味をする必要があるという点である。すなわち，そこに貢献する遺伝子のレパートリーをケース・バイ・ケースで検討する必要性があるのである。なぜなら非常にマイナーなコンポーネントの遺伝子発現がトキシコロジーの上で重要な意味を持つことがあり得ないとは言えないからである。

さらに，この時，それら個々のデータの持つ特性を理解することは重要である。データの中には，例数を増加させると限りなく数値が収束する性質のもの（図22右図の右）と，例数を増加させることによって一定の範囲内での明瞭な分散を示すものとがある（図22右図の左）。両者の値の相互関係の取り方は異なっており，後者のようなケースで器械的に平均値をとるとCV値の大きな値を得る結果となり，実態から離れるので，適用を誤ると違った結論を導き出すことになる。いろいろな実例があるが，因みに発生とか成長過程のパラメーターの多くは前者のデータをとる傾向があるのに対して，老化過程におけるパラメーターは，後者のような集束のないデータをとることが知られている。

6．トキシコロジーを寿命学的な見地から見る

最後に，4章でもふれた，寿命を尺度にしたトキシコロジーについて考えてみたい。

6.1 生体の異物応答と寿命

生体の外来異物に対する応答反応は，生活年齢の一時期の可逆性変化期を含めて，多くの物質で寿命全体に関わる変化を引き起こす蓋然性があることが分かってきた。このカテゴリーで放射線の晩発障害機構研究の歴史は古いが，化学物質を含む生体異物一般のそうした長期の寿命に及ぼす影響は，エピジェネティック効果と呼ばれ，近年，急速に統一概念が明らかになってきつつあり，実験的環境も整い始めた。先に紹介したGompertz表現による寿命関数が，腫瘍性疾患，非腫瘍性疾患の如何を問わず一元的に寿命に反映する性質を持っていることも相俟って，生体の異物応答と老年疾患との相互関係を視野に入れた人々の関心が亢まっている。

ヒトでの生活年齢に伴うテロメアの短縮が知られているのと同様に，骨髄移植のような強制増殖負荷を与えることにより，長く修復力の強いマウスのテロメアでさえも短縮傾向が検出可能なことが分かってきた。併せて，Gompertz表現（3.3 2)の項および図12参照）による寿命カーブに対して，外来異物がその傾きを急峻化させるという仮説的な命題も，当筆者らの長年の研究によって大きく進展してきた[49, 50]。ひと頃取りざたされたタバコ1本の喫煙が与えるリスクが何日分の寿命短縮に相当するかといった例をあげるまでもなく，トキシコロジーを“寿命に対するリスク学”として考えることにはいまや実質的な意味があり，それが本稿の終章でこれに言及しておこうと思う所以である。

6.2 放射線曝露における異物応答と寿命

いくつかの例を紹介しよう。図23は，種々の線量の放射線曝露をうけた野生型（Wで表記した右の3本：0, 3, 5Gy）とp53ヘテロ欠失マウス（Hで表記した左の4本：0, 1, 3, 5Gy）における死率曲線を比較して示している。照射線量の増加にともなって，ゴンパーツ曲線（Gompertz Curve）の傾きが急峻化し，平均寿命[※19]が短縮していく様子が分かるであろう。p53遺伝子の欠失型では，放射線照射に伴って発がん性頻度が高くなる。これによってハイリスクとなる関係が，野生型よりも急峻で寿命の短縮に傾くことがこの結果から見て取ることができる。

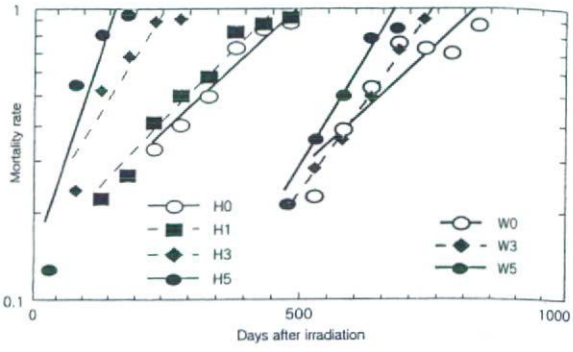

図23　放射線曝露による線量依存性の寿命短縮とゴンパーツ曲線の変化

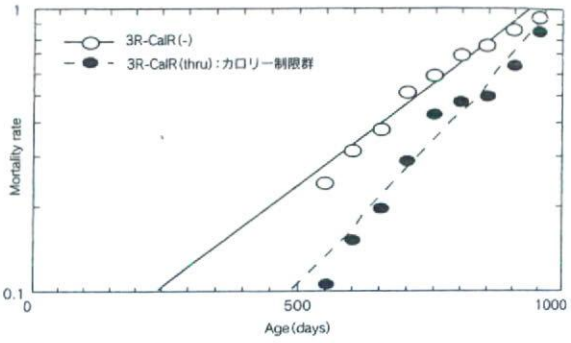

図24　放射線曝露による寿命短縮とカロリー制限によるその予防，およびゴンパーツ曲線の変化

こうして放射線照射によって傾きが急峻化する条件下でも，いわゆる発がんの予防・抑制効果で知られるカロリー制限のような処置を取ることによって，Gompertz Curveは改善（prevention）の方向に移動する。図24で示したように，3Gy照射のカロリー制限群の50%死亡寿命は15%あまりも延長し，白血病に対する予防効果も見られる（データ省略）が，最長寿命の短縮に対しての回復は見られなかった。

6.3 ベンゼン曝露における異物応答と寿命

ベンゼン曝露の場合も，図25のような，非曝露群に対して曝露群での傾きの急峻化と平均寿命の顕著な短縮が観察される。もちろんベンゼンは，単一の腫瘍や非腫瘍病変を引き起こすわけではない。Gompertz関数の法則的に面

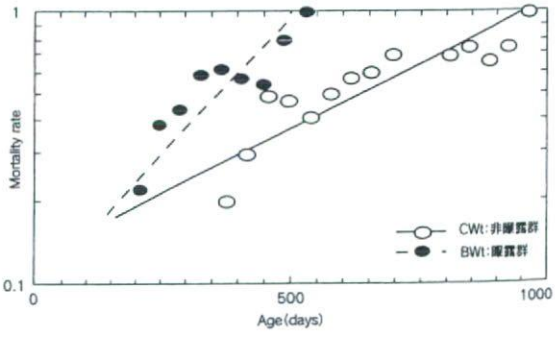

図25　ベンゼン曝露による寿命短縮とゴンパーツ曲線の変化

※19 Mortality rate 0.5における平均生存日数。

白い点は，各種の腫瘍病変のみならず非腫瘍病変全体がひとつの関数として動くことであり，それらの病変の発症にリンクするゲノムの不安定性の亢進といった共通因子の関与を示唆しているのであるが，現時点での知見は，現象論的事実にとどまる。図25に戻ってさらに興味深いことは，これに対して酸化的ストレス（ROS）の吸収を促進するチオレドキシン過剰発現マウスにベンゼンを曝露すると，図26のように平均寿命が大きく回復・延長するのみならず，最長寿命の延長傾向が見られる。

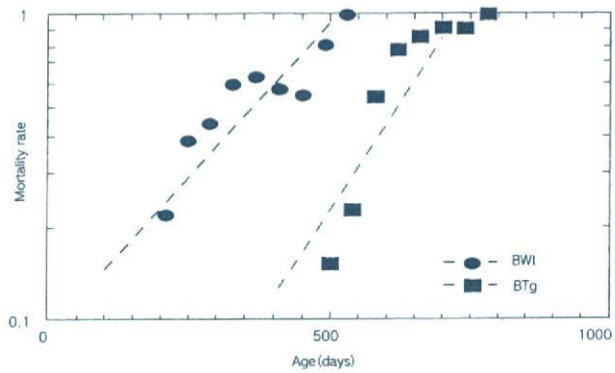

図26　ベンゼン曝露による寿命短縮とチオレドキシン過剰発現マウスにおけるベンゼン毒性の予防

6.4 生体の異物応答と種々の抗毒作用の寿命への影響

これらの毒性学的な寿命短縮要因とカロリー制限やチオレドキシン過剰発現による寿命延長（抗毒作用）の要因を図27に総合的にまとめた。ここに示すGompertz Curveは，実データ（real data）にもとづく寿命とトキシコロジーの相互関係である。これらの背景を構成する生体に対する異物の反応性を明らかにすることは，寿命に対する異物の要因の究極的なメカニズムを明らかにすることになるものと考えられる[50]。

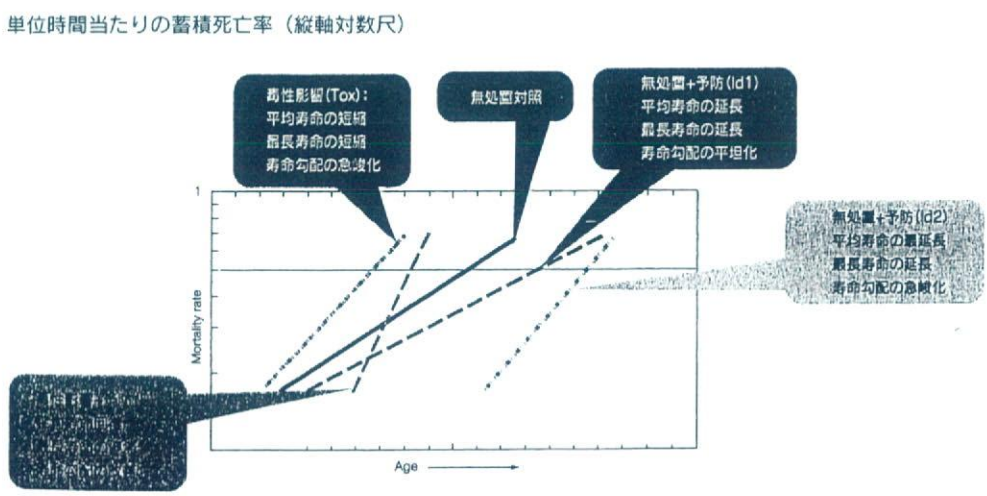

図27　寿命とトキシコロジー：まとめ

おわりに

本稿では，発生，成長，老化の三つに区切ってトキシコロジーの将来について筆者の考え方を述べた。それらのことがらは，いずれも，まだガイドラインの作られていない事柄である。ガイドラインのないこれらの事柄に注目するモチーフが何処にあったかというと，はじめに気づくことは，今回取り上げた周辺の事柄には分からないことが山積していて，現在進められている試験の解釈にしても，しばしばその辺のところを検討しないと，うまく説明がつかないということがあるということに起因している。そしてもっと大枠のところから考えてみると，私たちをとりまく社会一般のニーズがそれを求めるようになっている，ということではないかと思われる。

すなわち，少子高齢化社会は限りなく子供の安全に対する取り組みを求めており，WHOでも国連でも大きな問題になっているところである。思春期の問題も，胎生期に次ぐ第2の形態形成期にあって，不明の点がたくさんあり，この時期の過ごし方が一生を決めるところさえあるにも拘わらず，これまで，あまり顧みることがなかった。加えて，筆者の卒業後の研究生活は，まだ当時あたらしい分野であった老年学・老年病理学からはじまっている。いまや高齢者人口は，地域によっては25%にも及び，全国平均でも2割に達している状態にある。高齢者が自らの生活に対して，多くのものを云う人々になろうとしている。

私たちの携わる毒性学もそれらの変化に対応してゆく必要を求められている。もとよりキチンとした答えのある領域ではないが，本稿が議論のきっかけの一助とでもなれば，望外の喜びである。

謝辞

本稿は，去る7月4日（2006），名古屋国際会議場における日本トキシコロジー学会学術年会（堀井郁夫会長）にて行った教育講演の内容に加筆訂正を加えて書き下ろしたものである。講演で座長の労を取られた大野泰雄国立医薬品食品衛生研究所　副所長，及び原稿を起こすに当たって助言を戴いた安全性生物試験研究センター・平林容子博士に謝意を表す。

参考文献

1) 井上 達: 毒性学の現状と展望. 科学（岩波書店）, 74: 18-23（2004）
2) Lenz W: The History of Thalidomide, 1992. UNITHにおける講演録より。
3) 医薬品非臨床試験研究会編「医薬品非臨床試験ガイドライン」解説2002, 薬事日報社, pp. 287（2002）
4) Wilson JG. Environment and Birth Defects. New York, Academic Press, pp. 305（1973）
5) Chung F, et al: Thalidomide pharmacokinetics and metabolite formation in mice, rabbits, and multiple myeloma patients. Clin Cancer Res 10: 5949-5956（2004）
6) 井上 達, 野田哲生, 野本明男編「ヒト型モデル動物」, シュプリンガーフェアラーク東京社, pp. 218（2002）
7) レイチェル・カールソン著（青樹簗一訳）「沈黙の春」(Silent Spring), 新潮社（1962）

