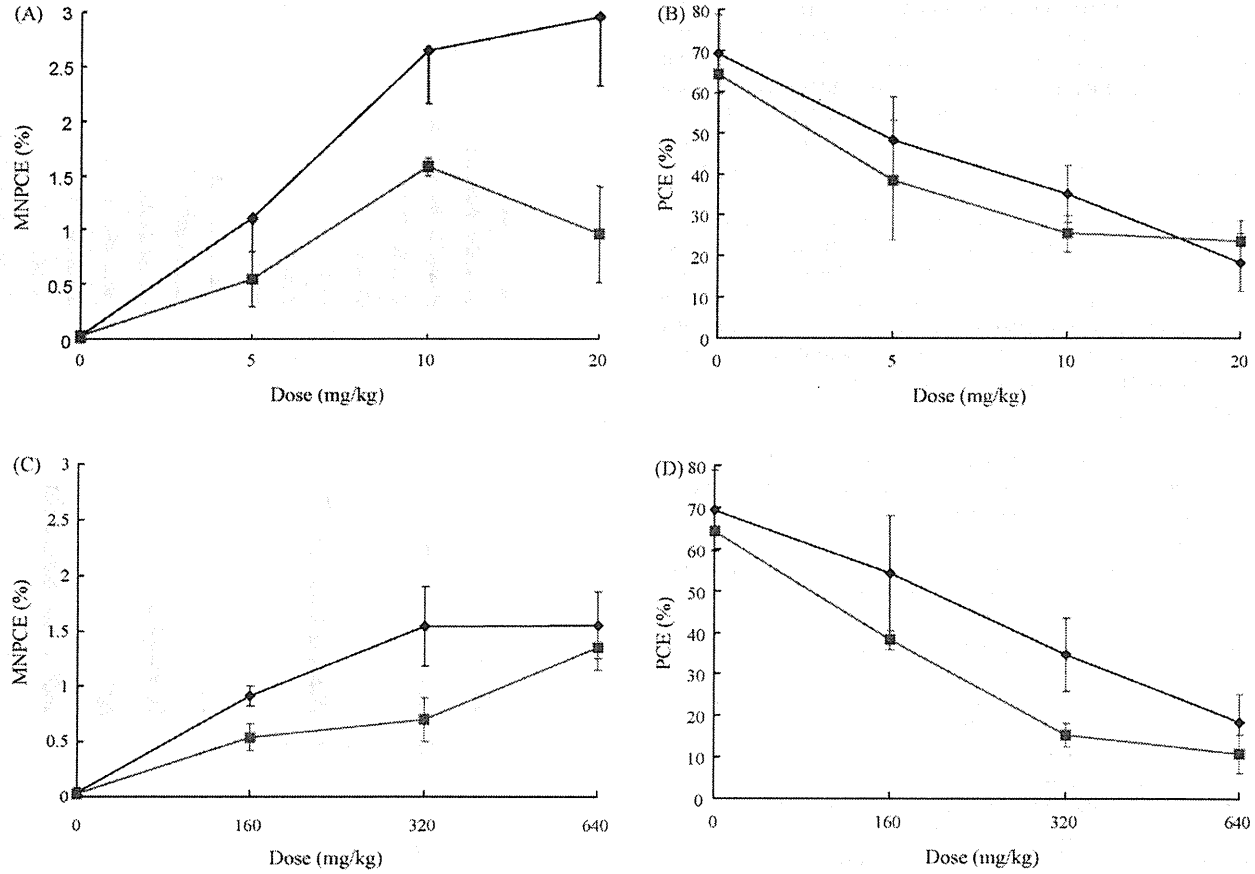

Fig. 3. Sex differences in the formation of MN in rats treated with cyclophosphamide (A and B) and cytarabine (C and D). Male (◆) and female (■) rats were dosed with test compounds by i.p. injection and the percentage MN in polychromatic erythrocytes (PCE) (A and C) and % PCEs among total erythrocytes (B and D) were measured 24 h after treatment.

Makoto Hayashi from the Biosafety Research Center in Shizuoka, Japan, presented data for CPA and Cytarabine (Ara-C) (Fig. 3) in male and female rats. Genotoxicity was clearly detected in rats of both sexes, the MN response being greater in males than in females. Cyrille Krul from TNO, The Netherlands, explained that the default strategy at TNO is to perform studies with one sex, namely males, if no substantial sex difference in toxicity occurs; otherwise the more sensitive sex is used. In their laboratory, exceptions have to be justified. Azeddine Elhajouji (Novartis) reported that 90% of all *in vivo* MN studies carried out in the last few years were single-sex studies, indicating that even under the premise that substantial sex differences in toxicity would trigger both sexes to be used, this can lead to huge savings in animal use. However, results from a recently conducted survey show that such a strategy is not current practice in industry. The ECVAM Task Force, as part of its activity to capture current status with regard to reduction, had asked laboratories performing *in vivo* MN and CA assays about the current practice in the industry [12], and the survey shows that the majority of the laboratories perform most or all of their studies using both sexes.

2.3. One treatment versus two or more treatments

Administering a test chemical to animals on two occasions and sacrificing those animals at one time point (for those endpoints where sacrifice is needed) obviously reduces the numbers

Table 1
Numbers of animals used in single-administration and two administration protocols (including positive controls), assuming single sex only.

Dose group	Two administration protocol[a]	One administration protocol[b]	
	No. of animals used	No. of animals used option 1	No. of animals used option 2
Vehicle	5	10	10
Low	5	5	10
Mid	5	5	10
High	5	10	10
Positive control	5	5	5
Total	25	35	45

Option 1 = all dose groups sampled at 24 h but only top dose and control at 48 h; option 2 = all dose groups sampled at 24 h and all dose groups except positive control sampled at 48 h.

[a] Single sampling time.

[b] Two sampling times.

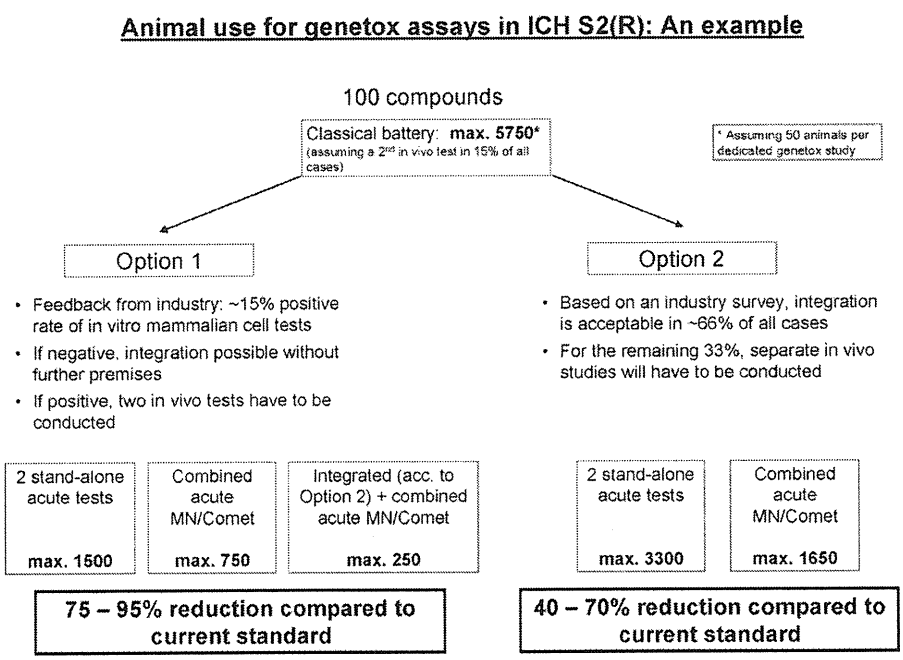

Fig. 4. Potential animal savings offered by the revised ICH S2 (R1) guideline, assuming the testing of 100 new compounds for their genotoxic properties.

of animals used compared with a single administration followed by two sacrifice times (Table 1). However, concerns had been expressed that more animals are needed to select the doses for a two-administration compared with a single-administration study. D. Kirkland therefore addressed this concern, and showed that for mean numbers of animals used in range finding the opposite is the case. Covance reviewed the numbers of animals used during range finding experiments for their last 12 studies using single- and double-administration protocols and found that for single-administration studies, the number of animals used ranged from 6 to 36 with a mean of 17, while for double-administration studies 3 to 36 animals were used, with a mean of 11. There was general consensus that for the MN and Comet assays no problems are expected when the test substance is dosed twice at the intervals suggested in OECD 474 and in the respective recommendations for the Comet assay [14,15]. Following this concept will lead to a substantial reduction in animal usage (see Table 1). However, the result of the ECVAM survey shows that to date, less than half of the labs actually utilize this option.

2.4. Omission of positive control groups

Philippe Vanparys from Altoxicon reported on the survey on the *in vivo* MN and CA test and on the use of positive and negative control animals [12]. The question whether concurrent positive controls are needed for every assay has long been debated and IWGT, following a meeting in Washington in 1999, has recommended that positive controls might not be needed [11]. The ECVAM TF survey shows where this stands today: 65% of those surveyed believed a concurrent positive control is not needed if the test facility has sufficient experience and 79% would accept a reduction of the number of animals in a concurrent positive control dose group. If regulators would accept the reduction in the number of animals, 100% of those surveyed would follow this, 59% would follow a total omission of the positive control animals if accepted by regulatory agencies. Amongst the workshop participants, Novartis has implemented the concept of a reduction in animal numbers for the positive control and has been using a positive control group of only three animals for the bone marrow MN test since 2001.

2.5. Integration of genotoxicity endpoints into repeat-dose toxicity testing

Stephen Dertinger from Litron Laboratories, USA, presented the latest status of flow-cytometric scoring of peripheral blood micronuclei, a prerequisite for efficient integration of the MN test into RDT testing. The efforts of the 9th and 13th Collaborative Study Group for the Micronucleus Test (CSGMT) were cited as important starting points that clearly established the feasibility of integration [16,17] and provided scientific proof that blood is a reliable tissue matrix for measuring micronuclei, even for rats that exhibit an efficient splenic filtration function [18,19]. Using flow cytometry, MacGregor and colleagues extended this work by demonstrating the feasibility of studying other species of toxicological interest – dogs and non-human primates [20,21]. In addition to this cross-species potential, other characteristics of the automated method that make it well-suited for integrated studies were outlined, including the scoring system's low blood-volume requirement, high throughput capacity, ability to discern mode of action based on MN size (fluorescence intensity) [22], and the objective nature of the scoring method, especially when biological standards are used for instrument calibration purposes, which controls intra- and inter-lab variability [23,24].

The joint industry initiative of pharmaceutical companies from the US, EU and Japan (Integration Initiative), which aims at investigating integration and animal saving options in the context of the revised ICH S2 guideline, was presented by Andreas Rothfuss from Bayer Schering Pharma, Germany. He emphasised the huge animal saving potential offered by the guideline options and provided an exemplary calculation for the ICH S2 (R) options 1 and 2 (Fig. 4). In the pharmaceutical industry, virtually no experience is available on the integration of the Comet assay into RDT studies with regard to sensitivity and specificity. A work plan defining experimental activities to address these concerns as well as questions related to the feasibility of integration has been suggested and is currently being discussed with the members of the Integration Initiative.

