define criteria for determining the relevance to human health. The workshop identified specific needs in two general categories, *i.e.*, improved testing and improved data interpretation and risk assessment. Recommendations to improve testing included: (1) re-examine the maximum level of cytotoxicity currently required for *in vitro* tests; (2) re-examine the upper limit concentration for *in vitro* mammalian studies; (3) develop improved testing strategies using current *in vitro* assays; (4) define criteria to guide selection of the appropriate follow-up *in vivo* studies; (5) develop new and more predictive *in vitro* and *in vivo* tests. Recommendations for improving interpretation and assessment included: (1) examine the suitability of applying the threshold of toxicological concern concepts to genotoxicity data; (2) develop a structured weight of evidence approach for assessing genotoxic/carcinogenic hazard; and (3) re-examine *in vitro* and *in vivo* correlations qualitatively and quantitatively. Conclusions from the workshop highlighted a willingness of scientists from various sectors to change and improve the current paradigm and move from a hazard identification approach to a "realistic" risk-based approach that incorporates information on mechanism of action, kinetics, and human exposure.

Keywords: Genotoxicity; *In vitro* assays; Carcinogenesis; Workshop report

1. Introduction

Human exposure to DNA-damaging agents is an important health issue because gene and chromosomal mutations can potentially lead to adverse health consequences, including cancer, reproductive impairment, developmental anomalies, or genetic diseases. Current regulatory practice is to use a battery of genetic toxicity tests to determine if a chemical has the potential to cause mutations or chromosomal damage. Tests conducted *in vitro* in bacteria and mammalian cells play an important role in this battery [1–13]. During the past 15 years, accumulated evidence has shown a high rate of positive results in the *in vitro* tests, especially in those assays performed in mammalian cells [14]. Importantly, a large number of the mammalian *in vitro* positive findings have not been confirmed in *in vivo* genotoxicity and/or carcinogenicity studies, and this raises the question of their specificity and relevance in human risk assessment [15].

Positive *in vitro* results generally lead to costly and time consuming additional testing, including mechanistic studies and *in vivo* genetic toxicity testing in rodent models. In the context of regulating pharmaceuticals, pesticides or industrial chemicals, such positive *in vitro* results could potentially lead to prohibiting the use and/or development of compounds of negligible concern for adverse human effects. It is increasingly accepted that positive results should not be considered in isolation, and that a weight of evidence approach considering all pertinent data should be the preferred approach [4,16–18]. As part of this weight of evidence approach, information on the mode of action, kinetics, and the extent of human exposure is useful for risk assessment.

The low dose portion of the dose–response curve is generally assumed to be linear with no threshold for compounds known to interact with DNA directly (*e.g.*, alkylating and intercalating agents). Other compounds can induce DNA damage as a secondary effect and act through non-DNA reactive mechanisms (*e.g.*, inhibition of topoisomerase, mitotic spindle disruption, inhibition of protein and DNA synthesis, imbalance of nucleotide pools). In this latter case, it is accepted that a threshold dose level exists below which no genotoxic effect is expected to occur. Despite recently accumulated information about possible indirect mechanisms of action [19–22], there is still a need to improve and enhance our understanding of these mechanisms and to provide clear recommendations on the approaches to be used to identify the mode of action and include these data in risk assessment.

It is agreed that *in vitro* models are imperfect models of *in vivo* biology. Nonetheless, business and regulatory decisions are often made on the basis of qualitative outcomes in these assays. In addition, extreme experimental conditions currently recommended in regulatory guidelines for the *in vitro* genotoxicity models are seen as a potential source of artifacts and irrelevant findings (*e.g.*, high level of cytotoxicity, precipitating concentrations, concentrations up to 5000 μg/ml or 10 mM that are very unlikely to be attained *in vivo*). Hence, there is a need for a better understanding of the limitations of the currently used *in vitro* models in order to more adequately interpret the *in vitro* findings, and to identify the key criteria for the development of better and more relevant predictive models for *in vivo* biology.

The need to re-consider the evaluation of *in vitro* positive findings and their impact on risk assessment has been recently highlighted by regulatory authorities [16,17,23–25], and organizations such as the International Workshops on Genotoxicity Testing (IWGT) and European Centre for the Validation of Alternative Methods (ECVAM) [18,26].

Taking all these questions/points into consideration, the Health and Environmental Sciences Institute (HESI), the global branch of the International Life Sciences Institute (ILSI) recently identified "the relevance and

follow-up of positive results in *in vitro* genetic toxicity" as an emerging issue. A HESI subcommittee was formed to address the following key objectives:

- improve the scientific basis of the interpretation of results from *in vitro* genetic toxicology tests for purposes of accurate human risk assessment,
- develop follow-up strategies for determining the relevance of *in vitro* test results to human health, and
- provide a framework for the integration of the *in vitro* testing results into a risk-based assessment of the effects of chemical exposures to human health.

In order to identify the actions to be initiated, HESI organized an international multi-sector workshop in Washington, DC, on June 21 and 22, 2006, which was attended by 45 experts in the fields of genetic toxicology, carcinogenesis and risk assessment. Participants included representatives from the United States Department of Agriculture (USDA), United States Environmental Protection Agency (USEPA), United States Food and Drug Administration (USFDA), Health Canada, Japan National Institute of Health Sciences (NIHS), and the European Food Safety Authority (EFSA), and over 15 companies from various industries involved in the development of products including industrial chemicals, agricultural chemicals, pharmaceuticals, and cosmetics. Several of the participants were also involved in other initiatives such as IWGT or ECVAM, and this workshop facilitated coordination between the different initiatives. Further information about the workshop participants and their affiliations can be found at http://www.hesiglobal.org/Committees/EmergingIssues/ToxTesting.

The program of this workshop consisted of a series of plenary lectures followed by break-out group sessions to address three main topics. These were:

- *Break-out group #1*: how to establish relevance of *in vitro* findings to humans using mechanistic and *in vivo* data.
- *Break-out group #2*: how to factor in quantitative consideration of the impact of dose–response.
- *Break-out group #3*: how to improve our testing for genetic toxicity.

This publication summarizes the three break-out group discussions and deliberations, and provides recommendations by the workshop participants to support the development of more reliable approaches to genetic toxicity risk assessment and risk management. No attempt was made at harmonizing the format of break-out group reports, but Table 1 captures the key recommendations and some of the commonalities shared between the three break-out groups. The conclusion statements will be used as a starting point for the next steps of the collaborative work. The different topics identified will be examined further in depth in the near future, in order to evaluate if new technical and scientific approaches can be used to address the identified issues and questions.

2. Summaries of the break-out group discussions

2.1. Break-out group #1: how to establish relevance of in vitro findings to humans using mechanistic and in vivo data

2.1.1. Break-out group #1: background

The positive results obtained in *in vitro* genotoxicity tests are often not confirmed in *in vivo* genotoxicity and carcinogenicity tests. To predict carcinogenesis from *in vitro* findings, there is a significant need to develop a weight of evidence approach that considers human exposure information and incorporates an understanding of the mechanism of action, metabolism and tissue distribution *in vivo*. The participants of this break-out group evaluated the relevance of *in vitro* data to humans by focusing on the concordance with *in vivo* genotoxicity and carcinogenicity data. This evaluation was done for different endpoints (*e.g.*, gene mutations, chromosome damage, primary DNA damage) with consideration of mechanistic information. The discussion also included the level of information needed to define a threshold and a potential safety margin in humans.

2.1.2. Break-out group #1: report

Genetic toxicology testing is an integral and essential part of the safety evaluation of chemicals (pharmaceuticals, pesticides, industrial chemicals, and consumer products). The primary focus of this testing is to assess the inherent potential of a substance to compromise the integrity of the genetic material. The types of heritable genetic events that are relevant to human risk assessment include gene mutations, structural chromosomal changes, and aneuploidy. Hence, all three end points should be included in human health assessments.

Currently, a battery of short-term tests is initially used to identify the genotoxic potential of a test material. There was a general consensus among the members of the break-out group that this initial battery of tests could be prescriptive. Towards this end, a battery, comprised of three or four tests (*e.g.*, a bacterial reverse mutation test, a test for chromosomal aberrations and/or mutations in mammalian cell cultures, and an *in vivo* test for

Table 1
IVGT break-out group key recommendations

Recommendation	Group #1—how to establish relevance of *in vitro* findings to humans using mechanistic and *in vivo* data	Group #2—how to factor in a quantitative consideration of the impact of dose–response	Group #3—how to improve our testing for genetic toxicity
Improving testing			
• Re-examine maximum level of cytotoxicity currently required for *in vitro* tests to determine scientific validity and evaluate appropriate measures for cytotoxicity	X		X
• Re-examine whether the current 10 mM upper limit concentration for *in vitro* mammalian studies is justified	X—conduct retrospective analysis using animal and human PK data and consider compatibility with cellular metabolic efficiency and enzyme saturation to identify an appropriate top concentration		X—evaluate the validity and applicability of current requirements which were originally based on early analysis of small databases
• Improve testing strategies using current *in vitro* assays so as to more reliably assess genotoxic hazard and predict carcinogenesis		X—understand limitations of *in vitro* tests	X—evaluate whether extended *in vitro* exposures (24 h) are contributing false positives, determine if human lymphocytes are more predictive compared to cell lines, and evaluate most appropriate metabolism system to use
• Define criteria to guide selection of the appropriate follow-up *in vivo* study(ies)	X	X	
• Develop new and more predictive *in vitro* and *in vivo* tests that could ultimately be used in addition or as a replacement of current models			X—p53 DNA repair proficient cells, metabolically relevant cells, systems with multi-endpoint analyses
Improving data interpretation and risk assessment			
• Examine the suitability of applying the threshold of toxicological concern (TTC) concepts to genotoxicity data	X—Examine the suitability of deriving benchmark doses, NOAELs, LOAELs from genotoxicity data	X—evaluate different mechanistic classes to identify thresholds or acceptable margins of exposure; evaluate available data to examine the scientific support for low-dose linearity	
• Develop a structured weight of evidence approach with robust qualitative and quantitative criteria for assessing genotoxic/carcinogenic hazard	X—consider structural alerts, assay strength/weakness, consistency and reproducibility of findings, *etc.*	X—utilize *in vivo* and *in vitro* dose–response data and human exposure information to characterize and bin levels of concern	
• Re-examine *in vitro* and *in vivo* correlations and assess the ability of current genotoxicity assays to predict carcinogenicity outcomes using thorough analysis and robust criteria	X—mine databases and conduct retrospective analyses to determine the value of *in vitro* and *in vivo* tests in predicting the outcomes of animal cancer studies	X—expand retrospective analysis to examine dose–response (*i.e.*, vivo potencies with *in vitro* concentration effects) by chemical class and type of damage	X—retrospective analysis should be based on a compiled dataset that includes ONLY current acceptance and interpretation criteria

X: Recommendation made by break-out group.

cytogenetic damage in bone marrow of rodents), as prescribed by a number of expert/regulatory bodies around the globe (*e.g.*, ICH, EEC, U.S.EPA, JMHLW, *etc.*), are considered to be still valid. For certain molecules, one or more of the above tests may not be relevant or useful (*e.g.*, bacterial reverse mutation test for peptides). Positive findings (*i.e.*, identification of genotoxicity) in one or more of the initial tests may require further investigation and usually trigger additional testing. However, the selection of additional test(s) at this stage cannot be prescriptive and should be handled on a case-by-case basis.

There was also a general consensus that the protocols and data interpretation strategies currently used in the conduct of the initial battery of tests needed improvement. For example, the rationale for the selection of the top concentration including the required levels of cytotoxicity currently prescribed by various regulatory guidelines for *in vitro* tests may need further examination. In this context, the group identified a need to undertake a retrospective analysis of any available internal exposure data from animal toxicokinetic and, where available, human pharmacokinetic studies to help establish a general guidance to limit the highest concentrations that need to be evaluated in *in vitro* genotoxicity studies utilizing mammalian cell cultures. Using information from such an analysis, one might be able to identify an appropriate top concentration to be used in *in vitro* assays (instead of the current 10 mM limit) that could generally be agreed not to be excessively above the typical pharmacologically active range for most drugs, above the K_ms for most relevant enzymes including those involved in metabolic activation/detoxification, and above the typical blood and tissue levels expected at the most extreme human exposures that would occur in actual usage situations. While this approach might be suitable in most cases, there could be instances where higher concentrations may need to be evaluated with certain agents which may have potentially extreme human exposures.

