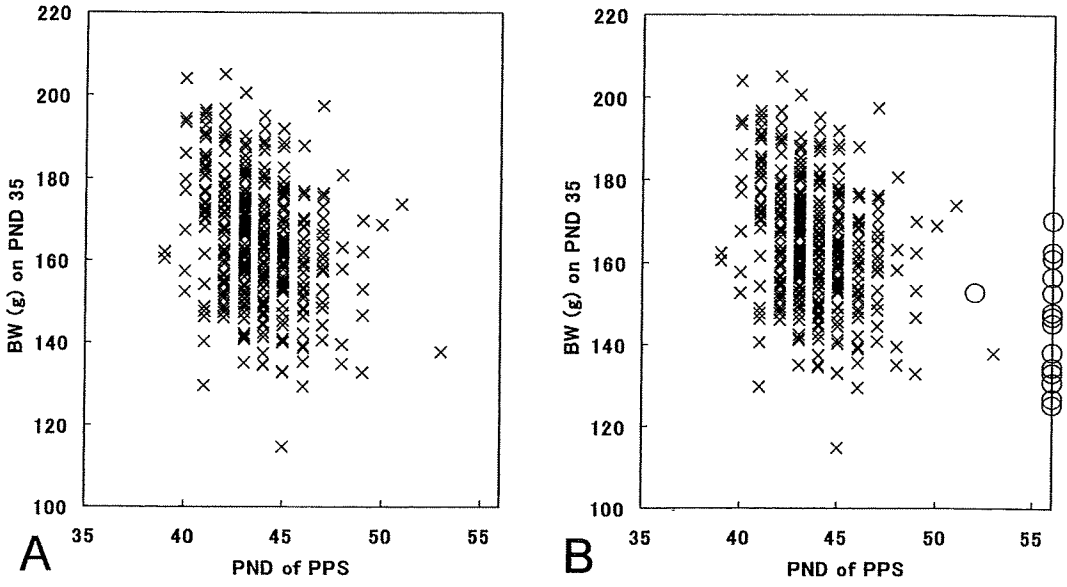

Fig. 5. Correlation diagram between the day of preputial separation and body weight of control and dosed rats. A: There is a negative correlation between PND of preputial separation (PPS) and body weight on PND 35. Number of cases = 366, Pearson's correlation coefficient: r = –0.30. B: Preputial separation (PPS) of males exposed to 3 mg/kg tamoxifen on PND 1–5 is delayed and the data are plotted as "O" outside the range of control data "X".

3) Pathological examination

Unilateral or bilateral cloudy white discoloration of the efferent ductule as well as testis enlargement was increased in the groups exposed to EE and DES on PND 1–5. Histopathological examination revealed retention of sperm or inflammatory cells in the lumen of the efferent ductule and edema in adipose tissue surrounding the dilated ductule. Organs of the EE or TAM groups showed no remarkable histological changes except for a relative atrophic appearance.

Correlation between preputial separation and body weight

The control animal data in this study (n=366) are symbolized by "X" on the scattergram of Fig. 5A. PND of preputial separation and body weight on PND 35 showed a negative correlation (Pearson's correlation coefficient: r = –0.30). The data from animals treated with 3 mg/kg TAM, symbolized by "O", are compared in Fig. 5B. Body weight on PND 35 in this group showed significant reduction. Preputial separation of TAM treated males was delayed, and the data were outside the range of the control data.

Discussion

Preputial separation in untreated rats initiates from cornification of the epithelium lying between the glans penis and prepuce[1,2]. The cornification progresses from the tip of the glans penis towards its base and from the dorsal surface to ventral aspect of the glans penis. Preputial separation is considered complete when cornification reaches the ventral end of the glans penis. However, complete separation was not observed in animals exposed to FLU or VZ in their fetal period, since they had a cleft phallus and hypospadias. Induction of hypospadias is reportedly caused by FLU, VZ and finasteride[9,14–19]. FLU, a well-known potent androgen receptor antagonist, is used as a non-steroidal, anti-androgen drug for the treatment of prostate cancer. FLU inhibits TS and DHT binding to the intracellular androgen receptor, and prenatal/perinatal FLU exposure induces abnormalities in the genital tract of rats such as hypospadias and agenesis of the prostate, epididymis, and vas deferens[9,14,15]. The fungicide, VZ, is also an androgen receptor antagonist. It induces hypospadias in rats after perinatal or prenatal administration[16,17]. Finasteride, which inhibits 5α-reductase conversion of TS to DHT, also induces hypospadias in male rats exposed to it from GD 15 to day 21 postpartum[18] or GD 6–20[19], and based on this finding, DHT is thought to be involved in the development of the external genitalia.

The most sensitive period to induce hypospadias is reportedly GD 15–16 with 400 mg/kg of VZ[17], while only weak sensitivity was found with treatment on GD 17–18. Finasteride-exposed rats also showed similar results[18]. In our present study, exposure to 100 mg/kg of VZ on GD 14–17 induced hypospadias and a cleft phallus with cleft prepuce, but exposure on GD 18–21 did not induce any abnormalities in the external genitalia. At a higher dose of 200 mg/kg, VZ caused the death of the pregnant females or newborn pups (data not shown). Exposure to 10 mg/kg of FLU on GD 14–17 induced hypospadias with a cleft phallus and cleft prepuce, while exposure on GD 18–21 induced hypospadias and a cleft phallus without a cleft prepuce.

Higher doses of FLU induced the same abnormality as our previous study[2]. Although both FLU and VZ are androgen receptor antagonists, their sensitive periods differed: cleft phallus was caused by FLU administration until later in pregnancy. FLU exposed males without the malformation showed a hypoplastic penis and a delay in preputial separation, but prenatal exposure to other chemicals did not affect the preputial separation.

Postnatal chemical exposure influenced preputial separation in a variety of ways. Anti-androgen, FLU caused a delay when administered on PND 35–39, and statistical analyses using individual data of DDE and VZ also showed significant delays of preputial separation, but neonatal exposure to these chemicals on PND 1–5 did not influence the time of preputial separation. On the other hand, neonatal administration of EE and TAM induced a marked delay or incomplete separation. DES exposure during both PND 1–5 and PND 35–39 caused a delay in separation. Preputial separation is thought to be dependent on the continued presence of androgen after PND 35, since castration on PND 35 blocks preputial separation and the addition of TS or DHT reverses the effects of castration[3]. Male rat serum testosterone levels have been reported to decrease after birth and to increase at the prepubertal stage[11]. Testosterone during fetal life is thought to act to masculinize the genitalia, and testosterone at puberty may act on the maturation of the target organ. Prenatal exposure to the anti-androgens used in the present study resulted in malformation of the external genitalia, and the effect of their prepubertal exposure was a delay in preputial separation. These results may indicate that anti-androgens effect on male rats in a relatively higher level of serum testosterone, and that male rats are not sensitive to neonatal exposure to anti-androgenic chemicals.

