To Whom It May Concern:

This document is a critique of "RESPONSE OF LARVAL *XENOPUS* LAEVIS TO ATRAZINE: ASSESSMENT OF GROWTH, METAMORPHOSIS, AND GONADAL AND LARYNGEAL MORPHOLOGY" by Carr *et* al. and a review of the findings and differences in the effects of atrazine in *Xenopus laevis* as described in "HERMAPHRODITIC, DEMASCULINIZED FROGS AFTER EXPOSURE TO THE HERBICIDE, ATRAZINE, AT ECOLOGICALLY RELEVANT DOSES. Below I address several concerns with the methods used in the Carr study. In brief, there are several issues regarding the husbandry and methodology that may explain the differences in our findings, but there are also concerns regarding normal husbandry and care, and several items that simply are not clear. Most significantly poor husbandry practices (over crowding, underfeeding, and unclean water and tanks) in the Carr study resulted in animals with poor health (50% or more of the animals do not metamorphose by 80 days, and those animals that do reach metamorphosis show a negative growth curve and poor development). Thus the data on development have a limited usefulness and reveal little about the true effects of atrazine on amphibian growth and development.

Table 1. Brief Summary of differences in experimental treatments

	LISAB	Carr *et al.*
Animal Source	UCB, Nasco (lab reared)	Xenopus Express (Wild caught)
Medium	0.1 x Holtfreter's	FETAX
Atrazine amount	0.04-1000 µg	variable
Co-solvent	ethanol	none for atrazine treatments
Larval selection	apportioned in groups of five	sorted before hatching
Replicate #	3 per treatment	variable (6-11)
Exp replicated	4 times (including 99XLATZ2)	1
Container size	5 L tank	variable (250, 1 L beaker or 10 L aquaria)
Container type	plastic mouse box	variable (glass beakers or glass aquaria)
Tank covered	yes	no/not reported
Water volume	4 L	variable (100 mls to 4 L)
Larvae per repl	30	variable (60-65)
Larval density	30/4L	variable 60-65/(100 ml, 2L or 4L)
Food type	Purina rabbit chow	Nasco Frog Brittle
Food amount	increased with growth (.32- 1.6 g)	0.4 g (did not increase with growth)
Feed frequency	daily	once per 72 hours
Water change	100%, every 72 hours	50% every 72 hours (possibly variable)
Tank clean	every 72 hours	Never
Exp end	after all animals metamorphose	terminated at day 80 (unspecified number of larvae euthanized)
Blindness	Dosing, lab work and analysis	dosing only
Time to met	based on all animals	based only on animals that completed metamorphosis by day 80 (seems to be about 40%)
Sex ratio	based on all animals	based only on animals that completed metamorphosis by day 80 (seems to be about 40%)
Laryngeal size	measured as cross –sect area	measured as cross sectional area, but manipulated in various ways: proportion body weight transformed by square root arcsine with some values substituted with 1/4n (justification unknown).
	Based on a min of 5 animals of each sex per replicate (10 in the PNAS exp).	variable sample sizes, seems to be n = 1 per replicate for females, or some replicates were not sampled (not clear).

Table 2. Brief Summary of similarities in experimental treatments

	LISAB	Carr *et al.*
Species	*X. laevis*	*X. laevis*
Temperature	21-22°C	21-22°C
Light:dark cycle	12:12	12:12
Aeration	continuous	continuous

SUMMARY

Tables 1 and 2 demonstrate that our experiments are not comparable in any way. The ONLY similarities are the species (although different populations), the temperature, light cycle and the aeration. Certainly, the conclusion that my studies (LISAB) cannot be replicated by three studies in two separate laboratories (as stated in the press and to the EPA) is not warranted. The current paper represents a single experiment, in a single laboratory at a single University, but coauthored by individuals from several universities. Further, their studies do not test their own stated hypotheses. There is apparently a concern over the co-solvent ethanol, yet they do not test atrazine alone against atrazine + ethanol in any of their treatments. In fact, I have tested atrazine in the presence and absence of ethanol and there are no interactions with the solvent. In addition, their current work does not support the hypothesized interaction between ethanol and atrazine either. Their FETAX controls and atrazine-treated groups had hermaphrodites, but not their ethanol treated groups (they did not test atrazine with ethanol). The second concern was that there was some effect of Holtfreter's solution. They did not test this question either. To test the effect of Holtfreter's my experiment should be replicated with the only difference being the use of Holtfreter's vs FETAX. Table 1 clearly shows that the rearing medium was not the only difference between the LISAB and NSE studies, so no conclusions can be drawn with regards to the rearing medium. The most likely explanation for their findings is poor husbandry: dirty tanks, unchanged water, and most importantly, starved animals (see below). Starvation is well known to inhibit growth, metamorphosis, and sexual differentiation (which explains their background intersex/undifferentiated controls, absence of atrazine effects, and inability to achieve 100% females with estrogen).

DETAILED ANALYSIS OF CONCERNS WITH THE CARR STUDY

Below I address concern in the order that they appear in the ms.

Page 6, lines 14-16: "...there are no standardized approaches to assessing developmental exposure effects on development in *X. laevis*, and differences in strain of animals, the type of rearing medium used, and replicate sample size can add to the normal variability that is associated with differences in testing methodologies."

> Relevance to LISAB studies: All of our protocols and SOPs were received and approved by Novartis/Syngenta/Ecorisk. So, in fact, there are "standardized approaches". It is not clear why they deviated (see Table 1) from every aspect of the established protocols and SOPs (see below). Further, if the intention was to test parameters such as strain, rearing medium, and replicate size, then why were all parameters changed with no controls? Changing all of the husbandry and dosing patterns makes it impossible to replicate LISAB studies and impossible to test any one factor.
>
> Husbandry and Care Concerns: Many of the husbandry practices do not promote growth and development of larvae.

General Concerns/Questions: Carr *et al.* used untested non-standardized husbandry practices. The many parameters that have been altered make it impossible to evaluate their results, let alone compare to LISAB studies.