The group acknowledged that toxicokinetic and toxicodynamic considerations dictate that findings from a well conducted *in vivo* genetic toxicity test that evaluates relevant endpoints and target tissues should carry more emphasis or weight than conflicting results from corresponding *in vitro* assays. At this time, however, there are no validated *in vivo* protocols amenable to assess all genetic events of human relevance (*i.e.*, mutations, chromosomal aberrations, and aneuploidy) in multiple tissues. Because of this limitation an integrative approach of different *in vitro* genetic toxicology studies will continue to play an important role in safety assessment programs. Furthermore, although analyses performed to date suggest that the results from *in vivo* tests correlate better than *in vitro* tests in predicting the outcomes of animal carcinogenicity studies, an exhaustive analysis of all available databases has not yet been performed that would allow one to make a definitive conclusion on this issue. Therefore, this group identified such an analysis as a worthwhile future activity.

Genetic toxicology is an integral part of the field of toxicology and as such the general weight of evidence principles of data interpretation widely accepted for other toxicities should be equally applicable to genotoxicity (*i.e.*, considering all pertinent information when available including, metabolism, kinetics, mechanism, dose–response and human exposure, placing emphasis on reliable *in vivo* results over *in vitro* findings, and acknowledging data limitations). The central dogma in toxicology is that it is the dose that determines the risk of toxicity. Accordingly, an experimentally derived no-adverse-effect-level (NOAEL) or mathematically modeled "bench mark dose (BMD)" in conjunction with a set of uncertainty factors usually forms the basis to establish a human exposure level to toxicants without an expectation of an adverse outcome. The group had a cursory discussion on the applicability of such an approach to mutagenicity data, irrespective of the mode of action of an agent (*i.e.*, even for DNA reactive mutagens which are currently excluded from this approach), for setting acceptable exposure levels. The following uncertainties were discussed in the use of experimental data to derive allowable human exposure levels: (1) extrapolation from *in vitro* to *in vivo* situations, (2) extrapolation from non-human species to humans, (3) existence of susceptible subpopulations among humans, (4) severity of the effect studied, and (5) deficiency in the database used to derive the NOAEL or BMD. The group concluded that further discussion is needed on the suitability of this approach as well as the use of factors (*e.g.*, 3× or 10×) to account for each of the identified uncertainties to be used in deriving the permissible exposure levels.

2.1.3. Break-out group #1: conclusions and recommendations

Based on the deliberations described above, break-out group #1 proposed the following conclusions and recommendations:

- Critically examine the currently required maximum level of cytotoxicity in *in vitro* mammalian assays.
- Re-evaluate the current 10 mM upper limit concentration for *in vitro* mammalian studies using a retrospective analysis, taking into account the following:

- animal and human pharmacokinetic data;
- metabolic efficiency;
- enzyme saturation;
- typical blood and tissue levels at the most extreme human exposure situations.

- Apply general weight of evidence principles of data interpretation accepted for other types of toxicity to genotoxicity data, considering:
  - metabolism;
  - kinetics;
  - mechanism;
  - dose–response and human exposure;
  - placing emphasis on reliable *in vivo* results over *in vitro* findings;
  - acknowledging data limitations.
- Critically examine the suitability of applying the concepts of benchmark dose, NOAELs, LOAELs, and uncertainty factors to genotoxicity data.
- Conduct a retrospective in-depth review of the available genotoxicity databases to better understand the respective contribution of *in vitro* and *in vivo* assays to the prediction of carcinogenic potential.

2.2. *Break-out group #2: how to factor in a quantitative consideration of the impact of dose–response*

2.2.1. *Break-out group #2: background*

The participants of this break-out group focused on how to use knowledge of the *in vivo* factors that determine genotoxic responses (including exposure, pharmacokinetics, metabolism, and mechanism) to interpret responses in *in vitro* and *in vivo* laboratory genetic toxicology tests, and to improve estimation of the risk of genetic damage and/or adverse health outcomes in humans. The discussion was focused on situations in which decisions must be made in the absence of carcinogenicity data, such as (a) the stage of pharmaceutical development at which *in vitro* and limited *in vivo* genetic toxicology information, but no carcinogenicity data, are available or (b) screening of industrial chemicals, and included (c) the near future of cosmetic and health-care product regulation in which decisions may need to be made primarily on the basis of *in vitro* data or at most with some limited *in vivo* data.

2.2.2. *Break-out group #2: report*

The general question addressed by this group was: "What information on exposure and genotoxicity such as potency, nature of genetic lesion, shape of dose–response curve and mode of action is needed to define acceptable exposure levels or levels of no concern for exposed humans". The specific questions addressed included:

- Is there a quantitative relationship between potency *in vitro* and potency *in vivo* for induction of the types of damage of interest (*e.g.*, adducts, strand breaks, nucleotide alterations, mutations, chromosomal aberrations, *etc.*), for (1) agents that do not require metabolic activation and (2) agents that do require metabolic activation?
- Can a combination of *in vitro* potency data, with or without *in vivo* potency data, and human exposure data provide an index of risk that supports regulatory decision-making in the absence of carcinogenicity data?
- By using such an index, can a level of risk be defined that is considered inconsequential or acceptable for a given human exposure?

Additionally, the general default assumptions about the shape of the dose–response curves for "genotoxic" *versus* "non-genotoxic" agents were discussed, including the assumptions that (1) agents that react with or "directly" damage, DNA should be assumed to have linear dose–response relationships as a conservative default, whereas (2) DNA non-reactive agents, *i.e.*, genotoxicants that act through a primary target or mechanism other than direct reaction with DNA, are considered likely to have non-linear dose–response relationships with a definable "threshold" below which *in vivo* risk of damage can be considered negligible. In particular, the extent of the scientific data available that supports these presumptions was questioned. It was felt that some generalizations can be made with regard to type of damage, mechanism of action, and/or class of agents, but that a more rigorous evaluation of the situations and conditions involved was needed. For certain classes of chemicals, such as specific DNA synthesis inhibitors or agents that interact very specifically with known non-DNA targets, it was agreed that a threshold below which significant DNA damage would not occur could be defined. Systematic approaches to evaluate available data that supports improved categorization of chemical classes and supports appropriate assumptions about expected dose–response relationships are recommended (see below).

Some data were presented suggesting that there are practical thresholds even for DNA-reactive genotoxic agents [27,28] and showing also that agents considered non-genotoxic, such as sucrose, can produce significant effects even *in vivo* if sufficient exposure is achieved [29]. The limited data presented suggests that a more comprehensive survey and analysis of results in the liter-

ature and those available from HESI member companies is warranted in order to determine whether practical thresholds can be defined for DNA-reactive genotoxic agents. Based on this analysis, an informed decision can be made as to whether it is necessary to move from the current practice of decision-making on the basis of qualitative or semi-quantitative characterization of agents to a more quantitative assessment of genotoxic risk under defined exposure conditions *in vivo*. As noted above, the group felt that sufficient data were already available to document that some classes of non-DNA reactive genotoxicants, such as most aneugens (based on known modes of action/dose–response data), agents that cause nucleotide pool imbalance or glutathione depletion, and DNA synthesis inhibitors, have a non-linear dose–response curve and that safe thresholds or margins of exposure can be defined for such agents. The group recommended a systematic compilation and analysis of data for both DNA reactive and non-reactive mutagens that explores the dose–response and modes of action to more thoroughly examine the default presumption of low dose linearity. A logical mechanism for achieving this would be via an expert committee charged with producing a "white paper" and subsequent journal publication. It was noted that it would likely also be necessary to build a consensus on acceptable methods for describing the shape of the dose–response curve and for evaluating the mode of action.

Considerable attention was placed on the question of whether information on the extent of human exposure (magnitude, duration, and route) can be used to define levels of concern about genotoxic damage. The "level of concern" and "threshold of toxicological concern" concepts (LOC, TTC) used for assessing environmental risks, direct and indirect food additives and pharmaceutical impurities [2,30–33] were cited as examples of how this is already being done in certain cases. The possibility of extending these concepts by combining human exposure information with information about *in vitro* dose–response relationships and/or *in vivo* animal genotoxicity information was discussed and it was concluded that this could be a profitable area of focus within the current HESI project.

The group recommended that *in vivo* potency information (both genotoxicity and carcinogenicity information), information about *in vitro* concentration–effect relationships in relation to effects in *in vivo* models, and likely concentrations achieved in anticipated human exposure situations should be evaluated as a basis for future recommendations. It was recommended that this analysis include a determination of whether available information allows development of semi-quantitative categories (bins) of concern (*e.g.*, low, intermediate, high) for some classes of chemicals based on:

- human exposure data;
- *in vivo* potency test data (*e.g.*, tumor data, genetic toxicity data);
- *in vitro* concentration far exceeding achievable *in vivo* exposure (*e.g.*, blood/tissue concentrations, DNA adducts);
- mode of action and metabolism/pharmacokinetics.

It was suggested that a weight of evidence approach that considers structural alerts, weaknesses and strengths of each assay, and consistency and reproducibility of the findings would be needed. Along with human exposure information, the biological plausibility for the response to occur in humans should be considered. Because positive findings in *in vitro* assays, particularly in mammalian cell systems, can be problematic, correlative *in vivo* data are preferred to evaluate the potential for human risk. However, the limitations of *in vivo* assays (*e.g.*, ability to measure relevant events in potential target cell populations, sensitivity of certain endpoints when exposure is short-term, metabolic and pharmacokinetic differences among species) need to be considered in developing an appropriate weight of evidence approach. The above recommendations for activities to be undertaken are directed at providing evidence-based approaches to these issues.

The group noted that any effort undertaken should be coordinated with the existing effort of the International Workshops on Genotoxicity Testing (IWGT) directed at improved genetic toxicology testing strategies, and especially the IWGT working group on appropriate follow-up testing when *in vitro* positive genotoxic responses occur [18].

2.2.3. *Break-out group #2: conclusions and recommendations*

Based on the discussions described above, break-out group #2 proposed the following conclusions and recommendations:

- Although *in vitro* assays are useful, recent analyses of expanded datasets have illustrated the limitations of these tests. Improved approaches are needed that allow the results of these *in vitro* assays to be better used in assessing genotoxic hazard.
- An evaluation of *in vivo* and *in vitro* genetic toxicology data including dose–response, by chemical class and type of damage, is needed to determine the feasibil-

ity of developing a tiered or quantitative classification system for genotoxic hazard.
  - The evaluation should include examination of the relationship between *in vitro* and *in vivo* responses, for different mechanistic classes of genotoxicants, analyzed separately by whether the agent is directly active or requires metabolic activation for genotoxic activity.
  - The evaluation should include correlation of tissue exposure *in vivo* with genetic damage *in vivo* (including tumor response) and *in vitro*, to support development of (semi) quantitative estimates of levels of concern.
  - An analysis is recommended to determine if we can develop different bins of concern (*e.g.*, low, intermediate, high) for some classes of chemicals based on human exposure data; *in vivo* potency test data (*e.g.*, tumor data, genetic toxicity data); and *in vitro* concentration in relation to achievable *in vivo* exposure (*e.g.*, blood/tissue concentrations, DNA adducts).
- An evaluation of the literature and available data bases (pesticides, drugs, NTP, *etc.*) is needed to determine the scientific support for low dose linearity *versus* practical thresholds for different classes of genetic toxicants.
- Whenever possible, *in vivo* dose–response and human exposure information should be used in a weight of the evidence approach to evaluate the potential for human risk.
  - Because concerns were raised over limitations of currently available *in vivo* methods, a review should be undertaken of available information to define these limitations so that the combination of *in vitro* and *in vivo* information can be used more effectively.

