Butylparaben multigenerational reproductive assessment by continuous breeding in Hsd:Sprague Dawley SD rats following dietary exposure

Troy D. Hubbard[a], **Amy Brix**[b], **Chad R. Blystone**[a], **Barry S. McIntyre**[a], **Keith Shockley**[c], **Helen Cunny**[a], **Suramya Waidyanatha**[a], **Katie J. Turner**[d,1], **Sandra McBride**[e], **Georgia K. Roberts**[a,*]

[a]Division of the National Toxicology Program, National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, NC, USA

[b]Experimental Pathology Laboratories, Inc., Research Triangle Park, NC, USA

[c]Division of Intramural Research, National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, NC, USA

[d]RTI International, Research Triangle Park, NC, USA

[e]Social & Scientific Systems, Inc., Research Triangle Park, NC, USA

Abstract

Butylparaben (BP) is an antimicrobial agent utilized for decades as a preservative in numerous consumer products. The safety of parabens has recently come under scrutiny based on reports of estrogenic activity and suggested adverse effects upon the reproductive system. Due to the limited availability of studies that address the potential for BP exposure to induce reproductive toxicity, and clear evidence of human exposure, the National Toxicology Program conducted a multigenerational continuous breeding study to evaluate the impact of dietary BP-exposure at 0, 5000, 15,000, or 40,000 ppm on reproductive and developmental parameters in Hsd:Sprague Dawley SD rats. BP-exposure was not associated with adverse alterations of fertility, fecundity, pubertal attainment, or reproductive parameters in F0, F1, or F2 generations. Exposure-dependent increases in liver weights, and incidences of non-neoplastic liver lesions suggest the liver is a target organ of BP toxicity. No findings were observed that would support the purported mechanism of BP-induced endocrine disruption in perinatally-exposed rodents.

*Corresponding author at: Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA. Georgia.Roberts@nih.gov (G.K. Roberts).
[1]Current location Janssen R&D LLC, Spring House, Pennsylvania, USA.

Keywords

Butylparaben; Reproduction; Development; Toxicity

1. Introduction

Butylparaben (*n*-butyl-p-hydroxybenzoate; BP) is a member of the paraben class, which is composed of differentially substituted esters of varying alkyl chain lengths (methyl-, ethyl-, propyl-, butyl-, heptyl-, benzyl-paraben) at the C-4 position of *p*-hydroxybenzoic acid. Parabens are chemically stable, and possess antimicrobial and antifungal properties, likely attributed to their capacity to inhibit bacterial/fungal membrane transport, impair mitochondrial function, and disrupt DNA/RNA/protein synthesis [1,2].

In general, parabens are classified as non-irritating, non-sensitizing, and non-toxic with a high benefit to risk ratio relative to other commercial preservatives. The FDA has approved the use of butylparaben as a direct food additive in food for human consumption at minimum levels required to produce its intended antibacterial effect [3]. In a survey of 267 food samples, parabens (mean total parabens, 10 ng/g) were detected in greater than 90 % of analyzed samples [4]. In a screening of 215 cosmetic products by European and Danish regulatory agencies found that 77 % of these products contained detectable concentrations of parabens ranging from 0.01 % to 0.87 % (w/w), with BP levels reaching a maximum concentration of 0.07 % (w/w) [5]. Parabens have been used as excipients in drug formulations in a wide range of products and delivery routes since 1924 [6]. In the United States, BP is currently approved for use to a maximum concentration of 0.4 % (w/w) in topical creams, 0.08mg/mL in oral solutions, and 28.4mg/mL in oral syrups [7].

Detectable levels of four parabens (methyl, ethyl, propyl, and butyl) have been quantitated in human urine samples from the National Health and Nutrition Examination Survey (NHANES). The NHANES data reported the 95th percentile of urinary BP concentrations in 2012 and 2014 of 10.0 and 4.39 μg/g creatinine respectively, suggesting that exposure to BP is declining in the U.S. population [8,9]. The 2014 NHANES data also reported 24-fold higher levels of urinary BP (95th percentile) among female respondents (10.2 μg/g creatinine) when compared to males (0.417 μg/g creatinine) [9]. Gender-bias in BP exposure may be associated with disproportionate use of cosmetics between men and women, and is further supported by higher concentrations of BP in cosmetics relative to food products or pharmaceuticals [10].

Numerous *in vitro* and *in vivo* studies have reported parabens to possess estrogenic activity [11,12]. Additional studies found that multiple parabens were competitive estrogen receptor ligands and induced transactivation of ER-dependent gene expression in MCF-7 cells [13–16]. However, estrogenic potency of parabens was estimated to be 1,000-1,000,000 fold less than that of 17-β-estradiol [17]. Structure activity relationship assessments found parabens containing longer or increasingly branched side chains possessed increased relative estrogenic potential [16]; which was observed in recombinant yeast assays, in which methyl-, ethyl-, propyl-, butyl-, and benzylparaben, displayed relative responses of 1/3,000,000, 1/200,000, 1/30,000, 1/8000, and 1/4000 the potency of estrogen, respectively

[18]. I*n vivo* rodent uterotrophic assays indicate BP-associated estrogenic effects are many orders of magnitude less potent than estrogen, although this disparity was lessened following subcutaneous injection relative to oral administration [11].

Speculation regarding paraben safety was prompted by numerous reports of severe adverse effects on the male reproductive system in rodents [19–21]. In one study, male Wistar rats fed diets containing up to 10,000 ppm BP (1 % w/w) for 8 weeks presented dose-related reductions in epididymal weights, testis and cauda epididymal sperm concentrations and serum testosterone levels [21]. A similar study in Crj:CD-1 mice found BP exposure decreased serum testosterone levels and daily sperm production at the highest dose tested (1 % w/w) [19]. Male Sprague-Dawley offspring from dams administered 100 or 200 mg BP/kg/day daily subcutaneous injections from gestational day (GD) 6 until weaning on postnatal day (PND) 20 displayed reductions in cauda epididymal sperm concentration and motility in males on PND 90 [22]. In contrast to these findings, a study in male Wister (Crl: (WI) BR) rats exposed to BP in feed (up to 1 % w/w) for 8 weeks, beginning at postnatal day 22, found no impact on similar endpoints and reported no significant decreases in male reproductive organ weights, sperm parameters, or serum testosterone [50]. Epidemiological studies have been unable to find significant correlations between paraben exposure and reduced semen quality in males among sub-fertile couples [23,24].

Currently, the available data on the endocrine disrupting potential and corresponding reproductive toxicity of parabens is conflicting, demonstrating the need for an appropriately powered and comprehensive reproductive toxicity study that included an in utero exposure would inform assessment of human risk. Due to potential endocrine disruption, evidence of human exposure, and reported effects upon male reproductive development, BP was selected for further toxicological characterization by the National Toxicology Program (NTP). The NTP conducted a reproductive assessment by continuous breeding (RACB) study to evaluate the potential toxicity of continuous BP exposure on reproductive fitness and male/female reproductive development over multiple generations. The successive breeding methodology provides enhanced statistical power to the evaluate overall fecundity compared to standard multigeneration studies and characterizes potential perturbations of pubertal development.