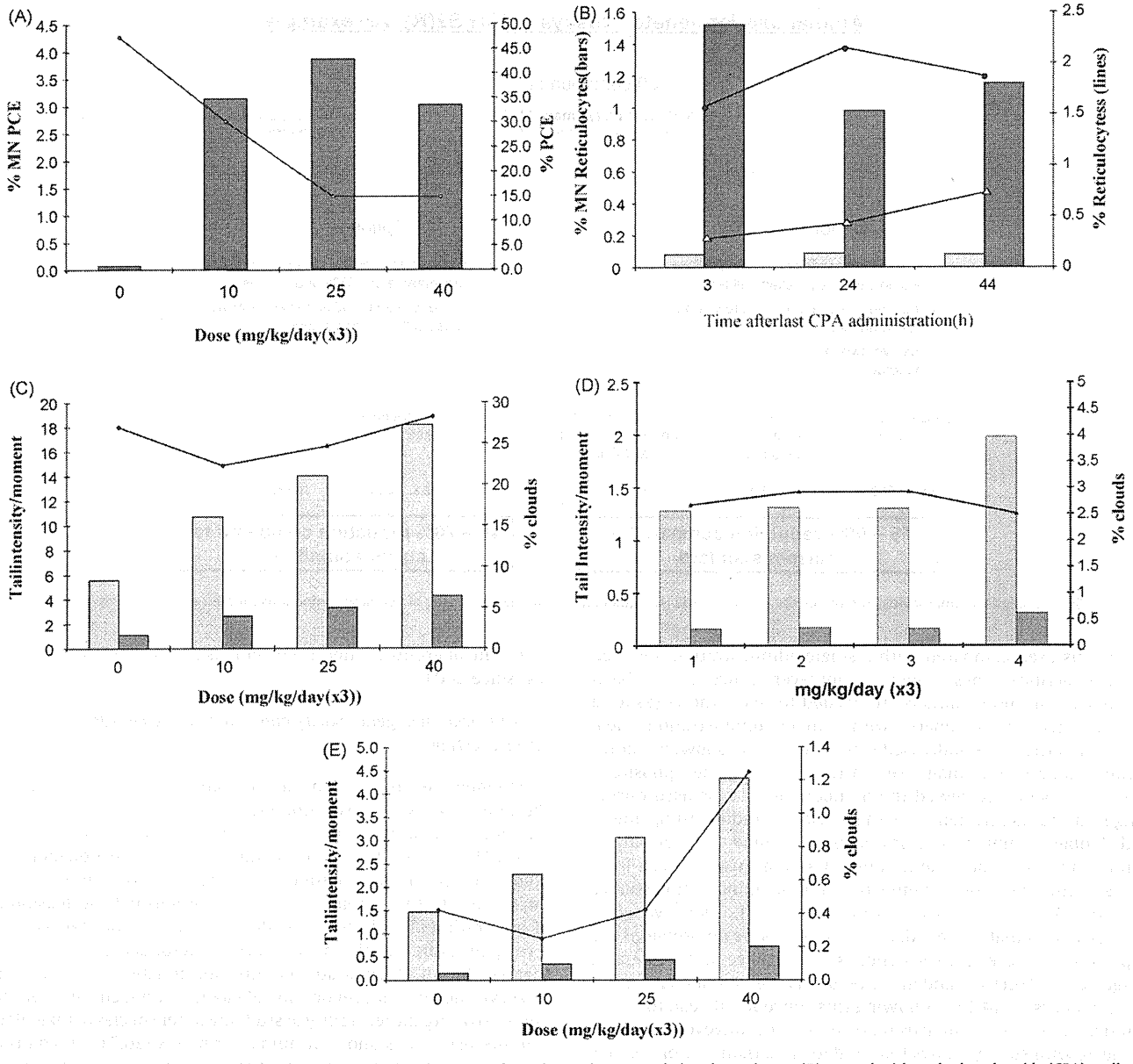

Fig. 5. MN/Comet data from a combination study performed with groups of 6 male rats (A, C, D and E) and 5 male rats (B) treated with cyclophosphamide (CPA) orally by gavage three times (48, 24 and 3 h before sacrifice). (A) MN data showing the percentage of MN polychromatic erythrocytes PCE (bars) and the percentage of PCE (line) in bone marrow; (B) MN data in rat reticulocytes measured by flow cytometry at different time points after the last administration of CPA. Rats were treated with vehicle control (□, ○) or 10 mg/kg CPA (■, ●). Bars represent % MN in reticulocytes and lines represent the % reticulocytes among all erythrocytes. (C to E) Comet assay data. The percentage of clouds serves as indicator for cytotoxic effects (line). Tail intensity (□) and tail moment (■) were analysed as indicators of DNA damage. Rats were treated with vehicle control or CPA. The graphs show data from the rat stomach (C), liver cells (D), and from peripheral blood cells.

Azeddine Elhajouji reported about Novartis's efforts to integrate the MN into 2-week toxicity studies. They have performed six pilot studies to date and have not observed any interference with clinical chemistry/haematology, toxicokinetics (TK) or with the sampling for histopathology. The sampling for the evaluation of micronuclei is very simple as freezing is possible after a short fixation procedure, which makes this compatible with multi-site studies.

2.6. Combination of standard assays into one study

Covance's efforts to check whether bone marrow MN and comets in stomach/liver/blood can be combined in one assay using a single sampling time were presented by D. Kirkland. Groups of six male rats were dosed on three consecutive days at 0, 24 and 45 h. A range of different chemicals (CPA, DMN, 2AAF, B(a)P, EMS, ENU, MMC and carbendazim (CBZ)) were tested, three dose levels per chemical. At 48 h (i.e. 3 h after the final dose), stomach, liver and blood were removed for Comet assay, and bone marrow was removed for analysis of MN in polychromatic erythrocytes. Additional groups of five male rats were also dosed as above, but blood samples were taken at 3, 24 and 44 h after the last dose in order to see if MN in peripheral blood reticulocytes will also give a positive response at the same (or later) sampling time. Samples were split, one being sent to Litron and the other kept at Covance for flow cytometry analysis of MN in young reticulocytes. Preliminary data for two chemicals, cyclophosphamide (CPA) and carbendazim

Table 2
Animal savings at Novartis related to the performance of combination studies of the Comet assay and the micronucleus test compared to pairs of studies for the individual assays. The impact of the use of one gender only is also demonstrated.

Information from studies	Total	Male/female	Male
Number of combination studies	7	1	6
Number of compounds tested	11	2	9
	Number of studies	Number of animals used (number of studies)	
In combination	7	51 (1)	210 (6)
Isolated MNT	4	43 (1)	93 (3)
Isolated Comet	4	38 (1)	75 (3)
Total number of animals used		**132 (2)**	**378 (9)**
Calculations			
Number of male and female animals per compound per MNT study			43
Number of male and female animals per compound per Comet study			38
Number of male and female animals needed to test 11 compounds			=(43+38) × 11 = 891
Number of male and female animals per compound for a combination of MNT and Comet			51
Number of male and female animals needed to test 11 compounds in combination studies			=51 × 11 = 561; saved **330** animals **(37% reduction)**
Number of male animals per compound for a combination of MNT and Comet			=210/6 = 35
Number of male animals needed to test 11 compounds			=35 × 11 = 385; saved **506** animals **(57% reduction)**

(CBZ) were shown (see Fig. 5A–E for CPA data). With this limited data set, positive genotoxicity responses can be found for MN in bone marrow and comets 3 h after a three administration protocol. The micronucleus frequency in reticulocytes from peripheral blood also increased at this 3-h sample for both compounds, which may be the better option than MN in bone marrow. These initial results suggest that different endpoints (MN, comets) can be measured at a single sample time in the same animals. Data have subsequently been obtained for DMN, EMS, ENU, and MMC. In all cases positive responses have been found either for MN in bone marrow, or in reticulocytes from peripheral blood at the 3-h sample, or for comets. The experiments with 2AAF and B(a)P are ongoing. When complete, these data will be submitted separately for publication.

A. Rothfuss reported that within the Integration Initiative repeat-dose toxicity studies are planned, amongst others, with liver-specific genotoxins to investigate the possibility of integration of the Comet assay as a complement to the MN assay and to profile the *in vivo* Comet assay against the UDS *in vivo*.

Novartis is already successfully using the combination of MN and Comet assays. Out of 11 Comet assays performed under GLP, seven were combination studies. Most of these studies (9) were carried out using males only. Performance of all studies as combination studies would have led to a 37% reduction in animal numbers, while a 57% reduction can be achieved by combining this with the use of one gender only (see Table 2).

A multiple-endpoint genotoxicity assay, combining Comet assay, MN test and gene mutation in transgenic animals was presented by M. Hayashi (Fig. 6, Table 3). Two to three male and female MutaMice were treated with three different compounds (etoposide, bleomycin and procarbazine). Etoposide, a topoisomerase inhibitor, showed a clear increase in micronuclei in the peripheral

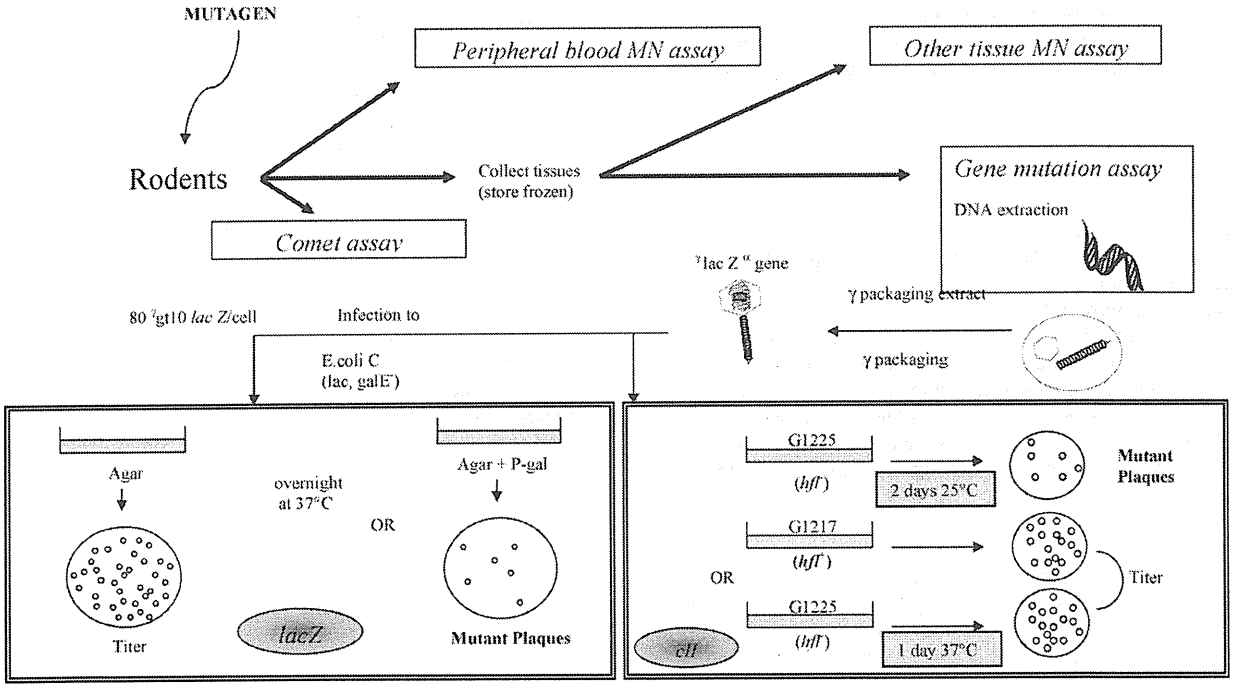

Fig. 6. Schematic of a multiple-endpoint assay combining the MN and Comet assays with the endpoint transgenic mutation (see Table 3).

Table 3
Multiple-endpoint assay in male and female transgenic animals (Muta™Mouse) for three test chemicals. DNA damage in blood was evaluated 3 h after administration of the test compound using the Comet assay (tail moment), induction of micronuclei 48 h after administration in reticulocytes (% micronucleated reticulocytes) and the frequency of *lacZ* mutations was determined in peripheral blood 14 days after the initial i.p. injection. Two to three male and female animals were used per group. Mice were treated with a single i.p. injection of test compound. Values are mean ± SD.

Test compound	Multiple-endpoint assay		
	Comet assay (3 h) Olive tail moment	MNT (48 h) MN/1000 reticulocytes	Mutation (14 day) mutation frequency (10^{-6})
Control (10 ml/kg saline)	1.2 ± 0.21	0.75 ± 0.5	9.24 ± 6.99
Etoposide (4 mg/kg)	1.16 ± 0.15	45.0 ± 14.5[a]	10.3 ± 3.19
Procarbazine HCl (200 mg/kg)	4.76 ± 0.18[a]	57.6 ± 16.8[a]	36.43 ± 36.88[a]
Bleomycin HCl (36 mg/kg)	2.55 ± 1.59[a]	4.2 ± 2.86[a]	34.81 ± 34.18[a]

[a] Represents a positive response.

blood while the other two compounds showed positive effects for all tree endpoints investigated.

3. Discussion and recommendations

Discussion of the issues and presentations took place in three break-out groups, followed by a final plenary session. There was general agreement that the current level of use of reduction opportunities is unsatisfactory. There are several actions that need to be taken in order to improve this situation, and the consensus recommendations are summarised in the following paragraphs. Some of the actions are short-term, i.e. to promote the use of reduction options that the participants considered ready to be implemented. Other actions are medium or long-term as they will require changes to the current OECD guidelines.