2.3. Break-out group #3: how to improve our testing for genetic toxicity

2.3.1. Break-out group #3: background

The participants in this break-out group started with the premise that we can not throw out the 'tried and tested' approaches without having something with which to replace them. It was recognized that the 'Ames test' has a very robust database and would be difficult to throw out, and that the most problematic tests are currently the *in vitro* chromosome damage tests, as they demonstrate the higher rates of positives. Additionally, it has been suggested that the Ames results generally correspond with structure-activity models based on electrophilicity [34]. This break-out group focused their attention on the need to develop *in vitro* models that are more predictive models for *in vivo* biology, and that reduce artifacts.

2.3.2. Break-out group #3: report

All *in vitro* systems are at best imperfect models for the biological effects seen *in vivo*. This generalization holds true for *in vitro* genotoxicity tests used as hazard identification tools in the prediction of carcinogenicity, especially in view of our current understanding that epigenetic events play a key role in carcinogenicity. One of the challenges in using *in vitro* genotoxicity assays as predictors of carcinogens was highlighted in a recent analysis by Kirkland et al. [15,35] of over 700 chemicals that have rodent carcinogenicity data, which found that 75–95% of non-carcinogens were positive in one or more of the standard *in vitro* genotoxicity assays. In this analysis, the false positive rate (defined as positive in mutagenicity assay but negative in a rodent cancer bioassay) was highest in mammalian cell tests such as the chromosomal aberration assay in Chinese hamster cells or the tk gene mutation assay in L5178Y mouse lymphoma cells. As a consequence of such positive *in vitro* genotoxicity data, numerous animal studies and mechanistic research projects are conducted in order to determine whether effects seen *in vitro* are biologically relevant *in vivo*. These studies are costly, time consuming, utilize many animals, and do not always give a definitive answer.

The findings reported by Kirkland et al. [15,35] and in several other earlier analyses [36–38] have recently been confirmed in an analysis by Matthews et al. [39,40] of a larger database of FDA and EPA chemicals. These recent analyses confirm earlier analyses on smaller data sets in the late 1980s to early 1990s. It is recognized that efforts to correlate the genetic toxicology assays with the cancer bioassay data are complicated by the fact that not all the genetic toxicology and cancer data have been evaluated according to current standards of acceptability and interpretation. Despite this, there was consensus that there is great value in developing new tests and/or approaches for predicting *in vivo* genotoxins and potential carcinogenic chemicals. This topic was the focus of the break-out group.

In vitro genotoxicity assays are used for a variety of purposes, from the rapid screening of potential drugs or other chemicals of interest to the detailed mode of action analyses for carcinogenicity risk assessment [16,23]. The participants focused their discussions on the use of genetic toxicity tests for predicting whether a chemical has the potential to cause carcinogenicity via a mutagenic mechanism, *i.e.*, hazard identification. This use of genetic toxicity tests to determine whether the mode of

action (MOA) of a known carcinogen is via a mutagenic mechanism was not addressed by this workgroup.

Over the years, it has been become apparent that the *in vitro* genotoxicity tests, particularly the mammalian cell assays, detect some non-DNA reactive agents (*i.e.*, the primary target of the chemical or its metabolite(s) is not DNA, for example, topoisomerase inhibitors) in addition to DNA-reactive agents. In fact, there has been an effort over the years to expand the spectrum of genetic events detected in each assay, for instance by extending the length and types of chemical treatment. Furthermore, there has been pressure to increase the numbers and types of assays in various genotoxicity testing batteries in order to detect the full spectrum of genetic events and/or as many rodent carcinogens as possible. The majority of the working group felt that this proliferation of testing was contributing to the generation of "false positive results" with respect to predicting whether a chemical will be a carcinogen. As such, we discussed the necessity of refocusing genetic toxicology tests on the detection of DNA reactive carcinogens. Other members expressed the view that genotoxicity assays detect genetic damage and would be expected to respond to all insults that damage DNA (regardless as to whether the damage is caused "directly" or "indirectly"). Both short-term and long-term solutions to address these issues were discussed.

2.3.2.1. Possible short-term solutions. There is a need to identify potential sources of false positive results (for predicting carcinogenicity) obtained with the current *in vitro* genotoxicity assays. This issue relates to determining what assay conditions cause biologically irrelevant positive responses—artifacts of the *in vitro* conditions. This issue was also viewed as essential for developing any new tests or longer term approaches. To this end, we addressed the question "What kind of research or efforts can be used to improve current *in vitro* tests?" The following possible activities were discussed:

- Re-examine whether a top concentration of 10 mM is justified. The original guidance to use a top concentration of 10 mM in the mammalian cell assays when there is no toxicity is based on early analyses of small databases which showed that there was a need to test up to10 mM to detect some mutagens. Because such mutagens may be detected in the bacterial gene mutation (Ames) assay, there was agreement that there may be inadequate justification for routine use of 10 mM in mammalian cell assays. Another factor in setting a top concentration of 10 mM was to avoid osmolality effects in these assays (*i.e.*, effects due to osmotic conditions that cannot be achieved *in vivo*). Changes in osmolality are controlled for in assays conducted by today's standards.
- Re-examine the maximum level of cytotoxicity needed and the appropriate measures of cytotoxicity. By virtue of being *in vitro* tests, high, non-physiological concentrations of test chemicals can be added to *in vitro* genotoxicity assays. Similarly to the above, this group felt there was a need to determine whether detection of *in vivo* mutagens and/or DNA reactive carcinogens required routine testing up to the cytotoxic levels used in current protocols.
- Determine whether both long exposures as well as short exposures in the mammalian cell assays are required to detect *in vivo* mutagens and/or DNA reactive carcinogens, particularly those not detected by the bacterial reverse mutation assay.
- Determine if induced rat liver S9 is the most appropriate metabolic activation for *in vivo* mutagens and/or DNA reactive carcinogens. Investigate other metabolic activation systems.
- Determine if cytogenetic assays in human lymphocytes are better predictors of human hazard and more relevant to human risk assessment than currently used mammalian cell lines. Anecdotal data as well as recent publications [41] have been discussed at various meetings to suggest fewer "irrelevant" positive results occur in the chromosome aberration assay and/or micronucleus assay when conducted in human lymphocytes than in other mammalian cell lines. There is a need to determine whether this possibility can be confirmed since this could lead to a simple solution to the problem of false positives for predicting carcinogenicity. It was, however, unclear whether there was sufficient data available with human cells for this analysis.
- After the meeting there was a proposal to conduct a thorough analysis of the existing genotoxicity and cancer databases to create a dataset that includes only data that meets current acceptance and interpretation criteria. Once such an analysis has been completed, the information can be used to more accurately access the ability of the current genotoxicity assays to predict whether a chemical will be a carcinogen. This effort would also provide a sound foundation for addressing and perhaps modifying some of the assay parameters (top dose, required cytotoxicity level, *etc.*).

2.3.2.2. Approaches for the possible short-term solutions. To initiate these activities, the following approaches were discussed:

- Form an expert panel to identify a list of definitive *in vivo* genotoxins and/or DNA reactive mutagenic carcinogens which we expect *in vitro* genotoxicity tests to detect, and then search these chemicals to answer the above questions.
- Form an expert panel to analyze the role of metabolism in the mutagenicity and/or carcinogenicity of *in vivo* genotoxins and/or of DNA reactive mutagenic carcinogens.
- Initiate a collaborative experimental study to analyze different measures of cytotoxicity to determine if appropriate measures are being used.
- Search existing databases to determine whether fewer false positive results occur in human lymphocyte cytogenetics assays than in other mammalian cell lines.
- Collect HESI member company data to determine whether there are fewer "false positive" results in human lymphocyte cytogenetics assays. This data collection exercise should address the following points:
  - Include data to address whether there is increased variability of human lymphocytes relative to other commonly used cell types (data collection from contract and testing labs that use HPBL).
  - Focus on chemicals negative for bacterial gene mutation.
  - Include data comparing rat lymphocytes to other commonly employed cell lines to address the possibility that primary lymphocytes yield more relevant results.
- Initiate a collaborative experimental effort to compare cytogenetic results between different cell types. It was recognized that this effort will take the largest amount of resources, but would be the most definitive way to address the question since analyses of databases are complicated by the quality of the studies that were not designed for this purpose.

2.3.2.3. Possible mid-term solutions. Based on the analyses of current databases, the bacterial reverse mutation assay has been shown to have the highest specificity for prediction of rodent carcinogenicity of the currently used *in vitro* genotoxicity assays [15,35]. Based on this, the group discussed whether the *in vitro* mammalian cell assays could be replaced by tests or approaches that compliment the bacterial reverse mutation assay, *i.e.*, that detect *in vivo* genotoxins and/or DNA reactive mutagenic carcinogens that are negative in the bacterial assay [38]. Because changing the standard genotoxicity testing battery would require changes in regulations, this was viewed as a possible mid-term solution.

As a first step, we discussed the question "Can we accept a battery of the bacterial reverse mutation assay and an *in vivo* MN or another assay for routine testing acknowledging that some chemical classes may require alternative testing?" Types of chemicals that are potential hazards that are known to be negative in the bacterial assay include: metals, steroids/hormones, topoisomerase inhibitors, nucleoside analogs, mammalian receptor-specific chemicals and chemicals whose primary activity is the induction of large deletions and other chromosomal damage. It is the detection of this latter class of chemicals (chromosomal mutagens) that led to the establishment of the current battery.

2.3.2.4. Approaches for possible mid-term solutions. One suggested approach to address this question is to conduct a database analysis of bacterial reverse mutation and *in vivo* MN tests (or other assays) to see if these detect relevant *in vivo* genotoxins or/and DNA reactive mutagenic carcinogens. Classes not detected by this battery could be identified and appropriate testing recommendations determined. This would involve the following steps: search existing databases like the database used by Kirkland et al. [15,35]; the database used by Matthews et al. ([40,41], EPA GENE-Tox [42], and others; and obtain HESI member Company data. This is best accomplished using databases that have been thoroughly evaluated to include only data that meets all current criteria for acceptability and interpretation.

2.3.2.5. Possible long-term solutions. While the above approaches have the potential to reduce problems with current *in vitro* genotoxicity tests and their interpretation in the near term, which was the primary focus of this workshop, there was also some discussion about the need for the development of new generation tests that could be used in the future—tests that could be specifically designed to address the features that the current tests lack. Features of new tests that were discussed included use of mammalian cells/cell lines to insure appropriate mammalian cell targets are present, use of p53 and DNA repair proficient cells that are metabolically relevant, and development of assays that would allow multi-endpoints analyses.

There was also discussion of the use of only *in vivo* genotoxicity tests in the future and/or of the need for the development of a new generation of *in vivo* tests that measure the full spectrum of mutagenic events. For optimal utility, such new systems should provide for rapid mutant detection and not require the *in vitro* growth of cells to enumerate mutants. While deemed scientifically appropriate, it was noted that the use of *in vivo* tests alone would not, under some current regulatory guidelines, be acceptable for some testing purposes, including indus-

trial chemicals, cosmetics, *etc.*, but their use could be valuable for some applications.

2.3.2.6. Approach for the possible long-term solutions. We discussed the utility of holding a workshop to discuss new generation *in vitro* and *in vivo* tests. In a workshop recently sponsored by ECVAM, some of the newer *in vitro* assays were discussed [26]. A workshop that discussed both new *in vitro* and new *in vivo* tests would be valuable.

2.3.3. Break-out group #3: conclusions and recommendations

This workgroup approached its discussion with the goal of capturing a wide variety of opinions and generating a number of options for improving the identification of chemicals that are carcinogens prior to the completion of any cancer bioassays. While there were diverse opinions concerning the utility of the current tests and approaches, there was general agreement that new tests and approaches are needed. The workgroup also agreed that, to make significant progress on this issue in a reasonable length of time, a variety of parallel activities would be required. As such, we encourage partnering of the various interested stakeholders in these initiatives.

3. Overall workshop conclusions

Table 1 summarizes the key recommendations of the workshop, and identifies some of the commonalities shared between the three break-out groups.