Neonatal exposure of estrogenic chemicals is known to induce marked effects on male rats. Delay of preputial separation and decrease of testis and prostate weight with a reduction of plasma testosterone levels have been reported in rats neonatally administered estradiol benzoate[20]. Neonatal treatment of DES or EE has also been reported to cause dose-dependent reductions in plasma testosterone levels and testis weights in adulthood[21]. An estrogen receptor is found in the male reproductive tract[22,23]. These findings suggest that estrogen is relevant to the growth of the male reproductive organs. The delay of preputial separation induced by DES, EE and TAM exposures on PND 1–5 in the present study is thought to be caused by estrogen-related effects, and the delay may have not only been caused by the direct effects on the genital tract but also by effects on the systemic endocrine function. Prepubertal exposure to DES induced a delay of preputial separation in male rats in the present study, the same as observed for anti-androgenic chemicals. Serum testosterone levels in adult or prepubertal male rats have been reported to be increased by FLU[24,25], DDE[5] and VZ[24] treatment, and the increase is thought to be caused by their anti-androgenic effects. It has been reported that serum testosterone levels in adult male rats treated with DES are decreased[26], thus, testosterone reduction may be the reason for the delay of preputial separation seen in the DES group prepubertally treated in our present study. EE and TAM have also been reported to decrease serum testosterone levels[25,27], but these chemicals did not delay preputial separation in our present study. The reason for the different effects of DES and EE/TAM treatments were not revealed in our study.

A negative correlation between the body weight on PND 35 and the day of preputial separation was demonstrated in the control males (Fig. 5A). This diagram shows a tendency for males of higher body weight to complete preputial separation earlier than males of lower body weight. Ashby and Lefevre[8] thought that there was a marked dependence of the day of preputial separation on the initial body weight of the test animals, and that delays in preputial separation can only be interpreted with confidence when they are not accompanied by losses in body weight. In our present study, males exposed to 3 mg/kg TAM on PND 1–5 showed significantly lower body weight on PND 35, and their preputial separation was delayed. The delay may not depend on the reduced weight gain, since 14 males out of 15 did not show complete separation of the prepuce on the day of autopsy, PND 56, and their data were outside the range of control data.

AGD was reduced by FLU and VZ exposures on GD 14–17. Reduction of AGD by anti-androgen has been reported, and AGD is thought to have high sensitivity to anti-andorogens[17]. In the present study, however, there were no apparent changes in AGD after DDE exposure. DES also caused a decrease in AGD with GD 18–21 exposure, and there was a difference between the results of DES and EE/TAM exposure. Histological examination revealed a tortuous and bent cavernous body of the penis seen in the sagittal section of males prenatally exposed to FLU or VZ and sacrificed on PND 6. This morphological change may be a reason for the reduction in AGD.

The relative weights of the ventral prostate in males prenatally exposed to VZ or DES were decreased significantly. In these groups, prostate aplasia or hypoplasia was observed in the group exposed to VZ on GD 14–17 and DES on GD 18–21. Postnatal FLU exposure decreased the relative weight of the ventral prostate in the group dosed on PND 35–39. No apparent effect was detected in males postnatally exposed to DDE or VZ. PND 1–5 exposure of EE or TAM induced a reduction in the weight of the testis and other reproductive organs. Although the statistical analyses using the individual data showed significant results in the lower dose group and the other organs, there were no significant differences in the groups exposed to VZ. These findings indicate that preputial separation is more useful than the measurement of organ weight as an endpoint in detecting endocrine active chemicals under the conditions used in this study.

Prenatal exposure of anti-androgens FLU and VZ induced hypospadias, and the time of preputial separation could not be determined. Although GD 18–21 exposure was expected to produce a very low incidence of hypospadias,

FLU exposure on GD 18–21 induced the abnormality. Other chemicals did not induce hypospadias or delays of the preputial separation. From these results, it is unclear if preputial separation after prenatal exposure is a useful way for detecting endocrine active chemicals. On the other hand, prepubertal exposure to FLU, DDE and VZ caused delays in preputial separation, and neonatal exposure to EE and TAM induced delays in separation with a reduction in organ and body weight gain. Both neonatal and prepubertal exposure to DES caused a delay in preputial separation. These results indicate that neonatal (PND 1–10) and prepubertal exposure may be useful for detecting endocrine active chemicals by observing preputial separation, and that continuous administration of chemicals from PND 35 to the day of preputial separation may be more effective for prepubertal exposure.

In conclusion, the usability of preputial separation to detect endocrine active chemicals after prenatal exposure to them is still unclear. Postnatal exposure, however, may be a useful method for a screening assay to detect endocrine active chemicals by preputial separation, and postnatal exposure is dependent on both neonatal (PND 1–10) and prepubertal continuous exposure.

Acknowledgements: This work was supported by Health and Labour Sciences Research Grants from the Ministry of Health, Labour and Welfare of Japan.

References

1. Lyons WR, Berlin I, and Friedlander S. Cornification of balano-preputial epithelium in normal rats and in castrated rats treated with testosterone propionate. Endocrinology. **31**: 659–663. 1942.
2. Yoshimura S, Yamaguchi H, Konno K, Ohsawa N, Noguchi S, and Chisaka A. Hypospadias and incomplete preputial separation in male rats induced by prenatal exposure to an anti-androgen, flutamide. J Toxicol Pathol. **17**: 113–118. 2004.
3. Korenbrot CC, Huhtaniemi IT, and Weiner RI. Preputial separation as an external sign of pubertal development in the male rat. Biol Reprod. **17**: 298–303. 1977.
4. Yamada T, Sunami O, Kunimatsu T, Kamita Y, Okuno Y, Seki T, Nakatsuka I, and Matsuo M. Dissection and weighing of accessory sex glands after formalin fixation, and a 5-day assay using young mature rats are reliable and feasible in the Hershberger assay. Toxicology. **162**: 103–119. 2001.
5. O'Connor JC, Frame SR, Davis LG, and Cook JC. Detection of the environmental antiandrogen *p,p'*-DDE in CD and Long-Evans rats using a tier I screening battery and a Hershberger assay. Toxicol Sci. **51**: 44–53. 1999.
6. Sunami O, Kunimatsu T, Yamada T, Yabushita S, Sukata T, Miyata K, Kamita Y, Okuno Y, Seki T, Nakatsuka I, and Matsuo M. Evaluation of a 5-day Hershberger assay using young mature male rats: methyltestosterone and *p,p'*-DDE, but not fenitrothion, exhibited androgenic or antiandrogenic activity *in vivo*. J Toxicol Sci. **25**: 403–415. 2000.
7. Yamasaki K, Sawaki M, Ohta R, Okuda H, Katayama S, Yamada T, Ohta T, Kosaka T, and Owens W. OECD validation of the Hershberger assay in Japan: phase 2 dose response of methyltestosterone, vinclozolin, and *p,p'*-DDE. Environ Health Perspect. **111**: 1912–1919. 2003.
8. Ashby J and Lefevre PA. The peripubertal male rat assay as an alternative to the Hershberger castrated male rat assay for the detection of anti-androgens, oestrogens and metabolic modulators. J Appl Toxicol. **20**: 35–47. 2000.
9. McIntyre BS, Barlow NJ, and Foster PMD. Androgen-mediated development in male rat offspring exposed to flutamide *in utero*: permanence and correlation of early postnatal changes in anogenital distance and nipple retention with malformations in androgen-dependent tissues. Toxicol Sci. **62**: 236–249. 2001.
10. Gaytan F, Pinilla L, Romero JL, and Aguilar E. Differential effects of the administration of human chorionic gonadotropin to postnatal rats. J Endocrinol. **142**: 527–534. 1994.
11. Paz GF, Winter JSD, Reyes FI, and Faiman C. Developmental pattern of testosterone production by the rat testis. Steroids. **36**: 675–688. 1980.
12. Watanabe C, Kuwagata M, Yoshimura S, Azegami J, Kojima K, Ono H, and Nagao T. An improved technique for repeated gavage administration to rat neonates. Congenit Anom (Kyoto). **43**: 177–179. 2003.
13. Wada K. Method and problem in measurement of AGD and preputial separation (in Japanese). Annual Report of Hatano Research Institute. **25**: 109–114. 2002.
14. Mylchreest E, Sar M, Cattley RC, and Foster PMD. Disruption of androgen-regulated male reproductive development by di(*n*-butyl) phthalate during late gestation in rats is different from flutamide. Toxicol Appl Pharmacol. **156**: 81–95. 1999.
15. Miyata K, Yabushita S, Sukata T, Sano M, Yoshino H, Nakanishi T, Okuno Y, and Matsuo M. Effects of perinatal exposure to flutamide on sex hormones and androgen-dependent organs in F1 male rats. J Toxicol Sci. **27**: 19–33. 2002.
16. Gray LE Jr, Ostby JS, and Kelce WR. Developmental effects of an environmental antiandrogen: the fungicide vinclozolin alters sex differentiation of the male rat. Toxicol Appl Pharmacol. **129**: 46–52. 1994.
17. Wolf CJ, LeBlanc GA, Ostby JS, and Gray LE Jr. Characterization of the period of sensitivity of fetal male sexual development to vinclozolin. Toxicol Sci. **55**: 152–161. 2000.
18. Clark RL, Anderson CA, Prahalada S, Robertson RT, Lochry EA, Leonard YM, Stevens JL, and Hoberman AM. Critical developmental periods for effects on male rat genitalia induced by finasteride, a 5α-reductase inhibitor. Toxicol Appl Pharmacol. **119**: 34–40. 1993.
19. Anderson CA and Clark RL. External genitalia of the rat: normal development and the histogenesis of 5α-reductase inhibitor-induced abnormalities. Teratology. **42**: 483–496. 1990.
20. Bellido C, Gaytan F, Aguilar R, Pinilla L, and Aguilar E. Prepuberal reproductive defects in neonatal estrogenized male rats. Biol Reprod. **33**: 381–387. 1985.
21. Atanassova N, McKinnell C, Walker M, Turner KJ, Fisher JS, Morley M, Millar MR, Groome NP, and Sharpe RM. Permanent effects of neonatal estrogen exposure in rats on reproductive hormone levels, Sertoli cell number, and the

efficiency of spermatogenesis in adulthood. Endocrinology. **140**: 5364–5373. 1999.
22. Fisher JS, Millar MR, Majdic G, Saunders PT, Fraser HM, and Sharpe RM. Immunolocalisation of oestrogen receptor-α within the testis and excurrent ducts of the rat and marmoset monkey from perinatal life to adulthood. J Endocrinol. **153**: 485–495. 1997.
23. Saunders PTK, Fisher JS, Sharpe RM, and Millar MR. Expression of oestrogen receptor beta (ER β) occurs in multiple cell types, including some germ cells, in the rat testis. J Endocrinol. **156**: R13–17. 1998.
24. Yu WJ, Lee BJ, Nam SY, Ahn B, Hong JT, Do JC, Kim YC, Lee YS, and Yun YW. Reproductive disorders in pubertal and adult phase of the male rats exposed to vinclozolin during puberty. J Vet Med Sci. **66**: 847–853. 2004.
25. Okuno Y and Miyata K. Effects of an endocrine disruptor on male reproductive organs. J Toxicol Pathol. **14**: 79–82. 2001.
26. Goyal HO, Braden TD, Mansour M, Williams CS, Kamaleldin A, and Srivastava KK. Diethylstilbestrol-treated adult rats with altered epididymal sperm numbers and sperm motility parameters, but without alterations in sperm production and sperm morphology. Biol Reprod. **64**: 927–934. 2001.
27. Gill-Sharma MK, Balasinor N, and Parte P. Effect of intermittent treatment with tamoxifen on reproduction in male rats. Asian J Androl. **3**: 115–119. 2001.

J Toxicol Pathol 2004; 17: 113–118

Original

Hypospadias and Incomplete Preputial Separation in Male Rats Induced by Prenatal Exposure to an Anti-androgen, Flutamide

Shinsuke Yoshimura[1], Hajime Yamaguchi[1], Kazunori Konno[1], Noriko Ohsawa[1], Satoshi Noguchi[1], and Akiko Chisaka[1]

[1]Hatano Research Institute, Food and Drug Safety Center, 729–5, Ochiai, Hadano, Kanagawa 257–8523, Japan

Abstract: Hypospadias was induced in male Sprague-Dawley rats by prenatal exposure to 30 mg/kg/day of flutamide from gestational days 14 to 17, or from 18 to 21. Their external genitalia were examined histopathologically and compared to untreated controls. On postnatal day 6, the glans penis of untreated controls was bordered with epithelium. On postnatal day 22, papillae were recognized on the glans penis side, and cornification started close to the tip of the papilla on postnatal day 35. On postnatal day 42 cornification spread to the surface of the glans penis. The cornification progressed from tip to base of the glans penis, and from dorsal to ventral. When cornification reached the base of the glans penis, separation of the double layered epithelia was complete and the animal was considered sexually mature. Flutamide treatment on gestational days 14–17 induced a defect in the ventral half of the glans penis (cleft phallus) and cleft in the ventral prepuce (cleft prepuce) in the male pups, while treatment on gestational days 18–21 induced cleft phallus without apparent abnormalities in the prepuce. The external urethral orifice opened at the ventral end of the glans penis (hypospadias) in both treatment groups. In male pups with cleft phallus, cornification of the dorsal epithelium followed by separation of the prepuce occurred, while separation of the ventral part of glans penis did not occur because epithelium was not formed at the ventral part of the glans penis. Consequently, the onset of puberty was not decided in these animals. These findings indicate that the defect of the ventral half of the phallus is the reason why the time of sexual maturation was not decided, and that there is a difference between the phallus and prepuce in the sensitive period concerning the development of flutamide-induced malformations.
(J Toxicol Pathol 2004; **17**: 113–118)