2. Materials and methods

2.1. Reproductive assessment by continuous breeding study design overview

The RACB study design utilizes a multiple breeding approach, described by Morrissey et al. (1989) and Chapin et al. (1997) [25,26]. In summary, exposure starts with the F0 generation and continues through the F2 generation. Multiple successive pairings (3 per generation) in both the F0 and F1 generations are conducted to evaluate test article induced reproductive toxicity (Fig. 1). In this design, the successive number of matings and evaluation of offspring provides increased statistical power to identify test article-related toxicities compared to standard multigeneration studies [25]. Additionally, maturation of F1c offspring to adulthood allows for the evaluation of the potential attenuation (or enhancement of) test article-related effects on fertility and fecundity.

2.2. Chemical procurement and characterization

Butylparaben (CAS# 94-26-8, Lot No. IF100205) was procured from Ivy Fine Chemicals Corporation (Cherry Hill, NJ). Chemical identity was confirmed by infrared spectroscopy, nuclear magnetic resonance spectroscopy, and gas chromatography-mass spectrometry. Chemical purity determined by high performance liquid chromatography with ultraviolet detection (HPLC/UV) and gas chromatography with flame ionization detection was 99.7 %.

2.3. Animals and husbandry

Nine-week old male and female Sprague Dawley (Hsd:Sprague Dawley® SD®) rats (n = 117/sex) were obtained from Harlan, Inc. (Now Enivgo; Dublin, VA). An additional 29 females were procured during the study for a cross-over mating assessment. Animals were randomly assigned to exposure groups by stratified randomization of body weights using Provantis™ and quarantined for 10 days upon arrival. This study was conducted in compliance with FDA Good Laboratory Practice Regulations established by the Food and Drug Administration. Animals were handled, cared for, and used in compliance with the study protocol approved by the RTI International IACUC, and National Research Council *Guide for the Care and Use of Laboratory Animals* (2011) [27]. Rats were housed in solid-bottom, polycarbonate cages (8"xl9"xl0.5") suspended on automatic watering racks with filter sheets. All cages were bedded with Sani-chips® hardwood bedding (P.J. Murphy Forest Products, Montville, NJ) for the duration the study. Nesting sheets or other enrichment materials, which might contain endocrine-disrupting chemicals that could affect study outcomes, were not placed in the cages.

All F0 male and female rats were individually housed except during periods of cohabitation (mating). Following weaning, F1c offspring were housed up to 3/sex/cage until approximately PND 90. Following PND 90, F1c non-parental (F1cNP) interim animals were removed from study for interim evaluations. F1c parental (F1cP) male and female rats were housed individually except during periods of cohabitation. F0 and F1c dams were cohoused with offspring until weaning. Irradiated NIH-07 Certified Mouse/Rat Diet (Zeigler Bros., Inc., Gardners, PA) was available *ad libitum* throughout the study. Feed was delivered to animals via glass feed jars with stainless steel lids. Tap water (City of Durham, NC) was available *ad libitum* via an automatic water delivery system (Edstrom Industries Inc., Waterford, WI) except during sanitization of the waterlines. The animal rooms were maintained on a 12 -h light/dark cycle per day (~0600/1800 on/off), excluding daylight savings time adjustments. Temperature and relative humidity (%) were continuously monitored and specified set points and ranges (i.e., 72 ± 3 °F and 50 % RH ± 15 %) were observed.

2.4. Exposure

Throughout the study, unless otherwise noted, animals were exposed to BP via supplementation in NIH-07 powdered feed at levels of 0, 5000, 15,000, or 40,000 ppm (Zeigler Brothers, Inc., Gardners, PA). Feed formulations were analyzed using a validated high-performance liquid chromatography method with ultraviolet detection at 256 nm (r > 0.99; relative standard deviation (RSD), 10 %; relative error (RE), ± 10 %); all

formulations were within 10 % of target concentration and homogeneous (RSD, 5 %). Prior to studies, stability of BP in the diet up to 43 d was established when stored at ambient temperature, protected from light. All feed formulations were stored under these conditions during study conduct.

F0 adults were exposed to BP during a 2-week pre-breed exposure period, during cohabitation, and gestation and lactation for the F1a, F1b, and F1c generations, until necropsy. The F1c generation was exposed throughout life. The F2c generation was exposed to BP via the mother during gestation and lactation until study completion on PND 21. Body weights and feed consumption were measured throughout the study (pre-cohabitation, cohabitation, gestation, lactation) and used to calculate chemical consumption (mg/kg/day).

2.5. Exposure concentration selection rationale

A preliminary study was conducted to select doses for the definitive reproductive and developmental toxicity study. Males (aged 10 weeks) and females (aged 8 weeks) (n = 8/sex/dose at 10,000/20,000 ppm, n = 11/sex/dose at 0/5000/40,000 ppm), were administered 0, 5000, 10,000, 20,000, or 40,000 ppm BP in the diet for 2 weeks prior to cohabitation and during cohabitation, until evidence of mating or up to 15 days. Upon evidence of mating, females were continually exposed throughout gestation and lactation, until study termination on PND 4. There were no observed effects of BP exposure on viability, fertility or litter size during the course of the study. These data supported the selection of 40,000 ppm as the highest exposure concentration used in the definitive RACB study. All of the preliminary study data are publicly available via the National Toxicology Program's Chemical Effects in Biological Systems (CEBS) database (https://doi.org/10.22427/NTP-DATA-NTP-DATA-RACB-BP).

2.6. RACB study

Sprague Dawley rats, F0 (aged 11 weeks, n = 22/sex/group) and F1c parental (F1cP) animals (aged 12–13 weeks, n = 26–40) were assigned a non-sibling mating partner within the same dose group (Fig. 1). The same pair of rats was paired each time to allow reproductive performance to be evaluated over the three pairings (15 days per cohabitation), to allow assessment of fecundity and fertility parameters. Females were examined daily by vaginal smears during cohabitation for the presence of a copulation plug or sperm; the presence of either was considered positive evidence of mating. The day of confirmed mating was designated as Gestation Day (GD) 0, at which point cohabitation of males and mated females was concluded.

Both F0 and F1c females were allowed to litter generating a F1a, F1b, and F1c or F2a, F2b, and F2c offspring, respectively. The day of littering was designated as PND 0. Litter parameters, such as litter size, numbers of live and dead pups, sex ratio, and pup weights were recorded for all litters from PND 0 to PND 4. Anogenital distance (AGD) was measured on PND 1 in all litters using a stereomicroscope with a reticule. On PND 4, the designated a and b offspring (F1a, F1b, F2a, F2b) of each generation were euthanized, whereas the F1c and F2c offspring were standardized to a litter size of 12 pups (6/sex/litter where possible). Following euthanasia of F1a, F1b, F2a, and F2b litters, mating pairs were

re-cohabitated. PND 4 offspring were euthanized by intraperitoneal injection of euthanasia solution or decapitation.