There was general agreement among workshop participants that the rate of *in vitro* positive findings not confirmed *in vivo* is too high to justify using qualitative outcomes as the sole basis of regulatory decision-making and that there is a critical need for an improved evaluation process and for better predictive models. The active participation of the workshop attendees in the discussions during the break-out groups highlighted a willingness to change and improve the current paradigm and to move from a hazard identification approach to a risk based approach that considers both toxicity and human exposure information. A general consensus was reached that the following points should be considered in the near future:

- Genotoxicity data should be considered along with other pertinent information, including extent of human exposure and dose–response relationships, in line with other toxicology end points. A weight of evidence approach should be widely applied that considers genotoxic exposure (*e.g.*, reproducibility, presence of cytotoxicity, corroborative data between studies evaluating the same end point), the relative potencies of these responses (by chemical class and type of damage), as well as the route, magnitude and duration of human exposure. When available, the weight of evidence approach should also integrate information on mode of action (*e.g.*, presence of DNA adducts/strand-breaks), metabolism and tissue concentrations *in vivo*, and tumor-related response such as relevant non-neoplastic and preneoplastic lesions. Moreover, whether the genotoxicity observed with a given chemical is a key event in the multistep process of carcinogenesis and the role of other key events (*e.g.*, regenerative proliferation, mitogenic stimulation of preneoplastic foci) should be further evaluated in case of tumor findings.
- Protocols need to be improved to reduce and possibly avoid the generation of artifacts and the unnecessary and extensive use of animal studies and resources, to minimize extreme high dosing conditions that would never be achieved *in vivo*, and to incorporate dosing conditions that are more realistic to human exposure situations to enable better extrapolation of the results. A collaborative effort was suggested to compare the results obtained with different cell types (*e.g.*, primary human lymphocytes *versus* cell lines), to evaluate the limits of different cytotoxicity measurements *in vitro*, to re-consider the rationale for the selection of the top concentration levels (*e.g.*, level of cytotoxicity, precipitates, and 10 mM limit), to review data obtained after short- and long-term exposures, to reconsider the metabolic activation in the *in vitro* systems, and in the case of *in vivo* tests to develop the possibility of evaluating multiple genotoxic end points from the same treated animals. This could be accomplished by examining existing databases, private and public, and by determining if certain assays could be eliminated or substituted. It is likely that some experimental work would be needed to obtain the information needed.
- The appropriateness of non-linear low dose–response extrapolations for both DNA reactive and non-reactive carcinogens should be further evaluated. A white paper should be prepared to examine the scientific validity or lack thereof of the low dose linear extrapolation for genotoxicity/carcinogenicity. Moreover, guidance should be given to clarify the acceptable approaches to define dose–response relationships, and to establish the existence of non-linearity. The development of uncertainty factors for establishing the "thresholds" or (semi) quantitative estimates of levels of concern was suggested.

The workshop participants stressed the importance of developing a risk-based paradigm for evaluating genotoxicity data that incorporates dose–response and human exposure information. Specific needs were identified in two general categories, *i.e.*, improving testing, and improving data interpretation and risk assessment. Recommendations to improve testing included (1) re-examine and evaluate the maximum level of cytotoxicity currently required for *in vitro* tests; (2) re-examine the current 10 mM upper limit concentration for *in vitro* mammalian studies; (3) develop improved testing strategies using current *in vitro* assays to more reliably assess genotoxic hazard and predict carcinogenesis; (4) define criteria to guide selection of the appropriate follow-up *in vivo* studies; (5) develop new and more predictive *in vitro* and *in vivo* tests, that could ultimately be used in addition or in replacement of the current models. Recommendations for improving data interpretation and risk assessment included: (1) examine the suitability of integrating threshold concepts in the assessment of genotoxicity data; (2) develop a structured weight of evidence approach for assessing genotoxic/carcinogenic hazard; and (3) re-examine *in vitro* and *in vivo* correlations. Additionally, the participants identified the critical need for support and coordination of an international collaborative effort to address these issues. The HESI subcommittee will facilitate this coordination, address the recommendations of this workshop, and identify specific research projects that will facilitate the development of a framework for the integration of *in vitro* testing results into a risk-based assessment of the effects of chemical exposure on human health.

References

[1] COM (Committee on Mutagenicity of Chemicals in Food, Consumer Products, and the Environment), COM Guidance on a Strategy for Testing of Chemicals for Mutagenicity, United Kingdom, December 2000 (http://www.advisorybodies.doh.gov.uk/com/).

[2] K.L. Dearfield, M.C. Cimino, N.E. McCarroll, I. Mauer, L.R. Valcovic, Genotoxicity risk assessment: a proposed classification strategy, Mutat. Res. 521 (2002) 121–135.

[3] U.S. FDA (Food and Drug Administration), Toxicological Principles for the Safety of Food Ingredients. IV.C.1. Short-Term Tests for Genetic Toxicity, Redbook, 2000.

[4] U.S. FDA (Food and Drug Administration), Guidance for Industry Recommended Approaches to Integration of Genetic Toxicology Study Results. Draft Guidance. Genotoxicity, Center for Drug Evaluation and Research, November 2004.

[5] Health Protection Board (Canada), The Assessment of Mutagenicity. Health Protection Branch Mutagenicity Guidelines, Environ. Mol. Mutagen. 21 (1992) 15–37.

[6] ICH (International Cooperation on Harmonization), ICH Harmonized Tripartite Guideline. S2A. Guidance on Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals, July 1995 (http://www.ich.org).

[7] ICH (International Cooperation on Harmonization), ICH Harmonized Tripartite Guideline. S2B. Genotoxicity: A Standard Battery for Genotoxicity Testing of Pharmaceuticals, July 1997 (http://www.ich.org).

[8] Council Directive Concerning the Placing of Plant Protection Products on the Market (91/414/CEE), 15 July 2001.

[9] P.G.N. Kramers, A.G.A.C. Knaap, C.A. van der Heijden, R.D.F.M. Taalman, G.R. Mohn, Role of genotoxicity assays in the regulation of chemicals in The Netherlands: considerations and experiences, Mutagenesis 6 (1991) 487–493.

[10] OECD (Organisation for Economic Co-operation and Development), OECD Guidelines for Testing of Chemicals. Introduction to the OECD Guidelines on Genetic Toxicology Testing and Guidance on the Selection and Application of Assays, OECD, Paris, France, 1986.

[11] T. Sofuni, Japanese guidelines for mutagenicity testing, Environ. Mol. Mutagen. 21 (1993) 2–7.

[12] TGD, Technical Guidance Document, second ed., European Chemicals Bureau, Joint Research Centre, Italy, December 2003 (http://ecb.jrc.it/new-chemicals/documents).

[13] M.C. Cimino, Comparative overview of current international strategies and guidelines for genetic toxicology testing for regulatory purposes, Environ. Mol. Mutagen. 47 (2006) 362–390.

[14] R.D. Snyder, J.W. Green, A review of the genotoxicity of marketed pharmaceuticals, Mutat. Res. 488 (2001) 151–169.

[15] D. Kirkland, M. Aardema, L. Henderson, L. Muller, Evaluation of the ability of a battery of three genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. 1. Sensitivity, specificity and relative predictivity, Mutat. Res. 584 (2005) 1–256.

[16] U.S. Environmental Protection Agency (EPA), Guidelines for Carcinogen Risk Assessment, EPA/630/P-03/001F, March 2005 (http://www.epa.gov/cancerguidelines).

[17] K.L. Dearfield, M.M. Moore, Use of genetic toxicology information for risk assessment, Environ. Mol. Mutagen. 46 (2005) 236–245.

[18] V. Thybaud, M. Aardema, J. Clements, K. Dearfield, S. Galloway, M. Hayashi, D. Jacobson-Kram, D. Kirkland, J.T. MacGregor, D. Marzin, W. Ohyama, M. Schuler, H. Suzuki, E. Zeiger, Strategy for genotoxicity testing: Hazard identification and risk assessment in relation to *in vitro* testing, Mutat. Res. 627 (2007) 41–58.

[19] D.J. Kirkland, L. Müller, Interpretation of the biological relevance of genotoxicity test results: the importance of thresholds, Mutat. Res. 464 (2000) 137–147.

[20] L. Müller, P. Kasper, Human biological relevance and the use of threshold-arguments in regulatory genotoxicity assessment: experience with pharmaceuticals, Mutat. Res. 464 (2000) 19–34.

[21] S. Madle, W. Von der Hude, L. Broschinski, G.R. Jänig, Threshold effects in genetic toxicology: perspective of chemicals regulation in Germany, Mutat. Res. 464 (2000) 117–121.

[22] M. Kirsch-Volders, A. Vanhauwaert, U. Eichenlaub-Ritter, I. Decordier, Indirect mechanism of genotoxicity, Toxicol. Lett. 140–141 (2003) 63–74.

[23] U.S. Environmental Protection Agency (EPA), Supplemental Guidance for Assessing Cancer Susceptibility from Early-life Exposure to Carcinogens, EPA/630/R-03/003F, March 2005 (http://www.epa.gov/cancerguidelines).

[24] D. Jacobson-Kram, A. Jacobs, Use of genotoxicity data to support clinical trial or positive genetox findings on a candidate pharmaceutical or impurity ... Now what? Int. J. Toxicol. 24 (2005) 129–134.

[25] FDA (Food and Drug Administration) CDER (Center for Drug Evaluation and Research), Guidance for Industry and Review Staff: Recommended Approaches to Integration of Genetic Toxicology Study Results, January 2006.

[26] D. Kirkland, S. Pfuhler, D. Tweats, M. Aardema, R. Corvi, F. Darroudi, A. Elhajouji, H. Glatt, P. Hastwell, M. Hayashi, P. Kasper, S. Kirchner, A. Lynch, D. Marzin, D. Maurici, J.-R. Meunier, L. Müller, G. Nohynek, J. Parry, E. Parry, V. Thybaud, R. Tice, J. van Benthem, P. Vanparys, P. White, How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM Workshop, Mutat. Res. 628 (2007) 31–55.

[27] N. Asano, D.K. Torous, C.R. Tometsko, S.D. Dertinger, T. Morita, M. Hayashi, Practical threshold for micronucleated reticulocyte induction observed for low doses of mitomycin C, Ara-C and colchicines, Mutagenesis 21 (2006) 15–20.

[28] T. Sofuni, T. Nohmi, T. Ohta, M. Hayashi, Genotoxicity: Is a threshold concept applicable to evaluate the mutagenic activity of DNA-targeting substances? Environ. Mutagen Res. 27 (2005) 61–73.

[29] L.O. Dragsted, B. Daneshvar, U. Vogel, H.N. Autrup, H. Wallin, L. Risom, P. Møller, A.M. Mølck, M. Hansen, H.F. Poulsen, S. Loft, A sucrose-rich diet induces mutations in the rat colon, Cancer Res. 62 (2002) 4339–4345.

[30] R. Kroes, A.G. Renwick, M. Cheeseman, J. Kleiner, I. Mangelsdorf, A. Piersma, B. Schilter, F. van Schothorst, J.G. Vos, G. Wurtzen, Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet, Food Chem. Toxicol. 42 (2004) 65–83.

[31] M.A. Cheeseman, Thresholds as a unifying theme in regulatory toxicology, Food Addit. Contam. 22 (2005) 900–906.

[32] L. Müller, R.J. Mauthe, C.M. Riley, M.M. Andino, D. De Antonis, C. Beels, J. DeGeorge, A.G.M. De Knaep, D. Ellison, J.A. Fagerland, R. Frank, B. Fritschel, S. Galloway, E. Harpur, C.D.N. Humfrey, A.S. Jacks, N. Jagota, J. Mackinnon, G. Mohan, D.K. Ness, M.R. O'Donovan, M.D. Smith, G. Vudathala, L. Yotti, A rationale for determining, testing, and controlling specific impurities in pharmaceuticals that possess potential for genotoxicity, Regul. Toxicol. Pharmacol. 44 (2006) 198–211.

[33] U.S. FDA (Food and Drug Administration). Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food, Chapter III. Concern Levels and Recommended Toxicity Tests, Center for Food Safety and Applied Nutrition, Office of Food Additive Safety, 1993.

[34] D. McGregor, Comments on the proposed use of Salmonella and structural alerts for the predictions of rodent carcinogens, Mutat. Res. 222 (1989) 299–306.

[35] D. Kirkland, M. Aardema, L. Müller, M. Hayashi, Evaluation of the ability of a battery of three *in vitro* genotoxicity tests to discriminate rodent carcinogens and non-carcinogens: II. Further analysis of mammalian cell results, relative predictivity and tumour profiles, Mutat. Res. 608 (2006) 29–42.

