Journal of Reproduction and Development, Vol. 57, No. 5, 2011,

—Technology Report—

Promoting the "3Rs" Principle in Developmental Biology with Early and Convenient Diagnosis of Pregnancy in Mice

Tommy SEABORN[1,3], Julie-Alexandra MOULIN[1,3], Mélissa CÔTÉ[1,3], Yves TREMBLAY[1–3]

[1)]Reproduction, Perinatal and Child Health Axis, [2)]Department of Obstetrics and Gynecology and [3)]Centre de Recherche en Biologie de la Reproduction (CRBR), Centre Hospitalier Universitaire de Québec (CHUQ), Centre Hospitalier de l'Université Laval (CHUL), Québec, QC, Canada, G1V 4G2

Abstract. The 3Rs stand for *R*eplacement of animals in experiments, *R*eduction in the number of experimental animals, and *R*efinement of experiments to minimize animal pain and stress. We propose to address *R*eduction and *R*efinement in the use of mice as experimental models in developmental research. This study focuses on the maternal percentage of weight increase at gestational day 8 ($\%WI_{GD8}$) to diagnose pregnancy early in BALB/c mice. We documented sensitivity, specificity, false positive and negative rates and probability of pregnancy associated with $\%WI_{GD8}$. This predictive model of pregnancy allows for significant reduction in the number of mice to be sacrificed in developmental research. Reported observations and literature suggest that this model is independent of litter size and should be applicable to other mice strains. This procedure allows mice pregnancy detection before midgestation and proposes an ethically sound approach to experimental animal use by optimizing the number of mice used and refining animal manipulation.

Key words: Mating, Pregnancy prediction, Reduction, Refinement, Weight increase

(**J. Reprod. Dev. 57:** 655–659, 2011)

The 3Rs principle proposed by Russell and Burch [1] and now actively promoted in animal-based science by worldwide ethical committees stands for *R*eplacement of animals in experiments, *R*eduction in the number of experimental animals and *R*efinement of procedures to minimize animal pain and stress. The increasing use of mice as experimental models in developmental research prompted us to further address reduction and refinement in experimental processes. This study focuses on early diagnosis of pregnancy, which could permit a significant reduction and refinement of animal usage in developmental research.

In several types of developmental studies, it is important to diagnose pregnancy as early as possible. Vaginal plug observation is an easy method of pregnancy prediction, but it presents a limited specificity, since false positives are often observed. Abdominal palpation of mated mice is also convenient but is only reliable from gestational day 11 (GD11). Assessment of the progesterone concentration in maternal serum is highly sensitive but requires blood harvesting and a quantification method [2], and more importantly, progesterone production significantly increases only in the second half of gestation [3]. Moreover, progesterone is increased in pseudogestation and could thus result in false positives. Biochemical pregnancy-specific markers can also be helpful in pregnancy diagnosis. For instance, pregnancy-associated murine protein-2 (PAMP-2) [2] and murine α-fetoprotein (m-AFP) [2] can be detected in maternal serum from GD10 and GD8, respectively. However, detection of these markers involves sample harvesting that is stressful for animals and time-consuming quantification methods. More recently, ultrasound was used for pregnancy detection in mice [4, 5]. Although this method is highly sensitive and allows pregnancy diagnosis from GD7.5, it requires specialized and expensive equipment.

Since it represents an easy experimental procedure and necessitates unstressful manipulation for animals, maternal weight gain has also been investigated as a tool for pregnancy diagnosis in mice [2]. The authors of this study concluded that pregnancy can be determined with 99% certainty on GD12 by using this method. Reported results suggest that GD8 is the earliest gestational age at which a difference in weight gain is observable between pregnant and nonpregnant mice. However, the possibility of using maternal weight gain at GD8 for pregnancy diagnosis was not specifically discussed. Since the GD8-GD10 interval represents a pivotal period during mice development, we considered it of great interest for developmental studies to characterize more specifically the probability of pregnancy associated with maternal weight gain at the onset of this crucial period. For example, cardiogenesis [6], pulmonary morphogenesis [7], hematopoietic stem cells lineage determination [8], cranial neurulation [9] and several other developmental processes start between GD8 and GD9. At the molecular level, several important developmental key players are modulated during this period of time. For instance, the peak of expression of *N-myc* occurs around GD9.5 [10], and *Notch1* [11] and *Neurogenin-3* [12] expression starts around GD9, to mention only a few.