Following litter standardization of the F1c and F2c litters, pup body weights were recorded during the lactational period (PND 4 to PND 21). The F1c and F2c male offspring were evaluated for areolae/nipples beginning on PND 13 and testes descent beginning on PND 14. All F1 and F2 offspring that exhibited severe toxicity and were removed from study or died received a gross necropsy. Initially, F1c offspring were weaned on PND 21. However, weaning was postponed to PND 28 for approximately half of the 40,000 ppm group due to increased incidence of mortality associated with inability to thrive following separation from their respective dam. At the time of weaning, some of the F1c offspring were randomly assigned to an interim assessment cohort (F1cNP) (up to 3/sex/litter). Remaining animals were paired three times as mentioned above.

Following weaning of the F1c generation and prior to the crossover mating, vaginal smears were collected from the F0 females for 16 consecutive days for evaluation of estrous cyclicity. Vaginal smears were also collected from F1c females for 16 consecutive days (PND 52–67) for evaluation of estrous cyclicity. Estrous stage was determined using previously established guidelines for evaluation of vaginal cytology in rodent species [28]. The acquisition of Vaginal Opening (VO) was evaluated in all F1c females beginning on PND 23, and acquisition of Balaneo-pre-preputial Separation (BPS) was evaluated in all F1c males beginning on PND 35.

2.7. Crossover mating

Due to a reduction in mean live litter size of F0 females at the highest exposure concentration, a crossover mating was conducted to determine if BP effects on reproduction were due to susceptibility in a single sex. The crossover mating entailed pairing 40,000 ppm F0 males (aged 37 weeks) with unexposed nulliparous females (aged 11 weeks), and 40,000 ppm F0 females with 0 ppm F0 males (both aged 37 weeks) (Fig. 1). Animals were cohabitated for seven days and females examined for evidence of mating each morning. All animals were provided *ad libitum* access to control NIH-07 diet during this period, and upon confirmation of mating, females were returned to their respective diets. Females were permitted to litter and their offspring were euthanized on PND 4. Live and dead pup numbers and associated sex were recorded daily from PND 0 through PND 4, and live pups were weighed on PND 1 and PND 4. Dams which did not deliver were euthanized 24 days after the last day of cohabitation and received an examination to determine pregnancy status. The uterus was examined for implantation sites, and stained with ammonium sulfide if none were visible. Following crossover mating, F0 males and females, and naïve females were euthanized.

2.8. Necropsy

The F0 males and females were necropsied at 43–14 weeks of age. F1cNP (interim) males and females were necropsied at 13–14 weeks of age. F1cP (terminal) males and females were euthanized at 30–31 weeks of age and 34–35 weeks of age, respectively. Adult rats and PND 21 offspring were euthanized by CO_2 asphyxiation. F0 and F1c (parental and

non-parental) terminal body weights were recorded and the respective weights of the adrenal glands, liver, kidney, spleen, thymus, thyroid, ovaries, seminal vesicles with coagulating glands, epididymides, dorsal prostate, ventral prostate, seminal vesicles, and testes collected. In addition to the aforementioned tissues, the pituitary gland, cervix, uterus, and vagina were retained from females for histopathological evaluations. Tissues were fixed in neutral buffered 10 % formalin except the right testis and epididymis and left and right ovaries, which were initially fixed in modified Davidson's fixative for approximately 24 h prior to transfer to formalin. Tissues were embedded in paraffin and stained with hematoxylin and eosin except for the right testis and epididymis, which were stained with periodic acid-Schiffs stain and counterstained with hematoxylin. Andrology measures including: left cauda epididymal spermatozoa measurements (number, density, motility) and left testicular spermatid counts were evaluated in F0 and F1c (parental and non-parental) males. F2 offspring were necropsied at 3 weeks of age, and gross lesions collected for microscopic evaluation.

Microscopic evaluations of the above tissues were completed by the study laboratory pathologist to a no effect level, and the study laboratory report, slides, paraffin blocks, residual wet tissues, and pathology data were sent to the NTP Archives for inventory, slide/ block match, wet tissue audit, and storage. The slides, individual animal data records, and pathology tables were evaluated by an independent quality assessment (QA) laboratory. The QA pathologist evaluated slides from all potential target organs (as identified by the study pathologist and existing literature) in the control, highest dose group with 90 % survival, and in lower dose groups to a no effect level; higher dose groups with less than 90 % survival were also evaluated. The tissues examined by the QA pathologist included the liver of F0 and F1 males and females, the testes and epididymides from F1 parental males, the ovaries from F1 parental females, and the mammary glands from F1 non-parental females. The QA report and the reviewed slides were submitted to the NTP pathologist, who reviewed and addressed any inconsistencies in the diagnoses made by the study laboratory and QA pathologist. The QA pathologist, who served as the coordinator of the Pathology Working Group (PWG), presented representative histopathology slides containing examples of lesions related to test agent administration, examples of disagreements in diagnoses between the laboratory and QA pathologist, or lesions of general interest to the PWG for review. The PWG consisted of the NTP pathologist and other pathologists experienced in rodent toxicologic pathology. This group examined the tissues without any knowledge of dose groups. When the PWG consensus differed from the opinion of the study laboratory pathologist, the diagnosis was changed. Final diagnoses for reviewed lesions represent a consensus between the study laboratory pathologist, QA pathologist, and the PWG. Details of these review procedures have been described, in part, by Maronpot and Boorman (1982) [29] and Boorman et al.(1985, 2002) [30,31].

2.9. Statistical analyses

All endpoints were tested for a trend across dose groups, followed by pairwise tests for each dose group against the negative control group, unless otherwise noted. Statistical significance was reported at the 0.05 level. In most cases, statistical methods differed for

F0 and F1 animals, since methods for the F1 animals needed to account for within litter correlation where present.

For F0 reproductive performance endpoints (CEBS, RACB R02), statistical analysis was performed by Cochran-Armitage (trend) and Fisher Exact (pairwise) two-sided tests [41]. For F1 reproductive performance endpoints, analysis was performed using the Rao-Scott Cochran-Armitage procedure for both trend and pairwise tests [43]. F0 litter size and survival endpoints (CEBS, RACB R03) were analyzed using Jonckheere's test for trend [33]. For pairwise comparison of dosed groups to controls, Shirley's [34] (as modified by Williams [35]) or Dunn's [36] method was used depending on detection of a significant trend. F1 litter size and survival endpoints were analyzed using a permutation test based on the Jonckheere trend test [33], and pairwise comparisons were made using a modified Wilcoxon test [39] with the Hommel procedure [40] to adjust for multiple comparisons.