[36] J. Ashby, R.W. Tennant, E. Zeiger, S. Stasiewicz, Classification according to chemical structure, mutagenicity to Salmonella and level of carcinogenicity of a further 42 chemicals tested for carcinogenicity by the U.S. National Toxicology Program, Mutat. Res. 223 (1989) 73–103.

[37] J. McCann, B. Ames, Detection of carcinogens as mutagens in the Salmonella/microsome test: assay of 300 chemicals: discussion, Proc. Natl. Acad. Sci. U.S.A. 73 (1976) 950–954.

[38] R.W. Tennant, B.H. Margolin, M.D. Shelby, E. Zeiger, J.K. Haseman, J. Spalding, W. Caspary, M. Resnick, S. Stasiewicz, B. Anderson, et al., Prediction of chemical carcinogenicity in rodents from *in vitro* genetic toxicigy assays, Science 236 (1987) 933–941.

[39] E.J. Matthews, N.L. Kruhlak, M.C. Cimino, R.D. Benz, J.F. Contrera, An analysis of genotoxicity, reproductive and developmental toxicity, and carcinogenicity data: I. Identification of carcinogens using surrogate endpoints, Regul. Toxicol. Pharmacol. 44 (2006) 83–96.

[40] E.J. Matthews, N.L. Kruhlak, M.C. Cimino, R.D. Benz, J.F. Contrera, An analysis of genotoxicity, reproductive and developmental toxicity, and carcinogenicity data: II. Identification of genotoxicants, reprotoxicants, and carcinogens using *in silico* methods, Regul. Toxicol. Pharmacol. 44 (2006) 97–110.

[41] C. Hilliard, R. Hill, M. Armstrong, C. Fleckenstein, J. Crowley, E. Freeland, D. Duffy, S.M. Galloway, Chromosome aberrations in Chinese hamster and human cells: a comparison using compounds with various genotoxicity profiles, Mutat. Res. 616 (2007) 103–118.

[42] NLM (National Library of Medicine) National Institutes of Health, National Library of Medicine. U.S. Environmental Protection Agency, Office of Prevention, Pesticides and Toxic Substances—GENE-TOX Agent Registry. Online computer database—TOXNET file. US Dept. Health and Human Services, Bethesda, MD, 2003 (http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?GENETOX).

Bull.Natl.Inst.Health Sci., 125, 35-50 (2007)　　Review

有機スズ化合物の生殖発生毒性

江馬 眞

Reproductive and Developmental Toxicity of Organotin Compounds

Makoto Ema

Organotin compounds are chemicals widely used in agriculture and industry. Widespread use of organotins has caused increasing amounts to be released into the environment. Organotins show many aspects of toxicity, such as immunotoxicity, neurotoxicity, and reproductive/developmental toxicity. However, the reproductive and developmental toxicity of organotins is not well understood. The findings of the studies on reproductive and developmental effects of organotin compounds in mammals were summarized in this review.

Keywords: Organotin, reproductive toxicity, developmental toxicity, implantation failure, teratogenicity

1. はじめに

有機スズ化合物は農業や工業の分野で広く使われている[1, 2]．四価のスズ化合物は主に他の有機スズ化合物生産の中間体として使用されている．三価の有機スズ化合物は殺生物作用を有しており，防黴剤，ダニ駆虫剤，ネズミ駆散剤，軟体動物駆除剤等として，また，船底防汚剤として広く用いられている．特に，トリフェニルスズ（TPT）とトリブチルスズ（TBT）は藻類駆除剤，軟体動物駆除剤として，防汚剤製品中によく使われてきた．二価の有機スズ化合物は商業上で最も重要な誘導体であり，主にプラスティック工業の分野でポリマーの劣化を防止するためにポリ塩化ビニル（PVC）プラスティックの熱，光安定剤として使われている．一価の有機スズ化合物はPVCの安定剤として使用されている．有機スズ化合物の生産量をTable 1に示した。

近年の有機スズ化合物の広範な使用により有機スズ化合物による環境汚染の懸念が高まっている．農薬としての使用以外の有機スズ化合物の環境汚染の経路は，PVCプラスティックの安定剤として使われた有機スズ化合物の水中への溶出であり[3]，また，船底防汚剤としての使用が水環境汚染の原因となっている[4]．海棲生物[5-7]や海産物[8-12]からTBTやTPTが検出されており，カキ[13]，泥ガニ[14]，ムールガイ[15]，チヌークサーモン[16]，イルカ，マグロ及びサメ[17]における食物連鎖によるTBTの生物濃

To whom correspondence should be addressed:
Makoto Ema; Kamiyoga 1-18-1, Setagaya, Tokyo 158-8501, Japan; Phone: +81-3700-9878; Fax: +81-3-9700-1408; E-mail: ema@nihs.go.jp

Table 1. スズ化合物の生産量

物　質　名	CAS	生産量（トン）
Dibutyltin dichloride	683-18-1	10,000 - 15,000
Dibutyltin dilaurate	77-58-7	1000 - 5000
Dibutyltin maleate	78-04-6	500 - 1000
Dibutyltin oxide	818-08-6	1000 - 5000
Dibutyltin bis (2-ethylhexylmercap-acetate)	10584-98-2	7,500 - 12,500
Dibutyltin bis (isooctyl mercap-acetate)	25168-24-5	Not available
Dimethyltin dichloride	753-73-1	1,000 - 5,000
Dimethyltin bis (2-ethylhexyl mercap-acetate)	57583-35-4	5,000 - 10,000
Dimethyltin bis (isooctyl mercap-acetate	26636-01-1	Not available
Dioctyltin dichloride	3542-36-7	5,000 - 10,000
Dioctyltin bis (2-ethylhexyl mercap-acetate)	15571-58-1	7,500 - 12,500
Dioctyltin bis (isooctyl mercap-acetate)	26401-86-5	Not available
Monobutyltin trichloride	1118-46-3	10,000-15,000
Monobutyltin tris (2-ethylhexyl mercap-acetate)	26864-37-9	2,500-7,500
Monobutyltin tris (isooctyl mercap-acetate)	25852-70-4	Not available
Monomethyltin trichloride	993-16-8	1,000 - 5,000
Methyltin Reverse Ester Tallate	201687-57-2	7,500 - 10,000
Monomethyltin tris (2-ethylhexylmercap-acetate)	57583-34-3	5,000 - 10,000
Monomethyltin tris (isooctylmercap-acetate)	54849-38-6	Not available
Mono-octyltin trichloride	3091-25-6	1,000 - 5,000
Mono-octyltin tris (2-ethylhexylmercap-acetate)	27107-89-7	2,500 - 7,500
Mono-octyltin tris (isooctylmercap-acetate)	26401-86-5	Not available
Tributyltin chloride	1461-22-9	2500 - 3000
Tetrabutyltin	1461-25-2	10,000 - 12,500
Tetraoctyltin	3590-84-9	2,500 - 7,500
Tin Tetrachloride	7646-78-8	20,000 - 25,000

出典：ORTEP Association. 2004. Global production data

縮，またコイ[16]及びカブトガニ[18]における食物連鎖によるTPTの生物濃縮が報告されている．ヒトは海産物を通じて有機スズを摂取しており，滋賀県人のTBTの1日摂取量は4.7-6.9 μg（1991年），2.2-6.7 μg（1992年），TPTの1日摂取量は4.7-6.9 μg（1991年），2.2-6.7 μg（1992年）であり[19]，また，1998年のトータルダイエット方式調査による日本人の1日摂取量は，TPT: 0.09 μg，ジフェニルスズ（DPT）：0 μg，TBT: 1.7 μg，ジブチルスズ（DBT）：0.45 μgと報告されている[19]．これらの値はFAO/WHO合同残留農薬専門家会議によるTPTの許容1日摂取量（25 μg）[20]及びtributyltin oxide（TBTO）の経口曝露指針値（18 μg）[21]よりも低く，海産物中の有機スズ濃度はヒトの健康に悪影響を及ぼすほど高くはない[11, 19]と考えられるが，Belfoidら（2000）[9]は，ヒトの健康影響の可能性について結論を下すためには海産物中のTBT含量についての更なる研究が必要であると述べている．

近年，環境汚染物質による内分泌系の障害の結果，野生動物の生殖に対する悪影響が惹起される可能性が指摘されている[22]．TBT及びTPTは内分泌撹乱作用が疑われる物質とされており[23]，低濃度で巻貝のインポセックス（imposex: 雌にペニスと輸精管が形成される現象），さらに繁殖障害を引き起こす[24]ことから，哺乳類における生殖発生毒性が懸念されている．

有機スズ化合物の一般毒性については古くから比較的よく知られている[2, 21, 25-28]が，生殖発生毒性の理解は十分ではない．本稿では，Ema M and Hirose A (2006) [29] Reproductive and developmental toxicity of organotin compounds. In Metals, Fertility, and Reproductive Toxicity, CRC Pressの実験動物における生殖発生毒性の項を基に最近の新たな知見を加えて，有機スズ化合物の生殖発生毒性について概説した．

2. フェニルスズ化合物の生殖毒性

2-1 トリフェニルスズ（TPT）の生殖毒性

TPTは昆虫の不妊化剤として知られている[30]．Table 2にTPTの生殖毒性試験の結果を示した．雄に対する影響として，100または200 ppmのtriphenyltin hydroxide（TPTH）を含む飼料を64日間与えた雄Sharmanラットを無処置雌ラットと繰り返し5回交配させたところ，体重増加及び摂餌量の著しい低下とともに，受精率，出産生児数及び交配あたりの生児数の低下が認められたが，摂餌量の回復とともに受精率が回復したことが報告されている[31]．Holtzmanラットに20 mg/kgのtriphenyltin acetate（TPTA）またはtriphenyltin chloride（TPTCl）を19日間混餌投与したとき，体重及び精巣重量への影響が顕著であった．精巣では，精細管の精上皮細胞層の減少，ステージの進行した精上皮細胞の減少及び精細管腔の狭小化等の精巣の退行性変化がみられ，TPTAを投与したときにより強い毒性が観察されている[32]．同様に，20 mg/kgのTPTAまたはTPTClのHoltzmanラットへの20日間混餌投与により精子形成が障害されたが，70日間正常飼料を与えると精子形成の完全な回復がみられた[33]．ICR/Ha SwissマウスにTPTA（2.4,12 mg/kg）またはTPTH（1.3,8.5 mg/kg）を単回腹腔内投与，もしくはTPTA（6mg/kg）またはTPTH（11 mg/kg）を5日間連続強制経口投与した後に，無処置雌と交配させ，妊娠13日に剖検した結果，優性致死作用は認められなかった[34]．

TPTの雌動物における生殖毒性についても報告がある．20 mg/kgのTPTAまたはTPTClのHoltzmanラットへの4日間混餌投与により，成熟卵胞の減少，初期卵胞の閉鎖の増加，黄体の著しい減少が観察されている[35]．このような現象は排卵の減少，延いては受胎率の低下の原因となる．

ラットの妊娠初期にTPTClを投与したときの妊娠の成立及び維持に対する影響が検討されている[36]．Wistarラットの妊娠0-3日に3.1, 4.7, 6.3 mg/kgまたは妊娠4-6日に6.3, 12.5, 25.0 mg/kg のTPTClを強制経口投与したところ，用量依存的な着床阻害が引き起こされ，妊娠0-3日の4.7 mg/kg以上，妊娠4-6日の12.5 mg/kg以上で妊娠率の低下が観察された．着床前胚死亡率の増加は妊娠0-3日の4.7 mg/kg以上でみられたが，TPTCl投与群の妊娠が成立した雌における着床数，生存胎児数，着床前及び着床後の胚死亡率は対照群と同様であった．これらの結果は妊娠初期に投与したTPTClは着床阻害作用を示し，着床前に投与した方が強い影響を及ぼすことを示している．

