Workshop on the Refinement of Mutagenicity / Genotoxicity Testing 23-24 April 2007, Malta

Workshop Report No. 9

Brussels, September 2007

ECETOC WORKSHOP REPORT No. 9

European Centre for Ecotoxicology and Toxicology of Chemicals
4 Avenue E. Van Nieuwenhuyse (Bte 6), B-1160 Brussels, Belgium.

The workshop was sponsored by the

Cefic Long-range Research Initiative (LRI)

LRI
The Long-range Research initiative

Refinement of Mutagenicity / Genotoxicity Testing

CONTENTS

1. EXECUTIVE SUMMARY

A workshop was convened by ECETOC to review current principles and methods of *in vitro* testing for mutagenicity and genotoxicity to identify areas for refinements and where there is scope for application of new methodologies. Currently, *in vitro* tests can show positive results with chemicals that are subsequently found to be negative in *in vivo* studies. This presents a considerable problem as the use of *in vivo* studies for safety testing of chemicals will be increasingly challenged under the new European chemicals legislation, REACh, and the 7th Amendment to the Cosmetics Directive precludes the use of animal testing. Thus, alternative testing strategies are being sought for all toxicological endpoints, including mutagenicity and genotoxicity. This has been made a topic under the Cefic Long-range Research Initiative which sponsored this meeting.

It was agreed that current *in vitro* testing methods were effective in detecting genotoxic carcinogens, albeit with some false positive findings. Non-genotoxic carcinogens are only poorly detected in these tests, and safety testing would benefit from introducing new methods. Emerging technologies and non-standard materials also require refined testing methods. Work groups evaluated new methods which use modified cells and human-specific metabolic activation systems. Furthermore, methodologies applying systems biology such as genomic technologies were considered. Related aspects were also raised including a better understanding of the role of population polymorphisms on susceptibility to carcinogens and the value of DNA-adduct biomarkers. The workshop identified areas where new improved methods are required and made some specific recommendations for future research on method development which are documented in this report.

2. WORKSHOP OVERVIEW

The use of alternative testing strategies to reduce the use of animals is becoming increasingly important within the context of the new European chemicals legislation, REACh, and the 7th Amendment to the Cosmetics Directive. Alternative approaches have been proposed for some toxicological endpoints but others, such as that for mutagenicity or genotoxicity, still need to be developed for more widespread application.

The testing for mutagenicity and genotoxicity often gives rise to positive *in vitro* genetic toxicity results, which then require additional *in vivo* testing. Further, several classes of materials such as chelating agents, heavy metals and some surfactants are not suitable for *in vitro* testing due to historic false positive results and/or incompatibility with *in vitro* test systems. Consequently, Cefic LRI (Long-range research initiative) proposed a Workshop to review the possibilities of refining mutagenicity / genotoxicity tests.

Therefore, the objective of the Workshop was to review the challenges of current *in vitro* testing and development of future *in vitro* type tests for mutagenicity and genotoxicity. This included looking at both new and refined biological systems and new methodology including those involving systems biology.

The Workshop participants were some 40 invited scientific experts from industry, academia, and governmental agencies. The Workshop began with detailed presentations followed by discussions on specific topics in five breakout groups and a final plenary session on the conclusions from the breakout groups.

3. PLENARY LECTURES AND BREAKOUT GROUPS

3.1 Plenary Lectures

Introducing the Workshop, Professor Parry outlined the objective and context within the framework of the LRI programme. On the first day, six invited experts gave plenary lectures, the topics of which had been selected to stimulate the discussions for the subsequent breakout groups.

Professor Parry reviewed **what brought us to the current state of mutagenicity / genotoxicity testing**. He said that the discipline of genetic toxicology has developed as a science which assesses the potential of chemicals and radiation to induce mutations in the genome of exposed organisms. Key driving factors in the development of the science were the realisation of the roles of both gene and chromosome mutations in birth defects and in the initiation and progression of cancer. To detect mutagenic activity a wide range of both *in vitro* and *in vivo* assay systems have been developed. The identification of effective assays has involved extensive national and international validation projects, which has resulted in advisory and regulatory bodies being able to recommend both batteries of assays and their application in test schemes for compounds of different types and uses.

When the packages of recommended assays are used we have considerable confidence that the vast majority of DNA reactive genotoxins can be identified and their potential hazards assessed. However, there is now evidence that some individual assays produce false positive results that are not always predictive of carcinogenic potential and therefore their usage should be reconsidered. Political influences such as requirements to reduce or eliminate *in vivo* studies for some products raise important questions as to whether and how we can continue to provide the public with the same level of reassurance with regard to the genotoxic safety of all compounds.

Mechanistic understandings of the relationship between the formation of DNA lesions, repair and mutant fixation has enabled the development of extremely sensitive methods such as adduct quantification which enable the detection of early events in mutagenic pathways. There are many developments in areas such as genomics leading to improved understanding of human variation and in areas such as image and flow analyses that may provide us with the basis for improved assays. Professor Parry concluded that it is now an appropriate time to consider what methods we may need now and in the future and to initiate developments which may fill critical gaps in our safety assessment processes.

Dr. Kirkland spoke about the **problems of false positives and false negatives in current *in vitro* testing**. In a recent review of methodology, he has shown an acceptable sensitivity of the

current methodology for mutagenic carcinogens[1], and this sensitivity improves slightly with a battery of two or three *in vitro* tests. However, many non-carcinogens and non-genotoxic carcinogens are falsely identified. The tests in mammalian cells have a very poor specificity. This means that though the tests pick out genotoxic carcinogens with success this is often at the expense of a high false positive rate. Other analyses and reports[2] [3] also indicate a high rate of positive results for mammalian cell tests *in vitro*. High frequencies of false positive results lead to unnecessary animal testing and delayed development of new products. A better understanding of the reasons for false positive results and how performance may be improved is needed, for example:

- Control of reactive oxygen species that may be produced by interaction between test chemical and culture medium.
- Deficiencies in detoxification, p53 function and DNA repair capability in rodent cell lines (e.g. CHO, CHL, V79, L5178Y) contribute to their high false positive rate. Human lymphocytes may have a lower rate of false positives than the rodent cell lines[2] [4].
- Better understanding of the likely mechanisms of genotoxicity and interpretation of positive results irrelevant for humans is urgently needed both for practitioners and regulatory reviewers.
- More appropriate use of concentrations for testing based on some *a priori* knowledge of the chemical, its metabolism and likely concentrations in use.
- Control of the level of cytotoxicity which can lead to false positive results. Various measures are currently allowable under OECD guidelines. A detailed comparison of multiple measures of cytotoxicity is needed to see if different measures would select different concentrations for test.
- Human cells that are p53 and DNA-repair proficient, and have defined phase 1 and phase 2 metabolism may offer the best hope for reduced false positives in the future. Human lymphocytes, HepG2 and TK6 cell systems also show promise for future testing approaches. Other human cell lines, such as HepaRG, should be evaluated for use in genotoxicity testing. A collaborative research programme is needed to review, identify and evaluate new cell systems with appropriate sensitivity but improved specificity.

[1] Kirkland D, Aardema M, Henderson L, Müller L. 2005. Evaluation of the ability of a battery of three *in vitro* genotoxicity tests to discriminate rodent carcinogens and non-carcinogens: I. Sensitivity, specificity and relative predictivity. *Mutat Re*s 584:1-256. Also see Lhasa database: www.lhasalimited.org.cgx.

[2] Kirkland D, Pfuhler S, Tweats D, Aardema M, Corvi R, Darroudi F, Elhajouji A, Glatt H, Hastwell P, Hayashi M, Kasper P, Kirchner S, Lynch A, Marzin D, Maurici D, Meunier JR, Müller L, Nohynek G, Parry J, Parry E, Thybaud V, Tice R, van Benthem J, Vanparys P, White P. 2007. 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* 628:31-55.

[3] Matthews EJ, Kruhlak NL, Cimino MC, Benz RD, Contrera JF. 2006. An analysis of genetic toxicity, reproductive and developmental toxicity, and carcinogenicity data: I. Identification of carcinogens using surrogate endpoints. *Regul Toxicol Pharmacol* 44:83-96.

[4] Hilliard C, Hill R, Armstrong M, Fleckenstein C, Crowley J, Freeland E, Duffy D, Galloway SM. 2007. Chromosome aberrations in Chinese hamster and human cells: A comparison using compounds with various genotoxicity profiles. *Mutat Res* 616:103-118.

Dr. Gant gave a presentation on the **potential application of toxicogenomics to mutagenicity / genotoxicity testing**. He explained that genomics has been part of the toxicologists toolbox for nearly a decade. During that time there has been much excitement and expectation of its potential to inform on molecular toxicity particularly using *in vitro* systems and over shorter time periods of exposure. To date, general expectation has been greater than delivery and this has led to some disillusionment and frustration with the technique[5]. To a very large extent this has been driven by: (i) exaggerated expectation, (ii) underestimation of the sheer complexity of the data and analysis required, (iii) difficulty in variable control and (iv) inappropriate experimental design. Recent data from the US FDA has indicated a large proportion of the technical variability has been controlled. Concurrent with technical improvement there has been a proliferation of bioinformatic tools. These have increased the ease of data analysis but also its complexity in terms of choice of analysis methods. We are now at a position, where as many of the technical barriers to effective implementation have been overcome, exploration of the true toxicological applicability of the technique can commence. Careful experimental design and considered bioinformatics will always be required.

Despite the difficulties of the application of genomic techniques in toxicology, discernment of genotoxic and non-genotoxic carcinogens is one of the success stories. Seminal studies carried out by many individuals in different groups[6] [7] [8] have shown the future potential. There is though more work to be done. While the published work shows that genotoxic and non-genotoxic molecules can be effectively separated, often these studies are at one dose or time point, and many of the genotoxins used form bulky DNA adducts. Other genotoxic carcinogens operating through different mechanisms of action need to be tested.