Key words: flutamide, anti-androgen, rat, prenatal exposure, hypospadias

Introduction

Preputial separation, which is observed as separation of the prepuce from the glans penis, has been used as a sign of puberty in the male rat. Histological observation on the progress of preputial separation after cornification at the lining of prepuce and surface of glans penis was well described using Long-Evans rats in 1942[1]. Preputial separation is thought to be dependent on androgens, because castration blocked preputial separation, and the addition of testosterone (TS) or dihydrotestosterone (DHT) recovered the effect of castration[1,2]. In recent years, preputial separation has been used as an endpoint to evaluate endocrine disrupting chemicals. Although the observation of preputial separation is a useful tool for detecting sexual maturation, anti-androgenic chemicals induce hypospadias in male rats by intrauterine exposure. The time of sexual maturation is determined by complete separation of the prepuce from the ventral surface of the glans penis, but in males with hypospadias, puberty is undetermined because this complete separation in the glans penis is not evident[3].

The purpose of this study was to reveal the histological process of normal and abnormal preputial separation, as well to reveal the reason why the time of sexual maturation cannot be decided in males with hypospadias induced by prenatal exposure to flutamide (FLU), an anti-androgenic chemical, in Sprague-Dawley rats. FLU was administered on gestational days (GD) 14–17 (expected to be the most sensitive period for hypospadias) or on GD 18–21 (thought to be a less sensitive period for hypospadias).

Received: 26 January 2004, Accepted: 15 April 2004
Mailing address: Shinsuke Yoshimura, Laboratory of Toxicology, Hatano Research Institute, Food and Drug Safety Center, 729–5, Ochiai, Hadano, Kanagawa 257–8523, Japan
TEL: 81-463-82-4751 FAX: 81-463-82-9627
E-mail: yoshimura.s@fdsc.or.jp

Materials and Methods

Sprague-Dawley rats (Crj:CD (SD) IGS), 30 males and 33 females, 11 weeks of age, were obtained from Charles

River Japan, Inc. (Atsugi, Japan). All animals were acclimatized to laboratory conditions and quarantined for one week before mating. Rats used for this study were selected based upon general condition, appearance and behavior during the acclimatization period. Animals were housed individually in wire-bottom metal cages (220 × 270 × 190 mm) and kept in a barrier sustained animal room that was maintained at 21.0 – 25.0°C and 40.0 – 75.0% relative humidity with a 12-hour artificial light cycle (lighting from 7:00 to 19:00). Fifteen changes of room air per hour were provided. Commercial diet CE-2 (CLEA Japan, Inc., Tokyo) and water (Hadano City) were available ad libitum throughout the study. The protocol of the present study was approved by the Animal Use Committee of the Hatano Research Institute.

The untreated control group consisted of 18 females. While the second group, consisting of 9 females, was treated orally with 30 mg/kg/day of FLU (Sigma Chemical Co., St. Louis, USA) dissolved in corn oil (Nacalai Tesque, Inc., Kyoto, Japan) from GD 14 to 17, the third group, consisting of 6 females, was treated with the same dose from GD 18 to 21. From the result of a preliminary study the dosage was decided as 30 mg/kg/day, because pregnant rats died after administration of 100 mg/kg/day of FLU. To obtain pregnant animals, 12-week-old females were cohabited overnight on a 1:1 basis with males 12 weeks of age or older. Females were considered to be at GD 0 when daily examination revealed a vaginal plug. All pregnant animals were housed in cages with animal bedding (PAPER CLEAN®, Japan SLC, Inc., Shizuoka) from GD 18 until postpartum day 10, and allowed to give birth. On postnatal day (PND) 6 (PND 0 is the day of delivery) all female pups were discarded. Body weights of male pups were measured on PND 0, 6, 22, 35 and 56. Progress of preputial separation of male pups was observed macroscopically from PND 35.

Control male pups were sacrificed by exsanguination under anesthesia on PND 6, 22, 35, 42 and 56 (number of pups in group 1 were 6, 16, 15, 4 and 24, respectively). Male pups from FLU-treated females were sacrificed under anesthesia on PND 6 and 56 (number of pups in group 2 were 18 and 41, respectively, and those in group 3 were 8 and 21, respectively). After macroscopic examination, the prepuce and penis were dissected and fixed with 0.1 mol/L phosphate buffered 10% formalin solution. Sagittal slices of the prepuce and penis were embedded in paraffin, and sections were stained with hematoxylin-eosin (H & E) for histopathological examination.

Results

On macroscopic examination, the glans penis of control males was covered with prepuce, and the prepuce could be completely retracted to expose the glans penis until PND 46 (Figs. 1A, 1B). Prepuce of males prenatally exposed to FLU on GD 14–17 had a cleft at the ventral part (cleft prepuce), and the glans penis was observed from the cleft (Fig. 1C). The ventral part of the glans penis of these males was incompletely formed (cleft phallus) and the os penis was often exposed (Fig. 1D). The incidence of cleft prepuce was 80% (33/41), and the incidence of cleft phallus was 90% (37/41). Cleft prepuce is usually observed with cleft phallus. Another 4 males showed no cleft on their prepuce or phallus, while preputial separation was delayed or incomplete on PND 56. Although there was no cleft at the prepuce of males exposed to FLU on GD 18–21 (Fig. 1E), the ventral part of the glans penis was incompletely formed (cleft phallus, Fig. 1F), and incidence of the cleft phallus in this group was 100% (21/21). Body weight gains of males were not affected by FLU exposure.

Upon histological examination of untreated controls on PND 6, the glans penis was bordered with specific epithelium (Fig. 2A). The epithelium consisted of outer and inner basal layers (Fig. 2B). The outer layer lined the inside of the prepuce and the inner layer covered the glans penis. The urethra was located in the center of the glans penis and the os penis was observed between the dorsal surface of the glans penis and the urethra (Fig. 2A). On PND 22, there were many papillary processes from the glans penis (arrows in Fig. 2C), and the surface of these processes was covered with squamous epithelial cells. At this point the two basal layers lost their parallel arrangement. On PND 35, the epithelial layer consisted of stratified squamous epithelium, and the surface of the papillary processes (arrows in Fig. 2D) was covered with cornified cells. The cornified layer was limited to the surface of these processes. On PND 42, epithelial cells between the papillae also cornified, and both surfaces of the penis and prepuce consisted of keratinized stratified squamous epithelium. Cornification and separation were incomplete at the basal part of the glans penis. Separation at the ventral surface of the glans penis was more delayed than that at the dorsal surface. On PND 56 cornification was complete from the tip to the base of the glans penis and the ventral surface also showed cornified layers (Figs. 2E, 2F). The preputial separation was complete across the entire surface of the glans penis.

