GUIDANCE FOR THE IMPLEMENTATION OF REACH

Guidance on information requirements and chemical safety assessment Chapter R.7a: Endpoint specific guidance

Draft version 3.0
February 2014

NOTE

Please note that the present document is a proposed amendment to specific extracts **only** of the *Guidance on IR&CSA, Chapter R.7a*. This document was prepared by the ECHA Secretariat for the purpose of this consultation and includes only the parts open for the current consultation, i.e. sub-sections R.7.7.1 to R.7.7.7 only.

The full document (version before proposed amendments) is available on the ECHA website at http://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf (version 2.4 published in February 2014).

The numbering and headings of the sub-sections that are displayed in the document for consultation correspond to those used in the currently published guidance document; this will enable the comparison of the draft revised sub-sections with the current text if necessary.

After conclusion of the consultation and before final publication the updated sub-sections will be implemented in the full document.

R.7.7 Mutagenicity and carcinogenicity

R.7.7.1 Mutagenicity

R.7.7.1.1 Definition of mutagenicity

Mutagenicity refers to the induction of permanent transmissible changes in the amount or structure of the genetic material of cells or organisms. These changes may involve a single gene or gene segment, a block of genes or chromosomes. The term clastogenicity is used for agents giving rise to structural chromosome aberrations. A clastogen can cause breaks in chromosomes that result in the loss or rearrangements of chromosome segments. Aneugenicity (aneuploidy induction) refers to the effects of agents that give rise to a change (gain or loss) in chromosome number in cells. An aneugen can cause loss or gain of chromosomes resulting in cells that have not an exact multiple of the haploid number. For example, three number 21 chromosomes or trisomy 21 (characteristic of Down syndrome) is a form of aneuploidy.

Genotoxicity is a broader term and refers to processes which alter the structure, information content or segregation of DNA and are not necessarily associated with mutagenicity. Thus, tests for genotoxicity include tests which provide an indication of induced damage to DNA (but not direct evidence of mutation) via effects such as DNA strandbreaks, unscheduled DNA synthesis (UDS), sister chromatid exchange (SCE), DNA adduct formation or mitotic recombination, as well as tests for mutagenicity.

The chemical and structural complexity of the chromosomal DNA and associated proteins of mammalian cells, and the multiplicity of ways in which changes to the genetic material can be effected make it difficult to give more precise, discrete definitions.

In the risk assessment of substances it is necessary to address the potential effect of *mutagenicity*. It can be expected that some of the available data will have been derived from tests conducted to investigate potentially harmful effects on genetic material (*genotoxicity*). Hence, both the terms *mutagenicity* and *genotoxicity* are used in this document.

R.7.7.1.2 Objective of the guidance on mutagenicity

The aims of testing for genotoxicity are to assess the potential of substances to induce genotoxic effects which may lead to cancer or cause heritable damage in humans. Genotoxicity data are used in risk characterisation and classification of substances. Genotoxicity data are useful for the determination of the general mode of action, *i.e.* thresholded *vs.* non-thresholded effects, and thus for the way the risk assessement can be approached. Expert judgement is necessary at each stage of the testing strategy to decide on the relevance of a result based on the data available for each endpoint.

Alterations to the genetic material of cells may occur spontaneously endogenously or be induced as a result of exposure to ionising or ultraviolet radiation, or genotoxic substances. In principle, human exposure to substances that are mutagens may result in increased frequencies of mutations above background.

Mutations in somatic cells may be lethal or may be transferred to daughter cells with deleterious consequences for the affected organism (*e.g.* cancer may result when they occur in proto-oncogenes, tumour suppressor genes and/or DNA repair genes) ranging from trivial to detrimental or lethal.

Heritable damage to the offspring, and possibly to subsequent generations, of parents exposed to substances that are mutagens may follow if mutations are induced in parental germ cells. To date, all known germ cell mutagens are also mutagenic in somatic cells *in vivo.* Substances

that are mutagenic in somatic cells may produce heritable effects if they, or their active metabolites, have the ability to interact with the genetic material of germ cells. Conversely, substances that do not induce mutations in somatic cells *in vivo* would not be expected to be germ cell mutagens.

There is considerable evidence of a positive correlation between the mutagenicity of substances *in vivo* and their carcinogenicity in long-term studies with animals. Genotoxic carcinogens are substances for which the most plausible mechanism of carcinogenic action involves genotoxicity.

R.7.7.2 Information requirements on mutagenicity

The information requirements on mutagenicity are described by REACH Annexes VI-XI, that specify the information that must be submitted for registration and evaluation purposes. The information is thus required for substances produced or imported in quantities of >1 t/y (tons per annum). When a higher tonnage level is reached, the requirements of the corresponding Annex have to be considered. However, factors including not only production volume but also pre-existing toxicity data, information about the identified use of the substance and exposure of humans to the substance will influence the precise information requirements. The REACH Annexes must thus be considered as a whole, and in conjunction with the overall requirements of registration, evaluation and the duty of care.

Column 1 of REACH Annexes VII-X informs on the standard information requirements for substances produced or imported in quantities of >1 t/y, >10 t/y, >100 t/y, and >1000 t/y, respectively.

Column 2 of REACH Annexes VII-X lists specific rules according to which the required standard information may be omitted, replaced by other information, provided at a different stage or adapted in another way. If the conditions are met under which column 2 of these Annexes allows adaptations, the fact and the reasons for each adaptation should be clearly indicated in the registration dossier.

The standard information requirements for mutagenicity and the specific rules for adaptation of these requirements are presented in Table R.7.7.1.

Table R.7.7-1 REACH information requirements for mutagenicity

COLUMN 1 STANDARD INFORMATION REQUIRED	COLUMN 2 SPECIFIC RULES FOR ADAPTATION FROM COLUMN 1
Annex VII: 1. *In vitro* gene mutation study in bacteria.	Further mutagenicity studies shall be considered in case of a positive result.
Annex VIII: 1. *In vitro* cytogenicity study in mammalian cells or *in vitro* micronucleus study.	1. The study does not usually need to be conducted - if adequate data from an *in vivo* cytogenicity test are available or - the substance is known to be carcinogenic category 1A or 1B or germ cell mutagenic category 1A, 1B or 2.
2. *In vitro* gene mutation study in mammalian cells, if a negative result in Annex VII, 1 and Annex VIII, 1.	2. The study does not usually need to be conducted if adequate data from a reliable *in vivo* mammalian gene mutation test are available. Appropriate *in vivo* mutagenicity studies shall be considered in case of a positive result in any of the genotoxicity studies in Annex VII or VIII.
Annex IX:	If there is a positive result in any of the *in vitro* genotoxicity studies in Annex VII or VIII and there are no results available from an *in vivo* study already, an appropriate *in vivo* somatic cell genotoxicity study shall be proposed by the registrant. If there is a positive result from an *in vivo* somatic cell study available, the potential for germ cell mutagenicity should be considered on the basis of all available data, including toxicokinetic evidence. If no clear conclusions about germ cell mutagenicity can be made, additional investigations shall be considered.
Annex X:	If there is a positive result in any of the *in vitro* genotoxicity studies in Annex VII or VIII, a second *in vivo* somatic cell test may be necessary, depending on the quality and relevance of all the available data. If there is a positive result from an *in vivo* somatic cell study available, the potential for germ cell mutagenicity should be considered on the basis of all available data, including toxicokinetic evidence. If no clear conclusions about germ cell mutagenicity can be made, additional investigations shall be considered.

In addition to these specific rules, the required standard information set may be adapted according to the general rules contained in Annex XI. In this case as well, the fact and the reasons for each adaptation should be clearly indicated in the registration.

In some cases, the rules set out in Annex VII to XI may require certain tests to be undertaken earlier than or in addition to the tonnage-triggered requirements. See Section R.7.7.6 for further guidance on testing requirements.

R.7.7.3 Information and its sources on mutagenicity

To be able to evaluate the mutagenic potential of a substance in a comprehensive way, information is required on its capability to induce gene mutations, structural chromosome aberrations (clastogenicity) and numerical chromosome aberrations (aneugenicity). Many test methods are available by which such information can be obtained. Non-testing methods, such as SAR, QSAR and read-across approaches, may also provide information on the mutagenic potential of a substance.

Typically, *in vitro* tests are performed with cultured bacterial cells, human or other mammalian cells. The sensitivity and specificity of tests will vary with different classes of substances and, if adequate data are available for the class of substance to be tested, these data can guide the selection of the most appropriate test systems to be used. In order to detect mutagenic effects also of substances that need to be metabolically activated to become mutagenic, an exogenous metabolic activation system is usually added in *in vitro* tests. For this purpose the post-mitochondrial 9000 x *g* supernatant (S-9 fraction) of whole liver tissue homogenate containing a high concentration of metabolising enzymes and extracted from animals that have been induced to raise the oxidative P450 levels is most commonly employed. In the case when information is required on the mutagenic potential of a substance *in vivo,* several test methods are available. In *in vivo* tests whole animals are used, in which metabolism and toxicokinetic mechanisms in general exist as natural components of the test animal. It should be noted that species-specific differences in metabolism are known. Therefore, different genotoxic responses may be obtained. Some *in vivo* genotoxicity tests such as the TGR and comet assays employ methods by which any tissue (containing nucleated cells) of an animal can in theory be examined for effects on the genetic material, giving the possibility to examine target tissues (possibly including germ cells) and site-of-contact tissues (*i.e.* skin, epithelium of the respiratory or gastro-intestinal tract).

Some test methods, but not all, have an officially adopted EU and/or OECD test guideline (TG) for the testing procedure. In cases where no adopted EU or OECD TG is available for a test method, rigorous and robust protocols should be followed, such as those defined by internationally recognised groups of experts like the International Workshop on Genotoxicity Testing (IWGT) under the umbrella of the International Association of Environmental Mutagen Societies. Furthermore, modifications to OECD TGs have been developed for some classes of substances and may serve to enhance the accuracy of test results. Use of such modified protocols is a matter of expert judgement and will vary as a function of the chemical and physical properties of the substance to be evaluated. Similarly, use of standard test methods for the testing of tissue(s) not covered by those standard test methods should be scientifically justified and validity of the results will depend on the appropriateness of the acceptability criteria, which should have been specifically developed for this (these) tissue(s) based on sufficient experience and historical data.

R.7.7.3.1 Non-human data on mutagenicity

Non-testing data on mutagenicity

Non-test information about the mutagenicity of a substance can be derived in a variety of ways, ranging from simple inspection of the chemical structure through various read-across techniques, the use of expert systems, metabolic simulators, to *global* or *local* (Q)SARs. The usefulness of such techniques varies with the amount and nature of information available, as well as with the specific regulatory questions under consideration.

Regarding substances for which testing data exist, non-test information can be used in the *Weight of Evidence* approach, to help confirm results obtained in specific tests, or to help develop a better understanding of mutagenicity mechanisms. The information may be useful in deciding if, or what, additional testing is required. At the other extreme, where no testing data are available, similar alternative sources of information may assist in setting test priorities. In cases where no testing is likely to be done (low exposure, <1 t/y) they may be the only options available to establish a hazard profile.

Weight of Evidence approaches that use expert judgement to include test results for close chemical analogues are ways of strengthening regulatory positions on the mutagenicity of a substance. Methods that identify general *structural alerts* for genotoxicity such as the Ashby-Tennant super-mutagen molecule (Ashby and Tennant, 1988) may also be useful.