To analyze attainment of developmental endpoints (testes descent (TD), vaginal opening (VO), Balaneo-pre-preputial Separation (BPS) in Table 5), trend and pairwise tests were based on mixed models for age at attainment with dose as a covariate and a random effect for litter followed by a Dunnett-Hsu adjustment for multiple comparisons. For VO and BPS, weaning weight was included as a covariate in mixed models. To calculate age at attainment adjusted for pup body weight, a linear model was fit to attainment age as a function of body weight. Then the estimated coefficient for body weight was used to adjust each pup's observed attainment age based on the difference between its body weight and the overall mean body weight. For epididymal sperm endpoints (Table 6), F0 animals were analyzed using the Jonckheere trend test followed by Shirley's or Dunn's method for pairwise comparisons. F1 animals with littermates were analyzed with a bootstrapped Jonckheere trend test where litters were permuted across dose groups, and animals with the same maternal dam were sampled with replacement. Modified Wilcoxon tests were used for pairwise comparisons, with the Hommel adjustment for multiple comparisons.

Organ and body weights (Tables 7; Supplemental Tables 4 and 5) in F0 animals were analyzed using a Jonckheere trend test and Williams or Dunnett (pairwise) tests. For organ and body weight endpoints in F1 animals, mixed models were fit with a random litter effect and a Dunnett-Hsu adjustment for both trend and pairwise analyses.

Lesion incidence for F0 animals (Table 8) was analyzed using the Poly-3 trend and pairwise statistics, as described below. The Poly-k test was used to assess neoplasm and nonneoplastic lesion prevalence [37–39]. This test is a survival-adjusted quantal-response procedure that modifies the Cochran-Armitage linear trend test to account for survival differences. Following Bailer and Portier [39], a value of k = 3 was used in the analysis of site-specific lesions. Variation introduced by the use of risk weights, which reflect differential mortality, was accommodated by adjusting the variance of the Poly-3 statistic as recommended by Bieler and Williams [40]. Poly-3 tests used the continuity correction described by Nam [41]. Lesion incidence for F1c parental animals was analyzed using the Cochran-Armitage test with a poly-3 adjustment for survival [39] and a Rao-Scott modification for litter effect. Lesion incidence for F1c non-parental animals were analyzed using Cochran-Armitage test for trend and the Fisher Exact test for pairwise testing, since no

survival corrections or littermate adjustments were needed [42]. All p-values calculated for histopathology data are one-sided.

Vaginal cytology data for F0 animals (Supplemental Table 3) were analyzed using the nonparametric multiple comparison methods of Shirley [34] (as modified by Williams [35]) and Dunn [36] (cycle length, number of cycles). Jonckheere's test [33] was used to assess the significance of the dose-related trends. F1 animals were analyzed using the modified Wilcoxon test with Hommel adjustment.

A cumulative logit model was fit to the number of litters as a function of generation, with a random effect for F0 dam, to compare the number of litters per pair between F0 and F1 dams within each dose (Table 3; CEBS, RACB R22). Average live litter size per pair between F0 and F1 dams were analyzed using Datta-Satten modified Wilcoxon test with a Hommel p-value adjustment [32] (Table 4; CEBS, RACB R22).

For continuous endpoints, extreme values were identified by the outlier test of Dixon and Massey [43] for small samples (n   20), and Tukey's outer fences method [30] based on three interquartile range intervals, for large samples (n > 20). To identify outliers for continuous endpoints with litter effects, all observations across dose groups were fit to a linear mixed effects model with random litter effect, and the residuals were tested by dose group for outliers using Tukey's outer fences method [30]. All flagged outliers were examined by NTP personnel, and implausible values were eliminated from the final analyses. Discrete count endpoints were manually reviewed for unusual values.

3. Results

Due to the extensive amount of data generated during the course of the study, key findings are provided herein. All study data are publicly available via the National Toxicology Program's database: Chemical Effects in Biological Systems (CEBS) (https://doi.org/10.22427/NTP-DATA-NTP-DATA-RACB-BP).

3.1. Butylparaben intake levels

BP consumption (mg/kg/day) during multiple life stages is summarized in Table 1 (CEBS, RACB 108). An average daily dose (mg/kg/day) was calculated for the three gestation (GD 0–21) and lactation periods (PND 1–4) that produced the F1 and F2 litters. Feed consumption (g/kg body weight/animal/day) in the 40,000 ppm group was increased ~10–37 % at various time points in F0 animals and ~14 – 61 % in the F1c animals (CEBS, RACB I06). Higher consumption in these groups did not coincide with increased growth rates or increased absolute body weights relative to controls. In some cases, increased food consumption corresponded to non-consumptive spillage of powdered feed and may reflect palatability issues due to the high level of BP in the diet (4 %). Additionally, increased feed consumption relative to body weight may be related to the magnitude of bodyweight reduction observed in the 40,000 ppm exposure groups.

3.2. Viability and clinical observations

In the F0 generation there was no effect of BP exposure on survival (Table 2, CEBS, RACB I01). High mortality was observed in F1c pups in the 40,000 ppm group weaned at PND 21 (Table 2). Approximately one-half of the 40,000 ppm litters were weaned on PND 21, with 100 % mortality/moribund removal of those animals due to small size leading to failure to thrive. Clinical observations of hypoactivity, dehydration, and cold to the touch primarily occurred in 40,000 ppm pups that were unable to thrive following weaning on PND 21 (CEBS, RACB I05 P). Weaning was extended to PND 28 for the remaining 40,000 ppm litters, relative to PND 21 for all other groups, resulting in 100 % survival to scheduled removal. As a result, the 40,000 ppm F1c parental group had 35 % fewer animals than controls (n = 26 per sex, relative to n = 40 per sex) so a limited number of animals were assigned to the interim evaluation cohort (F1cNP). There were subsequent BP exposure-related adverse clinical observations requiring F1c adults or F2 offspring removal.

3.3. Body weights

Dose-related decreases in mean body weights relative to controls were observed across generations in groups administered 15,000 ppm or greater (CEBS, RACB I04/I04 G). F0 male body weights were up to ~5 % and ~11 % lower than controls in the 15,000 and 40,000 ppm groups, respectively (Fig. 2A). F0 40,000 ppm female body weights were up to 7 % lower than controls prior to mating (Fig. 2A). Body weights of F0 40,000 ppm dams were up to 11 %, 12 %, and 14 % lower than controls during gestational intervals A, B, and C, respectively (Fig. 2B). This effect corresponded to lower gestational body weight gains relative to controls. Smaller litter sizes compared to the control group may have contributed to lower gestation weight. During lactation (Lactation day (LD) 1-2 1, Litter C), F0 female body weights in the 40,000 ppm group were up to 16 % lower than controls (Fig. 2C). Perinatal exposure to BP resulted in a dose-related decrease in F1c pup weights during their respective preweaning interval (PND 1–21). The effect on pup weight in the 15,000 ppm group was statistically significant (−11 %, relative to controls) just prior to weaning (PND 19). However, significantly lower pup weight in the 40,000 ppm group (−10 %, relative to controls) was observed earlier, on PND 4 (CEBS, RACB R19/R19 G). At PND 21, F1c male pup body weights were 11 % and 36 % lower than controls, and F1c female pups were 15 % and 37 % lower than controls in the 15,000 and 40,000 ppm exposure groups, respectively (Fig. 2D). The considerable decrease in size of the F1c pups exposed to 40,000 ppm BP was associated with decreased body weight gain, suggesting overt developmental delay and toxicity at the highest administered dose. Notably the average pup weight of 40,000 ppm-exposed F1c pups was approximately 31.7 g on PND21, corresponding to a similar weight observed in control pups between PND 13 and 16. This delay likely contributed to the observed reduction in survival of this group following weaning on PND21.