There is therefore a need to expand the number of chemicals assayed to gain a greater understanding of the mechanisms involved in genotoxic and non-genotoxic carcinogenicity, and look at the toxicological mantra of dose and time responses where data are lacking. Controversially, there may also be a need to undertake more work *in vivo* so that the data gained *in vitro* can be fully correlated with physiological response and pathological change. There is a lack of understanding of biological relevance of gene expression biomarkers particularly at low levels of exposure typically seen in exposed populations. As much genomic data is analysed by reference to the phenotype it is important that the phenotype be measured in the normal physiologic environment of the whole system before referencing to the *in vitro* situation. In this way, the application of genomics technologies to assess different adduct types, particularly at low

[5] Albertini S. 2005. Toxicogenomics in the pharmaceutical industry: hollow promises or real benefit? *Mutat Res* 575:102-115.
[6] van Delft JHM, van Agen E, van Breda SGJ, Herwijnen MH, Staal YCM, Kleinjans JCS. 2004. Discrimination of genotoxic from non-genotoxic carcinogens by gene expression profiling. *Carcinogenesis* 25:1265-1276.
[7] Ellinger-Ziegelbauer H, Stuart B, Wahle B, Bomann W, Ahr HJ. 2005. Comparison of the expression profiles induced by genotoxic and non-genotoxic carcinogens in rat liver. *Mutat Res* 575:61-84.
[8] Hockley SL, Arlt VM, Brewer D, Giddings I, Phillips DH. 2006. Time- and concentration-dependent changes in gene expression induced by benzo(a)pyrene in two human cell lines, MCF-7 and HepG2. *BMC Genomics* 16:260-283.

exposure levels, can be determined and phenotypically anchored. These data can then be correlated with the *in vitro* findings to allow an objective assessment of the utility of the *in vitro* models. With this information the best *in vitro* models and genomics assessment tools can be chosen for analysis of the effects of genotoxin exposure.

Further details on such new technologies provided by Dr. Gant are given in Appendix 2. In addition to transcriptomics the other new genomic technologies presented in Appendix 2 have much to offer in the development of carcinogenicity testing, particularly for non-genotoxic carcinogens. In this regard, attention should be given to the use of global technologies for the assessment of epigenetic DNA alteration and assessment of copy number variation for the further refinement of carcinogenicity testing.

Professor Oesch in his presentation reviewed the **metabolism and genetic susceptibility in *in vitro* testing for mutagenicity / genotoxicity**. He said that Professor Bruce Ames, and others at about the same time, recognised that mammalian metabolism is needed for most mammalian carcinogens to be toxified to the ultimate carcinogens. As a result, Ames found that addition of mammalian liver homogenates to bacterial mutagenicity assays allows detection of most mutagenic carcinogens. This has revolutionised early fast warning possibilities for potential carcinogenicity.[9] Nevertheless the ingenious fast and simple Ames-test has several fundamental deficiencies which can be overcome by refinement when needed. The standard Ames test is optimised for CYP activity whereas many predominantly detoxifying systems are little or not operative. Professor Oesch mentioned that his research group discovered that several of these are simple enough to be readily taken into account[10]. However, hitherto unknown systems additionally need to be considered[11] which in some cases outweigh in terms of activity the known systems[12]. They also discovered that the predictive power of mutagenicity tests, at least in the investigated cases, is best when using intact cells instead of cell homogenates[13]. Further, such intact cells from several species including human can be cryopreserved and used simply as chemical reagents[14]. Other groups came out with genetically modified animals and cells which

[9] Ames BN, Durston WE, Yamasaki E, Lee FD. 1973. Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. *PNAS* 70:2281-2285.
[10] Oesch F, Bentley P. 1976. Antibodies against homogeneous epoxide hydratase provide evidence for a single enzyme hydrating styrene oxide and benz(*a*)pyrene 4,5-oxide. *Nature* 259:53-55.
[11] Glatt HR, Vogel K, Bentley P, Oesch F. 1979. Reduction of benzo(*a*)pyrene mutagenicity by dihydrodiol dehydrogenase. *Nature* 277:319-320.
[12] Glatt HR, Cooper CS, Grover PL, Sims P, Bentley P, Merdes M, Wächter F, Vogel K, Guenthner T, Oesch F. 1982. Inactivation of a diol epoxide by dihydrodiol dehydrogenase but not by two epoxide hydrolases. *Science* 215:1507-1509.
[13] Glatt HR, Billings R, Platt KL, Oesch F. 1981. Improvement of the correlation of bacterial mutagenicity with carcinogenicity of benzo(*a*)pyrene and four of its major metabolites by activation with intact liver cells instead of cell homogenate. *Cancer Res* 41:270-277.
[14] Hengstler JG, Utesch D, Steinberg P, Platt KL, Diener B, Ringel M, Swales N, Fischer T, Biefang K, Gerl M, Bottger T, Oesch F. 2000. Cryopreserved primary hepatocytes as a constantly available *in vitro* model for the evaluation of human and animal drug metabolism and enzyme induction. *Drug Metab Rev* 32:81-118.

have a given drug metabolising enzyme gene knocked out and replaced by a homologous or otherwise closely related human gene[15] [16].

All these studies have clearly shown that mutagenicity / genotoxicity is crucially dependent on the metabolism of the system and animal species used. It is therefore imperative to choose an *in vitro* test system which is sufficiently close to humans. Further, extrapolation to humans required acknowledge of the differences of metabolic control between the *in vitro* test systems and those of humans. This will then allow a rational extrapolation of these data. Refinement of mutagenicity / genotoxicity tests taking these facts into account will improve their predictive power. There is still though a need to further research and refine *in vivo* metabolic activation systems to ensure they mimic the human for the chemical under test. By definition this means that *a priori* knowledge of human metabolism of a chemical is required before testing for mutagenicity.

Professor Speit discussed the **measurement of DNA damage and its relevance to mutagenicity / genotoxicity testing**. He started with mentioning that indicator tests (tests for DNA effects) are frequently used in mutagenicity testing. Mutagenicity testing is commonly subdivided into 'basic testing' and 'follow-up testing'. While mutation endpoints are almost always investigated in basic testing, indicator tests are frequently used in follow-up testing. They also play an important role as screening tests and in human biomonitoring. Indicator tests measure DNA alterations or other endpoints (e.g. repair, recombination) which are related to mutations. The main advantage of tests measuring DNA damage such as DNA adducts (e.g. postlabelling) or DNA strand breaks (e.g. comet assay) are their sensitivity and potential to detect organ-specific effects. The main limitation is the assessment of the biological relevance of findings – in particular of marginal effects in the low dose range.

The *in vivo* comet assay is increasingly being used in mutagenicity / genotoxicity testing[17]. This is due to its applicability to various tissues and cell types, its sensitivity for detecting low levels of DNA damage, its requirement for small numbers of cells per sample, general ease of test performance and the short time to complete a study at relatively low cost. International expert groups have published recommendations for an appropriate test performance in accordance with OECD guidelines for other genotoxicity tests[18] [19]. The main applications for the *in vivo* comet

[15] Gonzalez FJ. 2003. Role of gene knockout and transgenic mice in the study of xenobiotic metabolism. *Drug Metab Rev* 35:319-335.
[16] Gonzalez FJ, Yu. 2006. Cytochrome P450 and xenobiotic receptor humanized mice. *Annu Rev Pharmacol Toxicol* 46:41-64.
[17] Brendler-Schwaab S, Hartmann A, Pfuhler S, Speit G. 2005. The *in vivo* comet assay: use and status in genotoxicity testing. *Mutagenesis* 20:245-254.
[18] Hartmann A, Agurell E, Beevers C, Brendler-Schwaab S, Burlinson B, Clay P, Collins A, Smith A, Speit G, Thybaud V, Tice RR. 2003. Recommendations for conducting the *in vivo* alkaline Comet assay. 4th International Comet Assay Workshop. *Mutagenesis* 18:45-51.

assay are: (i) follow-up testing of positive *in vitro* findings, (ii) assessment of site specific genotoxicity and (iii) evaluation of the contribution of genotoxicity to the formation of neoplastic changes in rodent bioassays.

A positive result in an appropriately performed test has to be taken as an indication for a mutagenic potential of the test compound. It gains particular significance when a mutagenic potential of the test compound has already been demonstrated *in vitro*. Besides, the biological significance of the observed effect (e.g. exclusion of secondary effects due to cytotoxicity), the quality of the test performance and the plausibility of the result should be critically evaluated in the context of existing genotoxicity and toxicokinetic data.

In the context of new requirements for genotoxicity testing (e.g. the 7th Amendment of the European Cosmetics Directive) the comet assay may gain additional importance as an *in vitro* test with primary cells, with genetically engineered cell lines and with organ models (e. g. 3D skin models). These applications may lead to improved hazard characterisation and a reduction in animal testing. Further to this, Professor Speit mentioned advantages and limitations of the comet assay in relationship to other indicator tests.

Dr. van Benthem gave a detailed presentation on the **relevance of transgenic animals to mutagenicity / genotoxicity testing**. He explained that the assessment of the potential genotoxicity of chemicals *in vivo* plays an important role both for verification and confirmation of intrinsic mutagenicity and for establishing the mode of action of chemical carcinogens. There are appropriate tests to investigate *in vivo* chromosome aberrations and aneuploidy. However, until recently, the absence of reliable *in vivo* gene mutation tests hampered a justified assessment of the genotoxic potential of chemicals. The *in vivo* gene mutation tests with transgenic animals, which carry many copies of the transgene, lacI, lacZ or gpt of E. coli, as reporter gene may fill the gap. As the reporter gene is transmitted through the germ cells, it is present in all cells and allows the detection and quantification of mutants and the sequencing of mutations in all kinds of somatic or germ cells.

Recently, two validation studies on the predictive value of the *in vivo* gene mutation test with transgenic animals for carcinogenicity, using a database of 105 compounds with carcinogenicity data, were published[20] [21]. The concordance with rodent carcinogenicity was 77% which is higher than for classical *in vivo* genotoxicity tests. The sensitivity and specificity values suggest that

[19] Burlinson B, Tice RR, Speit G, Agurell E, Brendler-Schwaab S, Collins AR, Escobar P, Honma M, Kumaravel TS, Nakajima M, Sasaki YF, Thybaud V, Uno Y, Vasquez M, Hartmann A. 2007. Fourth International Workshop on Genotoxicity testing: results of the *in vivo* Comet assay workgroup. *Mutation Res* 627:31-35.
[20] Lambert IB, Singer TM, Boucher SE, Douglas GR. 2005. Detailed review of transgenic rodent mutation assays. *Mutation Research* 590:1-280.
[21] IPCS. International Programme on Chemical Safety. 2006. Environment Health Criteria 233. *Transgenic Animal Mutagenicity Assays*. WHO, Geneva, Switzerland.

most of the carcinogens were positive and many of the non-carcinogens were negative in the *in vivo* gene mutation test with transgenic animals. Positive predictivity was comparable to other *in vivo* genotoxicity tests; however, the negative predictivity was low, still leading to an excess false positive rate. This suggests that a positive result in the *in vivo* gene mutation test with transgenic animals is indicative for carcinogenicity, but a negative result is not necessarily indicative of non-carcinogenicity (high false positive rate). Limited data are available to evaluate the results in known target tissues for carcinogenicity.

For legislation purposes a properly standardised (e.g. OECD) guideline is essential. This guideline should, among others, give clear directives with regard to the duration of treatment, the sampling time, the number and level of test concentrations as well as the route of exposure. Next, it should be allowed to decide for each test compound on a case by case basis which tissues or cell types need to be evaluated taking into account all known toxicological data.