3.1. How to reduce the number of animals in standard genotoxicity tests?

3.1.1. Weight range

Many animals are wasted because laboratories order animals of a specific weight. When choosing the weight range for a study, the standard operating procedure (SOPs) should be as flexible as possible, to avoid unnecessary waste of animals at randomisation. We therefore encourage companies not to define too narrow a range for body-weight comparability and not to order an excessive number of animals for a study. In general, the criteria for a study should not be so strict that it becomes difficult to calculate the necessary number of animals.

3.1.2. Range finding study

If toxicity data are available, can range finder (RF) data from another facility be relied upon? With this in mind, although it may use additional animals, a small number of confirmatory RF experiments (e.g. 1–3 males and females each, as appropriate) may be worthwhile to confirm the maximum tolerated dose (MTD) and avoid possible failure of the main study.

Range finding (with or without prior data) should be undertaken by starting with the most likely dose to cause toxicity, using a small number of animals. If this does not define the MTD, then there is a need to increase or decrease the dose depending on the clinical effects of the first dose. It is a waste of animals to begin a range finder study with several dose groups and larger numbers of animals.

3.1.3. One sex versus two sexes

In the absence of data suggesting the use of both sexes, studies should use animals of one sex only. The available data suggest a slightly higher sensitivity to some positive control chemicals when males are used ([25] and Figs. 2 and 3). If at the time of the study, data show no substantial differences in toxicity between sexes, then use of a single sex should be sufficient (OECD 474 [26], paragraph 18). It is generally agreed that differences in toxicity of up to two-fold are often due to normal variability and are therefore not sufficient to demand the use of both sexes. Also, different severities of clinical effects in the two sexes would not usually justify the use of both sexes in the main study if the clinical signs themselves are similar. If available, other factors should also be taken into account, such as comparison of metabolic profiles and plasma concentrations. Some participants of the workshop suggested that where sex differences occur, the most sensitive sex can be used alone. However, if there are substantial differences between the sexes, we recommend that both sexes be used in the main study because there are insufficient data to date to identify which sex would be most likely to exhibit a positive genotoxic response. We do, however, encourage industry to evaluate available data to see whether the use of the most sensitive sex could be an option.

3.1.4. One administration and two sampling times versus two/three administrations and one sampling time

According to published literature and data presented at the workshop, the sensitivity of a protocol including a single administration with two sampling times is considered comparable with a protocol with two or three administrations and one sampling time. Whilst most comparisons of one administration *versus* two administration protocols have focused on clastogens [27], the aneugen carbendazim does elicit a positive response after three daily administrations, which provides further support for this assumption. The use of a two or three administration protocol can substantially reduce the number of animals (Table 1). Furthermore, this is a prerequisite for the combination of genotoxicity endpoints into a single study, since the kinetics of events leading to the measured endpoints differ significantly (e.g. primary DNA damage, as detected by the Comet assay, is an earlier event than MN formation, which requires the cells to go through cell division). Further saving of animals can be realized by measuring MN in blood (either through manual scoring or flow cytometry), which would allow for the use of pre-treatment samples to serve as negative controls instead of a concurrent vehicle control.

One concern voiced at the workshop was that dose-range finding for a two administration study may use more animals than for a single-administration study. However, according to the data presented at the workshop, this is not the case. We therefore strongly recommend the protocol using two or three administrations with one sampling time for acute MN and CA assays.

3.1.5. Positive controls

There was a consensus that, with sufficient experience of a laboratory, it is possible to omit the use of concurrent positive controls in routine MN and CA assays. For some other assays, e.g. the Comet assay and DNA adduct assays, it was the opinion of the workshop that at present, the technical and processing demands are such that concurrent positive controls should be used in these cases.

Table 4
Summary of results from the Japanese 13th Collaborative Study Group for the Micronucleus Test (CSGMT). The outcome of the testing of 15 different mutagens after daily administration of the compounds to rats for 28 days is shown for bone marrow (B.M.) and peripheral blood (P.B.), and for an initial measurement in P.B. at day 4.

Chemical name	Toxicity data/dose	Dose-range in this study (mg/kg/day)	Detection in the dose for general toxicity study			
			B.M. (4 week)	P.B. (4 week)	P.B. (4 day)	Overall call
2AAF	Mouse LD_{50}: 810 mg/kg	12–380; 60–240	○	○	△	○
B[a]P		31.3–250	○	△	X	○
CP	LD_{50}: 94 mg/kg	1–4; 15	●	●	○	●
DMH	LD_{50}: 100 mg/kg	5–50	X	X	X	X
2-Methylaziridine	LD_{50}: 19 mg/kg	2.5–20	○	○	○	○
MMS	LD_{50}: 225 mg/kg	3–30	●	●	●	●
MMC	LD_{50}: 30 mg/kg	0.08–0.5	X	X	△	△
Monocrotaline	LD_{50}: 66 mg/kg	3.75–30	X	X	○	○
Oil orange SS	LD_{50}: mg/kg	100–500	○	○	X	○
Phenacetin	LD_{50}: 3600 mg/kg	100–900	○	○	△	○
$KBrO_3$	LD_{50}: 321 mg/kg	10–80	○	○	X	○
Urethane	LD_{50}: 1809 mg/kg	89–300	●	●	●	●
Ara-C	Mouse LD_{50}: >5 g/kg	40–320	●	●	●	●
Colchicine	Mouse LD_{50}: 5886 μg/kg	0.7–6	○	X	X	○
6-MP	Mouse LD_{50}: 260 mg/kg	3.13–12.5	●	●	●	●

Symbols relate to the detection of general toxicity and represent "excellent correlation" (●), "good correlation" (○), "borderline between detectable and undetectable" (△) and "not measured" (**X**).

Periodically performed positive controls (e.g. every few months) for MN and CA assays would still be needed for quality control and furthermore would provide slides/samples that can be evaluated for experiments performed without a concurrent positive control, provided that the slides/samples were stored and coded appropriately. Hence, such slides/samples would allow checking the critical aspects such as the quality of staining and the ability of the evaluator or instrument used (e.g. flow cytometer, image analyser) to distinguish positive from negative responses. The key criteria that any laboratory will have to fulfil before it can consider omitting a concurrent positive control with every experiment will be to demonstrate consistent and reproducible detection of positive control responses in multiple independent experiments, as well as the presence of a well-defined historical negative/solvent control range in the specific strain of animals used [11]. Examples of positive control chemicals are given in OECD 474 and 475. Doses should be chosen so that the effects are clear but not excessive.

3.1.6. Toxicokinetics

Demonstration of exposure of the target tissue is important for all studies, and those performed without knowledge of whether the test substance will reach the tissue investigated for the genotoxic effect may be a waste of animals. Toxicokinetic (TK) data are routinely generated for pharmaceuticals but they may not be available for other substance classes, e.g. industrial chemicals. In these cases, other factors such as clinical symptoms or discoloured urine may help to demonstrate that systemic exposure occurred. In order to avoid the use of extra satellite animals when TK data are needed, *in vivo* genotoxicity testing should be conducted in the species for which TK data are generated. Genotoxicity protocols involving intravenous administration of test compounds may be an alternative to carry out TK studies in cases where quantitative systemic exposure data are not readily obtained.

3.2. Can the number of animals be reduced by applying smarter testing strategies?

3.2.1. Integration of the MN endpoint into repeat-dose toxicity studies

Integration of genotoxicity endpoints into repeat-dose studies (14- to 90-days RDT studies, dose-range finding studies, teratogenicity, etc.) is possible providing the highest dose tested is appropriate (i.e. represents a limit dose, MTD, maximum feasible dose (MFD), saturation, accumulation or accounts for at least 50% of the top dose of the acute toxicity test). The preferred method for MN analyses of blood sampled during the study is flow cytometry, however, acridine-orange staining with manual scoring is also considered acceptable.

The workshop participants addressed the question of whether integration of genotoxicity endpoints into acute toxicity testing would be possible. Acute toxicity testing uses a very limited number of animals and would test the dose groups in a step-wise manner. Therefore, no concurrent negative or positive controls would be included, which would be neither scientifically acceptable nor compliant with the current guidelines.

The influence of additional blood sampling on other parameters of repeat-dose toxicity (RDT) should be considered, but initial experience from pharmaceutical companies indicates that it is most probably not an issue. The integration of genotoxicity endpoints offers the possibility to cross-reference the results with routine haematology parameters evaluated during the RDT study for effects such as haematotoxicity or accelerated erythropoiesis, confound-

Table 5
Reduction options and their compatibility with current guidelines/recommendations.

	OECD	Pharmaceuticals e.g. ICH	Chemicals e.g. REACH ITS	IWGT and others
Dose range finding	No barrier	No barrier	No barrier	No barrier
Single sex	Covered adequately	Covered adequately Usually males	Covered adequately	Covered adequately Usually males
Omission positive control	No provision	Mentioned as option	Accepted, see OECD text and Art 13 par 3 REACH	Recommended
Combination of endpoints	Compliant except for sampling times	Recommended where appropriate	Scope available	Possible
	Applies to MN; UDS	Applies to MN; Comet; UDS; TG; etc.		Applies to MN; Comet; UDS; TG; etc.
Integration into repeated dose	No barrier	Recommended	Scope available	Proposed

ing factors known to have the potential to generate artefacts in the MN test [10].

Integration of the MN endpoint into RDT studies is in compliance with the OECD guideline for the MN test [26], strongly encouraged by ICH and an option in the REACH ITS [6] and, most importantly, scientifically credible. The workshop concluded that integration of the MN test should be the standard if RDT studies are foreseen for the test compound. In these cases, the performance of a stand-alone acute study should be justified. Such an integrated MN study should include an early blood sample taken 3 to 4 days after the start of dosing for highly clastogenic/toxic compounds (Table 4) [11].

3.2.2. Integration of the Comet assay endpoint into repeat-dose toxicity studies

The participants of the workshop viewed the integration of the Comet assay into RDT studies differently. To date, no published data are available to support this approach and industry experience is limited. There are several aspects that raise concern whether this is an appropriate option. As the Comet assay measures primary DNA effects which may be quickly repaired, according to the expert recommendations [14,15] the inclusion of a dose 3–6 h before sacrifice is needed unless the TK profile of the compound indicates accumulation. However, this may challenge the acceptability of the RDT study by influencing the general toxicity parameters or organ toxicity. If accumulation of the test compound over time occurs, this may not present a problem and the additional treatment may not be needed. However, the addition of a positive control group will still be required and therefore the full animal savings potential cannot be realized (see Section 3.1.5). Therefore, at this point in time, the workshop participants were unable to recommend this option. New information from the previously described initiative of EFPIA/PhRMA is expected and may alter this conclusion.

3.2.3. Combination of MN and Comet assay endpoints and integration into RDT studies

Ideally both the Comet and MN assays could be integrated into RDT studies but there is only limited scientific evidence available to support this. Most uncertainties are related to the Comet assay (see point 2b), however, the EFPIA/PhRMA initiative is planning to investigate this topic.

3.2.4. Combination of acute MN and Comet assay studies

In order to obtain the maximum information from one *in vivo* acute genotoxicity study, a combination of the *in vivo* MNT and the *in vivo* Comet assay is proposed. Such a combination can cover systemic genotoxic effects as well as local effects (site of contact tissue and target organ for toxicity) and different genotoxic mechanisms [28]. The test performance, including evaluation and interpretation of the results, should be in accordance with MN OECD guideline 474 and should meet state-of-the-art criteria regarding the performance of the *in vivo* Comet assay [14,15]. A test design that matches these criteria would involve administration of the test substance three times to each animal, at 48, 24 and 3–6 h prior to sacrifice. However, the third administration can be modified according to TK information, if available. It is recognised, however, that the third administration is not in compliance with OECD 474 as it is stated therein that the last administration should be within 18–24 h before sacrifice. While this is a formal deviation from the guideline, from a scientific point of view this is acceptable: the last administration is done to comply with the requirements for the Comet assay and no unfavourable impact on the outcome of the micronucleus test is expected. The last application relevant for the evaluation of micronuclei is done 24 h before sacrifice, which is in compliance with OECD 474.