Following weaning (PND 21 or 28 to scheduled removal), F1c body weights (F1cP and F1cNP combined) recovered partially in the 15,000 and 40,000 ppm groups, decreasing the magnitude of difference from controls observed at earlier timepoints (Fig. 3A). After the F1cP pairings, body weights of F1cP 40,000 ppm dams were up to 17 %, 16 %, and 17 % lower than controls during gestational intervals A, B, and C, respectively (Fig. 3B). This effect corresponded to lower gestational body weight gains relative to controls. Smaller litter

sizes in the 40,000 ppm group relative to the control group may have contributed to lower gestation weight. During lactation (LD 1–2 1, Litter C), F1c female body weights were up to 7 % and 20 % lower than controls in the 15,000 and 40,000 ppm groups, respectively (Fig. 3C). Similar to F1 litter observations, perinatal exposure to BP was associated with significant decreases in F2c pup body weights during their respective preweaning interval (PND 1–21). On PND 21, F2c male pup body weights were 11 % and 36 % lower than controls in the 15,000 and 40,000 ppm exposure groups, respectively (Fig. 3D). F2c female pup PND 21 body weights were 6 %, 15 %, and 40 % lower than controls in the 5000, 15,000, and 40,000 ppm exposure groups, respectively (Fig. 3D). There was minimal effect of BP-exposure on pup weight on PND 1, indicating that the reduction in pup body weight on PND 21 was due to lower weight gain during early postnatal development, presumably through maternal transfer and then direct exposure through feed consumption [44].

3.4. Reproductive performance and littering parameters

There was no effect of dietary BP exposure on indices of fertility (mated/pair, littered/pair, littered/mated) in any exposure group, relative to concurrent controls (Table 3, CEBS, RACB R02). Positive trends observed in the F1c "A" pairing were likely due to abnormally low mating/pair in the control group (CEBS, RACB R02). When the A,B,C pairings were averaged for each generation, no exposure related effects were apparent (Table 3) (CEBS, RACB R21). Additionally, there were no significant effects on the duration of the precoital or gestational intervals (CEBS, RACB R02).

A significant decreasing trend in total litter size occurred in F0 A and B pairings and F1cP A pairing with increasing exposure concentration (Table 4, CEBS, RACB R03/R22). Total litter size was significantly lower (~ 2 pups/litter) in the initial litters (litter A) produced by 40,000 ppm-exposed F0 and F1cP dams relative to controls. Additionally, a significant decreasing trend in live litter size was observed in litter A of the F0 breeding cohort. Smaller litter sizes were also observed in 40,000 ppm exposed F0 litter C and F1cP litters B and C relative to concurrent controls, however these did not attain statistical significance. No significant effects upon live litter size or pup survival were observed in 40,000 ppm F0 and F1cP litters. When looking at average live litter size across pairs, litter sizes were marginally lower in the 40,000 ppm group.

3.5. Crossover-Mating

Due to the purported spermatotoxic effects of BP in the literature, and the observed reduction in mean litter size of 40,000 ppm F0 dams, a crossover-mating was conducted to further evaluate reproductive performance and evaluate potential sex-related susceptibilities. Treated males (40,000 ppm, aged ~37 weeks) were bred with unexposed nulliparous females (aged ~11 weeks), and treated females (40,000 ppm, aged ~37 weeks) were bred with control males (0 ppm, aged ~37 weeks). There was no overt effect on fertility (e.g. litter/ mated, litter size, etc.) associated with male or female exposure to BP, when comparing the results to the C littering (Supplemental Table 1, CEBS, RACB R03). Reduced litter sizes were evident in the 40,000 ppm females/0 ppm male crossover-matings; however, it appeared to be consistent with a trend of decreasing litter size with the aging F0 females

(e.g. total litter was 11.7, 11.8, 9.4, 9.4 for the A, B, C and crossover mating of exposed females respectively).

3.6. Developmental and pubertal markers

Anogenital distance, a marker of developmental androgen activity, was not affected by perinatal BP exposure in male or female offspring (Supplemental Table 2, CEBS, RACB R04). Two male pups from separate litters in the F1c control group and two male pups derived from a single litter in the F2c 40,000 group presented with retained areolae/nipples (Table 5, CEBS, RACB R14). This observation was considered unrelated to BP-exposure. A significant decreasing trend in time to onset of testicular descent occurred in F2c males with increasing exposure concentration (CEBS, RACB R14). However, this effect was only significant in the 15,000 ppm group by pairwise comparison (Table 5), and did not occur in F1c generation.

Evaluation of markers of pubertal onset were limited to F1c males and females (CEBS, RACB R16). Balaneo-pre-preputial Separation (BPS) and Vaginal Opening (VO) were delayed in F1c male and female off-spring, respectively (Table 5). A significant delay in the onset of BPS (~ 5 days) occurred in 40,000 ppm F1c males relative to controls. Additionally, a significant delay in the onset of VO of ~ 3 days and ~ 9 days occurred in 15,000 and 400,000 ppm exposed F1c females, respectively. When evaluating effects upon BPS and VO it is important to consider additional parameters such as body weights at attainment and weaning, which are known to influence attainment independent of potential chemical hormonal activity [45]. Male body weight at attainment of BPS was significantly lower (~12 %) in the 40,000 ppm group relative to controls. Male weaning weights were ~ 14 % and ~30 % lower relative to controls in the 15,000 and 40,000 ppm groups, respectively. When adjusted using body weight at weaning (PND 21) as a covariate, a significant trend of delayed BPS with increasing exposure concentration was observed without any corresponding significant pairwise differences between vehicle and BP-exposed groups. Female body weight at attainment of VO was similar to controls at all exposure levels tested. Female weaning weights were ~13 %, ~18 % and ~29 % lower relative to controls in the 5000, 15,000 and 40,000 ppm groups, respectively. When adjusted using body weight at weaning (PND 21) as a covariate, a trend of delayed VO with increasing exposure concentration was observed and significant delays of ~2 days and ~7 days occurred respectively in 15,000 and 40,000 ppm groups relative to controls. Given the magnitude of the effect of BP exposure on body-weight, observed delays in attainment of BPS and VO are likely secondary to growth retardation.