Experimental data now indicate that transgene mutations do not give a selective advantage or disadvantage to the *in vivo* gene mutation test compared to other genotoxicity tests (e.g. *in vivo* micronucleus test) or even general toxicity tests (e.g. sub-acute 28-day toxicology test). This lack of variance may give added statistical power which should allow a reduction in the number of animals used.

Dr. van Benthem concluded that the results obtained until now indicate that the *in vivo* gene mutation tests with transgenic mice are suitable for mechanistic and fundamental studies and can be employed for legislative studies.

3.2 Breakout Groups

Following the plenary lectures, five topics were discussed in breakout groups on the second day. The briefing papers distributed to the participants in the breakout groups prior to the Workshop are given in Appendix 1.

Report from breakout group I: When are *in vitro* genotoxicity data sufficient for the assessment and classification of compounds?

1. *To what extent will limiting genotoxicity testing to* in vitro *studies result in an increase in the number of chemicals classified as positive/false positive? There are a number of reviews, both published and 'in press' evaluating false positives which may allow us to identify those assays that should no longer be used and/or can we recommend protocol changes to eliminate the potential for false positives?*

    - It was considered that better characterisation of the cell lines used is required and that this might be achieved by karyotyping and/or fingerprinting. Genotyping might also be used for this purpose in the future. There should also be general awareness of the importance of good housekeeping and the possibility of media variation effects. It is essential to confirm the identity and genetic stability, with time, of cells used for testing purposes.
    - The use of primary cells was discussed and the majority were in favour of their use in principle (human lymphocytes mainly). Problems though exist in the assessment of cytotoxicity which would hinder their use.
    - Although there was criticism of some of the established rodent cell lines currently and commonly used, it was considered that this is not the time to generate a new panel of cells.
    - Reducing the current maximum dose should be considered as a way to eliminate some false positive results induced by excessive toxicity.
    - It is also essential to improve the estimation cytotoxicity and it was recommended that for this purpose studies were required to specifically improve the current assessment of cytotoxicity.

2. *Can we estimate the number of chemicals we will incorrectly label as 'non mutagenic' if studies are limited to* in vitro *assays alone? Can we increase the range of our* in vitro *metabolic activation preparations to more effectively model the* in vivo *situation? Should we encourage the development and use of genetically engineered cell lines?*

    - The development of model systems for testing cosmetics e.g. 3D skin models has been described. These require further assessment and validation.

- Mechanistic models are required for ADME.
- It was recommended that a literature survey should be conducted to establish exactly what organ models are being developed for *in vitro* testing and to investigate the status of stem cells or others to develop specific organ models for specific investigations.

3. *Are there particular product groups (e.g. cosmetics) and chemical types for which we consider* in vitro *testing only will be inadequate? It will be important during the meeting for us to define adequate sets.*

- This question had also been addressed to some extent under question 2 (see above). The recommendation was to investigate current status of stem cell/organ models and their potential for use in genotoxicity testing *in vitro* with a priority for mammary and 3D liver culture systems -lattice structures.

4. *Are there particular endpoints (e.g. aneuploidy induction) for which there are significant advantages in undertaking* in vivo *work?*

- At the present time there is no particular advantage to undertake *in vivo* work for the known aneugens. However, the database is relatively small, and as knowledge advances this question may need to be revisited.

5. *How effective are the current range of models for studying potential organ specific effects, for example skin models? Can we indicate priorities for new model development?*

- A survey of this subject would be desirable. But it was also remarked that ECVAM is performing this currently.
- It is not clear at the present time how well they mimic the *in vivo* situation.

6. *To what extent can we model* in vitro *the cancer process for example with cell transformation assays (current status)?*

- These assays detect both genotoxic and non-genotoxic compounds but better methods are required to differentiate between genotoxic and non-genotoxic compounds. New genomic methods may also offer a solution for this problem.
- Various cell types are used e.g. SHE, BalB and other half-transformed cells. However, the ultimate test is tumour formation in an animal.
- There is uncertainty about the endpoint being measured in morphologically transformed cells. There are imaging systems to detect these cells and some groups are looking for markers of cell transformation.

- It might be possible to develop a suitable human epithelial cell system + telomerase, to immortalise the cells and use these for investigations of transformation (tumour formation in nude mice for validation not routine). This possibility should be investigated.
- The group raised the question whether a central repository of cell types should be established to ensure that cells verified for identity are being used.

7. *Can we model germ cell effects* in vitro*, for example using amphibian larvae and/or superovulated oocytes?*

- In the case of oocytes it is possible, although it was questioned to what extent would it be useful. It might be of value in an integrated reprotoxicity situation. However, the ability to perform this specialised technology is not widespread. This is not a major priority at this stage.
- The comet assay may be a good method of assessing genotoxicity of sperm cells.

8. *To what extent will we need to modify* in vitro *testing to adequately evaluate mixtures?*

- Special studies will be required in cases where information on LOEL and NOELs are known.
- Complex mixtures, such as hair dyes, will be a challenge requiring a good characterisation of individual components.
- Jim Parry presented data on this topic. The conclusion he drew from the experimental work that he described is:
  - In the case of different mechanisms of action, effects below LOEL are not additive. Within the safe NOEL exposure levels, mixtures will have no greater effects than single compounds.
  - In the case of the same mechanism, there are simple additive effects if you reach active concentrations. Combinations will need to be tested to determine how they add up to reach the combined LOEL.

9. *How will the lack of* in vivo *data influence our ability to provide the public with scientifically secure risk estimates?*

- The group felt comfortable with negative *in vitro* data.
- In the case of positives in existing tests, risk estimation is challenging without using *in vivo* data.
- SARs etc. might be used for similar chemical structures. This is of less use though for positives which may be due to impurities.

10. *To what extent can the use of molecular modelling and computerised structure / activity studies aid our predictions of potential genotoxicity?*

    - This will be entirely dependent upon database size.

Report from breakout group II: When are *in vivo* methodologies necessary and how can they be made more animal efficient?[22]

The questions below were discussed.

1. *When are* in vivo *methodologies necessary?*

    - Using current methodologies, there is general agreement that *in vivo* assays are required for genotoxicity assessment, especially given the known high rate of positive *in vitro* mammalian cell genotoxicity assays. Where the *in vitro* assays are clearly negative or clearly positive, *in vivo* follow-up is not absolutely recommended. In the case of clear positive *in vitro* tests and additional positive toxicity endpoints, the result of an *in vivo* genotoxicity assay would probably not greatly impact the risk management. On the other hand, *in vivo* systems are considered as the most integrated systems that can only be covered by several different *in vitro* systems trying to mimic the complex *in vivo* situation. In the case of unclear positive *in vitro* results (unexplained MOA or questionable relevance) *in vivo* assays are recommended, if feasible, in tissues relevant for human exposure and with an appropriate mechanistic basis.
    - A positive result in somatic cells of any tissue is an alert both for a possible carcinogenic activity and for possible mutagenic activity in germ cells (unless it can be proven that germ cells are not exposed). There may be a need to apply more relevant *in vivo* systems (target tissues) with higher sensitivity. An additional *in vivo* test should be performed only if likely to provide useful information.

2. *Which endpoints / assays are presently available for assessment of genotoxicity in different tissues? Where are the gaps?*

    - The three types of mutations, *i.e.* gene mutations, structural chromosome aberrations and numerical chromosome aberrations can be detected *in vivo* by the current techniques. In addition, DNA damage can also be assessed. Since we are facing specificity and not sensitivity problems with some of the current methodologies, one could argue that no

[22] Several arguments made in the report of this breakout group were not supported by some members of the group, in particular the opinion that *in vivo* testing is not absolutely recommended in case of clearly positive *in vitro* assays. This point was raised because it is a regulatory requirement, e.g. in REACh. But it is to be noted that the subject of this workshop had a broader scope.

additional endpoints are required. For some of the suggested additional endpoints (i.e. recombination and amplification) no *in vitro* models are available, so their *in vivo* need is questionable.

- The breakout group agreed that at present, tests for adequate endpoints are available but not in all tissues. In the future, other *in vitro* / *in vivo* endpoints such as amplification and recombination could be looked at, in a tiered approach and with a mechanism-based driver (i.e. disease process).
- At present, there is only limited overlap between the compounds tested in transgenic models for gene mutations and in the comet assay. The three assays which can be performed in virtually all tissues are comet assay, transgenic mutation assay and DNA adducts, but further investigation is required to estimate how well these endpoints predict the induction of the different types of mutation.

3. *Which route of exposures should be recommended (depending on human exposure?) and according to the route, which tissues should be examined in the same animal?*

- Inhalation, dermal and oral exposure are the most common routes. The breakout group agreed that in case the anticipated human route of exposure is known and it is feasible to perform, then this route should be applied. For industrial chemicals, often different routes are possible. In that case, the recommended route should be oral. The tissue(s) of interest should be the most relevant one(s) in terms of site of contact, high exposure levels and site(s) of metabolic activation. An additional requirement is that the particular route is practical and feasible to perform and that the methods are scientifically sound.

4. *Has the* in vivo *bone marrow still to be considered as the decisive* in vivo *tissue to assess hazard in somatic cells? If not, which other tissues do you recommend to study?*

- In many cases, yes, unless there is indication of non-relevance, e.g. in case the target tissue is not sufficiently exposed. Moreover, historical data and experience is available. If not, tissues relevant for human exposure should be used, if feasible.

5. *Which rodent species should be studied? And do we need both sexes?*

- One sex is mostly sufficient. The male rat should be considered as the default, unless there are reasons for studying another species/sex (*e.g.* metabolic differences, mechanistic reasons).

6. *Can assays on germ cells be performed directly on the same animal as the assay in somatic cells or should one first proceed through assessment on somatic cells to select the appropriate assay in germ cells?*

   - The induction of mutagenic events in any organ or tissue of an animal comprises an alert for both possible carcinogenic activities and for possible induction of inherited genetic damage. It is technically not possible to combine germ cell and somatic cell genotoxicity assessment (for micronuclei and chromosome aberrations). Therefore, a tiered approach (first somatic, second germ cells) appears logic. Only the in case of a positive *in vivo* somatic cell genotoxicity test, a test in germ cells should be considered. However, if there is evidence that germ cells are not exposed, there is no need to perform a test in germ cells. Nevertheless, in general, a test on germ cells is not required, unless for classification purposes.

7. *Which practical experimental modification of design and techniques might help reducing the use of animals?*

   - Study different tissues per animal, study different endpoints per tissue (if feasible) and integrate genotoxicity tests into standard toxicity studies.

8. *Should we consider both chronic and acute exposure in animals depending on the type of compound and its intended use?*

   - In some cases, a ten-fold difference in dose leads only to a two-fold increase in C_{max} or AUC. This phenomenon is typically compound dependent and influenced by e.g. bioaccumulation and activation/detoxification. In addition, the purpose of the test may also determine the type of exposure; i.e. for hazard identification, an acute exposure appears adequate, but for risk assessment, a repeated (chronic) exposure could be more relevant, although there might also be an increased risk of secondary effects. The breakout group agreed that it should be considered on a case by case basis, depending on the compound. For example, in case it is demonstrated that exposure of the target tissue is higher after repeated dosing, this treatment regimen should be followed.