According to the MN OECD guideline, three dose levels should be tested, while for the Comet, two (high and mid-dose) might be sufficient, whereas only one dose level is needed for MN and Comet if the limit test is performed. In order to avoid the inclusion of separate positive control animals for each endpoint, it is highly recommended to use a positive control compound that is suitable for both endpoints. In cases where a positive control might not be needed for the MN endpoint (see respective chapter), the Comet standard positive control can be used. Further studies are currently ongoing to confirm the compatibility of a three administration approach. However, based on the data presented at the workshop, the participants were convinced that such a combination of tests can already be used in laboratories that can demonstrate sufficient experience in both assays.

Other approaches for combinations of assays such as Comet and MN tests integrated into a transgenic mouse study, as presented by Makoto Hayashi, were also discussed and were considered to be valid from a scientific point of view, but currently there is a lack of experience.

3.3. How can the reduction possibilities be promoted and implemented?

The reduction options described above differ with regard to their guideline compliance (see Table 5). Fortunately, the majority of options, namely the use of single sex, one administration and two sampling times *versus* two administrations and one sampling time, and integration of genotoxicity endpoints into RDT studies, are in compliance with the OECD guidelines. Furthermore, their use is supported by the draft ICH S2 (R) guideline, encouraged by IWGT and covered in the REACH ITS. So why are these options not widely used today? The key factors seem to be uncertainty about regulatory compliance/acceptance and a simple lack of awareness. In addition, the workshop recognised that there is uncertainty regarding the scientific acceptance of reduction options, which may not be justified in all cases. However, a considerable amount of information generated by CROs and industry has not yet been published. For reduction options that can be implemented without generating conflict with the existing guidelines, the workshop participants were of the opinion that broader implementation can be achieved by better communication and motivation. The workshop agreed that the following actions may help to raise awareness and build consensus in the scientific and regulatory community:

- Collect more data regarding feasibility and use of reduction options and present these at scientific meetings.
- Generate and communicate data for those options that currently lack scientific proof (e.g. Comet integration into RDT studies).
- Use data to convince the national OECD Coordinators and encourage them to urge OECD to speed up the acceptance of amendments to guidelines and/or revisions of guidelines.
- Encourage people who are already using reduction possibilities to present their results and publish data in peer-reviewed literature.
- Use scientific, industry and regulatory societies and networks to promote reduction options.
- Use the distribution list from the ECVAM survey [12] to follow-up questions from the workshop.

It was the opinion of the workshop participants that the implementation of reduction possibilities that are not in full compliance with the guidelines, such as the omission of the positive control group, will depend on the interest/willingness of the regulators and the quality of the supporting evidence provided. The European Parliament has repeatedly voiced concerns about animal use in general and more specifically about the increase in animal numbers expected as a result of the REACH regulation. The EU Commission,

e.g. DG ENTR, DG ENV, and ECVAM (DG JRC), and the Community agencies such as ECHA, EFSA or EMEA through their involvement in guideline generation and acceptance, risk assessment and risk management, and steering of validation of non-animal methods, should play a key supporting role. However, a simple/single solution seems not to be realistic, and multiple approaches will be needed. The pressure to change may have to come from researchers whereby peer pressure and best practices could drive guidance and publications through industrial platforms and scientific societies such as ICH, the International workshop on Genotoxicity testing (IWGT), the Society of Toxicology (SOT) and the European Partnership on Animal Alternatives (EPAA) which then would help push forward changes in law/guidelines (OECD). Taking into account that even in a best-case scenario, changes in guidelines take years, the good will of the parties involved may have to be relied upon or alternative routes will have to be explored in order to move forward. One suggestion made at the workshop was that amendments to the OECD guidelines may be a faster way forward than revising the guidelines.

4. Conclusions

The premise under which the discussions have taken place was that any poorly designed or conducted experiment is a waste of animals and that the modification of methods should not compromise safety standards.

There was agreement amongst the workshop participants that there are many options available to reduce the numbers of animals in *in vivo* genotoxicity studies, and that most of them are in compliance with regulations and scientifically credible, i.e. ready for use. These options include the use of one sex only, one administration and two sampling times *versus* two/three administrations and one sampling time for MN, CA and Comet assays; the omission of a concurrent positive control in routine CA and MN tests; the combination of acute MN and Comet assay studies in case information from more than one endpoint or tissue is needed; and the integration of the MN endpoint into repeat-dose toxicity studies. These options are, to date, not sufficiently utilized and the workshop strongly encouraged the use and promotion of these options. The workshop participants want to encourage the scientific community to present and publish data related to reduction opportunities in order to boost the acceptance level of these approaches. Furthermore, experimental proof is needed and under way to demonstrate the credibility of additional options for reduction, such as the integration of the Comet assay into RDT studies.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

[1] Directive 67/548/EEC and draft EU Classification, Labelling and Packaging of Substances and Mixtures (CLP) Regulation, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52007PC0355:EN:NOT.

[2] Regulation (EC) No. 1336/2008 of the European Parliament and of the Council of 16 December 2008 amending Regulation (EC) No. 648/2004 in order to adapt it to Regulation (EC) No. 1272/2008 on classification, labelling and packaging of substances and mixtures, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:354:0060:01:EN:HTML.

[3] UN Globally Harmonised System (GHS), http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/02files_e.html.

[4] Regulation (EC) No. 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No. 793/93 and Commission Regulation (EC) No. 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32006R1907:EN:NOT.

[5] EU, Directive 2003/15/EC of the European Parliament and the Council of 27 February 2003 amending Council Directive 76/768/EEC on the approximation of the laws of the Members States relating to cosmetic products, Off. J. Eur. Union L66 (2003) 26–35.

[6] Guidance on information requirements and chemical safety assessment. Ch. R.7a: Endpoint Specific Guidance. See chapter R.7.7.1 Mutagenicity Testing, 374–402, http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_r7a_en.pdf?vers=20_08_08.

[7] D. Kirkland, M. Aardema, L. Henderson, L. Mueller, Evaluation of the ability of a battery of 3 *in vitro* genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. I. Sensitivity, specificity and relative predictivity, Mutat. Res. 584 (2005) 1–256.

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

[9] 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.

[10] M. Hayashi, R. Tice, J.T. MacGregor, D. Anderson, D. Blakey, M. Kirsch-Volders, F.B. Oleson, F. Pacchierotti, F. Romagna, H. Shimada, S. Sutou, B. Vannier, *In vivo* rodent erythrocyte micronucleus assay, Mutat. Res. 312 (1994) 293–304.

[11] M. Hayashi, J.T. MacGregor, D.G. Gatehouse, I.D. Adler, D.H. Blakey, S.D. Dertinger, G. Krishna, T. Morita, A. Russo, S. Sutou, 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 (2000) 234–252.

[12] P. Vanparys et al., in preparation.

[13] K. van der Jagt, S. Munn, J. Torslov, J. de Bruijn, Alternative approaches can reduce the use of test animals under REACH, http://ecb.jrc.ec.europa.eu/documents/REACH/PUBLICATIONS/Reducing_the_use_of_test_animals_under_REACH_IHCP_report.pdf.

[14] A. Hartmann, E. Agurell, C. Beevers, S. Brendler-Schwaab, B. Burlinson, P. Clay, A. Collins, A. Smith, G. Speit, V. Thybaud, R.R. Tice, Recommendations for conducting the in vivo alkaline comet assay, Mutagenesis 18 (2003) 45–51.

[15] B. Burlinson, R.R. Tice, G. Speit, E. Agurell, S.Y. Brendler-Schwaab, A.R. Collins, P. Escobar, M. Honma, T.S. Kumaravel, M. Nakajima, Y.F. Sasaki, V. Thybaud, Y. Uno, M. Vasquez, A. Hartmann, In vivo Comet Assay Workgroup, part of the fourth international workgroup on genotoxicity testing, Mutat. Res. 627 (2007) 31–35.

[16] S. Hamada, S. Sutou, T. Morita, A. Wakata, S. Asanami, S. Hosoya, S. Ozawa, K. Kondo, M. Nakajima, H. Shimada, K. Osawa, Y. Kondo, N. Asano, S. Sato, H. Tamura, N. Yajima, R. Marshall, C. Moore, D.H. Blakey, L.M. Schechtman, J.L. Weaver, D.K. Torous, R. Proudlock, S. Ito, C. Namiki, M. Hayashi, 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 (2001) 93–110.

[17] Y. Wakata, S. Miyamae, T. Sato, T. Suzuki, N. Morita, T. Asano, K. Awogi, Kondo, M. Hayashi, Evaluation of the rat micronucleus test with bone marrow and peripheral blood: summary of the 9th collaborative study by CSGMT/JEMS/MMS, Environ. Mol. Mutagen. 32 (1998) 84–100.

[18] M. Hayashi, J.T. MacGregor, D.G. Gatehouse, D.H. Blakey, S.D. Dertinger, L. Abramsson-Zetterberg, G. Krishna, T. Morita, A. Russo, N. Asano, H. Suzuki, W. Ohyama, D. Gibson, In Vivo Micronucleus Assay Working Group, IWGT. In vivo erythrocyte micronucleus assay III. Validation and regulatory acceptance of automated scoring and the use of rat peripheral blood reticulocytes, with discussion of non-hematopoietic target cells and a single dose-level limit test, Mutat. Res. 627 (2007) 10–30.

[19] J.T. MacGregor, M.E. Bishop, J.P. McNamee, M. Hayashi, N. Asano, M. Nakajima, A. Aidoo, M.M. Moore, S.D. Dertinger, Flow cytometric analysis of micronuclei in peripheral blood reticulocytes. II. An efficient method of monitoring chromosomal damage in the rat, Toxicol. Sci. 94 (2006) 92–107.

[20] S.B. Harper, S.D. Dertinger, M.E. Bishop, A.M. Lynch, M. Lorenzo, M. Saylor, J.T. MacGregor, Flow cytometric analysis of micronuclei in peripheral blood reticulocytes. III. An efficient method of monitoring chromosomal damage in the beagle dog, Toxicol. Sci. 100 (2007) 406–414.

[21] C.E. Hotchkiss, M.E. Bishop, S.D. Dertinger, W. Slikker Jr., M.M. Moore, J.T. MacGregor, Flow cytometric analysis of micronuclei in peripheral blood reticulocytes. IV. An index of chromosomal damage in the rhesus monkey (*Macaca mulatta*), Toxicol. Sci. 102 (2008) 352–358.

[22] K.L. Witt, E. Livanos, G.E. Kissling, D.K. Torous, W. Caspary, R.R. Tice, L. Recio, Comparison of flow cytometry- and microscopy-based methods for measuring micronucleated reticulocyte frequencies in rodents treated with nongenotoxic and genotoxic chemicals, Mutat. Res. 649 (2008) 101–113.

[23] A.M. Tometsko, D.K. Torous, S.D. Dertinger, Analysis of micronucleated cells by flow cytometry. 1. Achieving high resolution with a malaria model, Mutat. Res. 292 (1993) 129–135.

[24] S.D. Dertinger, D.K. Torous, N. Hall, C.R. Tometsko, T. Gasiewicz, Malaria-infected erythrocytes serve as biological standards to ensure reliable and consistent scoring of micronucleated erythrocytes by flow cytometry, Mutat. Res. 464 (2000) 195–200.
[25] CSGMT, Micronucleus test with mouse peripheral blood erythrocytes by acridine orange supravital staining: the summary report of the 5th collaborative study by CSGMT/JEMS.MMS. The Collaborative Study Group for the Micronucleus Test, Mutat. Res. 278 (1992) 83–98.
[26] OECD TG 474, OECD Guidelines for the Testing of Chemicals - Guideline 474, Mammalian Erythrocyte Micronucleus Test (adopted 21st July, 1997).
[27] CSGMT, Single versus multiple dosing in the micronucleus test: the summary of the fourth collaborative study by CSGMT/JEMS.MMS. Collaborative Study Group for the Micronucleus Test, the Mammalian Mutagenesis Study Group of the Environmental Mutagen Society, Japan (CSGMT/JEMS.MMS), Mutat. Res. 234 (1990) 205–222.
[28] S. Pfuhler, S. Albertini, R. Fautz, B. Herbold, S. Madle, D. Utesch, A. Poth, Gesellschaft fuer Umwelt-Mutationsforschung. Genetic toxicity assessment: employing the best science for human safety evaluation. Part IV: recommendation of a working group of the Gesellschaft fuer Umwelt-Mutationsforschung (GUM) for a simple and straightforward approach to genotoxicity testing, Toxicol. Sci. 97 (2007) 237–240.

