

between the h-Trx-Tg mice and Wt mice (right). In contrast to data shown in Figure 2A, the percentage of CFU-GMs in the Trx<sup>+/-</sup> mice more significantly decreased than that in the Wt mice (hatched columns) (62.4% vs 74.4%,  $p < 0.05$ ; with respect to each sham-exposed group), whereas that in the h-Trx-Tg mice did not decrease at all compared with that in the Wt mice (hatched columns) (100% vs 74.4%,  $p < 0.05$ ; with respect to each sham-exposed group).

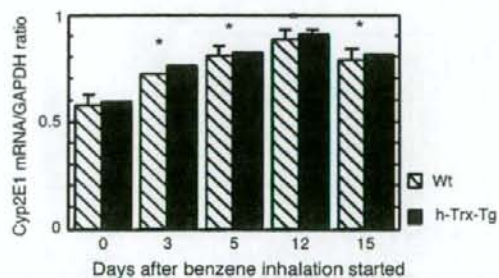
#### CYP2E1 gene expression in BM

The oxidation of benzene metabolites is mediated by CYP2E1, which is a prerequisite for cellular toxicity caused by benzene exposure [35,36]. Whether AhR induction modulates the expression of CYP2E1 remain unclarified [28]. To evaluate any differences in CYP2E1 expression between the Wt and Tg mice exposed to benzene at 300 ppm for 6 hours/day, 5 days/week, for 2 weeks, the expression level of CYP2E1 mRNA was determined by RT-PCR analysis (Fig. 3). There were no significant differences in expression level between the two genotypes. Thus, the induction of hematotoxic benzene metabolites by the increase in CYP2E1 mRNA expression level with time after benzene inhalation started regardless of the genotype, that is, Wt mice or h-Trx-Tg mice. Furthermore, Trx overexpression is not expected to play a role in benzene metabolism in relation to CYP2E1 expression in BM.

#### Expression of transgenic

##### h-Trx and endogenous m-Trx genes

Because the transgene was constructed such that it was constitutively expressed, the expression level of h-Trx in the Tg

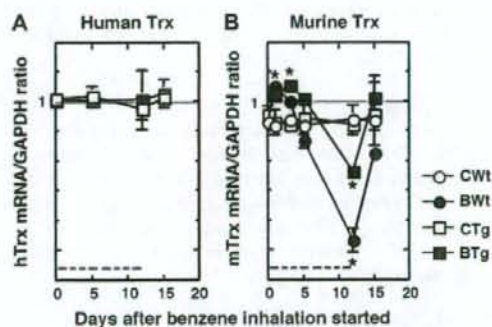


**Figure 3.** Total RNA was extracted from BM cells from three mice of each genotype at each time point. The expression levels of CYP2E1 mRNA during benzene exposure in the Wt mice and h-Trx-Tg mice were determined by RT-PCR analysis. The ordinate (y-axis) shows the ratio of mRNA expression level of CYP2E1 to that of GAPDH on 0, 3, 5, and 12 days of benzene exposure, and 15 days, that is, 12 days of exposure plus 3 days recovery, in the Wt and h-Trx-Tg mice. Solid columns indicate the expression level in h-Trx-Tg mice, and hatched columns, in the Wt mice. Continuous and statistically significant increases in CYP2E1 mRNA expression level from day 0 to day 12 were observed in both Wt mice and h-Trx-Tg mice, although there were no significant differences in the level between Wt mice and h-Trx-Tg mice. Asterisks indicate statistically significant differences with respect to the steady state, that is, day 0.

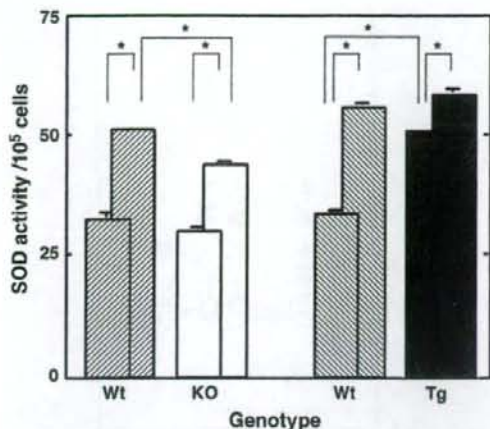
mice was stable during the steady state, as shown in Figure 4A (open squares), and remained stable during and after the benzene inhalation, as shown in Figure 4A (closed squares). During benzene inhalation, m-Trx gene expression level, on the other hand, increased transiently, then decreased until day 12, and started to return to the baseline after the cessation of benzene exposure in the Wt mice (Fig. 4B). During benzene exposure, h-Trx-Tg mice showed a temporary increase in the expression level of the endogenous m-Trx gene on day 3 post benzene inhalation as compared with the Wt mice on day 1, and then showed a slight decrease by day 12 as compared with the Wt mice, which showed a marked decrease. Thus, changes in endogenous m-Trx expression level in the h-Trx-Tg mice were more attenuated than those in the Wt mice during benzene exposure. Moreover, the total Trx activity normalized by GAPDH expression level in the h-Trx-Tg mice during the steady state was nearly twice that in the Wt mice, because both hTrx and mTrx activities on day 0 in the h-Trx-Tg mice were 0.97 and 0.95, respectively.

#### Induction of SOD activity by benzene inhalation in BM

Because the expression level of endogenous m-Trx decreased with benzene exposure duration and it was also shown that the expression of SOD was upregulated by Trx in an in vitro study [37], it is of interest to determine SOD activity in the steady state and during benzene exposure. Figure 5 shows SOD activities in the Trx<sup>+/-</sup> mice (left) and Tg mice (right) with respect to those in the counterpart Wt mice (front columns are for the steady state and back columns are for the benzene exposure). Analogous to the upregulation of SOD activity observed in vitro



**Figure 4.** h-Trx (A) and m-Trx (B) gene expressions during benzene exposure. Total RNA was extracted from BM cells from three mice of each genotype at each time point. The y-axis indicates the ratio of mRNA expression level of h-Trx to that of GAPDH (A) and that of m-Trx to that of GAPDH (B), vs the days after benzene exposure indicated by a dotted line at the bottom. Benzene-exposed groups (closed squares, h-Trx-Tg mice; BTg; closed circles, Wt mice; BWt) and sham-exposed groups (open squares, h-Trx-Tg mice; CTg; open circles, Wt mice; CWt). Asterisks indicate a significant difference between the benzene-exposed and sham-exposed groups ( $p < 0.05$ ).



**Figure 5.** Superoxide dismutase (SOD) activity in BM. During the steady state (left, front columns) and benzene exposure (right, back columns), SOD activities in three mice of each of *Trx*<sup>+/-</sup> mice (open columns) and h-Trx-Tg mice (closed column) are compared with those in the Wt mice (hatched columns). A significant upregulation of SOD activity was observed in the h-Trx-Tg mice compared with Wt in the steady state (\*:  $p < 0.05$ ), whereas all the benzene-exposed groups showed a significant increase in SOD activity compared with their counterparts in the steady state (\*:  $p < 0.05$ ). Furthermore, the relative increase in SOD activity in the Wt mice is significantly larger than that in the *Trx*<sup>+/-</sup> mice. Note that the upregulation of SOD activity in the benzene-exposed *Trx*<sup>+/-</sup> mice is limited compared with that in the Wt mice. Moreover, the percent upregulations of SOD activity in the steady state and during benzene exposure in the h-Trx-Tg mice are lower than those in the Wt mice (113.3% vs 166.3%).

attributable to the upregulation of *Trx* [37], we observed this upregulation in the h-Trx-Tg mice at the in vivo level in the steady state as compared with the Wt control (Fig. 5, front columns; Tg, far right, 51.4% vs Wt, second from right, 33.4% inhibition of NTB). During benzene exposure (shown by back columns), SOD activity increased as compared with the steady state, both in the Wt and h-Trx-Tg mice, although the ratio of increase relative to the steady state was larger in the Wt mice than in the h-Trx-Tg mice (Fig. 5, right four columns). Such SOD activity induction by benzene inhalation was observed but to a significantly lesser extent when endogenous *Trx* was deficient (Fig. 5, back columns; KO, second from left, 43.0% vs Wt, far left, 51.0% inhibition of NTB).

#### Evaluation of ROS amount in tissue during benzene inhalation

Various defense mechanisms against ROS function together, including those involving enzymes such as *Trx* and SOD; thus, the total amount of ROS accumulated in tissues needs to be evaluated in an in vivo system. We evaluated the intracellular redox status using an oxidation-sensitive dye, DCFH-DA. A significant increase in ROS

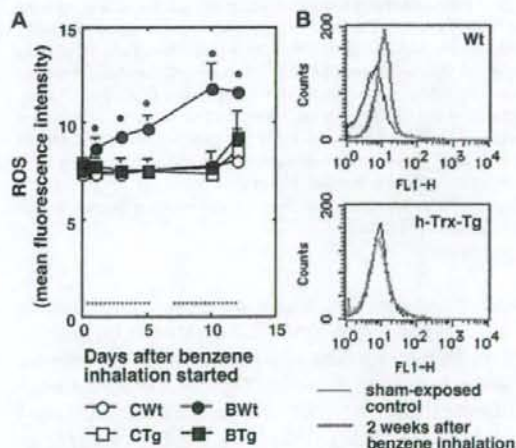
generation was observed in Wt mice erythrocytes but not in h-Trx-Tg mice erythrocytes, as shown in Figure 6.

#### Attenuation of cell cycle-related gene expression, *p21*, by *Trx*

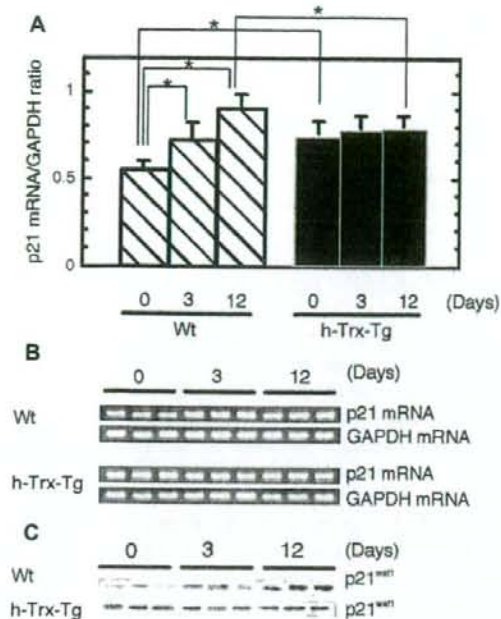
During benzene exposure, the upregulation of *p21*, a cyclin-dependent kinase inhibitor, followed by *p53* activation suppresses the cell cycle of hemopoietic progenitors, which is known to lead to the repair of damaged DNA [23,38]. However, in mice overexpressing h-Trx, the expression levels of both *p21* mRNA and its protein did not increase, but rather remained the same. This implies that the repair of DNA damage caused by benzene exposure does not require *p21* expression induction in the h-Trx-Tg mice. On the other hand, the expression levels of *p21* mRNA and its protein in Wt mice showed a continuous and statistically significant increase with exposure duration as reported previously (Fig. 7) [23]. In the steady state, on the other hand, it is of particular interest that the expression level of *p21* mRNA was higher in the h-Trx-Tg mice than in the Wt mice, which may be due to the observed enhancement of *p53* transcription by *Trx*, resulting in *p21* overexpression in the steady state [39].

#### Decrease in incidence of micronucleated reticulocytes by *Trx* overexpression

In Figure 8, the incidence of micronucleated reticulocytes (MN-Rets) is shown in percentage along with the duration

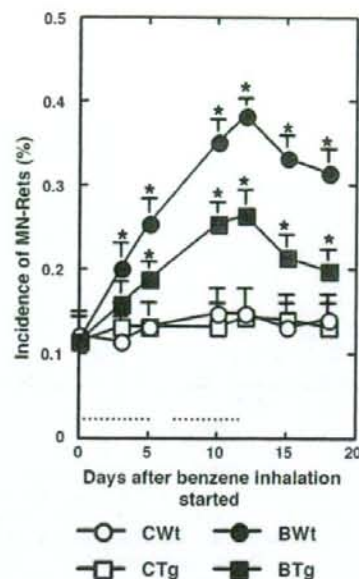


**Figure 6.** ROS in peripheral RBCs collected from the same three mice of each genotype during benzene inhalation. The total amount of ROS accumulated in RBCs was measured by labeling with DCFH-DA in an in vivo system (A) (see Materials and methods). Sample data obtained on day 12, Wt mice on the top and h-Trx-Tg at the bottom (B). With benzene exposure, fluorescence intensity (F.I.) for the BWt mice (closed circles) steadily increased until day 12, whereas F.I. for all other groups remained unchanged. All the data points for the BWt mice showed statistically significant increases as compared with those for the sham-exposed control CWt (\*:  $p < 0.05$ ).



**Figure 7.** (A,B) Expression levels of p21 mRNA in Wt mice and h-Trx-Tg mice during benzene exposure (B) and their relative signal intensities (A). The y-axis shows the ratio of mRNA expression level of p21 to that of GAPDH 0, 3, and 12 days after benzene exposure started in the Wt and h-Trx-Tg mice. A continuous and statistically significant increase in p21 mRNA expression level from day 0 to day 12 was observed in the Wt mice ( $p < 0.05$ ), whereas a less significant stable expression was observed in the h-Trx-Tg mice. The level of p21 mRNA expression in h-Trx-Tg mice was higher than that in the Wt mice at the steady state, which may be explained by the observation that *Trx* enhanced the transcription of p53, resulting in this overexpression at the steady state (A,B). Furthermore, the attenuation of DNA damage caused by benzene exposure did not seem to induce p21 expression in the h-Trx-Tg mice, although p21 mRNA expression level in Wt mice continuously and statistically significantly increased with exposure duration as reported previously (\*:  $p < 0.05$ ). (C) Comparable protein expression levels are observed at the steady state and during benzene exposure.

of benzene exposure. The mice were exposed to benzene at a dose of 300 ppm for 6 hours/day, 5 days/week, for 2 weeks as shown in the exposure schedule indicated by a dotted line at the bottom of Figure 8. On day 12, all the mice were removed from the inhalation chamber and maintained without benzene inhalation until day 18. The incidence of MN-Rets in the Wt group, indicated by closed circles (top), increased with the number of days of exposure to benzene until day 12. It then started to return to the baseline after the termination of benzene exposure. Interestingly, the h-Trx-Tg mice (closed squares) showed a lower incidence of about 52% than the Wt group on day 12, as shown by a lower peak (0.12% and 0.23% increases in sham-exposed h-Trx-Tg and Wt mice, respectively), which then started to decrease after the termination of benzene exposure. Open

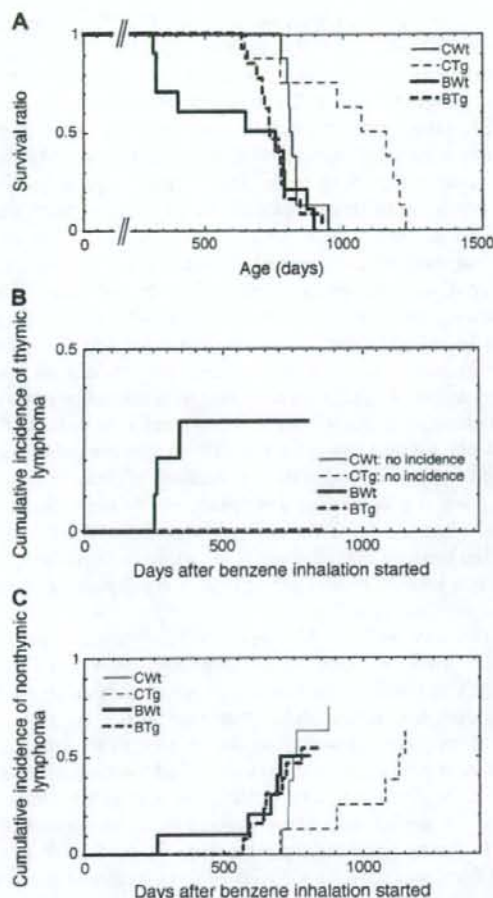


**Figure 8.** Time course of sequential changes in incidence of MN-Rets. Closed symbols indicate mice exposed to benzene at 300 ppm for 6 h/d, 5 d/wk, for 2 weeks as indicated by a dotted line at the bottom of the graph. After day 12, all the mice were removed from the inhalation chambers until day 18 and compared with sham-exposed groups (open symbols). Circles and squares indicate Wt groups and h-Trx-Tg groups, respectively. Both the benzene-exposed groups, BWt and BTg, showed significantly higher incidences of MN-Rets than the sham-exposed groups, CWt and CTg, 3, 5, 10, 12, 15, and 18 days after starting benzene exposure (asterisks on the closed circle, BWt:  $p < 0.05$ ). The BWt mice (top, closed circles) showed a higher incidence (%) of MN-Rets, whereas the BTg mice (second from top, closed squares) showed about 48% decrease throughout the observation period. A significantly lower MN-Ret incidence in benzene-exposed BTg than in BWt on days 5, 10, 12, 15, and 18 (asterisks on the closed squares, BTg:  $p < 0.05$ ) was observed.

squares and circles represent sham-exposed Wt and h-Trx-Tg groups, respectively.

#### Lifetime incidence of lympho-hematopoietic malignancies

To investigate a possible inhibition of the development of benzene-induced lympho-hematopoietic neoplasms after benzene inhalation, a lifetime bioassay of h-Trx-Tg mice exposed to intermittent benzene inhalation at 300 ppm for 26 weeks was conducted in comparison with Wt mice. The survival curves are shown in Figure 9A. The curve of sham-exposed h-Trx-Tg mice shows a survival of 1111.5 days; that is, 298 days longer than the median lifetime as compared with that in Wt mice (thinner dotted line vs thinner solid line, respectively). In the groups exposed to benzene, when the survival curve of h-Trx-Tg mice (bold dotted line) is compared with that of the Wt mice, the former again shows a longer survival curve, although in the



**Figure 9.** For lymphomagenicity studies, the numbers of mice per group, namely, Wt mice with (designated as B in the figure) or without (designated as C in the figure) benzene exposure, and Tg mice with or without benzene exposure, were 10 BWt mice, 8 CWt mice, 13 BTg mice, and 8 CTg mice. Survival curve (A) and incidences of thymic lymphoma (B) and nonthymic lymphoma (C) are shown. Thin solid line, sham-exposed Wt group (CWt); bold solid line, benzene-exposed Wt group (BWt); thin dotted line, sham-exposed Trx group (CTg); and bold dotted line, benzene-exposed Trx group (BTg). No thymic lymphomas were observed in the CWt, CTg, and BTg groups.

later part, after 750 days of age, the survival curves of the h-Trx-Tg mice and Wt mice are not significantly different from each other. As shown in Figure 9B and C, benzene exposure at 300 ppm for 26 weeks induced hemopoietic malignancies, namely, thymic lymphomas (Fig. 9B) and nonthymic lymphomas (Fig. 9C). Whereas the cumulative incidence of thymic lymphomas reached the maximum of 30% in the Wt mice by day 350, surprisingly, no thymic lymphomas developed in the h-Trx-Tg group (Fig. 9B). Figure 9C, on the other hand, shows comparable cumulative

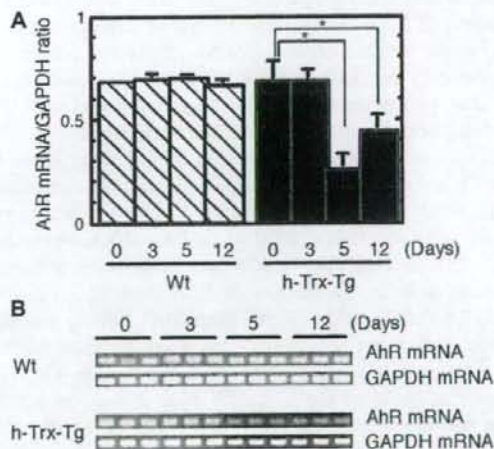
incidences of benzene-induced nonthymic lymphomas (bold dotted line vs bold solid line, respectively), between the h-Trx-Tg mice (53.8%) and the Wt mice (50.0%), whereas the sham-exposed groups of h-Trx-Tg mice and Wt mice show slightly higher incidences than the benzene-exposed groups (62.5% and 75.0%, respectively) with identical onsets but a much larger time lag (thinner dotted line vs thinner solid line, respectively).

#### AhR mRNA expression in BM

Our previous study showed that benzene toxicity is transmitted by AhR [28]. It is of interest to examine AhR gene expression in the h-Trx-Tg mice after benzene inhalation. The expression level of AhR mRNA significantly decreased to one-half of the steady-state level on days 5 and 12 of benzene inhalation only in the h-Trx-Tg mice (Fig. 10A, right, and B, bottom), whereas there were no significant differences in AhR mRNA expression level at the steady state between the Wt mice and h-Trx-Tg mice (Fig. 10A, left, and B, top). Furthermore, the expression level of AhR mRNA in the Wt mice did not change throughout the exposure period (Fig. 10A, left, and B, top).

#### Discussion

The thioredoxin gene-modified mice,  $Trx^{+/-}$  mice and h-Trx-Tg mice, responded to ROS in a *Trx* gene-dose-dependent manner not only in mature BM cells but also in progenitor cells, CFU-GMs (Fig. 2B). The exacerbation



**Figure 10.** (A,B) Expression levels of AhR mRNA in Wt mice and h-Trx-Tg mice during benzene exposure (A), and their relative signal intensities (B). The y-axis shows the ratio of mRNA expression level of AhR to that of GAPDH 0, 3, 5, and 12 days after benzene exposure started in the Wt and h-Trx-Tg mice. The expression of AhR mRNA in the h-Trx-Tg mice was downregulated after benzene exposure started, and the expression levels on days 5 and 12 significantly decreased to 35.9% and 63.4% of those in the steady state, respectively (\*:  $p < 0.05$ ).

of benzene-induced hemato-lymphoid toxicity in the Trx<sup>+/-</sup> mice in combination with the attenuation of that in the h-Trx-Tg mice has confirmed for the first time at the whole-animal level that the hemato-lymphoid toxicity of benzene is caused by oxidative stress from ROS (Fig. 2B). These correlations were not observed for RBCs (Fig. 2A), because of the very short period (2 weeks) of benzene exposure for RBC turnover to induce anemia in peripheral blood. For WBCs and BM cellularity (BMC), significant correlations were not observed either, but because of heterozygosity (see Materials and methods; homozygosity is fatal) in giving rise to a knockout phenotype, although attenuation in hemizygous overexpression (h-Trx-Tg) was significant for both blood parameters.

Whether the different benzene toxicities observed in h-Trx-Tg and Wt mice may have been caused by a difference in ROS generation or a difference in repair potential between h-Trx-Tg and Wt mice, may be clarified by examining the expression level of CYP2E1 in the h-Trx-Tg and Wt mice. The identical CYP2E1 expression levels not only in the Wt mice but also in h-Trx-Tg mice suggest a possible induction of ROS even in h-Trx-Tg mice (Fig. 3).

However, interestingly, ROS level evaluated using DCFH-DA did not increase at all in the h-Trx-Tg mice as shown in Figure 6 (BTg, closed squares; benzene-exposed Tg mice). *How were generated ROS removed in the h-Trx-Tg mice?* The major factor for ROS scavenging in the h-Trx-Tg mice appears to be transgenic h-Trx, which is constitutively expressed throughout the benzene exposure period as shown in Figure 4A. In the h-Trx-Tg mice, SOD activity was upregulated owing to the constitutive expression of h-Trx even without benzene exposure, and in the case of benzene exposure, SOD activity was further upregulated to the same level as that in the Wt mice (Fig. 5, Tg, dark pair columns on the far right compared with Wt, shaded pair columns on the second from far right). The reasons for the limited upregulation of SOD activity in the h-Trx-Tg mice with benzene exposure may be as follows. First, SOD activity was evaluated on day 12, i.e., corresponding to the lowest level of m-Trx-mRNA expression (Fig. 4B). Second, there could be a sufficient scavenging capacity in h-Trx-Tg mice when ROS level was evaluated using DCFH-DA (Fig. 6). A higher SOD activity was observed during the steady state of h-Trx-Tg than the Wt mice, because the total Trx activity normalized by GAPDH expression level in the h-Trx-Tg mice during the steady state was nearly twice (hTrx + mTrx) that in the Wt mice (0.96 + 1.01 vs 0.93, respectively). This may agree with the extent of upregulation of SOD activity in the h-Trx-Tg mice during the steady state (51.4%).

On the other hand, ROS production continued during benzene exposure in the Wt mice (Fig. 6). During this period, de novo m-Trx was induced initially on day 1 and then its level decreased even during the second week of benzene exposure, followed by a rapid return to the base-

line after the termination of exposure (Fig. 4B, closed circles). This trend of m-Trx production in the Wt mice was attributable to the increase in SOD activity in the steady state compared with SOD activity during benzene exposure of mice despite the rapid downmodulation of Trx from the steady-state level. Because SOD activity is induced by Trx, the induction of SOD in the Trx<sup>+/-</sup> mice is enhanced significantly by benzene exposure but to a lesser extent than that in the steady state, which is significantly lower than that in the benzene-exposed Wt mice (Fig. 5, open back column vs corresponding shaded back column). (Note that the Trx<sup>+/-</sup> mice showed an expression level 50% of that of m-Trx owing to hemizygosity; see Fig. 1). When one examines the ROS level of benzene-exposed Wt mice (Fig. 6, closed circles), it increases with time after benzene inhalation, indicating that the above-mentioned de novo Trx and SOD inductions were not effective in removing such ROS. Despite the complete elimination of ROS in h-Trx-Tg mice, it is interesting that milder and delayed responses were observed in m-Trx expression in the h-Trx-Tg mice during benzene inhalation (Fig. 4B, closed squares), implying that local ROS may not be readily eliminated, but there may be some time lag.

The scavenging of ROS in h-Trx-Tg mice was supposed not to show any increase in the expression level of p21-mRNA throughout the benzene exposure period, although the level was rather higher than that in the Wt mice in the steady state. This may be due, in part, to the upregulation of the transcriptional activity of p53 known to be induced by Trx [39], contributing to a possible efficient repair of steady-state DNA damage by a decelerated cell cycle during benzene exposure (Fig. 7). In the Wt mice, p21 level was upregulated with benzene exposure duration as compared with the level of that in the steady state. This is consistent with a previous study revealing the expression of p21 in the Wt mice followed by the upregulation of p53, which is, vice versa, nullified in the p53 KO mice [23]. The expression level of p21 on day 12 in the Wt mice was higher than that in the h-Trx-Tg mice, implying that DNA damage caused by accumulated ROS is milder in the h-Trx-Tg mice.

In the present study, it was surprising that the incidence of thymic lymphoma was 0% in the h-Trx-Tg mice (Fig. 9B). Nonthymic lymphoma development, on the other hand, was not attenuated (Fig. 9C), which does not necessarily suggest that the lymphomagenesis of nonthymic lymphomas is due to ROS. The nontreated control mice are genetically predisposed to developing nonthymic lymphomas as they age. Regardless of the genotype, that is, Wt or Trx-Tg mice, this process may be accelerated in benzene-treated animals in an h-Trx-independent manner; but this plausibility cannot be confirmed in the present study. Also, whether significant lymphoid cell depletion after benzene exposure and its attenuation in Trx-Tg mice has any possible relevance to the incidence of lymphomagenesis

and its attenuation in Trx-Tg mice cannot be clarified in the present study.

Because thymic lymphoma generally does not develop spontaneously in our breeding colony of C57BL/6 mice, and benzene exposure preferentially induces thymic lymphoma rather than nonthymic lymphoma [40], these imply that h-Trx overexpression in the present study likely attenuates solely the benzene-induced increase in the incidence of hemopoietic malignancies [41]. Along with the decrease in the incidence of MN-Rets in the h-Trx-Tg mice (Fig. 8), the prevention of benzene-induced thymic and nonthymic lymphomagenicities in the Tg mice as shown in Figure 9B also confirmed that thymic lymphoma was induced by oxidative stress.

The last unresolved question is why the incidence of nonthymic lymphoma did not decrease in the h-Trx-Tg mice, which seemed to be related to the insufficient decrease in MN-Ret incidence in the h-Trx-Tg mice. There seems to be an inconsistency between this insufficient decrease in MN-Ret incidence and the almost complete clearance of ROS as shown in Figure 6. An interpretation of this inconsistency remains hypothetical, but thymic lymphoma cells may preferentially develop under leukocyte chemotaxis and chemokine exposure, which may be specifically induced by benzene-induced oxidative stress. When Trx is overexpressed, such leukocyte chemotaxis as well as chemokine expression may be suppressed and thymic lymphoma may not develop under such circumstances anymore [42]. In another aspect, Trx overexpression may preferentially enhance p53-dependent apoptosis even when thymic lymphoma developed. These possibilities may be relevant to the observations that thymic lymphoma seems to be a neoplasm only induced by benzene exposure, and nonthymic lymphoma is considered to develop spontaneously [43,44]. Consistently, the incidence of MN-Rets increased in the Wt mice during benzene exposure, which then decreased after the cessation of exposure, whereas in the h-Trx-Tg mice, on the other hand, the incidence of MN-Rets markedly decreased (Fig. 8). Also, a possible explanation is due to immunosuppression induced by benzene-related pro-hapten, and its consequent increase in stress level, which may lead to ROS-independent lymphomagenicity.

It is of interest to examine AhR gene expression. Because our previous study showed that benzene-induced hemato-lymphoid toxicity is transmitted via AhR and that AhR-KO mice surprisingly do not show any benzene-induced hemato-lymphoid toxicity [28], the significant downregulation of AhR observed solely in the h-Trx-Tg mice may be another factor contributing to the attenuation of benzene-induced hemato-lymphoid toxicity in the h-Trx-Tg mice. This downmodulation of AhR, specifically observed in the h-Trx-Tg mice, did not make any difference in the expression level of CYP2E1 in the present study; therefore, the mechanistic background underlying the dif-

ference in ROS production between Wt mice and h-Trx-Tg mice in relation to this downmodulation is yet to be determined.

### Acknowledgments

This work was supported in part by the Human Sciences of Japan (KH31034), Grant-in-Aid for Scientific Research C (Nos. 15510064 and 16590329), the fund from Nuclear Research of the MEXT, Japan, and also the MHLW-Research Fund (H16-Chemistry0002) National Institute of Health Sciences.

We thank Ms. E. Tachihara, Ms. Y. Usami, Ms. Y. Shinzawa, Ms. M. Uchiyama, Mr. K. Terasaka, and Ms. C. Aoyagi for excellent technical assistance. We also thank Drs. M. Hayashi, H. Matsutani, and K. Igarashi for valuable advice, and Ms. Y. Kikuchi, Ms. N. Katsu, Ms. M. Yoshizawa, and Ms. M. Hojo for secretarial assistance.

### References

- Golding BT, Watson WP. Possible mechanisms of carcinogenesis after exposure to benzene. *IARC Sci Publ.* 1999;75–88.
- Madl AK, Paustenbach DJ. Airborne concentrations of benzene due to diesel locomotive exhaust in a roundhouse. *J Toxicol Environ Health A.* 2002;65:1945–1964.
- Mitacek EJ, Brunneemann KD, Polednak AP, Limsila T, Bhothisuwan K, Hummel CF. Rising leukemia rates in Thailand: the possible role of benzene and related compounds in cigarette smoke. *Oncol Rep.* 2002; 9:1399–1403.
- Aksoy M, Erdem S, DinCol G. Leukemia in shoe-workers exposed chronically to benzene. *Blood.* 1974;44:837–841.
- Yin SN, Hayes RB, Linet MS, et al. An expanded cohort study of cancer among benzene-exposed workers in China. *Benzene Study Group. Environ Health Perspect.* 1996;104:1339–1341.
- Lan Q, Zhang L, Li G, et al. Hematotoxicity in workers exposed to low levels of benzene. *Science.* 2004;306:1774–1776.
- Ross D. The role of metabolism and specific metabolites in benzene-induced toxicity: evidence and issues. *J Toxicol Environ Health A.* 2000;61:357–372.
- Lewis JG, Stewart W, Adams DO. Role of oxygen radicals in induction of DNA damage by metabolites of benzene. *Cancer Res.* 1988; 48:4762–4765.
- Parke DV. Personal reflections on 50 years of study of benzene toxicology. *Environ Health Perspect.* 1996;104(Suppl 6):1123–1128.
- Dean BJ. Genetic toxicology of benzene, toluene, xylenes and phenols. *Mutat Res.* 1978;47:75–97.
- Dean BJ. Recent findings on the genetic toxicology of benzene, toluene, xylenes and phenols. *Mutat Res.* 1985;154:153–181.
- Creek MR, Mani C, Vogel JS, Turteltaub KW. Tissue distribution and macromolecular binding of extremely low doses of [<sup>14</sup>C]-benzene in B6C3F1 mice. *Carcinogenesis.* 1997;18:2421–2427.
- Mazzullo M, Bartoli S, Bonora B, et al. Benzene adducts with rat nucleic acids and proteins: dose-response relationship after treatment in vivo. *Environ Health Perspect.* 1989;82:259–266.
- Lutz WK. Quantitative evaluation of DNA binding data for risk estimation and for classification of direct and indirect carcinogens. *J Cancer Res Clin Oncol.* 1986;112:85–91.
- Snyder R, Hedli CC. An overview of benzene metabolism. *Environ Health Perspect.* 1996;104(Suppl 6):1165–1171.
- Siou G, Conan L, el Haitem M. Evaluation of the clastogenic action of benzene by oral administration with 2 cytogenetic techniques in mouse and Chinese hamster. *Mutat Res.* 1981;90:273–278.

17. Winn LM. Homologous recombination initiated by benzene metabolites: a potential role of oxidative stress. *Toxicol Sci.* 2003;72:143–149.
18. Takagi Y, Mitsui A, Nishiyama A, et al. Overexpression of thioredoxin in transgenic mice attenuates focal ischemic brain damage. *Proc Natl Acad Sci U S A.* 1999;96:4131–4136.
19. Mitsui A, Hirakawa T, Yodoi J. Reactive oxygen-reducing and protein-refolding activities of adult T cell leukemia-derived factor/human thioredoxin. *Biochem Biophys Res Commun.* 1992;186:1220–1226.
20. Hirabayashi Y. p53-dependent gene profiling for reactive oxygen species after benzene inhalation: special reference to genes associated with cell cycle regulation. *Chem Biol Interact.* 2005;153–154:165–170.
21. Matsui M, Oshima M, Oshima H, et al. Early embryonic lethality caused by targeted disruption of the mouse thioredoxin gene. *Dev Biol.* 1996;178:179–185.
22. Cronkite EP, Bullis J, Inoue T, Drew RT. Benzene inhalation produces leukemia in mice. *Toxicol Appl Pharmacol.* 1984;75:358–361.
23. Yoon BI, Hirabayashi Y, Kawasaki Y, et al. Mechanism of action of benzene toxicity: cell cycle suppression in hemopoietic progenitor cells (CFU-GM). *Exp Hematol.* 2001;29:278–285.
24. Hirabayashi Y, Matsuda M, Aizawa S, Kodama Y, Kanno J, Inoue T. Serial transplantation of p53-deficient hemopoietic progenitor cells to assess their infinite growth potential. *Exp Biol Med (Maywood).* 2002;227:474–479.
25. Hayashi M, Morita T, Kodama Y, Sofuni T, Ishidate M Jr. The micronucleus assay with mouse peripheral blood reticulocytes using acridine orange-coated slides. *Mutat Res.* 1990;245:245–249.
26. Frith CH, Ward JM, Harleman JH, et al. Hematopoietic system. In: Mohr U, ed. *International Classification of Rodent Tumor: The Mouse.* Heidelberg: Springer; 2001. p. 417–451.
27. Hirabayashi Y, Inoue T, Suda Y, Aizawa S, Ikawa Y, Kanisawa M. Hemopoietic neoplasms in lethally irradiated mice repopulated with bone marrow cells carrying the human c-myc oncogene: a repopulation assay. *Exp Hematol.* 1992;20:167–172.
28. Yoon BI, Hirabayashi Y, Kawasaki Y, et al. Aryl hydrocarbon receptor mediates benzene-induced hematotoxicity. *Toxicol Sci.* 2002;70:150–156.
29. Fridovich I. Quantitative aspects of the production of superoxide anion radical by milk xanthine oxidase. *J Biol Chem.* 1970;245:4053–4057.
30. Bass DA, Parce JW, Dechatelet LR, Szejda P, Seeds MC, Thomas M. Flow cytometric studies of oxidative product formation by neutrophils: a graded response to membrane stimulation. *J Immunol.* 1983;130:1910–1917.
31. Royall JA, Ischiropoulos H. Evaluation of 2', 7'-dichlorofluorescein and dihydrorhodamine 123 as fluorescent probes for intracellular H<sub>2</sub>O<sub>2</sub> in cultured endothelial cells. *Arch Biochem Biophys.* 1993;302:348–355.
32. Frank J, Biesalski HK, Dominici S, Pompella A. The visualization of oxidant stress in tissues and isolated cells. *Histol Histopathol.* 2000;15:173–184.
33. Imrich A, Kobzik L. Flow cytometric analysis of macrophage oxidative metabolism using DCFH. *Methods Mol Biol.* 1998;91:97–108.
34. Amer J, Goldfarb A, Fibach E. Flow cytometric measurement of reactive oxygen species production by normal and thalassaemic red blood cells. *Eur J Haematol.* 2003;70:84–90.
35. Smith MT. The mechanism of benzene-induced leukemia: a hypothesis and speculations on the causes of leukemia. *Environ Health Perspect.* 1996;104:1219–1225.
36. Bauer AK, Faiola B, Abernethy DJ, et al. Genetic susceptibility to benzene-induced toxicity: role of NADPH: quinone oxidoreductase-1. *Cancer Res.* 2003;63:929–935.
37. Das KC, Lewis-Molock Y, White CW. Elevation of manganese superoxide dismutase gene expression by thioredoxin. *Am J Respir Cell Mol Biol.* 1997;17:713–726.
38. Yoon BI, Li GX, Kitada K, et al. Mechanisms of benzene-induced hematotoxicity and leukemogenicity: cDNA microarray analyses using mouse bone marrow tissue. *Environ Health Perspect.* 2003;111:1411–1420.
39. Ueno M, Masutani H, Arai RJ, et al. Thioredoxin-dependent redox regulation of p53-mediated p21 activation. *J Biol Chem.* 1999;274:35809.
40. Kawasaki Y, Hirabayashi Y, Yoon BI, et al. Benzene inhalation induced an early onset and a high incidence of leukemias in the p53-deficient C57BL/6 mice. *Jpn J Cancer Res.* 2001;92:71.
41. Husbeck B, Powis G. The redox protein thioredoxin-1 regulates the constitutive and inducible expression of the estrogen metabolizing cytochromes P450 1B1 and 1A1 in MCF-7 human breast cancer cells. *Carcinogenesis.* 2002;23:1625–1630.
42. Burke-Gaffney A, Callister ME, Nakamura H. Thioredoxin: friend or foe in human disease? *Trends Pharmacol Sci.* 2005;26:398–404.
43. Ravi D, Muniyappa H, Das KC. Endogenous thioredoxin is required for redox cycling of anthracyclines and p53-dependent apoptosis in cancer cells. *J Biol Chem.* 2005;280:40084–40096.
44. Zhou Y, Kok KH, Chun AC, et al. Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits p53-induced apoptosis. *Biochem Biophys Res Commun.* 2000;268:921–927.

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

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

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

### 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 (house-keeping 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 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

**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

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

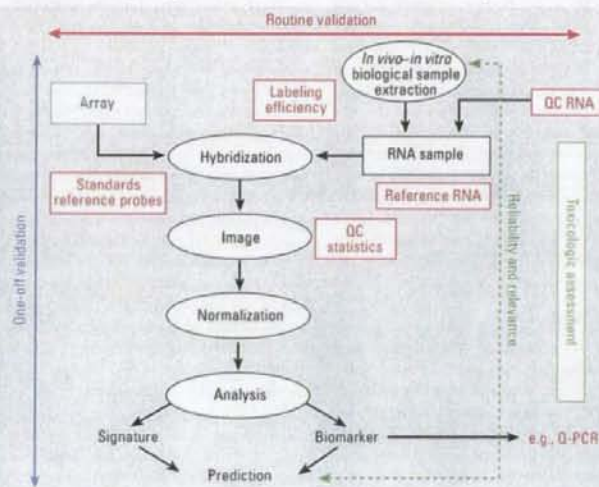
procedures and control measures and make experimental data open to regulatory review. "Best practices" for toxicogenomics can be established until formal procedures are adopted. This may be a more realistic solution that permits the advancement of science while addressing the need for QA and QC.

### 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 validated, 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/tox-nutri/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 technological 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

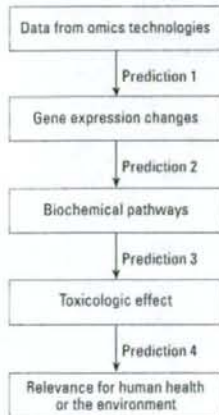
**Table 2.** Recommendations: focus on technology.

- 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 vivo* 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 practitioners 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.

## REFERENCES

**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

- 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, Blaubeer BJ, Fentem JH, Bruner L, Combes RD, Etwall 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* 3(2):259-271.
- Boess F, Kamber M, Romers S, Gassei 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, Ahbeygunawardena N, et al. 2003. ArrayExpress—a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res* 31:68-71.
- Brindle JT, Arnti 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 <sup>1</sup>H-NMR-based metabolomics. *Nat Med* 12:1439-1445.
- Brown HR, Ni H, Benevides 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* 112:445-455.
- 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, Sistara 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.mfprog.org.uk/themes/theme\\_publications\\_item.asp?intThemeID=22&intPublicationID=961](http://www.mfprog.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):228-230.
- 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: forecost 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* 415:323.
- 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\\_00.html](http://www.oecd.org/document/63/0,2340,en_2649_34381_2346175_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\\_00.html](http://www.oecd.org/document/29/0,2340,en_2649_34373_34704669_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/tet-toc.htm](http://altweb.jhsph.edu/publications/humane_exp/tet-toc.htm) [accessed 8 July 2005].
- Sensone SA, Morrison N, Rocca-Serra P, Foster 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, Chevret S, et al. 2002. Design and implementation of microarray gene expression markup language (IMAGE-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* 112:1236-1248.
- 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, Fusco J, et al. 2003. ArrayTrack—supporting toxicogenomic research at the U.S. Food and Drug Administration National Center for Toxicological Research. *Environ Health Perspect* 111:1819-1828.
- 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/oa/spc/html/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: <http://www.fda.gov/cder/guidance/6400n1.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, Curiñonis 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, Curiñonis R, Lum PY, Praestgaard JT, Moritt 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.

## ORIGINAL ARTICLE

## Nestin expression in odontoblasts and odontogenic ectomesenchymal tissue of odontogenic tumours

S Fujita, K Hideshima, T Ikeda

*J Clin Pathol* 2006;59:240-245. doi: 10.1136/jcp.2004.025403

See end of article for authors' affiliations

Correspondence to:  
Dr S Fujita, Division of  
Oral Pathology and Bone  
Metabolism, Department  
of Developmental and  
Reconstructive Medicine,  
Nagasaki University  
Graduate School of  
Biomedical Sciences, 1-7-  
1 Sakamoto, Nagasaki  
852-8588, Japan;  
fujitsu@net.nagasaki-u.ac.  
jp

Accepted for publication  
14 June 2005

**Background:** Nestin, one of the intermediate filaments constituting the cytoskeleton, is a marker of neural stem cells or progenitor cells. Its expression is also related to tooth development and repair of dentine.

**Aims:** The aim of this study was to investigate nestin expression in various odontogenic tumours and evaluate its usefulness for histopathological diagnosis.

**Methods:** We studied formalin fixed, paraffin embedded specimens from 129 cases of odontogenic tumours and 9 of mandibular intraosseous myxoma. After characterisation of odontogenic ectomesenchymal tissues in these tumours using antibodies to vimentin, desmin, neurofilament, and glial fibrillary acidic protein, we immunohistochemically examined nestin expression.

**Results:** No differentiation towards muscle and nervous tissues was found in the odontogenic ectomesenchymal tissues. Although almost all the ameloblastomas and malignant ameloblastomas were negative for nestin, odontogenic ectomesenchyme in the odontogenic mixed tumours demonstrated nestin immunolocalisation, particularly in the region adjacent to the odontogenic epithelium. Odontoblasts and their processes, pulp cells near the positive odontoblasts, and flat cells adhering to the dentine showed immunoreaction with nestin in the odontomas and odontoma-like component in the ameloblastic fibro-odontomas. Neoplastic cells in almost half cases of jaw myxoma and one case of odontogenic fibroma expressed nestin.

**Conclusions:** The distribution of nestin in the odontogenic mixed tumours suggests that nestin expression in the odontogenic ectomesenchyme is upregulated by stimulation from odontogenic epithelium. In addition, nestin may also be involved in the differentiation from pulp cells to odontoblasts in odontogenic tumours. Therefore, nestin is a useful marker for the odontogenic ectomesenchyme and odontoblasts in odontogenic tumours.

Nestin, one of the intermediate filaments constituting the cytoskeleton, is a marker of neural stem cells or progenitor cells. Its expression is also related to tooth development and repair of dentine.

Nestin is one of the intermediate filaments constituting the cytoskeleton and is known as a marker of neural stem cells or progenitor cells.<sup>1</sup> This protein appears at an early developmental stage of both the central nervous system and muscles. Although the function of nestin in cell differentiation is not clearly understood, nestin is replaced with glial fibrillary acidic protein (GFAP) and desmin, accompanied by developmental processing of nervous and muscular tissues, respectively.<sup>1-3</sup> Nestin expression has been reported in other tissues and pathological conditions, such as neural crest,<sup>4</sup> heart,<sup>5</sup> testis,<sup>6</sup> reactive astrocytes after brain injury,<sup>6,7</sup> and central and peripheral nervous system tumours.<sup>8-12</sup>

With respect to tooth development and dental pathological conditions, odontoblasts producing dentine in the human and rodent tooth germ express nestin.<sup>13-16</sup> Dentine fibres and pulp cells adjacent to the odontoblasts at the tooth cusp regions also indicate localisation of nestin. Human nestin expression continues during tooth development and disappears when development is complete.<sup>17</sup> In the mature tooth with caries or after cavity formation, nestin is re-upregulated in odontoblasts and their processes when they are stimulated by caries or when they form the secondary dentine after injury, respectively.<sup>13,18,19</sup> Thus, nestin is considered to be related to the differentiation toward odontoblasts and to the acquisition of the function to produce dentine.

In this study, we immunohistochemically investigated nestin expression in varied odontogenic tumours. Odontogenic neoplasms are histologically divided into several categories depending on their constituting components by

WHO as detailed in table 1.<sup>17</sup> The neoplasms indicated within parentheses were the objective materials of our study, although there is no conclusive evidence that the jaw myxomas examined in our study are derived from odontogenic organs. We found a characteristic nestin distribution in all these tumours and discussed the implication of its expression in the odontogenic tumours.

## MATERIALS AND METHODS

We investigated 129 cases of human odontogenic tumours: ameloblastoma (44), malignant ameloblastoma (3), odontoma (compound type 40; complex type 22), ameloblastic fibroma (2), ameloblastic fibro-odontoma and fibrodentinitoma (7), ameloblastic fibrosarcoma (2), adenomatoid odontogenic tumour (AOT) (6), and odontogenic fibroma (3). In addition, nine cases of intraosseous myxoma in the mandible were also examined. These materials were retrieved from the histopathological files of Division of Oral Pathology and Bone Metabolism, Department of Developmental and Reconstructive Medicine, Nagasaki University Graduate School of Biomedical Sciences. The specimens were fixed in 10% neutral buffered formalin fixative, decalcified with formic acid where necessary, embedded in paraffin, cut into 3 µm thick sections, and stained with haematoxylin and eosin. Table 2 summarises the clinical data of all the patients.

To confirm that the odontogenic ectomesenchymal tissues in these tumours contain no nervous and muscular tissues,

**Abbreviations:** AOT, adenomatoid odontogenic tumour; GFAP, glial fibrillary acidic protein

**Table 1** Categories of tumour

<b>Benign</b>
Tumours comprising odontogenic epithelium without odontogenic ectomesenchyme (for example, ameloblastoma)
Tumours comprising odontogenic epithelium with odontogenic ectomesenchyme, with or without dental hard tissue formation (for example, ameloblastic fibroma, ameloblastic fibro-odontoma, ameloblastic fibradentinoma, odontoma and adenomatoid odontogenic tumour)
Tumours comprising odontogenic ectomesenchyme with or without odontogenic epithelium (for example, odontogenic fibroma and myxoma)
<b>Malignant</b>
Odontogenic carcinomas (for example, malignant ameloblastoma)
Odontogenic sarcomas (for example, ameloblastic fibrosarcoma)
Odontogenic carcinosarcoma

**Table 2** Examined odontogenic tumours and jaw myxoma

Histology	No. of cases	Sex		Average age at first visit
		Male	Female	
Ameloblastoma	44	26	18	35.4
Malignant ameloblastoma	3	2	1	53.0
Odontoma	62	29	33	23.2
Compound type	40	17	23	21.1
Complex type	22	12	10	26.7
Ameloblastic fibroma	2	1	1	21.0
AFO and AFD	7	5	2	13.3
Ameloblastic fibrosarcoma	2	1	1	26.5
AOT	6	2	4	19.8
Odontogenic fibroma	3	1	2	36.7
Myxoma	9	5	4	41.7
<b>Total</b>	<b>138</b>	<b>72</b>	<b>66</b>	

AFO and AFD, ameloblastic fibro-odontoma and fibradentinoma, AOT, adenomatoid odontogenic tumour.

immunostaining for vimentin, desmin, neurofilament, and GFAP were performed, because muscular and neural cells frequently react with antibodies to nestin. All monoclonal antibodies against these proteins were purchased from DakoCytomation Co. Ltd, Kyoto, Japan. The immunohistochemical assay was performed using the EnVision+ system (DakoCytomation) with goat anti-mouse immunoglobulin conjugated to peroxidase labelled dextran polymer.

For the investigation of nestin expression, we used rabbit anti-human nestin polyclonal antibody (Chemicon International Inc., CA, USA) as a primary antibody, and the EnVision+ system (DakoCytomation) for polyclonal antibody.

The chromogen 3,3'-diaminobenzidine (Sigma-Aldrich Japan, Tokyo, Japan) was used to reveal the sites of peroxidase activity. After the sections were counterstained with haematoxylin, they were dehydrated and mounted in a synthetic medium.

## RESULTS

### Characterisation of odontogenic ectomesenchymal tissues

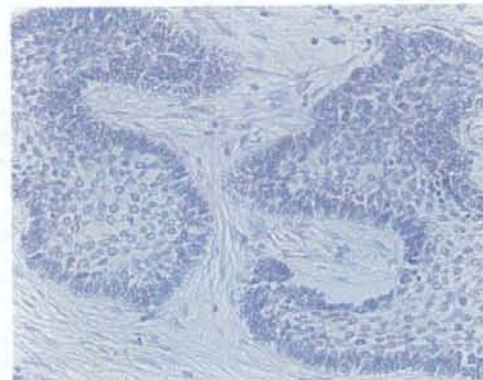
Odontogenic ectomesenchyme and pulp tissue included in the odontogenic mixed tumours, and the odontomas were positive for vimentin and negative for desmin, neurofilament, and GFAP (supplemental fig 1; <http://www.jclinpath.com/supplemental>). We found that intrinsic muscle and nervous tissues showed desmin expression, and neurofilament and GFAP expression, respectively (supplemental fig 2; <http://www.jclinpath.com/supplemental>).

### Immunohistochemical assay for nestin

A brief summary of the results is shown in table 3.

#### Ameloblastoma

The parenchyma of almost all the cases (43 of 44) of ameloblastoma displayed no nestin immunoreactivity (fig 1). However, one case showed the expression in the cytoplasm of the focal peripheral columnar epithelium bordering the fibrous stroma.



**Figure 1** Parenchyma and stroma cells of ameloblastoma showed no nestin immunoreactivity.

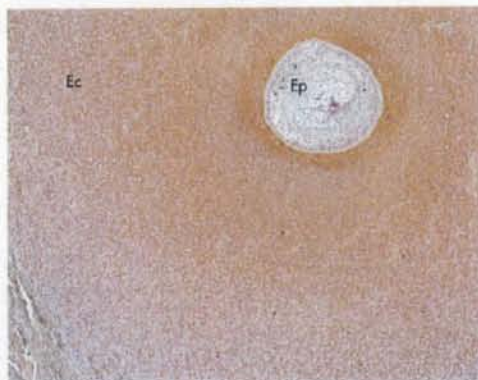
**Table 3** Expression of nestin in odontogenic tumours and jaw myxoma

Histology (no. of cases)	No. of positive cases	Localisation in parenchyma
Ameloblastoma (44)	1 (2%)	Peripheral columnar cells
Malignant ameloblastoma (3)	0 (0)	No localisation
Odontoma (62)	49 (79%)	
Compound type (40)	33 (82%)	Odontoblasts, dentinal fibres, pulp cells associated with odontoblasts
Complex type (22)	16 (73%)	Cells adhering to dentine, their processes in dentinal tubules
Ameloblastic fibroma (2)	2 (100%)	Ectomesenchyme > epithelium
AFO (7) and AFD (6)	6 (86%)	Ectomesenchyme > epithelium
Ameloblastic fibrosarcoma (2)	2 (100%)	Ectomesenchyme
AOT (6)	5 (83%)	Small nodular foci and rosette patterns
Odontogenic fibroma (3)	1 (33%)	Ectomesenchyme
Myxoma (9)	4 (44%)	Tumour cells homogeneously in positive cases

AFO and AFD, ameloblastic fibro-odontoma and fibradentinoma, AOT, adenomatoid odontogenic tumour.



**Figure 2** Odontoma showed nestin expression in the odontoblasts and their processes in the dental tubules. Pulp cells adjacent to the odontoblasts gave a positive reaction (arrows).



**Figure 3** Ameloblastic fibroma, consisting of odontogenic epithelial follicle (Ep) and ectomesenchymal tissue (Ec). Nestin was expressed in the odontogenic ectomesenchymal cells, particularly around the epithelial component.

#### Malignant ameloblastoma

No localisation of nestin was observed in the parenchyma of all three cases of malignant ameloblastoma.

#### Odontoma

Nestin immunoreactivity was found in the columnar odontoblasts adjacent to the dentine matrix, particularly in the compound-type odontomas. The positive odontoblasts were arranged in a palisading pattern on the predentine or dentine. Dentinal fibres through the dental tubules extending from the odontoblasts also showed immunoreactivity, and pulp cells adjacent to the odontoblasts were positive for nestin in some cases (fig 2). In addition, in complex-type odontomas, sparse flat cells adhering to the dentine and their processes in the irregular dental tubules were positive. Of 40 compound odontomas, 33 cases were positive, as were 16 of 22 complex odontomas.

#### Ameloblastic fibroma

Odontogenic ectomesenchymal tissues in the ameloblastic fibromas demonstrated immunoreactivity for nestin,



**Figure 4** Dentinoid matrix included in the ameloblastic fibrodentinoma. The odontogenic epithelial component (Ep) is adjacent to the dentinoid. Intense reactivity with nestin was found in the plump cells adhering to the dentinoid (arrows) and dentinal fibres (arrow heads).

although they did not contain the differentiated odontoblasts. Foci of aggregation of positive cells were noted in the ectomesenchymal element. As an intriguing finding, the ectomesenchymal cells around the neoplastic follicular odontogenic epithelium showed intense reaction with nestin, and the intensity weakened with greater distance from the odontogenic epithelium (fig 3). Some follicular epithelial components showed a cytoplasmic reaction in the peripheral columnar and inner stellate cells.

#### Ameloblastic fibro-odontoma and fibrodentinoma

The localisation of nestin in the ameloblastic fibro-odontomas and ameloblastic fibrodentinomas mimicked that in the ameloblastic fibromas and odontomas. The focal areas in the odontogenic ectomesenchymal tissue, particularly near the epithelial elements, showed strong reactivity. Some of the follicular patterns of epithelial elements demonstrated nestin immunoreactivity. Odontoma components showed nestin expression in differentiated odontoblasts, their processes, and adjacent pulp cells. Nestin localisation was also seen in the flat to plump cells adhering the dentinoid matrix and in sparse dentinal fibres in ameloblastic fibrodentinomas (fig 4).

#### Ameloblastic fibrosarcoma

Odontogenic ectomesenchymal cells were positive for nestin, although epithelial components were negative in all cases.

#### AOT

Five of six cases of AOT indicated intense nestin expression in epithelial elements, unlike the odontogenic mixed tumours, which showed immunoreactivity principally in the ectomesenchymal tissue. Nestin was intensely expressed in small nodular areas and in rosette arrangements in the epithelial elements of AOT (fig 5A). These positive cells localised near the droplet extracellular depositions, and haematoxylin and eosin revealed amorphous eosinophilic characteristics of the droplets (fig 5B, C).

#### Odontogenic fibroma

Of the three cases of odontogenic fibroma, one had odontogenic ectomesenchymal cells that were positive for nestin. This specimen contained abundant odontogenic epithelial strands, and the positive cells presented near the odontogenic epithelium.