Report from breakout group III: How important is human variation in terms of assessing risk and hazard of chemical exposure to the human population? What are the variables to be considered in testing strategies?

The following is the summary of the discussions in this breakout group.

General:

- The human genome consists of 3.2 billion of DNA base pairs, of which approx. 0.1% is considered to show genetic variability (corresponding to approx. 3,000,000 variable bp per genome).
- The majority of genetic variability is due to single nucleotide polymorphisms (SNPs) which occur in the human genome in a frequency of 1/300 to 1/10000 base pairs. In most cases, it is not known whether a variant (even in coding or regulatory sequences) has phenotypic consequences. Each variant in one of the multiple susceptibility genes apparently may only have a minor effect on the phenotype. The genes involved (e.g. in metabolic activation and DNA repair) interact in large molecular networks and only the net effect of allelic variation in multiple interacting genes determines the phenotypic response of a cell/organism to chemical exposure. Furthermore, variants conferring susceptibility are not sufficient to cause cancer alone but modulate the risk in combination with other alleles, carcinogen exposure and other environmental factors (e. g. nutrition). These factors render the overall patterns of response to carcinogen exposure very complex.

Summary of main discussion points and conclusions:

- The group did not anticipate an impact from the presently known polymorphisms on the outcome of current genotoxicity tests. However, vigilance needs to be maintained as new variants are being continuously detected. No known examples could be identified, where an absence/presence of polymorphisms would qualitatively change the outcome of hazard identification, making it unlikely that we miss a genotoxic hazard of a compound. Thus, at present, polymorphisms do not need to be included into current basic genotoxicity testing.
- As long as we rely on a qualitative evaluation of a genotoxic potential, there is, at present, no need to consider polymorphisms in the genotoxicity testing strategy. If however in the future we want to assess potency/quantitative effects and/or new test systems (primary human cells, organ models), the influence of polymorphisms might become more significant. With respect to the future use of new test systems, their experimental variability is considered as a main limiting factor (see research needs).
- Generally, analysis of single (individual) polymorphisms is restricted to and meaningful only for those genes that exhibit clear single-gene effects. If polymorphisms are to be included in quality controlled and validated new test systems, combinations of several polymorphisms known to influence cancer susceptibility (from literature data and/or mathematical models) are clearly favoured.

Recommendations of the breakout group for future research needs:

- Identification and elimination of main factors contributing to the experimental variability of primary human cells and organ models to be used as new model systems in genotoxicity testing.
- Identification of pathways, genes and/or cellular phenotypes (metabolism, cell cycle regulation, DNA repair) that are relevant and need to be controlled in order to ensure the validity of a new test system.
- Generation of stable cells from human primary cells that have increased genetic stability for *in vivo* phenotype (with respect to metabolism, cell cycle and DNA repair).
- There is still a need to maintain recognition of polymorphisms that may have relevant phenotypic effects and to define their influence on mutagenicity and carcinogenicity.

Report from breakout group IV: How are difficult (non-standard) materials tested and what would be the biological significance of reactive oxygen species in terms of assessing genotoxicity of such materials?

The following questions were discussed.

1. *Can certain categories of chemicals and/or materials be concluded unsuitable for* in vitro *testing and the testing requirement waived? Are there examples of this?*

    - Unusual physical and chemical properties adherent with certain categories of chemicals and/or materials can cause them to be considered unsuitable for *in vitro* testing in mammalian cells. Examples of this are ultra-fine particles and insoluble polymers. The group raised the question whether testing requirements could be waived when the *in vitro* delivery of the chemical is questionable and does not mimic the *in vivo* delivery.
    - According to existing testing guidelines, precipitation of a material does not preclude testing. It is not, however, logical to test precipitating compounds since the real dose concentration or the treatment/exposure time cannot be defined. There are also data from one study indicating that cell monolayers can engulf particles and precipitates, leading to a false positive genotoxic response based on the internalisation rather than the compound. The group considered it worthwhile to revisit these and similar issues, and possible to have a waiver based on whether the compound can get into the cell.
    - No specific future research needs were identified.

2. *Is specific guidance needed for handling and testing of difficult materials, including modifications of the test systems and interpretation of the results, in addition to the directives provided by current testing guidelines?*

- The group concluded that improved guidance is needed concerning the testing of 'difficult to test compounds'. One of the challenges is to ensure that the delivery to the test system is accurate and biologically meaningful. It may be possible, through further research, to develop guidance for specific classes of compounds (i.e. for specific classes of compounds such as insoluble precipitates). For instance, further research may deliver a standard disaggregation protocol for 'clumped' particles.

3. *What degree of reliability can be assigned to results from* in vitro *tests with materials that are difficult to formulate as dosing solutions?*

- The group considered that the results with 'difficult to test materials' will be test system dependent and as reliable/unreliable as for fully testable compounds. A central question to ask is whether exposure to the target cell is achieved or not. This is particularly important for producing evidence of exposure for an observed negative genotoxic effect. The group's recommendation is to adhere to the current testing guidelines.
- No specific future research needs were identified.

4. *What is a reasonable requirement for testing of materials that require modified external metabolic activation systems?*

- The group did not consider any recommendations with respect to this point that would be specific to 'difficult to test compounds'. The problems regarding modified metabolic activation systems are the same for all compounds; for 'difficult to test' as well as for 'readily testable' compounds.
- The group considered it desirable to do further work into introducing Phase II co-factors into the external metabolic systems. Furthermore, the group considered it necessary to conduct further testing with a compound if a human metabolite is identified.
- The use of human S9 was discussed. For the first time, the use of human S9 appears as recommended in an OECD draft guideline. If still included in the adopted guideline, it will set precedence for other *in vitro* genotoxicity testing. Although human S9 is already commercially available several diverse questions need to be answered before reaching consensus regarding the routine use of it. For example, is there any demographic specificity associated with the source of origin, and would a positive result with rat S9 be overturned by a negative result with human S9?

5. *When positive* in vitro *results are generated by compounds with genotoxic intermediate metabolites, such as formaldehyde or peroxide, should some guidance for data interpretation and regulatory acceptance be given to alleviate the need for follow-up animal testing and supplementary tests?*

- The group considered there being a difference between the approach regarding known and novel or unknown intermediates. For substances with known daily human exposure and known *in vitro* only positives, there should be sufficient data to establish a risk without additional (repetitive) *in vitro* testing. However, an animal test might be called for in order to show evidence for no genotoxic *in vivo* responses. It may also be recommendable to test other known metabolites *in vitro*. However, when an agent is obviously genotoxic through a well established genotoxic intermediate/metabolite, there could/should be no further need for follow-up animal testing and supplementary tests. For novel/unknowns intermediates, follow-up is required.
- No specific future research needs were identified.

6. *How should intracellular generation of highly reactive metabolic products with oxidative DNA damaging activity, such as ROS or other species, be measured and considered in the interpretation of genotoxicity data?*

- Further research is definitely needed into the extra-cellular, cell media-based production of ROS. The group considered it important to identify/develop an ideal culture medium (for the ideal mammalian cell) that does not lead to adventitious, non-relevant ROS formation, and that mimics the *in vivo* environment.
- Since extra-cellular production of ROS can happen *in vivo* and can be of biological significance we may not wish to identify/develop a medium that will negate this totally; we need to identify/develop a medium that best mimics the *in vivo* situation.
- Key questions are: What are the features/circumstances of extra-cellular ROS production? Must key features be redox-active metal-ion levels?
- The best measurement of this would be via the measurement of a specific oxidative DNA lesion; the most pertinent/measured lesion being 8-oxo-dG. The best ways to measure this are the analytical methods of LC-MS or HPLC-ECD, but one has to take precautions to avoid adventitious (isolation/protocol-derived) oxidative damage.
- Alternatively, use of the FPG/hOGG-comet assay has many advantages over LC-MS/MS or HPLC-ECD; however, there are some questions as to assay specificity. Alternative lesions include thymine glycols, M1dG and SSBs. Each method and the lesion measured have advantages and disadvantages.
- It is important to identify/delineate the false oxidative effects from real oxidative effects plus the agent's intrinsic genotoxicity (i.e. its DNA adduct forming ability).

7. *Are there other induced or endogenous cytotoxicity mechanisms that may interfere with testing and data interpretation of certain materials? Are there other biological mechanisms with possible causative effects (such as inflammation) that are not captured?*

- As current regulatory test systems may not pick up these 'other biological mechanisms', further research work is needed to understand cytoxicity resulting from them. The ongoing development of a more robust transformation assay (i.e. the Syrian Hamster Embryonic [SHE] assay) may encompass these mechanisms in the near future.
- There might be possible causative effects (such as inflammation) that are not captured in *in vitro* testing. Since these functions and their possible relationship to genotoxicity are not defined, current 'regulatory' test systems may not pick them up. Further work would be needed to understand cytoxicity resulting from these 'other biological mechanisms'.
- There might also be other features that 'deliver' a (false) positive result or that confound the test. For example, an agent may promote DNA repair (a benefit) giving strand breaks that are detectable (by the comet assay) and are 'negative'.

8. *How can the challenges of safety assessment be met for new technology, such as nanomaterials or electromagnetic fields (mobile communication technology)?*

- One of the challenges posted by materials created by new technology is to ensure that the delivery to the test system best mirrors the *in vivo* exposure route. Further research may be needed to deliver superior guidance as to the 'handling' of these 'materials' in their regulatory testing.
- The group did not reach consensus on this question since there is yet not enough substance for forming an agreement.
- No specific future research needs were identified.

9. *Can the assessment of difficult-to-test materials be aided by using methodology such as (Q)SAR, modelling or comprehensive databases?*

- Presently most of the existing databases do not have sufficient information pertaining to difficult-to-test compounds. In the future, if there is a consistent prediction from numerous/multiple/independent packages, this may provide sufficient weight of evidence to overrule a positive *in vitro* test. It was felt that we are still a long way from this point.

Report from breakout group V: What progress has been made in developing new and relevant models and how can human data be incorporated into assessments?

In considering this question, group discussions focussed on the progress that has been made in human biomonitoring and the role of humanised *in vitro* and *in vivo* systems. The following issues were identified.

1. *Biomarker validation*.
   - With the development of molecular epidemiology comes the prospect of using intermediate biomarkers as a means of assessing human exposure to carcinogens and potentially evaluating cancer risk. However, whilst a variety of assessment and biomonitoring tools exist there are relatively few fully validated predictive biomarkers of risk (e.g chromosomal aberrations, micronuclei). A greater number of exposure biomarkers are currently available but these are for specific chemical exposures and need to be validated on an individual basis. Continued effort is therefore needed in the area of biomarker development and validation.