Histological examination of males prenatally exposed to FLU on GD 14–17 revealed a cleft at the ventral surface of genital tubercle on PND 6 (arrow in Fig. 3A). The urethra was not located in the center of the glans penis, but instead was observed at the ventral surface of the glans penis. The dorsal part of the glans penis was bordered by epithelium as observed in controls, while the ventral part was covered with urethral epithelium. This finding indicates that the ventral half of glans penis was not formed (comparison with controls as shown by an asterisk (*) in Fig. 2A). The cavernous body of the penis was tortuous and also observed in PND 56 males exposed to FLU on GD 14–17 (Fig. 3B). The dorsal surface of the glans penis and prepuce of PND 56 males were covered with keratinized stratified squamous epithelium, and the prepuce was separated from the glans penis (Fig. 3B). The ventral part of the glans penis and ventral epithelium were not formed between the urethra and subcutis, the ventral surface of the glans penis was not covered with squamous epithelium and the preputial

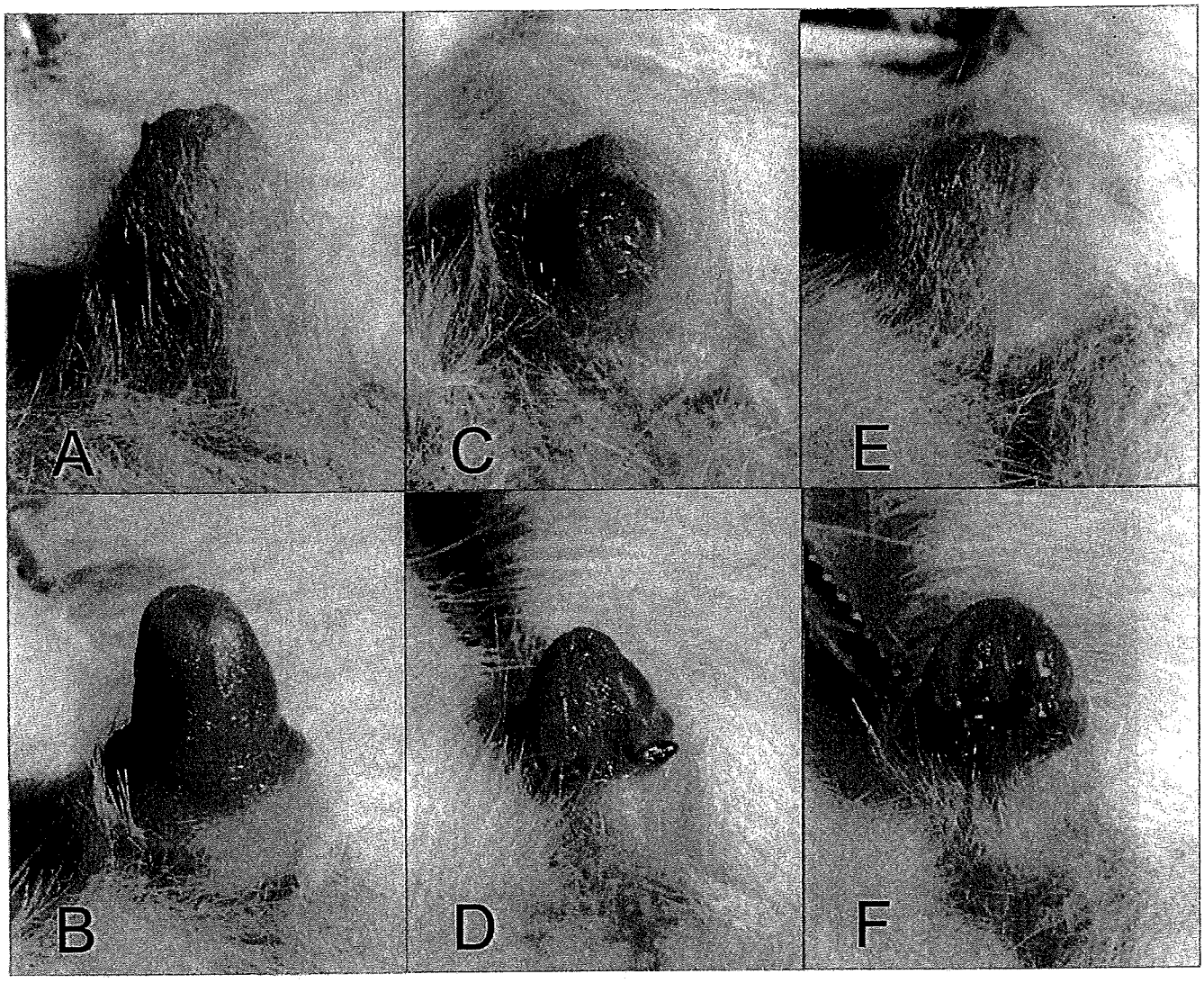

Fig. 1. Ventral surface of genital tubercle (A, C and E) and glans penis (B, D and F) of males at PND 56. A and B: Control rat. Prepuce is completely retracted. C and D: Male rat prenatally exposed to FLU on GD 14–17. Ventral side of the prepuce has a cleft, and the glans penis is observed from the cleft. Ventral part of the glans penis is incompletely formed (cleft phallus) and os penis is observed. E and F: Male rat prenatally exposed to FLU on GD 18–21. Prepuce does not have a cleft at the ventral side. Glans penis shows cleft phallus and os penis is observed.

separation did not progress at the ventral part. The external urethral orifice opened at the ventral surface of the glans penis (hypospadias). The preputial tissue was hypoplastic and the tip of the penis was not overlain with prepuce.

The glans penis of PND 6 males exposed to FLU on GD 18–21 was covered with skin, and a cleft was not observed at the ventral part of genital tubercle (Fig. 3C). The dorsal part of the glans penis was bordered by epithelium, but the ventral part of the glans penis and ventral epithelium were not formed between the urethra and subcutis (comparison with controls as shown by an asterisk (*) in Fig. 2A). The tortuous structure of the cavernous body was indistinct. The prepuce overlaid the glans penis of PND 56 males exposed on GD 18–21 (Fig. 3D). The dorsal surface of the glans penis of these males was covered with keratinized stratified squamous epithelium, and the prepuce was separated from the glans penis. The ventral part of the glans penis and ventral epithelium were not formed between the urethra and subcutis, and preputial separation did not progress at the ventral part. In these rats the external urethral orifice opened at the ventral surface of the glans penis.