8) 井上 達：エンドクリン問題の最近の動向, 季刊環境研究 106: 24-35 (1997)
9) Faustman EM, et al: Risk Assessment. In: Eds. Casarett & Doull's. Toxicology - the basic science of poisons 6th ed. Curtis D. Klaassen, pp. 97 (1996)
10) Global Assessment of the State-of-the-Science of Endocrine Disruptors. Eds: Damstra T, et al: IPCS, International Programme of Chemical Safety, WHO, pp. 180 (1992)
11) 第3節 低用量問題 In: 厚生労働省医薬食品局化学物質安全対策室編「内分泌攪乱化学物質問題の現状と今後の取組—検討会中間報告追補その2」,pp. 75-85 (2004)
12) Gray LE Jr, et al: Environmental antiandrogens: Low doses of the fungicide vinclozolin alter sexual differentiation of the male rat. Toxicol Ind Health 15(1-2): 48-64 (1999)
13) Wolf CJ, et al: Effects of prenatal testosterone propionate on the sexual development of male and female rats: a dose-response study. Toxicol Sci 65: 71-86 (2002)
14) Ankley GT, et al: Effects of the androgenic growth promoter 17-beta-trenbolone on fecundity and reproductive endocrinology of the fathead minnow. Environ Toxicol Chem 22: 1350-1360 (2003)
15) 井上 達, 井口泰泉編:「生体統御システムと内分泌攪乱」, シュプリンガーフェアラーク東京, pp. 321 (2005)
16) Hirabayashi Y, et al: Evaluation of nonthreshold leukemogenic response to methylnitrosourea in p53-deficient C3H/He mice. J Toxicol Applied Pharmacol 190: 251-261 (2003)
17) Terracini B, et al: The roles of age at treatment and dose in carcinogenesis in C3Hf/Dp mice with a single administration of N-nitroso-N-methylurea. Br J Cancer 33: 427-439 (1976)
18) Gompertz B: On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Pholos Trans Royal Soc. (London) 115: 513-585 (1825)
19) Soffritti M, et al: First experimental demonstration of the multipotential carcinogenic effects of aspartame administered in the feed to Sprague-Dawley rats. Environ Health Perspect 114: 379-385 (2006)
20) Nuslein-Volhard C, et al: Mutations affecting segment number and polarity in Drosphila. Nature 287: 795-801 (1980)
21) C. elegans Sequencing Consortium. Genome sequence of the nematode C. elegans: A platform for investigating biology. Science 182: 2012-2018 (1998)
22) Mounkes LC, et al: A progeroid syndrome inmice is caused by detects in A-type lamins. Nature 423: 293-298 (2003)
23) Kuro-o, et al: klotho null. Nature 390: 45-51 (1997)
24) Abbott BD, et al: AH receptor, ARNT, glucocorticoid receptor, EGF receptor, EGF, TGF alpha, TGF beta 1, TGF beta 2, and TGF beta 3 expression in human embryonic palate, and effects of 2,3,7,8-tetrachlorobibenzo-p-dioxin (TCDD). Teratology 58: 30-43 (1998)
25) Romero P, et al: Computational prediction of human metabolic pathways from the complete human genome. Genome Biol, R2 (2005)
26) Brown NA, et al: Screening chemicals for reproductive toxicity: the current alternatives. ATLA 23: 865-882 (1995)
27) Sarkar P, et al: Activation of telomerase in BeWo cells by estrogen and 2,3,7,8-Tetrachlorodibenzo-p-dioxin in co-operation with c-Myc. Int J Oncol 28: 43-51 (2006)
28) Barrett JC, et al: Diethylstilbestrol induces neoplastic transformation without measurable gene mutation at two loci. Science 212: 1402-1404 (1981)
29) Huff J: Carcinogenicity of bisphenol-A in Fischer rats and B6C3F1 mice. Odontology 89: 12-20 (2001)
30) Wang Z, et al: Differentiation therapy for acute promyelocytic leukemia with all-trans retinoic acid: 10-yar experience of its clinical application. Chin Med J (Engl.) 112: 963-967 (1999)
31) Schindler T, et al: Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. Science 289: 1938-1942 (2000)
32) Hirabayashi Y, et al: Toxicogenomics Applied to Hematotoxicology. In: Ed by Juergen Borlak, Handbook of Toxicogenomics. Wiley-VCH Verlag GmbH. Weinheim, pp. 583-608 (2005)
33) Hirabayashi Y, et al: Protective role of connexin 32 in steady-state hematopoiesis, regeneration state and leukemogensis. Expt'l Biol Med, in press (2007)
34) Hirabayashi Y, et al: Benzene-induced hematopoetic toxicity transmitted by AhR in the wild-type mouse was nullified by repopulation of AhR deficient bone marrow cells: time after benzene treatment and a recovery. Chemosphere. in press (2006)
35) Ivanova, et al: A stem cell molecular signature. Science 298: 601-604 (2002)
36) Hirabayashi Y, et al: Serial transplantation of p53-deficient hemopoietic progenitor cells to assess their infinite growth potential. Expt' Biol Med (Maywood) 227: 474-479 (2002)
37) Hirabayashi Y, et al: Cell kinetics of hemopoietic colony-forming units in spleen (CFU-S) in young and old mice. Mech Ageing Dev 101: 221-231 (1998)
38) 平林容子,井上 達: 総説・新技術の毒性学への応用,造血幹細胞動態解析法—BUUV法. J Toxicol Sci 23: app. 55-61 (1998)

39) Yoon BI, et al: Mechanism of action of benzene toxicity: Cell cycle suppression in hemopoietic progenitor cells (CFU-GM). Exptle Hematol 29: 278-285 (2001b)
40) Inoue T: Toxicogenomics - a new paradigm of toxicology. In: Eds. T. Inoue, W.D. Pennie. Toxicogenomics. Springer-Verlag Tokyo, pp. 3-11 (2003)
41) Hu T, et al: Identification of a gene expression profile that discriminates indirect-acting genotoxins from direct-acting genotoxins. Mutation Res 549: 5-27 (2004)
42) Gary A: Churchill. Fundamentals of exprerimental design for cDNA microarray. Nature Genetics 32: 490-495 (2002)
43) Schadt EE, et al: Freature extraction and normalization algorithms for high-density oligonucleotibe gene expression array data. J Cell Biochem; Supple 37: 12-125 (2001)
44) Kroll TC, et al: Ranking: a closer look on globalization methods for normalization of gene expression arrays. Nucleic Acids Res 30: e50 (2002)
45) Stoyanova R, et al: Normalization of single-channel DNA array data by principal component analysis. Bioinformatics 20: 1772-1784 (2004)
46) Astrand M: Contrast normalization of oligonucleotide arrays. J Comput Biol 10: 95-102 (2003)
47) Hill AA, et al: Evaluation of normalization procedures for oligonucleotide array data based on spided cRNA controls. Genome Biology 2: 0055.1-13 (2001)
48) Bahar R, et al: Increased cell-to-cell variation in gene expression in ageing mouse heart. Nature 441: 1011-1014 (2006)
49) Trosko JE, et al: Oxidative stress, signal transduction, and intercellular communication in radiation carcinogenesis. Stem Cells 15 (Suppl. 2): 59-67 (1997)
50) Hirabayashi Y, et al: Implication of hemopoietic progenitor cell kinetics and leukemogenesis: Relevance to Gompertzean mortality as possible hematotoxicological endpoint. Expt'l Hematol. In press (2007)

Meeting Report: Validation of Toxicogenomics-Based Test Systems: ECVAM–ICCVAM/NICEATM Considerations for Regulatory Use

Raffaella Corvi,[1] Hans-Jürgen Ahr,[2] Silvio Albertini,[3] David H. Blakey,[4] Libero Clerici,[5] Sandra Coecke,[1] George R. Douglas,[4] Laura Gribaldo,[1] John P. Groten,[6] Bernd Haase,[7] Karen Hamernik,[8] Thomas Hartung,[1] Tohru Inoue,[9] Ian Indans,[10] Daniela Maurici,[1] George Orphanides,[11] Diana Rembges,[5] Susanna-Assunta Sansone,[12] Jason R. Snape,[13] Eisaku Toda,[14] Weida Tong,[15] Joost H. van Delft,[16] Brenda Weis,[17] and Leonard M. Schechtman[18,19]

[1]European Centre for the Validation of Alternative Methods (ECVAM), Institute for Health and Consumer Protection (IHCP), Joint Research Centre of the European Commission (JRC), Ispra, Italy; [2]Bayer HealthCare AG, Wuppertal, Germany; [3]Hoffmann-La Roche, Basel, Switzerland; [4]Environmental Health Centre, Health Canada, Ottawa, Ontario, Canada; [5]Physico-Chemical Exposure, IHCP, JRC, Ispra, Italy; [6]TNO, Utrecht, the Netherlands; [7]QIAGEN, Hilden, Germany; [8]U.S. Environmental Protection Agency, Washington, DC, USA; [9]National Institute of Health Sciences, Tokyo, Japan; [10]Health Safety Executive, London, United Kingdom; [11]Syngenta, Macclesfield, United Kingdom; [12]European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, Cambridge, United Kingdom; [13]AstraZeneca, Brixham, United Kingdom; [14]Organisation for Economic Co-operation and Development, Paris, France; [15]Food and Drug Administration, National Center for Toxicological Research, Jefferson, Arkansas, USA; [16]University of Maastricht, Maastricht, the Netherlands; [17]National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA; [18]U.S. Interagency Coordinating Committee on the Validation of Alternative Methods, Research Triangle Park, NC, USA; [19]U.S. Food and Drug Administration, National Center for Toxicological Research, Rockville, Maryland, USA

This is the report of the first workshop "Validation of Toxicogenomics-Based Test Systems" held 11–12 December 2003 in Ispra, Italy. The workshop was hosted by the European Centre for the Validation of Alternative Methods (ECVAM) and organized jointly by ECVAM, the U.S. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). The primary aim of the workshop was for participants to discuss and define principles applicable to the validation of toxicogenomics platforms as well as validation of specific toxicologic test methods that incorporate toxicogenomics technologies. The workshop was viewed as an opportunity for initiating a dialogue between technologic experts, regulators, and the principal validation bodies and for identifying those factors to which the validation process would be applicable. It was felt that to do so now, as the technology is evolving and associated challenges are identified, would be a basis for the future validation of the technology when it reaches the appropriate stage. Because of the complexity of the issue, different aspects of the validation of toxicogenomics-based test methods were covered. The three focus areas include *a)* biologic validation of toxicogenomics-based test methods for regulatory decision making, *b)* technical and bioinformatics aspects related to validation, and *c)* validation issues as they relate to regulatory acceptance and use of toxicogenomics-based test methods. In this report we summarize the discussions and describe in detail the recommendations for future direction and priorities. *Key words:* acceptance, alternatives, biomarker, predictive test, regulatory use, standardization, toxicogenomics, toxicology, validation. *Environ Health Perspect* 114:420–429 (2006). doi:10.1289/ehp.8247 available via *http://dx.doi.org/* [Online 17 August 2005]

Toxicogenomics, an emerging field in molecular toxicology, offers the promise of new approaches to identify and characterize such factors as the biologic activity of new and existing chemicals and drugs and could play an important role in hazard assessment for human health. This revolutionary field can potentially affect many scientific and medical areas, including the development of a new generation of alternative predictive testing and screening methods that could lend themselves to the reduction, refinement, and replacement of animals used for such purposes.

The European Centre for the Validation of Alternative Methods (ECVAM), the U.S. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) are currently investigating the specific considerations necessary for adequate validation of toxicogenomics-based test methods. The primary objective of ECVAM and ICCVAM/NICEATM is to facilitate development, validation, and regulatory acceptance of new, revised, and alternative test methods that reduce, refine, and replace the use of animals (referred to as the three Rs; Russell and Burch 1959) in testing while maintaining and promoting scientific quality and the protection of human health, animal health, and the environment. The efforts of such organizations as ICCVAM/NICEATM and ECVAM have helped foster the principles of the three R's and have contributed to progress in the use of alternative methods for regulatory, research, and educational purposes.

Experience in the validation of conventional alternative test methods has led to an understanding that new and innovative approaches likely will be necessary to standardize test methods based on toxicogenomics and to evaluate the scientific validity and regulatory applicability of such test methods. It is envisioned that the entire validation process will be more complex and challenging than that typically encountered thus far for other alternative test methods. This is because not only will the technology itself need to be standardized and validated, but the methods that are based upon the technology and their predictive aspects will also need to undergo validation if they are to be employed in regulatory decision-making processes. In addition the validation process must be able to accommodate the anticipated rapid changes in technology that could affect the performance of the test method and its reliability for a specific purpose.

Toxicogenomics-based methods are being widely applied in toxicology and biomedical research. Because data are already being generated using these technologies, it is both timely and important to address the subject of validation now with the aim of establishing a foundation that will facilitate future regulatory acceptance of scientifically validated

Address correspondence to R. Corvi, European Centre for the Validation of Alternative Methods, IHCP, Joint Research Centre of the European Commission, Via E. Fermi 1, 21120 Ispra, Italy. Telephone: 39-0332-785266. Fax: 39-0332-785845. E-mail: raffaella.corvi@jrc.it

Supplemental Material is available online (http://ehp.niehs.nih.gov/members/2005/8247/suppl.pdf).

We gratefully acknowledge W.S. Stokes, Director of the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, for contributions to the planning and organization of the workshop as well as insightful and constructive comments.

This document represents the consensus of the participants' views expressed as individual scientists and does not necessarily represent the policies and procedures of their respective institutions.

The authors declare they have no competing financial interests.

Received 22 April 2005; accepted 17 August 2005.

toxicogenomics-based test methods. By addressing the critical validation issues early, and in parallel with the evolutionary and maturation phases of the technologic development of toxicogenomics-based methods, it should be possible to preempt many potential pitfalls and data gaps encountered with retrospective method evaluations that could impede validation of this promising research and regulatory tool. Such a strategy will also facilitate early buy-in and confidence in the technologies by the regulatory arena in its quest for new, improved, and relevant methods by which to help ensure human health, protect the environment, and demonstrate responsiveness to animal welfare issues.