子宮内膜の正常な機能は胚生存のために重要であり，子宮の脱落膜化は正常な着床及び胎盤形成，その後の正常妊娠の維持に必須である．偽妊娠動物における内膜創傷による子宮内膜の変化は，胚の着床によって惹起される妊娠子宮における脱落膜反応と同様であり[37, 38]，着床に関連した母体の生理学的変化のモデルとなりうる[37]．この方法を用いて脱落膜反応を誘起することにより化学物質の生殖発生毒性を母体と胚/胎児とに分けて検討することが可能となる[37-41]．TPTClの着床阻害作用の原因を明らかにするために，子宮機能に対する作用が偽妊娠Wistarラットを用いて検討されている[42]．ラットの偽妊娠0-3日にTPTCl（3.1, 4.7, 6.3 mg/kg）を強制経口投与し，偽妊娠4日の11:00から13:00の間に麻酔下で偽妊娠ラットの子宮内膜を創傷することにより脱落膜反応を誘起し，偽妊娠9日の子宮重量を子宮脱落膜化の指標として測定した[43]．その結果，子宮重量の低下（子宮脱落膜化の抑制），偽妊娠4及び9日の血清中プロゲステロン低下が4.7 mg/kg以上の投与でみられた．この投与量は妊娠0-3日に投与したときには着床前胚致死を引き起こす

Table 2 フェニルスズ化合物による生殖毒性

物質名	動物種	投与量	投与日	投与経路	生殖発生毒性	著者
TPTH	Sherman ラット	100-200 ppm	64-238 日	経口（混餌）	↓生存児のある母体数 ↓同腹児数/交配	Geiss & Kimbrough (1968)
TPTA TPTCl	Holtzman ラット	20 mg/kg	19 日	経口（混餌）	↓精巣サイズ ↑精巣の形態学的変化	Pate & Hays (1968)
TPTA TPTCl	Holtzman ラット	20 mg/kg	20 日	経口（混餌）	↑精子形成過程の障害	Snow & Hays (1983)
TPTA	ICR/Ha Swiss マウス	2.4-12 mg/kg 6 mg/kg	1 日 5 日	腹腔内 強制経口	優性致死作用なし 優性致死作用なし	Epstein et al. (1972)
TPTH		1.3-8.5 mg/kg 11 mg/kg	1 日 5 日	腹腔内 強制経口	優性致死作用なし 優性致死作用なし	
TPTA TPTCl	Holtzman ラット	20 mg/kg	4-24 日	経口（混餌）	↓成熟卵胞数 ↑初期卵胞の閉鎖数 ↓黄体数	Newton & Hays (1968)
TPTCl	Wistar ラット	4.7-6.3 mg/kg 12.5-25 mg/kg	妊娠 0-3 日 妊娠 4-6 日	強制経口 強制経口	↓妊娠率，↓胎児体重 ↓妊娠率	Ema et al. (1997a)
TPTCl	Wistar ラット	4.7-6.3 mg/kg	偽妊娠 0-3 日	強制経口	↓子宮内膜脱落膜化 ↓血清プロゲステロン	Ema et al. (1999a)
DPTCl	Wistar ラット	16.5-24.8 mg/kg	妊娠 0-3 日	強制経口	↓妊娠率，着床前胚死亡 ↓胎児体重	Ema et al. (1999b)
		33.3 mg/kg	妊娠 4-7 日	強制経口	同上，↑着床後胚死亡	
DPTCl	Wistar ラット	4.1-24.8 mg/kg	偽妊娠 0-3 日	強制経口	↓子宮内膜脱落膜化 ↓血清プロゲステロン	Ema & Miyawaki (2002)

TPTH: Triphenyltin hydroxide, TPTA: Triphenyltin acetate, TPTCl: Triphenyltin chloride, DPTCl: Diphenyltin dichloride

量であった[36]．これらの結果は，TPTClはプロゲステロン低下を伴う子宮内膜の脱落膜化抑制を惹起し，これらがTPTClによる着床阻害に関与していることを示唆している．TPTClの子宮の脱落膜化抑制及び着床阻害作用に対する卵巣ホルモンの作用を検討したところ，プロゲステロンとエストロンの投与はTPTClを投与した卵巣摘出ラットの脱落膜化を維持し，4.7 mg/kg以上のTPTClとプロゲステロンを併用投与したラットの妊娠率及び着床数はTPTClを単独投与したラットよりも高かった[44]．これらの結果から，TPTClによる子宮内膜の脱落膜化抑制は，少なくとも部分的には，卵巣ホルモンを介しており，またプロゲステロンはTPTClによる着床阻害を防御することが示された．

2-2 ジフェニルスズ（DPT）の生殖毒性

ラットに経口摂取されたTPTは，ジフェニルスズ（DPT），モノフェニルスズ（MPT）さらに無機スズに代謝される[45-47]．DPT化合物の生殖毒性試験の結果をTable 3に示した．Diphenyltin dichloride（DPTCl）の妊娠成立及び妊娠維持に対する影響についてラットを用いて検討した[48]．DPTClをWistarラットの妊娠0-3日に4.1，8.3，16.5，24.8 mg/kg，妊娠4-7日に8.3，16.5，24.8，33.0 mg/kgを強制経口投与したところ，妊娠率の低下が妊娠0-3日の24.8 mg/kg，妊娠4-7日の33.0 mg/kgの投与でみられた．妊娠0-3日の16.5mg（48 μmol）/kg以上の投与で着床前の胚致死が増加したが，妊娠の成立した雌の着床前胚死亡率は対照群と同様であった．着床後胚死亡率は妊娠4-7日の33.0 mg/kg 投与で上昇した．これらの結果から，妊娠初期に投与したDPTClは着床阻害を引き起こし，着床前の投与は着床中及び着床直後の投与よりも作用が強く発現することが明らかになった．妊娠0-3日の投与ではDPTClの親化合物であるTPTClも4.7 mg（12 μmol）/kg以上で着床前胚致死作用を示す[49]．モル投与量の比較により，TPTClの作用がDPTClよりも強いことが明らかなので，DPTClまたはその代謝物がTPTClの着床阻害作用の原因物質である可能性は低いと考えられる．しかしながら，TPT化合物はDPTClを投与したラットの肝で生成される[47]ので，投与されたDPT化合物の一部がTPTとして有害作用を発現している可能性があり，DPTの毒性研究の際にはこのことを考慮する必要がある．TPTとDPTによる生殖毒性の差異を明らかにし，その原因物質を明らかにするためには更なる研究を要する．

DPTClの子宮機能に対する影響について偽妊娠ラットを用いて検討されている．Wistarラットの偽妊娠0-3日に4.1，8.3，16.5，24.8 mg/kgのDPTClを強制経口投与した結果，16.5 mg/kg以上の投与で子宮内膜脱落膜化の抑制，偽妊娠4日及び9日の血清プロゲステロンの低下が観察された[50]．これらの投与量はラットの妊娠0-3日に投与したときには着床前胚致死作用を示す量であった[48]．これらの知見は，DPTClはプロゲステロン低下を伴う子宮内

膜の脱落膜化抑制を惹起し，これらがDPTClによる着床阻害の要因であることを示唆している．DPTClの子宮内膜脱落膜化抑制及び着床阻害作用に対する卵巣ホルモンの作用を検討したところ，DPTClを与えた卵巣摘出ラットにおける子宮脱落膜化がプロゲステロンとエストロンの投与で維持された[50]．また，16.5 mg/kg以上のDPTClとプロゲステロンを併用投与したラットの妊娠率及び着床数はDPTClを単独投与したラットよりも高かった．これらの結果から，DPTClによる子宮内膜脱落膜化の抑制は，少なくとも部分的には，卵巣ホルモンを介しており，プロゲステロンはDPTClによる着床阻害を防御することが示唆された．

3. フェニルスズ化合物の発生毒性

Table 3にフェニルスズ化合物の発生毒性試験の結果を示した．妊娠6-15日のSDラットにTPTA（5，10，15 mg/kg）を強制経口投与した実験では，10 mg/kg以上で母体重増加抑制，15 mg/kgで着床後胚死亡の増加，5 mg/kg以上で胎児の骨化遅延の増加が観察されているが，明らかな母体毒性を発現する投与量でも催奇形性は検出されていない[51]．同様に，妊娠7-17日のWistarラットへのTPTA（1.5，3.0，6.0，9.0，12.0 mg/kg）の強制経口投与により，9.0 mg/kg以上の投与で母体重増加抑制，着床後胚死亡率の上昇及び胎児の骨化遅延がみられているが，催奇形性は認められていない[52]．TPTAの出生前投与による児の生後の行動変化が報告されている．CFYラットの妊娠6-14日に6 mg/kgのTPTAを強制経口投与した結果，母ラットに明確な毒性徴候はみられなかったが，児の自発運動の一過性の増加及び離乳前の死亡率の上昇が観察されている[53]．妊娠6-20日にTPTA（4，8 mg/kg）を強制経口投与したTokai High Avoiders（THA）ラットの児では，シドマン回避学習試験での低回避率，E型水迷路学習試験の逆転試験でのエラー数増加と到達点までの時間の延長等の学習獲得への影響が観察されている[54]．この実験では母体死亡及び体重増加抑制は8 mg/kgでみられたが，児の奇形はいずれの投与量でも観察されなかった．

SDラットにVancide KS（TPTH）を強制経口投与した実験では，妊娠1-7日の20 mg/kg投与では吸収胚の増加はみられなかったが，妊娠8-14日の15 mg/kgの投与では6母体中2母体でしか生児が得られず，妊娠14-20日の15 mg/kgの投与では6母体中4母体で生児が得られたことが報告されている[55]．しかし，この実験で使用した動物数は少なく，実験方法の詳しい報告がなされていない．妊娠6-15日のSDラットにTPTH（13 mg/kg）を強制経口投与した実験では，母体重増加抑制及び着床後胚死亡の増加が認められ，母体毒性と胎児体重及び胚/胎児死亡率との相関性がみられたが，TPTHによる胎児奇形の発現はなかった[56]．

Wistarラットの器官形成期にTPTClを強制経口投与した実験では，母ラットの体重と摂餌量の低下が妊娠7-9日の3.1 mg/kg以上，妊娠10-12日または妊娠13-15日の6.3 mg/kg以上でみられた．着床後胚死亡率の上昇が妊娠7-9日の6.3 mg/kg以上，妊娠10-12日または妊娠13-15日の9.4 mg/kg以上でみられ，器官形成期の遅い時期ほど胚致死作用が強く発現する傾向がみられた．さらに，妊娠10-12日の12.5 mg/kgまたは妊娠13-15日の9.4 mg/kg以上で低体重胎児が認められたが，いずれの投与日及び投与量でも奇形胎児の発現率の上昇はみられていない[57]．

Table 3 フェニルスズ化合物による発生毒性

物質名	動物種	投与量	投与日	投与経路	生殖発生毒性	著者
TPTA	Wistar ラット	5-15 mg/kg	妊娠 6-15 日	強制経口	↑着床後胚死亡，↓胎児骨化	Giavini et al. (1980)
TPTA	Wistar ラット	9-12 mg/kg	妊娠 7-17 日	強制経口	↑着床後胚死亡，↓胎児骨化	Noda et al. (1991a)
TPTA	CFY ラット	6 mg/kg	妊娠 6-14 日	強制経口	↑生後児死亡， ↑児の自発運動（一過性）	Lehotzky et al (1982)
TPTA	THA ラット	4-8 mg/kg	妊娠 6-20 日	強制経口	↓児の学習獲得	Miyake et al.(1991)
TPTH	SD ラット	20 mg/kg 15 mg/kg 15 mg/kg	妊娠 1-7 日 妊娠 8-14 日 妊娠 14-12 日	強制経口 強制経口 強制経口	↓妊娠率 ↑着床後胚死亡，↓胎児体重 同上	Winek et al. (1978)
TPTH	SD ラット	13 mg/kg	妊娠 6-15 日	強制経口	↑着床後胚死亡	Chernoff et al.(1990)
TPTCl	Wistar ラット	6.3-12.5 mg/kg 9.4-12.5 mg/kg	妊娠 7-9 日 妊娠 10-12 日・13-15 日	強制経口 強制経口	↑着床後胚死亡 ↑着床後胚死亡，↓胎児体重	Ema et al. (1999c)

TPTA: Triphenyltin acetate, TPTH: Triphenyltin acetate, TPTCl: Triphenyltin chloride.