Discussion

As described above, preputial separation in untreated rats initiated from cornification of the epithelium on the penile side lying in the dual phasic epithelium between the glans penis and prepuce. Cornification began at the surface very close to the apex of the papillary process from the glans penis, and when the cornification reached the next papilla the prepuce separated from the glans penis. Preputial separation progressed from the tip of the glans penis towards its base,

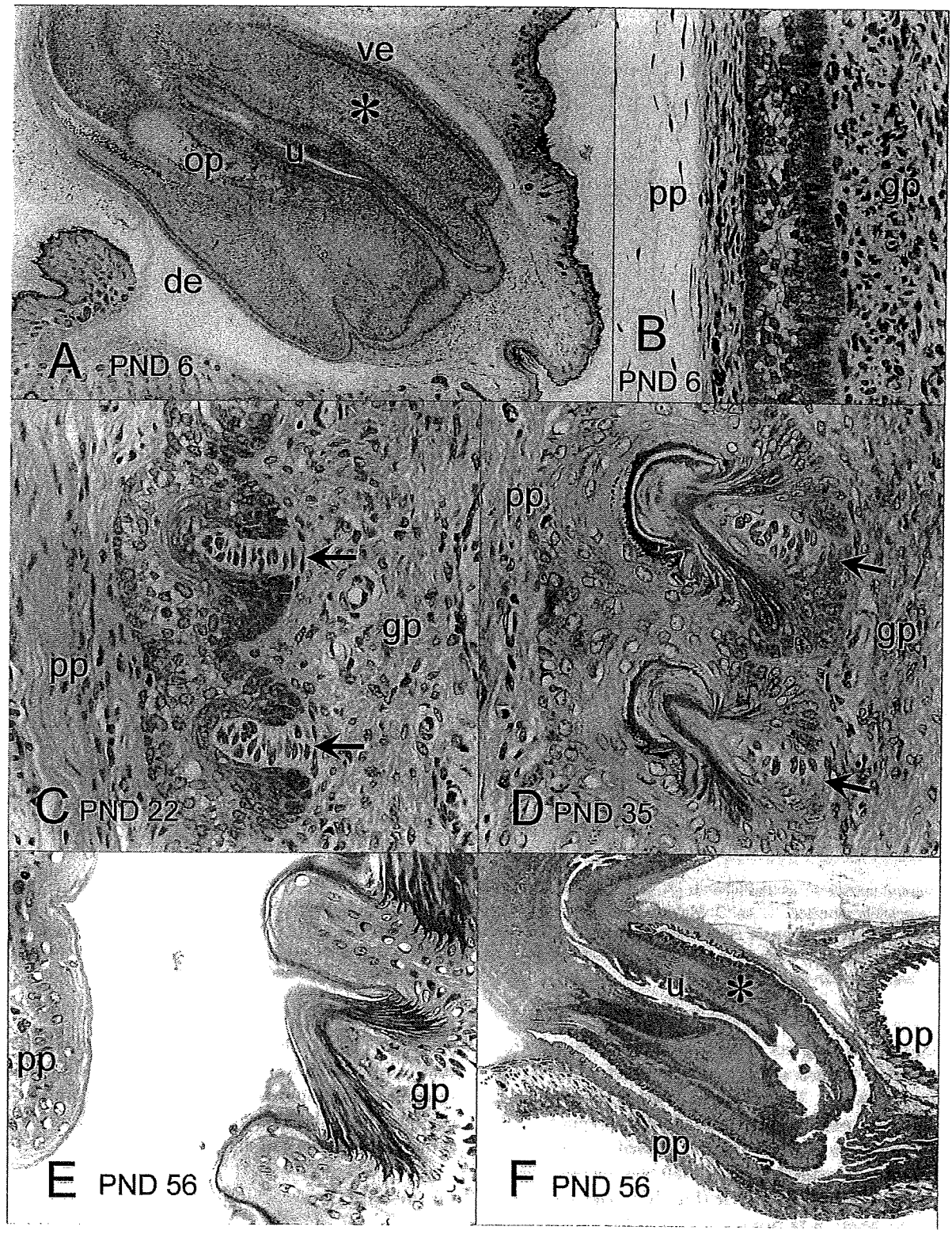

Fig. 2. Sagittal sections of the genital tubercle from control males.
A: Genital tubercle of a male on PND 6. Glans penis is bordered with dorsal epithelium (de) and ventral epithelium (ve). Urethra (u) is located in the center of the glans penis. op: os penis. *: ventral half of the glans penis.
B, C, D and E: Epithelium between the dorsal part of the glans penis (gp) and prepuce (pp) on PND 6, 22, 35 and 56, respectively. Epithelium of PND 6 consisted of outer and inner basal layers (B). Squamous epithelial cell of PND 22 is covering the papillary processes (arrows) from the glans penis (C). Epithelial layer of PND 35 consists of stratified squamous epithelium, and cornified layer is covering the papillary processes (arrows in D). Whole surface of both the glans penis and prepuce is covered with cornified layer on PND 56 (E).
F: Glans penis and prepuce on PND 56. Preputial separation is completed. Urethra (u) is located in the center of the glans penis.
*: ventral half of the glans penis. H & E. Magnification, A: × 35, B, C and D: × 400, E: × 330, F: × 9.

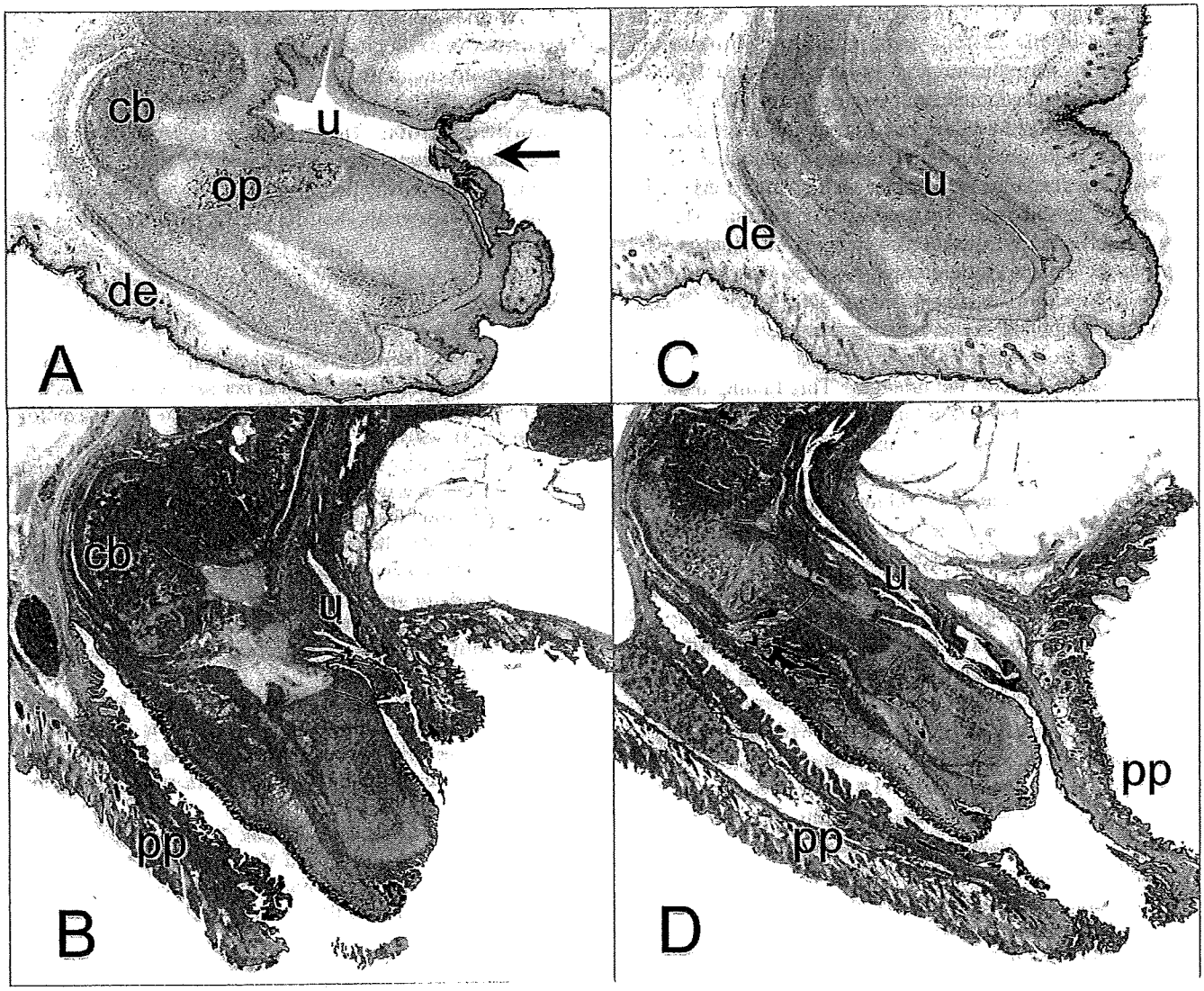

Fig. 3. Sagittal sections of genital tubercle from males prenatally exposed to FLU on GD 14–17 (A, B) or GD 18–21 (C, D), sacrificed on PND 6 (A, C) and PND 56 (B, D).
A: There is a cleft (arrow in the figure) at the ventral side of genital tubercle. Urethra (u) is observed at the ventral surface of the glans penis. Dorsal part of the glans penis is bordered with epithelium (de). Cavernous body (cb) shows tortuous structure.
B: Prepuce (pp) is separated from the glans penis at the dorsal part. Prepuce is hypoplastic, and the glans penis is not completely overlain with the prepuce. Urethra is located between the glans penis and subcutis.
C: Cleft is not formed at the ventral side of genital tubercle. Dorsal part of the glans penis is bordered with epithelium (de). Ventral part of the glans penis and ventral epithelium is not formed. Urethra is located between the glans penis and subcutis.
D: Prepuce is separated from the glans penis at the dorsal part. The ventral part of the glans penis and ventral epithelial layer is not formed. The glans penis is completely overlain with the prepuce.
H&E. Magnification, A and C: × 27; B and D: × 10.

and also from the dorsal to ventral surface of the glans penis. Histological features observed in controls of this study were almost the same as shown in Long-Evans rats[1]. Complete separation was not observed in animals exposed to FLU in their fetal period, since they had a cleft phallus at their ventral surface of the glans penis. Histopathological examination revealed defects in the ventral part of the glans penis and lack of an epithelial layer at the ventral part in newborn rats.

Induction of hypospadias has been reportedly caused by various chemicals, which include anti-androgens such as FLU, vinclozolin and finasteride. FLU is a well-known potent androgen receptor antagonist and is used as a nonsteroidal anti-androgen drug for the treatment of prostate cancer. FLU inhibits TS and DHT binding to the intracellular androgen receptor and prenatal/perinatal exposure to FLU induces abnormalities in the genital tract such as hypospadias, agenesis of the prostate, epididymis and vas deferens[2,4,5]. Vinclozolin, a fungicide, is also an androgen receptor antagonist, and induces hypospadias in

rats by perinatal (GD 14 to day 3 postpartum)[6] or prenatal (for 2 days in GD 12–21)[7] administration. Finasteride, which inhibits 5α-reductase conversion of TS to DHT, also induces hypospadias in male rats exposed from GD 15 to day 21 postpartum, and based on this finding DHT is thought to be involved in the development of external genitalia[8]. While the Wolffian ducts are dependent on TS, their derivatives such as the epididymis, vas deferens or seminal vesicles were not affected by intrauterine exposure to finasteride[8]. Hypoplastic change in the genital tubercle was reported in fetuses exposed to finasteride from GD 6 to 20[9]. Wedge shaped mesenchymal tissue between rectum and urogenital sinus failed to develop in these fetuses, and the urethra opened near the base of the tubercle. This finding indicates that the mesenchymal wedge may be the most sensitive area to loss of the effect of DHT in male fetuses. Our study also revealed a defect in the ventral part of the glans penis, in which preputial separation could not progress, and this is the reason why the time of sexual maturation could not be decided in males with hypospadias.

Androgen receptors are detectable in the mesenchymal cells of the rat urogenital tubercle from fetal day 14 onwards[10]. In many studies of sexual differentiation and male reproductive organ malformation, dosing starts from GD 12 or 14. The most sensitive period to induce hypospadias is reportedly GD 15–16 with 400 mg/kg of vinclozolin exposure, and the incidence was 42% (10/24), while only weak sensitivity (11%, 1/9) was found with treatment on GD 17–18[7]. Similar results were obtained in our study of vinclozolin (unpublished data)[11]. Exposure to 100 mg/kg of vinclozolin on GD 14–17 induced cleft phallus with cleft prepuce, and the incidence was 85%, but exposure on GD 18–21 induced no abnormalities in external genitalia. Pregnant females or newborn pups died after exposure to 200 mg/kg of vinclozolin. Finasteride exposed rats also showed similar results[8]. Male pups exposed to 20 mg/kg of finasteride on GD 16–17 showed hypospadias, and the incidence was 39% (14/36), while incidence in the GD 18–19 group was 0% (0/36).

Many reports describe the malformation of the phallus as hypospadias, but details and the incidence of the prepuce malformations are not clear. In our study, pregnant rats were administered with 30 mg/kg of FLU on GD 14–17, which was considered to be the sensitive period, and the incidence of cleft phallus and cleft prepuce was compared to the exposure on GD 18–21. Male pups exposed to FLU on GD 14–17 showed cleft phallus with hypospadias (90%) and cleft prepuce (80%), while males exposed on GD 18–21 had cleft phallus with hypospadias (100%) without apparent abnormality in the prepuce. A lower dose of FLU also showed similar results with lower incidence (unpublished data)[11]. Exposure of 10 mg/kg of FLU on GD 14–17 induced cleft phallus (58%) and cleft prepuce (25%), while exposure on GD 18–21 induced cleft phallus (25%) without apparent abnormality in the prepuce. These findings show that the period of sensitivity to FLU in terms of phallus malformation is different from vinclozolin and finasteride, and also that there are differences among the sensitive periods between the phallus and prepuce concerning FLU-induced malformations.

Acknowledgments: This work was supported by Health and Labour Sciences Research Grants from the Ministry of Health, Labour and Welfare of Japan. We would also like to thank Dr. Masayoshi Kanisawa for proofreading this manuscript.

References

1. Lyons WR, Berlin I, and Friedlander S. Cornification of balano-preputial epithelium in normal rats and in castrated rats treated with testosterone propionate. Endocrinology 1942; **31**: 659–663.
2. Korenbrot CC, Huhtaniemi IT, and Weiner RI. Preputial separation as an external sign of pubertal development in the male rat. Biol Reprod 1977; **17**: 298–303.
3. McIntyre BS, Barlow NJ, and Foster PMD. Androgen-mediated development in male rat offspring exposed to flutamide *in utero*: permanence and correlation of early postnatal changes in anogenital distance and nipple retention with malformations in androgen-dependent tissues. Toxicol Sci 2001; **62**: 236–249.
4. Mylchreest E, Sar M, Cattley RC, and Foster PMD. Disruption of androgen-regulated male reproductive development by di(*n*-butyl) phthalate during late gestation in rats is different from flutamide. Toxicol Appl Pharmacol 1999; **156**: 81–95.
5. Miyata K, Yabushita S, Sukata T, Sano M, Yoshino H, Nakanishi T, Okuno Y, and Matsuo M. Effects of perinatal exposure to flutamide on sex hormones and androgen-dependent organs in F1 male rats. J Toxicol Sci 2002; **27**: 19–33.
6. Gray LE Jr, Ostby JS, and Kelce WR. Developmental effects of an environmental antiandrogen: the fungicide vinclozolin alters sex differentiation of the male rat. Toxicol Appl Pharmacol 1994; **129**: 46–52.
7. Wolf CJ, LeBlanc GA, Ostby JS, and Gray LE Jr. Characterization of the period of sensitivity of fetal male sexual development to vinclozolin. Toxicol Sci 2000; **55**: 152–161.
8. Clark RL, Anderson CA, Prahalada S, Robertson RT, Lochry EA, Leonard YM, Stevens JL, and Hoberman AM. Critical developmental periods for effects on male rat genitalia induced by finasteride, a 5α-reductase inhibitor. Toxicol Appl Pharmacol 1993; **119**: 34–40.
9. Anderson CA and Clark RL. External genitalia of the rat: normal development and the histogenesis of 5α-reductase inhibitor-induced abnormalities. Teratology 1990; **42**: 483–496.
10. Bentvelsen FM, Brinkmann AO, van der Schoot P, van der Linden JETM, van der Kwast TH, Boersma WJA, Schröder FH, and Nijman JM. Developmental pattern and regulation by androgens of androgen receptor expression in the urogenital tract of the rat. Mol Cell Endocrinol 1995; **113**: 245–253.
11. Yoshimura S. Effects of prenatal or postnatal exposure to chemicals on preputial separation of male rats (unpublished).

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

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

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

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

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

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

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

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

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

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

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

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

The authors declare they have no competing financial interests.

Received 22 April 2005; accepted 17 August 2005.

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

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

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

Validation of Toxicogenomics: Focus on the Biological Systems

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

One-Off Validation

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

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

Table 1. Recommendations: focus on biological systems.

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

Routine Validation

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

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

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

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

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

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

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

Validation as a Result of Procedural Changes

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

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

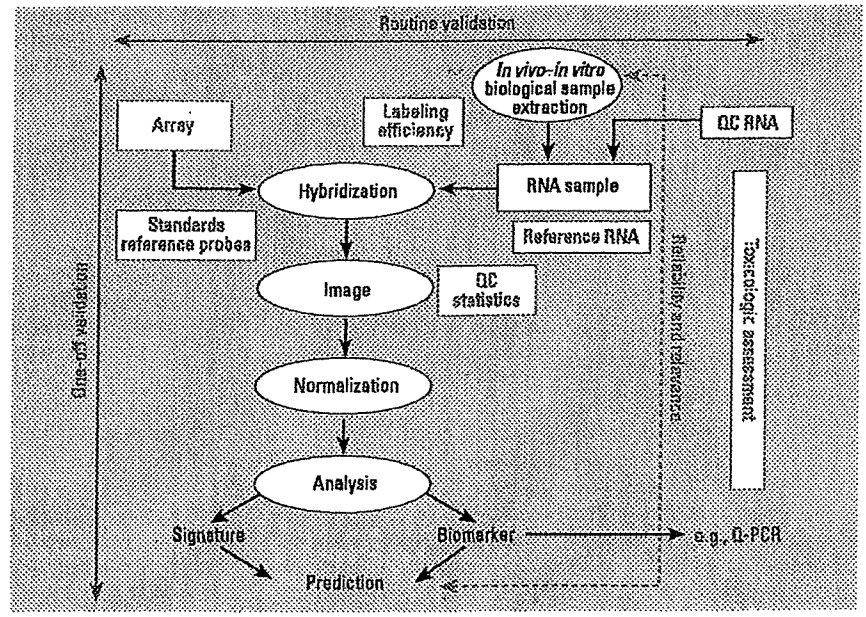

Figure 1. Scheme of the different steps in a toxicogenomics-based test. Three distinct levels were identified where validation is necessary: one-off validation (left), which should be performed once and is mainly related with the quality of the microarray and the instrumentation (blue); routine validation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

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

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

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

Data Management

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

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

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

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

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

Regulatory Acceptance of Validated Toxicogenomics-Based Methods

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

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

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

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

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

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

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

Table 2. Recommendations: focus on technology.

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

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

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

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

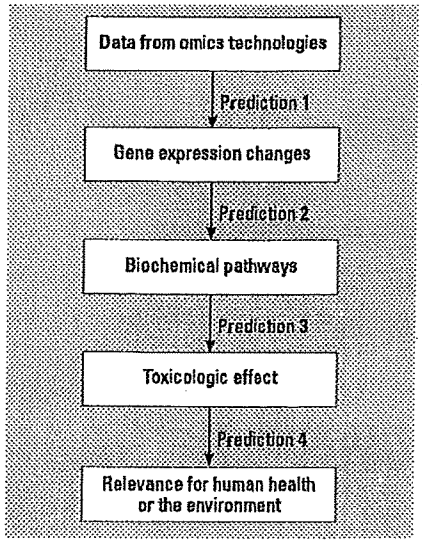

Figure 2. Process flow showing different independent prediction levels considered important in assessing validity of a toxicogenomics-based test method.

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

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

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

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

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

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

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

Conclusions

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

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

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

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

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Kuang Y, Biran I, Walt DR. 2004. Simultaneously monitoring gene expression kinetics and genetic noise in single cells by optical well arrays. Anal Chem 76(21):6282–6288.

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

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

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

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

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

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

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

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

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

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

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

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

Steiner G, Suter L, Boess F, Gasser R, de Vera MC, Albertini S, et al. 2004. Discriminating different classes of toxicants by transcript profiling. Environ Health Perspect 112:1236–1248.

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

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

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

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

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

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

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

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