2. *Biomarker persistence*.
   - Persistence varies depending on the specific biomarker being measured and it is important to establish the lifetime in order to obtain meaningful data. Current evidence indicates all known biomarkers (e.g. protein or DNA adducts) are relatively short lived. Therefore, studies should be designed on the basis of this assumption, unless there are data to the contrary.

3. *Exposure data*.
   - High quality exposure data are important for risk assessments. However, it is appreciated that it can be difficult to find good control groups as other lifestyle and environmental factors differing between populations from different areas may confound the effects due to occupational or environmental chemical exposures. Furthermore, exposure needs to be properly evaluated by appropriate analytical methods, which can become costly.

4. *New methods in molecular genetics and bioinformatics*.
   - The value of new methods such as proteomics and genomics in the identification of biomarkers for safety evaluation and exposure assessment was considered. It was felt that application of these techniques, especially proteomics, is still at relatively early stage and although they have enormous potential they must be supported by high quality studies in animal models and *in vitro* human systems, in order to gain the necessary mechanistic information required to define whether the biomarker is truly indicative of exposure or risk.

5. *Humanised* in vitro *systems*.
   - Humanised *in vitro* systems such as primary cells or the use of human liver S9 in bacterial mutation assays can provide valuable mechanistic information and assist in hazard identification. However, unless pooled samples or large panels are used they introduce the element of individual variation, which complicates risk assessment for the general population. Conversely, such systems allow the effects of gene polymorphisms to be studied and offer the possibility of identifying high risk groups.
   - Concern was raised that substituting human for rat S9 in the Ames test may produce different results and it was considered worthwhile investing time collating the available data comparing human and rat systems for the metabolic activation of compounds. The availability and variability of pooled human S9 should also be assessed.

6. *Humanised* in vivo *systems*.
   - For example knock-in mice expressing human p53 or xenobiotic metabolism genes, can be extremely useful on a case-by-case basis but care must be taken in extrapolating data to humans where exposure is to mixtures rather than defined single agents.
   - It was recognised that the use of humanised systems for hazard identification and mechanistic studies are two different situations and any data generated must be evaluated at several levels; in terms of its value in screening, providing mechanistic understanding and the possibility of extrapolating to humans.
   - It was proposed that a review be conducted to assess the value of humanised *in vitro* and *in vivo* systems, to identify the methods currently available and examine whether they can be used for extrapolation.

4. SUMMARY AND RECOMMENDATIONS

The workshop concluded that the current *in vitro* strategies are relatively effective at detecting genotoxic carcinogens. However, their power is at the expense of a number of false positives which is not without consequences. Further, while the current methods can detect genotoxic chemicals, they are relatively ineffective at detecting non-genotoxic carcinogens. The workshop discussed various new methods including modified cells and human specific activation systems. While there was a conclusion that genetically modified cells were much preferred in testing strategies, many of these are not validated at the present time and there was reluctance in the interim to abandon those currently in use. There was also recognition that there are potential problems in the maintenance of cell lines with genetic instability. It is essential to continuously monitor cell lines for genetic variance and contamination. This could be done by establishing a central database of DNA fingerprints and analysis of copy number variation answered by genomic analysis. These data could be held in a central database aligned to a repository of cells from which validated cell cultures could be obtained. There was little discussion of the value of humanised mice in the testing strategy but these also may offer a way forward and further investigation is required.

New genomic technologies were felt to have much to offer and have already proven their worth in terms of differentiating genotoxic from non-genotoxic carcinogens. Further evaluation of these methods and ongoing vigilance of the research in this area is required. The application of other genomic technologies such as analysis of gene copy number variance also has potential in genotoxicity testing and its application needs to be further evaluated.

The role of population polymorphisms, particularly in respect of QTLs (quantitative trait loci) in conveying susceptibility to carcinogens is poorly understood at the present time. However, this is also a complex area in which to initiate research. Given the rapid advances in genetics, continued monitoring of research in the area is required to recognise when polymorphisms are identified that may have an impact on carcinogenicity testing. A literature based study to review the field should be considered; there is one report of an association of CYP3A4 phenotype and susceptibility to drug genotoxicity in man[23]. Such polymorphisms may define susceptible sub-populations in man which depending on their prevalence may need to be taken into account during the testing process.

The rapid evolution of methods for assessing DNA adducts in the absence of biological applicability has highlighted how advances in technology can potentially influence risk perception. The need to further understand the relationship between DNA adduct biomarkers and biological recurrence was agreed to be of importance.

[23] Lopes L, Piccoli F, Paixão V, Latorre M, Camargo B, Simpson A, Caballero O. 2004. Association of CYP3A4 genotype with detection of Vγ/Jβ *trans*-rearrangements in the peripheral blood leukocytes of pediatric cancer patients undergoing chemotherapy for ALL. *Leukemia Research* 28(12):1281-1286.

The use and role of metabolic activation systems in carcinogenicity testing is a controversial area. There is a need to ensure that activation systems are relevant and are also balanced in test systems by appropriate deactivation mechanisms modelling the human situation. The use of both S9 fractions from human and genetically modified cells offers new potential advances. For both though the question of validation needs to be addressed. The stability of cells needs to be ensured, and variance between human S9 fractions further assessed. Additionally, there is an obvious *a priori* need to understand how the chemical is metabolised in humans so the most appropriate systems are chosen.

There was an agreement that cytotoxicity can be a major confounding factor in the testing strategy and some of the methods for assessing cytotoxicity are not appropriate. There is a need for more work in this area to standardise results and to ensure that cytotoxicity is not a confounding factor in the testing process.

Overall the workshop addressed some difficult questions successfully. On some there was not an overall consensus and a majority opinion is reflected. There were some specific recommendations and these are listed below.

Desk /Literature Research

1. Comparative study of *in vitro* cytotoxicity methods, identification of most relevant method(s) and new recommendations on top-dose setting in mammalian cell assays.

2. Survey of available methods for *in vitro* organ specific models to identify research priorities. More specifically, identification of the critical factors i.e. metabolism, repair capacity, p53 status etc. necessary for the development of new test systems based upon the use of human cells. Ideally, this would include the potential influence of polymorphisms.

3. Development of recommendations for the 'good house keeping' of cell cultures. Storage of identity data in a database to include genomic stability and a repository of validated cells.

4. Evaluation of the available methods to test 'difficult' compounds, particularly those that are insoluble in aqueous media.

5. Comparative review of available information on exogenous human and rodent metabolic activation systems with an emphasis upon the identification of effectiveness of the preparations in hazard and risk estimations.

6. Identification and evaluation of all humanised *in vitro* and *in vivo* systems currently available. This would include a review of the appropriateness of humanised mouse models.

7. A review of the current genomic and bioinformatic technologies particularly for the recognition of carcinogens. A literature, and potentially laboratory based project, to asses the values of array-based comparative genomic hybridisation for both the validation of the genetic stability of cell lines and in the testing process.

8. An evaluation of the hierarchy of *in vivo* and *in vitro* models for testing, their roles and position in the process.

Laboratory Research

1. Development of new *in vitro* model(s) of cancer induction and progression. The general consensus was that this should be based upon human cells with a transformation-based endpoint, although for specific industries (e.g. cosmetics) the focus should be made on the development of new chemical carcinogenesis biomarkers for specific tissues, e.g. skin.

2. Assessment of the effects of culture conditions and genetic background on the production and influence of ROS (reactive oxygen species) on *in vitro* systems. Selection or development of media which eliminate ROS-derived artefacts and more effectively model the *in vivo* situation.

3. Evaluation of new technology biomarkers in genotoxicity testing and the development of more innovative risk assessment and weight of evidence approaches in the interpretation of genotoxicity data and chemical exposure.

ABBREVIATIONS

ADME	Absorption, distribution, metabolism, excretion
arrayCGH	Array comparative genome hybridisation
AUC	Area under curve
BalB	'Mouse cell line'
bp	Base pairs
C_{max}	Maximum concentration
Cefic LRI	European Chemical Industry Council Long-range Research Initiative
CHL	'Chinese hamster cell line'
CHO	'Chinese hamster cell line'
CNV	Copy number variation
CYP	Cytochrome
3D	Three dimensional
DNA	Deoxyribonucleic acid
E.coli	*Escherichia coli*
ECVAM	European Centre for the Validation of Alternative Methods
ESTR loci analysis	Expanded single tandem repeat loci analysis
FPG/hOGG	Formamidopyrimidine DNA glycosylase / human 8-oxoguanine DNA glycosylase
gpt	Guanine phosphoribosyl transferase
HepaRG	'Human cell line'
HepG2	'Human cell line'
HPLC-ECD	High performance liquid chromatography-electrochemical detection
HPV program	High production volume program
lacI	'Bacterial gene'
lacZ	'Bacterial gene'
LC-MS	Liquid chromatography-mass spectrometry
L5178Y	'Mouse cell line'
LOEL	Lowest observed effect level
M1dG	'Type of DNA lesion'
MOA	Mode of action
MS	Mass spectrometry
NOEL	No observed effect level
8-oxo-dG	8-Oxo-7,8-dihydro-2'-deoxyguanosine
OECD	Organisation for Economic Co-operation and Development

p53	'Tumour suppressor gene'
QSAR	Quantitative structural activity relationship
QTL	Quantitative trait loci
3Rs	Reduction, refinement and replacement
REACh	Registration, evaluation, authorisation and restriction of chemicals
RNA	Ribonucleic acid
ROS	Reactive oxygen species
SSB	Single strand break in DNA
S9	Fraction of liver enzymes used in metabolic preparations
SAR	Structure activity relationship
SHE	Syrian hamster embryonic cell line
SHE assay	Syrian hamster embryonic cell line assay
SNP	Single nucleotide polymorphisms
TK6	'Human cell line'
US EPA	United States Environmental Protection Agency
US FDA	United States Food and Drug Administration
V79	'Chinese hamster cell line'

APPENDIX 1: BRIEFING PAPERS FOR BREAKOUT GROUPS

Briefing Paper for Breakout Group I

Topic: When are *in vitro* genotoxicity data sufficient for the assessment and classification of compounds?

Chair: Jim Parry
Rapporteur: Elizabeth Parry
Raffaella Corvi
Bernd Herbold
Mike Kenrick
Martin Kuster
Claire Moore
Stefan Pfuhler
Andrew Smith

We now have available over 20 years of experience in the development and application of *in vitro* methodologies which can be used to detect and evaluate chemicals capable of inducing genetic changes and which may lead to the production of both somatic and germ cell changes. Thus far, we have been able to further evaluate both negative and positive *in vitro* results in intact animals and base our estimates of hazard and risk upon a combination of both *in vitro* and *in vivo* data sets. It has generally been considered that the confirmation of an *in vitro* result in an appropriate *in vivo* assay ensures the relevance of our genotoxicity results to the human population.

For a range of products such as cosmetics the ability to undertake *in vivo* studies will shortly end. These changes in the methods that can be used in the assessment of compounds for potential genotoxicity suggests that this is an appropriate time to consider both how current *in vitro* methods may need to be modified and how we can incorporate developing technologies to provide scientifically justified conclusions. This is also an appropriate time to consider to what extent we can reduce our dependence on *in vivo* testing for more chemical types.

The currently available *in vitro* methods are capable of detecting the ability of chemicals to induce gene and chromosomal (both structural and numerical) mutations and may potentially be used to detect other endpoints such as methylation changes which are related to the induction of cancer and inherited diseases. In this workshop group we will aim to evaluate the current status of *in vitro* genotoxicity testing and provide recommendations on a number of key areas related to both the current and future testing of chemicals. To undertake this exercise we request the members of the working group to undertake some literature search and assessments prior to the

meeting and, thus, facilitate our aim of producing a basic outline of our evaluations and suggestions during our time in Malta.

We ask that the members of this breakout group allocate themselves to the topics listed below and volunteer to undertake pre-meeting work in the suggested areas. The areas indicated are not definitive and we encourage group participants to suggest further topics for inclusion.

Our suggestions for the breakout topics are as follows:

i. To what extent will limiting genotoxicity testing to *in vitro* studies result in an increase in the number of chemicals classified as positive/false positive? There are a number of reviews, both published and 'in press' evaluating false positives which may allow us to identify those assays that should no longer be used and/or can we recommended protocol changes to eliminate the potential for false positives?

ii. Can we estimate the number of chemicals we will incorrectly label as 'non mutagenic' if studies are limited to *in vitro* assays alone? Can we increase the range of our *in vitro* metabolic activation preparations to more effectively model the *in vivo* situation? Should we encourage the development and use of genetically engineered cell lines?

iii. Are there particular product groups (e.g. cosmetics) and chemical types for which we consider *in vitro* testing only will be inadequate? It will be important during the meeting for us to define adequate sets.

iv. Are there particular endpoints (e.g. aneuploidy induction) for which there are significant advantages in undertaking *in vivo* work?

v. How effective are the current range of models for studying potential organ specific effects, for example skin models? Can we indicate priorities for new model development?

vi. To what extent can we model *in vitro* the cancer process for example with cell transformation assays (current status)?

vii. Can we model germ cell effects *in vitro,* for example using amphibian larvae and/or superovulated oocytes?

viii. To what extent will we need to modify *in vitro* testing to adequately evaluate mixtures?

ix. How will the lack of *in vivo* data influence our ability to provide the public with scientifically secure risk estimates?

x. To what extent can the use of molecular modelling and computerised structure / activity studies aid our predictions of potential genotoxicity (good examples if anyone has them)?

Briefing Paper for Breakout Group II

Topic: When are *in vivo* methodologies necessary and how can they be made more animal efficient?

Chair: Micheline Kirsch-Volders
Rapporteur: Marlies De Boeck
Norbert Banduhn
Neil Carmichael
Håkan Cederberg
Peter Kasper
Mike O'Donovan
David Owen
Véronique Thybaud

Despite the major progress made for the validation and the predictivity of *in vitro* systems, *in vivo* tests often remain the final level to decide about the potential mutagenic/carcinogenic hazard of a compound.

In vivo tests in rodents play a pivotal role in the assessment of genotoxicity and carcinogenicity.

- They integrate the balance between protective and stressing effects.

- They include metabolism activation/detoxification of promutagens / carcinogens.

- They permit comparison of different routes of exposure, according to the type of hazard studied (inhalation, dermal, *per os*, intraperitoneal).

- They potentially allow the analysis in different cell types, tissues and somatic versus germ cells.

- They offer the possibility to focus on the target cells of given cancer types.

- They allow the assessment of systemic effects (e.g. blood).

- *In vitro* genotoxicity tests are often considered purely for hazard identification purpose, while *in vivo* tests contribute, at least to some extent, also to risk assessment given the 'real life' situation.

However, difficulties for the extrapolation to the human *in vivo* situation make interpretation sometimes difficult.

- Metabolic activation can differ very significantly between mouse, rat and human (e.g. butadiene).

- Some critical organs for human cancer (e.g. prostate) are absent in mouse and rat.

- Hormone dependent and other non-genotoxic mechanisms of carcinogenesis show major discrepancies between rodents and humans (e.g. breast cancers).

- Functional activity of enzymes presenting genetic polymorphisms in DNA repair, metabolism, cell cycle checkpoints can differ significantly.

The present questions related to *in vivo* methodologies are:

- When are they necessary?
- How can they be made more animal efficient?

We ask that the members of this breakout group allocate themselves to the topics listed below and volunteer to undertake pre-meeting work in the suggested areas. The areas indicated are not definitive and we encourage group participants to suggest further topics for inclusion.

Our suggestions for the breakout topics are as follows:

i. Which endpoints / assays are presently available for assessment of genotoxicity in different tissues? Where are the gaps?

ii. Which route of exposures should be recommended (depending on human exposure?) and according to the route, which tissues should be examined in the same animal?

iii. Has the *in vivo* bone marrow still to be considered as the decisive *in vivo* tissue to assess hazard in somatic cells? If not, which other tissues do you recommend to study?

iv. Which rodent species should be studied? And do we need both sexes?

v. Can assays on germ cells be performed directly on the same animal as the assay in somatic cells or should one first proceed through assessment on somatic cells to select the appropriate assay in germ cells?

vi. Which practical experimental modification of design and techniques might help reducing the use of animals?

vii. Should we consider both chronic and acute exposure in animals depending on the type of compound and its intended use?

Briefing Paper for Breakout Group III

Topic: How important is human variation in terms of assessing risk and hazard of chemical exposure to the human population? What are the variables to be considered in testing strategies?

Chair: Günter Speit
Rapporteur: Andreas Rothfuss
Birte Dreeβen
Peter Farmer
Tim Gant
Soterios Kyrtopoulos
Franz Oesch
Dirk Schwartz

Numerous studies have reported associations between genetic variation (e. g. polymorphisms of xenobiotic-metabolising genes or DNA repair genes) and the occurrence of various cancers. However, the potential carcinogenic exposure behind such findings is usually not well-defined. Genetic polymorphisms are also frequently considered in human biomonitoring studies for genotoxic / mutagenic effects. Such studies have the advantage that the exposure and the endpoint studied are nearer in time and better information might be obtained on susceptibility to specific mutagens / carcinogens. Associations between mutagen sensitivity and specific genotypes have actually been found in some cases, although inconsistent results were published for the majority of exposures. It is generally believed that the consideration of genetic polymorphisms as modulators of genotoxicity and cancer will improve hazard characterisation and risk assessment.

The pathways leading to mutations and cancer are very complex and hundreds of genes are involved. It is well known that (rare) mutations in genes responsible for genome integrity (DNA

repair genes and tumour suppressor genes) lead to clearly increased individual cancer risks. However, there is also considerable variability in the human genome and hundreds of polymorphisms have been identified in genes involved in the metabolism of xenobiotics and in DNA repair. The majority of genetic variability is due to single nucleotide polymorphisms (SNPs) which occur in the human genome in a frequency of 1/300 to 1/10000 base pairs. In most cases, it is not known whether a variant (even in coding or regulatory sequences) has phenotypic consequences. Each variant in one of the multiple susceptibility genes apparently only has a minor effect on the phenotype. The genes involved interact in large molecular networks and only the net effect of allelic variation in multiple interacting genes determines the phenotypic response of a cell / organism to chemical exposure. Furthermore, variants conferring susceptibility are not sufficient to cause cancer alone but modulate the risk in combination with other alleles, carcinogen exposure and other environmental factors (e. g. nutrition). Therefore, the predictive value of genetic variants to estimate individual cancer risks will be restricted to those genes that exhibit single-gene effects. The successful application of genetic polymorphisms in gene-environment studies will be small unless established environmental risk factors are tested on proven candidate genes.

To better characterise such genes and possible interactions, new experimental strategies have to be developed. Besides allele-disease association studies, functional studies of allelic variants are needed. *In vitro* and *in vivo* systems capable of tracking different gene activities may elucidate the possible role of allelic variants in mutagen sensitivity and certain cancers. Genetically engineered cell lines and animal models (‘humanised mice’) are useful tools to evaluate the environmental relevance of selected polymorphisms and to establish (more or less complex) gene variant – environment interactions. This information may then be used to appropriately design studies for molecular epidemiology and might lead to modifications of test systems and current strategies for genotoxicity testing.

Our suggestions for the breakout topics are as follows:

i. What are generally accepted examples of genetic polymorphisms that influence mutagenicity / carcinogenicity and how are they currently considered in hazard / risk assessment?

ii. Do we have appropriate strategies to identify relevant (i.e. functional) polymorphisms and how should they be considered in risk assessment?

iii. Do we have appropriate strategies to identify gene-environment interactions and how effective are *in vitro* methods?

iv. Do we need specific organ models to appropriately consider organ-specific differences in metabolism and DNA repair?

v. How can we measure the impact of genetic polymorphisms on the outcome of genotoxicity tests *in vitro* and *in vivo*?

vi. Will our knowledge on genetic polymorphisms lead to modified testing strategies in the near future?

vii. Do we have mathematical models to estimate the contribution of several polymorphisms on the outcome of mutagenicity / carcinogenicity studies?

Briefing Paper for Breakout Group IV

Topic: How are difficult (non-standard) materials tested and what would be the biological significance of reactive oxygen species in terms of assessing genotoxicity of such materials?

Chair: Maria Donner
Rapporteur: Don Jones
Christa Hennes
David Kirkland
Douglas McGregor
Thomas Petry
Markus Schulz
Wera Teubner

Existing regulatory testing guidelines and international harmonisation rely extensively on *in vitro* tests for genotoxicity and mutagenicity. Product stewardship initiatives (i.e. the US EPA High Production Volume (HPV) program, the 7th Amendment to the European Cosmetics Directive, REACh) in connection with the principles for Reduction, Refinement, and Replacement of animals (3Rs) contribute to an increasing requirement for the scientific community to base weight-of-evidence and mode of action reasoning on *in vitro* tests. Both hazard identification and gap analyses for human risk assessment are initially attempted without *in vivo* verification of *in vitro* results. The picture is further complicated, not the least in the chemical industry, by the difficult-to-test materials and intracellular metabolic considerations.

The unusual physical and chemical properties of some test materials can cause practical and test-system-specific difficulties, and compromise the outcome of the test (i.e. give false negative or positive results). Conventionally used external metabolic activation systems are not optimised for biotransformation of certain chemicals. Alternatively the material may be metabolised *in vitro* to mutagenic intermediates without *in vivo* effects or relevance for humans. It is known that

intracellular endogenous metabolic pathways may create reactive molecules, such as reactive oxygen species (ROS) that lead to DNA damage that can be expressed as mutations and genetic alterations. Finally, innovative products of new technology may not be compatible with testing in established tests and testing schemes.

Our suggestions for the breakout topics are as follows:

i. Can certain categories of chemicals and/or materials be concluded unsuitable for *in vitro* testing and the testing requirement waived? Are there examples of this?

ii. Is specific guidance needed for handling and testing of difficult materials, including modifications of the test systems and interpretation of the results, in addition to the directives provided by current testing guidelines?

iii. What degree of reliability can be assigned to results from *in vitro* tests with materials that are difficult to formulate as dosing solutions? Examples include: insoluble materials that require excessive manipulations and still precipitate when applied to aqueous testing conditions; materials that chemically change the properties of the culture media, or are non-removable after the completed treatment period; and materials that due to physical properties (high molecular weight, fibres, mixtures…) will remain intact in the media surrounding the cells, never reaching the target cells.

iv. What is a reasonable requirement for testing of materials that require modified external metabolic activation systems?

v. When positive *in vitro* results are generated by compounds with genotoxic intermediate metabolites, such as formaldehyde or peroxide, should some guidance for data interpretation and regulatory acceptance be given to alleviate the need for follow-up animal testing and supplementary tests?

vi. How should intracellular generation of highly reactive metabolic products with oxidative DNA damaging activity, such as ROS or other species, be measured and considered in the interpretation of genotoxicity data?

vii. Are there other induced or endogenous cytotoxicity mechanisms that may interfere with testing and data interpretation of certain materials? Are there other biological mechanisms with possible causative effects (such as inflammation) that are not captured?

viii. How can the challenges of safety assessment be met for a new technology, such as nanomaterials or electromagnetic fields (mobile communication technology)?

ix. Can the assessment of difficult-to-test-materials be aided by using methodology such as (Q)SAR, modelling or comprehensive databases?

Briefing Paper for Breakout Group V

Topic: What progress has been made in developing new and relevant models and how can human data be incorporated into assessments?

Chair: David Phillips
Rapporteur: Karen Brown
Paul Carmichael
Barry Elliott
Radim Šrám
Jan Van Benthem
Joost Van Delft
Paul White

With the development of molecular epidemiology has come the prospect of using intermediate biomarkers as a means of assessing human exposure to carcinogens and even as a means of assessing cancer risk. Current issues are concerned with the extent to which various biomarkers are validated and have predictive value. Confidence in extrapolating from *in vitro* and non-human *in vivo* systems to humans is the goal, requiring consideration of where matters currently stand and where they might usefully be advanced in the future.

This breakout group will consider the progress that has been made in human biomonitoring and the extent to which data from *in vitro* and non-human *in vivo* systems has, or has not, been borne out in scenarios of known human exposure. Participants are asked to come up with examples of where this has, or has not, been demonstrated, and how newer methodologies might be applied to the process of human risk assessment.

Our suggestions for the breakout topics are:

i. What can be and has been learned from known human exposures to genotoxins?

ii. What are the useful / validated biomarkers of exposure and/or risk? What are the limitations?

iii. What is known about the timescale / persistence of observed effects?

iv. How important are good exposure data?

v. How might human or humanised *in vitro* systems assist in hazard identification and mechanistic understanding?

vi. Can humanised *in vivo* models (e.g. rodents) increase our confidence in extrapolating animal data to humans?

vii. What is the potential of new methods in molecular genetics and bioinformatics for safety evaluation and exposure assessment? What can 'omics' offer? How should such methods be used?

APPENDIX 2: FURTHER INFORMATION ON NEW TECHNOLOGIES

Genotoxicity has for many years relied on a few core methods, detection of DNA breaks, chromosomal rearrangements, unscheduled DNA synthesis, micronucleus formation and gene mutations leading to a change in phenotype such as gain or loss of resistance. For some molecules, such as tamoxifen, DNA adducts have been detected using techniques like ^{32}P-postlabelling. All of these systems detect substantial changes in DNA absolutely none detect epigenetic changes now recognised as key regulators of inheritance, and possibly of transgeneration genetic effects. Furthermore, there has always been a question over the *in vitro* activation systems which have relied in the main on S9 activation fractions for metabolic conversions.

The advent of high throughput molecular techniques that have been developed over the last two decades has given rise to profound new opportunities to take genotoxicity testing into a new era. There are two major areas in which the new technologies can be applied, in detection of genotoxicity and in the provision of appropriate cell systems.

Cell systems

How appropriate are the cell systems that are currently being used for genotoxicity testing? Cells kept in culture are often genetically unstable, are easily contaminated and very susceptible to environmental changes such as altered media or confluence. Most have limited metabolic activation systems, or have differing abilities to repair DNA damage. The use of transcriptomic technologies giving a whole cell view using data that is highly multidimensional in nature indicates cell differences and responses in a manner not been previously possible. Primary cells in particular undergo profound changes when placed into culture[24]. As such, choosing a time point after establishment in culture for use of primary cells as a test system can be very problematic. There are three viable alternatives for *in vitro* systems which need to be explored as sources of biological material for the testing of chemicals. First, genetically altered cells; these cells will totally mimic a true *in vivo* system. They have the advantage that the human forms of the metabolic enzymes can be stably expressed. With knowledge of the metabolism of the chemical, an appropriate cell system can potentially be chosen with the requisite metabolic profile. Second, stem cells which are little explored to date but offer potential for use as differentiated derivatives, and, finally, the use of activation systems constructed from the membranes of bacteria which over-express the appropriate enzymes. As with genetically altered cells, some prior knowledge of the metabolism of the chemical is required to choose the appropriate system.

[24] Gant TW and Zhang SD. 2005. In pursuit of effective toxicogenomics. *Mutat Res* 575:4-16.

Characterisation

All cells used for genotoxicity testing need to be genomically characterised. Cells in culture are intrinsically genetically unstable and can also be easily contaminated. It is therefore essential that cell lines are characterised by minisatellite profiling, karyotyping and arrayCGH (array comparative genome hybridisation) for amplifications and deletions compared with a standard control. This is absolutely essential to ensure that quality control is maintained in cells during their subsequent use and maintenance in culture.

Cells in culture

Profound differences can occur in cells with density in culture. Apart from the differences in gene expression and sensitivity, which can occur when the cells are either dividing or confluent, a recent study has shown that confluent cells can develop genetic instability[25]. Therefore, cells may alter both their responsiveness and genetic stability depending on how they are maintained.

Compared with an inbred animal strain, cell lines are a difficult biological system to work with that may show high variability. If cells *in vitro* are going to be used, care needs to be exercised in their characterisation and culture.

Methods

There are many new methods on the horizon which offer the possibility of not just detection of genotoxicity, but of increased understanding. These are covered in brief below. All have developed from the cancer field where their application has had a profound impact on the understanding of genomic instability in tumours. Detection of genetic instability in tumours has many parallels with genotoxicity testing. Little application of these new methods has been made in genotoxicity testing. It is now essential that some studies are commissioned to study the applicability of these methods, else there is the very real possibility that application will occur in an inappropriate manner to which an industry unprepared will not be in a position to respond.

This is particularly the case for the assessment of epigenetic change because in genotoxicity testing procedures there is currently no assessment made of the influence of chemicals on epigenetic status. It is absolutely essential that this deficiency is addressed.

The following lists the methods that should now be assessed with appropriate cell lines and chemicals to ascertain their relevance to genotoxicity testing:

[25] Polyzos A, Parfett C, Healy C, Douglas G, Yauk C. 2006. A single-molecule PCR approach to the measurement of induced expanded simple tandem repeat instability *in vitro*. *Mutat Res* 594 93-100.

1. Methylation status to determine epigenetic effects. This can be done with bisulfite sequencing, methylcytosine immunoprecipitation with microrarray analysis or methylation sensitive restriction enzyme analysis. Changes in methylation are potentially associated with transgenerational genotoxicity as well as with non-genotoxic carcinogenicity making this a vital area to be explored in the context of genotoxicity testing.
2. Expanded single tandem repeat (ESTR) loci analysis. This can be carried out easily and cheaply on single molecules. This makes it viable for analysis of sperm to assess inherited genome instabilities or for single cells from a culture. The method looks for changes in the length of ESTR regions which are subject to mutation during homologous recombination. A change in the length of these regions with chemical exposure could indicate a genotoxic effect.
3. ArrayCGH. Genome copy number variation (CNV) is a feature of life and a recent study has indicated just how much variation there is in normal populations[26]. Changes in CNV are a feature of genotoxicity. ArrayCGH works by detecting these changes via labelling of the DNA and then hybridisation to genomic sequence probes which are immobilised on a microarray. There is an absolute requirement to conduct some studies to assess the applicability of this technique in genotoxicity testing procedure with relevance to current protocols.
4. Transcriptomics. The RNA profile of a cell has been shown to be a good marker for both cell phenotypes and for the action of chemicals. It has been employed in genotoxicity testing to differentiate between genotoxins and non-genotoxins[27] [28]. As this technique has shown potential there is the need to conduct further studies with an aim to develop a gene expression signature indicative of genotoxic agents. This is also, with the possible exception of epigenetic methods, the only way to recognize potential non-genotoxic carcinogens. There is a fundamental need to build further datasets on these important chemicals.
5. Derived from knowledge gained from transcriptomics with regard to gene activation during DNA damage there is the possibility of using constructs of the promoter of a responsive gene couple with a reporter gene such as that coding for green fluorescent protein. These constructs can be transfected stably into cells. The cells will then respond by activation of the reporter to genotoxins and fluoresce green. This approach has been reported in one paper[29].

[26] Redon R, Ishikawa S, Fitch KR, Feuk L, George H, Perry GH, Daniel Andrews T, Fiegler H, Shapero MH, Carson AR, Chen W, Kyung Cho E, Dallaire S, Freeman JL, González JR, Gratacòs M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Stephen W. Scherer SW, Hurles ME. 2006. Global variartion in copy number in the human genome. *Nature* 444:444-454.

[27] van Delft JHM, van Agen E, van Breden SGJ, Herwijnen MH, Staal YCM, Kleinjans JCS. 2005. Comparison of supervised clustering methods to discriminate genotoxic from non-genotoxic carcinogens by gene expression profiling. *Mutat Res* 575:17-33.

[28] Ellinger-Ziegelbauer H, Stuart B, Wahle B, Bomann W, Ahr HJ. 2005. Comparison of the expression profiles induced by genotoxic and nongenotoxic carcinogens in rat liver. *Mutat Res* 575:61-84.

[29] Hastwell PW, Chai LLC, Roberts KJ, Webster TW, Harvey JS. 2006. High-specificity and high-sensitivity genotoxicity assessment in a human cell line: Validation of the GreenScreen HC *GADD45a-GFP* genotoxicity assay. *Mutat Res* 607:160-175.

6. Francis Collins who was responsible for the sequencing of the whole genome has recently stated that the $1000 genome will be with us by 2010[30]. Even if this prediction is not fulfilled there is no doubt that sequencing costs are falling. There is then the very real possibility of detecting SNPs and other genomics changes in regions of DNA including epigenetic analysis by sequencing methods. These data would also require an analysis with regard to whether SNPs observed after exposure to a chemical were in those regions of DNA where they may be harmful, for example oncogenes. This and all the other above techniques dictate that more effort is going to have to be expended in the analysis and interpretation of data. This analysis is an order of magnitude more complex than counting reversions or micronuclei but has the potential to be far more informative.

There is an understandable concern that application of these techniques will be too sensitive and lead to a greater false positive rate that is currently observed. However this is not necessarily the case. Most current assays are narrow in terms of the target they offer the chemical, for example reversion assays which depend on one gene. Alternatively, they detect gross effects such as aneuploidy or micronuclei. These endpoints give no indication of mechanism or relevance to carcinogenicity of the regions of DNA affected. The genomics methods are broad and generate a lot of information. Correct analysis or genomic data, coupled with appropriate cells and knowledge of metabolism would enable a mechanistically based assessment to be made of genotoxic potential that should lead to a much smaller false positive rate without a compromise in the detection sensitivity.

[30] Collins F. Presentation at the American Association for Cancer Research (AACR) meeting, Los Angeles, USA, April 2007.

APPENDIX 3: WORKSHOP PROGRAMME

Monday 23 April 2007

13.00-14.00	Registration and sandwich lunch	
14.00-14.15	**Welcome and Introduction**	Dr. Neil Carmichael, ECETOC

BACKGROUND TO THE TOPIC OF THE WORKSHOP

14.15-14.45	**What brought us to the Current State of Mutagenicity / Genotoxicity Testing**	Prof. Jim Parry University of Wales Swansea
14.45-15.15	**Problems of False Positives and False Negatives in Current *in vitro* Testing**	Dr. David Kirkland Covance
15.15-15.45	Coffee Break	

NEW METHODOLOGY IN MUTAGENICITY / GENOTOXICITY TESTING

15.45-16.15	**Potential Application of Toxicogenomics to Mutagenicity / Genotoxicity Testing**	Dr. Tim Gant University of Leicester
16.15-16.45	**Metabolism and Genetic Susceptibility in *in vitro* Testing for Mutagenicity / Genotoxicity**	Prof. Franz Oesch University of Mainz
16.45-17.15	**Measurement of DNA Damage and its Relevance to Mutagenicity / Genotoxicity Testing**	Prof. Günter Speit University of Ulm
17.15-17.45	**Relevance of Transgenic Animals to Mutagenicity / Genotoxicity Testing**	Dr. Jan van Benthem RIVM
17.45-18.00	**Organisation of the following day**	Chair: Prof. Jim Parry
18.15-18.30	**Information on Workshop Follow-up [Speakers, Breakout Group Chairs and Rapporteurs, and Organising Committee only]**	Dr. Christa Hennes, ECETOC
19.00-22.00	Dinner	

Tuesday 24 April 2007

08.30-11.30 **Breakout Groups:**

I When are *in vitro* data sufficient for the assessment and classification of compounds?

Chair: **Prof. Jim Parry, University of Wales Swansea**
Rapporteur: **Dr. Elizabeth Parry, University of Wales Swansea**

II When are *in vivo* methodologies necessary and how can they be made more animal efficient?

Chair: **Prof. Micheline Kirsch-Volders, Vrije Universiteit Brussels**
Rapporteur: **Dr. Marlies de Boeck, Johnson & Johnson Pharmaceutical**

III How important is human variation in terms of assessing risk and hazard of chemical exposure to the human population? What are the variables to be considered in testing strategies?

Chair: **Prof. Günter Speit, University of Ulm**
Rapporteur: **Dr. Andreas Rothfuss, Schering**

IV How are difficult (non-standard) materials tested and what would be the biological significance of reactive oxygen species in terms of assessing genotoxicity of such materials?

Chair: **Dr. Maria Donner, DuPont Haskell Laboratory**
Rapporteur: **Dr. Don Jones, University of Leicester**

V What progress has been made in developing new and relevant models and how can human data be incorporated into assessments?

Chair: **Prof. David Phillips, Institute for Cancer Research, UK**
Rapporteur: **Dr. Karen Brown, University of Leicester**

All groups will also address the following question:
How does mutagenicity testing fit into an overall chemical safety testing strategy?

(coffee in breakout rooms)

11.00-12.30 Buffet Lunch

Report of Breakout Groups and Plenary Discussion:

12.30-12.45 Breakout Group I
12.45-13.00 Breakout Group II
13.00-13.15 Breakout Group III
13.00-13.30 Breakout Group IV
13.30-13.45 Breakout Group V

13.45-14.00 **Conclusions of Workshop** Chair: Prof. Jim Parry

Close of Workshop

APPENDIX 4: LIST OF PARTICIPANTS

Name	*E-mail*	*Affiliation*
N. Banduhn	norbert.banduhn@henkel.com	Henkel KGaA, Germany
K. Brown	kb20@le.ac.uk	Leicester University, UK
N. Carmichael	neil.carmichael@ecetoc.org	ECETOC, Belgium
P. Carmichael	paul.carmichael@unilever.com	Unilever SEAC, UK
H. Cederberg	hakan.cederberg@kemi.se	KemI (Swedish Chemicals Agency), Sweden
R. Corvi	raffaella.corvi@jrc.it	JRC, ECVAM - IHPC, EU
M. De Boeck	mdboeck@prdbe.jnj.com	Johnson & Johnson Pharmaceutical R&D, Belg.
M. Donner	maria.donner@usa.dupont.com	DuPont Haskell Laboratory, USA
B. Dreeβen	birte.dreessen@de.sasol.com	Sasol Germany GmbH, Germany
B. Elliott	barry.elliott@syngenta.com	Syngenta, UK
P. Farmer	pbf1@le.ac.uk	Leicester University, UK
T. Gant	twg1@le.ac.uk	Leicester University, UK
C. Hennes	christa.hennes@ecetoc.org	ECETOC, Belgium
B. Herbold	bernd.herbold@bayerhealthcare.com	Bayer HealthCare, Germany
G.D. Jones	gdj2@le.ac.uk	Leicester University, UK
P. Kasper	p.kasper@bfarm.de	BfArM (Fed. Inst. for Drugs & Med. Devices), Germany
M. Kenrick	mike.kenrick@ge.com	GE HealthCare, UK
D. Kirkland	david.kirkland@covance.com	Covance Laboratories Limited, UK
M. Kirsch-Volders	mkirschv@vub.ac.be	Vrije Universiteit Brussel, Belgium
M. Kuster	martin.kuster@novartis.com	Novartis International AG, Switzerland
S. Kyrtopoulos	skyrt@eie.gr	National Hellenic Research Foundation, Greece
D. McGregor	mcgregortec@btinternet.com	Toxicity Evaluation Consultants, UK
C. Moore	claire.moore@unilever.com	Unilever SEAC, UK
M. O'Donovan	mike.odonovan@astrazeneca.com	AstraZeneca R&D, UK
F. Oesch	oesch@uni-mainz.de	University of Mainz, Germany
D. Owen	david.owen@shell.com	Shell Chemicals, UK
E. Parry	e.m.parry@swansea.ac.uk	University of Wales Swansea, UK
J. Parry	jmp@swansea.ac.uk	University of Wales Swansea, UK
T. Petry	thpe@weinberggroup.com	The Weinberg Group LLC, Belgium
S. Pfuhler	pfuhler.s@pg.com	Cosmital SA (P&G), Switzerland
D. Phillips	david.phillips@icr.ac.uk	Institute of Cancer Research, UK
A. Rothfuss	andreas.rothfuss@schering.de	Schering, Germany
M. Schulz	markus.schulz@basf.com	BASF, Germany
D. Schwartz	dirk.schwartz@nycomed.com	Nycomed GmbH, Germany
A. Smith	andrew.smith@hse.gsi.gov.uk	Health and Safety Executive, UK
G. Speit	guenter.speit@uni-ulm.de	University of Ulm, Germany
R. Šrám	sram@biomed.cas.cz	Institute of Experimental Medicine AS CR, Czech Republic
W. Teubner	wera.teubner@cibasc.com	Ciba Specialty Chemicals, Switzerland
V. Thybaud	veronique.thybaud@sanofi-aventis.com	Sanofi Aventis, France
J. van Benthem	j.van.benthem@rivm.nl	RIVM (Dutch National Institute of Public Health and the Environment), The Netherlands
J. van Delft	j.vandelft@grat.unimaas.nl	Maastricht University, The Netherlands
P. White	paul_white@hc-sc.gc.ca	Health Canada, Dept. of Biology, Canada

APPENDIX 5: ORGANISING COMMITTEE

Professor Jim Parry
University of Wales Swansea
Medical School
Swansea
SA2 8BG
United Kingdom

Dr. Paul Carmichael
Unilever SEAC
Sharnbrook
MK44 1LQ
United Kingdom

Dr. Maria Donner
DuPont Haskell Laboratory
Newark
DE 19714
USA

Dr. Tim Gant
University of Leicester
Medical Research Council
Leicester
LE1 9HN
United Kingdom

Dr. Christa Hennes
ECETOC
B-1160 Brussels
Belgium

Dr. David Owen
Shell Chemicals
London
SE1 7NA
United Kingdom

ECETOC WORKSHOP REPORTS

No.	Title
No. 1	Workshop on Availability, Interpretation and Use of Environmental Monitoring Data. 20-21 March 2003, Brussels
No. 2	Strategy Report on Challenges, Opportunities and Research Needs Arising from the Definition, Assessment and Management of Ecological Quality Status as Required by the EU Water Framework Directive Based on the Workshop EQS and WFD versus PNEC and REACh - Are They Doing the Job ? 27-28 November 2003, Budapest
No. 3	Workshop on Use of Human Data in Risk Assessment. 23-24 February 2004, Cardiff
No. 4	Influence of Maternal Toxicity in Studies on Developmental Toxicity. 2 March 2004, Berlin
No. 5	Workshop on Alternative Testing Approaches in Environmental Risk Assessment. 7-9 July 2004, Crécy-la-Chapelle
No. 6	Workshop on Chemical Pollution, Respiratory Allergy and Asthma. 16-17 June 2005, Leuven
No. 7	Workshop on Testing Strategies to Establish the Safety of Nanomaterials. 7-8 November 2005, Barcelona
No. 8	Workshop on Societal Aspects of Nanotechnology, 9 November 2005, Barcelona