In consideration of all these related issues, ECVAM and ICCVAM/NICEATM held the first of a planned series of workshops to address the validation principles that lend themselves to toxicogenomics-based test methods, for example, gene expression technologies and associated bioinformatics. Given the complexity of the rapidly evolving toxicogenomics field, a variety of issues were addressed. These included but were not limited to *a*) differences in and evolution of technology platforms including changes in genome coverage for model species; *b*) quality assurance (QA) and Good Laboratory Practice (GLP) compliance; *c*) technology standardization, transferability, and reproducibility; *d*) relevance to *in vivo* biological responses; *e*) yardsticks against which toxicogenomics responses should be measured; *f*) data evaluation, statistical approaches, and databases; *g*) validation approaches; and *h*) regulatory acceptability.

To begin to examine these complex issues, three breakout groups were formed. Each group concentrated on different aspects of the validation of toxicogenomics-based test methods, and the discussions were shared with the other participants in plenary sessions. The three focus areas were *a*) biological validation of toxicogenomics-based test methods for regulatory decision making, *b*) technical and bioinformatics aspects related to validation, and *c*) validation issues as they relate to regulatory acceptance and use of toxicogenomics-based test methods.

Validation of Toxicogenomics: Focus on the Biological Systems

The biological issues related to the validation of toxicogenomics-based test methods involved two strategies proposed for developing and validating such methods so that they can be employed to support regulatory decision making. One strategy involves phenotypic anchoring of gene expression changes to identify molecular mechanisms and candidate biomarkers of toxicity (i.e., single genes, proteins, or biological pathways). A second strategy involves the identification and validation of predictive gene expression signatures of toxicity. Validation considerations specific to data quality and cross-platform and interlaboratory variability that are common to both strategies were identified. It is acknowledged that any new toxicogenomics-based methods will need to address established validation criteria for determination of reliability and relevance (Balls et al. 1995; ICCVAM 1997, 2003) as well as articulate the advantages and limitations of a given toxicogenomics-based test method. In addition biological validation of such a test method, that is, assessment of the concordance of gene changes with biological events, is essential but is contingent upon validation of the technology itself, which is addressed elsewhere in this article.

Strategy 1: use of toxicogenomics data to define mechanism and identify biomarkers. Toxicogenomics offers the opportunity to enhance existing toxicity prediction strategies through elucidation of biological mechanisms around critical events. This sentiment is captured in the recent U.S. Environment Protection Agency (EPA) and U.S. Food and Drug Administration (FDA) strategies regarding the inclusion of genomics data in submissions of regulated substances (U.S. EPA 2002; U.S. FDA 2005). Although these agencies currently preclude basing regulatory decision making on genomics data alone, they do encourage the voluntary submission of well-documented, quality genomics data. Both agencies are considering the use of submitted data on a case-by-case basis for assessment purposes (e.g., to help elucidate mechanism of action or contribute to a weight-of-evidence approach) or for populating relevant comparative databases by encouraging parallel submissions of genomics data and traditional toxicologic test results. This approach is appropriate given the state of scientific knowledge of toxicogenomics and the requisite need for a clear understanding of the toxicologic relevance of the gene expression signals detected by this technology. There is a small but rapidly increasing number of published reports demonstrating a linkage between gene expression changes and adverse phenotypic changes (Huang et al. 2003; Orphanides 2003). These reports provide qualitative evidence of the power of genomics to link phenotype with gene expression, thereby contributing to an understanding of mechanism of action. Some such reports demonstrate the predictive power of these data to classify compounds. However, they fail to address adequately quantitative dose- and time-dependent (e.g., threshold) responses that are the hallmark of toxicologic evaluation, making their immediate acceptance in regulatory arenas circumspect.

Nonetheless, toxicogenomics data may eventually be useful in hazard and risk assessment if data quality and validity can be adequately substantiated. Some regulators are finding that these data have the potential to add to the body of knowledge about compound mechanism of action. With appropriate dose- and time-dependent measurements, gene and protein changes can be used to mark the molecular events that occur as an organism moves through the continuum from exposure to response. The obvious benefit is the identification of early markers of response, including responses that mark the point of departure from adaptation to toxicity. In addition, it may be possible to detect unforeseen effects at very low doses or in unexpected tissues (Brown et al. 2002). This is important because changes in gene or protein expression alone are not sufficient to differentiate toxicity from biologic adaptation after exposure to an exogenous compound. The challenge for predictive toxicology is to link changes in gene and protein expression to sequential changes in phenotype, both adaptive and adverse, in a manner that is consistent with the underlying biologic mechanisms. For example, gene expression profiling has been used to classify hepatotoxins based on mechanism of action and to differentiate early, presumably adaptive, responses from later responses that are reflective of toxicity (Hamadeh et al. 2002a, 2002b; Waring et al. 2001, 2003). The gene expression changes correlated well with changes in histopathology and clinical chemistry, supporting the liver as target organ for the test compounds.

Although good technical progress has been made in recent years, additional proof-of-principle studies are needed for the regulatory community to become more accepting of the use of toxicogenomics data as part of the regulatory decision-making process. It would be important to demonstrate, for instance, that toxicogenomics not only can confirm what is already known about specific compounds and toxic end points (i.e., phenotypic anchoring) but also can accurately predict toxicity for unknown compounds. The task is to present regulatory scientists with new knowledge gained from toxicogenomics approaches in a familiar context. Ideally, at least in the short term, the focus will be the identification of single, or small sets of, genes or proteins that serve as biomarkers of response, as opposed to signatures of response that are the typical output of microarray experiments. Simple biomarkers of response are favored over complex expression signatures because they are familiar in toxicology assessment, are easy to maintain over time (e.g., are independent of the microarray platform), and can be readily validated. Validation strategies for toxicogenomics-based markers can be modeled after protocols for existing biomarkers. Thus, global gene expression technologies such as microarrays can be used to identify a specific gene marker,

or a suite of markers, that can then be validated by conventional methods such as Northern blot analysis, *in situ* hybridization, and quantitative polymerase chain reaction. This approach has advantages because regulatory agencies such as the U.S. FDA have proposed procedures to address gene and protein biomarkers, and other organizations, such as the Organisation for Economic Co-operation and Development (OECD 2005), are embarking on establishing similar guidance (Supplemental Material, Section 1; http://ehp.niehs.nih.gov/members/2005/8247/suppl.pdf).

Proof-of-principle studies could be conducted concurrently with existing regulatory test methods using similar samples of test compounds. In such situations, it may be appropriate to use *in vivo* systems, which are widely accepted by the regulatory community. Parallel *in vitro* studies could be conducted in situations where an appropriate test system is available. It may be wise to focus initial efforts on defining relationships between gene expression changes and toxicity for individual compounds or compound classes with well-defined end points. The experimental design should address conventional aspects of dose and time (dose response), species and strain susceptibility, group size and sex, and selection of end points for study (e.g., histopathology, clinical chemistry). Numerous commercial microarray platforms offer genomewide coverage for model systems such as rat, mouse, *Caenorhabditis elegans*, and humans. Commercial microarrays are also available for genes that are highly expressed in specific tissues (e.g., liver, breast) and during specific biological processes such as metabolism (e.g., P450 enzymes). Both genomewide and dedicated arrays can be used with RNA samples from *in vivo* and *in vitro* (tissue and cell culture) systems, enabling parallel studies to be conducted with a single microarray platform. This is important because the results of microarray experiments can vary depending on the array design and the selection and performance of gene probes on the array. Encouraging results on cross-platform comparisons and between-laboratory reproducibility are now emerging (Bammler et al. 2005; Chu et al. 2004; Irizarry et al. 2005; Larkin et al. 2005; Yauk et al. 2004). Toxicogenomics studies conducted in parallel and comparative systems can demonstrate the biologic relevance of *in vitro* models as surrogates for *in vivo* models without the need to address cross-platform (technologic) issues (Boess et al. 2003; Huang et al. 2003). Although initial efforts should focus on defining simple gene and protein biomarkers for specific compound classes, end points, and model systems, the end goal is to establish a compendium of compound-specific knowledge that transcends technology platform. Ideally, the markers should be robust enough to withstand technologic advances in toxicology that add to the existing knowledge about the compound. Once sufficient and adequately validated data are available, toxicogenomics can become part of a hierarchical approach to compound assessment.

The use of toxicogenomics to identify (screen) compounds with the potential to cause adverse effects may present opportunities to reduce the need for full animal tests, or perhaps refine animal use, and/or reduce the numbers of animals needed when *in vivo* tests are necessary. Of course, the statistical power of any test will influence the number of animals used in an *in vivo* test as well. Screening-type assessments may be appropriate for priority setting, dose setting, chemical ranking, and so forth. The extent of validation required for screening tests may be different than that required for full replacement tests because negative compounds might still undergo full animal testing. Establishing a compendium of compound-specific information will enable regulators and sponsors to access what is known about a compound across multiple test systems, species, and end points, thereby improving the biological relevance of regulatory decisions to safeguard human health and the environment.

Strategy 2: use of gene expression signatures to predict toxicity. Toxicogenomics holds great promise for improving predictive toxicologic assessments. Gene expression profiling has been used to classify compounds by chemical class and mechanism (Hughes et al. 2000; Scherf et al. 2000; Steiner et al 2004; Thomas et al. 2001), tumors by origin and type (Chung et al. 2002), and breast cancer patients for follow-up chemotherapy (van 't Veer et al. 2002). In all cases, classification was based on a set of discriminatory gene elements, between 10 and several hundred, identified from a larger pool of genes on a microarray. The pattern of gene expression, not the measurement of a single or a small set of genes, was the basis for classification. A variety of gene expression analysis algorithms were used to discriminate samples based on gene expression signature. In all cases, the compound class or tumor status was known *a priori*, and gene expression signatures for known samples were used to predict classification for other known but blinded samples (Blower et al. 2002; Brindle et al. 2002). Such models are currently being developed in the private sector (e.g., Gene Logic, Iconix) and are commercially available but cannot, as yet, be exploited by regulators and the scientific community because the underlying data sets and algorithms have not been made available outside the private sector.

Predictive model development will require an extensive "training" set of gene expression measurements for classes of model compounds in a variety of test systems, both *in vivo* and *in vitro*, at multiple doses and time points. Initial studies can be conducted concurrently with conventional testing systems as a way to confirm model predictions. In the short term, it is unlikely that sufficient data will be available for gene expression signatures to replace conventional approaches. Until then, such data can be used as part of a hierarchical approach to toxicity testing in conjunction with accepted methods routinely used for regulatory purposes. In the long-term, sufficient data should accumulate from well-designed validation studies such that gene expression signatures could be part of a battery of tests that reduce or replace animal procedures.

Model validation will necessitate multiple independent data sets and application of sophisticated statistical approaches. Acceptance of these models will require that research and regulatory communities have access to the data analysis tools used to build the models, and that they become familiar with the limitations and uncertainties of using these complex computational models. Confidence in and acceptance of these models will also require rigorous performance standards and appropriate controls to ensure reproducibility and stability over time (see below) and adequate sensitivity and specificity to discriminate toxic from nontoxic responses. Initial model development could easily be accelerated through coordinated sector-spanning efforts. Coordinated efforts across academia, government, and industry partnerships will accelerate progress in defining gene sets that are robust and discriminatory both within and across technology platforms. This is an ideal scenario given the rapidly advancing pace of technology development.

An important aspect of any toxicogenomics validation strategy is the need to measure the range of biological variability of gene responses for a given test system. Ideally, this should be accomplished by one species, tissue, and end point at a time, in order to adequately assess cross-species differences that often hamper risk assessments. Measurements of biologic variability under baseline and toxicant-challenged conditions will enable regulators to better discriminate biologically relevant responses from baseline homeostatic fluctuation. This is an important issue for toxicogenomics, as studies conducted on cell culture populations demonstrate a wide range of biological variability in gene expression measurements for individual cells under both baseline and challenged conditions (Kuang et al. 2004). Therefore, it is necessary to define criteria to adequately address biological variability in a data submission and to establish whether the burden of maintaining these data is that of the regulator or sponsor.

The recommendations related to the biological validation of toxicogenomics-based test methods are listed in Table 1.

Standardization and Validation of Toxicogenomics-Based Methods: Focus on the Technology

Considerations given to validation of the technology encompassed the technical and bioinformatics issues related to the validation of toxicogenomics-based test methods. The starting premise adopted was that with the availability of bioinformatics expertise, biological data generated from toxicogenomics studies could be interpreted with a high degree of confidence. The ultimate aim was to identify a strategic approach that would enable credible biological observations and consequential judicious regulatory decisions, and that this approach would be independent of the toxicogenomic platform used. Moreover, standardization and validation of toxicogenomic platforms were seen as essential for identifying and reducing technologic artifacts. Standardization would also be required to increase the certainty by which biological observations could be extrapolated across and between different microarray platforms. It is therefore important to build on the learning of previous and ongoing efforts in standardization of toxicogenomics (reviewed by Sansone et al. 2004).

Three distinct levels where validation is necessary were identified (see Figure 1 and discussion below). The first level of validation is the responsibility of the array manufacturer or provider and has to be performed only once. This can be seen as a "one-off validation" and relates to both the microarray quality and the instrumentation. The second level of validation is the responsibility of both the experimental toxicologist and the array manufacturer or provider. This can be seen as "routine validation" or best practice to allow data comparability. It encompasses quality control (QC) aspects of the critical experimental components and is a process that occurs on a regularly scheduled basis. The third level of validation, that is, determination of reliability and relevance, is needed every time a change is introduced into the test procedure. Performance standards developed based upon the original test method would serve as the criteria against which the revised method would be compared. Despite these multilevel validation needs, it was repeatedly emphasized that significant technologic development and progress in microarray platforms are still under way and that efforts to validate and standardize these technologic platforms must not be at the expense of innovation.