Page 7, line 16-17: "Sexually mature male and female *X. laevis* imported from South Africa were purchased from Xenopus Express"

Relevance to LISAB studies: A different vendor was used for animal purchase. LISAB sources (as cited in PNAS) were from UCB or from NASCO Inc. (both sources are laboratory reared for more than five generations). Different populations may have different sensitivities to atrazine. Further, the source used by Carr *et al.* is from South Africa. Sources suggest that atrazine levels are very high in many aquatic habitats in South Africa. Ongoing studies by LISAB examine whether historically exposed populations may develop atrazine-resistance. It is very possible that animals from this wild population display atrazine resistance and would not show the same sensitivity. The authors were obviously aware of NASCO, because they used this company as a food source. In addition, many times (more than six per year) I offered to supply the authors with animals from the UCB colony free of charge. In, fact, on several occasions I suggested that we each share embryos so that we would work on the same animals. So why did they choose not to obtain animals from NASCO (or from me), which would allow comparisons between their study and LISAB?
Husbandry and Care Concerns: None
General Concerns/Questions: This source is not commonly used. I am curious why imported animals from an unknown source and history were used. Where in South Africa were they collected, what habitat, history of contaminant exposure, etc? More common sources are NASCO, XENOPUS 1.

Page 8, lines 3-5: "Starting on posthatch day 5, larvae were fed 0.4 grams of powdered frog brittle (Nasco, Ft. Atkinson, WI)...every 72 h"

Relevance to LISAB studies: Animals were starved for the first five days and a different food and a different amount of food was used. Food type alone can affect growth and development (Kupferberg, '97; Buchholz and Hayes, '00). Some differences in food content are shown below. In addition, the Nasco frog brittle contains animal meal whereas the Purina rabbit chow is plant-based (but does contain animal fat). Further, our work on another species (Buchholz and Hayes, '00) showed that Purina rabbit chow promoted better growth and development when compared with a fish flake food with a composition similar to Nasco frog brittle (Table 1). Thus, it is possible that the quality of the food, in addition to the low amount and frequency of feeding contributed to poor growth and metamorphosis (discussed later).

Table 3. Nutritional content of tadpole food

	Nasco Frog Brittle	Purina Rabbit Chow	Tetramin Fish Flakes
Protein	44%	14%	46%
Crude Fiber	2%	18.8%	2%
Fat	6%	1.5%	7%

Furthermore, the animals were fed less frequently than in the LISAB study. In the LISAB study, animals were fed daily, but larvae were fed once every 72 hours only in the Carr study. Also, we fed larvae a prescribed increasing amount of food to match the growth of the animals (LISAB SOP#30). By day 17 post-hatching, we fed 30 tadpoles 1.6 g ground rabbit chow every day, compared with 0.4 g per 60 tadpoles once every three days in the Carr study. Thus, over a three-day period, the LISAB animals were fed 24 times more food (grams food/tadpole/day) than the animals in the Carr study. In addition, Carr's animals were starved for the first five days. The data provided in 99XLFOOD2 study show that the food level used in the Carr study would result in severe delays in metamorphosis, slow growth, and high mortality. In fact, LISAB SOP#30 warns under "Animal Health" considerations:

"**B. Animal Health**

1. The amount of chow provided to *Xenopus laevis* should be increased as animals grow. Underfeeding will result in retarded growth, development and inability to metamorphosis, as well as increased mortality."
(LISAB, SOP#30, page 1)

The consequences of such poor feeding are demonstrated by the figures below. Their food level would result in retarded growth and development, low metamorphic rates and eventually high mortality.

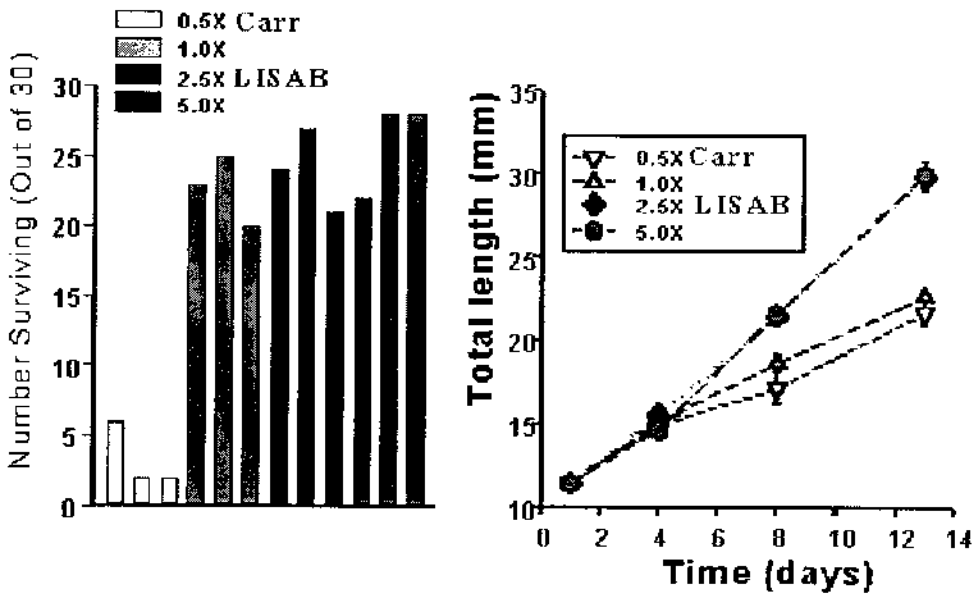


Figure 1. Effect of food level on growth (left) and mortality (right). Individual bars (left panel) represent replicate tanks for each treatment (see legend). Feeding levels similar to those of the LISAB and Carr studies are indicated in the legends. In the original experiment (98XLATZ1) shown here, four food levels were tested to determine why we experienced unusually high mortality in an experiment (98XLATZ1). This experiment (98XLATZ1) differed from our usual husbandry in that the water was changed and tanks were cleaned every 72 hours as opposed to every 24 hours as was the practice before our interactions with Novartis/ Ecorisk. As a result of this change in husbandry (which was required by Novartis), we reduced the amount of food to (1×) to avoid fouling the water (previously, we fed animals *ad lib*). Food levels of 1× to 5× allowed normal survivorship (right panel), but only 2.5-5× allowed normal growth. Note that 1× (0.15 g / 30 tadpoles/ day) is similar to the amount of food used in 98XLATZ1, 2.5× (0.3 g / 30 tadpoles/ day) reflects the starting amount of food that was used in 99XLATZX2, 00XLATZ46, 5, and 7 and is now standard in our laboratory. The amount of food used in the Carr study (0.4 g / tadpoles), is similar to 2.5× but delivered to twice as many tadpoles once every three days. So, the food amount used by Carr *et al.* is equal to 0.5× (g food/ tadpole/ day), a dose that resulted in retarded growth and development, poor health and high mortality in a study provided to Novartis/Syngenta/Ecorisk (and signed off). Further, in their current study, Carr did not increase the feeding amount as animal grew (as prescribed by the LISAB SOP#30).

