September 2019

OECD GUIDELINES FOR THE TESTING OF CHEMICALS

Draft TG 488 on Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays

INTRODUCTION

1. The OECD Test Guidelines for the testing of chemicals are periodically reviewed in light of scientific progress, changing regulatory needs and animal welfare considerations. The original Test Guideline 488 was adopted in 2011. In 2013, a revised guideline was adopted that updated: the age range of animals at the start of the treatment; the sections on reproductive tracts to be sampled for sperm collection; and the time for rodent spermatogonial stem cells to become mature sperm and reach the cauda epididymis. This present version of the Test Guideline (TG) focuses on updating the recommended regimens for the analysis of mutations in germ cells.

2. OECD Test Guidelines (TGs) are available for a wide range of *in vitro* mutation assays that are able to detect chromosomal and/or gene mutations. There are Test Guidelines for several *in vivo* genotoxic endpoints (*i.e.* chromosomal aberrations, micronuclei, unscheduled DNA synthesis, and DNA strand breaks using the comet assay); however, these do not measure gene mutations. Transgenic Rodent (TGR) mutation assays fulfil the need for practical and widely available *in vivo* tests for gene mutations.

3. Data from the TGR mutation assays have been reviewed extensively (1) (2). They use transgenic rats and mice that contain multiple copies of chromosomally integrated plasmid or phage shuttle vectors. The transgenes contain reporter genes for the detection of various types of mutations induced *in vivo* by test chemicals.

4. Mutations arising in a rodent are scored by recovering the transgene and analysing the phenotype of the reporter gene in a bacterial host deficient for the reporter gene. TGR gene mutation assays measure mutations induced in genetically neutral genes recovered from virtually any tissue of the rodent. These assays, therefore, circumvent many of the existing limitations associated with the study of *in vivo* gene mutation in endogenous genes (e.g. limited tissues suitable for analysis, negative/positive selection against mutations).

5. The weight of evidence suggests that transgenes respond to mutagens in a similar manner to endogenous genes, especially with regard to the detection of base pair substitutions, frameshift mutations, and small deletions and insertions (1).

6. The International Workshops on Genotoxicity Testing (IWGT) have endorsed the inclusion of TGR gene mutation assays for *in vivo* detection of gene mutations, and have recommended a protocol for their implementation (3) (4). This TG is based on these recommendations. Further analysis supporting the use of this protocol can be found in (5).

7. It is anticipated that in the future it may be possible to combine a TGR gene mutation assay with a repeat dose toxicity study (TG 407). However, data are required to ensure that the sensitivity of TGR gene mutation assays is unaffected by the shorter one day period of time between the end of the administration

period and the sampling time, as used in the repeat dose toxicology study, compared to three days used in TGR gene mutation assays. Data are also required to indicate that the performance of the repeat dose assay is not adversely affected by using a transgenic rodent strain rather than traditional rodent strains. When these data are available, this TG will be updated.

8. Definitions of key terms are set out in the Annex.

INITIAL CONSIDERATIONS

9. TGR gene mutation assays for which sufficient data are available to support their use in this TG are: *lacZ* bacteriophage mouse (MutaMouse); *lacZ* plasmid mouse; *gpt* delta (*gpt* and Spi$^-$*)* mouse and rat; *lacI* mouse and rat (Big Blue®), as performed under standard conditions. In addition, the *cII* positive selection assay can be used for evaluating mutations in the Big Blue® and MutaMouse models. Mutagenesis in the TGR models is normally assessed as mutant frequency; if required, however, molecular analysis of the mutations can provide additional information (see Paragraph 25).

10. These rodent *in vivo* gene mutation tests are especially relevant to assessing mutagenic hazard in that the assays' responses are dependent upon *in vivo* metabolism, pharmacokinetics, DNA repair processes, and translesion DNA synthesis, although these may vary among species, among tissues and among the types of DNA damage. An *in vivo* assay for gene mutations is useful for further investigation of a mutagenic effect detected by an *in vitro* system, and for following up results of tests using other *in vivo* endpoints (1). In addition to being causally associated with the induction of cancer, gene mutation is a relevant endpoint for the prediction of mutation-based non-cancer diseases in somatic tissues (6) (7) as well as diseases transmitted through the germline.

11. If there is evidence that the test chemical, or a relevant metabolite, will not reach any of the tissues of interest, it is not appropriate to perform a TGR gene mutation assay.

PRINCIPLE OF THE TEST METHOD

12. In the assays described in paragraph 9, the target gene is bacterial or bacteriophage in origin, and the means of recovery from the rodent genomic DNA is by incorporation of the transgene into a λ bacteriophage or plasmid shuttle vector. The procedure involves the extraction of genomic DNA from the rodent tissue of interest, *in vitro* processing of the genomic DNA (i.e. packaging of λ vectors, or ligation and electroporation of plasmids to recover the shuttle vector), and subsequent detection of mutations in bacterial hosts under suitable conditions. The assays employ neutral transgenes that are readily recoverable from most tissues.

13. The basic TGR gene mutation experiment involves treatment of the rodent with a chemical over a period of time. Agents may be administered by any appropriate route, including implantation (*e.g.* medical device testing). The total period during which an animal is dosed is referred to as the administration period. Administration is usually followed by a period of time, prior to sacrifice, during which the agent is not administered and during which unrepaired DNA lesions are fixed into stable mutations. In the literature, this period has been variously referred to as the manifestation time, fixation time or expression time; the end of this period is the sampling time (3) (4). After the animal is sacrificed, genomic DNA is isolated from the tissue(s) of interest and purified.

14. Data for a single tissue per animal from multiple packaging/ligations are usually aggregated, and mutant frequency is generally evaluated using a total of between 10^5 and 10^7 plaque-forming or colony-forming units. When using positive selection methods, total plaque-forming units are determined with a

separate set of non-selective plates.

15. Positive selection methods have been developed to facilitate the detection of mutations in both the *gpt* gene [*gpt* delta mouse and rat, *gpt*$^-$ phenotype (8) (9) (10)] and the *lacZ* gene [MutaMouse or *lacZ* plasmid mouse (11) (12) (13) (14)]; whereas, *lacI* gene mutations in Big Blue® animals are detected through a non-selective method that identifies mutants through the generation of coloured (blue) plaques. Positive selection methodology is also in place to detect point mutations arising in the *cII* gene of the λ bacteriophage shuttle vector [Big Blue® mouse or rat, and MutaMouse (15)] and deletion mutations in the λ *red* and *gam* genes [Spi$^-$ selection in *gpt* delta mouse and rat (9) (10) (16)]. Mutant frequency is calculated by dividing the number of plaques/plasmids containing mutations in the transgene by the total number of plaques/plasmids recovered from the same DNA sample. In TGR gene mutation studies, the mutant frequency is the reported parameter. In addition, a mutation frequency can be determined as the fraction of cells carrying independent mutations; this calculation requires correction for clonal expansion by sequencing the recovered mutants (1).

16. The mutations scored in the *lacI, lacZ, cII* and *gpt* point mutation assays consist primarily of base pair substitution mutations, frameshift mutations and small insertions/deletions. The relative proportion of these mutation types among spontaneous mutations is similar to that seen in the endogenous *Hprt* gene. Large deletions are detected only with the Spi$^-$ selection and the *lacZ* plasmid assays (1). Mutations of interest are *in vivo* mutations that arise in the mouse or rat. *In vitro* and *ex vivo* mutations, which may arise during phage/plasmid recovery, replication or repair, are relatively rare, and in some systems can be specifically identified, or excluded by the bacterial host/positive selection system.

DESCRIPTION OF THE METHOD

Preparations

Selection of animal species

17. A variety of transgenic mouse gene mutation detection models are currently available, and these systems have been more widely used than transgenic rat models. If the rat is clearly a more appropriate model than the mouse (e.g. when investigating the mechanism of carcinogenesis for a tumour seen only in rats, to correlate with a rat toxicity study, or if rat metabolism is known to be more representative of human metabolism) the use of transgenic rat models should be considered.

Housing and feeding conditions

18. The temperature in the experimental animal room ideally should be 22ºC (± 3ºC). Although the relative humidity should be at least 30% and preferably not exceed 70% other than during room cleaning, the goal should be to maintain a relative humidity of 50-60%. Lighting should be artificial, with a daily sequence of 12 hours light, followed by 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical when administered by this route. Animals should be housed in small groups (no more than five) of the same sex if no aggressive behaviour is expected. Animals may be housed individually if scientifically justified.

Preparation of the animals

19. Healthy young sexually mature adult animals (8-12 weeks old at start of treatment) are randomly assigned to the control and treatment groups. The animals are identified uniquely. The animals are acclimated to the laboratory conditions for at least five days. Cages should be arranged in such a way that possible effects due to cage placement are minimized. At the commencement of the study, the weight variation of

animals should be minimal and not exceed ± 20% of the mean weight of each sex.

Preparation of doses

20. Solid test chemicals should be dissolved or suspended in appropriate solvents or vehicles or admixed in diet or drinking water prior to dosing of the animals. Liquid test chemicals may be dosed directly or diluted prior to dosing. For inhalation exposures, test materials can be administered as gas, vapour, or a solid/liquid aerosol, depending on their physicochemical properties. Fresh preparations of the test chemical should be employed unless stability data demonstrate the acceptability of storage.

Test Conditions

Solvent/vehicle

21. The solvent/vehicle should not produce toxic effects at the dose volumes used, and should not be suspected of chemical reaction with the test chemical. If other than well-known solvents/vehicles are used, their inclusion should be supported with reference data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle should be considered first.

Positive Controls

22. Concurrent positive control animals should normally be used. However, for laboratories that have demonstrated competency (see Paragraph 24) and routinely use these assays, DNA from previous positive control treated animals may be included with each study to confirm the success of the method. Such DNA from previous experiments should be obtained from the same species and tissues of interest, and properly stored (see Paragraph 46). When concurrent positive controls are used, it is not necessary to administer them by the same route as the test chemical; however, the positive controls should be known to induce mutations in one or more tissues of interest for the test chemical. The doses of the positive control chemicals should be selected so as to produce weak or moderate effects that critically assess the performance and sensitivity of the assay. Examples of positive control substances and some of their target tissues are included in Table 1.

Table 1: Examples of positive control substances and some of their target tissues

Chemical and CAS No.	**Characteristics**	**Mutation Target Tissue**	
		Rat	**Mouse**
N-Ethyl-N-nitrosourea [CAS no. 759-73-9]	Direct acting mutagen	Liver, lung	Bone marrow, colon, colonic epithelium, intestine, liver, lung, spleen, kidney, ovarian granulosa cells, male germ cells
Ethyl carbamate (urethane) [CAS no. 51-79-6]	Mutagen, requires metabolism but produces only weak effects		Bone marrow, forestomach, small intestine, liver, lung, spleen
2,4-Diaminotoluene [CAS no. 95-80-7]	Mutagen, requires metabolism, also positive in the Spi$^-$ assay	Liver	Liver
Benzo[a]pyrene [CAS no. 50-32-8]	Mutagen, requires metabolism	Liver, omenta	Bone marrow, breast, colon, forestomach, glandular stomach, heart, liver, lung, male germ cells

Negative controls

23. Negative controls, treated with solvent or vehicle alone, and otherwise treated in the same way as the treatment groups, should be included for every sampling time. In the absence of historical or published control data showing that no deleterious or mutagenic effects are induced by the chosen solvent/vehicle, untreated controls should also be included for every sampling time in order to establish acceptability of the vehicle control.

Verification of laboratory proficiency

24. Competency in these assays should be established by demonstrating the ability to reproduce expected results from published data (1) for: 1) mutant frequencies with positive control substances (including weak responses) such as those listed in Table 1, non-mutagens, and vehicle controls; and 2) transgene recovery from genomic DNA (*e.g.* packaging efficiency).

Sequencing of mutants

25. For regulatory applications, DNA sequencing of mutants is not required, particularly where a clear positive or negative result is obtained. However, sequencing data may be useful when high inter- individual variation is observed. In these cases, sequencing can be used to rule out the possibility of jackpots or clonal events by identifying the proportion of unique mutants from a particular tissue. Sequencing approximately 10 mutants per tissue per animal should be sufficient for simply determining if clonal mutants contribute to the mutant frequency; sequencing as many as 25 mutants may be necessary to correct mutant frequency mathematically for clonality. Sequencing of mutants also may be considered when small increases in mutant frequency (*i.e.* just exceeding the untreated control values) are found. Differences in the mutant spectrum between the mutant colonies from treated and untreated animals may lend support to a mutagenic effect (4). Also, mutation spectra may be useful for developing mechanistic hypotheses. When sequencing is to be included as part of the study protocol, special care should be taken in the design of such studies, in particular with respect to the number of mutants sequenced per sample, to achieve adequate power according to the statistical model used (see Paragraph 53).

PROCEDURE

Number and Sex of Animals

26. The number of animals per group should be predetermined to be sufficient to provide statistical power necessary to detect at least a doubling in mutant frequency. Group sizes will consist of a minimum of five animals; however, if the statistical power is insufficient, the number of animals should be increased as required. Male animals should normally be used. There may be cases where testing females alone would be justified; for example, when testing human female-specific drugs, or when investigating female-specific metabolism. If there are significant differences between the sexes in terms of toxicity or metabolism, then both males and females will be required.

Administration Period

27. Based on observations that mutations accumulate with each treatment, a repeated-dose regimen is necessary, with daily treatments for a period of 28 days. This is generally considered acceptable both for producing a sufficient accumulation of mutations by weak mutagens, and for providing an exposure time adequate for detecting mutations in slowly proliferating organs. Alternative treatment regimens may be appropriate for some evaluations, and these alternative dosing schedules should be scientifically justified in the protocol. Treatments should not be shorter than the time required for the complete induction of all the

relevant metabolising enzymes, and shorter treatments may necessitate the use of multiple sampling times that are suitable for organs with different proliferation rates. In any case, all available information (*e.g.* on general toxicity or metabolism and pharmacokinetics) should be used when justifying a protocol, especially when deviating from the above standard recommendations. While it may increase sensitivity treatment times longer than 8 weeks should be explained clearly and justified, since long treatment times may produce an apparent increase in mutant frequency through clonal expansion (4).

Dose Levels

28. Dose levels should be based on the results of a dose range-finding study measuring general toxicity that was conducted by the same route of exposure, or on the results of pre-existing sub-acute toxicity studies. Non-transgenic animals of the same rodent strain may be used for determining dose ranges. In the main test, in order to obtain dose response information, a complete study should include a negative control group (see Paragraph 23) and a minimum of three, appropriately-spaced dose levels, except where the limit dose has been used (see Paragraph 29). The top dose should be the Maximum Tolerated Dose (MTD). The MTD is defined as the dose producing signs of toxicity such that higher dose levels, based on the same dosing regimen, would be expected to produce lethality. Chemicals with specific biological activities at low non-toxic doses (such as hormones and mitogens), and chemicals which exhibit saturation of toxicokinetic properties may be exceptions to the dose-setting criteria and should be evaluated on a case-by-case basis. The dose levels used should cover a range from the maximum to little or no toxicity.

Limit Test

29. If dose range-finding experiments, or existing data from related rodent strains, indicate that a treatment regimen of at least the limit dose (see below) produces no observable toxic effects, and if genotoxicity would not be expected based upon data from structurally related substances, then a full study using three dose levels may not be considered necessary. For an administration period of 28 days (*i.e.* 28 daily treatments), the limit dose is 1000 mg/kg body weight/day. For administration periods of 14 days or less, the limit dose is 2000 mg/kg/body weight/day (dosing schedules differing from 28 daily treatments should be scientifically justified in the protocol; see Paragraph 27).

Administration of Doses

30. The test chemical is usually administered by gavage using a stomach tube or a suitable intubation cannula. In general, the anticipated route of human exposure should be considered when designing an assay. Therefore, other routes of exposure (such as, drinking water, subcutaneous, intravenous, topical, inhalation, intratracheal, dietary, or implantation) may be acceptable where they can be justified. Intraperitoneal injection is not recommended since it is not a physiologically relevant route of human exposure. The maximum volume of liquid that can be administered by gavage or injection at one time depends on the size of the test animal. The volume should not exceed 2 mL/100g body weight. The use of volumes greater than this should be justified. Except for irritating or corrosive chemicals, which will normally reveal exacerbated effects at higher concentrations, variability in test volume should be minimized by adjusting the concentration to ensure a constant volume at all dose levels.

Sampling Time

Somatic Cells

31. The sampling time is a critical variable because it is determined by the period needed for mutations to be fixed. This period is tissue-specific and appears to be related to the turnover time of the cell population, with bone marrow and intestine being rapid responders and the liver being much slower. A suitable compromise for the measurement of mutant frequencies in both rapidly and slowly proliferating tissues is 28

consecutive daily treatments (as indicated in Paragraph 27) and sampling three days after the final treatment; although the maximum mutant frequency may not manifest itself in slowly proliferating tissues under these conditions. If slowly proliferating tissues are of particular importance, then a longer sampling time of 28 days following the 28 day administration period may be more appropriate (4) (5). In such cases, the 28-day sampling time would replace the 3-day sampling time, and would require scientific justification.

Germ Cells

32. TGR assays are well-suited for the study of gene mutation induction in male germ cells (17) (18) (19) (20), in which the timing and kinetics of spermatogenesis have been well-defined (21) (22) (23). The low numbers of ova available for analysis, even after super-ovulation, and the fact that there is no DNA synthesis in the oocyte, preclude the determination of mutation in female germ cells using transgenic assays (27). The available germ cell mutagenicity data obtained with TGR assays have been recently reviewed (24) together with modelling of mouse and rat spermatogenesis (25) to inform on the selection of an appropriate experimental design for assessing mutagenicity in germ cells. The modelling considered that the mitotic phase of spermatogenesis (i.e. stem cells, proliferating and differentiating spermatogonia) is the only spermatogenic phase where both DNA replication and cell proliferation, which are necessary to fix mutations into the transgene (26), are occurring.

33. Male germ cells can be collected as either mature sperm from the cauda epididymis or as developing germ cells from the seminiferous tubules. Developing germ cells from the seminiferous tubules can be collected by simply removing the tunica albuginea that encapsulates the testis, or by extruding them from the seminiferous tubules using either enzymatic or physical separation (29). The latter approach is preferred as it enriches the collected population for germ cells because somatic cells (e.g., Leydig and Sertoli cells) present in the testis cannot be easily separated from the tubules.

34. The timing of spermatogenesis in both mouse (23) and rat (21) (22) is well established. The time for the progression of developing germ cells from exposed spermatogonial stem cells to mature sperm reaching the cauda epididymis is ~49 days for the mouse (25) and ~70 days for the rat (21) (22). Therefore, sampling of caudal mouse and rat sperm at 28+3d does not provide meaningful mutagenicity data because these cells represent a population of germ cells that has not undergone DNA replication during the exposure, and should thus not be conducted. For the mouse, there is also experimental data that this 28+3d design does not detect the strong germ cell mutagens N-ethyl-N-nitrosourea (24) and benzo(a)pyrene (28). Sampling of caudal sperm should be conducted only at a minimum of 49 days (mouse) or 70 days (rats) after the end of the 28 day administration period in those cases where it is important to assess mutations in spermatogonial stem cells (24) (25).

35. Germ cells extruded from seminiferous tubules comprise a mixed population of spermatogonia, spermatocytes and spermatids (17) (18) (25). The composition of the germ cell population collected from mouse and rat seminiferous tubules, according to the number of days of treatment received during the proliferative phase of spermatogenesis, has been described in detail for various sampling times taking into account the known kinetics of spermatogenesis (25). While positive results in tubule germ cells after a 28+3d regimen are informative, a negative result after a 28+3d regimen is insufficient to negate the possibility that a chemical is a germ cell mutagen because only a limited fraction of collected germ cells have received continuous treatment for the full 28 day administration period during the proliferative phase of spermatogenesis (24) (25).

36. Based primarily on extensive modelling of spermatogenesis (25) and limited experimental data (24), collection of germ cells from the seminiferous tubules at a sampling time longer than 3 days is better for the assessment of germ cell mutagenicity. According to the modelling, the 28+28d regimen, as currently recommended for slowly proliferating somatic tissues (Paragraph 31), enables the evaluation of mutations

in a population of mouse germ cells that has received 99.6% of the 28 days of treatment during the proliferative phase of spermatogenesis, versus only 42.2% with the 28+3d regimen (25). Thus, both positive and negative results in mouse germ cells obtained with this 28+28d regimen are considered informative. Accordingly, the 28+28d regimen permits the assessment of mutations in somatic tissues and tubule germ cells from the same animals.

37. Because of the longer duration of spermatogenesis in the rat versus the mouse, the 28+28d regimen in the rat does not provide the same degree of exposure of proliferating cell stages as in the mouse using the same regimen (Paragraph 36). The modelling of rat spermatogenesis indicates that the 28+28d regimen enables the evaluation of mutations in a population of cells that has received 80.3% of the 28 day administration period during the proliferative phase of spermatogenesis versus only 21.6% with the 28+3d regimen (25). According to the modelling (25), 28+44d would be needed to reach 100% effective exposure of proliferating cell stages. While positive results in rat tubule germ cells after 28+28d are informative, negative results may be insufficient to negate the possibility that a chemical is a germ cell mutagen, based on the reduced exposure of proliferating cell stages, as indicated above.

38. The spermatogenesis model (25) is based on the assumptions that (i) all cell types are equally recoverable when collecting tubule germ cells; and (ii) the exposure does not produce a significant induction of germ cell apoptosis or delays in the progression of spermatogenesis. It should be noted that the impact of toxicity, or metabolic factors, on cell-cycle delay are not considered when recommending the sampling times for somatic tissues, irrespective of proliferation rate (see paragraph 31). Nevertheless, according to the model, if such effects were to occur in germ cells, longer sampling times, such as provided by the 28+28d regimen, would enable recovery by allowing the testes to be repopulated with surviving stem cells and differentiating spermatogonia that have received the full 28 day administration period during the proliferative phase of spermatogenesis.

39. Other sampling times for germ cells around 28 days may also be acceptable, but would require scientific justification.

40. When a sufficient number of studies become available to ascertain the benefit of any other germ cell regimen, the Test Guideline will be reviewed and, if necessary, revised in light of the experience gained.

Observations

41. General clinical observations should be made at least once a day, preferably at the same time(s) each day and considering the peak period of anticipated effects after dosing. The health condition of the animals should be recorded. At least twice daily, all animals should be observed for morbidity and mortality. All animals should be weighed at least once a week, and at sacrifice. Measurements of food consumption should be made at least weekly. If the test chemical is administered via the drinking water, water consumption should be measured at each change of water and at least weekly. Animals exhibiting non-lethal indicators of excess toxicity should be euthanatized prior to completion of the test period (30).

Tissue Collection

42. The rationale for tissue collection should be defined clearly. Since it is possible to study mutation induction in virtually any tissue, the selection of tissues to be collected should be based upon the reason for conducting the study and any existing mutagenicity, carcinogenicity or toxicity data for the chemical under investigation. Important factors for consideration should include the route of administration (based on likely human exposure route(s)), the predicted tissue distribution, and the possible mechanism of action. In the absence of any background information, several somatic tissues as may be of interest should be collected. These should represent rapidly proliferating, slowly proliferating and site of contact tissues. In addition,

spermatozoa from the cauda epididymis or developing germ cells from the seminiferous tubules (as described in Paragraphs 33 and 36) could be collected and stored in case future analysis of germ cell mutagenicity is required and an appropriate sample time has been used. Organ weights should be obtained, and for larger organs, the same area should be collected from all animals.

Storage of Tissues and DNA

43. Tissues (or tissue homogenates) should be stored at or below −70 ºC and be used for DNA isolation within 5 years. Isolated DNA, stored refrigerated at 4 ºC in appropriate buffer, should be used optimally for mutation analysis within 1 year.

Selection of Tissues for Mutant Analysis

44. The choice of tissues should be based on considerations such as: 1) the route of administration or site of first contact (*e.g.* glandular stomach if administration is oral, lung if administration is through inhalation, or skin if topical application has been used); and 2) pharmacokinetic parameters observed in general toxicity studies, which indicate tissue disposition, retention or accumulation, or target organs for toxicity. If studies are conducted to follow up carcinogenicity studies, target tissues for carcinogenicity should be considered. The choice of tissues for analysis should maximize the detection of chemicals that are direct-acting *in vitro* mutagens, rapidly metabolized, highly reactive or poorly absorbed, or those for which the target tissue is determined by route of administration (31).

45. In the absence of background information and taking into consideration the site of contact due to route of administration, the liver and at least one rapidly dividing tissue (*e.g.* glandular stomach, bone marrow) should be evaluated for mutagenicity. In most cases, the above requirements can be achieved from analyses of two carefully selected tissues, but in some cases, three or more would be needed. If there are reasons to be specifically concerned about germ cell effects, including positive responses in somatic cells, germ cell tissues should be evaluated for mutations.

Methods of Measurement

46. Standard laboratory or published methods for the detection of mutants are available for the recommended transgenic models: *lacZ* lambda bacteriophage and plasmid (14); *lacI* mouse (32) (33); *gpt* delta mouse (9); *gpt* delta rat (10); *cII* (15). Modifications should be justified and properly documented. Data from multiple packagings can be aggregated and used to reach an adequate number of plaques or colonies. However, the need for a large number of packaging reactions to reach the appropriate number of plaques may be an indication of poor DNA quality. In such cases, data should be considered cautiously because they may be unreliable. The optimal total number of plaques or colonies per DNA sample is governed by the statistical probability of detecting sufficient numbers of mutants at a given spontaneous mutant frequency. In general, a minimum of 125,000 to 300,000 plaques is required if the spontaneous mutant frequency is in the order of 3×10^{-5} (3). For the Big Blue® *lacI* assay, it is important to demonstrate that the whole range of mutant colour phenotypes can be detected by inclusion of appropriate colour controls concurrent with each plating. Tissues and the resulting samples (items) should be processed and analysed using a block design, where items from the vehicle/solvent control group, the positive control group (if used) or positive control DNA (where appropriate), and each treatment group are processed together.

DATA AND REPORTING

Treatment of Results

47. Individual animal data should be presented in tabular form. The experimental unit is the animal.

The report should include the total number of plaque-forming units (pfu) or colony-forming units (cfu), the number of mutants, and the mutant frequency for each tissue from each animal. If there are multiple packaging/rescue reactions, the number of reactions per DNA sample should be reported. While data for each individual reaction should be retained, only the total pfu or cfu need be reported. Data on toxicity and clinical signs as per paragraph 44 should be reported. Any sequencing results should be presented for each mutant analysed, and resulting mutation frequency calculations for each animal and tissue should be shown.

Statistical Evaluation and Interpretation of Results

48. There are several criteria for determining a positive result, such as a dose-related increase in the mutant frequency, or a clear increase in the mutant frequency in a single dose group compared to the solvent/vehicle control group. At least three treated dose groups should be analysed in order to provide sufficient data for dose-response analysis. While biological relevance of the results should be the primary consideration, appropriate statistical methods may be used as an aid in evaluating the test results (3) (34) (35) (36) (37). Statistical tests used should consider the animal as the experimental unit.

49. A test chemical for which the results do not meet the above criteria in any tissue is considered non-mutagenic in this assay. For biological relevance of a negative result, tissue exposure should be confirmed.

50. For DNA sequencing analyses, a number of statistical approaches are available to assist in interpreting the results (38) (39) (40) (41).

51. Consideration of whether the observed values are within or outside of the historical control range can provide guidance when evaluating the biological significance of the response (42).

Test report

52. The test report should include the following information:

Test chemical:
- identification data and CAS n°, if known;
- source, lot number if available;
- physical nature and purity;
- physicochemical properties relevant to the conduct of the study;
- stability of the test chemical, if known;

Solvent/vehicle:
- justification for choice of vehicle;
- solubility and stability of the test chemical in the solvent/vehicle, if known;
- preparation of dietary, drinking water or inhalation formulations;
- analytical determinations on formulations (e.g., stability, homogeneity, nominal concentrations);

Test animals:
- species/strain used and justification for the choice;
- number, age and sex of animals;
- source, housing conditions, diet, etc.;
- individual weight of the animals at the start of the test, including body weight range, mean and standard deviation for each group;

Test conditions:
- positive and negative (vehicle/solvent) control data;
- data from the range-finding study;

- rationale for dose level selection;
- details of test chemical preparation;
- details of the administration of the test chemical;
- rationale for route of administration;
- rationale for tissues/cell type analysed
- methods for measurement of animal toxicity, including, where available, histopathological or hematological analyses and the frequency with which animal observations and body weights were taken;
- methods for verifying that the test chemical reached the target tissue, or general circulation, if negative results are obtained;
- actual dose (mg/kg body weight/day) calculated from diet/drinking water test chemical concentration (ppm) and consumption, if applicable;
- details of food and water quality;
- detailed description of treatment and sampling schedules and justifications for the choices;
- method of euthanasia;
- procedures for isolating and preserving tissues;
- methods for isolation of rodent genomic DNA, rescuing the transgene from genomic DNA, and transferring transgenic DNA to a bacterial host;
- source and lot numbers of all cells, kits and reagents (where applicable);
- methods for enumeration of mutants;
- methods for molecular analysis of mutants and use in correcting for clonality and/or calculating mutation frequencies, if applicable;

Results:
- animal condition prior to and throughout the test period, including signs of toxicity;
- body and organ weights at sacrifice;
- for each tissue/animal, the number of mutants, number of plaques or colonies evaluated, mutant frequency;
- for each tissue/animal group, number of packaging reactions per DNA sample, total number of mutants, mean mutant frequency, standard deviation;
- dose-response relationship, where possible;
- for each tissue/animal, the number of independent mutants and mean mutation frequency, where molecular analysis of mutations was performed;
- concurrent and historical negative control data with ranges, means and standard deviations;
- concurrent positive control (or non-concurrent DNA positive control) data;
- analytical determinations, if available (e.g. DNA concentrations used in packaging, DNA sequencing data);
- statistical analyses and methods applied;

Discussion of the results;

Conclusion.

LITERATURE

(1) OECD (2009), *Detailed Review Paper on Transgenic Rodent Mutation Assays*, Series on Testing and Assessment, N° 103, ENV/JM/MONO(2009)7, OECD, Paris.

(2) OECD (2011), *Retrospective Performance Assessment of OECD Test Guideline on Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays*, Series on Testing and Assessment, N° 145, OECD, Paris.

(3) Heddle, J.A., S. Dean, T. Nohmi, M. Boerrigter, D. Casciano, G.R. Douglas, B.W. Glickman, N.J. Gorelick, J.C. Mirsalis, H.-J Martus, T.R. Skopek, V. Thybaud, K.R.Tindall and N. Yajima (2000), "In vivo Transgenic Mutation Assays", *Environ. Mol. Mutagen*., 35: 253-259.

(4) Thybaud, V., S. Dean, T. Nohmi, J. de Boer, G.R. Douglas, B.W. Glickman, N.J. Gorelick, J.A. Heddle, R.H. Heflich, I. Lambert, H.-J. Martus, J.C. Mirsalis, T. Suzuki and N. Yajima (2003), "*In vivo* Transgenic Mutation Assays", *Mutation Res*., 540: 141-151.

(5) Heddle, J.A., H.-J. Martus and G.R. Douglas (2003), "Treatment and Sampling Protocols for Transgenic Mutation Assays", *Environ. Mol. Mutagen.*, 41: 1-6.

(6) Erikson, R.P. (2003), "Somatic Gene Mutation and Human Disease other than Cancer", *Mutation Res.*, 543: 125-136.

(7) Erikson, R.P. (2010), "Somatic Gene Mutation and Human Disease other than Cancer: an Update", *Mutation Res.*, 705: 96-106.

(8) Nohmi, T., M. Katoh, H. Suzuki, M. Matsui, M. Yamada, M. Watanabe, M. Suzuki, N. Horiya, O. Ueda, T. Shibuya, H. Ikeda and T. Sofuni (1996), "A new Transgenic Mouse Mutagenesis Test System using Spi$^-$ and 6-thioguanine Selections", *Environ. Mol. Mutagen*., 28(4): 465–470.

(9) Nohmi, T., T. Suzuki and K.I. Masumura (2000), "Recent Advances in the Protocols of Transgenic Mouse Mutation Assays", *Mutation Res.*, 455(1–2): 191–215.

(10) Toyoda-Hokaiwado, N., T. Inoue, K. Masumura, H. Hayashi, Y. Kawamura, Y. Kurata, M. Takamune, M. Yamada, H. Sanada, T. Umemura, A. Nishikawa and T. Nohmi (2010), "Integration of *in vivo* Genotoxicity and Short-term Carcinogenicity Assays using F344 *gpt* delta Transgenic Rats: *in vivo* Mutagenicity of 2,4-diaminotoluene and 2,6-diaminotoluene Structural Isomers", *Toxicol. Sci.*, 114(1): 71-78.

(11) Boerrigter, M.E., M.E. Dollé, H.-J. Martus, J.A. Gossen and J. Vijg (1995), "Plasmid-based Transgenic Mouse Model for Studying *in vivo* Mutations" *Nature*, 377(6550): 657–659.

(12) Gossen, J.A., W.J. de Leeuw, C.H. Tan, E.C. Zwarthoff, F. Berends, P.H. Lohman, D.L. Knook and J. Vijg (1989), "Efficient Rescue of Integrated Shuttle Vectors from Transgenic Mice: a Model for Studying Mutations *in vivo", Proc. Natl. Acad. Sci. USA,* 86(20): 7971–7975.

(13) Gossen, J.A. and J. Vijg (1993), "A Selective System for lacZ-Phage using a Galactose-sensitive *E. coli* Host", *Biotechniques*, 14(3): 326, 330.

(14) Vijg, J. and G.R. Douglas (1996), "Bacteriophage λ and Plasmid *lacZ* Transgenic Mice for studying Mutations *in vivo*" in: G. Pfeifer (ed.), *Technologies for Detection of DNA Damage and Mutations, Part II,* Plenum Press, New York, NY, USA, pp. 391–410.

(15) Jakubczak, J.L., G. Merlino, J.E. French, W.J. Muller, B. Paul, S. Adhya and S. Garges (1996), "Analysis of Genetic Instability during Mammary Tumor Progression using a novel Selection-based Assay for *in vivo* Mutations in a Bacteriophage λ Transgene Target", *Proc. Natl. Acad. Sci. USA*, 93(17): 9073–9078.

(16) Nohmi, T., M. Suzuki, K. Masumura, M. Yamada, K. Matsui, O. Ueda, H. Suzuki, M. Katoh, H. Ikeda

and T. Sofuni (1999), "Spi⁻ Selection: an Efficient Method to Detect γ-ray-induced Deletions in Transgenic Mice", *Environ. Mol. Mutagen.*, 34(1): 9–15.

(17) Douglas, G.R., J. Jiao, J.D. Gingerich, J.A. Gossen and L.M. Soper (1995), "Temporal and Molecular Characteristics of Mutations Induced by Ethylnitrosourea in Germ Cells Isolated from Seminiferous Tubules and in Spermatozoa of *lacZ* Transgenic Mice", *Proc. Natl. Acad. Sci. USA*, 92: 7485-7489.

(18) Douglas, G.R., J.D. Gingerich, L.M. Soper and J. Jiao (1997), "Toward an Understanding of the Use of Transgenic Mice for the Detection of Gene Mutations in Germ Cells", *Mutation Res.*, 388(2-3): 197-212.

(19) Singer, T.M., I.B. Lambert, A. Williams, G.R. Douglas and C.L. Yauk (2006), "Detection of Induced Male Germline Mutation: Correlations and Comparisons between Traditional Germline Mutation Assays, Transgenic Rodent Assays and Expanded Simple Tandem Repeat Instability Assays", *Mutation. Res.*, 598: 164-193.

(20) Olsen, A.K., A. Andreassen, R. Singh, R. Wiger, N. Duale, P.B. Farmer and G. Brunborg (2010), "Environmental exposure of the mouse germ line: DNA adducts in spermatozoa and formation of de novo mutations during spermatogenesis", PLoS One, 5:e11349.

(21) Clermont, Y. (1972), "Kinetics of spermatogenesis in mammals: seminiferous epithelium cycle and spermatogonial renewal". *Physiol. Rev. 52:* 198-236.

(22) Robaire, B., Hinton, B.T., and Oregbin-Crist, M.-C. (2006), "The Epididymis", in Neil, J.D., Pfaff, D.W., Chalis, J.R.G., de Kretser, D.M., Richards, J.S., and P. M, Wassarman (eds.), *Physiology of Reproduction*, Elsevier, the Netherlands, pp. *1071-1148*.

(23) Oakberg, E.F. (1956), "Duration of spermatogenesis in the mouse and timing of the stages of the cycle of the seminiferous epithelium", *Am. J. Anat.*, 99: 507–516.

(24) Marchetti, F., M. Aardema, C. Beevers, J. van Benthem, R. Godschalk, C.L. Yauk, B. Young, A. Williams and G.R. Douglas (2018), "Identifying germ cell mutagens using OECD test guideline 488 (transgenic rodent somatic and germ cell mutation assay) and integration with somatic cell testing", *Mutat. Res.*, 832-833: 7-18. Corrigendum: *Mutat. Res.*, 2019, 844: 70-71.

(25) Marchetti, F., M. Aardema, C. Beevers, J. van Benthem, G.R. Douglas, R. Godschalk, C.L. Yauk, B. Young and A. Williams (2018), "Simulation of mouse and rat spermatogenesis to inform genotoxicity testing using OECD test guideline 488", *Mutat. Res.*, 832-833: 19-28. Corrigendum: *Mutat. Res.*, 2019, 844: 69.

(26) Bielas, J.H. and J.A. Heddle (2000), "Proliferation is necessary for both repair and mutation in transgenic mouse cells", *Proc. Natl. Acad. Sci. U.S.A.*, 97: 11391-11396.

(27) Yauk, C.L., J.D. Gingerich, L. Soper, A. MacMahon, W.G. Foster and G.R. Douglas (2005), "A *lacZ* Transgenic Mouse Assay for the Detection of Mutations in Follicular Granulosa Cells", *Mutation Res.*, 578(1-2): 117-123.

(28) O'Brien, J.M., M.A. Beal, C.L. Yauk and F. Marchetti (2016), "Benzo(a)pyrene is mutagenic in mouse spermatogonial stem cells and dividing spermatogonia", *Toxicol. Sci.*, 152: 362-371.

(29) O'Brien, J.M., M.A. Beal, J.D. Gingerich, L. Soper, G.R. Douglas, C.L. Yauk and F. Marchetti (2014), "Transgenic rodent assay for quantifying male germ cell mutant frequency", *J. Vis. Exp.*, 90: e51576.

(30) OECD (2000), *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation*, Series on Testing and Assessment, N°19, ENV/JM/MONO(2000)7, OECD, Paris.

(31) Dean, S.W., T.M. Brooks, B. Burlinson, J. Mirsalis, B. Myhr, L. Recio and V. Thybaud (1999), "Transgenic Mouse Mutation Assay Systems can Play an important Role in Regulatory Mutagenicity

Testing *in vivo* for the Detection of Site-of-contact Mutagens", *Mutagenesis*, 14(1): 141–151.

(32) Bielas, J.H. (2002), "A more Efficient Big Blue® Protocol Improves Transgene Rescue and Accuracy in an Adduct and Mutation Measurement", *Mutation Res.*, 518: 107–112.

(33) Kohler, S.W., G.S. Provost, P.L. Kretz, A. Fieck, J.A. Sorge and J.M. Short (1990), "The Use of Transgenic Mice for Short-term, *in vivo* Mutagenicity Testing", *Genet. Anal. Tech. Appl.*, 7(8): 212–218.

(34) Carr, G.J. and N.J. Gorelick (1995), "Statistical Design and Analysis of Mutation Studies in Transgenic Mice*", Environ. Mol. Mutagen*, 25(3): 246–255.

(35) Fung, K.Y., G.R. Douglas and D. Krewski (1998), "Statistical Analysis of *lacZ* Mutant Frequency Data from Muta™Mouse Mutagenicity Assays", *Mutagenesis*, 13(3): 249–255.

(36) Piegorsch, W.W., B.H. Margolin, M.D. Shelby, A. Johnson, J.E. French, R.W. Tennant and K.R. Tindall (1995), "Study Design and Sample Sizes for a *lacI* Transgenic Mouse Mutation Assay", *Environ. Mol. Mutagen.*, 25(3): 231–245.

(37) Piegorsch, W.W., A.C. Lockhart, G.J. Carr, B.H. Margolin, T. Brooks, G.R. Douglas, U.M. Liegibel, T. Suzuki, V. Thybaud, J.H. van Delft and N.J. Gorelick (1997), "Sources of Variability in Data from a Positive Selection *lacZ* Transgenic Mouse Mutation Assay: an Interlaboratory Study", *Mutation. Res.*, 388(2–3): 249–289.

(38) Adams, W.T. and T.R. Skopek (1987), "Statistical Test for the Comparison of Samples from Mutational Spectra", *J. Mol. Biol*., 194: 391-396.

(39) Carr, G.J. and N.J. Gorelick (1996), "Mutational Spectra in Transgenic Animal Research: Data Analysis and Study Design Based upon the Mutant or Mutation Frequency", *Environ. Mol. Mutagen,* 28: 405–413.

(40) Dunson, D.B. and K.R. Tindall (2000), "Bayesian Analysis of Mutational Spectra", *Genetics*, 156: 1411–1418.

(41) Lewis P.D., B. Manshian, M.N. Routledge, G.B. Scott and P.A. Burns (2008), "Comparison of Induced and Cancer-associated Mutational Spectra using Multivariate Data Analysis", *Carcinogenesis,* 29(4): 772-778.

(42) Hayashi, M., K. Dearfield, P. Kasper, D. Lovell, H.-J. Martus, V. Thybaud (2011), "Compilation and Use of Genetic Toxicity Historical Control Data", *Mutation Res.*, doi:10.1016/j.mrgentox.2010.09.007.

ANNEX

DEFINITIONS

Administration period: the total period during which an animal is dosed.

Base pair substitution: a type of mutation that causes the replacement of a single DNA nucleotide base with another DNA nucleotide base.

Capsid: the protein shell that surrounds a virus particle.

Clonal expansion: the production of many cells from a single (mutant) cell.

C*olony-forming unit* (cfu): a measure of viable bacterial numbers.

Concatamer: a long continuous biomolecule composed of multiple identical copies linked in series.

Cos site: a 12-nucleotide segment of single-stranded DNA that exists at both ends of the bacteriophage lambda's double-stranded genome.

Deletion: a mutation in which one or more (sequential) nucleotides is lost by the genome.

Electroporation: the application of electric pulses to increase the permeability of cell membranes.

Endogenous gene: a gene native to the genome.

Extrabinomial variation: greater variability in repeat estimates of a population proportion than would be expected if the population had a binomial distribution.

Frameshift mutation: a genetic mutation caused by insertions or deletions of a number of nucleotides that is not evenly divisible by three within a DNA sequence that codes for a protein/peptide.

Insertion: the addition of one or more nucleotide base pairs into a DNA sequence.

Jackpot: a large number of mutants that arose through clonal expansion from a single mutation.

Large deletions: deletions in DNA of more than several kilobases (which are effectively detected with the Spi^- selection and the lacZ plasmid assays).

Ligation: the covalent linking of two ends of DNA molecules using DNA ligase.

Mitogen: a chemical that stimulates a cell to commence cell division, triggering mitosis (i.e. cell division).

Neutral gene: a gene that is not affected by positive or negative selective pressures.

Packaging: the synthesis of infective phage particles from a preparation of phage *capsid* and tail proteins and a *concatamer* of phage DNA molecules. Commonly used to package DNA cloned onto a lambda vector (separated by *cos sites*) into infectious lambda particles.

Packaging efficiency: the efficiency with which packaged bacteriophages are recovered in host bacteria.

Plaque forming unit (pfu): a measure of viable bacteriophage numbers.

Point mutation: a general term for a mutation affecting only a small sequence of DNA including small insertions, deletions, and base pair substitutions.

Positive selection: a method that permits only mutants to survive.

Reporter gene: a gene whose mutant gene product is easily detected.

Sampling time: the end of the period of time, prior to sacrifice, during which the agent is not administered and during which unprocessed DNA lesions are fixed into stable mutations.

Shuttle vector: a vector constructed so that it can propagate in two different host species; accordingly, DNA inserted into a shuttle vector can be tested or manipulated in two different cell types or two different organisms.

Transgenic: of, relating to, or being an organism whose genome has been altered by the transfer of a gene or genes from another species.