We describe here a simple procedure that does not require any special equipment and allows for pregnancy detection in BALB/c mice as early as GD8 (term=GD19) based on the percentage of maternal weight increase at GD8 ($\%WI_{GD8}$). Retrospective compilation and statistical analysis of almost 500 matings provided a logistic regression model for pregnancy prediction in mice (Fig. 1). In addition to the significant improvements in ethical qualities

Received: June 7, 2011
Accepted: June 15, 2011
Published online in J-STAGE: July 16, 2011

Correspondence: Y Tremblay (e-mail: yves.tremblay@crchul.ulaval.ca)

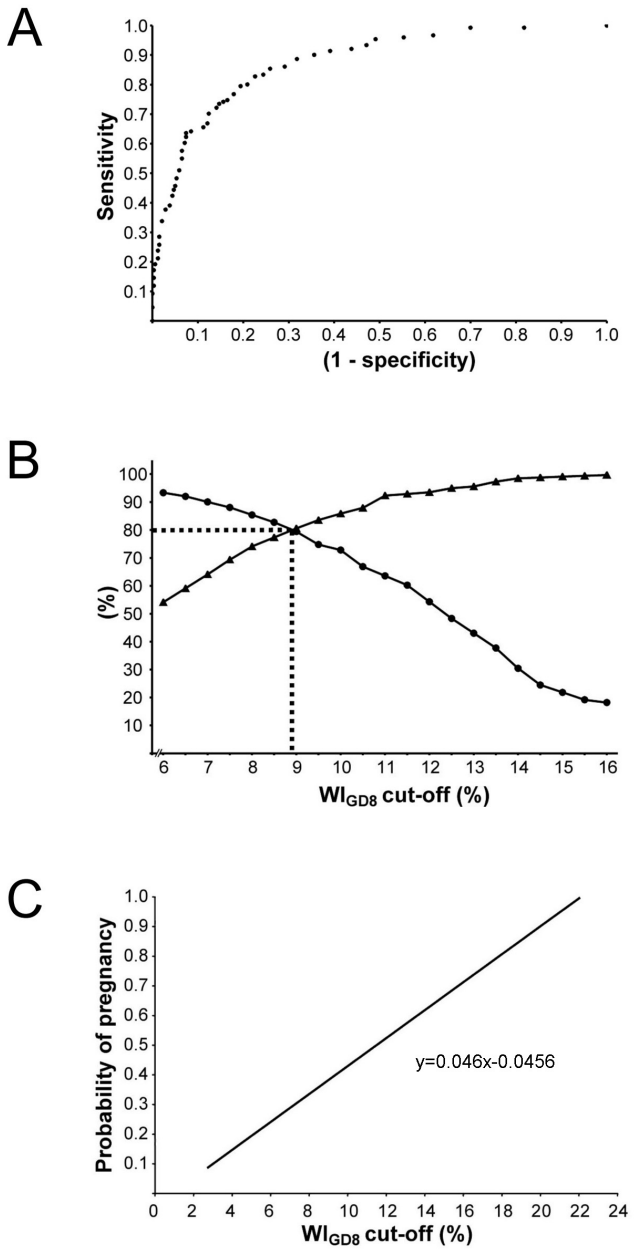

Fig. 1. Statistical characterization of the pregnancy prediction model based on the percentage of weight increase in the mated mice at GD8 (%WI$_{GD8}$). A: Receiver Operating Characteristic (ROC) curve showing the trade off between sensitivity (the proportion of actual pregnant mice that are correctly identified as such) and specificity (the proportion of nonpregnant mice that are correctly identified) and confirming the accuracy of the model with a close to 1 area under the curve (c=0.881). B: Sensitivity (circles) and specificity (triangles) associated with the %WI$_{GD8}$. The dotted line represents the %WI$_{GD8}$ threshold (8.74%) for which the better compromise between sensitivity and specificity is observed. C: Linear segment of the logistic regression plot of probability of pregnancy as a function of the observed %WI$_{GD8}$.

achieved by promoting refinement in the animal manipulations, this procedure also contributes to the reduction in the number of sacrificed animals. Indeed, the use of %WI$_{GD8}$ as a predictive model of pregnancy allows a diminution in the number of mice to be sacrifice by more than 50% (Table 1). More precisely, a %WI$_{GD8}$ cut-off ranging from 6 to 16% reduces the number of mice used by 55 to 69% in order to obtain the desired number of litters. Indeed, the number of mice spared by this procedure will decrease with an increasing pregnancy rate. For instance, if estimated on the basis of a pregnancy rate of 60% (which would be very high in a 1-h window mating protocol such as the one we used), the number of mice used to obtain the desired number of litters would be reduced by 12 to 40% for a %WI$_{GD8}$ cut-off ranging from 6 to 16% (166 females with vaginal plug should be sacrificed to obtain 100 litters). Therefore, althought the number of females spared decreases with an increasing pregnancy rate, this number remains significant and ethically important.

Table 1 summarizes values of both sensitivity (proportion of actual pregnant mice correctly predicted) and specificity (proportion of nonpregnant mice correctly predicted) associated with 0.5% increments of the %WI$_{GD8}$ as well as the proportions of false positives (nonpregnant mice incorrectly predicted as pregnant) and false negatives (pregnant mice incorrectly predicted as nonpregnant) for a %WI$_{GD8}$ cut-off ranging from 6 to 16.0%. This simple procedure represents a versatile tool adaptable to a wide range of protocols depending on the required sensitivity and specificity. For instance, some experiments necessitate that false negatives are minimized and false positives are limited at the same time, such as protocols needing a large sample size and/or in which multiple injections or chronic treatment have to be administered from GD8-GD9 to fetuses or pregnant mice. In such cases, the %WI$_{GD8}$ threshold presenting the best compromise between sensitivity and specificity should be considered. In our predictive model of mice pregnancy, this threshold is 8.74%, and it presents a sensitivity of 80% and a specificity of 79% (Fig. 1B). More specifically, the 95% confidence interval for this threshold is [8.09–9.39%]. If the treatment to administer in similar experiments is particularly time consuming or if the injected substance is scarce or expensive, false positives should be avoided, and a %WI$_{GD8}$ threshold providing higher specificity should be preferred. Other experiments may require detection of the greatest number of pregnancies. This would be the case for a study of null-mutant mice resulting in embryonic lethality around GD9-GD10 (e.g. mice deficient for *connexin-45* [13], *suppressor of fused* (*Sufu*) [14], *N-myc* [15], etc.). In such transgenic mice, the experimenter interested in acute treatment administration, rescue treatment or phenotypic characterization prior to embryonic death has to use a %WI$_{GD8}$ threshold associated with high sensitivity to detect pregnant females as much as possible. Alternatively, this experimenter could use the logistic regression plot of the probability of pregnancy as a function of %WI$_{GD8}$ (Fig. 1C) defined by a linear equation (y=0.0476x–0.0456) indicating that a mouse having a %WI$_{GD8}$ higher than 21.8% presents a 99% probability of pregnancy. By adjusting the pregnancy detection protocol with the specific experimental requirements, the versatility of the method gives the ethical possibility to limit the number of animals to be sacrificed at the required

Table 1. Reduction in the number of mice to be sacrificed by the pregnancy detection procedure associated with a 0.5% increment in the percentage of weight increase of the mated mice at GD8 (%WI_{GD8})

%WI_{GD8} cut-off	False positive rate (α)[a]	False negative rate (β)[b]	Specificity (1-α)	Sensitivity (1-β)	Number of females with vaginal plug to be sacrificed to obtain 100 litters		Females spared by using %WI_{GD8} to obtain 100 litters	
(%)	(%)	(%)	(%)	(%)	With pregnancy detection[c]	Without pregnancy detection[d]	Number	(%)
6.0	46	7	54	93	146		179	55
6.5	41	8	59	92	141		184	57
7.0	36	10	64	90	136		189	58
7.5	31	12	69	88	131		194	60
8.0	26	15	74	85	126		199	61
8.5	23	17	77	83	123		202	62
9.0	19	21	81	79	119		206	63
9.5	16	25	84	75	116		209	64
10.0	14	27	86	73	114		211	65
10.5	12	33	88	67	112		213	66
11.0	8	36	92	64	108	325	217	67
11.5	7	40	93	60	107		218	67
12.0	6	46	94	54	106		219	67
12.5	5	52	95	48	105		220	68
13.0	4	57	96	43	104		221	68
13.5	3	62	97	38	103		222	68
14.0	1	70	99	30	101		224	69
14.5	1	75	99	25	101		224	69
15.0	1	78	99	22	101		224	69
15.5	1	81	99	19	101		224	69
16.0	0	81	100	18	100		225	69

[a] Nonpregnant mice incorrectly identified as pregnant: α=False positive / (False positive + True negative). [b] Pregnant mice incorrectly identified as non-pregnant: β=False negative / (True positive + False negative). [c] Number of mice to sacrifice=100 + α. [d] Based on the observed rate of pregnancy of 30.8% among the mice showing a vaginal plug (151 / 491).

level. However, there are some cases in mouse developmental biology research for which this procedure could not be applicable. For instance, protocols involving embryo transfer experiments in which the operation damages the recipient mice by bleeding and anesthesia result in a weight body decrease during the first days after the operation. Although these implanted females usually recover their weight and further gain weight until they give birth, they present a specific pattern of weight increase during pregnancy for which the statistical model corresponding to our procedure may not work as described.

In developmental studies, accurate timing of pregnancy is essential to obtain a precise estimate of the gestational age and to reduce experimental variability [16, 17]. Overnight mating periods generally used lead to a sizeable lack of precision in gestational age (±7–8 h). Such imprecision is sufficient to introduce a major interlitter variability that could potentially hide subtle biological differences between experimental groups and result in an increase in the required sample size. Embryogenesis and morphogenesis being ongoing processes, organs and tissues are developmentally evolving throughout gestation, and the pace of this maturation is particularly fast within species showing a short pregnancy duration such as mice. For this reason, procedures using a mating window of 4 h [18] or 2 h [16] have been previously reported to generate accurately timed pregnancy in mice. We attempted to improve the reduction in the incertitude of gestational age at a level of ± 30 min by using mating windows of only 1 h with the hope of contributing to minimization of the experimental variability and then potentially to reduction in the number of animals needed to obtain statistically-relevant data. The apparently low pregnancy rate, i.e. the percentage of pregnant mice (151) over the total number of mice showing a vaginal plug (491), we observed (30.8%) could probably be explained by the 1-h mating window we used, since it has been demonstrated in mice that matings with a longer duration (i.e. more intromissions) produce more pregnancies than matings with a shorter duration (i.e. less intromissions) [22].

The mean litter size we observed (4.3 ± 0.3 fetuses) is lower than obtained with classic overnight mating protocols. Indeed, overnight mating previously done with BALB/c mice in our laboratory showed an average of 7.0 ± 0.2 pups per litter ([19] and unpublished data), which is similar to the results reported by others [20, 21]. Since the number of intromissions correlates positively with litter size in mice [22], the limited number of intromissions allowed during the short mating period we used likely explains the lower litter size we observed. Importantly, we did not observed significant correlation between %WI_{GD8} and litter size among pregnant mice (P=0.36 and R^2=0.13). In addition, no statistical difference was detected between the mean litter size of false negatives and well-predicted pregnant mice (P=0.14), suggesting that false negatives

are observable at the same rate of occurrence regardless of the litter size. These results are in agreement with a previous report demonstrating that the total placental and fetal weight become significant only from GD12 in Bom:NMRI mice, which show a mean litter size of 11.7 [2]. Taken together, these observations suggest that the maternal weight gain is solely responsible for the measured $\%WI_{GD8}$, without significant contribution from fetuses or extraembryonic annexes. Therefore, the relevance of the described protocol seems to be independent of the litter size and, to a greater extent, to the time window used for mating.

In this study, we observed a mean $\%WI_{GD8}$ of 13.2 ± 0.5% in pregnant BALB/c mice, which is consistent with the previously reported mean $\%WI_{GD8}$ of approximately 15% in the Bom:NMRI mice strain [2]. Moreover, another study performed using the Rockland-Swiss strain reported a weight difference of 13.3% between pregnant and nonpregnant mice at GD8 [23]. This apparent homogeneity in the maternal weight gain at GD8 among different mice strains crucially increases the potential of the procedure described here by suggesting its reliability for other mice strains. From an ethical point of view, this study represents at least an interesting proof-of-concept in the worthiness of the described procedure in the attempt to optimize the number of mice used in developmental studies. To a greater extent, such statistical models could be characterized at different gestational ages as well as in other rodent species such as the rat.

The present study provides a versatile and convenient tool allowing early pregnancy diagnosis in mice as early as GD8 on the basis of the percentage of maternal weight gain. This protocol has the potential to facilitate and optimize a large spectrum of functional *in vivo* experiments such as the evaluation of the effect of a midgestation treatment on fetal and/or maternal parameter(s) assessed immediately or later in gestation. Moreover, this procedure allows ethical improvement for developmental studies by offering the possibility to significantly reduce the number of mice to be sacrificed by using a method that is totally unstressful for animals.

Methods

BALB/c mice were purchased from Charles River Laboratories (St-Constant, QC, Canada). Upon arrival, females were 63–70 days of age and weighed 18–21 g, and males were 63–70 days of age and weighed 21–24 g. The mice were housed in a same-sex manner (5 females or 1 male per cage) in a room maintained at 22 ± 3 C with a relative humidity of 50 ± 20% and a 12-h light cycle. Commercial diet (Global 18% Protein Rodent Diet, Teklad, Montréal, QC, Canada) and water treated by reverse osmosis were provided *ad libitum*. Animals were acclimatized to these conditions for at least 7 days prior to mating.

On the morning of GD0, nulliparous females were screened for estrus on the basis of vulvar appearance [24]. Females showing vulvar characteristics of estrus (gapped vagina with swollen dorsal lip showing lightly pink and moist tissues) were weighed and placed individually in a male's cage for 1 h [16, 18]. After this breeding period, females showing a vaginal plug were secluded individually for the pregnancy duration (n=491). On the morning of GD8, females were weighed and the percentage of weight increase at GD8 ($\%WI_{GD8}$) was calculated (Equation 1) and used as the criteria for pregnancy prediction. The females that were actually pregnant at GD18 (n=151) underwent cesarean section and were characterized by their litter size.

Equation 1 $$\%WI_{GD8}=(W_{GD8}-W_{GD0}) / W_{GD0} \times 100$$

Descriptive statistics concerning rate of pregnancy and litter size were reported as the mean ± SEM. Statistical analyses regarding gestation prediction were performed with SAS® software using a logistic regression fitted on a binary logit model (Equation 2) in which $\%WI_{GD8}$ values of the dataset were ordered as *1* for pregnant females (n=151) and *0* for nonpregnant females (n=340).

Equation 2 $$\log it(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + (\beta_1 \bullet \%WI_{GD8})$$

In Equation 2, p is the probability of a mouse being pregnant considering a given value of $\%WI_{GD8}$, β_0 is the intercept and β_1 is the slope of the function. Fitting of the binary logit model and test accuracy were confirmed by the high value of the area under the curve (c=0.881) of the Receiver Operating Characteristic (ROC) curve (Fig. 1A) and highly significant p-value of the Hosmer and Lemeshow test (P=0.9541). Both of these statistics represent goodness-of-fit tests, and the closer to 1 are their values, the better the statistical model fits.

Acknowledgments

This work was supported by grant # MOP74531 from the Canadian Institutes of Health Research (CIHR) to YT and by Wyeth Pharmaceuticals & CIHR/Rx&D Research Fellowship Program awarded to TS. We gratefully thank E Boucher and Dr M Simard for helpful discussions. We also wish to thank Dr R Charbonneau (D.M.V., director of veterinary services at the CHUQ Research Center) for his critical reading of the manuscript.

References

1. **Russell WMS, Burch RL**. The Principles of Humane Experimental Technique. London: Methuen & Co. Ltd.; 1959: 238.
2. **Hau J, Skovgaard Jensen HJ**. Diagnosis and monitoring of pregnancy in mice: correlations between maternal weight, fetal and placental mass and the maternal serum levels of progesterone, pregnancy-associated murine protein-2 and alpha-fetoprotein. *Lab Anim* 1987; 21: 306–310.
3. **Cerruti RA, Lyons WR**. Mammogenic activities of the midgestational mouse placenta. *Endocrinology* 1960; 67: 884–887.
4. **Russo M, Meomartino L, Greco A, Catone G, Cocchia N, Tortora G, Brunetti A**. Pregnancy detection in mice using ultrasound. *Vet Rec* 2007; 160: 446–447.
5. **Mu J, Slevin JC, Qu D, McCormick S, Adamson SL**. *In vivo* quantification of embryonic and placental growth during gestation in mice using micro-ultrasound. *Reprod Biol Endocrinol* 2008; 6: 34.
6. **Kumai M, Nishii K, Nakamura K, Takeda N, Suzuki M, Shibata Y**. Loss of connexin45 causes a cushion defect in early cardiogenesis. *Development* 2000; 127: 3501–3512.
7. **Perl AK, Whitsett JA**. Molecular mechanisms controlling lung morphogenesis. *Clin Genet* 1999; 56: 14–27.
8. **Auerbach R, Huang H, Lu L**. Hematopoietic stem cells in the mouse embryonic yolk sac. *Stem Cells* 1996; 14: 269–280.
9. **Brook FA, Estibeiro JP, Copp AJ**. Female predisposition to cranial neural tube defects

is not because of a difference between the sexes in the rate of embryonic growth or development during neurulation. *J Med Genet* 1994; 31: 383–387.

10. **Sawai S, Shimono A, Wakamatsu Y, Palmes C, Hanaoka K, Kondoh H**. Defects of embryonic organogenesis resulting from targeted disruption of the N-myc gene in the mouse. *Development* 1993; 117: 1445–1455.
11. **Lutolf S, Radtke F, Aguet M, Suter U, Taylor V**. Notch1 is required for neuronal and glial differentiation in the cerebellum. *Development* 2002; 129: 373–385.
12. **Lee CS, De Leon DD, Kaestner KH, Stoffers DA**. Regeneration of pancreatic islets after partial pancreatectomy in mice does not involve the reactivation of neurogenin-3. *Diabetes* 2006; 55: 269–272.
13. **Kruger O, Plum A, Kim JS, Winterhager E, Maxeiner S, Hallas G, Kirchhoff S, Traub O, Lamers WH, Willecke K**. Defective vascular development in connexin 45-deficient mice. *Development* 2000; 127: 4179–4193.
14. **Svard J, Heby-Henricson K, Persson-Lek M, Rozell B, Lauth M, Bergstrom A, Ericson J, Toftgard R, Teglund S**. Genetic elimination of suppressor of fused reveals an essential repressor function in the mammalian Hedgehog signaling pathway. *Dev Cell* 2006; 10: 187–197.
15. **Charron J, Malynn BA, Fisher P, Stewart V, Jeannotte L, Goff SP, Robertson EJ, Alt FW**. Embryonic lethality in mice homozygous for a targeted disruption of the N-myc gene. *Genes Dev* 1992; 6: 2248–2257.
16. **Endo A, Watanabe T**. Interlitter variability in fetal body weight in mouse offspring from continuous, overnight, and short-period matings. *Teratology* 1988; 37: 63–67.
17. **Watanabe T, Endo A**. Digit development and embryonic weight in mice: analysis of sex-related time difference and mating period-related interlitter variability. *Teratology* 1988; 38: 157–163.
18. **Scharmann W, Wolff D**. Production of timed pregnant mice by utilization of the Whitten effect and a simple cage system. *Lab Anim Sci* 1980; 30: 206–208.
19. **Provost PR, Simard M, Tremblay Y**. A link between lung androgen metabolism and the emergence of mature epithelial type II cells. *Am J Respir Crit Care Med* 2004; 170: 296–305.
20. **Wahlsten D, Bulman-Fleming B**. The magnitudes of litter size and sex effects on brain growth of BALB/c mice. *Growth* 1987; 51: 240–248.
21. **Ruiz-Luna AC, Salazar S, Aspajo NJ, Rubio J, Gasco M, Gonzales GF**. Lepidium meyenii (Maca) increases litter size in normal adult female mice. *Reprod Biol Endocrinol* 2005; 3: 16.
22. **deCatanzaro D**. Duration of mating relates to fertility in mice. *Physiol Behav* 1991; 50: 393–395.
23. **Biggerstaff S, Mann M**. Consummatory behaviors and weight regulation in pregnant, lactating, and pregnant-lactating mice. *Physiol Behav* 1992; 52: 485–491.
24. **Champlin AK, Dorr DL, Gates AH**. Determining the stage of the estrous cycle in the mouse by the appearance of the vagina. *Biol Reprod* 1973; 8: 491–494.