These protocols, SOPs, and final reports were delivered to Novartis/Ecorisk in 1999. So, it is not clear why their reported feeding protocol was adapted. Poorly fed animals are not likely to have gonads and secondary sex characters that are as

developed as well fed animals. The reportedly small size of their animals (less than 0.3 g for the latter half), and inhibition of development and metamorphosis (only about 40% metamorphosed in 78 days, compared to 80-90% metamorphosis in 50 days in the LISAB studies) reported in their results are testimony to the poor condition of their animals. In fact, the effect of starvation on growth, development and metamorphosis is well-established. Swingle ('18) reported on the effect of starvation (inanition) on amphibian development in 1918. In "The effect of inanition upon the development of the germ glands and germ cells of frog larvae", he reports in his conclusions:

"1. Total starvation inhibits indefinitely the growth and the metamorphosis of larval frogs.
2. It prevents the development of germ glands and delays any increase in the number of germ cells, interstitial cells and other tissue elements in the gonads.
3. Starvation greatly retards the normal cycle of development of the germ cells.
4. It prevents the onset of sexual differentiation" (Swingle, '18; page 564).

In fact, this study from 1918 explains many of the disparate findings in the NSE study: The incomplete sex reversal and differentiation of the gonads in the estrogen-treatments, the low incidence of hermaphroditism in atrazine-treated frogs (they simply are not differentiated yet), and the "background" occurrence of intersexes (most likely undifferentiated) gonads in controls.

Husbandry and Care Concerns: The Carr feeding regime is inappropriate and has been shown to compromise growth, development and metamorphosis. The data reported in this study reflect this as well. The animals produced in the Carr study were half the size of LISAB animals. In my laboratory, *X.* laevis typically metamorphose at approximately 0.6 g. In fact, the mean weight of the slower animals in the Carr study (less than 0.3 g) is the size of the smallest metamorph that we have ever produced in our laboratory (out of 10,000s in over ten years). Even larvae dosed with DDT are larger than this at metamorphosis (Noriega and Hayes, submitted). Furthermore, mean time to complete metamorphosis (Stage 66) was 48 days in our study and ranges from 40-50 in general (multiple animal sources over the last ten years). Their reported mean time of 80 days indicates that the animals were starving, unhealthy and not developing properly. Even animals inhibited by estrogen-treatment metamorphose within 50 days at this temperature. Similarly, DDT-dosed animals metamorphose on average within 50 days. Furthermore, I am not sure how they can report time to metamorphosis as they report that "The experiment was terminated on post-hatch day 80." (page 11, lines 1-2). Also, "Any remaining tadpoles were processed as described"... and not included in any of the data. Thus, the authors have biased the data set against the slower developing animals. How many animals have been excluded from the study? What stages were the animals that were terminated? How can mean time to metamorphosis be calculated, if some unspecified numbered of animals (although I believe that as many as 60% of the animals did not metamorphose, see below) were terminated prior to metamorphosis?

General Concerns: There are many unanswered questions. See above.

Page 9, line 18: "60-65 *X. laevis* larvae per replicate beaker"

Relevance to LISAB studies: A different number of animals per replicate was used. We always rear larvae at 30 animals per replicate (4 L) from hatching until metamorphosis. Tadpole density affects growth and development even if food is *ad lib* (Gromko et al, '73; Newman, '94; Tejedo and Reques, '94; Hayes, '97a; Buchholz and Hayes, '00; and many others, I could probably pull out 100 references in the last ten years). At the start of the experiment, the animals density was 80 times higher than the LISAB study, eight times higher, followed by two times higher at the end of the their study.

Tadpole density affects steroid metabolism and alters the effective dose (μg / g animal) (Hayes and Licht, '95). Depending on the species and the compound, two tadpoles are capable of completely metabolizing steroids eight time faster than one (Hayes and Licht, '95). Similar density effects have been shown for DDT exposure where effective doses are higher when animal density is increased (Licht, '76; Cooke, '79; Hayes, '97b). Thus, the threshold dose for effects on the gonads may very well be higher in their study because the effective dose was lowered by loading more tadpoles into each replicate. This difference is analogous to giving a tablespoon of medicine to one child vs dividing the same dose between two children.

Husbandry and Care Concerns: The reported density is very high. One becomes concerned about tadpole waste, water quality, etc.

General Concerns: Why is the density 60-65 larvae per replicate? Were the animals not counted accurately? This means that different replicates differed in animals density? If so (see concerns above), each replicate is confounded by density differences. Also, were the animal apportioned to prevent bias? For example, LISAB swirls a large pool of animals and they are apportioned to tanks five at a time until all tanks are filled to avoid biases in behavior or size between the first and last replicate filled (see LISAB SOP#25).

Page 9, line 22: "diluted in FETAX medium"

Relevance to LISAB studies: LISAB used 0.1× Holtfreter's solution (made with filtered, UV-treated, deionized double distilled water) not FETAX. It is not clear how this difference might affect results, but the differences in ions are shown below:

Table 4. Difference between 0.1 × Holtfreter's[1] and FETAX[2] media

Component	0.1 × Holtfreter's (mg)	FETAX (mg)
NaCl	350	625
KCl	5	30
$CaCl_2$	10	15
$NaHCO_3$	20	96
$CaSO_4 \cdot 2H_2O$	0	60
$MgSO_4$	0	75
ddH_2O	1 liter	1 liter

1. Holtfreter, '31
2. ASTM, 1991

It is possible that the difference in ions could affect growth and development, but no studies are available to show how (although effects of osmotic pressure are known, I know of no studies that specifically test Holtfreter's against FETAX). It is not possible that ions/nutrients in the LISAB water altered the effectiveness of Holtfreter's, however, as we used deionized, UV-treated, filtered, distilled water when making Holtfreter's solution.

Husbandry and Care Concerns: FETAX medium (Frog embryonic teratogenic assay: Xenopus) is designed for pre-hatching frog embryos. It is not clear if larvae would develop optimally in this medium. In fact, in information that I obtained from John Bantle (the developer of FETAX), he states "Disadvantages of FETAX: Occasionally, some tests have to be terminated because of excessively high numbers of malformations in controls." Bantle (1991; page 3). My laboratory has reared 10,000s of larvae (more than 30 species) in 0.1× Holtfreter's per year for the last ten years and we have never had to terminate an experiment because of "excessive malformations in controls". Holtfreter's was developed in 1931 (at least) and has been used widely. In particular, I note that Kristen Lopez (1989) used 0.1 X Holtfreter's solution in her work on sex differentiation in *Bombina orientalis*. This work was done at the University of Colorado, Boulder in the same department that Carr received his PhD. In fact, Carr overlapped with Lopez during her study, which was published in 1989 and Carr and Norris ('88) used Holtfreter's solution for rearing amphibians.

General Concerns: Why was 0.1× Holtfreter's not used? I know that there is a claim of high mortality with Holtfreter's, however, as described in this review (Table 1), animal density, water change frequency, volume of water, food type, feeding frequency, food amount, animal source, apportioning strategy, length of exposure, experimental termination, etc. all differed from the LISAB studies. So, was there really a problem with Holtfreter's or was the problem one of the many other factors or a combination of other factors that differed?

Page 10, lines 1-4: " On post hatch day 5, larvae were transferred from 250 ml glass beakers to 2 L glass beakers containing 1 L of treatment matched solutions for 14 d. On day 21, larvae were transferred to treatment-matched 10L glass tanks containing 4 L FETAX medium and appropriate test or reference solutions."

Relevance to LISAB studies: These changes in volume mean that the animal density and available water volume were all changing every few days. It is not clear why this was practiced, especially because the recent work of this group (Coleman et al., '02a,b) did not use such a confusing rearing regime: Larvae in this published study were reared at a reasonable density (56 tadpoles per 8 L, similar to the density in our study). The consequences of the current husbandry are that they initially reared animals at a density of 600 tadpoles / L (60 larvae in 0.1 ml) which is 80 times higher than the density in LISAB studies! Larvae were then changed to a density that was eight times higher than our study, followed by 60 days at a density that was two times higher than the LISAB study (and with a food levels that was 24 times less). As written, it appears that there was no established protocol and the husbandry was changed throughout the experiment. Husbandry for LISAB studies has remained consistent and practiced since 1990 (with the exception of the water change renewal practice which was introduced by Novartis in 1998). The table below summarizes the variability in their husbandry:

Table 5. Variability in animal rearing condition in the Carr study

Days	Volume (L)	Density[1] (Larvae/L)	Relative to to LISAB[2]	Container (type)
2-5	0.1	600	80	glass beakers (250 ml)
5-19	1	60	8	glass beakers (2 L)
21-80?	4	15	2	glass/silicone aquaria (10 L)

1. Although the number of larvae per replicate was not controlled (reported as 60-65), I chose 60 here to produce the table.
2. Calculated as: (density of animals in Carr) / (density of LISAB animals = 7.5 larvae/ L)
? There is some confusion here. The authors state that "On post hatch day 5, larvae were transferred from 250 ml glass beakers to 2 L glass beakers containing 1 L of treatment matched solutions for 14 d." which would mean that they were maintained under these condition until day 19 (5 + 14 = 19). However, they do not report moving the animals to the aquaria until day 21. It is not clear where the animals were maintained between days 19-21.

In addition, the varying volumes and 50% water change affect the available dose of atrazine. Although activity and metabolism of atrazine may be concentration-dependent, the ability of the animals to deplete the water of compounds (metabolize them) is dependent on the absolute amount of compound. Hayes and

Licht ('95) showed that the rate of steroid metabolism by *X. laevis* larvae is concentration dependent, whereas the ability to completely deplete the water is dependent on the absolute amount of hormone added to the water. This factor is of particular concern given that the water is only 50% changed, meaning that the exposure amount is only half of what is reported. In addition to the effective lower doses, the parent compound is potentially competing with the accumulating metabolites. The table below summarizes the true doses. At best (day 21), Carr's tadpoles are exposed to 50% of the atrazine dose that LISAB used, but otherwise are exposed to as little as 0.6% of the dose that LISAB used (day 5). Thus, their 25ppb dose is equal to 1.565 ppb relative to LISAB.

Table 6. Variation in the total amount of atrazine added to larvae at the reported 25 ppb (25 µg /L) dose. [1]

Days	Atrazine[2] (µg) (Carr)	Atrazine (µg) (LISAB)	Atrazine[3] (µg/tadpole) (Carr)	Atrazine (µg/tadpole) (LISAB)	Relative[4] to LISAB
1	2.5	100	0.04	3.33	0.012
5	1.25	100	0.02	3.33	0.006
8-19	25	100	0.42	3.33	0.126
11-12	12.5	100	0.21	3.33	0.063
21	100	100	1.67	3.33	0.500
24-80	50	100	0.83	3.33	0.250

1. Here, we calculate difference for the 25 ppb dose only, but similar problems exist for all doses of all compounds. Carr's values are shown in red.
2. The total weight of atrazine added to each replicate (µg/L × total volume of medium). We assumed that 100% of the medium is changed when the containers are changed, afterwards, the amount of new compound added decreases by 50% because only 50% of the water is changed. The total amount of atrazine is controlled for LISAB and variable for Carr *et al.*
3. Amount of atrazine per tadpole (µg atrazine / total number of tadpoles). The number of tadpoles was 30 for LISAB, but not controlled for Carr *et al.* (we used 60 to calculate values).
4. Comparison of amount of atrazine added:
(µg/tadpole for Carr *et al.*) / (µg/tadpole for LISAB).

Husbandry and Care Concerns: The densities in the Carr study are all very high and likely contributed to the small size and inhibition of metamorphosis in these animals. In addition, clear glass, which lets in more light likely contributed to the growth of algae and perhaps other organisms that may have grown in tanks as a result of not cleaning them (see below). Also, Niewkoop-Faber points out, "The larvae do not need much light; they develop just as satisfactorily in total darkness. Certainly too much light interferes with normal development." (Nieuwkoop and Faber.'94). Opaque covered tanks (as we used) are a better alternative.

Finally, there is concern that steroids and other contaminants may stick to silicone in the aquaria. Certainly, it has been our experience that contaminants and steroids may be sequestered by the silicone.

General Concerns: Another concern is that the initial small volume (100 mls) contained 60-65 larvae and was aerated. The aeration must have really stirred the water and the animals. Other concerns are the evaporative loss from aerating this small volume. Covering the containers (as LISAB did) would prevent this. Further, how much atrazine and steroid are aerosolized? We cover all treatments to prevent cross contamination associated with aerating treatment solutions. In addition, although aquatic, in our experience *X. laevis* are able to crawl up the sides of containers and escape and/or enter other treatment cages. If tanks were not covered this is a concern.

Page 10, lines 4-6: "Eleven replicates of each atrazine and untreated FETAX control treatment were performed. Six replicates each were used for the ethanol vehicle control, DHT, or estradiol treatments."

Relevance to LISAB studies: The replicate size is higher than the LISAB study and should provide more statistical power.

Husbandry and Care Concerns: None.

General Concerns: Why 11 replicates? What was the justification for this number of replicates? Also, why does the number of replicates differ between the tanks with no co-solvent and the tanks with ethanol at the end of the experiment? Further, why does the number of replicates change? In the preceding paragraph (page 9, lines 18-21), the authors state that "At 48 h post-hatch, 60-65 *X. laevis* larvae per replicate tank (11 replicates per treatment) were exposed to a single concentration of atrazine (nominal concentrations of 0, 1, 10, or 25 µg atrazine/L diluted in FETAX medium), DHT (100 mg/L, nominal), estradiol (100 mg/L, nominal) or ethanol alone in FETAX medium", yet by day 21 there were only six replicates of ethanol and hormone-treatments. What happened to the other five replicates of these treatments? Why were the sample sizes not equal? How were the unequal sample sizes addressed in the statistical analysis?

Page 10, lines 7-8: "A 50% change of test and control solutions was performed every 72h."

Relevance to LISAB studies: The LISAB study changed 100% of the water and solutions every 72 hours. A 50% water change (as in the Carr study) means that decaying food and other matter were maintained in the cages. Also residual atrazine and metabolites remain in the tank and possibly interfere with the new parent compound that is added. Also, when LISAB personnel conduct complete water changes every 72 hours, each tank is cleaned (scrubbed, wiped, and rinsed)

(see LISAB SOP#29). If only a 50% water change was conducted by Carr *et al.*, then tanks were not cleaned. Algae, fungus and other organisms likely grew on the tanks and may have contributed to metabolism or even sequestering the test compound.

Husbandry and Care Concerns: Not cleaning the tanks likely contributed to low growth and development / poor health of these animals.

General Concerns: See above.

Page 10, lines 19-20: "Each animal was given a unique identification number that included the study number, treatment color, replicate tank letter, and animal number."

Relevance to LISAB studies: In the LISAB studies, all lab procedures and analyses were conducted blindly. Animals were given ID numbers and tanks and treatment data were maintained elsewhere. If treatment and tank information were maintained with the animals, there may have been some bias in how the animals were analyzed/ treated during laboratory analyses.

Husbandry and Care Concerns: None

General Concerns: None

Page 11, lines 1-3: "The experiment was terminated on post-hatch day 80. Any remaining tadpoles were processed as described for NF Stage 66 animals above. Only data from NF-stage 66 animals were used for the present studies."

Relevance to LISAB studies: The LISAB study was carried out until all animals metamorphosed. In the Carr study, as few as 50% of the animals (fewer than 40% in some cases) reached metamorphosis (complete tail reabsorption, stage 66) and thus, most of their animals were not included in the analysis. What stage were the animals that were terminated? As discussed above, terminating the experiment biased all of the data. Basically, the faster developing animals have been selected. The gonads of the faster developing animals are not as developed as the later animals. In fact, in *Rana pipiens* the earlier larvae to metamorphose have less severe hermaphroditism than the latter (Hayes et al., unpubl). Further, killing animals that have not metamorphosed biased the estimation of time to metamorphosis (the slower animals have been eliminated). This is both alarming and telling, as the time to metamorphosis in the Carr studies is already two times longer than expected. In LISAB studies, animals always metamorphose between 40-50 days post hatching (even DDT and E2-treated animals metamorphose within this time) at 21-22°C. Niewkoop-Faber ('94) states that *X. laevis* reared at 22-24°C reach stage 66 by 51 days post-hatching.

Page 11, line 13: "In the initial analysis, 40-45 males and 11 females were analyzed per treatment"

Relevance to LISAB studies: No specific comments. See below, however?

Husbandry and Care Concerns: NA

General Concerns: How many animals per replicate were analyzed? Why is there a range for males (40-45), but a specific number for females (11)? Why were unequal sample sizes analyzed for males and females and how were these sample sizes chosen? How was the unequal sample size dealt with statistically? The statement describes "initial analysis", what was the sample size for the final analysis? Also, is this statement supposed to say "...were analyzed per replicate?" If the statements is indeed correct as written, then only one female was sampled from each replicate ("11 females were analyzed per treatment") as there were 11 replicates for the atrazine treatments. Alternatively, does this mean that not all replicates were sampled? If an equal number of animals were sampled per treatment, does this mean that more animals per replicate were sampled in treatments that had only six replicates? If so, how was this addressed statistically?

Page 12, lines 6-9: "Sex ratios were determined by direct visual inspection of 276 to 334 animals per treatment for the atrazine and FETAX medium controls (1,215 animals total) and 135-160 animals per group from the ethanol vehicle controls and steroid treatments (446 animals total) as described previously."

Relevance to LISAB studies: Sex ratios in the LISAB studies were based on analysis of the entire population of treated animals. Why was the population sub-sampled by Carr to determine sex ratio? Fewer than 50% of the animals dosed are used here, why? On page 11, line 3, the authors state "Only data from NF-stage 66 animals were used for the present studies." Does this mean that the 60% of the animals not included in the sex ratio determination were animals that did not metamorphose and were terminated at day 80? If so, then the sex ratio only reflects the faster developing animals. The sex ratio in the Carr study was based on only 41-46% of the animals (based on the exact number of animals reported in the analysis divided by the calculated total number of animals per replicate—I used 60-- times the number of replicates), yet they reported little mortality. Do the missing 2059 animals not included in their analysis reflect the number that were terminated on day 80? If not, why was the sex ratio based on less than 50% of the animals in the study? Using only 50% of the animals would likely bias the detection of hermaphrodites. Furthermore, if we consider the number of replicates, it appears that fewer than 40% and (when variation between tanks is considered) as few as only 30% of the animals metamorphosed from each replicate. In fact, this is likely the case as they report that only about 56% of the animals developed to foreleg emergence and only about 32% reached complete tail reabsorption by day 70 in their recent studies (Coleman, 02a,b).

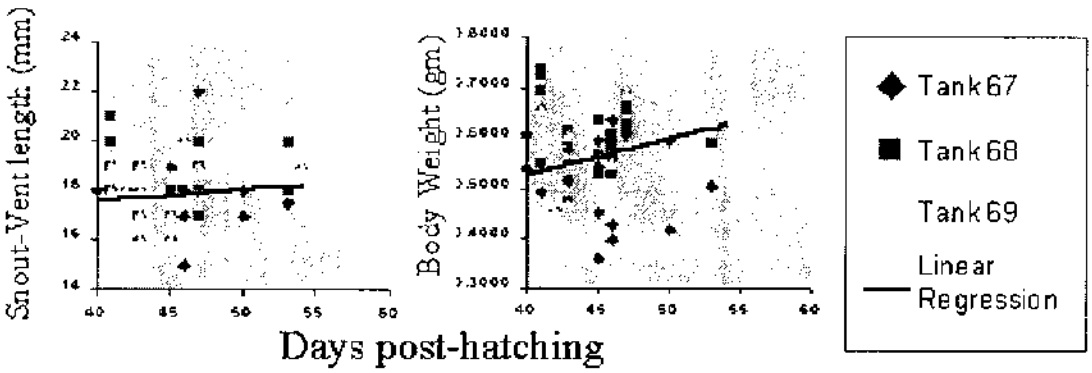
Also, hermaphrodites were more likely and more severe in slower developing animals in our current work (Hayes *et al* submitted). It is likely that the slower developing larvae in *X. laevis* (which would be exposed for a longer time) were more likely to show abnormalities at metamorphosis, but these animals were eliminated from the current study, by termination at day 80.

Husbandry and Care Concerns: NA

General Concerns: See above. I am concerned that more than 50% of the animals in the experiment were not used when determining the sex ratio.

Page 14, line 15-16: "...time to complete metamorphosis (NF-Stage 66) varied inversely with body weight and SVL (r^2 = 0.36 for body weight, n = 1,211, and SVL, n = 1,215)."

Relevance to LISAB studies: This finding is quite troubling. In fact, the first animals to metamorphose are almost twice the body weight and 130% longer than the last animals to metamorphose (Carr et al, Fig, 1). Healthy animals should be larger as they take longer to metamorphosis (as the larval period is a growth period). By definition, growth is an increase in size over time, the opposite of what Carr *et al.* describe. Figure 29 in the LISAB-99XLATZ2 final report shows the appropriate relationship, animals that take longer to metamorphose are larger.



Also, the figure below shows data from control *X. laevis* taken from 00XLATZ46.

Figure 2. Size (snout-vent length; left and body weight; right) vs time to metamorphosis (days post-hatching) in control *Xenopus laevis*). Slower animals with longer larval periods should increase in size when healthy.

Although there is a trend towards increasing size at metamorphosis, there is no statistical difference between the first half and the last half of the animals to metamorphose in LISAB studies. In fact, this noted difference in the Carr studies is an indication that the animals are unhealthy and not growing. Below I compare

the differences in size of the first and last half of controls for LISAB and the Carr study.

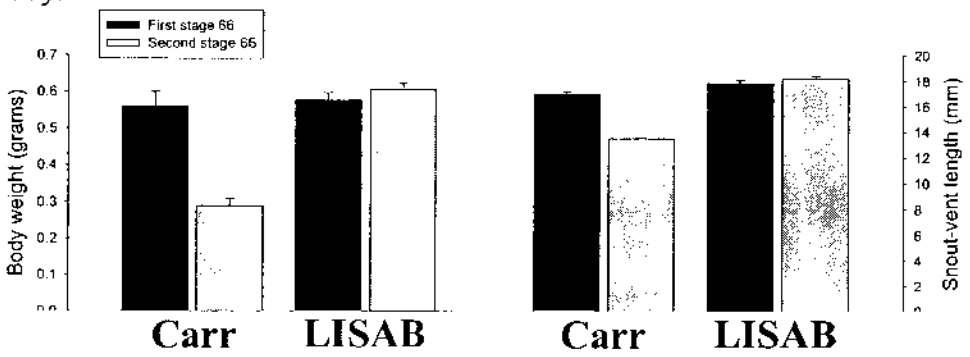


Figure 3. Body size (weight, left and snout-vent length, right) for the first and last animals to metamorphose in the Carr and LISAB control *Xenopus laevis*. Carr's data were taken from Figures 1 (Carr *et al.*, submitted). In the Carr study there is a dramatic reduction in size (50% reduction in body weight) rather than the expected increase. For LISAB the first stage 66 represent animals reaching metamorphosis between days 40 and 45 (mean = 42.11, n = 36) and the last stage 66 represent animals reaching metamorphosis between days 45 and 56 (mean = 47.92, n =37; only 2-3 animals per replicate metamorphose after day 50). The total number of animals reaching metamorphosis represents 81% (19% mortality in this group). Neither the difference in body weight nor snout vent length were statistically significant (SVL: F = 1.319, df = 1, $P = 0.285$; BW" F = 1.029, df = 1, $P = 0.314$).

Further, we offer data from *Rana pipiens* that shows a positive relationship between time to metamorphosis and size, unless the animals are stressed by pesticide exposure. When animals are exposed to a mixture of pesticides which compromise their immune system and inhibit metamorphosis by 50% we see a relationship similar to that observed in the Carr study (see below). These data further suggest that the Carr animals are in poor health and reared under poor conditions.

Page 15, Lines 4-6: "Despite weak trends toward fewer atrazine-exposed animals reaching FLE (ANOVA trend test, $F_{1,40} = 5$, $p = 0.03$) and completing tail resorption (ANOVA trend test, $F_{1,40} = 4.5$, $p = 0.04$)..."

Relevance to LISAB studies: LISAB did not observe effects on metamorphosis.

Husbandry and Care Concerns: NA

General Concerns: There is a general concern that Carr *et al.* have a biased interpretation. Here, an effect with a P value of 0.03 and 0.04 are referred to as "weak trends" when statistical significance is typically accepted at $P < 0.05$. Further, at other times in the paper they use $P < 0.05$, and even two lines up from the current statement, they refer to an effect on size as "significantly longer" when they have a P value of $p = 0.02$. Further up they say that an decrease in size with time to metamorphosis is significant ($P < 0.05$). Thus, they are not consistent in their interpretations. It appears that $P < 0.05$ is a weak trend when the result is potentially negative for atrazine, but significant otherwise.

Page 15, Lines 10-12: "Although incidence of edema was weakly correlated with atrazine concentration (Cochran-Armitage trend test, $Z = -2.3$, $p = 0.02$).... Incidence of abnormal swimming was also correlated with atrazine concentration (Cochran-Armitage trend test, $Z = -2.9$, $p = 0.004$, Table 1).... "

Relevance to LISAB studies: Here, they have identified effects that we did not observe in our studies. Possibly the combined stress of starvation, dirty water, and atrazine exposure has produced these abnormalities. Contrary to their statements, effects with P values of 0.02 and 0.004 are not weak correlations.

Husbandry and Care Concerns: The described developmental abnormalities are further evidence that the NSE animals are in poor health.

General Concerns: Statistical significance is typically accepted at $P < 0.05$ ($P < 0.1$ in some ecology/field studies). The authors should explain why they have chosen to refer to effects at $P = 0.02$ as weak. A 98% probability that an effect is associated with atrazine exposure is not weak in my opinion. Furthermore, apparently $P < 0.05$ is accepted as significance in some cases as the exact P values are given when they refer to a low P value as weak, but $P < 0.05$ is given when they discuss other data (such as the effect of estradiol).

Page 16, lines 3- : Exposure to 25 µg atrazine/L or estradiol significantly increased the percentage of individuals with intersex gonads in each tank replicate as determined by gross morphology of the gonads (Fig. 3). Although incidence of intersex increased with increasing atrazine concentration (Cochran-Armitage, $Z = 3.6$, $p = 0.0003$), only 25 mg atrazine/L significantly increased the average incidence of intersex animals per tank compared to FETAX medium controls ($n = 11$ per treatment, KW = 12.4, $p = 0.0061$, Fig. 3). Exposure to 25 µg atrazine/L resulted in a total number of 14 of 296 animals

(4.7%) with intersex gonads versus a total of 2 intersex animals out of 334 (0.6%) in the FETAX medium controls, a total of 3 intersex animals out of 309 (0.97%) in the 1 μg atrazine / L group and a total of 1 intersex out of 276 (0.36%) in the 10 μg atrazine / L treatment group. There was a significant correlation between incidence of discontinuous gonads and atrazine concentration (Cochran-Armitage, $Z = 2.9$, $p = 0.0042$), although 25 μg/L was the only atrazine concentration to increase incidence of discontinuous gonads compared to FETAX medium controls. "

Relevance to LISAB studies: Here, I am very confused at to what the problem is. Despite poor husbandry and altered conditions described above, Carr *et al.* obtains the same results that we do. There was a very strong correlation between atrazine and intersexes and with discontinuous gonads. Although they only report 4.7% intersex, our data in the PNAS article reported total gonadal abnormalities (including discontinuous gonads). In our PNAS article (Hayes *et al.*, 2002), we clearly state "...atrazine produced gonadal abnormalities. Up to 20% of the animals (16-20%) had multiple gonads (up to 6 in a single animal) or were hermaphrodites (with multiple testes and ovaries)" (page *). We did not find "16-20% intersex" as claimed by Carr *et al.* (page 19, line 19). When we add the app. 8% discontinuous gonads to the app. 5% intersex reported in the Carr study, we get 13% gonadal abnormalities (similar to our findings of 16-20% reported in our PNAS article). So, there is not much difference in our findings. When one considers the dosage corrections described in Table 6, at best, Carr's 25 ppb dose is a little more atrazine than our 1 ppb dose. Considering the strong correlations that Carr obtained in their studies and the biased sampling and poor conditions, there really is very little difference in our findings.

Further, I interpret some of the animals depicted in their figures differently. The animal pictures in 4D appears to be an underdeveloped male (without histology, I cannot tell). This morphology resembles a morphology seen when animals are developmentally arrested such as by compounds that block thyroid hormone synthesis (thiourea).

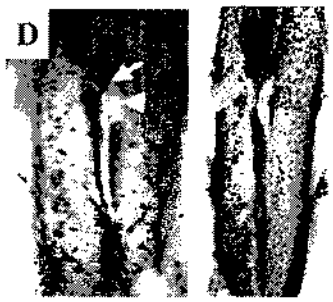


Figure 5. Gonad described as "intersex" by Carr (Figure 4D). The animal shown is most likely Undifferentiated (as a result of inhibited development associated with poor husbandry). The morphology of the Carr animal is similar to the morphology associated with animals inhibited with thiourea (shown on the right, Hayes *et al*, unpubl.).

In addition, the histology shown in Figure 5D appears to be underdeveloped testes, whereas the left gonad in Figure 5E appears to be undifferentiated (intact medulla) rather than intersex. My guess is that some of the intersexes described may in fact be undifferentiated animals and may explain what has been interpreted as a low incidence of intersex in the controls.

Husbandry and Care Concerns: The small body size and inhibited development likely made scoring the gonads more difficult.

General Concerns: See above.

Page 16, lines 16-18: "The percentage of intersex gonads per tank also was significantly (KW = 8.1, $P< 0.01$, Fig. 3) greater in the estradiol treated group (10 out of 135 or 7.4%)"

Relevance to LISAB studies: This is disturbing. They report 7.4% intersex and about 24% males in a group that should be 100% females. We have produced this effect many times with a sample size in the 1000s. In addition, Gallien ('53, '57) Chang and Witschi, ('55a,b) and many others since have shown that estrogen produces 100% females (Marchant Larios in Mexico, Werner in Germany, etc.). The only conditions where LISAB does not obtain 100% females with estradiol treatments is when the treatments are given at low doses or are discontinuous and do not overlap the period during gonadal differentiation.

Husbandry and Care Concerns: Most likely the failure in their positive control group is related to the dosing issues laid out in Table 6. The increased number of tadpoles / dose of estrogen, 50% water change, low volume, competition by estradiol metabolites accumulating in the unchanged water etc all contribute to doses that are lower than the intended dose.

General Concerns: See above.

Page 17, line 4: "Atrazine did not decrease laryngeal dilator muscle size in developing *X. laevis*.

Relevance to LISAB studies: I am concerned that the data reported here differ from other presentations of the data from this laboratory. Previous presentations reported an increase in laryngeal size with the 1 ppb dose. Furthermore, there is concern that 50% or more of the animals (the slower developing larvae) have been eliminated from the study. The size of the larynx varies with time to metamorphosis, depending on treatment. So, by examining only the fastest 40% of the animals in an inhibited population, they have biased their laryngeal data set as well. Even if they chose animals for analysis randomly, the random set is chosen from a biased sample because only the first half of the animals to metamorphose were sampled (The last half were killed on day 80). The figure below shows the bias that was likely created.

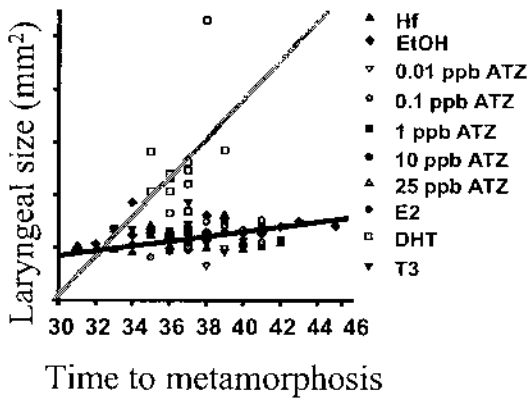


Figure 7. Correlation between laryngeal size and time to complete metamorphosis in males. Relationships are apparent for some treatments, but not others and the relationship varies depending on treatment. The grey line shows the relationship for androgen-treated animals and the brown line the relationship for controls (other treatment groups show a relationship similar to that for controls).

Husbandry and Care Concerns: Still some concern that they have chosen the faster 40% of the animals and that even these were inhibited.

General Concerns: See above.

REFERENCES CITED

Bantle, J. (1991) FETAX, a developmental toxicity screen.

Buchholz, D.R. and T.B. Hayes (2000) Larval comparison for the spadefoot toad *Scaphiopus couchi* and *Spea multiplicata* (Pelobatidae: Anura). *Herpetologica*, 56(4): 455-468.

Carr, J.A. and D.O. Norris (1988) Interrenal activity during metamorphosis of the tiger salamander *Ambystoma tigrinum*. *Gen. Comp. Endocrinol.* 71:63-69.

Chang, C.-Y. and E. Witschi (1955a) Breeding of sex-reversed *Xenopus laevis* Daudin. *Soc. Exp. Biol.* 89: 150-151.

Chang, C.-Y. and E. Witschi (1955b) Genic and hormonal control reversal of sex differentiation in *Xenopus*. *Proc. Soc. Exp. Biol. Med.* 93: 140-144.

Coleman, W.L., L.J. Urqyuidi, T.A. Anderson, E.S. Smith, R. Kendall, and J.A. Carr. (2002a) Environmentally relevant concentrations of ammonium perchlorate inhibit development and metamorphosis in *Xenopus laevis*. *Environm. Toxicol. Chem.* 21(2): 424-430.

Coleman, W.L., J. Carr, and T.A. Anderson (2002b) Environmentally relevant concentrations of ammonium perchlorate inhibit thyroid function and alter sex ratios in developing *Xenopus laevis*. *Environm. Toxicol. Chem.* 21(3): 590-597.

Cooke, A.S. (1979) The influence of rearing density on the subsequent response to DDT

dosing for tadpoles of the frog *Rana temporaria.E nviron. Pollut. Environm. Contam. Toxicol.* 21:837-841.

Gaillien, L. Inversion totale du sexe chez *Xenopus laevis* Daud. à la suite d'un traitment gyoène par le benzoate of oestradiol, administer è pendant la vie larvaire. *C.R. Acad. Sci.* 237:1565-1566.

Gaillien, L. Inversion du phènotype sexuel, altèrations de l a gonadogenèse et determinisme de la differenciation des gonies chez le amphibians. *J. Fac. Sci. Hokkaido Univ. Ser. VI (Zool).* 13: 373-378.

Gromko, M.J., F.S. Mason, and S.J. Smith-Gill (1973) Analysis of the crowding effect in *Rana pipiens* tadpoles. *J. Exp. Zool.* 186:63-72.

Hayes, T.B. (1997) Steroid as potential modulators of thyroid hormone activity in anuran metamorphosis. *Am. Zool.* 37:185-194.

Hayes, T.B., Collins, A., Lee, M., Mendoza, M., Noriega, N., Stuart, A.A. & Vonk, A. (2002) Hermaphroditic, demasculinized frogs after exposure to the herbicide atrazine at low ecologically relevant doses. *Proc. Natl. Acad. Sci. USA* 99, 5476-5480.

Hayes, T.B., T.H. Wu, and T.N. Gill (1997) DDT-like effects as a result of corticosterone treatment in an anuran amphibian: Is DDT a corticoid mimic or stressor? *Environm. Toxicol. Chem.* 16(9): 1948-53.

Hayes, T.B. and P. Licht (1995) Factors influencing testosterone metabolism by anuran larvae. *J. Exp. Zool.* 271(2): 112-119.

Holtfreter (1931) Uber die Aufzucht isolierter Teile des Amphbian Keimes II. *Arch. F. Ent, Mech.* 124: 404-465.

Kupferberg, S.J. (1997) The role of larval diet in anuran metamorphosis. *Amer. Zool.* 37: 146-159.

Licht, L. (1976) Postmetamorphic retention of ^{C14}DDT by wood frog tadpoles. *Comp. Biochem. Physiol. C.* 55: 119-121.

Lopez, K. (1989) Sex differentiation and early gonadal development in *Bombina orientalis* (Anura: Discoglossidae). *J. Morphol.* 199:299-311.

Newman, R.A. (1994) Effects of changing density and food level on metamorphosis of a desert amphibian, *Scaphiopus couchii. Ecology.* 75:1085-1096.

Nieuwkoop, P.D. and J. Faber (1994) Normal Table of *Xenopus laevis* (systematical and chronological survey of development from the fertilized egg till the end of metamorphosis). Garland Publ. Co., NY.

ASTM: American Society for Testing Materials. Standard Guide for Conducting the Frog Embryo Teratogenesis Assay Xenopus (FETAX). Designation: E 1439-91.

Swingle, W.W. (1918) The effect of inanition upon the development of the germ glands and germ cells of frog larvae. *J. Exp. Zool.* 24(3): 545-565.

Tejedo, M. and R. Reques (1994) Plasticity in metamorphic traits of natterjack tadpoles: the interactive effects of density and pond duration. *Oikos.* 71:295-304.