3.7. Sperm analysis and estrous cyclicity

Sperm quality parameters were evaluated at terminal necropsy in the F0 (~Study day (SD) 225), F1cNP interim (~PND 95), and F1cP terminal (~PND 210) groups (Table 6, CEBS, RACB R06). Exposure to BP did not affect any of the assessed sperm parameters in any test group. Additionally, estrous cyclicity was not affected in either F0 (~SD 161–176) or F1c (PND 52–67) females at all exposure levels tested (Supplemental Table 3, CEBS, RACB Vaginal Cytology Data).

3.8. Organ weights

Organ weights were evaluated in male and female rats of the F0 (Males ~SD 225, Females ~SD 218), F1cNP (~PND 91), and F1cP (Males ~PND 210, Females ~PND 235) groups at necropsy (CEBS, RACB PA06R). Alteration of liver weights in male and female rats was associated with BP-exposure (Table 7). In male rats, there were decreasing trends in mean absolute liver weights with increasing exposure concentration in all assessed generations, and significant decreases in mean absolute liver weights were observed in the 15,000 (−3 % and −8 %) and 40,000 ppm (−12 and −15 %) groups of the interim and terminal timepoints, respectively. This effect was not observed in female rats. However, liver weights adjusted to body weights (relative liver weights) showed significant increasing trends in all assessed male and female groups with increasing exposure concentration. Mean relative liver weights were significantly increased in all 40,000 ppm-exposed male cohorts, and 15,000 ppm-exposed males of the F0 and F1cNP cohorts relative to controls. Group mean relative liver weights were significantly increased in the female F0/F1cP 40,000 ppm-exposure groups and F1cP/F1cNP 15,000 ppm-exposure groups relative to controls.

Decreased reproductive organ weights in male and female rats were associated with exposure to BP, but were generally considered to be secondary to lower body weights (Supplemental Table 4/5, CEBS, RACB PA06R). Significant decreases in mean absolute weights of the testes, epididymides, seminal vesicles, dorsolateral prostate, and ventral prostate were noted in the 40,000 F1cNP cohort. However, these results were confounded by the limited number of test animals assessed in this group (n = 6) relative to the F0/F1cP cohorts (n = 22–26). Absolute ventral prostate weights were decreased 15 % and 37 % in the 15,000 ppm and 40,000 ppm F1c interim males, respectively, and decreased 14 % in the 40,000 ppm terminal male group. There was a decreasing trend in relative prostate weights in the interim males, although without a significant pairwise comparison, a similar response was not observed in terminal (parental) males. It is not clear if these effects were secondary to body weight changes or direct action of the chemical. In female rats, significantly decreased mean absolute left ovary weights in the F1cP cohort coincided with significantly decreased terminal body weights relative to controls. Significantly increased mean relative ovary weights (+ 70 %) occurred in the F0 40,000 ppm group. However, it was determined that this effect was unlikely treatment-related, and related to the differential completion of a crossover mating by the F0 40,000 ppm group, relative to the other assessed F0 female groups. F0 40,000 ppm females were necropsied 4-days post-parturition of the crossover mating litter, whereas all other F0 females had completed delivery of their final litter (F1c) weeks prior. Therefore, it is plausible that the recent completion of parturition and corresponding presence of residual *corpora lutea* in the ovary caused the increased ovary weights.

3.9. Histopathology

Histopathologic lesions related to BP-exposure were limited to the liver in male rats, and the liver and adrenal gland of female rats (Table 8, CEBS, RACB PA02R/03R/08R/10R).

3.9.1. Liver—Increased incidences of bile duct hyperplasia were identified in the 15,000 and 40,000 ppm F1c parental males and in the 40,000 ppm F1c parental females, relative to

controls. Bile duct hyperplasia consisted of increased profiles of bile ducts within the portal region; often the hyperplastic bile ducts were surrounded by a thin rim of fibrous connective tissue. On occasion, mononuclear cell infiltrates were found in association with bile duct hyperplasia; these were not recorded separately.

Incidence of mononuclear cell infiltrates were increased in the F0, F1cNP, and F1cP males and females exposed to 15,000 or 40,000 ppm BP. They were generally of minimal severity, and characterized by small clusters of lymphocytes, with fewer plasma cells and macrophages, either randomly distributed in the liver, or located within the portal region, distinct from areas of bile duct hyperplasia. While mononuclear cell infiltrates are a common background finding in rats, they were only recorded when clusters of mononuclear cells were present in numbers above what was considered a background level.

In F0 males and F1c parental males, there were increased incidences of hepatocyte cytoplasmic vacuolation in the 15,000 and 40,000 ppm groups; there was also a significant increasing trend in F1c parental females with increasing exposure concentration. This change was characterized by hepatocytes with small vacuoles consistent with microvesicular lipid accumulation. Most of the affected hepatocytes were found in the periportal region, especially in females, but in males, small to coalescing, randomly located areas of cytoplasmic vacuolization were also present.

Increased incidences of hypertrophy of the periportal hepatocytes were recorded in F0 females exposed to 15,000 and 40,000 ppm BP, and in F1c parental females exposed to 5000, 15,000, and 40,000 ppm BP relative to controls; females in the F1c non-parental cohort had a significant increasing trend with exposure concentration, but no pairwise significance when compared to controls. Increased incidences were not seen in males. This change was characterized by hepatocytes, in the portal region, that had increased amounts of pale eosinophilic to amphiphilic cytoplasm.

When compared to controls, there were an increased number of F1c parental males in the 40,000 ppm group with increased hepatocytes undergoing mitosis. There was a significant increasing trend in increased mitoses in F0 females and males and males with increasing exposure concentration. While occasional mitotic figures can be found in the normal rat liver, the number observed was increased over background levels in animals in which increased mitoses were recorded. However, the severity of the increases was minimal in all cases.

In F1c parental females, there was an increased incidence of animals in the 40,000 ppm group with cytoplasmic inclusions in the hepatocytes compared to the control group. These inclusions were found within periportal hepatocytes and were small, round and brightly eosinophilic; typically, one to three inclusions were present in the cytoplasm of an affected hepatocyte.

3.9.2.   Adrenal cortex—Increased incidences of cytoplasmic vacuolation of the zona glomerulosa were present in the adrenal cortices of the 15,000 and 40,000 ppm groups of F1c parental females. The vacuolation was bilateral and was a diffuse change throughout

the zona glomerulosa; the zona fasciculata and zona reticularis were not affected. Within the zona glomerulosa the cells had a distended cytoplasm due to prominent vacuoles, which caused an overall thickening of the zona glomerulosa. There was no apparent increase in the number of cells in this region.

4. Discussion

Public concern regarding potential endocrine disrupting effects of the paraben class of chemicals have been on the rise. Numerous studies of butylparaben report estrogenic activity and possible adverse effects upon the reproductive system, while other findings are contradictory and indicate little concern surrounding the safe use of these chemicals [19,21,46–50]. Due to the inconclusiveness surrounding the potential developmental and reproductive toxicity of butylparaben and wide-spread human exposure, this test article was evaluated by the NTP using the multigenerational RACB testing paradigm.