4. ブチルスズ化合物の生殖毒性

4-1 トリブチルスズ（TBT）の生殖毒性

Table 4にブチルスズ化合物の生殖毒性試験の結果を示した．雄ICRマウスにTBTO（2, 10 mg/kg）を2回/週の頻度で4週間強制経口投与したところ，精子の減少及びセルトリ細胞の空胞化が認められている[58]．Wistarラットを用いた2世代繁殖試験において，F0の妊娠0日からF1の離乳まで，さらに交配前，交配中，妊娠中，授乳中を通じてF2の生後91日までのtributyltin chloride（TBTCl: 5, 25, 125 ppm: 0.4, 2.0, 10.0 mg/kgに相当）の混餌投与により，雄児に対する影響が報告されている[59]．125 ppmのF1及びF2雄で体重増加が抑制され，精巣及び精巣上体の重量低下，精子細胞及び精子数の減少が125 ppmでみられた．さらに，腹部前立腺重量低下及び精子細胞減少がF1の125 ppm，F2の25及び125 ppmで観察され，F2世代に対する影響はF1世代よりも大きかった．血清エストラジオールの低下が125 ppmでみられたことから，著者らはこれらの変化はアロマターゼ抑制による影響であり，TBTClは雄ラットにおいて弱いアロマターゼ抑制因子として作用していると述べている．

上記の2世代繁殖試験における雌児ラットへの影響が報告されている[60]．125 ppm のF0及びF1雌動物において，膣開口遅延，性周期の乱れ，妊娠中の体重増加，児数，児体重及び生児分娩率の低下が観察されている．肛門生殖突起間距離（AGD）の体重による補正値[61]は，5 ppm以上のF1の生後1日，125 ppmのF1の生後4日及びF2の生後1日及び4日で高かった．これらの結果は，生涯にわたるTBTCl曝露が雌ラットの性発生と生殖機能に影響する可能性を示しており，著者らは雌のAGD延長はTBTClの男性化作用を示唆していると述べている．

妊娠の成立及び維持に対するTBTClの影響についてHarazonoら（1996;1998ab）[62, 63, 64]，HarazonoとEma（2000）[65]によりWistarラットを用いて詳しく調べられている．妊娠0-7日にTBTCl（8.1, 12.2, 16.3 mg/kg）を強制経口投与したところ，12.2 mg/kg以上で母体重の増加抑制，8.1 mg/kg以上で摂餌量低下がみられ，着床阻害は母体毒性が認められた12.2 mg/kg以上で観察されたが，妊娠の成立した雌においては黄体数，着床数及び胚死亡数にTBTClの影響は認められなかった[62]．妊娠阻害がTBTClそのものによるのか，母体の摂餌量低下によりもたらされた栄養不良によるものかを確認するために，ペア・フィーディング（PF）試験を行ったとこ

Table 4 ブチルスズ化合物による生殖毒性

物質名	動物種	投与量	投与日	投与経路	生殖発生毒性	著者
TBTO	ICRマウス	2-10 mg/kg	4週間 (2回/週)	強制経口	↓精子頭部数 ↑セルトリ細胞空胞化	Kumasaka et al. (2002)
TBTCl	Wistar ラット	25-125 ppm	2世代	経口（混餌）	↓精巣・精巣上体重量 ↓精子細胞数，↓血清エストラジオール ↓雄児の体重増加	Omura et al. (2001)
TBTCl	Wistar ラット	5-125 ppm	2世代	経口（混餌）	↓生児分娩率，↓児数・児の体重 ↓膣開口，↑雌AGD ↓雌児の体重増加	Ogata et al. (2001)
TBTCl	Wistar ラット	12.2-16.3 mg/kg	妊娠 0-7日	強制経口	↓妊娠率 ↓胎児体重	Harazono et al. (1996)
TBTCl	Wistar ラット	16.3-32.5 mg/kg 16.3-65.1 mg/kg	妊娠 0-3日 妊娠 4-7日	強制経口 強制経口	↓妊娠率，↓胎児体重 同上，↑着床後胚死亡	Harazono et al. (1998b)
TBTCl	Wistar ラット	16.3-32.5 mg/kg	偽妊娠 0-3日	強制経口	↓子宮内膜脱落膜化 ↓血清プロゲステロン ↑血清エストラジオール	Harazono & Ema (2000)
		16.3-65.1 mg/kg	偽妊娠 4-7日	強制経口	↓子宮内膜脱落膜化 ↓血清プロゲステロン	
DBTCl	Wistar ラット	7.6-15.2 mg/kg	妊娠 0-3日・4-7日	強制経口	↓妊娠率 ↑着床前後胚死亡，↓胎児体重	Ema & Harazono (2000)
DBTCl	IRC マウス	7.6-30.4 mg/kg	妊娠 0-3日・4-7日	強制経口	↓妊娠率 ↑着床前後胚死亡，↓胎児体重 ↓血清プロゲステロン	Ema et al. (2007a)
DBTCl	Wistar ラット	7.6-15.2 mg/kg	偽妊娠 0-3日・4-7日	強制経口	↓子宮内膜脱落膜化 ↓血清プロゲステロン	Harazono & Ema (2003)
MBTCl	Wistar ラット	903 mg/kg	妊娠 0-3日・4-7日	強制経口	↓胎児体重	Ema & Harazono (2001)

TBTO: Tributyltin oxide, TBTCl: Tributyltin chloride, DBTCl: Dibutyltin dichloride, MBTCl: Butyltin trichloride.

ろ，TBTCl投与群の妊娠阻害はTBTClそのものによるものであり，母体の栄養不良によるものでないことが示された[63]．次に，TBTClの投与時期による影響を調べるために妊娠0-3日に4.1，8.1，16.3，32.5 mg/kg または妊娠4-7日に8.1，16.3，32.5，65.1 mg/kgを強制経口投与した結果，妊娠0-3日の16.3 mg/kg以上及び妊娠4-7日の65.1 mg/kgで妊娠率の低下及び着床前胚死亡の増加が認められた[64]．また，妊娠4-7日の16.3 mg/kg以上の投与では着床後胚死亡率の上昇が観察された．これらの結果は，TBTClによる着床に対する悪影響は投与した妊娠時期により異なり，着床前に投与したときには着床阻害を，着床中及び着床直後に投与したときには着床した胚の生存に悪影響を及ぼすことを示している．TBTClによる着床阻害の要因を調べるために，子宮機能に対する影響が偽妊娠ラットを用いて検討されている．偽妊娠0-3日の16.3 mg/kgの強制経口投与により，子宮重量低下（子宮内膜の脱落膜化の抑制）及び偽妊娠4日及び9日の血清中プロゲステロンの低下が認められた[65]．偽妊娠4-7日の16.3 mg/kg以上の投与により偽妊娠9日の血清中プロゲステロンの低下がみられた．偽妊娠ラットの子宮重量低下及びプロゲステロン低下を引き起こす投与量は，妊娠ラットにおいて着床前及び着床後の胚死亡を惹起する投与量と同じであった．これらの実験結果は，TBTClは子宮内膜の脱落膜化抑制とプロゲステロン低下を引き起こし，これらがTBTClによる着床阻害の要因となっていることを示唆している．

4-2 ジブチルスズ（DBT）及びモノブチルスズ（MBT）の生殖毒性

ラットに投与されたTBTはDBT及びモノブチルスズ（MBT）に代謝され，また投与されたDBTはMBTに代謝される[45, 66-68]．TBTの生殖毒性発現におけるdibutyltin dichloride（DBTCl）の役割を検討するために，DBTClの妊娠成立及び維持に対する影響についてWistarラットを用いて調べられている[69]．妊娠0-3日または妊娠4-7日に3.8，7.6，15.2 mg/kgを強制経口投与した．3.8 mg/kg以上で摂餌量の低下が観察されたため，PF群を設けた．妊娠0-3日の投与では，妊娠率は7.6 mg/kgで対照群より低く，15.2 mg/kgで対照群及びPF群よりも低かった．着床後胚死亡率は妊娠4-7日の7.6 mg/kg以上で対照群及びPF群よりも高くなった．これらの知見から，DBTClによる初期胚の死亡は摂餌量の低下による影響ではなく，DBTClによる直接的な作用であると考えられる．初期胚の死亡率上昇をもたらす最も低いDBTClの投与量は7.6 mg（25 μmol）/kgであった．DBTClの親化合物のTBTClは16.3 mg（50 μmol）/kg以上の投与で着床阻害を惹起させた[64]．DBTClはTBTClよりも低い投与量で初期胚の死亡を引き起こすことから，DBTClまたはその代謝物がTBTClによる胚死亡の原因物質である可能性がある．着床阻害を引き起こす投与量のDBTClを強制経口投与した偽妊娠ラットでは，プロゲステロン低下を伴った子宮内膜の脱落膜化抑制がみられ[70]，プロゲステロンの投与により，少なくとも部分的には，DBTClによる着床阻害が防御された[71]．これらのことはプロゲステロンの低下がDBTClによる着床阻害の第一の要因であることを示唆している．Wistarラットの妊娠0-3日または妊娠4-7日に903 mg（3200 μ mol）/kgのbutyltin trichloride（MBTCl）を強制経口投与しても着床前及び着床後の胚死亡率の上昇は認められなかった[72] ことから，MBTClまたはその代謝物がブチルスズによる着床阻害の原因物質であるとは考え難い．脱落膜反応の低下及びプロゲステロン低下をもたらすDBTClはモル比較でTBTClよりも低いことは，TBTClによるこれらの現象にDBTClが関与していることを示唆している．偽妊娠0-3日にTBTClを投与したときには血清エストラジオールが低下した[65] が，DBTClの投与ではこのような低下は観察されなかったことから，TBTClとDBTClの卵巣機能に及ぼす悪影響の機序は異なっている可能性もある．卵巣を含めて母体の内分泌系に対するTBTClとDBTClの影響については更なる検討を要する．また，ICRマウスにDBTClを強制経口投与して着床阻害作用が検討され，妊娠0-3日の30.4 mg/kgの投与により妊娠率の低下及び着床前胚死亡率の上昇，妊娠0-3日の15.2 mg/kg以上及び妊娠4-7日の7.6 mg/kg以上の投与により着床後胚死亡率の上昇が認められた[73]．妊娠0-3日または妊娠4-7日に30.4 mg/kgを投与したときには，妊娠ラット血清中プロゲステロンの低下がみられたことから，マウスにおけるDBTClによる着床阻害作用においてもプロゲステロン低下が要因となっており，ラットと同様の機序により着床阻害が惹起される可能性が示唆された．

5. ブチルスズ化合物の発生毒性

5-1 ブチルスズのin vivo発生毒性

ブチルスズの発生毒性試験の結果をTable 5に示した．TBTOの発生毒性についてはマウス及びラットを用いて検討されている．NMRIマウスの妊娠6-15日にTBTOを強制経口投与したとき，母体体重低下を引き起こす最も低い投与量は11.7 mg/kgであり，35 mg/kgでは吸収胚が59%の頻度でみられ，低胎児体重も観察されている[74]．口蓋裂が11.7 mg/kgで7%，35 mg/kgで48%の頻度で観察されたが，Davisら（1987）[74] は，口蓋裂はTBTOに非特異的な発現であり，TBTOによる発現ではないと結論した．Swiss マウスの妊娠6-15日にTBTOを強制経口投与した実験では，40 mg/kgで母体体重及び胎児体重低下，

Table 5 ブチルスズ化合物による発生毒性

物質名	動物種	投与量	投与日	投与経路	生殖発生毒性	著者
TeBT	Wistar ラット	1832 mg/kg	妊娠 13-15 日	強制経口	↑口蓋裂	Ema et al. (1996a)
TBTO	NMRI マウス	11.7-35 mg/kg	妊娠 6-15 日	強制経口	↑着床後胚死亡, ↓胎児体重 ↑口蓋裂	Davis et al. (1987)
TBTO	Swiss マウス	40 mg/kg	妊娠 6-15 日	強制経口	↑着床後胚死亡, ↓胎児体重	Baroncelli et al.(1990)
TBTO	Swiss マウス	10-30 mg/kg	妊娠 6-15 日	強制経口	↓同腹児数, ↓児体重 妊娠期間の変化, ↓雷鳥行動を示す母体	Baroncelli et al.(1995)
TBTO	Swiss マウス	5-20 mg/kg	妊娠 6-15 日	強制経口	↑末梢異的血液学的変化	Kamrar et al (1995)
TBTO	HsNMRI マウス	27 mg/kg	妊娠 6-17 日	強制経口	↓胎児体重, ↑口蓋裂 ↑骨格異常	Faqi et al. (1997)
TBTO	Long Evans ラット	2.5-16 mg/kg	妊娠 6-20 日	強制経口	↓児数・児体重 ↑口蓋裂, ↓ 出産後体重増加 ↓瞬開口, ↓脳重量, ↓児運動（一過性）	Crofton et al. (1989)
TBTO	THA ラット	5-10 mg/kg	妊娠 6-20 日	強制経口	↑生後児死亡, ↓学習獲得	Miyako et al. (1990)
TBTA	Wistar ラット	16 mg/kg	妊娠 7-17 日	強制経口	↑着床後胚死亡, ↑口蓋裂 ↓胎児体重	Noda et al. (1991b)
TBTCl	Wistar ラット	5-25 mg/kg	妊娠 7-15 日	強制経口	↑着床後胚死亡, ↓胎児骨化	Itami et al. (1990)
TBTCl	Wistar ラット	25-50 mg/kg 50-100 mg/kg 25-100 mg/kg	妊娠 7-9 日 妊娠 10-12 日 妊娠 13-15 日	強制経口 強制経口 強制経口	↑着床後胚死亡, ↓胎児体重 ↑着床後胚死亡, ↓胎児体重, ↑口蓋裂 ↓胎児体重, ↑口蓋裂	Ema et al. (1995a)
TBTCl	Wistar ラット	100-200 mg/kg	妊娠 7-15 日の 1 日	強制経口	↑着床後胚死亡, ↓胎児体重 ↑口蓋裂（妊娠 8, 11, 12, 13, 14 日の投与）	Ema et al. (1997b)
TBTCl	SD ラット	0.25-20 mg/kg	妊娠 0-19 日	強制経口	↑着床後胚死亡, ↓胎児体重 ↑雄 AGD, ↓胎児骨化 ↓血清チロキシン・トリヨードチロニン	Adenko et al. (2003)
		2.5-10 mg/kg	妊娠 8-19 日		↓血清チロキシン	
TBTCl	SD ラット	0.025-2.5 mg/kg	妊娠 8 日から離乳	強制経口	↓肝臓・肺臓・胸腺重量 ↓血清クレアチニン・トリグリセリド ↓アミラーゼ・チロキシン 成長プロファイルの変化	Cooke et al. (2004)
TBTCl	SD ラット	0.25-2.5 mg/kg	妊娠 8 日から離乳	強制経口	↑胸腺萎縮, ↑NK 細胞数 ↑IgM ・ IgG, ↑未成熟 T リンパ球数, ↓IgG2a	Tryphonas et al. (2004)
TBTCl	SD ラット	1-5 mg/kg	妊娠 6-20 日	強制経口	↑自発運動 ↓放射状迷路課題遂行能力獲得 ↑d-アンフェタミンによる活動亢進	Gärdlung et al. (1991)
TBTCl	Wistar ラット	40-80 mg/kg	妊娠 7-8 日	強制経口	↑着床後胚死亡, ↓胎児体重	Ema et al. (1995b)
TBTCl	Wistar ラット	54-108 mg/kg	妊娠 13-15 日	強制経口	↓胎児体重, ↑口蓋裂	Ema et al. (1996a)
DBTA	Wistar ラット	15 mg/kg	妊娠 0-19 日	強制経口	↑着床後胚死亡, ↓胎児体重, ↑下顎異常 ↑舌癒合・舌裂, ↑骨格変異	Noda et al. (1988)
DBTA	Wistar ラット	5-15 mg/kg	妊娠 7-17 日	強制経口	↑着床後胚死亡, ↓胎児体重 ↑下顎裂・下唇裂・舌癒合・舌裂 ↑尾異常・肋骨及び椎骨の奇形・骨格変異	Noda et al. (1992a)
DBTA	Wistar ラット	15 mg/kg 22 mg/kg	妊娠 7-9 日 妊娠 8 日	強制経口 強制経口	↑着床後胚死亡, ↓胎児体重 ↑下顎裂・下唇裂・舌癒合・舌裂 ↑尾異常・肋骨及び椎骨の奇形・骨格変異	Noda et al. (1992b)
DBTA	Wistar ラット	28.1 mg/kg	妊娠 8 日	強制経口	↑同上の奇形	Noda et al. (1993)
DBTA	Wistar ラット	10-22 mg/kg	妊娠 8 日	強制経口	↑同上の奇形	Noda et al. (2001)
DBTCl	Wistar ラット	5-10 mg/kg	妊娠 7-15 日	強制経口	↑着床後胚死亡, ↓胎児体重 ↑下顎裂・口蓋裂・舌癒合・臍帯ヘルニア ↑尾異常・肋骨及び椎骨の奇形	Ema et al. (1991)
DBTCl	Wistar ラット	20 mg/kg	妊娠 7-9, 10-12, 13-15 日	強制経口	↓胎児体重, ↑着床後胚死亡 ↑同上の奇形（妊娠 7-9 日の投与）	Ema et al. (1992)
		20-40 mg/kg	妊娠 6, 7, 8, 9 日	強制経口	↓胎児体重, ↑着床後胚死亡（妊娠 6, 7, 8 日の投与） ↑同上の奇形（妊娠 7, 8 日の投与）	
DBTCl	Wistar ラット	24.3 mg/kg	妊娠 8 日	強制経口	↓胎児体重, ↑下顎裂・下唇裂・舌癒合 ↑舌裂・臍ヘルニア・肋骨及び椎骨の奇形	Noda et al. (1993)

DBTCl	Wistar ラット	10-15 mg/kg	妊娠 7-8 日	強制経口	↓胎児体重, ↑口蓋裂	Ema et al. (1995b)
DBTCl	Wistar ラット	50-100 mg/kg	妊娠 13-15 日	強制経口	↓胎児体重	Ema et al. (1996a)
DBTCl	Wistar ラット	1-10 mg/kg	妊娠 6-15 日	強制経口	影響なし	Farr et al. (2001)
DBTCl	SD ラット	15 mg/kg	妊娠 6-15 日	強制経口	↓胎児体重, ↑着床後胚死亡 ↑水頭症・下顎異常・外脳・無眼・口蓋裂・舌癒合・無舌	Thullen & Holson (2006)
DBTCl	NZW ウサギ	5 mg/kg 0.4-1.0 mg/kg	妊娠 6-19 日 妊娠 6-28 日	強制経口 強制経口	↓胎児体重, ↑着床後胚死亡 ↑流産	Thullen & Holson (2006)
DBTCl	カニクイザル	2.5-3.8 mg/kg	妊娠 20-50 日	胃内(経鼻)	↑着床後胚死亡	Ema et al. (2007b)
DBTM	Wistar ラット	27.8 mg/kg	妊娠 8 日	強制経口	↑下顎裂・下唇裂・舌癒合, ↑舌裂・脳ヘルニア・肋骨及び椎骨の奇形	Noda et al. (1993)
DBTO	Wistar ラット	19.9 mg/kg	妊娠 8 日	強制経口	↑同上の奇形	Noda et al. (1993)
DBTL	Wistar ラット	50.0 mg/kg	妊娠 8 日	強制経口	↑同上の奇形	Noda et al. (1993)
3-OHDBTL	Wistar ラット	100 mg/kg	妊娠 8 日	強制経口	↓ 胎児体重, ↑実下顎	Noda et al. (1993)
MBTCl	Wistar ラット	50-400 mg/kg	妊娠 7-17 日	強制経口	影響なし	Noda et al. (1992a)
MBTCl	Wistar ラット	1000-1500 mg/kg	妊娠 7-8 日	強制経口	↓胎児体重	Ema et al. (1995b)

TeBT: Tetrabutyltin, TBTO: Tributyltin oxide, TBTA: Tributyltin acetate, TBTCl: Tributyltin chloride, DBTA: Dibutyltin diacetate, DBTCl: Dibutyltin dichloride, DBTM: Dibutyltin maleate, DBTO: Dibutyltin oxide, DBTL: Dibutyltin dilaurate, 3-OHDBTL: Butyl (3-hydroxybutyl)tin diaurate, MBTCl: Butyltin trichloride.

胚死亡率の上昇がみられたが，催奇形性は認められていない[75]．

児の生後観察に関する実験では，妊娠6-15日のSwissマウスへのTBTOの強制経口投与により，20 mg/kg以上で児数の低下及び児の低体重，10 mg/kg以上で母マウスの営巣行動不良，5 mg/kg以上で低体重母体，分娩時期の乱れが認められたが，児の奇形は観察されていない[76]．同様に，Swiss マウスの妊娠6-15日にTBTO (5, 10, 20 mg/kg) を強制経口投与したところ，児動物に非特異的血液学的変化，胸腺及び脾臓重量の低下が認められた[77]．Han:NMRIマウスの妊娠6-17日にTBTOを強制経口投与した実験では，27 mg/kgで11.4%の頻度で口蓋裂が観察され，2例の胎児では橈骨弯曲，8例の胎児で短顎，5例の胎児で後頭骨癒合がみられたが，13.5 mg/kg以下の投与では母体及び胎児に対する悪影響は認められなかった[78]．ラットを用いた実験では，妊娠6-20日にTBTO (2.5, 5, 10, 12, 16 mg/kg) を強制経口投与したLong Evansラットを自然分娩させ，出生後の児を調べたところ，10 mg/kg以上で母体重増加抑制，児数，児体重及び生後1日及び3日の児生存率の低下，12 mg/kgで3%の頻度で口蓋裂，10 mg/kgで膣開口遅延，全ての投与量で生後14日の児の運動低下が観察されている[79]．また，妊娠6-20日にTBTOを強制経口投与したTHAラットの児は，10 mg/kgでは生後3日までにすべて死亡し，5 mg/kgではシドマン回避学習試験，E型水迷路学習試験の逆転試験における学習獲得が障害されていた[80]．

Nodaら (1991b) [52] は，妊娠7-17日のWistarラットにtributyltin acetate (TBTA: 1, 2, 4, 8, 16 mg/kg) を強制経口投与したところ，16 mg/kgで子宮内死亡及び口蓋裂の頻度増加，低体重胎児がみられ，この投与量では母体重と摂餌量の著しい低下，4 mg/kg以上で妊娠ラットの胸腺重量の低下がみられたと報告している．彼らは，この実験で観察された胎児の奇形はDaivisら (1987) [74] により報告されたものと同様であることから，TBTAによる特異的な作用ではないと結論した．

TBTClについては比較的よく研究されている．Wistarラットの妊娠7-15日にTBTClを強制経口投与したところ，9 mg/kg以上で母体毒性，5 mg/kg以上で胎児の骨化遅延がみられたが，胎児奇形は観察されなかった[81]．この実験結果をより詳しく調べるために，器官形成期を三分割して，妊娠7-9日に25, 50 mg/kg，妊娠10-12日に50, 100 mg/kgまたは妊娠13-15日に25, 50, 100 mg/kgをWistarラットに強制経口投与して発生毒性を検討した[82]．投与日にかかわらず母体重増加抑制が認められ，着床後胚死亡率の上昇は，妊娠7-9日の25 mg/kg以上及び妊娠10-12日の100 mg/kgでみられたが，妊娠13-15日の投与では100 mg/kgでも認められなかった．低体重胎児は妊娠10-12日の50 mg/kg以上及び妊娠13-15日の100 mg/kgでみられた．奇形胎児の発現頻度は妊娠10-12日の100 mg/kg及び妊娠13-15日の25 mg/kg以上で上昇し，口蓋裂が最も高頻度で観察された．これらの結果は，TBTClによる発生毒性は投与時の胚の発生段階によって異なり，TBTClの催奇形性には時期特異性があることを示している．催奇形性の感受期を更に詳しく調べるために，Wistarラットの器官形成期のいずれか1日にTBTClを強制経口投与したところ，TBTClの催奇形性の発現頻度は2峰性を示し，妊娠8日の100 mg/kg以上，妊娠11日，12日，13日または14日の200 mg/kgの投与で外表奇形の