One-Off Validation

The one-off validation is the responsibility of the array manufacturer or array provider. This is required to ensure that the array platform being used is robust and that the inherent variability within the platform is transparent to the user and the regulator (Figure 1). The following were identified as being necessary for microarray-based toxicogenomics to be used in regulatory assessments:

- Microarrays should be fabricated in accordance with the principles of Good Manufacturing Practice (GMP).
- Specifications and performance criteria for all instrumentation and method components should be available.
- All quality assurance/quality control (QA/QC) procedures should be transparent, consistent, comparable, and reported.
- The array should have undergone sequence verification, and the sequences should be publicly available.
- All data should be exportable in a MAGE (MicroArray and Gene Expression)-compatible format.

Table 1. Recommendations: focus on biological systems.

- Encourage increased use of toxicogenomics-based approaches to define the mechanistic context of toxic responses to exogenous compounds
- Promote greater understanding of the relationships between gene expression responses and altered phenotype, considering the biological pathways affected, dose response, and the point of departure from adaptive to toxic response
- Favor the identification of biomarkers that are independent of technology platform but acknowledge the potential strengths of pathway analysis
- Characterize the range and extent of biological variability of responses for the test systems (e.g., diurnal effects, animal care and use, age-related context)
- Encourage the immediate use of toxicogenomics-based approaches in conjunction with conventional toxicity testing approaches
- Explore the extent to which toxicogenomics can address cross-species responses and specific disease states
- Promote the conduct of parallel and comparative *in vivo* and *in vitro* studies to identify *in vitro* systems that can serve as surrogates for *in vivo* systems
- Characterize predictive toxicology models with respect to parameters such as dose, time, study design, relevance; characterize the system to fulfill validation criteria
- Promote the identification of gene and protein biomarkers as early (prognostic) markers as a refinement to existing toxicity testing methods
- Establish a compendium of toxicant information based on gene expression responses for model compounds across multiple species, end points, and test systems
- Foster the development of effective partnerships between academic, government, and industry groups to promote collaborative efforts to validate toxicogenomics-based test methods and generate sufficient high-quality data to support regulatory decision making

Routine Validation

Routine validation is an ongoing process that is the responsibility of the experimental toxicologist and the array manufacturer or provider (Rockett and Hellmann 2004). Again, for microarray-based toxicogenomic assays to be used in regulatory decision making the following important factors were identified (Figure 1):

- Oligos, cDNAs, or clones that are arrayed should be randomly sequence-verified to ensure that no errors are introduced between batch syntheses. This verification process should be recorded and reported by the manufacturer
- All reagent components should be identified. Reagents should be prepared according to GMP and/or GLP as appropriate. Data regarding batch variability should also be recorded and reported
- Common reference RNA standards (housekeeping genes) should be adopted to facilitate comparison between array platforms. This may be achieved in collaboration with the international Microarray Gene Expression Data (MGED) Society and other related efforts (see below).

Biological standards. Performance standards, test component standards, and QC measures are key components of any validation strategy for a toxicologic test method. Establishing standards is particularly important for gene expression technologies due to the inherent technologic and biological "noise" in these systems. Commonly used biological standards are reference RNAs that are competitively hybridized with the sample of interest in two-channel array formats, and *in vitro* RNA transcripts that are "spiked into" RNA samples of interest in either one-channel or two-channel array formats. Establishing accepted RNA standards will address concerns of regulatory reviewers about data quality and variability within and between laboratories and across different technology platforms. The standards will also provide a common benchmark for regulators to assess platform performance over time. To achieve this goal, we must establish standards that maintain a defined level of accuracy, sensitivity, specificity, and reproducibility across platforms.

Reference RNAs can be derived from tissue extracts, cell lines, or both and serve a variety of purposes. Workshops sponsored by governments and industry have focused on defining the specifications for reference RNAs for clinical and regulatory applications (Joseph 2004). The consensus is a that multiple RNA standards are needed to measure the accuracy, dynamic range, sensitivity, and specificity of varied technology platforms under varied conditions. Important questions are whether regulatory agencies will define preferred sources of RNA standards, and, if so, who will generate and maintain baseline information about these

standards. Although the selection of a given RNA standard depends primarily on the purpose and application, all RNA standards should be tested for a clearly defined number of copies of a given sequence within an RNA preparation over some linear range (Cronin et al. 2004).

Some initiatives are raising awareness of the effects of variables that might hamper data comparability and are working toward developing best practice guidelines for microarray-based measurements (Hopkins et al. 2004). For example, recommendations for best practice in array normalization, together with performance characteristics in terms of sensitivity, accuracy, and comparability of different array platforms (cDNA and oligo, spotted and *in situ* synthesis), are beginning to emerge together with proposals for transparency and availability through publicly accessible databases (http://www.vam.org.uk). Other initiatives are considering the use of quality metrics for standardizing and validating array-based toxicogenomics measurements. The extent to which such efforts will be pursued and the impact they will have upon the standardization issues that are a necessary prerequisite to the validation exercises remain to be seen.

Quality assurance and Good Laboratory Practice. GLP is intended to promote proper documentation, quality, and authenticity of toxicity test data and is required for data acceptance by regulatory agencies (e.g., U.S. FDA, U.S. EPA). At the international level, GLP has been promulgated under the OECD guidelines program (OECD 1998). As part of the progression toward regulatory acceptance, toxicogenomics experiments should ideally be conducted in accordance with GLP. However, at present, most large-scale toxicogenomics efforts are not arising from GLP-compliant laboratories, and requiring compliance for data submission could greatly hamper the technical advancement of new technologies and retard their migration into the regulatory arena. To avoid discouraging technologic progress while maintaining a level of GLP conformity, it could be argued that for research and technical development and improvement purposes, it might be acceptable if array-based studies could at least measure up to the reporting standards required by GLP. However, with the adoption of the toxicogenomics-based technologies into regulatory decision-making practices, GLP compliance undoubtedly will be expected. Procedural aspects of GLP compliance not currently captured in MIAME-Tox (minimum information about a microarray experiment for toxicogenomics) will need to be identified but can be incorporated over time. Until then, it may be possible to allow for proof-of-principle and prevalidation studies to be conducted in accordance with the "intent" of GLP practices by requiring submitters to adequately document procedures and control measures and make experimental data open to regulatory review. "Best practices" for toxicogenomics can be established until formal procedures are adopted. This may be a more realistic solution that permits the advancement of science while addressing the need for QA and QC.

Validation as a Result of Procedural Changes

This third level of validation is necessary whenever a technical or methodologic change is introduced into the test. Such changes might, on one hand, be restricted to the microarray technology (e.g., modification or addition of sequences to a microarray, changes in data analysis procedures). Alternatively, they could involve the experimental design (e.g., dose, time, cell culture procedures). One consideration is that a distinction between minor and major procedural changes that might be incorporated into a test would help determine the extent of such validation necessary. Additionally, to facilitate the process, performance standards should be defined based upon the original validated test procedure. Minor changes would entail a demonstration of equivalence of results obtained with the modified test to that obtained from the validated test. Major changes would involve the need to define a new set of reference materials to be tested and a more extensive validation. Guidance on the use of performance standards and the elements comprising them have been published (ICCVAM 2003) and have been employed for *in vitro* dermal corrosion assessment methods (ICCVAM 2004). Such guidance can also help facilitate the establishment performance standards for toxicogenomics-based test methods in which procedural modifications have been introduced after an initial validation exercise, thereby providing a basis for the comparison of reliability and accuracy of the modified method relative to the validated and accepted reference test method.

The concept of performance standards was originally developed to evaluate the acceptability (accuracy and reliability) of proposed test methods that are based on similar scientific principles and that measure or predict the same biologic or toxic effect as an accepted (previously validated) test method. Because some regulatory authorities and international test guidelines programs (e.g., OECD) have restrictions regarding the use of proprietary test methods (methods that are copyrighted, trademarked, or patented), performance standards also allow for the development and validation of comparable nonproprietary methods based on performance standards derived from the corresponding proprietary antecedent method. Under these circumstances, performance standards allow the characteristics and functional attributes of a proprietary method or technique to be described and offer a procedure for evaluating the performance of methods claimed to be substantially similar. A method that meets the established performance standards is

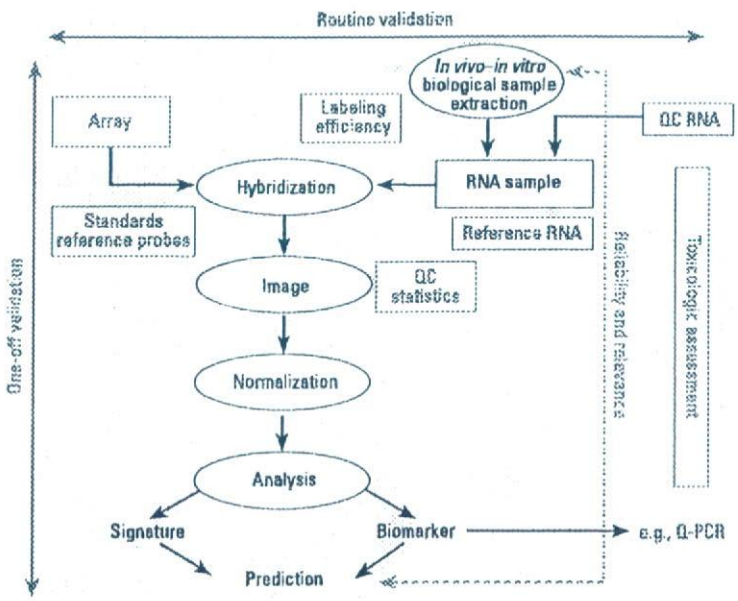

Figure 1. Scheme of the different steps in a toxicogenomics-based test. Three distinct levels were identified where validation is necessary: one-off validation (left), which should be performed once and is mainly related with the quality of the microarray and the instrumentation (blue); routine validation and QC (top), representing the ongoing requirements that are the responsibilities of the experimental toxicologist and the manufacturer (red); and the extent of validation necessary whenever a technical or methodologic change is introduced in the test (right): a method should meet the preestablished performance standards in order to be considered reliable and relevant as the original test method (green). Q-PCR, quantitative PCR.

considered sufficiently accurate and reliable for the specific testing purpose for which it is designed and is viewed as comparable with the original test method upon which it is based. If the correct performance standards have been developed, a method for which the results have the same accuracy and reliability as the original should by definition also be as relevant as the original method.

The conceptual framework and scope of performance standards could be expanded or adapted to include innovations or advancements in areas such as microarray or protein or metabolite separation and identification technology, where proposed improvements might or might not be generally or completely analogous to those in existing systems but would still enable similar applications. Performance standards could still provide a gauge for evaluating newer or revised technologies to ensure that their reliability and accuracy were at least comparable with that of existing acceptable techniques using similar chemicals even if essential test method components (i.e., structural, functional, and procedural elements of a validated test method to which a proposed, mechanistically and functionally similar test method should adhere) were not substantially similar.

This level of validation, which does not imply that a test needs to be completely revalidated, is of extreme importance for tests based on rapidly evolving technologies. It would be a mistake to immobilize these technologies by enforcement of a strict and inflexible validation approach that would hamper progress and test improvement. Finally, a periodic reassessment of a test method's performance (accuracy and reliability) employing established performance standards would help ensure adherence to essential test method components and the reliability and accuracy of the modified test method relative to the validated antecedent method (Hartung et al. 2004). Such assurance could be best established and reported by international validation bodies such as ECVAM and ICCVAM/NICEATM, which could track the history, performance, and validation status of a given test.

Data Management

The lack of robust QC procedures and capture of adequate metadata has caused problems with the analysis and reproducibility of array-based transcriptomics investigations. Consequently, the international MGED Society proposed standards for publication (Nature 2002) that were designed to clarify the MIAME guidelines (Brazma et al. 2001). As a result, a number of journals now require that articles containing microarray experiments must be compliant with the MIAME standard; some also require that the data integral to the article's conclusions be submitted to the ArrayExpress database at the EBI (European Bioinformatics Institute) (Brazma et al. 2003), GEO (Gene Expression Omnibus) at NCBI (National Center for Biotechnology Information) (Edgar et al. 2002), and CIBEX (Center for Information Biology Gene Expression database) at DDBJ (DNA Databank of Japan) (Ikeo et al. 2003)—the European, American, and Japanese database counterparts, respectively.

There is a critical need for public toxicogenomics databases because of the significant volume of data associated with these experiments, the complexity of comparing different gene annotations and splice variants across platforms, and the need for a resource for complex informatics analyses of the traditional toxicology and microarray data in parallel. However, to fully achieve the potential of this emerging interdisciplinary field, it is necessary that we move toward the establishment of a common public infrastructure for exchanging toxicogenomics data (Mattes et al. 2004). The infrastructure should address *a*) the technical problems involved in data upload, *b*) the demand for standardizing data models and exchange formats, *c*) the requirement for identifying minimal descriptors to represent the experiment, *d*) the necessity of defining parameters that assess and record data quality, and *e*) the challenge of creating standardized nomenclature and ontologies to describe biological data. The goal is also to create an internationally compatible informatics platform integrating toxicology/pathology data with transcriptomics, providing the scientific community with easy access to integrated data in a structured standard format, facilitating data analysis and data comparison, and enhancing the impact of the individual data sets and the comprehension of the molecular basis of actions of drugs or toxicants. Ultimately, such a knowledge-base could be maintained (respecting confidentiality as appropriate) as a reference for regulatory organizations to evaluate toxicogenomics and pharmacogenomics data submitted by registrants to those organizations.