In this study, BP exposure concentration was associated with effects on bodyweight in both sexes and across all three generations. The highest magnitude of lower body weights were noted in perinatally exposed rats of the F1c and F2c cohorts. Generally, the effect on weight occurred late in lactation (~PND 19) in the 15,000 ppm group, while the effects were present earlier (~PND 4–7) in lactation in the 40,000 ppm group. Weight decrements observed in the 40,000 ppm offspring on PND 21 corresponded to severely decreased body weight gain, such that their body weights were equivalent to that of PND 14 unexposed controls. The significant mortality observed in this group likely corresponded to the inability to thrive without further maternal intervention, following weaning on PND 21. This hypothesis is further supported by the 100 % survival of 40,000 ppm offspring whose weaning was delayed to PND 28. Interestingly, no significant differences in offspring birthweight were observed, indicating that lower body weight gain coincided with lactational exposure (40,000 ppm group) and the onset of direct consumption of dosed feed (15,000 and 40,000 ppm) in developing offspring. The NTP previously conducted studies to evaluate potential for gestational and lactational transfer of butylparaben to offspring in rats [44]. BP displayed low gestational/lactational transfer potential; however, higher concentrations of the BP parent molecule, relative to conjugated metabolites, were found in offspring relative to dams during the lactational interval. This suggests developing offspring may have limited capacity to metabolize this substance, resulting in higher internal exposure to the BP parent-molecule during the developmental period.

No effects on reproductive performance (e.g. mating or litter parameters) of the F0 or F1c were associated with BP-exposure in the current study. Although not always statistically significant, the average total and live litter size on PND 0 of the 40,000 ppm group was generally 1–2 pups fewer than controls across all three pairings of both the F0 and F1c. The marginal, but biologically significant decrease in litter size may have contributed to the lower maternal body weights during gestation. There were no corresponding effects on sperm parameters, and histological evaluation of the testis and epididymis did not reveal an effect that would explain the smaller litter size. A crossover mating of the F0 was conducted to investigate a potential sex effect of the smaller litter size. Litter size of the 40,000 ppm males mated with naïve females was consistent with the control F1a,b,c litter size (13.1 vs

12.5 pups/litter respectively), while 40,000 ppm females mated with unexposed males had a total litter size in line with the 40,000 ppm group total litter size (11.7, 11.8, 9.4, 9.4 pups/litter for the 40,000 ppm F1a,b,c, and female crossover respectively). This is suggestive of an effect in females, but not conclusive. Due to the reported rapid metabolism of BP, it is unlikely that detriments to reproduction reflect a bioaccumulation of the test article, but may be associated with the increased age of the dam at the time of third mating interval (litter C) [51]. Previous lifetime breeding studies in female CFY Sprague Dawley rats found that total litter size significantly decreases with increasing age of dams [52].

Markers of normal endocrine signaling during gestation (AGD, nipple retention) and pubertal development (BPS, VO) were unaffected by BP-exposure. While onset of BPS was delayed in F1c 40,000 ppm males, this effect was likely associated with the severe growth delay induced in this group and not due to endocrine disrupting activity of the test article. In support of this notion, when age at BPS was adjusted for weight at weaning, there were no longer statistically significant pairwise differences between exposed groups and controls [45,53]. Correspondingly, delays in age at VO of F1c 15,000 and 40,000 ppm females were also likely attributed to growth retardation based on the observation that all dose groups had nearly equivalent body weights at the age of vaginal opening [54,55]. Estrous cyclicity was also unaffected by BP exposure. Estrogenic agents would be anticipated to accelerate onset of VO and induce prolonged estrus in test animals, which did not occur in the current study, indicating BP did not induce prototypical *in vivo* estrogenic effects [56].

A recent human epidemiological study found that elevated urinary paraben levels had a significant positive association with diminished sperm count and motility measures among 501 male participants [57]. However, other studies did not find significant correlations between paraben exposure and reduced semen quality in males among sub-fertile couples [23,24]. Decreased serum testosterone and spermatotoxicity in rats following BP-exposure have been reported in numerous studies [19,21,22,58,59]. A 2002 study by Oishi reported dose related decreases in testicular and epididymal sperm counts, and decreased serum testosterone in Crj:CD-1 mice following exposure to BP in dosed feed from 4 to 14 weeks of age, establishing a lowest-observed adverse effect level (LOAEL) of 10 mg/kg dietary concentration (intake of 14.4 mg BP/kg/day). In contrast to these findings, all assessed sperm parameters in the current study, including testicular spermatid count, motility, and caudal sperm count were unaffected by dietary BP at daily exposures in excess of 300 mg BP/kg/day. Additionally, no histological findings and only sporadic weight effects were noted in assessed male reproductive organs in exposed groups. The lack of functional effects upon fertility throughout the RACB assessment and in the crossover mating of 40,000 ppm exposed males with naïve females provides further support to our conclusion that BP did not adversely affect sperm number or quality.

In concordance with our overall findings, lack of BP-mediated reproductive/developmental toxicity has been reported in two additional studies utilizing rats with exposure by diet or oral gavage. Young male Wistar rats exposed to BP in diet containing up to 10,000 ppm BP for eight weeks displayed no effects on any male reproductive organ or sperm parameters [50]. BP-exposure via oral gavage up to 1000 mg/kg/day during gestation in Sprague-Dawley Crl:CD®BR VAF/Plus® rats did not affect litter size or induce

developmental toxicity [48]. While, decreased sperm number/motility and altered testicular Leydig cell number has been reported in rats following subcutaneous dosing of BP (up to 200 mg/kg/day) during gestational and early postnatal development [22,59]. Different routes of BP exposure among studies may contribute to the disparate toxicity findings in the available literature. Metabolism and disposition studies found BP to be well absorbed and rapidly conjugated or hydrolyzed to 4-hydroxybenzoic acid and 4-hydroxyhippuric acid following oral gavage administration [60]. Subcutaneous injection of BP bypasses first-pass metabolism in the liver and the intestinal tract, and could facilitate heightened exposure to parent BP relative to oral-exposure [60]. The discordance of findings between this study and that of the rat and mouse studies conducted by Oishi [19,21] may be related to the different species or strain utilized.

Following necropsy, the liver was identified as the primary target organ of BP toxicity due to dose related increases in relative liver weight and increased incidences bile duct hyperplasia, mononuclear cell infiltration, periportal hepatocyte hypertrophy, hepatocyte cytoplasmic inclusions, and hepatocyte cytoplasmic vacuolation. Incidence patterns of some hepatic lesions varied across exposed F0, F1cNP, and F1cP cohorts. Increased incidence of mononuclear cell infiltration was the only dose related microscopic finding identified in F1cNP interim animals, suggesting onset of other hepatic lesions may require a longer duration of exposure. Notably, higher incidences of bile duct hyperplasia and periportal hepatocyte hyperplasia occurred predominantly in the exposed F1c cohort, relative to the F0 generation. These observations may be associated with a specific exposure window of susceptibility to BP-mediated hepatic toxicity or heightened exposure during the developmental period. To our knowledge, increased incidence of hepatocellular lesions and hepatomegaly have not been previously ascribed to BP-exposure in animal models. Evidence of hepatic toxicity in the current study may be associated with sustained adaptive responses in the liver as a result of developmental/long-term exposure to BP. Additionally, increased incidence of mild bilateral adrenal cortical vacuolization was noted in exposed F1cP females. This lesion may be caused by a perturbation of the hypothalamic-pituitary-adrenal hormone axis via altered secretion/metabolism of adrenocorticotropic hormone (ACTH) [61] or could related to overall conditions of stress. Previous assessments of BP disposition following oral exposure indicate low distribution of the radio-labeled test article to tissues (~1 % of administered dose), but high tissue/blood ratios in target tissues (liver, kidney/adrenal) identified in the current study [51].

Differences in BP safety assessment considerations among regulatory authorities have led to regional discrepancies in the levels of BP allowed in consumer products. In 2010 Denmark's environmental ministry announced that both propyl- and butylparabens are now banned in lotions and other products for children under age three years [62]. This change in policy was mainly due to concerns of endocrine disruption and the possibility of heightened exposure related to possible decreased carboxy-esterase activity in the skin of young children relative to adults. The European Union Scientific Committee on Consumer Safety (SCCS) recently established new cosmetic maximum concentration limits of 0.4 % for methyl paraben or ethyl paraben (single esters/salts), 0.19 % for propyl paraben or BP (single esters/salts), and 0.8 % for mixtures of these four parabens, wherein each paraben component does not exceed its individual limit. These levels were based on a no observed effect level (NOEL) of 2

mg/kg/day derived from a study by Fisher et al. (1999) in which juvenile rats were exposed to BP (2 mg/kg/day) via subcutaneous administration for 17 days (PND 2–18) exhibited no alteration in the structure of the testicular excurrent ducts [58]. Decreased limits of paraben use were established by the Scientific Committee on Consumer Safety to ensure a margin of safety of at least 100-times that of anticipated human exposure [63]. However, in a recent final safety assessment by the U.S. Cosmetic Ingredient Review (CIR) board, a no observed adverse effect level (NOAEL) of 160 mg/kg/day was used, given the reported endocrine effects in F1 rat offspring when dams were gavaged GD 7 through PND 21 [64]. Using this NOAEL and a systemic adult exposure dose estimate of 0.35 mg/kg/day, a conservative 457-fold margin of safety was quantitated [65]. The NOEL/NOAEL determinations for both assessments were based on observations of adverse effects on the reproductive system in developmentally exposed male offspring. The European Food Safety Authority (EFSA) and the Joint FAO/WHO Expert Committee on Food Additives (JECFA) concluded that more data are required before the safe use of BP as a food additive can be evaluated [66].

In the current study, there were limited findings to support BP-endocrine induced developmental or reproductive toxicity following dietary exposure. A lowest observed effect level (LOEL) of 5000 ppm (approximately 325–740 mg/kg/day) was determined based on F1 general toxicological findings of a dose-related increase in the incidence of mild periportal hepatocyte hypertrophy in perinatally-exposed female rats. This was observed at a lower exposure than the F0 LOEL (15,000 ppm). Although, a NOEL was not identified in the current study, all assessed *in vivo* measures of potential estrogenic and anti-androgenic activity were unperturbed at exposure levels that far exceed those currently used for BP risk assessments and margin of safety determinations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Abbreviations:

AGD	Anogenital distance
BPS	Balaneo-pre-preputial separation
BP	Butylparaben
GD	gestation day
LD	lactation day

LOAEL	lowest-observed adverse effect level
LOEL	lowest observed effect level
NOAEL	no observed adverse effect level
NOEL	no observed effect level
PND	postnatal day
RACB	reproductive assessment by continuous breeding
SD	study day
VO	vaginal opening

References

[1]. Freese E, Sheu CW, Galliers E, Function of lipophilic acids as antimicrobial food additives, Nature 241 (1973) 321–325 doi:10.1038/241321a0. [PubMed: 4633553]

[2]. Nes IF, Eklund T, The effect of parabens on DNA, RNA and protein synthesis in Escherichia coli and Bacillus subdlis, J. Appl. Bacteriol 54 (1983) 237–242, 10.1111/j.1365-2672.1983.tb02612.x. [PubMed: 6189812]

[3]. 21 CFR 172.515, FDA, 2017.

[4]. Liao C, Liu F, Kannan K, Occurrence of and dietary exposure to parabens in foodstuffs from the United states, Environ. Sci. Technol 47 (2013) 3918–3925, 10.1021/es400724s. [PubMed: 23506043]

[5]. Rastogi SC, Schouten A, de Kruijf N, Weijland JW, Contents of methyl-, ethyl-, propyl-, butyl-and benzylparaben in cosmetic products, Contact Derm. 32 (1995) 28–30.

[6]. Elder RL, 3 final report on the safety assessment of methylparaben, Ethylparaben, Propylparaben, and butylparaben, J. Am. Coll. Toxicol 3 (2007) 147–209, 10.3109/10915818409021274.

[7]. FDA, Ingredient Search for Approved Drug Products, (2017) (Accessed November 16, 2017), https://www.accessdata.fda.gov/scripts/cder/iig/index.cfm.

[8]. Centers for disease control and prevention (CDC), Fourth national report on human exposure to environmental chemicals, Fourth Natl. Rep. Hum. Expo, to Environ. Chem Updat (2013) Tables, Sept. 2013 http://www.cdc.gov/exposurereport/pdf/FourthReport_UpdatedTables_Sep2013.pdf.

[9]. Centers for disease control and prevention (CDC), Fourth National Report on Human Exposure to Environmental Chemicals, Updated Tables, 2019 (2019).

[10]. Soni MG, Carabin IG, Burdock GA, Safety assessment of esters of p-hydroxybenzoic acid (parabens), Food Chem. Toxicol 43 (2005) 985–1015, 10.1016/j.fct.2005.01.020. [PubMed: 15833376]

[11]. Routledge EJ, Parker J, Odum J, Ashby J, Sumpter JP, Some alkyl hydroxy benzoate preservatives (Parabens) are estrogenic, Toxicol. Appl. Pharmacol 153 (1998) 12–19, 10.1006/taap.1998.8544. [PubMed: 9875295]

[12]. Golden R, Gandy J, Vollmer G, A review of the endocrine activity of parabens and implications for potential risks to human health, Crit. Rev. Toxicol 35 (2005) 435–458, 10.1080/10408440490920104. [PubMed: 16097138]

[13]. Byford JR, Shaw LE, Drew MGB, