopathology. In this validation study, histopathology will be used as a primary endpoint to evaluate cytotoxicity, although both of the data will be collected for further analysis on the consistency between the percentage of "hedgehogs" and histopathology.

6. REFERENCES

Burlinson B, et al., Fourth International Workgroup on Genotoxicity Testing: result of the in vivo comet assay workgroup. Mutation Res., 627, 31-35, 2007.

Collins AR, et al., Direct enzymatic detection of endogenous oxidative base damage in human lymphocyte DNA. Carcinogenesis, 14, 1733-1735, 1993.

Hartmann A, et al., Recommendation for conducting the *in vivo* alkaline Comet assay. Mutagenesis, 18(1), 45-51, 2003.

Lovell DP, G Thomas G, R Dubow., Issues related to the experimental design and subsequent statistical analysis of in vivo and in vitro comet studies. Teratog Carcinog Mutagen. 19(2), 109-119, 1999.

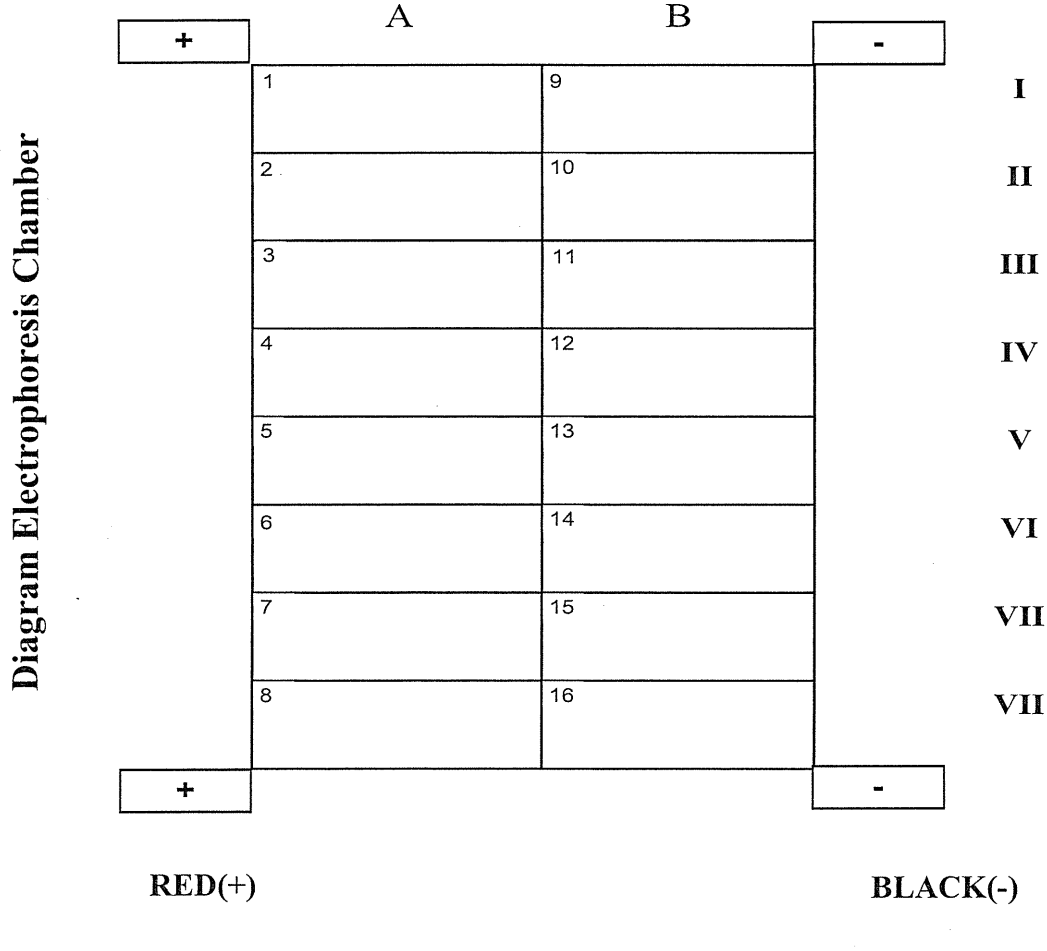
Olive PL, et al., Heterogeneity in radiation-induced DNA damage and repair in tumor and normal cell using the "comet" assay. Radiat. Res., 122, 86-94, 1990.
Tice RR et al., Single cell gel/Comet assay: guidelines for in vitro and in vivo genetic toxicology testing. Environ. Mol. Mutagen., 35, 206-221, 2000.

Wiklund SJ, E Agurell., Aspects of design and statistical analysis in the Comet assay. Mutagenesis 18(2):167-175, 2003.

Attachment 1:

SLIDES UNWINDING & ELECTROPHORESIS RECORDING SHEET

Electrophoresis Run #				**Initials & Date**
Approximate alkaline electrophoresis buffer volume in chamber				
Unwinding				
Time	Total	Start	End	
Buffer Temperature				
Electrophoresis				
Running time	Total	Start	End	
Volts				
Milliamperes				
Buffer Temperature				
Thermometer No.				
Electrophoresis chamber No.				
Power supply No.				



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

146 published literatures, in order to assist the dose selection process in each laboratory.
147 Exceptionally, the dose levels and vehicles of three chemicals, A4204, A4206 and
148 A4219 were directed by VMT because of the retests in another laboratory (Table 1).
149
150 Table 1 Test chemical code, test chemical name, category, laboratory tested, vehicle,
151 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