The potential exists for the international development of this public infrastructure. As part of the collaborative undertaking with the International Life Sciences Institute Health and Environmental Sciences Institute (ILSI–HESI) Technical Committee on the Application of Genomics to Mechanism Based Risk Assessment (http://www.hesiglobal.org/committees), the European Molecular Biology Laboratory of the European Bioinformatics Institute (EMBL–EBI; Brazma et al. 2003; http://www.ebi.ac.uk/microarray/Projects/toxnutri/index.html), the National Institutes of Health/National Institutes of Health National Institute of Environmental Health Sciences National Center for Toxicogenomics (NCT; Waters et al. 2003; http://www.niehs.nih.gov/nct/), and the U.S. FDA NCT (Tong et al. 2003; http://www.fda.gov/nctr/science/centers/toxicoinformatics/index.htm) have worked closely together. The respective databases are based on the international standards developed by the MGED Society (Brazma et al. 2001; Spellman et al. 2002). After the very favorable response that the MIAME received from the microarray community and key scientific journals (Ball et al. 2002, 2004; Nature 2002), the MIAME checklist was extended to describe array-based toxicogenomics experiments. The MIAME-Tox checklist (MGED 2004) is an attempt to define the minimum information required to interpret unambiguously and potentially reproduce and verify array-based toxicogenomics experiments. MIAME-Tox also supports a number of other objectives, for example, linking data from different experimental domains within a study and linking several studies from one institution and exchanging toxicogenomics data sets among public databases. The major objective of MIAME-Tox is to guide development of toxicogenomics databases and data management software. Without a sufficient depth of data in these resources, the scientific community's opportunity to develop consensus on analysis and application of these data for risk assessment or screening may be limited. The availability of this level of information regarding platform specification, appropriate common reference standards, and the toxicologic study alone will facilitate the predictive value of toxicogenomics across different array-based platforms. This, in turn, will result in a greater appreciation of and confidence in the value of toxicogenomics within a regulatory context, such that testing strategies can be optimized, predictive alternative models can be identified, and animal use can be reduced (Supplemental Material, Section 2; http://ehp.niehs.nih.gov/members/2005/8247/suppl.pdf).

Moreover, the long-term provision of a MIAME-Tox–compliant database with a MAGE-ML (Microarray Gene Expression Markup Language) export is required for the long-term storage of toxicogenomics data. This would directly support the role of ECVAM, ICCVAM/NICEATM, and other validation bodies in the validation of toxicogenomics-based test methods.

The recommendations related to the technical and bioinformatics aspects of validation are listed in Table 2.

Regulatory Acceptance of Validated Toxicogenomics-Based Methods

Regulatory scientists are increasingly being called upon to consider incorporation of toxicogenomics data in regulatory assessment processes that involve evaluation of potential human health or environmental hazard and risk. Those scientists will need to be able to

judge the level of confidence to place in both *in vivo* and *in vitro* toxicogenomics-based test methods and the resulting data that might be submitted in support of regulatory decision making. Whether a method has been determined to be valid for a specific purpose will be an important factor for the consideration of its use for regulatory purposes. Furthermore, the level of confidence held by regulators will influence regulatory acceptance of methods and data, and will affect both the further pursuit of toxicogenomics technologies and technologic improvements and the extent of industry application of these technologies.

Potential uses of toxicogenomics data in the regulatory area. The potential of toxicogenomics-based methods in contributing to regulatory assessment processes is broad. Examples might include, but would not be limited to, obtaining microarray data from individual *in vivo* bioassays or *in vitro* cell or tissue-based assays or from batteries of assays, using conventional or high-throughput approaches. In accordance with the current developing state of the science, realistic possibilities for initial uses of toxicogenomics data in regulatory settings might be first in the realm of hazard assessment, such as to support chemical mechanism of action arguments. Other early uses might include aiding individual chemical/chemical mixture screening or ranking exercises to set priorities for toxicity testing or to sort chemicals into batches. These types of applications might involve identification of individual genes or gene patterns associated with particular toxic effects or pathways, adaptive responses, or metabolic pathways. However, global pattern recognition–type techniques are, as yet, not considered to be ready to fully replace traditional bioanalytical methods for predicting toxicity or elucidating information on mechanism of action or biochemical pathway component identification.

Using only human or animal *in vitro* or *in vivo* data derived from toxicogenomics technology to estimate such parameters as adverse/no adverse effect levels or to determine dose–response relationships for conducting risk assessments is regarded as a much longer term goal. However, for hazard assessment purposes, the possibility of considering toxicogenomics data along with other types of toxicologic information and data [e.g., from *in vivo* and *in vitro* studies, determinations of quantitative structure–activity relationships (QSAR) or SAR] in a weight-of-evidence approach on a case-by-case basis was not discounted. Regulatory bodies have begun to craft preliminary proposals, policies, and guidance for the submission and use of omics-type data in regulatory deliberations and to provide encouragement for the use and further development of the technology (U.S. EPA 2002; U.S. FDA 2005). Additionally, organizations such as the OECD are actively working with member countries on approaches that seek to harmonize the use of omics-derived information for hazard assessment related to health and environmental effects.

Harmonization of toxicogenomics-based test methods will first necessitate the standardization and validation of the specific test protocol(s) developed for a specific purpose(s), as conducted by international validation bodies such as ECVAM and ICCVAM/NICEATM. It will then be important for such organizations to interface with the OECD to ensure the appropriate crafting of harmonized OECD toxicogenomics-based test guidelines that are based upon standardized, adequately validated procedures, that are considered practical, and that permit consistent regulatory judgments.

Case for a modular approach to validation. Because of the extraordinary rate at which toxicogenomics technologies are evolving, current validation processes might need to adapt so as to accommodate the rapidly developing changes and advancements while still observing the basic tried-and-true validation principles. To meet this anticipated need, a modular approach to validation (Hartung et al. 2004) was considered, not to abridge the process but to allow for more flexibility in data collection and evaluation throughout the progressive changes that the technology will undergo. Typically, in the conventional validation procedures for an alternative test method, a sequential approach to the process is taken. The test protocol is first optimized and its transferability is determined. The resulting standardized method is then evaluated for within-lab and between-lab reproducibility and for its accuracy. Thus, an optimized, standardized protocol linked to specific test method elements and a prediction of outcome for given classes of chemicals are evaluated together for performance characteristics and applicability. Such a linear validation model, although effectively employed for other test methods, might not be optimal for dynamic test methods in which changes are rapidly introduced that improve or alter the protocol or the technology incorporated in the protocol in any substantive way. The linear validation model might result in unnecessary delays in incorporating innovations into toxicogenomics-type test methods. In contrast, with a modular approach to validation, which capitalizes on the fundamental classic concepts of validation as defined by ECVAM and ICCVAM (Balls et al. 1995; ICCVAM 1997, 2003), the different steps in the validation process are subdivided into independent modules, each of which can be assessed individually so that those components that have been completed need not undergo repeated validation. Further validation activities would instead be directed to only that part of the process flow where needed. The proposed model would accommodate validation of innovation affecting only a particular part of the sequence such that incorporation of advancements in a particular sector into testing strategies would less likely be impeded. At the same time, a modular approach to validation could efficiently handle information/data gaps that could be filled over time without derailing the validation stages already achieved. The modular approach, complemented with the use of performance standards (see "Validation as a Result of Procedural Changes" above), is expected to facilitate and help expedite the validation of the toxicogenomics technology and test methods that are based on toxicogenomics.

The modular approach follows the fundamental classic concepts of validation as defined by ECVAM and ICCVAM. Validation is defined as the process by which the relevance and reliability of a test method for a specific purpose are determined (Balls et al. 1995; ICCVAM 1997, 2003). Adequate validation

Table 2. Recommendations: focus on technology.

- Validation and QA/QC should be mandatory during the manufacturing of the arrays
- The array should undergo sequence verification and sequences should be available in the public domain
- MIAME guidelines should be adhered to
- Initially, develop "best practices" for toxicogenomics, including the interpretation of data and how to manage uncertainties and limitations
- Subsequently develop guidance for and adherence to GLPs for toxicogenomics experiments
- Common reference standards should be considered
- A workshop should be convened to address the development of standards for RNA sample preparation (and other biologic aspects of microarray analyses)
- Develop a "common" RNA standard including developing consensus about sources and maintenance of baseline data for regulatory and research purposes
- Studies should be MIAME-Tox compliant
- Performance standards should be developed and implemented to evaluate reliability and accuracy of test methods incorporating procedural modifications
- An ongoing dialogue should be maintained between scientists in the various relevant disciplines, including bioinformaticians, through meetings, published papers, and advisory/discussion panels (e.g., ILSI-HESI committee, NCT consortium, OECD panel)
- Ensure that validation efforts and QA/QC criteria are not restrictive to the technology or its advancement
- Explore whether toxicogenomics measurements can define toxicologic effects quantitatively
- Develop prediction models (e.g., algorithms) for toxicogenomics-based test methods
- Develop a data infrastructure for capturing, storing, and reporting toxicogenomics data
- Ensure continuation of financial support for long-term public database maintenance

involves development of a standardized test method protocol and assessment of the protocol's within- and between-laboratory variability, predictive capacity/accuracy, usefulness and limitations, and adherence to performance standards.

Standards for comparison. As technologic advancements are made and new, modified, or revised toxicogenomics-type test methods are put forward for consideration, it will be necessary to have a means by which the performance of proposed methodologies can be compared with that of existing (traditional and nontraditional) methods, especially those that employ animals. The lack of an approach rooted firmly in high-quality science could jeopardize attempts to seek or gain regulatory acceptance of toxicogenomics-based test methods and strategies. Evaluations of test method performance might be based on comparisons made between particular parameters, as dictated by the specific intent for which the assay was developed. Examples include the following:

- *In vivo–in vivo* study comparisons to examine concordance of gene changes with such factors as onset, duration, severity, dose, age, possible temporal changes of effects, and species differences
- *In vitro–in vivo* study comparisons to explore gene changes associated with a critical event or end point in an *in vitro* cell-based assay and an established *in vivo* biomarker of toxicity
- *In vitro–in vitro* study comparisons to analyze the responses of human and animal cell systems to xenobiotics
- Technologic comparisons to evaluate the effects of proposed technical improvements (e.g., comparing gene changes using different techniques of array/platform preparation)

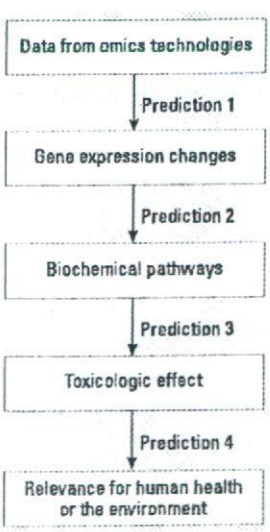

Figure 2. Process flow showing different independent prediction levels considered important in assessing validity of a toxicogenomics-based test method.

Accordingly, to determine the appropriate types of validation activity and comparison in a given situation, it is important that the specific purpose of the proposed methodology and a detailed description of all relevant procedures be clearly elaborated (Balls et al. 1995; Hartung et al. 2004; ICCVAM 1997, 2003).

Toxicogenomics data from in vitro *systems and data relevance.* At the present time, toxicogenomics data derived from *in vitro* systems have been considered to have limited utility in regulatory applications. However, a great deal of interest exists for the further development of *in vitro*–based toxicogenomics methods, for an examination of their potential applicability in the regulatory arena, and for an appraisal of their potential for contributing to improvements in animal welfare. It is anticipated that technologic advancements will ultimately facilitate the use of *in vitro*–based methods as adjuncts to or surrogates for *in vivo*–based methods. Possible areas where validated *in vitro*–based toxicogenomics test methods might play a future role include *a*) preliminary assessments (prescreens), *b*) complementary testing that might assist in obtaining additional (e.g., mechanistic) information, and *c*) surrogate tests that could help in the refinement, reduction, and replacement of animals used for omics-based or traditional testing methods. One exciting aspect of toxicogenomics technology is the prospect of being able to identify species differences and/or similarities in the response to a xenobiotic. Although this is not viewed as near-term prospect, it obviously has potential applications for hazard and risk assessment purposes and could also have an impact on previous regulatory decisions when the technology becomes sufficiently advanced to permit such uses for it.

Additional regulatory acceptance issues. In considering approaches to validation, achieving regulatory acceptance of toxicogenomics-based methods or acceptance of information/data derived from such methods is an important goal. Regulators will be asked to evaluate whether data submitted using omics technologies can be used in support of a particular or broader based toxicologic, pharmacologic, or physiologic premise. For example, experiments using microarrays demonstrated increased expression of a cluster of related genes that was associated with enhanced activity and production of a microsomal enzyme important in the metabolic activation of a chemical to a toxic entity, which in turn was associated with a histopathologic biomarker lesion in the liver with a known human cancer correlate. Each of the events in this example can be thought of as a sequence of separate critical steps or information levels (Figure 2) that progressively connect omics data (from microarrays) to gene expression changes (increased expression), to a biochemical pathway (liver enzyme induction leading to toxic metabolite formation), to a toxicologic effect *in vivo* (liver lesion) with human relevance (cancer). Moving between two levels involves a prediction of outcome linking both steps. At each of these prediction junctures, regulators would be looking for evidence to scientifically substantiate moving to the next step and whether the prediction linking the levels (e.g., in this example, prediction 1, 2, 3, or 4 in Figure 2) was adequately validated. Theoretically, with this type of system, validated links could be established between any two levels. Technologic advancements or new information could be independently incorporated into a given level and considered and evaluated for the specific relevant prediction juncture. In this way, each of the prediction levels can be assessed independently and the validity of the links determined.

In the future toxicogenomics-based test methods may be shown to have been adequately validated and technically suitable for certain specific purposes, but regulatory acceptability and implementation will depend partly on whether the methods validated can be used for a given regulatory agency or program, that is, they are applicable to the products that fall within their regulatory purview. Some regulatory bodies may have internal peer-review processes, specific regulatory mandates, and/or regulatory assessment procedures that also have a role in the determination of test method applicability in regulatory programs, even though a test method may have been appropriately validated.

The widespread use of omics technologies will also bring about increasing demands on the regulatory community in terms of training of regulatory personnel in areas such as potential applications; data QC, analysis, and interpretation; statistical analysis; limitations of the technology; and how the information might be incorporated into safety, hazard, and risk assessment processes. To satisfy these needs, regulatory agencies have been engaging in developing and implementing training procedures, hiring scientists with the necessary technical knowledge and experience, establishing centers of excellence and dedicated laboratories focused specifically on the various omics and related informatics areas [e.g., National Center for Toxicological Research (U.S. FDA), NCT (NIEHS), Minister of Health Labour and Welfare–National Institute of Health Sciences Project in Japan, Netherlands Genomics Initiative, and EMBL–EBI, where informatic scientists are working with experimental practictioners and the MGED Society to ensure that transcriptomic experiments can be mapped on to regulatory toxicology studies]. In addition the regulatory arena has found that maintenance of open lines of communication with appropriate external scientists facilitates cooperation and the sharing of technical aspects, skills,

and practical experiences that help to broaden the collective knowledge base. Regardless, as the technology evolves further and finds wider application and acceptance, it will be necessary to address such fundamental matters as *a*) the generation, management, and interpretation of massive amounts of data; *b*) the consequent complex questions that will undoubtedly arise (e.g., what constitutes an adverse effect as identified using the technology; how does a given gene pattern correlate with a particular toxic end point or relate to onset, duration, and severity of effects, and to age, dose, and species?); and *c*) the limitations to the technology. Addressing such issues efficiently will warrant an ongoing dialogue between regulators and practitioners and a willingness to share relevant experiential and theoretical knowledge. Standard submission and presentation formats compatible with electronic data submission likely would need to be developed. Programs and staff would need to learn how information from the new technologies might be incorporated in regulatory practices and decision-making processes and would also have to face possible incongruities between toxicogenomics-derived data and existing or future submissions of conventional toxicity data. A number of regulatory authorities have already begun to contemplate and make provisions for this enormous and challenging task, but others may not yet have committed the resources to do so.

The recommendations related to regulatory acceptance and use of toxicogenomics-based test methods are listed in Table 3.

Conclusions

This workshop was organized as a result of the rapid growth and technologic advancements in the field of toxicogenomics; the promise it offers for numerous scientific arenas, especially human health and the environment; and the interest demonstrated by regulatory agencies as well as by the industrial sector. Consequently, it has become apparent that a considerable effort needs to be invested in the appropriate validation of both the technology alone and those test methods that incorporate the technology. The workshop provided a platform for technical experts in the field to become cognizant of the validation principles and regulatory issues to be encountered and for regulators and principal validation bodies to gain a better sense of those technologic aspects that would lend themselves to standardization, harmonization, and validation. Thus, this workshop was an important initiative that fostered an exchange of information fundamental to the ultimate adoption of toxicogenomics-based test methods for regulatory decision-making purposes. It is envisioned that the conclusions and recommendations that resulted will be a basis for future validation considerations for test method applications of toxicogenomics technologies in the regulatory arena and evaluating their potential utility for hazard/safety/risk assessments.

Several aspects of the validation of toxicogenomics that were identified as needing further exploration to help facilitate regulatory acceptance of future toxicogenomics-based test methods are as follows:

- Conduct toxicogenomics-based tests and the associated conventional toxicologic tests in parallel to *a*) generate comparative data supportive of the use of the former in place of the latter or *b*) provide relevant mechanistic data to help define the biological relevance of such responses within a toxicologic context
- Determine and understand the range of biologic and technical variability between experiments and between laboratories and ways to bring about greater reproducibility
- In the short term, favor defined biomarkers that are independent from technology platforms, and therefore are easier to validate; in the longer term, focus on pathway analysis (i.e., system biology approach) rather than just on individual genes
- Harmonize reference materials, QC measures, and data standards and develop compatible databases and informatics platforms that are key components of any validation strategy for a toxicologic method; this can only be achieved by promoting partnerships and collaborations among ongoing initiatives in toxicogenomics, standardization, and validation
- Determine performance standards for toxicogenomics-based test methods that will serve as the yardsticks for comparable test methods that are based on similar operational properties
- Define further the modular validation scheme that would allow keeping up with methodologic improvements and innovations without having to repeat the entire validation process but would, however, integrate ECVAM and ICCVAM principles of validation and acceptance.

Table 3. Recommendations: focus on regulatory acceptance of toxicogenomics-based methods.

- Build on and/or learn from previous and ongoing efforts in toxicogenomics, standardization, validation, and harmonization efforts where possible (e.g., MIAME, ICCVAM, ECVAM, NCT, EMBL–EBI, ILSI–HESI, U.S. FDA, U.S. EPA, OECD)
- Fund pilot programs to test possible validation strategies and processes
- Identify training needs and assist in developing training vehicles and ways of presenting the state-of-the-science to regulators and the regulated community (including electronic means)
- Maintain transparency of validation processes
- Explore additions, amendments, and revisions to ICCVAM and ECVAM validation guidance that would accommodate new and rapidly changing technologies
- Implement the modular approach to validation to accommodate existing knowledge and future technical developments
- Establish performance standards for toxicogenomics-based test methods and have them accommodate rapid technologic advancements and procedural modifications
- Explore, develop, and support sector-spanning worldwide harmonization entities
- Create confidence among regulators by involving them early on in discussions and various scientific forums that would facilitate application of the technology for regulatory purposes
- Encourage industry and other parties to share data, in part, to support validation comparisons
- Promote high-quality science in supporting the use and development of the technology for regulatory purposes to further protection of human health and the environment
- Consider opportunities for synergy between QSAR, pharmacokinetic, and pharmacodynamic modeling, and other *in silico* efforts and the toxicogenomics communities

References

Ball C, Brazma A, Causton H, Chervitz S, Edgar R, Hingamp P, et al., on behalf of the MGED Society. 2004. Standards for microarray data: an open letter [Letter]. Environ Health Perspect 112:A666–A667.

Ball CA, Sherlock G, Parkinson H, Rocca-Sera P, Brooksbank C, Causton HC, et al. 2002. Standards for microarray data. Science 298:539.

Balls M, Blaauboer BJ, Fentem JH, Bruner L, Combes RD, Ekwall B, et al. 1995. Practical aspects of validation of toxicity test procedures. Altern Lab Anim 23:129–147.

Bammler T, Beyer RP, Bhattacharya S, Boorman GA, Boyles A, Bradford BU, et al. 2005. Standardizing global gene expression analysis between laboratories and across platforms. Nat Methods 2:351–356.

Blower PE, Yang C, Fligner MA, Verducci JS, Yu L, Richman S, et al. 2002. Pharmacogenomic analysis: correlating molecular substructure classes with microarray gene expression data. Pharmacogenomics J 2(4):259–271.

Boess F, Kamber M, Romers S, Gasser R, Muller D, Albertini S, et al. 2003. Gene expression in two hepatic cell lines, cultured primary hepatocytes and liver slices, compared to the *in vivo* liver gene expression in rats: possible implications for toxicogenomics use in *in vitro* systems. Toxicol Sci 73(2): 386–402.

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 29(4):365–371.

Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N, et al. 2003. ArrayExpress—a public repository for microarray gene expression data at the EBI. Nucleic Acids Res 31:68–71.

Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HW, et al. 2002. Rapid and non-invasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabolomics. Nat Med 12:1439–1445.

Brown HR, Ni H, Benavides G, Yoon L, Hyder K, Giridhar J, et al. 2002. Correlation of simultaneous differential gene expression in the blood and heart with known mechanisms of adriamycin-induced cardiomyopathy in the rat. Toxicol Pathol 30(4):452–469.

Chu TM, Deng S, Wolfinger R, Paules RS, Hamadeh HK. 2004. Cross-site comparison of gene expression data reveals high similarity. Environ Health Perspectives 112:449–455.

Chung CH, Bernard PS, Perou CM. 2002. Molecular portraits and the family tree of cancer. Nat Genet 32(suppl):533–540.

Cronin M, Ghosh K, Sistare F, Quackenbush J, Vilker V, O'Donnell C. 2004. Universal standards for gene expression. Clin Chem 50(8):1464–1471.

Edgar R, Domrachev M, Lash AE. 2002. Gene Expression

Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30:207–210.

Hamadeh HK, Bushel PR, Jayadev S, Martin K, DiSorbo O, Sieber S, et al. 2002a. Gene expression reveals chemical-specific profiles. Toxicol Sci 67:219–231.

Hamadeh HK, Bushel PR, Jayadev S, Martin K, DiSorbo O, Sieber S, et al. 2002b. Prediction of compound signature using high density gene expression profiling. Toxicol Sci 67:232–240.

Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, et al. 2004. A modular approach to the ECVAM principles on test validity. Altern Lab Anim 32:467–472.

Hopkins D, Gout S, Burns M, Nixon G, Andersen M, Foy C, et al. 2004. Comparability of Gene Expression Measurements on Microarrays. LGC/MFB Report. Available: http://www.mfbprog.org.uk/themes/theme_publications_item.asp?intThemeID=22&intPublicationID=961 [accessed 9 January 2006].

Huang E, Ishida S, Pittmann J, Dressman H, Bild A, Kloos M, et al. 2003. Gene expression phenotypic models that predict the activity of oncogenic pathways. Nat Genet 34(2):226–230.

Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, et al. 2000. Functional discovery via a compendium of expression profiles. Cell 102:109–126.

ICCVAM. 1997. ICCVAM Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publ no 97-3981. Research Triangle Park, NC:Interagency Coordinating Committee on the Validation of Alternative Methods, National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov [accessed 8 July 2005].

ICCVAM. 2003. ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods. NIH Publ no 03-4508. Research Triangle Park, NC:Interagency Coordinating Committee on the Validation of Alternative Methods, National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov [accessed 8 July 2005].

ICCVAM. 2004. ICCVAM Recommended Performance Standards for *In Vitro* Test Methods for Skin Corrosion. NIH Publ no 04-4509. Research Triangle Park, NC:Interagency Coordinating Committee on the Validation of Alternative Methods, National Institute of Environmental Health Sciences. Available: http://iccvam.niehs.nih.gov [accessed 8 July 2005].

Ikeo K, Ishi-i J, Tamura T, Gojobori T, Tateno Y. 2003. CIBEX: Center for Information Biology Gene Expression database. C R Biol 326(10–11):1079–1082.

Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, Frank BC, et al. 2005. Multiple-laboratory comparison of microarray platforms. Nat Methods 2:345–349.

Joseph LJ. 2004. RNA reference materials for gene expression studies. RNA metrology: forecast calls for partial clearing. Clin Chem 50(8):1290–1292.

Kuang Y, Biran I, Walt DR. 2004. Simultaneously monitoring gene expression kinetics and genetic noise in single cells by optical well arrays. Anal Chem 76(21):6282–6286.

Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. 2005. Independence and reproducibility across microarray platforms. Nat Methods 2:337–343.

Mattes WB, Pettit SD, Sansone SA, Bushel PR, Waters MD. 2004. Database development in toxicogenomics: issues and efforts. Environ Health Perspect 112:495–505.

MGED Society (Microarray Gene Expression Data Society). 2004. RSBI Tiered Checklist (RSBI TC). A Proposed Framework Structure for Reporting Biological Investigations. Available: http://www.mged.org/Workgroups/rsbi/rsbi.html [accessed 9 January 2006].

Nature. 2002. Microarray standards at last [Editorial]. Nature 419:323.

OECD. 1998. OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring. 1: OECD Principles on Good Laboratory Practice. Paris:Organisation for Economic Co-operation and Development. http://www.oecd.org/document/63/0,2340,en_2649_34381_2346175_1_1_1_1,00.html [accessed 8 July 2005].

OECD. 2005. OECD Activities to Explore and Evaluate Regulatory Application of Genomic Methods; Toxicogenomics. Paris:Organisation for Economic Co-operation and Development. Available: http://www.oecd.org/document/29/0,2340,en_2649_34373_34704669_1_1_1_1,00.html [accessed 2 February 2006].

Orphanides G. 2003. Toxicogenomics: challenges and opportunities. Toxicol Lett 140–141:145–148.

Rockett JC, Hellmann GM. 2004. Confirming microarray data—is it really necessary? Genomics 83:541–549.

Russell WMS, Burch RL. 1959. The Principles of Humane Experimental Technique. London:Methuen. Available: http://altweb.jhsph.edu/publications/humane_exp/het-toc.htm [accessed 8 July 2005].

Sansone SA, Morrison N, Rocca-Serra P, Fostel J. 2004. Standardization initiatives in the (eco)toxicogenomics domain: a review. Comp Funct Genom 5:633–64.

Scherf U, Ross DT, Waltham M, Smith LH, Lee JK, Tanabe L, et al. 2000. A gene expression database for the molecular pharmacology of cancer. Nat Gen 24:236–244.

Spellman PT, Miller M, Stewart J, Troup C, Sarkans U, Chervitz S, et al. 2002. Design and implementation of microarray gene expression markup language (MAGE-ML). Genome Biol. 3(9):RESEARCH0046. Available: http://genomebiology.com/2002/3/9/research/0046 [accessed 9 January 2006].

Steiner G, Suter L, Boess F, Gasser R, de Vera MC, Albertini S, et al. 2004. Discriminating different classes of toxicants by transcript profiling. Environ Health Perspect 112:1238–1248.

Thomas RS, Rank DR, Penn SG, Zastrow GM, Hayes KR, Pande K, et al. 2001. Identification of toxicologically predictive gene sets using cDNA microarrays. Mol Pharmacol 60(6):1189–1194.

Tong W, Cao X, Harris S, Sun H, Fang H, Fuscoe J, et al. 2003. ArrayTrack—supporting toxicogenomic research at the U.S. Food and Drug Administration National Center for Toxicological Research. Environ Health Perspect 111:1819–1826.

U.S. EPA. 2002. Interim Genomics Policy. U.S. Environmental Protection Agency Science Policy Council. Washington, DC: U.S. Environmental Protection Agency. Available: http://epa.gov/osa/spc/htm/genomics.htm [accessed 8 July 2005].

U.S. FDA. 2005. FDA Guidance for Industry—Pharmacogenomic (PG) Data Submissions (Final: March 2005). Washington, DC:U.S. Food and Drug Administration. Available: www.fda.gov/cder/guidance/6400fnl.htm [accessed 8 July 2005].

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. Nature 415(6871):530–536.

Waring JF, Cavet G, Jolly RA, McDowell J, Dai H, Ciurlionis R, et al. 2003. Development of a DNA microarray for toxicology based on hepatotoxin-regulated sequences. Environ Health Perspect 111:863–870.

Waring JF, Jolly RA, Ciurlionis R, Lum PY, Praestgaard JT, Morfitt DC, et al. 2001. Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. Toxicol Appl Pharmacol 175(1):28–42.

Waters MD, Boorman G, Bushel P, Cunningham M, Irwin R, Merrick A, et al. 2003. Systems toxicology and the chemical effects in biological systems knowledge base. Environ Health Perspect 111:811–824.

Yauk CL, Berndt ML, Williams A, Douglas GR. 2004 Comprehensive comparison of six microarray technologies. Nucleic Acid Res 32(15):e124.

Available online at www.sciencedirect.com
ScienceDirect
Reproductive Toxicology 25 (2008) 21–38
Reproductive Toxicology
Developmental Basis of Health and Disease
www.elsevier.com/locate/reprotox

Two-generation reproductive toxicity study of the rubber accelerator *N,N*-dicyclohexyl-2-benzothiazolesulfenamide in rats

Makoto Ema [a,*], Sakiko Fujii [b], Mariko Matsumoto [a], Mutsuko Hirata-Koizumi [a], Akihiko Hirose [a], Eiichi Kamata [a]

[a] *Division of Risk Assessment, Biological Safety Research Center, National Institute of Health Sciences, Tokyo 158-8501, Japan*
[b] *Safety Research Institute for Chemical Compounds Co. Ltd., Sapporo 004-0839, Japan*

Received 14 June 2007; received in revised form 6 August 2007; accepted 18 October 2007
Available online 25 October 2007

Abstract

Male and female Crl:CD(SD) rats were fed a diet containing rubber accelerator *N,N*-dicyclohexyl-2-benzothiazolesulfenamide (DCBS) at 0, 80, 600 or 4500 ppm throughout the study beginning at the onset of a 10-week pre-mating period and continuing through the mating, gestation, and lactation periods for two generations. At 4500 ppm, decreases in the body weight, body weight gain, and food consumption were found in F0 males and females. No changes in the estrous cyclicity, copulation index, fertility index, gestation index, delivery index, number of implantations, precoital interval, or gestation length were observed in any generation at any dose of DCBS. Delayed preputial separation at 4500 ppm as well as delayed vaginal opening and higher body weight at the age of vaginal opening at 600 and 4500 ppm were found in the F1 generation. A transient change in performance in a water-filled multiple T-maze was found at 600 and 4500 ppm in F1 females. There were no compound-related changes in number of pups delivered, sex ratio of pups, viability of pups, anogenital distance, surface righting reflex, negative geotaxis reflex, mid-air righting reflex, pinna unfolding, incisor eruption, or eye opening in the F1 and F2 generations. The body weight of F1 and F2 male and female pups was lowered at 4500 ppm. Reduced uterine weight of the weanlings was noted in the F1 generation at 4500 ppm and in the F2 generation at 600 and 4500 ppm. The data indicate that the NOAEL of DCBS for two-generation reproductive toxicity is 80 ppm (5.2 mg/kg bw per day) in rats.

Keywords: *N,N*-Dicyclohexyl-2-benzothiazolesulfenamide; Rubber accelerator; Two-generation reproductive toxicity; Developmental toxicity; Rat

1. Introduction

N,N-Dicyclohexyl-2-benzothiazolesulfenamide (DCBS) is a sulfenamide accelerator. The sulfenamide accelerator class of rubber accelerators has been manufactured in the USA for over 60 years [1]. Sulfenamide accelerator compounds are widely used in the manufacture of automotive compartments and industrial rubber products such as tires, hoses, conveyer belts, bushings seals, gaskets and windshield wiper blades, and the typical usage for sulfenamide accelerators is from 0.5 to 4 parts accelerator per every 100 parts of rubber [1]. Sulfenamide accelerator materials are shipped extensively throughout the world from manufacturing plants located in North America, South America, Europe, Asia and Africa [1]. DCBS was produced in Japan with an annual production level of about 1000 tonnes in 1990–1993 and 1900 tons in 2000–2003, and most of this amount was sold and handled domestically [2]. DCBS is used as an accelerator of vulcanization and is completely reacted in the vulcanizing process [2]. DCBS is regulated for use in articles in contact with food in Germany, but this compound is not regulated for use in FDA food contact applications [3]. Exposure of workers handling sulfenamide accelerator materials is likely to be highest in the area of materials packaging. During material packout at the manufacturing site and to a lesser degree during weigh-up activities at the consumer site, there is potential for skin and inhalation exposure. Although consumer exposure would be minimal, the most likely route of consumer exposure is skin contact from rubber or latex articles [1].

Only up to 6% biodegradation for DCBS was determined in a ready biodegradability test, and a measured log K_{ow} value of 4.8 suggests that DCBS may have a high bioaccumulation potential [2]. The possibility of such a chemical compound entering

* Corresponding author. Tel.: +81 3 3700 9878; fax: +81 3 3700 1408.
E-mail address: ema@nihs.go.jp (M. Ema).

0890-6238/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.reprotox.2007.10.006

into biological systems has aroused great concern regarding its toxicological potential. Generally, biological effects produced by chemicals should be studied in laboratory animals to investigate their possible influences on human health, and the results of animal tests of chemical toxicity are relevant to humans [4]. However, very little information on the toxicity of DCBS has been published. Vorobera (1969) [5] reported that the oral LD50 value was 8500 mg/kg bw in male mice and that repeated inhalation exposure of male rats for 15 days, daily, 2 h/day, at 350–400 mg/m^3 caused mucous membrane irritation. Although the toxic effects of DCBS have been briefly summarized by the European Chemical Bureau [6] and EPA [1], descriptions regarding the toxicity of DCBS are insufficient to assess the adverse effects of this compound. The EPA [1] noted that the oral LD50 values were 1077–10000 mg/kg bw in rats, the oral NOAEL for 44-day repeated dose toxicity was higher than 100 mg/kg bw per day in rats, and no effects on reproduction were observed at doses up to 400 mg/kg bw per day in rats. Toxicity studies including acute toxicity, in vitro genotoxicity, and repeat dose toxicity combined with reproductive/developmental toxicity studies of DCBS were performed as a part of the Safety Examination of Existing Chemical Substances and Chemical Safety Programmes by the Japanese Government [7]. These toxicity studies are summarized in the IUCLID Data Sets [8], OECD Screening Information Data Sets [2] and the Hazard Assessment Sheet [9]. We previously reported results of repeat dose toxicity combined with a reproductive/developmental toxicity screening test of DCBS showing that DCBS at 400 mg/kg bw per day possessed a deleterious effect on reproduction and development and caused a marked decrease in the number of live pups as well as a total loss of pups until postnatal day (PND) 4 [10]. The primary effects may be on the gestation index for dams and live birth index for pups, which both appear to be affected at multiple points along the female reproductive process; the viability of neonatal pups may also be affected. The previous study was performed in compliance with the OECD guideline for a Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test [11,12], but this screening test guideline does not provide complete information on all aspects of reproduction and development due to the relatively small numbers of animals in the dose groups and selectivity of the endpoints. In order to further evaluate the reproductive and developmental toxicity of DCBS in rats, a two-generation reproductive toxicity study was conducted. We examined reproductive and developmental endpoints such as sexual development, estrous cyclicity, anogenital distance (AGD), physical and functional development, serum hormone levels, and sperm count and motility.

2. Materials and methods

This study was performed in 2006–2007 at the Safety Research Institute for Chemical Compounds Co. Ltd. (Sapporo, Japan) in compliance with OECD guideline 416 Two-generation Reproduction Toxicity Study [13] and in accordance with the principles for Good Laboratory Practice [14], "Law for the Humane Treatment and Management of Animals" [Law No. 105, 1 October 1973, revised 22 December 1999, Revised Law No. 221; revised 22 June 2005, Revised Law No. 68], "Standards Relating to the Care, Management and Refinement of Laboratory Animals" [Notification No. 88 of the Ministry of the Environment, Japan, 28 April 2006] and "Fundamental Guidelines for Proper Conduct of Animal Experiment and Related Activities in the Testing Facility under the Jurisdiction of the Ministry of Health, Labour and Welfare" [Notification No. 0601005 of the Health Sciences Division, Ministry of Health, Labour and Welfare, Japan, 1 June 2006].

2.1. Chemical and dosing

N,N-Dicyclohexyl-2-benzothiazolesulfenamide (DCBS, CAS No. 4979-32-2) was obtained from Ouchishinko Chemical Industrial Co. Ltd. (Tokyo, Japan). DCBS in the form of off white to tan granules is very slightly soluble in water and methanol but soluble in oil, and its melting point is 100–105 °C, density at 21 °C is 1230 kg/m^3, and molecular weight is 347 [3]. The DCBS (Lot no. 508001) used in this study was 99.7% pure, and it was kept in a sealed container under cool (1–8 °C) and dark conditions. The purity and stability of the chemical were verified by analysis using high-performance liquid chromatography before and after the study. Rats were given dietary DCBS at a concentration of 0 (control), 80, 600 or 4500 ppm. The dosage levels were determined based on the results of our previous dose-finding study in male and female rats fed a diet containing DCBS at 0, 1500, 3000, 6000 or 10,000 ppm (0, 83, 172, 343 or 551 mg/kg bw per day in males and 0, 126, 264, 476 or 707 mg/kg bw per day in females) for a total of eight weeks beginning 16 days before mating in males and a total of nine weeks in females throughout the mating, gestation and lactation periods beginning 16 days before mating. In that study, we found reduced body weight gain in males at 6000 ppm and higher and females at 3000 ppm and higher, reduced number of implantations at 6000 ppm and higher, decreased absolute and relative weight of the spleen in females at 6000 ppm and higher, reduced number of pups born at 10000 ppm, lowered body weight of pups at 6000 ppm and higher, and decreased absolute and relative weight of the spleen in male weanlings at 1500 ppm and higher and female weanlings at 3000 ppm and higher [15]. Dosed diet preparations were formulated by mixing DCBS into an appropriate amount of a powdered basal diet (CRF-1, Oriental Yeast Co. Ltd., Tokyo, Japan) for each dietary concentration. The control rats were fed a basal diet only. Analysis showed that DCBS was homogeneous in the diet and stable for at least 21 days in a room temperature, and formulations were maintained in a room temperature for no more than 21 days. Generally, diet was replaced every 1 week.

2.2. Animals and housing conditions

Crl:CD(SD) rats were used throughout this study. Rats of this strain were chosen because they are the most commonly used in reproductive and developmental toxicity studies and historical control data are available. Male and female rats at 4 weeks of age were purchased from Hino Breeding Center, Charles River Laboratories Japan, Inc. (Yokohama, Japan). The males and females were acclimated to the laboratory for eight days prior to the start of the experiment. Male and female rats found to be in good health were selected for use. One hundred and ninety two rats were randomly assigned 24/sex/group and all animals were assigned a unique number and ear tattooed prior to the start of the experiment. Animals were housed individually in suspended aluminium/stainless steel cages except during the acclimation, mating and nursing periods. From day 17 of pregnancy to the day of weaning, individual dams and litters were reared using wood chips as bedding (White Flake; Charles River Laboratories Japan, Inc.).

Animals were reared on a basal diet or diet containing DCBS and filtered tap water *ad libitum* and maintained in an air-conditioned room at 22 ± 3 °C, with a humidity of 50 ± 20%, a 12-h light (8:00–20:00)/dark (20:00–8:00) cycle, and ventilation at 10–15 times/h.

2.3. Experimental design

Twenty-four rats (5-week-old males and females)/sex/group were fed a diet containing DCBS at 0, 80, 600 or 4500 ppm for 10 weeks prior to the mating period. Each female F0 rat was mated with a male rat of the same dosage group, with administration of DCBS in the diet continuing throughout the mating period. Administration of DCBS was continued throughout gestation and lactation. Twenty-four male and 24 female F1 weanlings (1 male and 1 female

in each litter) in each group were selected as F1 parents on PNDs 21–25 to equalize the body weights among groups. The day on which F1 parental animals were selected was designated as 0 week of dosing for the F1 generation. The administration of DCBS in the diet was not suspended during PNDs 21–25. F1-selected rats were administered DCBS in the diet with the respective formulation for 10 weeks prior to the mating period and mated as described above. Administration of DCBS in the diet was continued throughout the mating, gestation, and lactation periods. On PND 26, F1 weanlings not selected for breeding and all F2 weanlings were necropsied.

2.4. Mating procedures

Each female was mated with a single male of the same dosage group until copulation occurred or the mating period had elapsed. The mating periods for F0 and F1 animals were three weeks. During the mating period, daily vaginal smears were examined for the presence of sperm. The presence of sperm in the vaginal smear and/or a vaginal plug was considered evidence for successful mating. The day of successful mating was designated as day 0 of pregnancy. F1 females that did not mate during the 3-week mating period were cohabited with other males from the same group who had been proven to copulate. For F1 matings, cohabitation of siblings was avoided.

2.5. Parental data

All adult rats were observed twice a day for clinical signs of toxicity, and body weights and food consumption were recorded weekly. For females exhibiting evidence of successful mating, body weight and food consumption were recorded on days 0, 7, 14, and 20 of pregnancy and days 0, 4, 7, 14, and 21 of lactation. Daily vaginal lavage samples of each F0 and F1 female were evaluated for estrous cyclicity throughout the 2-week pre-cohabitation period and during cohabitation until evidence of copulation was detected. Females having repeated 4–6 day estrous cycles were judged to have normal estrous cycles. After weaning of their pups, parental female rats were necropsied at the proestrous stage of the estrous cycle. For each female, the number of uterine implantation sites was recorded.

2.6. Litter data

Once insemination was confirmed, female rats were checked at least three times daily at days 21–25 of pregnancy to determine the time of delivery. The females were allowed to deliver spontaneously and nurse their pups until PND 21 (the day of weaning). The day on which parturition was completed by 13:00 was designated as PND 0. Total litter size and the numbers of live and dead pups were recorded, and live pups were counted, sexed, examined grossly, and individually weighed on PNDs 0, 4, 7, 14, and 21. On PND 4, litters were randomly adjusted to eight pups comprising of four males and four females. No adjustment was made for litters of fewer than eight pups. Selected pups were assigned a unique number and limb tattooed on PND 4.

2.7. Developmental landmarks

All F1 and F2 pups were observed daily for pinna unfolding on PNDs 1–4, incisor eruption beginning on PND 8, and eye opening beginning on PND 12. One male and one female F1 and F2 pup selected from each dam was evaluated for the surface righting reflex on PND 5, negative geotaxis reflex on PND 8, and mid-air righting reflex on PND 18 [16]. All F1 offspring were observed daily for male preputial separation beginning on PND 35 or female vaginal opening beginning on PND 25. Body weight of the respective F1 rats was recorded on the day of preputial separation or vaginal opening. The AGD was measured using calipers on PND 4 in all F1 and F2 pups, and the AGD per cube root of body weight ratio was calculated [17].

2.8. Behavioral tests

Spontaneous locomotor activity was measured with a multi-channel activity monitoring system (Supermex; Muromachi Kikai Co. Ltd., Tokyo, Japan) in 10 male and 10 female F1 rats selected from each group at 4 weeks of age. Rats were placed individually in transparent polycarbonate cages ($27.6W \times 44.5D \times 20.4H$ cm, CL-0108-1, Clea Japan Inc., Tokyo, Japan) under an infrared sensor that detects thermal radiation from animals. Spontaneous motor activity was determined for 10 min intervals and for a total of 60 min.

A test in a water-filled multiple T-maze was conducted in 10 male and 10 female F1 rats selected from each group at 6 weeks of age. The apparatus was similar to that described by Biel [18]. The water temperature of the maze was kept 22–23 °C. As a preliminary swimming ability test, each rat was allowed to swim three times in a straight channel on the day before the maze trial, and then tested in the maze with three trials per day for the next consecutive three days. The elapsed time between entry into the water at the starting point and touching the goal ramp as well as the number of errors were recorded. To prevent exhaustion of the rats, no animal was allowed to remain in the water for more than 3 min in any trial.

2.9. Termination/necropsy-adults

Parental rats were necropsied: males after the parturition of paired female, and females after weaning of their pups. Ages on the day of the scheduled terminal sacrifice were 19–20 weeks old in F0 males, 21–22 weeks old in F0 females, 18 weeks old in F1 males and 19–20 weeks old in F1 females. The proestrous stage of the estrous cycle was characterized by examination of the vaginal smears of female rats on the day of necropsy. A complete necropsy was performed on all rats found dead and those killed at the scheduled terminal sacrifice. Live rats were euthanized by exsanguination under ether anesthesia. The external surfaces of the rats were examined. The abdomen and thoracic cavities were opened, and a gross internal examination was performed. Weights of the brain, pituitary, thyroid, thymus, liver, kidney, spleen, adrenal, testis, epididymis, seminal vesicle (with coagulating glands and their fluids), ventral prostate, uterus and ovary were recorded. Weights of the thyroid and seminal vesicle were measured after fixation. Major organs were stored in 10% neutral buffered formalin. The testis and epididymis were fixed with Bouin's solution and preserved in 70% ethanol.

Histopathological evaluations in F0 and F1 adults were performed on the tissues specified below after fixation, paraffin embedding, and sectioning and staining with hematoxylin and eosin: the liver, pituitary, thymus, thyroid, kidney, spleen, adrenal, bone marrow, mesenteric lymph node, Peyer's patches, testis, epididymis, seminal vesicle, coagulating gland, ventral prostate, ovary, uterus, vagina and mammary gland of all males and females in the control and highest dose (4500 ppm) groups and of females with abnormal estrous cycles, of males and females without evidence of copulation or insemination and of females with abnormal delivery or totally dead pups in all groups. Any organs or tissues of F0 and F1 adults showing gross alterations were evaluated histopathologically.

In ten each F1 females of the control and highest dose groups, the primordial follicles were counted [19]. The right ovary was fixed in 10% neutral buffered formalin and then dehydrated and embedded in paraffin in a longitudinal orientation by routine procedures. Sections were cut serially at 5 μm and every 20th one was serially mounted on slides and stained with hematoxylin and eosin. About 40 sections per ovary were used to determine the primordial follicles.

2.10. Termination/necropsy-pups

Following adjustment of litter size on PND 4, culled pups were euthanized by inhalation of carbon dioxide and subjected to a gross external and internal necropsy. No tissues from these pups were collected.

The weanlings not selected to become parents were euthanized and necropsied as described for the adults. Organ weights of one male and one female F1 and F2 weanling selected from each dam was measured as described above for adults. The weights of the pituitary and thyroid were not determined in weanlings. All pups found dead before weaning were also necropsied.

In all male and female F1 and F2 weanlings whose organs were collected, histopathological evaluations of the thymus, liver and spleen in the control and 4500 ppm groups were performed after fixation, paraffin embedding, and sectioning and staining with hematoxylin and eosin.

2.11. Hematological and blood biochemical parameters

On the day of the scheduled terminal sacrifice, blood samples were collected from the abdominal aorta of adult rats under ether anesthesia.

Hematological examinations were performed for 10 males and 10 females of F0 and F1 rats randomly selected from each group. Blood samples were analyzed for the following hematological parameters, using 2K-EDTA as an anticoagulant: white blood cell count (WBC) and differential leukocyte count.

Blood biochemical evaluations were performed for 10 males and 10 females of F0 and F1 rats randomly selected from each group. Serum samples obtained from centrifuged whole blood were analyzed for biochemistry parameters such as total protein, albumin and globulin.

2.12. Serum hormone levels

On the day of the scheduled terminal sacrifice, blood samples were collected from the abdominal aorta of adult rats. Eight males and eight proestrous females of the F0 and F1 generations from each group were selected randomly for blood collection. Hormone levels were determined by Panapharm Laboratories Co. Ltd. (Uto, Japan). Serum levels of testosterone, 5α-dihydrotestosterone (DHT), luteinizing hormone (LH), and follicle stimulating hormone (FSH) in males, and estradiol, progesterone, LH, and FSH in females were measured. The testosterone, DHT, estradiol, and progesterone concentrations were measured using a double antibody kit (Diagnostic Products Corp., Los Angeles, CA or Diagnostic Systems Laboratories Inc., Webster, TX). Serum concentrations of LH and FSH were measured using (rat LH)[125I] and (rat FSH) [125I] assay systems (GE Healthcare Bio-Sciences Corp., Piscataway, NJ), respectively.

2.13. Sperm parameters

Sperm parameters were determined for all F0 and F1 male adults, except dead males, on the day of the scheduled terminal sacrifice. The right testis was used to count testicular homogenization-resistant spermatid heads. The right cauda epididymis was weighed and used for sperm analysis. Sperm motility was analyzed using a computer-assisted cell motion analyzer (TOX IVOS, Hamilton Thorne Biosciences, Beverly, MA). The percentage of motile sperm and progressively motile sperm as well as their swimming speed and pattern were determined. After the recording of sperm motion, the cauda epididymal fluid was diluted and sperm were enumerated using a hemacytometer under a light microscope. A sperm count per gram of epididymal tissue was obtained by dividing the total count by the gram weight of the cauda epididymis. The sperm were stained with eosin and mounted on a slide glass. Two hundred sperm in each sample were examined under a light microscope, and the percentage of morphologically abnormal sperm was calculated.

2.14. Statistical analysis

Statistical analysis of offspring before weaning was carried out using the litter as the experimental unit.

Body weight, body weight gain, food consumption, length of estrous cycle, precoital interval, gestation length, numbers of implantations and pups delivered, delivery index, sperm parameters, hematological and blood chemical parameters, hormone levels, organ weight, organ/body weight ratio (relative organ weight), reflex response time, age displayed pinna unfolding, incisor eruption, and eye opening, age and body weight at sexual maturation, parameters of behavioral tests, AGD, AGD/cube root of body weight ratio, and the viability of pups were analyzed for statistical significance in the following way. Bartlett's test of homogeneity of variance was used to determine if the groups had equivalent variances. If the variances were equivalent, the groups were compared by one-way analysis of variance (ANOVA). If significant differences were found, Dunnett's multiple comparison test was performed. If the groups did not have equivalent variances, the Kruskal–Wallis test was used to assess the overall effects. Whenever significant differences were noted, pairwise comparisons were made by Mann–Whitney *U*-test. The incidence of pups with changes in clinical and gross internal observations, and reflex completion rate of pups were analyzed by Wilcoxon rank sum test. The number of primordial follicles in the control and highest dose groups was analyzed in the following way. Variance ratio was analyzed by *F*-test. Since the variance ratio was equivalent, the groups were compared by Student's *t*-test. The incidence of females with normal estrous cycles, copulation index, fertility index, gestation index, neonatal sex ratio, and completion rate of the reflex response were analyzed by Fisher's exact test.

The 0.05 level of probability was used as the criterion for significance.

3. Results

3.1. Clinical observations, body weight and food consumption during the pre-mating, mating, gestation, and lactation periods (F0 and F1)

There were no compound-related clinical signs of toxicity in either male or female F0 and F1 rats during the pre-mating, mating, gestation, or lactation periods. One F0 male at 80 ppm was euthanized in 11 weeks of dosing because of a moribund condition resulting from accidental injury in the home cage. One F1 female without any apparent clinical signs of toxicity died on day 5 of lactation in the control group, and no abnormal necropsy findings were found.

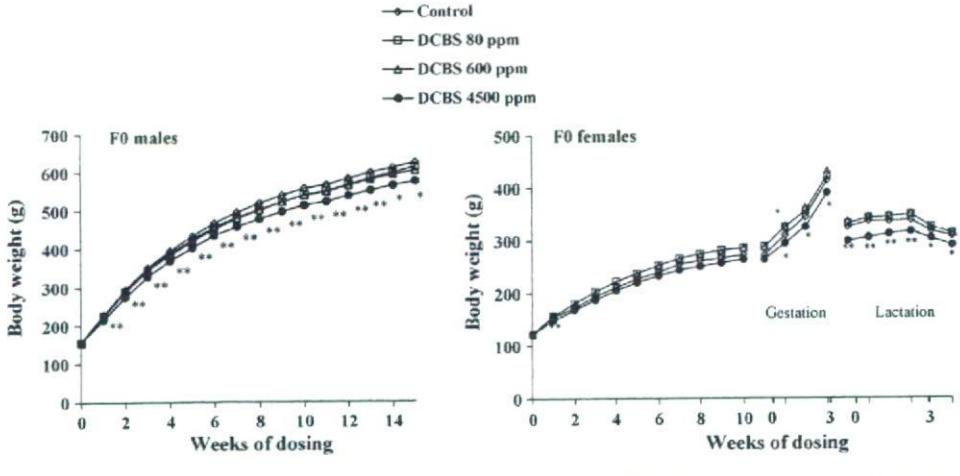

Fig. 1. Body weight of F0 males and females. *Significantly different from the control, $p < 0.05$. **Significantly different from the control, $p < 0.01$.