Prediction models for mutagenicity

There are hundreds of (Q)SAR models available in the literature for predicting test results for genotoxic endpoints for closely related structures (Naven *et al.*, 2012; Bakhtyari *et al.*, 2013). These are known as *local* (Q)SARs. When essential features of the information domain are clearly represented, these models may constitute the best predictive tools for estimating a number of mutagenic/genotoxic endpoints. However, quality of reporting varies from model to model and predictivity must be assessed case-by-case on the basis of clear documentation. Use of harmonised templates, such as the QSAR Model Reporting Format (QMRF) and the QSAR Prediction Reporting Format (QPRF) developed by the Joint Research Centre (JRC) of the European Commission (http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/qsar_tools/QRF), can help ensure consistency in summarising and reporting key information on (Q)SAR models and substance-specific predictions generated by (Q)SAR models. The JRC website also hosts the JRC (Q)SAR Model Inventory, which is an inventory of information on the validity of (Q)SAR models that have been submitted to the JRC (http://ihcp.jrc.ec.europa.eu/our_databases/jrc-qsar-inventory).

Generally, (Q)SAR models that contain putative mechanistic descriptors are preferred; however many models use purely structural descriptors. While such models may be highly predictive, they rely on statistical methods and the toxicological significance of the descriptors may be obscure.

(Q)SAR models for mutagenicity can apply to a limited set of congeneric substances (local models) or to a wide variety of non-congeneric substances (global models). Global (Q)SARs are usually implemented in computer programs and may comprise a set of local models; these global models first categorise the input molecule into the chemical domain it belongs to, and then apply the corresponding local prediction model. These are known as expert systems. Other global models apply the same mathematical algorithm on all input molecules without prior separation. It is generally observed that the concept of applicability domain is a useful one and the endpoints for substances inside the applicability domains of the models are better predicted than for substances falling outside.

Many global models for mutagenicity are commercial and some of the suppliers of these global models consider the data in their modelling sets to be proprietary. Proprietary means that the training set data used to develop the (Q)SAR model is hidden from the user. In other cases it means that it may not be distributed beyond use by regulatory authorities. The models do not always equal the software incorporating them, and the software often has flexible options for expert uses. Thus, the level of information available, from both (Q)SAR models and compiled databases, should be adequate for the intended purpose.

A list of the available (free and commercial) predictive software for ecotoxicological, toxicological and environmental endpoints, including mutagenicity models, has been compiled within the frame of the EU project Antares (http://www.antares-life.eu/).

The most common genotoxicity endpoint for global models has been to predict results of the Ames test. Some models for this endpoint include a metabolic simulator.

There are models for many other mutagenicity endpoints. For example, the Danish EPA and the Danish QSAR group at DTU Food (National Food Institute at the Technical University of Denmark) have developed a (Q)SAR database that contains predictions from a number of mutagenicity models. In addition to assorted Ames models, the database contains predictions of the following *in vitro* endpoints: chromosomal aberrations (CHO and CHL cells), mouse lymphoma/*tk*, CHO/hprt gene-mutation assays and UDS (rat hepatocytes); and the following *in vivo* endpoints: *Drosophila* SLRL, mouse micronucleus, rodent dominant lethal, mouse SCE in bone marrow and mouse comet assay data. The database is freely accessible via http://qsar.food.dtu.dk. The online database contains predictions for over 166,000 substances and includes a flexible system for chemical structure and parameter searching. A user manual with information on the individual models including training set information and validation

results is available at the website. The database is also integrated into the OECD (Q)SAR Toolbox. A major update of the database with consensus predictions by use of different QSAR models for each of the modelled endpoints for more than 600,000 structures, including over 70,000 REACH pre-registered substances, and with an improved user interface is scheduled for the beginning of 2015.

Another example of a database with predictions on mutagenicity is the Enhanced NCI Database Browser (http://cactus.nci.nih.gov) sponsored by the U.S. National Cancer Institute. It contains predictions for over 250,000 substances for mutagenicity as well as other non-mutagenic endpoints, some of which may provide valuable mechanistic information (for example alkylating ability or microtubule formation inhibition). It is also searchable by a wide range of parameters and structure combinations.

Neither of these two examples is perfect, but they illustrate a trend towards predictions of multiple endpoints and may assist those making *Weight of Evidence* decisions regarding the mutagenic potential of untested substances. More detailed information on the strengths and limitations of the different (Q)SAR models can be found elsewhere (Serafimova *et al.*, 2010).

OECD QSAR Toolbox

To increase the regulatory acceptance of (Q)SAR models, the OECD has started the development of a QSAR Toolbox to make (Q)SAR technology readily accessible, transparent and less demanding in terms of infrastructure costs (http://www.qsartoolbox.org/). The OECD QSAR Toolbox facilitates the practical application of grouping and read-across approaches to fill gaps in (eco-)toxicity data, including genotoxicity and genotoxic carcinogenicity, for chemical hazard assessment. In particular, the OECD QSAR Toolbox covers the *in vitro* gene mutation (Ames test), *in vitro* chromosomal aberration, *in vivo* chromosomal aberration (micronucleus test), and genotoxic carcinogenicity endpoints. The predictions are based on the implementation of a range of profilers connected with genotoxicity and carcinogenicity (to quickly evaluate substances for common mechanisms or modes of action), and the incorporation of numerous databases with results from experimental studies (to support read-across and trend analysis) into a logical workflow. The Toolbox and guidance on its use are freely available. A user manual "Strategies for chemicals to fill data gaps to assess genetic toxicity and genotoxic carcinogenicity" and various tutorials for categorisation of substances by use of the Toolbox in relation to protein- and DNA- binding and Ames test mutagenicity are also available on the OECD QSAR Toolbox web site.

The *Guidance on IR&CSA Chapter R.6: QSARs and grouping of chemicals* (available at http://echa.europa.eu/web/guest/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment) explains basic concepts of (Q)SARs and gives generic guidance on validation, adequacy and documentation for regulatory purposes. It also describes a stepwise approach for the use of read-across/grouping and (Q)SARs. Further information on the category formation and read-across approach for the prediction of toxicity can be found in Enoch (2010).

Testing data on mutagenicity

Test methods preferred for use are listed in tables R.7.7-2, R.7.7-3 and R.7.7-4. The introduction to the OECD TGs on genetic toxicity testing as well as some of the related OECD TGs are currently being revised under the OECD Test Guidelines Programme (TGP). In addition, an OECD Guidance Document on the selection and application of the assays for genetic toxicity is being developed. For further information, please see http://www.oecd.org/env/testguidelines.

In vitro data

Table R.7.7-2 *In vitro* test methods

Test method	GENOTOXIC ENDPOINTS measured/ PRINCIPLE OF THE TEST METHOD	EU/OECD guideline[a]
Bacterial reverse mutation test	Gene mutations / The test uses amino-acid requiring strains of bacteria to detect (reverse) gene mutations (point mutations and frameshifts).	EU: B.13/14 OECD: 471
In vitro mammalian cell gene mutation test – *hprt* test	Gene mutations / The test identifies substances that induce gene mutations in the *hprt* gene of established cell lines.	EU: B.17 OECD: 476[b]
In vitro mammalian cell gene mutation test – Mouse lymphoma assay	Gene mutations and structural chromosome aberrations / The test identifies substances that induce gene mutations in the *tk* gene of the L5178Y mouse lymphoma cell line. If colonies in a *tk* mutation test are scored using the criteria of normal growth (large) and slow growth (small) colonies, gross structural chromosome aberrations (*i.e.* clastogenic effect) may be measured, since mutant cells that have suffered damage to both the *tk* gene and growth genes situated close to the *tk* gene have prolonged doubling times and are more likely to form small colonies.	EU: B.17 OECD: 476[b]
In vitro mammalian chromosome aberration test	Structural and numerical chromosome aberrations / The test identifies substances that induce structural chromosome aberrations in cultured mammalian established cell lines, cell strains or primary cell cultures. An increase in polyploidy may indicate that a substance has the potential to induce numerical chromosome aberrations, but this guideline is not designed to measure numerical aberrations and is not routinely used for that purpose.	EU: B.10 OECD: 473[b]
In vitro micronucleus test	Structural and numerical chromosome aberrations / The test identifies substances that induce micronuclei in the cytoplasm of interphase cells. These micronuclei may originate from acentric fragments or whole chromosomes, and the test thus has the potential to detect both clastogenic and aneugenic substances.	EU: B.49 OECD: 487[b]

[a] For EU guidelines, see Regulation (EC) No 440/2008 (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008R0440:en:NOT) / for OECD guidelines see http://www.oecd.org/env/testguidelines

[b] OECD TGs 473, 476 and 487 are currently being revised (see http://www.oecd.org/env/testguidelines)

As noted earlier, accepted modifications to the standard test guidelines/methods have been developed to enhance test sensitivity to specific classes of substances. Expert judgement should be applied to judge whether any of these are appropriate for a given substance being registered. For example, protocol modifications for the Ames test might be appropriate for substances such as gases, volatile liquids, azo-dyes, diazo compounds, glycosides, and petroleum oil derived products, which should be regarded as special cases.

Animal data

Somatic cells

Table R.7.7-3 *In vivo* test methods, somatic cells

Test method	GENOTOXIC ENDPOINTS measured/ PRINCIPLE OF THE TEST METHOD	EU/OECD guideline[a]
In vivo mammalian bone marrow chromosome aberration test	Structural and numerical chromosome aberrations / The test identifies substances that induce chromosome aberrations in the bone-marrow cells of animals, usually rodents. An increase in polyploidy may indicate that a substance has the potential to induce numerical chromosome aberrations, but this guideline is not designed to measure numerical aberrations and is not routinely used for that purpose.	EU: B.11 OECD: 475[b]
In vivo mammalian erythrocyte micronucleus test	Structural and numerical chromosome aberrations / The test identifies substances that cause micronuclei in erythroblasts sampled from bone marrow and/or peripheral blood cells of animals, usually rodents. These micronuclei originate from acentric fragments or whole chromosomes, and the test thus has the potential to detect both clastogenic and aneugenic substances.	EU: B.12 OECD: 474 [b]
Unscheduled DNA synthesis (UDS) test with mammalian liver cells *in vivo*	DNA repair / The test identifies substances that induce DNA damage followed by DNA repair (measured as unscheduled "DNA" synthesis) in liver cells of animals, commonly rats. The test is usually based on the incorporation of tritium labelled thymidine into the DNA by repair synthesis after excision and removal of a stretch of DNA containing a region of damage.	EU: B.39 OECD: 486
Transgenic rodent (TGR) somatic and germ cell gene mutation assays	Gene mutations and chromosomal rearrangements (the latter specifically in the plasmid and Spi- assay models) / Since the transgenes are transmitted by the germ cells, they are present in every cell. Therefore, gene mutations and/or chromosomal rearrangements can be detected in virtually all tissues of an animal, including target tissues and specific site of contact tissues.	EU: none OECD: 488
In vivo alkaline single-cell gel electrophoresis assay for DNA strand breaks (comet assay)	DNA strand breaks / The DNA strand breaks may result from direct interactions with DNA, alkali labile sites or as a consequence of incomplete excision repair. Therefore, the alkaline comet assay recognises primary DNA damage that would lead to gene mutations and/or chromosome aberrations, but will also detect DNA damage that may be effectively repaired or lead to cell death. The comet assay can be applied to almost every tissue of an animal from which single cell or nuclei suspensions can be made, including specific site of contact tissues.	EU: none OECD: in development

[a] For EU guidelines, see Regulation (EC) No 440/2008 (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008R0440:en:NOT) / for OECD guidelines see http://www.oecd.org/env/testguidelines

[b] OECD TGs 474 and 475 are currently being revised (see http://www.oecd.org/env/testguidelines)

A detailed review of transgenic animal model assays, including recommendations on how to perform such assays in somatic cells, has been produced for the OECD (Lambert *et al.,* 2005; OECD, 2009).

Validation studies and recommendations have been published in recent years, identifying experimental factors which are of importance for improved harmonisation of data obtained in the alkaline single-cell gel electrophoresis assay for DNA strand breaks (comet assay) (Ersson *et al*., 2013; Azqueta *et al.*, 2013; Forchhammer *et al*., 2012; Azqueta *et al*., 2011a; Azqueta *et al*., 2011b; Forchhammer *et al*., 2009; Collins *et al*., 2008). Specifically, various

international groups have proposed protocols and recommendations for performing the *in vivo* alkaline comet assay (Tice *et al.,* 2000; Hartmann *et al.,* 2003; McKelvey-Martin *et al.*, 1993; Brendler-Schwaab *et al.*, 2005; Burlinson et al., 2007; Smith *et al.*, 2008; Rothfuss *et al.*, 2010; Burlinson, 2012; Vasquez, 2012; Johansson *et al.*, 2010; Kirkland and Speit, 2008; EFSA, 2012). An international validation study on the in vivo alkaline single-cell gel electrophoresis assay was coordinated by the Japanese Centre for the Validation of Alternative Methods (JaCVAM) from 2006 to 2012. The validation study report was peer reviewed by the OECD and an OECD expert group is currently working on the drafting of the TG, with a target date for adoption by the OECD Working Group of National Coordinators of the Test Guidelines Programme (WNT) in April 2014. While awaiting the adoption of the comet OECD TG, the minimum criteria for acceptance of the comet assay published by EFSA (2012) can be used.

Germ cells

Testing in germ cells has in the past been conducted only on very rare occasions (see Section R.7.7.6).

Table R.7.7-4 *In vivo* test methods, germ cells

Test method	GENOTOXIC ENDPOINTS measured/ PRINCIPLE OF THE TEST METHOD	EU/OECD guideline[a]
Mammalian spermatogonial chromosome aberration test	Structural and numerical chromosome aberrations / The test measures chromosome aberrations in mammalian, usually rodent, spermatogonial cells and is, therefore, expected to be predictive of induction of heritable mutations in germ cells. An increase in polyploidy may indicate that a substance has the potential to induce numerical chromosome aberrations, but this guideline is not designed to measure numerical aberrations and is not routinely used for that purpose..	EU: B.23 OECD: 483[b]
Rodent dominant lethal test	Structural and numerical chromosome aberrations / The test measures dominant lethal effects causing embryonic or foetal death resulting from inherited dominant lethal mutations induced in germ cells of an exposed parent, usually the male. It is generally accepted that dominant lethals are due to structural and numerical chromosome aberrations. Rats or mice are recommended as the test species.	EU: B.22 OECD: 478[b]
Transgenic rodent (TGR) somatic and germ cell gene mutation assays	Gene mutations and chromosomal rearrangements (the latter specifically in the plasmid and Spi- assay models) / Since the transgenes are transmitted by the germ cells, they are present in every cell. Therefore, gene mutations and/or chromosomal rearrangements can be detected in virtually all tissues of an animal including specific site of contact tissues and germ cells. Delayed sampling times may need to be considered in order to detect mutations in different stages of spermatogenesis.	EU: none OECD: 488
In vivo alkaline single-cell gel electrophoresis assay for DNA strand breaks (comet assay)	DNA strand breaks / The test measures DNA strand breaks in spermatocytes of an animal and may, therefore, be used to obtain information about the DNA-damaging activity of a substance in germ cells.	EU: none OECD: in development[c]

[a] For EU guidelines, see Regulation (EC) No 440/2008 (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008R0440:en:NOT) / for OECD guidelines see http://www.oecd.org/env/testguidelines

[b] OECD TGs 478 and 483 are currently being revised (see http://www.oecd.org/env/testguidelines)

[c] Applicability of the comet assay to germ cells is currently being discussed at the OECD (see http://www.oecd.org/env/testguidelines)

A detailed review of transgenic animal model assays, including recommendations on how to perform such assays in germ cells, has been produced for the OECD (Lambert *et al.,* 2005; OECD, 2009). The ability to include sampling of somatic and germ cells in a single study significantly reduces the need to perform additional studies to obtain such information, thereby conforming to the 3Rs principles. As specified in the OECD TG 488, additional sampling times may be needed to cover for the all the stages of spermatogenesis. The test can also be used to investigate transmission of mutations to the offspring since treatment of transgenic male mice can result in offspring carrying mutations (Barnett *et al*., 2002). An example of mutagenicity investigation in epididymal spermatozoa using a transgenic mouse model has been published (Olsen *et al.,* 2010).

Databases with experimental data

There are several open-source databases with experimental information on mutagenicity and carcinogenicity (the two endpoints can often not easily be separated). A review of these databases can be found in Serafimova *et al.* (2010).

R.7.7.3.2 Human data on mutagenicity

Occasionally, studies of genotoxic effects in humans exposed by, for example, accident, occupation or participation in clinical studies (*e.g.* from case reports or epidemiological studies) may be available. Generally, cells circulating in blood are investigated for the occurrence of various types of genetic alterations.

R.7.7.4 Evaluation of available information on mutagenicity

Genotoxicity is a complex endpoint and requires evaluation by expert judgement. For both steps of the effects assessment, *i.e.* hazard identification and dose (concentration)-response (effect) assessment, it is very important to evaluate the data with regard to their adequacy and completeness. The evaluation of adequacy should address the reliability and relevance of the data in a way as outlined in the introductory chapter. The completeness of the data refers to the conclusion on the comparison between the available adequate information and the information that is required under the REACH provisions for the applicable tonnage level of the substance. Such a conclusion relies on *Weight of Evidence* approaches, mentioned in Annex XI Section 1.2 of REACH, which categorise available information based on the methods used: *guideline tests*, *non-guideline tests*, and other types of information which may justify adaptation of the standard testing regime. Such a *Weight of Evidence* approach also includes an evaluation of the available data as a whole, *i.e.* both *over and across* toxicological endpoints.

This approach provides a basis to decide whether further information is needed on endpoints for which specific data appear inadequate or not available, or whether the requirements are fulfilled.

R.7.7.4.1 Non-human data on mutagenicity

Non-testing data on mutagenicity

In a more formal approach, documentation can include reference to a related substance or group of substances that leads to the conclusion of concern or lack of concern. This can either be presented according to scientific logic (read-across) or sometimes as a mathematical relationship of chemical similarity.

If well-documented and applicable (Q)SAR data are available, they should be used to help reach the decision points described in the section below. In many cases the accuracy of such methods will be sufficient to help, or allow either a testing or a specific regulatory decision to be made. In other cases the uncertainty may be unacceptable due to the severe consequences of a possible error. This may be driven by many factors including high exposure potential or toxicological concerns.

Substances for which no test-data exist or for which testing is technically not possible represent a special case in which reliance on non-testing data may be absolute. Many factors will dictate the acceptability of non-testing methods in reaching a conclusion based on no tests at all. It may be discussed whether *Weight of Evidence* decisions based on multiple genotoxicity and carcinogenicity estimates can equal or exceed those obtained by one or two *in vitro* tests, and whether general rules for adaptation of the standard testing regime as described in Annex XI to REACH may be invoked based on such estimates. This must be considered on a case-by-case basis.

Testing data on mutagenicity

Evaluation of genotoxicity test data should be made with care.

Regarding ***positive*** **findings**, particular points should be taken into account:

- are the testing conditions (*e.g.* pH, osmolality, precipitates) in *in vitro* mammalian cell assays relevant to the conditions *in vivo*?
- for studies *in vitro*, factors known to influence the specificity of mammalian cell assays such as the cell line used, the top concentration tested, the toxicity measure used or the metabolic activation system used, should be taken into consideration
- responses generated only at highly toxic/cytotoxic doses or concentrations should be interpreted with caution (*i.e.* taking into account the criteria defined in OECD guidelines)
- the presence or absence of a dose (concentration)-response relationship should be considered

Particular points to take into account when evaluating ***negative*** **test results** include:

- the doses or concentrations of test substance used (were they high enough? For studies *in vivo*, was a sufficienlty high dose level inducing signs of toxicity used? For studies *in vitro*, was a sufficient level of cytotoxicity reached?)
- was the test system used sensitive to the nature of the genotoxic changes that might have been expected? For example, some *in vitro* test systems will be sensitive to point mutations and small deletions but not to mutagenic events that create large deletions
- the volatility of the test substance (were concentrations maintained in tests conducted *in vitro*?)
- for studies *in vitro*, the possibility of metabolism not being appropriate in the test system including studies in extra-hepatic organs
- was the test substance taken up by the test system used for *in vitro* studies?
- were sufficient cells scored/sampled for studies *in vitro*? Has the appropriate number of samples/technical replicates been scored to support statistical significance of the putative negative result?

- for studies *in vivo*, did the substance reach the target organ? Or was the substance only in a position to act at the site of contact due to its high reactivity or insufficient systemic availability (taking also toxicokinetic data into consideration, *e.g.* rate of hydrolysis and electrophilicity may be factors that need to be considered)?

- for studies *in vivo*, was sampling appropriate? (Was a sufficient number of animals used? Were sufficient sampling times used? Was a sufficient number of cells scored/sampled?)

Different results between different test systems should be evaluated with respect to their individual significance. Examples of points to be considered are as follows:

- different results obtained in non-mammalian systems and in mammalian cell tests may be addressed by considering possible differences in substance uptake and metabolism, or in genetic material organisation and ability to repair. Although the results of mammalian tests may be considered of higher significance, additional data may be needed to explain differences

- if the results of indicator tests detecting putative DNA lesions (*e.g.* DNA binding, DNA damage, DNA repair; SCE) are not in agreement with results obtained in tests for mutagenicity, the results of mutagenicity tests are generally of higher significance provided that appropriate mutagenicity tests have been conducted. This is subject to expert judgement.

- if different findings are obtained *in vitro* and *in vivo*, in general, the results of *in vivo* tests indicate a higher degree of reliability. However, for evaluation of *negative* results *in vivo*, it should be considered whether the most appropriate tissues were sampled and whether there is adequate evidence of target tissue exposure

- the sensitivity and specificity of different test systems vary for different classes of substances. If available testing data for other related substances permit assessment of the performance of different assays for the class of substance under evaluation, the result from the test system known to produce more accurate responses would be given higher priority

Different results may also be available from the same test, performed by different laboratories or on different occasions. In this case, expert judgement should be used to evaluate the data and reach an overall conclusion. In particular, the quality of each of the studies and of the data provided should be evaluated, with special consideration of the study design, reproducibility of data, dose (concentration)-effect relationships, and biological relevance of the findings. The identity and purity of the test substance may also be a factor to take into account. In the case where an EU/OECD guideline is available for a test method, the quality of a study using the method is regarded as being higher if it was conducted in compliance with the requirements stated in the guideline, unless convincing scientific evidence can be provided to justify certain deviations from the standard test guideline for the specific substance evaluated. Furthermore, for the same assay, studies compliant with GLP may be regarded as being of a higher quality than non GLP studies since GLP studies generally provide more documentation and details of the study, which are important factors to consider when assessing study reliability/quality.

When making an assessment of the potential mutagenicity of a substance, or considering the need for further testing, data from various tests and genotoxic endpoints may be found. Both the strength and the weight of the evidence should be taken into account. The strongest evidence will be provided by modern, well-conducted studies with internationally established test guidelines/methods. For each test type and each genotoxic endpoint, there should be a separate *Weight of Evidence* analysis. It is not unusual for positive evidence of mutagenicity to be found in just one test type or for only one endpoint. In such cases the positive and negative results for different endpoints are not conflicting, but illustrate the advantage of using test

methods for a variety of genetic alterations to increase the probability of identifying substances with mutagenic potential. Hence, results from methods testing different genotoxic endpoints should not be combined in an overall *Weight of Evidence* analysis, but should be subjected to such analysis separately for each endpoint. Based on the whole data set one has to consider whether there are data gaps: if there are data gaps further testing should be considered, otherwise an appropriate conclusion/assessment can be made.

R.7.7.4.2 Human data on mutagenicity

Human data have to be assessed carefully on a case-by-case basis. The interpretation of such data requires considerable expertise. Attention should be paid especially to the adequacy of the exposure information, confounding factors, co-exposures and to sources of bias in the study design or incident. The statistical power of the test may also be considered. It may be mentioned that, to date, no germ cell mutagen has been identified based on human data.

R.7.7.4.3 Remaining uncertainty on mutagenicity

Reliable data can be generated from well-designed and conducted studies *in vitro* and *in vivo*. However, in addition to the uncertainty inherent in testing and due to the lack of human data available, a certain level of uncertainty remains when extrapolating these testing data to the effect in humans.

R.7.7.5 Conclusions on mutagenicity

R.7.7.5.1 Concluding on Classification and Labelling

In order to conclude on an appropriate classification and labelling position with regard to mutagenicity, the available data should be considered using the criteria according to Annex I to the CLP Regulation (EC) No 1272/2008 (See also Section 3.5 of the *Guidance on the application of the CLP criteria*, available at http://echa.europa.eu/web/guest/guidance-documents/guidance-on-clp).

R.7.7.5.2 Concluding on suitability for Chemical Safety Assessment

Considerations on dose (concentration)-response shapes and mode of action of mutagenic substances in test systems

Considerations on the dose (concentration)-response relationship and on possible mechanisms of action are important components of a risk assessment. The default assumption for genotoxic substances has for long been that they have a linear dose (concentration)-response relationship. However, this assumption has recently been challenged by experimental evidence showing that both direct and indirect acting genotoxins can possess non-linear or thresholded dose (concentration)-response curves.

Examples of non-DNA reactive mechanisms that may be demonstrated to lead to genotoxicity via non-linear or thresholded dose (concentration)-response relationships include inhibition of DNA synthesis, alterations in DNA repair, overloading of defence mechanisms (anti-oxidants or metal homeostatic controls), interaction with microtubule assembly leading to aneuploidy, topoisomerase inhibition, high cytotoxicity, metabolic overload and physiological perturbations (*e.g.* induction of erythropoeisis). The mechanisms underlying non-linear or thresholded dose (concentration)-response relationships for some DNA reactive genotoxic substances like alkylating agents seem linked to DNA repair capacity.

Assessment of the significance to be assigned to genotoxic responses mediated by such mechanisms would include an assessment of whether the underlying mechanism can be

induced at substance concentrations that can be expected to occur under relevant *in vivo* conditions.

In general, several concentrations/doses are tested in genotoxicity assays. At least three experimental concentrations/doses have to be tested as recommended in the OECD test guidelines for genotoxicity. Determination of experimental dose (concentration)-effect relationships is essential and may be used to assess the genotoxic potential of a substance, as indicated below. It should be recognised that not all of these considerations may be applicable to *in vivo* data.

- the OECD introduction to the genotoxicity test guidelines lists the relevant criteria for identification of clear positive findings: (i) the increase in genotoxic response is concentration- or dose-related, (ii) at least one of the data points exhibits a statistically significant increase compared to the concurrent negative control, and (iii) the statistically significant result is outside the distribution of the historical negative control data (*e.g.* 95% confidence interval). In practice, the criterion for dose (concentration)-related increase in genotoxicity will be most helpful for *in vitro* tests, but care is needed to check for cytotoxicity or cell cycle delay which may cause deviations from a dose (concentration)-response related effect in some experimental systems

- genotoxicity tests are not designed in order to derive no effect levels. However, the magnitude of the lowest dose with an observed effect (*i.e.* the Lowest Observed Effect Dose or LOED) may, on certain occasions, be a helpful tool in risk assessment. This is true specifically for genotoxic effects caused by thresholded mechanisms, like, *e.g.* aneugenicity. Further, it can give an indication of the mutagenic potency of the substance in the test at issue. Modified studies, with additional dose or concentration points and improved statistical power may be useful in this regard. The Benchmark dose (BMD) approach presents several advantages over the NOED/LOED approach and can be used as an alternative strategy for dose (concentration)-response assessment (see *Guidance on IR&CSA, Chapter R.8*)

- unusual shapes of dose (concentration)-response curves may contribute to the identification of specific mechanisms of genotoxicity. For example, extremely steep increases suggest an indirect mode of action or metabolic switching which could be confirmed by further investigation.

Considerations on genetic risks associated with human exposure to mutagenic substances

There are no officially adopted methods for estimating health risks associated with (low) exposures of humans to mutagens. In fact, most – if not all tests used today – are developed and applied to identify mutagenic properties of the substance, *i.e.* identification of the mutagenic hazard *per se*. In today's regulatory practice, the assessment of human health risks from exposure to mutagenic substances is considered to be covered by assessing and regulating the carcinogenic risks of these agents. The reason for this is that mutagenic events underlie these carcinogenic effects. Therefore, mutagenicity data is not used for deriving dose descriptors for risk assessment purposes and the reader is referred to this aspect in Section R.7.7.8 (Carcinogenicity) for guidance on how to assess the chemical safety for mutagenic substances.

R.7.7.5.3 Information not adequate

A *Weight of Evidence* approach, comparing available adequate information with the tonnage-triggered information requirements by REACH, may result in the conclusion that the requirements are not fulfilled. In order to proceed in gathering further information, the following testing strategy can be adopted:

R.7.7.6 Integrated Testing Strategy (ITS) for mutagenicity

R.7.7.6.1 Objective / General principles

This testing strategy describes a flexible, stepwise approach for hazard identification with regard to the mutagenic potential of substances, so that sufficient data may be obtained for adequate risk characterisation including classification and labelling. It serves to help minimise the use of animals and costs as far as it is consistent with scientific rigour. A flow chart of the testing strategy is presented in Figure R.7.7-1 and recommendations on follow up procedures based on different testing data sets are given in Table R.7.7-5. As noted later in this section, deviations from this strategy may be considered if existing data for related substances indicate that alternate testing strategies yield results with greater sensitivity and specificity for mutagenicity *in vivo*.

The strategy defines a level of information that is considered sufficient to provide adequate reassurance about the potential mutagenicity of most substances. As described below, this level of information will be required for most substances at the Annex VIII tonnage level specified in REACH, although circumstances are described when the data may be required for substances at Annex VII.

For some substances, relevant data from other sources/tests may also be available (*e.g.* physico-chemical, toxicokinetic, and toxicodynamic parameters and other toxicity data; data on well-investigated, structurally similar, substances). These should be reviewed because, sometimes, they may indicate that either more or less genotoxicity studies are needed on the substance than defined by standard information requirements; *i.e.* they may allow tailored testing/selection of test systems. For example, bacterial mutagenesis assays of inorganic metal compounds are frequently negative due to limited capacity for uptake of metal ions and/or the induction of large DNA deletions by metals in bacteria potentially leading to an increased death rate in mutants. The high prevalence of false negatives for metal compounds might suggest that mutagenesis assays with mammalian cells, as opposed to bacterial cells, would be the preferred starting point for testing for this class of Annex VII substances.

In summary, a key concept of the strategy is that initial genotoxicity tests and testing guidelines/methods should be selected with due consideration to existing data that has established the most accurate testing strategy for the class of compound under evaluation. Even then, initial testing may not always give adequate information and further testing may sometimes be considered necessary in the light of all available relevant information on the substance, including its use pattern. Further testing will normally be required for substances which give rise to positive results in any of the *in vitro* tests.

If negative results are available from an adequate evaluation of genotoxicity from existing data in appropriate test systems, there may be no requirement to conduct additional genotoxicity tests.

Substances for which there is a harmonised classification in category 1A, 1B or 2 for germ cell mutagenicity and/or category 1A or 1B for carcinogenicity according to Annex VI to the CLP Regulation (EC) No 1272/2008 will usually not require additional testing in order to meet the requirements of Annex VIII for the *in vitro* cytogenicity study in mammalian cells. Provided that appropriate risk management measures are implemented, the carcinogenicity study to meet the requirements of Annex X (see Section R.7.7.2 of this Guidance) and the reproductive toxicity studies to meet the requirements of Annexes VIII to X (see Section R.7.7.6 of this Guidance) may also be omitted for substances classified in category 1A or 1B for germ cell mutagenicity. In cases where a registrant is unsure of the formal position on the classification of a substance, or wishes to make a classification proposal himself, advice should be sought from an appropriate regulatory body before proceeding with any further testing.

In case additional testing is needed to meet the requirements of Annexes IX or X, the registrant must first submit a testing proposal to the European Chemicals Agency (ECHA) and obtain prior authorisation before any testing can be initiated.

R.7.7.6.2 Preliminary considerations

For a comprehensive coverage of the potential mutagenicity of a substance, information on gene mutations (base substitutions and deletions/additions), structural chromosome aberrations (breaks and rearrangements) and numerical chromosome aberrations (loss or gain of chromosomes, defined as aneuploidy) is required. This may be obtained from available data or tests on the substance itself or, sometimes, by prediction using appropriate *in silico* techniques (*e.g.* chemical grouping, read-across or (Q)SAR approaches).

It is important that whatever is known of the physico-chemical properties of the test substance is taken into account before devising an appropriate testing strategy. Such information may impact upon both the selection of test systems to be employed and/or modifications to the test protocols used. The chemical structure of a substance can provide information for an initial assessment of mutagenic potential. The need for special testing in relation to photomutagenicity may be indicated in some specific cases by the structure of a molecule, its light absorbing potential or its potential to be photoactivated. By using expert judgement, it may be possible to identify whether a substance, or a potential metabolite of a substance, shares or does not share structural characteristics with known mutagens. This can be used to justify a higher or lower level of priority for the characterisation of the mutagenic potential of a substance. Where the level of evidence for mutagenicity is particularly strong, it may be possible to make a conclusive hazard assessment in accordance with Annex I to REACH without additional testing on the basis of structure-activity relationships alone: in this case, the registrant still has to provide sufficient information to meet the requirements of Annexes VII to X but he may, if scientifically justified and duly documented in the registration dossier, invoke the general rules of Annex XI for adaptation of the standard testing regime by demonstrating, *inter alia*, that the results he wishes to use instead of testing in that context are adequate for the purpose of classification and labelling and/or risk assessment.

In vitro tests are particularly useful for gaining an understanding of the potential mutagenicity of a substance and they have a critical role in this testing strategy. They are not, however, without their limitations. Animal tests will, in general, be needed for the clarification of the relevance of positive findings and in case of specific metabolic pathways that cannot be simulated adequately *in vitro*.

The toxicokinetic and toxicodynamic properties of the test substance should be considered before undertaking, or appraising, animal tests. Understanding these properties will enable appropriate protocols for the standard tests to be developed, especially with respect to tissue(s) to be investigated, the route of substance administration and the highest dose tested. If little is understood about the systemic availability of a test substance at this stage, toxicokinetic investigations or modelling may be necessary.

Certain substances in addition to those already noted may need special consideration, such as highly electrophilic substances that give positive results *in vitro*, particularly in the absence of metabolic activation. Although these substances may react with proteins and water *in vivo* and thus be rendered inactive towards many tissues, they may be able to express their mutagenic potential at the initial site of contact with the body. Consequently, the use of test methods such as the comet assay or the gene mutation assays using transgenic animals that can be applied to the respiratory tract, upper gastrointestinal tract and skin may be appropriate. It is possible that specialised test methods will need to be applied in these circumstances, and that these may not have recognised, internationally valid, test guidelines. The validity and utility of such tests and the selection of protocols should be assessed by appropriate experts or authorities on a case-by-case basis.

Criteria for the evaluation and interpretation of results (e.g. how to define clear positive and clear negative results) are normally defined in the testing guidelines/methods. There is no requirement for verification of a clear positive or clear negative result. In cases where the response is neither clearly negative nor clearly positive and in order to assist in establishing the biological relevance of a result (e.g. a weak or borderline increase), the data should be evaluated by expert judgement and/or further investigations. A substance giving an equivocal test result, *i.e.* not all the requirements for a clear positive or clear negative result have been met, should be reinvestigated immediately, normally using the same test method, but varying the conditions to obtain conclusive results. Wherever possible, clear results should be obtained for one step in the strategic procedure before going on to the next. In cases where this does not prove to be possible and the study is inconclusive as a consequence of *e.g.* some limitation of the test or procedure, a further test should be conducted in accord with the strategy.

Tests need not be performed if it is not technically possible to do so, or if they are not considered necessary in the light of current scientific knowledge. Scientific justifications for not performing tests required by the strategy should always be documented. It is preferred that tests as described in OECD Guidelines or Regulation (EC) No 440/2008 are used where possible. Alternatively, for other tests, up-to-date protocols defined by internationally recognised groups of experts, *e.g.* International Workshop on Genotoxicity Testing (IWGT, under the umbrella of the International Association of Environmental Mutagen Societies), may be used provided that the tests are scientifically justified. It is essential that all tests be conducted according to rigorous protocols in order to maximise the potential for detecting a mutagenic response, to ensure that negative results can be accepted with confidence and that results are comparable when tests are conducted in different laboratories. At the time of writing this guidance, a standard test guideline/method is still to be established for the *in vivo* comet assay described below, so if this test is to be conducted consultation on the protocol with an appropriate expert or authority is advisable.

If a registrant wishes to undertake any tests for substances at the Annex IX or X tonnage levels that require the use of vertebrate animals, then there is a need to make a testing proposal to ECHA first. Testing may only be undertaken after ECHA has accepted the testing proposal in a formal decision.

R.7.7.6.3 Testing strategy for mutagenicity

Standard information requirement at Annex VII

A preliminary assessment of mutagenicity is required for substances at the REACH Annex VII tonnage level. All available information should be included but, as a minimum, there should normally be data from a gene mutation test in bacteria unless existing data for analogous substances indicates this would be inappropriate. For substances with significant toxicity to bacteria, not taken up by bacteria, or for which the gene mutation test in bacteria cannot be performed adequately, an *in vitro* mammalian cell gene mutation test may be used as an alternative test.

When the result of the bacterial test is positive, it is important to consider the possibility of the substance being genotoxic in mammalian cells. The need for further test data to clarify this possibility at the Annex VII tonnage level will depend on an evaluation of all the available information relating to the genotoxicity of the substance.

Standard information requirement at Annex VIII

For a comprehensive coverage of the potential mutagenicity of a substance, information on gene mutations, and structural and numerical chromosome aberrations is required for substances at the Annex VIII tonnage level of REACH.

In order to ensure the necessary minimum level of information is provided, at least one further test is required in addition to the gene mutation test in bacteria. This should be an *in vitro*

mammalian cell test capable of detecting both structural and numerical chromosome aberrations.

There are essentially two different methods that can be viewed as alternative options according to REACH for this first mammalian cell test:

- An *in vitro* chromosome aberration test (OECD TG 473), *i.e.* a cytogenetic assay for structural chromosome aberrations using metaphase analysis. An increase in polyploidy may indicate that a substance has the potential to induce numerical chromosome aberrations, but this guideline is not designed to measure numerical aberrations and is not routinely used for that purpose.

- An *in vitro* micronucleus test (OECD TG 487). This is a cytogenetic assay that has the advantage of detecting not only structural chromosomal aberrations but also aneuploidy. Use of a cytokinesis block, fluorescence *in situ* hybridisation with probes for centromeric DNA, or immunochemical labelling of kinetochore proteins can provide information on the mechanisms of chromosome damage and micronucleus formation. The labelling and hybridisation procedures can enable aneugens to be distinguished from clastogens. This may sometimes be useful for risk characterisation. If a substance is demonstrated to be an aneugen, it is assumed that its genotoxicity is thresholded, in contrast to non-thresholded genotoxicity. Both types of genotoxicity mechanisms trigger different ways to perform risk assessment.

Other *in vitro* tests may be acceptable as the first mammalian cell test, but care should be taken to evaluate their suitability for the substance being registered and their reliability as a screen for substances that cause structural and/or numerical chromosome aberrations. A supporting rationale should be presented for a registration with any of these other tests.

It is possible to present existing data from an *in vivo* cytogenetic test (*i.e.* a study or studies conducted previously) as an alternative to the first *in vitro* mammalian cell test. For instance, if an adequately performed *in vivo* micronucleus test is available already it may be presented as an alternative. There may however be specific cases where the *in vitro* mammalian cell test can still be justified even though *in vivo* cytogeneticity data exist, *e.g.* to understand whether the substance is clastogenic (or aneugenic) *in vitro*, and whether it requires a specific metabolism to be genotoxic. This should be considered on a case-by-case basis.

An *in vitro* gene mutation study in mammalian cells (OECD TG 476) is the second part of the standard information set required for registration at the Annex VIII tonnage level. For substances that have been tested already, this information should always be presented as part of the overall *Weight of Evidence* for mutagenicity. For other substances, this second *in vitro* mammalian cell test will normally only be required when the results of the bacterial gene mutation test and the first study in mammalian cells (*i.e.* an *in vitro* chromosome aberration test or an *in vitro* micronucleus test) are negative. This is to detect *in vitro* mutagens that give negative results in the other two tests.

Under specific circumstances it may be possible to omit the second *in vitro* study in mammalian cells, *i.e.* if it can be demonstrated that this mammalian cell test will not provide any further useful information about the potential *in vivo* mutagenicity of a substance, then it does not need to be conducted. This should be evaluated on a case-by-case basis as there may be classes of compound for which conclusive data can be provided to show that the sensitivity of the first two *in vitro* tests cannot be improved by the conduct of the third test.

The *in vitro* mammalian cell gene mutation test will not usually be required if adequate information is available from a reliable *in vivo* study capable of detecting gene mutations. Such information may come from a TGR gene mutation assay. A comet assay or a liver UDS test may also be adequate. However, these two tests being indicator assays detecting putative DNA lesions, their use should be justified on a case-by-case basis, *e.g.* the UDS should be restricted to the detection of primary DNA repair in liver cells.

Provided the *in vitro* tests have given negative results, normally, no *in vivo* tests will be required to fulfil the standard information requirements at Annex VIII. However, there may be rare occasions when it is appropriate to conduct testing *in vivo*, for example when it is not possible technically to perform satisfactory tests *in vitro*. Substances which, by virtue of, for example, their physico-chemical characteristics, chemical reactivity or toxicity cannot be tested in one or more of the *in vitro* tests should be considered on a case-by-case basis. In the same way, it may not always be possible with the S9 fraction used *in vitro* to mimic the *in vivo* metabolism of some substances, and the relevance of the *in vitro* negative results for those substances should be evaluated case by case. In addition, equivocal *in vitro* results or different results from different *in vitro* studies may require the consideration of further testing to reach a clear conclusion on mutagenicity. For those types of cases, expert judgement would be needed to determine whether *in vivo* testing is appropriate.

Requirement for testing beyond the standard levels specified for Annexes VII and VIII

Introductory comments

Concerns raised by positive results from *in vitro* tests usually require the consideration of further testing. The chemistry of the substance, data on analogous substances, toxicokinetic and toxicodynamic data, and other toxicity data will also influence the timing and pattern of further testing.

Unless there are appropriate results from an *in vivo* study already, testing beyond the standard set of *in vitro* tests is normally first directed towards investigating the potential for mutagenicity in somatic cells *in vivo*. Positive results in somatic cells *in vivo* constitute the trigger for consideration of investigation of potential expression of genotoxicity in germ cells. However, to avoid unnecessary testing of vertebrate animals and for cost reasons, as the TGR assays give the possibility to include sampling of somatic and male germ cells in a single study providing adapted sampling times (see OECD TG 488 for details), it is recommended to include such samples in the testing proposal for the TGR assays and to appropriately store the germ cell samples for later analysis in case there is a positive result in any of the somatic tissues tested.

Substances that are negative in the standard set of *in vitro* tests

In general, substances that are negative in the full set of *in vitro* tests specified in REACH Annexes VII and VIII are considered to be non-genotoxic. There are only a very limited number of substances that have been found to be genotoxic *in vivo*, but not in the standard *in vitro* tests. Most of these are pharmaceuticals designed to affect pathways of cellular regulation, including cell cycle regulation, and this evidence is judged insufficient to justify routine *in vivo* testing of industrial chemicals. However, occasionally, knowledge about the metabolic profile of a substance may indicate that the standard *in vitro* tests are not sufficiently reassuring and a further *in vitro* test, or an *in vivo* test, may be needed in order to ensure mutagenicity potential is adequately explored (*e.g.* use of an alternative to rat liver S9 mix, a reducing system, a metabolically active cell line, or genetically engineered cell lines might be judged appropriate).

Substances for which an *in vitro* test is positive

REACH Annex VII substances for which only a bacterial gene mutation test has been conducted and for which the result is positive should be studied further, according to the requirements of Annex VIII.

Regarding Annex VIII, when both the mammalian cell tests are negative but there was a positive result in the bacterial test, it will be necessary to decide whether any further testing is

needed on a case-by-case basis. For example, suspicion that a unique positive response observed in the bacterial test was due to a specific bacterial metabolism of the test substance could be explored further by investigation *in vitro.* Alternatively, an *in vivo* test may be required (see below).

In REACH Annex VIII, following a positive result in an *in vitro* mammalian cell mutagenicity test, adequately conducted somatic cell *in vivo* testing is required to ascertain if this potential can be expressed *in vivo*. In cases where it can be sufficiently deduced that a positive *in vitro* finding is not relevant for *in vivo* situations (*e.g.* due to the effect of the test substances on pH or cell viability, *in vitro*-specific metabolism: see also Section R.7.7.4.1), or where a clear threshold mechanism has been identified (*e.g.* damage to non-DNA targets at high concentrations that will not be reached *in vivo*), *in vivo* testing will not be necessary.

Annex VIII, Column 2 requires the registrant to consider appropriate mutagenicity *in vivo* studies already at the Annex VIII tonnage level, in cases where positive results in genotoxicity studies have been obtained. It should be noted that where this involves tests mentioned in Annexes IX or X, such as *in vivo* somatic cell genotoxicity studies, testing proposals must be submitted by the registrant and accepted by ECHA in a formal decision before testing can be initiated.

Standard information requirement according to Annexes IX and X

According to the requirements of Annexes IX and X, if there is a positive result in any of the *in vitro* studies from Annex VII or VIII and there are no appropriate results available from an *in vivo* study already, an appropriate *in vivo* somatic cell genotoxicity study should be proposed.

Before any decisions are made about the need for *in vivo* testing, a review of the *in vitro* test results and all available information on the toxicokinetic and toxicodynamic profile of the test substance is needed. A particular *in vivo* test should be conducted only when it can be reasonably expected from all the properties of the test substance and the proposed test protocol that the specific target tissue will be adequately exposed to the test substance and/or its metabolites. If necessary, a targeted investigation of toxicokinetics should be conducted before progressing to *in vivo* testing (*e.g*. a preliminary toxicity test to confirm that absorption occurs and that an appropriate dose route is used).

In the interest of ensuring that the number of animals used in genotoxicity tests is kept to a minimum, both males and females should not automatically be used. In accord with standard guidelines, testing in one sex only is possible when the substance has been investigated for general toxicity and no sex-specific differences in toxicity have been observed. If the test is performed in a laboratory with substantial experience and historical data, it should be considered whether a concurrent positive control and a concurrent negative control for all time points (*e.g*. for both the 24h and 48h time point in the micronucleus assay) will really be necessary (Hayashi *et al*., 2000).

For test substances with adequate systemic availability (*i.e*. evidence for adequate availability to the target cells) there are several options for the *in vivo* testing:

- A rodent bone marrow or mouse peripheral blood micronucleus test (OECD TG 474) or a rodent bone marrow chromosome aberration study (OECD TG 475). The micronucleus test has the advantage of detecting not only structural chromosomal aberrations (clastogenicity) but also numerical chromosomal aberrations (aneuploidy). Potential species-specific effects may also influence the choice of species and test method used.

- A transgenic rodent (TGR) mutation assay (OECD TG 488). TGR assays measure gene mutations and chromosomal rearrangements (the latter specifically in the plasmid and Spi- assay models) using reporter genes present in every tissue. In principle every tissue can be sampled, including target tissues and specific site of contact tissues.

- A comet (single cell gel electrophoresis) assay (OECD TG in development), which detects DNA strand breaks and alkali labile DNA lesions. This assay has the advantage of not being restricted to bone marrow cells. In principle every tissue from which single cell or nuclei suspensions can be prepared can be sampled, including specific site of contact tissues.

- Other DNA strand breakage assays may be presented as alternatives to the comet assay. All DNA strand break assays should be considered as surrogate tests, they do not necessarily detect permanent changes to DNA.

- A rat liver Unscheduled DNA synthesis (UDS) test (OECD TG 486). The UDS test is an indicator test measuring DNA repair of primary damage in liver cells but not a surrogate test for gene mutations *per se*. The UDS can detect some substances that induce *in vivo* gene mutation because this assay is sensitive to some (but not all) DNA repair mechanisms. However not all gene mutagens are positive in the UDS test and it is thus useful only for some classes of substances. A positive result in the UDS assay can indicate exposure of the liver DNA to and induction of DNA damage by the substance under investigation but it is not sufficient information to conclude on the induction of gene mutation by the substance. A negative result in a UDS assay alone is not a proof that a substance does not induce gene mutation.

Only the first two options for testing mentioned above can be used directly for providing evidence of *in vivo* chromosomal and gene mutagenicity, respectively. The other test methods require specific supporting information to be used for making definitive conclusions about *in vivo* mutagenicity and lack thereof.

In the framework of the 3Rs principles, the combination of *in vivo* genotoxicity studies or integration of *in vivo* genotoxicity studies into repeated dose toxicity studies, whenever possible and when scientifically justified, is strongly encouraged if this is to be performed to meet the requirements of the REACH Annex VIII tonnage level. All the above-mentioned *in vivo* tests for somatic cells are in principle amenable to such integration although sufficient experience is not yet available for all of the tests. It is possible for two or more endpoints to be combined into a single *in vivo* study, and thereby save on resources and numbers of animals used. The comet assay and the *in vivo* micronucleus test can be combined into a single acute study, although some modification of treatment and sampling times is needed (Hamada *et al*., 2001; Madrigal-Bujaidar *et al*., 2008; Pfuhler *et al.*, 2009; Bowen *et al*., 2011,). These same endpoints can be integrated into repeated dose (*e.g.* 28-day) toxicity studies (Pfuhler *et al*., 2009; Rothfuss *et al.*, 2011; EFSA, 2011).

Any one of these tests may be conducted, but this has to be decided using expert judgement on a case-by-case basis. The nature of the original *in vitro* response(s) (*i.e.* gene mutation, structural or numerical chromosome aberration) should be considered when selecting the *in vivo* study. For example, if the test substance showed evidence of *in vitro* clastogenicity, then it would be appropriate to follow this up with either a micronucleus test or chromosomal aberration test or a comet assay. However, if a positive result were obtained in the *in vitro* micronucleus test, the rodent micronucleus test would be appropriate to best address clastogenic and aneugenic potential.

For substances that appear preferentially to induce gene mutations, the TGR assays are the most appropriate and usually preferred tests to follow-up an *in vitro* gene mutation positive result and detect, *in vivo*, substances that induce gene mutation. With respect to the 3Rs principle and taking into account that a positive result in somatic cells triggers the need to consider the potential for germ cell testing, germ cells should always be collected, if possible, when a TGR study is performed. The rat liver UDS test has a long history of use and may in some specific cases be adequate to follow-up an *in vitro* gene mutation positive result, but not for tissues other than the liver. The sensitivity of the UDS test has been questioned (Kirkland and Speit, 2008) and the use of this test should be justified on a case-by-case basis, and take account of substance-specific considerations. Discussions on the recommended use of the comet assay are ongoing at the OECD level. The choice of any of these three assays can be

justified only if it can be demonstrated that the tissue(s) studied in the assay is (are) sufficiently exposed to the test substance (or its metabolites). This information can be derived from toxicokinetic data or, in case no toxicokinetic data are available, from the observation of treatment-related effects in the organ of interest. Another type of data that can support evidence of organ exposure is knowledge on the target organ(s) of specific classes of substances (*e.g.* the liver for aromatic amines). In case the *in vivo* comet assay is used or proposed by the registrant, and since no adopted OECD TG is yet available, the test protocol followed or suggested should be described in detail and be in accordance with current scientific best practice, so as to ensure acceptability of the generated data. In waiting for the adoption of the comet OECD TG the registrant should follow the EFSA guidance indicating the minimum criteria for acceptance of the comet assay (2012), as well as, for the combined comet-micronucleus test, the 3-day treatment schedule described by *e.g.* Bowen *et al.* (2011). The TGR and comet assays offer greater flexibility than the UDS test, most notably with regard to the possibility of selecting a range of tissues for study on the basis of what is known of the toxicokinetics and toxicodynamics of the substance. It should be realised that the UDS and comet tests are indicator assays: the comet assay detects DNA lesions whereas the UDS assay detects DNA repair patches (which depend on the DNA repair pathway involved and the proficiency of the cell type investigated), indirectly showing DNA lesions. In contrast, the TGR gene mutation assays measure heritable mutations.

Additionally, evidence for *in vivo* DNA adduct formation in somatic cells together with positive results from *in vitro* mutagenicity tests are sufficient to conclude that a substance is an *in vivo* somatic cell mutagen. In such cases, positive results from *in vitro* mutagenicity tests may not trigger further *in vivo* somatic tissue testing, and the substance would be classified at least as a category 2 mutagen. The possibility for effects in germ cells would need further investigation (see Section R.7.7.6.3, *Substances that give positive results in an in vivo test for genotoxic effects in somatic cells*).

Non-standard studies supported by published literature may sometimes be more appropriate and informative than established assays. Guidance from an appropriate expert or authority should be sought before undertaking novel studies. Furthermore, additional data that support or clarify the mechanism of action may justify a decision not to test further.

For substances inducing gene mutation or chromosomal aberration *in vitro*, and for which no indication of sufficient systemic availablity has been presented, or that are short-lived or reactive, an alternative strategy involving studies to focus on tissues at initial sites of contact with the body should be considered. Expert judgement should be used on a case-by-case basis to decide which tests are the most appropriate. The main options are the *in vivo* comet assay, TGR gene mutation assays, and DNA adduct studies. For any given substance, expert judgement, based on all the available toxicological information, will indicate which of these tests are the most appropriate. The route of exposure should be selected that best allows assessment of the hazard posed to humans. For insoluble substances, the possibility of release of active molecules in the gastrointestinal tract may indicate that a test involving the oral route of administration is particularly appropriate.

If the testing strategy described above has been followed and the first *in vivo* test is negative, the need for a further *in vivo* somatic cell test should be considered. The second *in vivo* test should only then be proposed if it is required to make a conclusion on the genotoxic potential of the substance under investigation; *i.e.* if the *in vitro* data show the substance to have potential to induce both gene and chromosome mutations and the first *in vivo* test has not addressed this comprehensively. In this regard, on a case-by-case basis, attention should be paid to the quality and relevance of all the available toxicological data, including the adequacy of target tissue exposure.

For a substance giving negative results in adequately conducted, appropriate *in vivo* test(s), as defined by this strategy, it will normally be possible to conclude that the substance is not an *in vivo* mutagen.

Substances that give positive results in an *in vivo* test for genotoxic effects in somatic cells

Substances that have given positive results in cytogenetic tests both *in vitro* and *in vivo* can be studied further to establish whether they specifically act as aneugens, and therefore whether thresholds for their genotoxic activity can be identified, if this has not been established adequately already. This should be done using *in vitro* methods and will be helpful in risk evaluation.

The potential for substances that give positive results in *in vivo* tests for genotoxic effects in somatic cells to affect germ cells should always be considered. The same is true for substances otherwise classified as category 2 mutagens under the CLP Regulation (EC) No 1272/2008 (for detailed information on the criteria for classification of substances for germ cell mutagenicity under the CLP Regulation (EC) No 1272/2008, see Section 3.5 of the *Guidance on the application of the CLP criteria*, available at http://echa.europa.eu/web/guest/guidance-documents/guidance-on-clp). The first step is to make an appraisal of all the available toxicokinetic and toxicodynamic properties of the test substance. Expert judgement is needed at this stage to consider whether there is sufficient information to conclude that the substance poses a mutagenic hazard to germ cells. If this is the case, it can be concluded that the substance may cause heritable genetic damage and no further testing is justified. Consequently, the substance is classified as a category 1B mutagen. If the appraisal of mutagenic potential in germ cells is inconclusive, additional investigation will be necessary. In the event that additional information about the toxicokinetics of the substance would resolve the problem, toxicokinetic investigation (*i.e.* not a full toxicokinetic study) tailored to address this should be performed. Although the hazard class for mutagenicity primarily refers to germ cells, the induction of genotoxic effects at site of contact tissues by substances which have no relevant systemic availability to reach the germ cells are also relevant and considered for classification. For such substances, at least one positive *in vivo* genotoxicity test in somatic cells can lead to classification in Category 2 germ cell mutagens and to the labelling as 'suspected of causing genetic defects' if the positive effect *in vivo* is supported by positive results of *in vitro* mutagenicity tests. Classification as Category 2 germ cell mutagen may also have implications for potential carcinogenicity classification.

If specific germ cell testing is to be undertaken, expert judgement should be used to select the most appropriate test strategy. Internationally recognised guidelines are available for investigating clastogenicity in rodent spermatogonial cells and for the dominant lethal test. Dominant lethal mutations are believed to be primarily due to structural or numerical chromosome aberrations.

Alternatively, other methods can be used if deemed appropriate by expert judgement. These may include the TGR gene mutation assays (with modified sampling times as indicated in the OECD TG 488 to detect effects at the different stages of spermatogenesis), or DNA adduct analysis. The use of the comet assay for germ cell testing is still under discussion at the OECD level. In principle, it is the potential for effects that can be transmitted to the progeny that should be investigated, but tests used historically to investigate transmitted effects (the heritable translocation test and the specific locus test) use very large numbers of animals. They are rarely used and should normally not be proposed for substances registered under REACH.

In order to minimise animal use, it is recommended to include cell samples from both relevant somatic and germ cell tissues (*e.g.* testes) in *in vivo* mutagenicity studies: the somatic cell samples can be investigated first and, if they are positive, germ cell tissues can then also be analysed. Finally, the possibility to combine reproductive toxicity testing with *in vivo* mutagenicity testing could be considered.

Figure R.7.7-1 Flow chart of the mutagenicity testing strategy

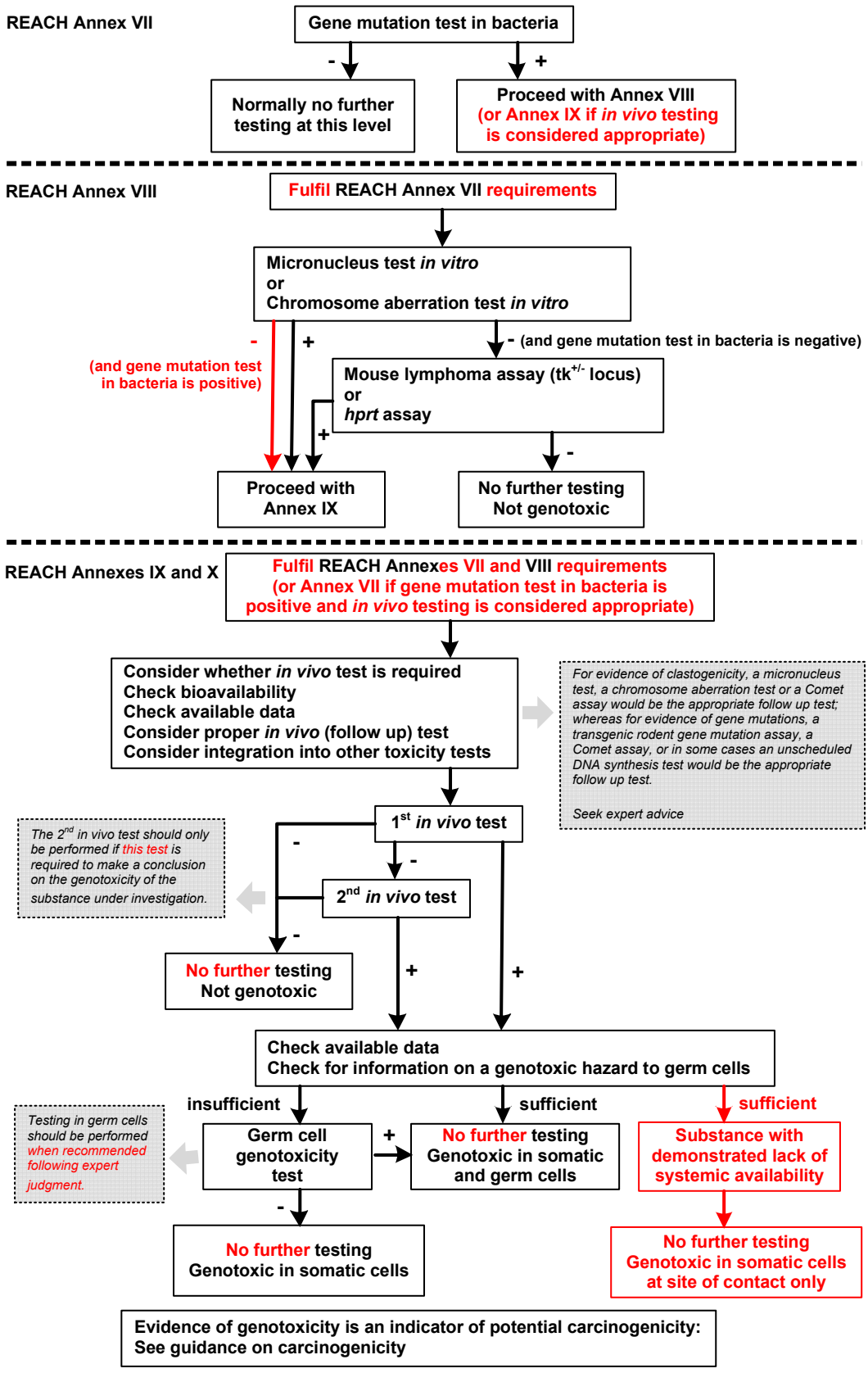

Table R.7.7-5 Examples of different testing data sets and follow-up procedures to conclude on genotoxicity/mutagenicity according to the mutagenicity testing strategy. Depending on the *in vitro* and *in vivo* test results available and the REACH Annex(es) of interest, further testing may be required to meet the standard information requirements for mutagenicity and allow for a conclusion on genotoxicity/mutagenicity to be reached. Recommendations on what should be done or particurlarly looked at in those different cases are mentioned in the table, together with specific rules for adaptation when applicable (for detailed guidance see also main text).

	GM bact	Cytvitro	GMvitro	Cytvivo	GMvivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation [for detailed guidance, incl. timing of the tests, see main text]	Comments
1	neg					Annex VII: no further tests are required. Annexes VIII, IX & X: conduct a CAbvitro or preferably a MNTvitro, and if this is negative, a GMvitro.	Annex VII: not genotoxic		Annexes VIII, IX & X: Select further tests in such a way that all the tests, together with other available information, enable thorough assessment for gene mutations and effects on chromosome structure and number.
2	neg	neg				Annex VII: no further tests are required. Annexes VIII, IX & X: conduct a GMvitro.	Annex VII: not genotoxic		Annexes VIII, IX & X: Select tests in such a way that all the tests, together with other available information, enable a thorough assessment for gene mutations and effects on chromosome structure and number.
3	neg		neg			Annex VII: no further tests are required. Annexes VIII, IX & X: conduct a CAbvitro or preferably a MNTvitro	Annex VII: not genotoxic		Annexes VIII, IX & X: Select tests in such a way that all the tests, together with other available information, enable a thorough assessment for gene mutations and effects on chromosome structure and number.

	GM bact	Cytvitro	GMvitro	Cytvivo	GMvivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation [for detailed guidance, incl. timing of the tests, see main text]	Comments
4	neg	neg	neg			Annexes VII, VIII, IX & X: no further tests are required.	not genotoxic		The available metabolic evidence may, on rare occasions, indicate that *in vitro* testing is inadequate*; in vivo* testing is needed. Seek expert advice. Annexes VIII, IX & X: Select tests in such a way that all the tests, together with other available information, enable a thorough assessment for gene mutations and effects on chromosome structure and number.
5	pos					Annexes VII, VIII, IX & X: Complete *in vitro* testing with a CAbvitro or preferably a MNTvitro.			Consider need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment, and to determine whether C&L is justified.

	GM bact	Cytvitro	GMvitro	Cytvivo	GMvivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation [for detailed guidance, incl. timing of the tests, see main text]	Comments
6	pos	neg				Annexes VII & VIII: *Complete in vitro* testing by conducting a GMvitro only under special conditions (see column 'Specific rules for adaption') Annexes IX & X: If systemic availability cannot be ascertained reliably, it should be investigated before progressing to *in vivo* tests. Select adequate somatic cell i*n vivo* test to investigate gene mutations *in vivo* (TGR, or if justified UDSvivo or comet). If the TGR is to be conducted, germ cell samples should be included as well in case of a positive result in somatic cells and germ cell mutagenicity testing would need to be considered. If necessary seek expert advice.		Suspicion that a positive response observed in the GMbact was due to a specific bacterial metabolism of the test substance could be explored further by investigation *in vitro*.	Ensure that all tests together with other available information enable thorough assessment for gene mutations and effects on chromosome structure and number. Consider on a case-by-case basis need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment, and to determine whether C&L is justified.
7	neg	pos				Annexes VII, VIII, IX & X: If systemic availability cannot be ascertained reliably, it should be investigated before progressing to *in vivo* tests. Select adequate somatic cell *in vivo* test to investigate structural or numerical chromosome aberrations (MNTvivo or comet for *in vitro* clastogens and/or aneugens or CAbvivo for *in vitro*-clastogens) If necessary seek expert advice.			Ensure that all tests together with other available information enable thorough assessment for gene mutations and effects on chromosome structure and number. Consider need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment and to determine whether C&L is justified.

	GM bact	Cytvitro	GMvitro	Cytvivo	GMvivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation [for detailed guidance, incl. timing of the tests, see main text]	Comments
8	pos	pos				Annexes VII, VIII, IX & X: If systemic availability cannot be ascertained with acceptable reliability, it should be investigated before progressing to *in vivo* tests. Select adequate somatic cell *in vivo* tests to investigate both structural or numerical chromosome aberrations and gene mutations. If necessary seek expert advice.		Generally, both genotoxic endpoints should be investigated. If the first *in vivo* test is positive, a second *in vivo* test to confirm the other genotoxic endpoint need not be conducted. If the first *in vivo* test is negative, a second *in vivo* test is required if the first test did not address the endpoints comprehensively.	Ensure that all tests together with other available information enable thorough assessment for gene mutations and effects on chromosome structure and number. Consider need for further tests to understand the *in vivo* mutagenicity hazard, to make a risk assessment, and to determine whether C&L is justified.
9	pos	neg			neg	Annexes VII, VIII, IX & X: no further tests are required.	not genotoxic		Further *in vivo* test may be necessary pending on the quality and relevance of available data.
	neg	pos		neg					
10	pos	neg			pos	Annexes VII, VIII, IX & X: No further testing in somatic cells is needed. Germ cell mutagenicity tests should be considered. If necessary seek expert advice on implications of all available data on toxicokinetics and toxicodynamics and on the choice of the proper germ cell mutagenicity test.	genotoxic	Expert judgement is needed at this stage to consider whether there is sufficient information to conclude that the substance poses a mutagenic hazard to germ cells. If this is the case, it can be concluded that the substance may cause heritable genetic damage and no further testing is justified.	If the appraisal of mutagenic potential in germ cells is inconclusive, additional investigation may be necessary. Risk assessment and C&L can be completed.
	neg	pos		pos					
	neg	neg	pos		pos				

	GM bact	Cytvitro	GMvitro	Cytvivo	GMvivo	Standard information required General follow up procedure	Conclusion	Specific rules for adaptation [for detailed guidance, incl. timing of the tests, see main text]	Comments
11	pos	pos	(pos)	pos		Annexes VII, VIII, IX & X: No further testing in somatic cells is needed. Germ cell mutagenicity tests should be considered. If necessary seek expert advice on implications of all available data on toxicokinetics and toxicodynamics and on the choice of the proper germ cell mutagenicity test.	genotoxic	Expert judgement is needed at this stage to consider whether there is sufficient information to conclude that the substance poses a mutagenic hazard to germ cells. If this is the case, it can be concluded that the substance may cause heritable genetic damage and no further testing is justified.	If the appraisal of mutagenic potential in germ cells is inconclusive, additional investigation may be necessary. Risk assessment and C&L can be completed.
	pos	pos	(pos)		pos				
12	pos	pos	(pos)	neg		Annexes VII, VIII, IX & X: Select adequate somatic cell *in vivo* tests to investigate both structural or numerical chromosome aberrations and gene mutations. If necessary seek expert advice.			
	pos	pos	(pos)		neg				
13	pos	pos	(pos)	neg	neg	Annexes VII, VIII, IX & X: no further tests are required.	not genotoxic	Further *in vivo* test may be necessary pending on the quality and relevance of available data.	Risk assessment and C&L can be completed.
14	pos	pos	(pos)	neg	pos	Annexes VII, VIII, IX & X: No further testing in somatic cells is needed. Germ cell mutagenicity tests should be considered. If necessary seek expert advice on implications of all available data on toxicokinetics and toxicodynamics and on the choice of the proper germ cell mutagenicity test.	genotoxic	Expert judgement is needed at this stage to consider whether there is sufficient information to conclude that the substance poses a mutagenic hazard to germ cells. If this is the case, it can be concluded that the substance may cause heritable genetic damage and no further testing is justified.	If the appraisal of mutagenic potential in germ cells is inconclusive, additional investigation will be necessary. Risk assessment and C&L can be completed.
	pos	pos	(pos)	pos	neg				

Abbreviations: pos: positive; neg: negative; (pos): the follow up is independent from the result of this test; GM_{bact}: gene mutation test in bacteria (Ames test); Cyt_{vitro}: cytogenetic assay in mammalian cells; CAb_{vitro}: *in vitro* chromosome aberration test; MNT_{vitro}: *in vitro* micronucleus test; GM_{vitro}: gene mutation assay in mammalian cells; Cyt_{vivo}: cytogenetic assay in experimental animals; GM_{vivo}: gene mutation assay in experimental animals; CAb_{vivo}: *in vivo* chromosome aberration test (bone marrow); MNT_{vivo}: *in vivo* micronucleus test (erythrocytes); UDS_{vivo}: *in vivo* unscheduled DNA synthesis test; TGR: *in vivo* gene mutation test with transgenic rodent; comet: comet assay.

R.7.7.7 References on mutagenicity

Ashby J and Tennant RW (1988) Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. Mutat Res 204: 17-115.

Azqueta A, Gutzkow KB, Brunborg G and Collins AR (2011a) Towards a more reliable comet assay: optimising agarose concentration, unwinding time and electrophoresis conditions. Mutat Res 724: 41-5.

Azqueta A, Meier S, Priestley C, Gutzkow KB, Brunborg G, Sallette J, Soussaline F and Collins A (2011b) The influence of scoring method on variability in results obtained with the comet assay. Mutagenesis 26: 393-9.

Azqueta A, Gutzkow KB, Priestley CC, Meier S, Walker JS, Brunborg G and Collins AR (2013) A comparative performance test of standard, medium- and high-throughput comet assays. Toxicol In Vitro 27: 768-73.

Bakhtyari NG, Raitano G, Benfenati E, Martin T and Young D (2013) Comparison of in silico models for prediction of mutagenicity. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev 31: 45-66.

Bowen DE, Whitwell JH, Lillford L, Henderson D, Kidd D, Mc Garry S, Pearce G, Beevers C and Kirkland DJ (2011) Work conducted at Covance Laboratories Ltd., Harrogate. Evaluation of a multi-endpoint assay in rats, combining the bone-marrow micronucleus test, the Comet assay and the flow-cytometric peripheral blood micronucleus test. Mutat Res 722: 7-19.

Brendler-Schwaab S, Hartmann A, Pfuhler S and Speit G (2005) The *in vivo* comet assay: use and status in genotoxicity testing. Mutagenesis 20: 245-54.

Burlinson B, Tice RR, Speit G, Agurell E, Brendler-Schwaab SY, Collins AR, Escobar P, Honma M, Kumaravel TS, Nakajima M, Sasaki YF, Thybaud V, Uno Y, Vasquez M and Hartmann A (2007) Fourth International Workgroup on Genotoxicity Testing: result of the in vivo Comet assay workgroup. Mutat Res 627:31-5.

Burlinson B (2012) The *in vitro* and *in vivo* comet assays. Methods Mol Biol 817:143-63.

Collins AR, Oscoz AA, Brunborg G, Gaivão I, Giovannelli L, Kruszewski M, Smith CC, Stetina R (2008) The comet assay: topical issues. Mutagenesis 23: 143-51.

EFSA, European Food Safety Authority (2011) Scientific Opinion of the Scientific Committee on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal 2011 9(9):2379 [69 pp.] Available online: www.efsa.europa.eu/efsajournal

EFSA, European Food Safety Authority (2012) Minimum Criteria for the acceptance of in vivo alkaline Comet Assay Reports. EFSA Journal 2012 10(11):2977 [12 pp.] Available online: www.efsa.europa.eu/efsajournal

Enoch SJ (2010) Chemical Category Formation and Read-Across for the Prediction of Toxicity *In:* Puzyn T, Leszczynski J and Cronin MTD (eds), Recent Advances in QSAR Studies, Volume 8, pp 209-219.

Ersson C, Møller P, Forchhammer L, Loft S, Azqueta A, Godschalk RW, van Schooten FJ, Jones GD, Higgins JA, Cooke MS, Mistry V, Karbaschi M, Phillips DH, Sozeri O, Routledge MN, Nelson-Smith K, Riso P, Porrini M, Matullo G, Allione A, Stepnik M, Ferlińska M, Teixeira JP, Costa S, Corcuera LA, López de Cerain A, Laffon B, Valdiglesias V, Collins AR and Möller L (2013) An ECVAG inter-laboratory validation study of the comet assay: inter-laboratory and intra-laboratory variations of DNA strand breaks and FPG-sensitive sites in human mononuclear cells. Mutagenesis 28: 279-86.

Johansson C, Møller P, Forchhammer L, Loft S, Godschalk RW, Langie SA, Lumeij S, Jones GD, Kwok RW, Azqueta A, Phillips DH, Sozeri O, Routledge MN, Charlton AJ, Riso P, Porrini M, Allione A, Matullo G, Palus J, Stepnik M, Collins AR and Möller L (2010) An ECVAG trial on assessment of oxidative damage to DNA measured by the comet assay. Mutagenesis 25: 125-32.

Forchhammer L, Johansson C, Loft S, Möller L, Godschalk RW, Langie SA, Jones GD, Kwok RW, Collins AR, Azqueta A, Phillips DH, Sozeri O, Stepnik M, Palus J, Vogel U, Wallin H, Routledge MN, Handforth C, Allione A, Matullo G, Teixeira JP, Costa S, Riso P, Porrini M and Møller P (2010) Variation in the measurement of DNA damage by comet assay measured by the ECVAG inter-laboratory validation trial. Mutagenesis 25: 113-23.

Forchhammer L, Ersson C, Loft S, Möller L, Godschalk RW, van Schooten FJ, Jones GD, Higgins JA, Cooke M, Mistry V, Karbaschi M, Collins AR, Azqueta A, Phillips DH, Sozeri O, Routledge MN, Nelson-Smith K, Riso P, Porrini M, Matullo G, Allione A, Stępnik M, Komorowska M, Teixeira JP, Costa S, Corcuera LA, López de Cerain A, Laffon B, Valdiglesias V and Møller P (2012) Inter-laboratory variation in DNA damage using a standard comet assay protocol. Mutagenesis 27: 665-72.

Hamada S, Sutou S, Morita T, Wakata A, Asanami S, Hosoya S, Ozawa S, Kondo K, Nakajima M, Shimada H, Osawa K, Kondo Y, Asano N, Sato S, Tamura H, Yajima N, Marshall R, Moore C, Blakey DH, Schechtman LM, Weaver JL, Torous DK, Proudlock R, Ito S, Namiki C and Hayashi M (2001) Evaluation of the rodent micronucleus assay by a 28-day treatment protocol: Summary of the 13th Collaborative Study by the Collaborative Study Group for the Micronucleus Test (CSGMT)/Environmental Mutagen Society of Japan (JEMS)-Mammalian Mutagenicity Study Group (MMS). Environ Mol Mutagen 37:93-110.

Hartmann A, Agurell E, Beevers C, Brendler-Schwaab S, Burlinson B, Clay P, Collins A, Smith A, Speit G, Thybaud V and Tice RR (2003) Recommendations for conducting the *in vivo* alkaline Comet assay. Mutagenesis 18: 45-51.

Hayashi M, MacGregor JT, Gatehouse DG, Adler ID, Blakey DH, Dertinger SD, Krishna G, Morita T, Russo A and Sutou S (2000) In vivo rodent erythrocyte micronucleus assay. II. Some aspects of protocol design including repeated treatments, integration with toxicity testing, and automated scoring. Environ Mol Mutagen 35: 234-52.

Kirkland D and Speit G (2008) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens III. Appropriate follow-up testing in vivo. Mutat Res 654:114-32.

Lambert IB, Singer TM, Boucher SE and Douglas GR (2005) Detailed review of transgenic rodent mutation assays. Mutat Res 590: 1-280.

Madrigal-Bujaidar E, Madrigal-Santillán EO, Alvarez-Gonzalez I, Baez R and Marquez P (2008) Micronuclei induced by imipramine and desipramine in mice: a subchronic study. Basic Clin Pharmacol Toxicol 103:569-73.

McKelvey-Martin VJ, Green MH, Schmezer P, Pool-Zobel BL, De Méo MP and Collins A (1993) The single cell gel electrophoresis assay (comet assay): a European review. Mutat Res 288:47-63.

Naven RT, Greene N and Williams RV (2012) Latest advances in computational genotoxicity prediction. Expert Opin Drug Metab Toxicol 8: 1579-87.

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

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

Pfuhler S, Kirkland D, Kasper P, Hayashi M, Vanparys P, Carmichael P, Dertinger S, Eastmond D, Elhajouji A, Krul C, Rothfuss A, Schoening G, Smith A, Speit G, Thomas C, van Benthem J and Corvi R (2009) Reduction of use of animals in regulatory genotoxicity testing: Identification and implementation opportunities—Report from an ECVAM workshop. Mutat Res 680: 31–42.

Rothfuss A, O'Donovan M, De Boeck M, Brault D, Czich A, Custer L, Hamada S, Plappert-Helbig U, Hayashi M, Howe J, Kraynak AR, van der Leede B-J, Nakajima M, Priestley C, Thybaud V, Saigo K, Sawant S, Shi J, Storer R, Struwe M, Vock E and Galloway S (2010) Collaborative study on fifteen compounds in the rat-liver Comet assay integrated into 2- and 4-week repeat-dose studies. Mutat Res 702: 40-69.

Rothfuss A, Honma M, Czich A, Aardema MJ, Burlinson B, Galloway S, Hamada S, Kirkland D, Heflich RH, Howe J, Nakajima M, O'Donovan M, Plappert-Helbig U, Priestley C, Recio L, Schuler M, Uno Y and Martus HJ (2011) Improvement of in vivo genotoxicity assessment: combination of acute tests and integration into standard toxicity testing. Mutat Res 723:108-20.

Serafimova R, Fuart Gatnik M and Worth A (2010) Review of QSAR models and software tools for predicting genotoxicity and carcinogenicity. JRC Technical Report EUR 24427 EN. Publications Office of the European Union, Luxembourg. http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/doc/EUR_24427_EN.pdf

Smith CC, Adkins DJ, Martin EA and O'Donovan MR (2008) Recommendations for design of the rat comet assay. Mutagenesis 23: 233-40.

Somers CM, Yauk C, White PA, Parfett CLJ and Quinn JS (2002) Air pollution induces heritable DNA mutations. Proc Natl Acad Sci USA 99: 15904-15907.

Tice RR, Agurell E, Anderson D, Burlinson B, Hartmann A, Kobayashi H, Miyamae Y, Rojas E, Ryu J-C and Sasaki F (2000) Single cell gel/Comet assay: guidelines for *in vitro* and *in vivo* genetic toxicology testing. Environ Mol Mutagen 35: 206-221.

Vasquez MZ (2012) Recommendations for safety testing with the *in vivo* comet assay. Mutat Res 747:142-56.

Yauk C (2004) Advances in the application of germline tandem repeat instability for in situ monitoring. Mutat Res 566: 169-182.