The Journal of Toxicological Sciences (J. Toxicol. Sci.)
Vol.34, Special Issue I, SP189-SP197, 2009

Morphological characterization of the ovary under normal cycling in rats and its viewpoints of ovarian toxicity detection

Midori Yoshida[1], Atsushi Sanbuissyo[2], Shigeru Hisada[3], Michihito Takahashi[4], Yasuo Ohno[5] and Akiyoshi Nishikawa[1]

[1]Division of Pathology, Biology Safety Research Center, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan
[2]Medical Safety Research Laboratories, R & D Division, Daiichi-Sankyo Co., Ltd., 717 Horikoshi, Fukuroi, Shizuoka 437-0065, Japan
[3]Safety Research Department, Aska Pharmaceutical Co. Ltd., 1604 Shimosakunobe, Taktsu-ku, Kawasaki 213-8522, Japan
[4]Pathology Peer Review Center, 3-45-11 Uehara, Shibuya-ku, Tokyo 151-0064, Japan
[5]National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

(Received December 10, 2008)

ABSTRACT — Identification of ovarian toxicity is very important for safety assessment of drugs and other environmental chemicals. The detection of interference with ovarian function is very hard without a thorough understanding of the normal ovarian morphology based on reproductive physiology. The focus of the present study was therefore a practical analysis in each stage of the estrous cycles using ovaries obtained from 143 rats demonstrating normal cycling. Transversely dissected maximum areas in the ovaries were examined microscopically for the two major features, follicles and corpora lutea (CL). Classification of growing follicles was in reference to Pedersen and Peters (1968), and functionally divided into follicular stimulating hormone (FSH)-independent and dependent categories. The former, small and medium-sized follicles, respectively primordial/primary and preantral follicles, could be readily detected by immunohistochemical staining for proliferating cell nuclear antigen (PCNA). The large antral and Graafian follicles and large sized atretic follicles showed sequential changes depending on the estrous cycle stage. CL could be divided into currently and previously formed examples. Currently formed CL underwent remarkable changes in their appearance with the cycle, reflecting ovulation and progesterone production. Thus morphological analysis that is synchronized the large antral follicle changes with recently formed CL ones allows the ovary to be classified into the each estrous cycle stage. Morphological deviation from any synchronized combination provides a first pointer of ovarian toxicity. PCNA immunohistochemical staining is also useful to detect small follicles.

Key words: Estrous cycle, Morphology, Ovarian toxicity, Rat

INTRODUCTION

Detection of ovarian toxicity in preclinical studies is very important for safety assessment of drugs and chemicals, because oocytes have no regenerative ability, and any abnormalities in the ovaries may be directly linked to impairment of female reproductive capacity. The U.S. Environmental Protection Agency (EPA), U.S. Food and Drug Administration (FDA) and the Organization for Economic Cooperation and Development (OECD) recommended qualitative and quantitative evaluation of primordial follicles in the ovary for regulatory guidelines for 2-generation reproductive toxicity studies (U.S. EPA, 1998; U.S. FDA, 2000; OECD, 2001). The ovary has a complicated structure and its appearance changes with the estrous cycle, so that detection of interference with ovarian function requires a comprehensive understanding of the normal variation in ovarian morphology. In particular, the growth of follicles and corpora lutea (CL) and their regression, as well as knowledge on the hypothalmo-pituitary-gonadal axis control system are crucial. Quantitative follicular analysis using serial sections is accepted

Correspondence: Midori Yoshida (E-mail: midoriy@nihs.go.jp)

for assessment of both biological reproduction and toxicology (Bolon *et al.*, 1997; U.S. EPA, 1998; Meredith *et al.*, 1999; OECD, 2001; U.S. FDA, 2000), although the approach is very laborious and time-consuming.

According to the guideline of non-clinical safety studies, the completion of a female fertility study is required prior to the enrollment of women with childbearing potential in phase I clinical trials conducted in Japan. On the other hand, the completion is required prior to phase III clinical trials only in the United States and the European Union. Currently, the ICM M3 program to harmonize these regional differences is progressing (The ICH Steering Committee to the regulatory authorities of the three ICH regions (the European Union, Japan and USA), 2008). For this harmonization, collaboration among the members of the Japan Pharmaceutical Manufacturers Association was organized to conduct validation studies (Sanbuissho *et al.*, 2009). In the validation studies, any improvement to increase attention to toxicity to the female reproductive organs or many conventional guides or references to consistent, reliable and cost-effective methods are required using rat repeated toxicity studies. Recently, immunohistochemical staining of proliferating cell nuclear antigen (PCNA) has been accepted as a useful aid to making ovarian follicle counts, particularly for primordial and primary follicles (Muskhelishvili *et al.*, 2005; Picut *et al.*, 2008). The Society of Toxicologic Pathology (STP) ovary evaluation working group has published a position paper on histopathological approaches for the assessment of rodent reproductive toxicity, and recommended qualitative evaluation of the ovary as the first-tier assessment (Regan *et al.*, 2005). The OECD (2008) is preparing a guidance document for histopathological evaluation with endocrine and reproductive tests to detect chemicals with hormonal potential, including estrogens and antiestrogens. Westwood (2008) has just produced a good practical histological guide of the female reproductive tract in normal cycling rats. However, there is still only limited information on ovarian morphology of relevance to toxicological assessment. The main purpose of the present study was therefore to provide a practical guide of ovarian morphology in rats undergoing normal cycling. In addition, several morphological and methodological points for detection of ovarian toxicity are highlighted.

MATERIALS AND METHODS

Animals

Forty-eight Fischer344 DuCrlCrlj (Fischer) female rats aged 3 months and 92 Crlj:DON (Donryu) female rats aged 3 or 4 months purchased from Charles River Laboratories Japan, Inc. (Kanagawa, Japan) were available for the present study. The animals were maintained in an air-conditioned animal rooms under constant conditions of 24 ± 2°C and 55 ± 10% humidity with a 12-hr light/dark cycle (lighting from 8:00 to 20:00; dark, 20:00 to 8:00), housed 3 or 4 rats per cage, and allowed free access to commercial rodent chow, CRF-1 (Oriental yeast Co., Ltd., Kanagawa, Japan) and drinking water throughout. The Donryu strain rats were checked for estrous cyclicity by vaginal cytology during 9:00-10:00am throughout the study. Animal care and use followed the guidelines for the Care and Use of Laboratory Animals in the National Institute of Health Sciences.

Histology

At necropsy, all animals were euthanized under deep ether anesthesia in the morning. After exsanguination, the ovaries were removed from attached tissues such as ovarian bursa and oviduct, weighed and fixed in 10 vol% neutral buffered formalin. At dissection, the bilateral ovaries were transversally halved to examine maximum areas microscopically. The uterine horn and the vagina also were fixed to allow confirmation of the estrous cycle stage. All of the organs were embedded in paraffin, sectioned at 4 μm and stained with hematoxylin and eosin (HE). All the animals used in the present study showed a normal estrous cycle and the estrous cycle stages at necropsy were determined by vaginal cytology and morphological features in the uterus and the vagina with reference to previous studies (OECD, 2008; Westwood, 2008; Yuan and Foley, 2002).

Immunohistochemistry

Serial sections in the ovaries were incubated with anti-PCNA antibody (Dako Japan, Kyoto, Japan) to find small follicles such as primordial and primary follicles or to distinguish growing follicles as reported by Muskhelishvili *et al.* (2005).

Classification of follicles

In the present study, we applied Pedersen's follicular classification in rodents (Pedersen and Peters, 1968), featuring division into 3 main categories, i.e., small, medium and large corresponding to primordial or primary, preantral, and antral or Graafian follicles, respectively (Table 1). These were further subdivided into 8 types (Types 1 to 8) according to the morphological appearance and follicular size. The classification was established for mouse ovary, but application to the rat or hamster ovary has been recommended (Greenwald and Roy, 1994). In the present

Table 1. Classification of growing follicles

This article	Pedersen and Peters (1968)	Characteristics and Synonyms
Small	Small, Type 1a to 3a	Primordial and primary follicles. Type 3a follicle is a primary follicle with complete ring surrounding small oocyte.
Medium	Medium, Type 3b to 5a	Type 3b follicle is a primary follicle with complete ring surrounding growing oocyte. Secondary follicles and preantral follicles less than 70 μ on the largest cross-section.
Large	Large, Type 5b to 8	Type 5b follicle is preantral one with a fully grown ooctyes surrounded by many layers of granulosa cells. Type 6 follicle is a small antral one with a large oocyte with many layers of granulosa cells and scattered areas of fluid. Type 7 follicle is a antral one with single cavity containing follicle fluid. The cumulus oohorus has formed, but not be formed the stalk. Type 8 follicle is a large antral follicle with a single cavity with follicle fluid and a well-formed cumulus stalk. Type 8 follicle is also termed as Graafian or preovulatory follicles.

study, large follicles without oocytes in the sections were distinguished by morphological features of the granulosa cells and the thickness of the theca cell layer as far as possible. Follicles with apoptotic cells in the granulosa cell layer or degeneration of oocytes were judged as atresia. Atretic follicles were also classified into small, medium or large sized. If necessary, more detailed descriptions based on the follicle size in Pedersen's classification were added.

Classification of CL

CL were divided into currently formed CL and previously formed ones. Currently formed CL, which are also termed as most recently formed or new CL, are defined within one estrous cycle after ovulation. The previously formed CL, which are also termed as old CL, remain in the ovary throughout several estrous cycles before their complete dissolution. In currently formed CL, the ones observed at estrus when ovulation occurred early in the morning were defined as newly formed CL in this article.

Counts of Follicles and newly formed CL

The numbers of each type of follicle, atretic follicles and newly formed CL were counted in unilateral ovaries in F344 rats. Small and medium follicles were recognized on the basis of PCNA positive nuclei. Large follicles and CL were identified primarily in HE-stained samples. Separation was made into each estrous cycle stage. Statistical analysis was with ANOVA and P Values less than 0.05 were considered to be significant.

RESULTS

Small and medium follicles were distributed unevenly throughout the ovary, and they were easily detectable with PCNA immunohistochemical staining (Fig. 1). Small follicles were occasionally aggregated in the cortex. The estrous cycle did not affect their distribution or number. Typical atretic follicles were usually easy to detect at low magnification in HE stained sections. Higher magnification observation after HE staining was often necessary for detection of atretic follicles at early stages, focusing on apoptotic cells in the granulosa cell layer in early stages with other granulosa cells still positive for PCNA (Fig. 2).

Histological characteristics of each estrous cycle stage were as follows:

Proestrus (Fig. 3)

Follicles

Large follicles such as Graafian or Type 8 follicles in Pedersen's classification were evident at this stage. The large follicles Type 8 were located in surface area of the ovary (Figs. 3A-B). Their granulosa cells were cuboidal and/or polygonal, resembling the luteal cells at estrus (Figs. 3C-D). Most large follicles without the cumulus oophorus could be recognized by a thickened theca cell layer or cuboidal or polygonal granulosa cells (Figs. 3E-F). Large follicles Type 7 and some of Type 6 were atretic.

CL

Currently formed CL were starting degenerative processes characterized by vacuoles in cytoplasm, and increased apoptotic cells. Necrosis in the central parts was sometimes observed (Figs. 3G-H). Fibrous tissue proliferation was noted in previously formed CL.

Estrus (Fig. 4)

Follicles

There was a lack of large follicles such as Type 7 or 8. Although distinction of large follicles from degenerated large follicles was sometimes difficult at lower magnification, apoptotic granulosa cells in the degenerated large follicles were recognized them as atretic follicles (Fig. 4C). A number of large follicles of Type 6 were observed.

CL

Newly formed CL were characteristically observed as currently formed ones after ovulation (Figs. 4A-D). They were composed of basophilic, small and spindle-shaped luteal cells, positive for PCNA antibody binding (Figs. 4B-C). In addition, the newly formed CL were easily and clearly distinguishable from the large follicles by angiogenesis between the luteal cells and the break down of the basement membrane between the granulosa cell layer and the theca cell layer (Fig. 4D). Previously formed CL had a similar size as those at proesturs, but degenerative processes including fibrosis were more advanced (Figs. 4E-F).

Metestrus (Fig. 5)

Follicles

There were no large follicles of Type 8, but many large follicles Type 7 were growing (Fig. 5F). Atretic follicles were observed in all types of follicles at a constant rate.

CL

Currently formed CL were characteristically observed. They were increased in size compared to those at estrus, but smaller than those at diestrus (Figs. 5B-D). The luteal cells had still basophilic but not foamy cytoplasm with large nuclei and nucleoli prominent (Fig. 5C). The CL sometimes contained fluid-filled central cavities of various sizes (Figs. 5G-H). PCNA levels in the luteal cells were lower than in that at estrus (Fig. 5I). Previously formed CL demonstrated advanced fibrosis, but their sizes were still similar to those of currently formed examples.

Diestrus (Fig. 6)

Follicles

Large follicles Types 7 and 8 were increased in number, but the latter were smaller than at proestrus (Fig. 6A). Atretic follicles were observed in all types at a constant rate.

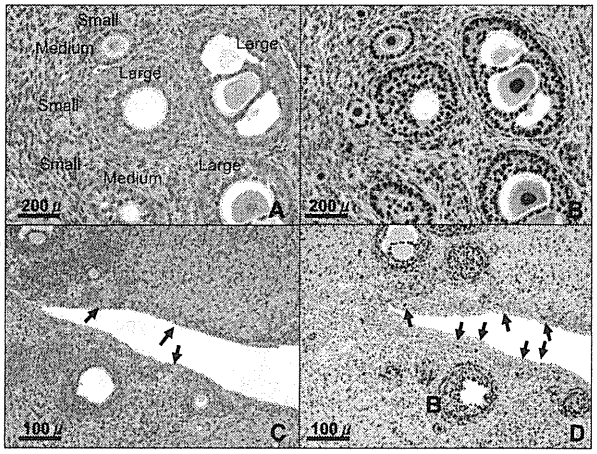

Fig. 1. Classification of follicles. **A**, Small follicles (small), medium follicles (medium), and large follicles are shown. **B**, Immunohistochemical staining for PCNA in a serial section to A. All oocytes and most granulosa cells are positive for PCNA. **C**, Small follicles (arrows) are not distributed uniformly and their identification with HE staining is difficult. **D**, Serial section to C. Note that small follicles (arrows) are easy to detect with PCNA immunohistochemical staining.

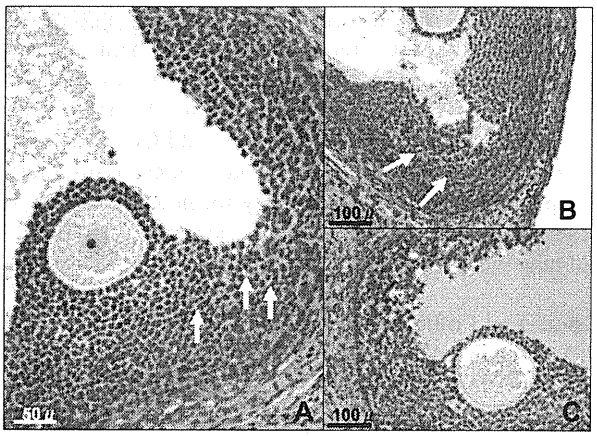

Fig. 2. Various types of atretic follicles. **A**, Early stage. A few apoptotic cells are apparent in the granulosa cell layer in a large follicle (arrows). **B**, Apoptotic cells are scattered in the granulosa cell layer (arrows). **C**, A number of apoptotic cells are present in the lumen and the granulosa cell layer. A-C, HE staining.

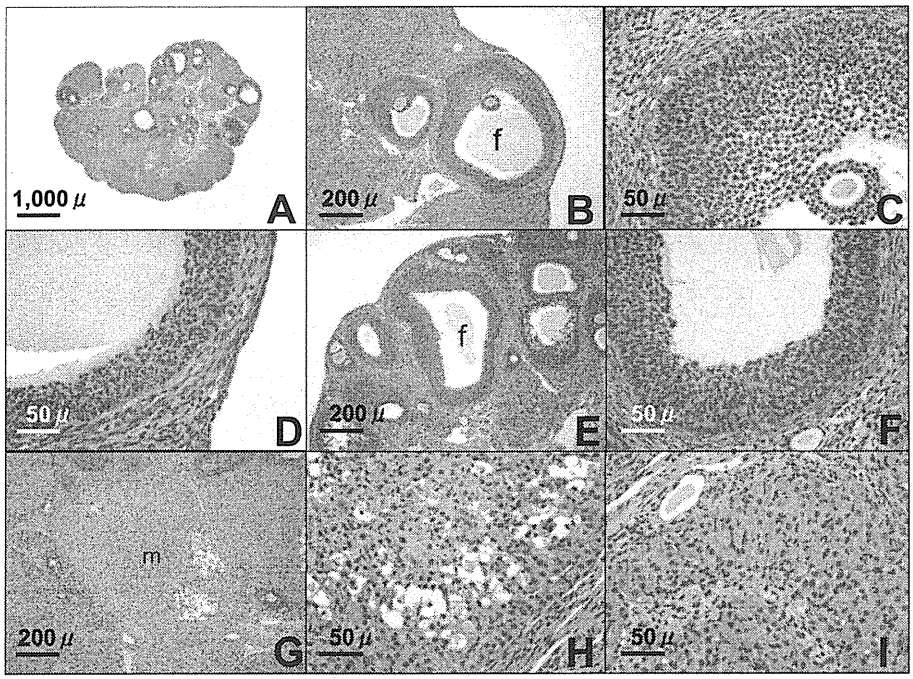

Fig. 3. The ovary at proestrus. **A**, Low-magnification. Large follicles are evident. **B**, A large follicle of Type 8 (f). **C**, Higher magnification of B. Cumulus oophorus formation. **D**, The same large follicle as B. The granulosa cells are round to cuboidal or polygonal in shape, indicating a luteinizing function **E**, A large follicle without cumulus oophorus (f). **F**, Higher magnification of E. Polygonal granulosa cells without apoptotic cells have characteristic of a large follicle Type 8. **G**, A currently formed corpus luteum. **H**, The corpus luteum contains vacuolated, degenerating and necrotic luteal cells in the central portion with some fibrous tissue. **I**, A previously formed corpus luteum. Note proliferation of fibrous tissue. A-I, HE staining.

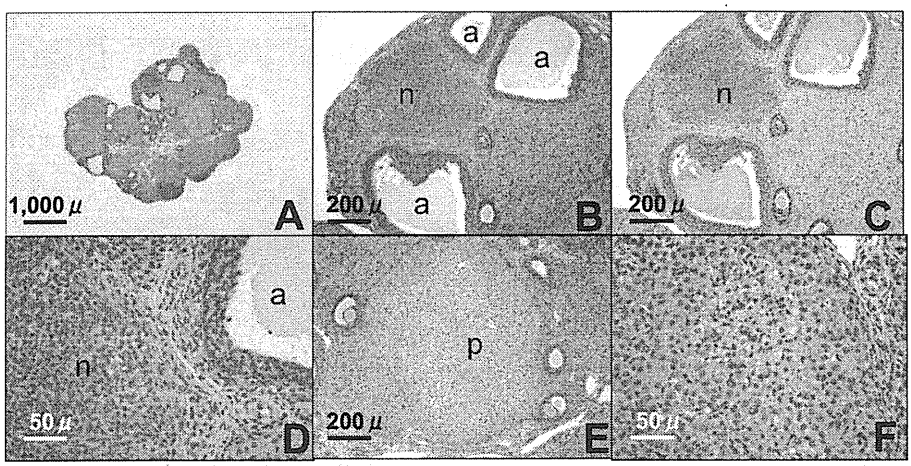

Fig. 4. The ovary at estrus. **A**, Low-magnification. **B**, A currently formed corpus luteum (n) is recognized as a newly formed one at this stage. Large sized atretic follicles are also observed (a). **C**, PCNA immunohistochemical staining of a serial section to B. A newly formed corpus luteum is strongly positive. **D**, Higher magnification of B. Spindle-shaped basophilic luteal cells and capillary formation are observed in a newly formed corpus luteum (n). A large sized atretic follicle is also detected at right upper part (a). **E**, Previously formed corpus luteum (p) of still large size. **F**, Higher magnification of E. Vacuolated and apoptotic cells are apparent. A, B, D-F, HE staining; C, PCNA immunohistochemical staining.

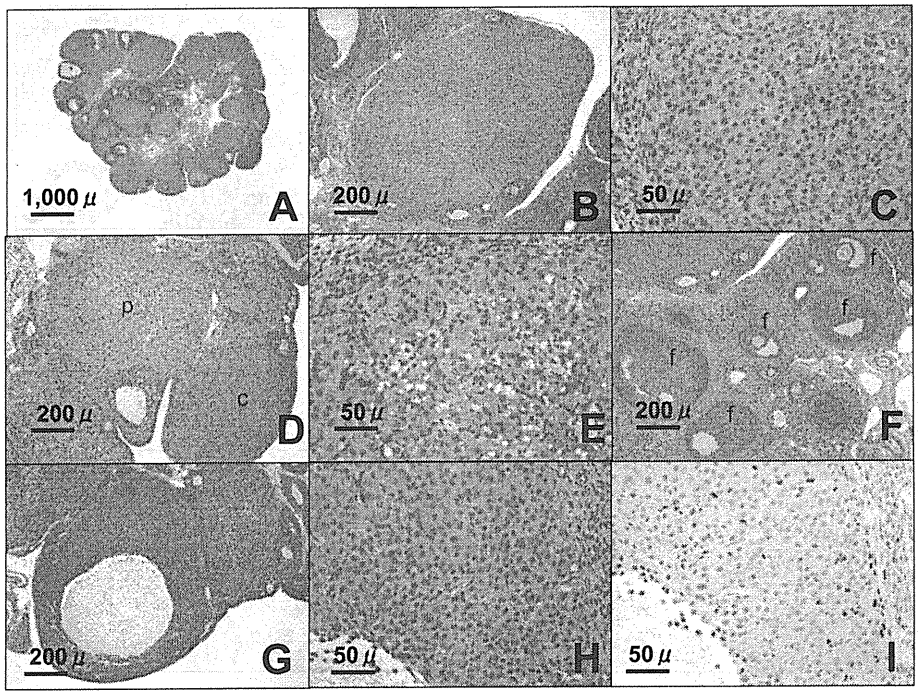

Fig. 5. The ovary at metestrus. **A**, Low-magnification. **B**, A currently formed corpus luteum. **C**, Higher magnification of B. The basophilic luteal cells sometimes exhibit prominent nucleoli. **D**, Currently formed (c) and previously formed (p) CL. Currently formed CL are still smaller than previously formed ones. **E**, A previously formed corpus luteum with fibrous tissue infiltration. **F**, Large follicles of Type 7 (f) appear at this stage. **G**, Currently formed copus luteum with a cavity. **H**, Higher magnification of G. Morphology of the luteal cells is very similar to 5C. **I**, PCNA immunohistochemical staining of a serial section to 5H. A-H, HE staining; I, PCNA immunohistochemical staining.

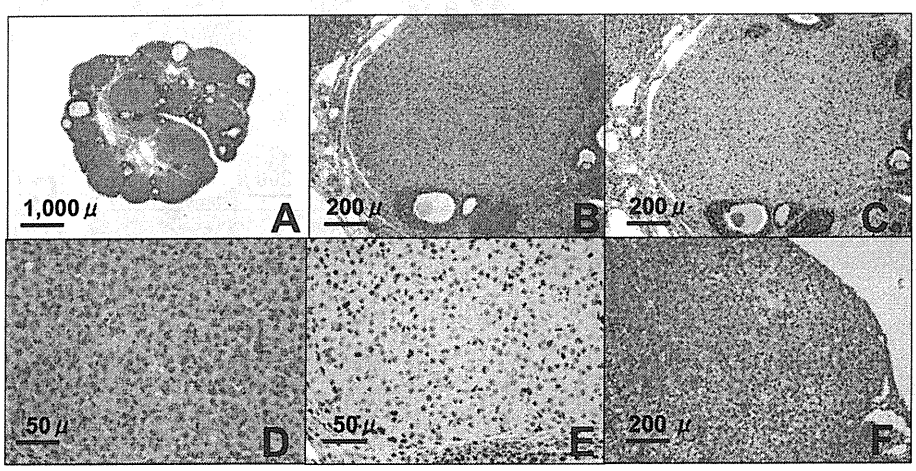

Fig. 6. The ovary at diestrus. **A**, Low-magnification **B**, A currently formed corpus luteum. **C**, A serial section of A immunohistochemically stained with PCNA antibody. **D**, Higher magnification of B. It is larger than at metestrus and the cytoplasm of their component cells is foamy and eosinophilic. **E**, Higher magnification of C. Most luteal cells are negative for PCNA, in contrast to fibrous tissue including fibroblasts or lymphoid cells. **F**, A previously formed corpus luteum with vacuolated luteal cells and fibrous tissue infiltration. A, B, D, F, HE staining. C, E, PCNA immunohistochemical staining.

CL

Currently formed CL had attained the maximum size (Figs. 6A-C). The luteal cells had foamy eosinophilic cytoplasm (Fig. 6D). Most luteal cells were negative for PCNA (Fig. 6E), and fibrous tissue were sometimes detected in the center of CL. Previously formed CL were vacuolated and fibrous tissue infiltration was advanced (Fig. 6F).

The average numbers of follicles and newly formed CL are shown in Table 2. Whereas the numbers of all follicular types and newly formed CL varied, the counts indicated some tendencies as follows: 1) small and medium follicles were found at constant ratios in all stages; 2) large follicles of Type 8 appeared at diestrus and reached a peak in proestrus; and 3) newly formed CL were found only in the estrus stage.

DISCUSSION

The present study demonstrated that the synchronized combination of morphological findings in large follicles, which are follicle stimulating hormone (FSH)-dependent (Greenwald and Roy, 1994), and currently formed CL is allowed as the ovary reliable clarification of each estrous cycle stage. In particular, classification of CL into 2 types appears very useful for staging. There were no morphological differences in either follicles or CLs between F344 and Donryu rats.

Our results for the estrous cycle are similar to those reported previously except for the detailed description of follicular changes (OECD, 2008; Westwood, 2008). In brief, in **proestrus**, the most characteristic feature is the presence of large follicles of Type 8 (Graafian follicles). In addition, cytoplasmic similarity of granulosa cells to luteal cells is characteristic of the Type 8 follicle. The increase in atretic change of Type 7 follicles indicated withdrawal of this type of follicle on the morning after ovulation. While currently formed CL at proestrus resembled those at diestrus, they appeared to be much more regressive. **Estrus** is the easiest stage to recognize due to newly formed CL. Confirmation of angiogenesis in CL in this stage is important to avoid misdiagnosing them from the edge of the Graafian follicles. A lack of large follicles of Types 7 and 8 is also a feature at estrus. Many large follicles of Type 6 might reflect the start of growing large follicles toward next ovulation (Watanabe *et al.*, 1990; Freeman, 2006). **At metestrus**, currently formed CL typicaly had luteal cells with basophilia and prominent nucleoli. This morphology might reflect progesterone production in rats classified into an ultrashort CL species (Freeman, 2006; Stouffer, 2006). **At diestrus,** currently formed CL were still characteristic. They reached the maximum size and contained luteal cells with eosinophilic and foamy cytoplasm. CL are not able to produce active progesterone at this stage (Watanabe *et al.*, 1990; Freeman, 2006).

CLs with cavities were encountered at metestrus in the present study. Their interpretation differs between OECD (2008) and Westwood (2008). Whereas OECD diagnoses them as ovarian luteal cysts which failed ovulation but undergoing luteinization of Graafian follicles in young adult rats (2008), Westwood described them as CL still containing occasional central fluid-filled cavities (2008). In domestic animals, CL with cavities have been accepted not to be confused with cystic follicles or luteinized cysts (Kenndey and Miller, 1993). On the other hand, the distinc-

Table 2. Follicular and newly formed CL counts per unilateral ovary (a) in F344 rats

	Estrous cycle			
Follicle/CL	Proestrus	Estrus	Metestrus	Diestrus
No. of rats examined	13	11	12	12
Follicles				
Small	15.9 ± 2.5 (b)	15.8 ± 3.3	17.0 ± 2.0	20.4 ± 4.4
Medium	2.2 ± 0.5	2.2 ± 0.5	2.0 ± 0.5	2.7 ± 0.4
Large (c)	16.7 ± 1.7	20.9 ± 0.1	18.0 ± 1.4	18.8 ± 2.2
Type 8	2.5 ± 0.3	0	0	0.8 ± 0.3
Atresia	10.8 ± 1.1	8.5 ± 1.4	11.9 ± 1.9	14.5 ± 1.6
CL				
Newly formed	0	1.6 ± 0.4	0	0

(a), the transverse section dissected to obtained maximum cut surface
(b), mean ± S.D.
(c), Including Type 8 follicles

tion between CL with cavity and luteinized cyst is sometimes difficult in rodents, multiovulated species. There are various description of similar morphology showing luteal tissue with cavity, such as CL containing cavities (Westwood, 2008), luteal cyst (OECD, 2008) or cystic CL (Yuan and Foley, 2002). In general, luteinized cysts are easy to distinguish from cystic follicles, which are very large atretic follicles and reflect impair of large follicle growth. They are caused by many chemicals and hormonal disturbances (Bogovich, 1991; Røste *et al.*, 2001; Baravalle *et al.*, 2006). In typical luteinized cyst arising from failure of ovulation and undergoing luteinization, the luteinized cells with angiogenesis formation are found in a thin cell layer or uneven distributed in crescent area in cows (Kenndey and Miller, 1993) and rats (Tamura *et al.*, 2009). The rupture of follicles at ovulation is mainly controlled by progesterone and prostaglandin, and disruptions of these hormones induce unruptured follicles with or without luteinization or formation of luteinized cysts (Davis *et al.*, 1999; Gaytán *et al.*, 2003; Shirota *et al.*, 1998; Tamura *et al.*, 2009; Tsubota *et al.*, 2009; Yuan and Foley, 2002). The CL containing small cavities seems to be normal range in normal cycling rats, because the cavity will be replaced to fibrous tissues next day, at diestrus. However, the toxicological significance of CL with cavities remains to be fully undetermined. We should pay attention an increase of CL containing cavities in treated groups. Further investigation or information is required. Sequential vaginal cytology might be very informative to check ovulation.

Detection of damage to small follicles is very important for detection of ovarian toxicity in preclinical studies. Small follicle counts using single ovarian sections are challenging and evaluation may be misleading (Meredith *et al.*, 1999; Bolon *et al.*, 1997). However, FSH-independent small and medium follicles (Greenwald and Roy, 1994) were observed in all sections examined in the present study and at a constant rate with the estrous cycle stage. The result indicates that lack or only a few visible small follicles in a single ovarian section might give a pointer to alarm for small follicular damage. In addition, PCNA immunohistochemical staining and microscopic observation of bilateral ovaries are helpful to increase the reliability of detection. The STP Ovarian evaluation workshop group (Regan *et al.*, 2005) recommended a qualitative morphological approach for ovarian sections by toxicologic pathologists familiar with normal reproductive cycle as the first tier approach for detection of ovarian toxicity.

In conclusion, the present results demonstrate that ovarian morphology are allowed each estrous cycle stage by combination of morphologic changes in large follicles and CL undergoing development and/or regression (Fig. 7). In the classification of CL into 2 categories, morphological changes in currently formed CL appear particularly useful for staging. In addition, a single ovarian section might be sufficient for analyzing each estrous stage. Our results indicate that any morphological deviation in follicles and CL from the synchronized combination might be the first indicator of ovarian toxicity. PCNA immunohistochemical staining is useful to detect small follicles. Vaginal cytology is also informative for detection of endocrinological condition.

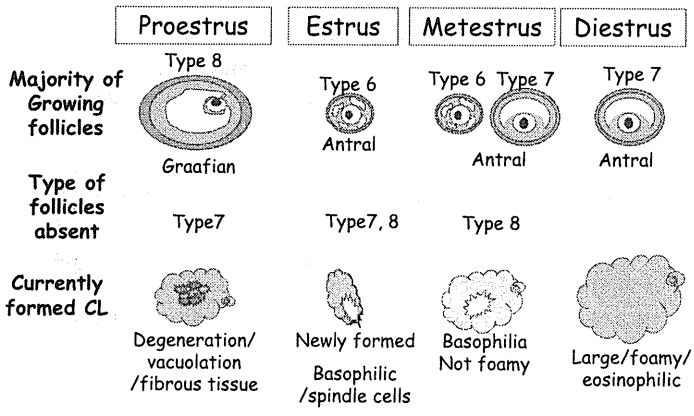

Fig. 7. A scheme for combination of morphologic features in follicles and CL at each estrous cycle in rats. The classification of follicles was referred to Pedersen and Peters (1968).

REFERENCES

Baravalle, C., Salvetti, N.R., Mira, G.A., Pezzone, N. and Ortega, H.H. (2006): Microscopic characterization of follicular structures in letrozole-induced polycystic ovarian syndrome in the rat. Arch. Med. Res., **37**, 830-839.

Bogovich, K. (1991): Induction of ovarian follicular cysts in the pregnant rat by human chorionic gonadotropin. Biol. Reprod., **45**, 34-42.

Bolon, B., Bucci, T.J., Warbritton, A.R., Chen, J.J., Mattison, D.R. and Heindel, J.J. (1997): Differential follicle counts as a screen for chemically induced ovarian toxicity in mice: results from continuous breeding bioassays. Fundam. Appl. Toxicol., **39**, 1-10.

Davis, B.J., Lennard, D.E., Lee, C.A., Tiano, H.F., Morham, S.G., Wetsel, W.C. and Langenbach, R. (1999): Anovulation in cyclooxygenase-2-deficient mice is restored by prostaglandin E2 and interleukin-1beta. Endocrinology, **140**, 2685-95.

Freeman, M.E. (2006): Neuroendocrine control of the ovarian cycle of the rat, In the Physiology of Reproduction, Third Edition (Knobil, E., Neill, J.D. eds.) pp.2327-2388, Raven Press Ltd., New York.

Gaytán, F., Bellido, C., Gaytán, M., Morales, C. and Sánchez-Criado, J.E. (2003): Differential effects of RU486 and indomethacin on follicle rupture during the ovulatory process in the rat. Biol. Reprod., **69**, 99-105.

Greenwald, G.S. and Roy, S.K. (1994): Follicular development and its control. In The Physiology of Reproduction, Second Edition (Knobil, E., Neill, J.D. eds.) pp.629-724, Raven Press Ltd., New York.

Kennedy, P.C. and Miller, R.B. (1993): The female genital system. In Pathology of Domestic Animals. Fourth Edition. (Jubb, K. V. F., Kennedy, P. C., Palmer, N. eds.) pp.349-470, Academic Press, San Diego.

Meredith, S., Dudenhoeffer, G. and Jackson, K. (1999): Single-section counting error when distinguishing between primordial and early primary follicles in sections of rat ovary of different thickness. J. Reprod. Fertil., **117**, 339-343.

Muskhelishvili, L., Wingard, S.K. and Latendresse, J.R. (2005): Proliferating cell nuclear antigen--A marker for ovarian follicle counts. Toxicol. Pathol., **33**, 365-368.

Organization of Economic Cooperation and Development (2001): Proposal for updating Guideline 416: two generation reproduction toxicity study. In OECD Guideline for Testing of Chemicals, pp. 1-13. OECD, Paris.

Organization of Economic Cooperation and Development (OECD) (2008): Draft Guidance and Review Documents/Monographs, Guidance Document for Histopathologic Evaluation of Endocrine and Reproductive Tests. Part 3: Female reproductive system. http://www.oecd.org/document/12/0,3343,en_2649_34377_1898188_1_1_1_1,00.html

Pedersen, T. and Peters, H. (1968): Proposal for a classification of oocytes and follicles in the mouse ovary. J. Reprod. Fertil., **17**, 555-557.

Picut, C.A., Swanson, C.L, Scully, K.L., Roseman, V.C., Parder, R.F. and Remick, A.K. (2008): Ovarian follicle counts using proliferating cell nuclear antigen (PCNA) and semi-automated image analysis in rats. Toxicol. Pathol., **36**, 674-679.

Regan, K.S., Cline, J.M., Creasy, D., Davis, B., Foley, G.L., Lanning, L., Latendresse, J.R., Makris, S., Morton, D., Rehm, S., Stebbins, K. and STP Ovary Evaluation Working Group (2005): STP Position Paper: Ovarian follicular counting in the assessment of rodent reproductive toxicity. Toxicol. Pathol., **33**, 409-412.

Røste, L.S., Taubøll, E., Berner, A., Isojärvi, J.I.T. and Gjerstad, L. (2001): Valproate, but not lamotrigine, induces ovarian morphological changes in Wistar rats. Exp. Toxicol. Pathol., **52**, 545-552.

Sanbuissho, A., Yoshida, M., Hisada, S., Sagami, F., Kudo, S., Kumazawa, T., Ube, M., Komatsu, S. and Ohno, Y. (2009): Collaborative work on evaluation of ovarian toxicity by repeated-dose and fertility studies in female rats. J. Toxicol. Sci., **34** (Special Issue I), SP1-SP22.

Shirota, M., Watanabe, G., Taya, K. and Sasamoto, S. (1998): Effects of indomethacin on the selective release of follicle-stimulating hormone during the period of ovulation in the rat. J. Vet. Med. Sci., **60**, 1059-65.

Smith, B.J., Plowchalk, D.R., Sipes, I.G. and Mattinson, D.R. (1991): Comparison of random and serial sections in assessment of ovarian toxicity. Reprod. Toxicol., **5**, 379-383.

Stouffer, R.L. (2006): Structure, function, and regulation of the corpus luteum. In the Physiology of Reproduction, Third Edition (Knobil, E., Neill, J.D. eds.), pp.475-526. Raven Press Ltd., New York.

Tamura, T., Yokoi, R., Okuhara, Y., Harada, C., Terashima, Y., Hayashi, M., Nagasawa, T., Onozato, T., Kobayashi, K., Kuroda, J. and Kusama, H. (2009): Collaborative work to evaluate ovarian toxicity 2) Two- or four-week repeated dose studies and fertility study of mifepristone in female rats. J. Toxicol. Sci., **34** (Special Issue I), SP31-SP42.

The ICH Steering Committee to the regulatory authorities of the three ICH regions (the European Union, Japan and USA) (2008): Guidance on nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals. M3(R2). http://www.pmda.go.jp/ich/m/Step3_m3(r2)_08 _07_14_e.pdf

Tsubota, K., Kushima, K., Yamauchi, K., Matsuo, S., Saegusa, T., Ito, S., Fujiwara, M., Matsumoto, M., Nakatsuji, S., Seki, J. and Oishi, Y. (2009): Collaborative work on evaluation of ovarian toxicity 12) Effects of 2- or 4-week repeated dose studies and fertility study of indomethacin in female rats. J. Toxicol. Sci., **34** (Special Issue I), SP129-SP136.

U.S. Environmental Protection Agency (1998): Health Effects Test Guidelines, OPPTS 870.3800, Reproduction and Fertility Effects. EPA 712-C-98-239. U.S. Environmental Protection Agency, Office of Prevention, Pesticides and Toxic Substances, Washington, DC.

U.S. Food and Drug Administration (2000): Redbook 2000. Toxicological Principles for the Safety Assessment for Food Ingredients. IV.C.9.a. Guidelines for Reproductive Studies. U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, Washington, DC.

Watanabe, G., Taya, K. and Sasamoto, S. (1990): Dynamics of ovarian inhibin secretion during the oestrous cycle of the rat. J. Endocrinol., **126**, 151-157.

Westwood, F.R. (2008): The female rat reproductive cycle: A practical histological guide to staging. Toxicol. Pathol., **36**, 375-384.

Yuan, Y.D. and Foley, G.L. (2002): Female Reproductive System. In Handbook of Toxicologic Pathology. Second Edition (Haschek, W.M., Rousseaux, C.G., Walling, M.A. eds.), pp.847-894. Academic Press, San-Diego.

Bull. Natl. Inst. Health Sci., **128**, 39-43 (2010) **Special Report**

3）構造活性相関による遺伝毒性の予測

本間正充

Structural Activity Relationship Approaches for Assessing Genotoxicity

Masamitsu Honma

The focus of the latest legislative and governmental efforts is to establish simple screening tools for identifying those chemicals most likely to cause adverse effects without experimental testing of all chemicals of regulatory concern. The use of structure-activity relationship (SAR) models is a powerful *in silico* technique that should be considered for prioritizing chemicals for subsequent experimental verification. Because carcinogenicity and genotoxicity are among the toxicological endpoints that pose the highest concern for human health, efforts in SAR models for them have been much more pronounced than for any of the other human health end points. This review paper overviews the historical background of SAR models for predicting carcinogenicity and genotoxicity, the current status of capacity and usefulness of some *in vitro* genotoxicity SAR models, and their perspective.

Keywords; Structural activity relationship (SAR), Genotoxicity, Ames tests, Sensitivity, Specificity

はじめに

化学物質の規制に関わる国際機関や，各国規制当局の最近の関心の焦点は，規制の対象となるすべての化学物質を実験的に試験することなく，有害作用を引き起こす化学物質を同定するための単純なスクリーニングツールを確立することにある．構造活性相関（Structure Activity Relationship; SAR）は，コンピュータトキシコロジーの重要な研究分野であり，有害作用を引き起こす可能性が高い化学物質を，その化学構造から*in silico*で予測する手法である．SARは統合型毒性評価システムの重要な構成要素の１つであり，安全性評価が必要とされる化学物質の優先順位付けや絞り込みに有用である．また，動物実験の代替，もしくは最小化にも貢献できる．SARは，創薬における探索試験段階での医薬品候補化合物の選択や，実際の試験が困難な不純物の安全性評価にも利用されている．現在，工業化学物質，農薬，食品添加物，化粧品材料，医薬品候補化合物の毒性予測のため多くのエキスパートシステムやSARツールが開発されている．

To whom correspondence should be addressed:
Masamitsu Honma: 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan; Tel: +81-3-3700-9847; Fax: +81-3-3700-2348; E-mail: honma@nihs.go.jp

発がん性，遺伝毒性予測とSAR

発がん性は，ヒトの健康にとって最も関心の高い毒性の一つであり，日常生活において暴露する可能性のある発がん性物質に関するSARモデル化の試みは，その他のヒトにおける健康関連のエンドポイントのいずれに対する試みよりも多大な努力が払われている．化学物質の構造から発がん性を予測する研究の歴史も古い．すでに1930年頃には強力な発がん物質として知られているベンツピレンの物性と発がん性との関連を明らかにするための研究が行われている．当初は，吸収，蛍光，赤外，NMRスペクトル，イオン化ポテンシャル，電位，磁気異方性，化学反応性などと，発がん性との関係が詳細に調べられたが，発がん性を規定する性質を見つけることはできなかった．1940年代にSchmidtおよびPullmanらはベンツピレン分子内の電子分布と発がん性の関係に注目し，一部の芳香族炭化水素の発がん性を合理的に説明することに成功した[1,2]．これらの研究が発端となり，いわゆるベンツピレン分子のK領域や，Bay領域と言われる部分構造と，発がん性との相関が明らかとなった．

1970年代，James & Elizabeth Millerらは発がん性アルキル化剤の求電子性に注目し，多くの発がん性化学物質は，求電子性誘導体か，もしくは生体内でそれらに代謝されて，発がん標的組織においてDNAやタンパク質などの求核性基と結合し，がんを引き起こすという理論を唱えた[3]．それ以降，化学発がんに関する研究は急速

に進展した．重要な進歩の一つは，化学発がん研究の重要なツールであるげっ歯類を用いた発がん性試験に対して，より安価で短期間の代替バイオアッセイが確立されたことである．Amesは発がん性化学物質（アルキル化剤，インターカレーターなど）に感受性をもつ一連のサルモネラ変異株を開発し，いわゆるエームス試験を確立した[4]．エームス試験は，発がん性化学物質を検出する*in vitro*モデルといえる．当時の既知の発がん物質の多くは遺伝毒性機序によるものでありエームス試験で陽性を示す化学物質の作用は，ほとんどMillers仮説の範疇で妥当と考えられた．現在では作用機序の観点から，発がん性物質は以下の二つに分類される．1)遺伝毒性発がん物質であり，DNAに直接損傷を与え，突然変異を誘発し，これが発がんの第1ステップになりうる．2)エピジェネテックな発がん物質であり，これはDNAとは共有結合はせず，直接的にDNA損傷を起こすことは無い．またエームス試験のような標準的な遺伝毒性試験では通常陰性を示す．1)のカテゴリーに属す遺伝毒性発がん物質がMiller仮説に従った特徴を持ち，それ自体が求電子性であるか，その代謝産物や代謝中間体が求電子性を持つ．

Millerの求電子理論に続いて，AshbyとTennantは発がん化学物質に対する構造アラート（Structural Alert; SA）と，発がん性予測のコンパイル（SARモデル）を開発した[5]．発がん性に対するSAは，化学物質の発がん性活性に関連した分子官能基，または部分構造と定義され，同時に発がんの主要なステップであるDNAへの損傷や，突然変異の誘発をもたらす遺伝毒性のSAとも考えられた．Ashbyは米国National Toxicology Programの222の化学物質の中からげっ歯類発がん性試験陽性と強い相関性を示す18種類のSAを同定した．Fig. 1に"Ashby's polycarcinogen"と呼ばれる18の全てのSAを持つ仮想の究極発がん物質を示す[5]．Toxnet公開データベース，Gold/ZeigerのCarcinogenic Potency Database，イタリアのInstituto Superiore di Santaの動物がん原性ISSCAデータベースに存在する698のエームス試験データと，878のげっ歯類発がん性試験データを用いてAshbyのSARモデルの検証を行ったところ，発がん性試験結果とは65%の一致に留まったのに対して，エームス試験結果とは78%が一致した[6]．このことは，エームス試験用サルモネラ菌株はDNA応答性に対して正しくデザインされており，Millerの提唱する発がん性物質の求電子理論を適切に反映することを示している．一方，他のバイオマーカーについてはSARモデルとの相関性が低い．このような事実から，SARモデルの予測率の検証にはエームス試験が使われることが多い．

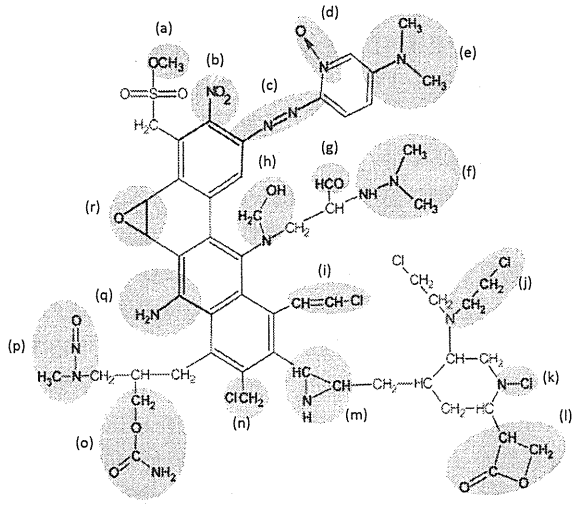

Fig. 1　"Ashby's poly-carcinogen"：Major structural units which are classified as structure-active positive in the Ames tests. The structures are as follows: (a) alkyl ester of either phosphonic or sulphonic acids; (b) aromatic nitro group; (c) aromatic azo group; (d) aromatic ring N-oxides; (e) aromatic mono-and dialkylamino group; (f) alkyl hydrazines; (g) alkyl aldehyde; (h) N-methylol derivatives; (i) monohaloalkanes; (j) a large family of N and S mustards; (k) N-chloramines; (l) propiolactones and propiosultones; (m) aromatic and aliphatic aziridinyl derivatives; (n) both aromatic and aliphatic substituted primary alkyl halides; (o) derivative of urethane; (p) alkyl N-nitrosoamines; (q) aromatic amines; (r) aliphatic and aromatic epoxide.

代表的なSARモデルによるエームス試験の予測

先に述べたようにエームス試験の予測に関しては多くのSARモデルが開発されている．そのモデルはアプローチ法により2つに大別される．1つは，知識ベース，規則ベースのエキスパートシステムで，Ashbyらが行ったように，既知データから陽性をもたらす特徴的な部分構造を定義し，ルール化された経験則から，定性的にエームス試験結果の予測を行うものである．もう一つは，化学物質の構造をフラグメントに分解後，パラメータ（数値データ）に変換し，エームス試験陽性と相関性の高いパラメータを用いて，多変量解析，パターン認識により試験結果を予測する人工知能型アプローチである．数値データから定量的な毒性の予測が可能であり，こちらはQSAR（Quantitative SAR）モデルである．前者としては英国ラーサ社のDEREK，後者のQSARとしては富士通が開発したADMEWORKSなどが代表的である．また，その中間型として，化学物質の構造と特徴を表す構造記述子と，多数の部分的構造（Biophore）を機械的に検出し，統計理論からエームス試験陽性と相関する構造記述子を選別し，予測を行うMultiCASEがある．3つのSARモデルの一般的な特徴に関しては本特論集の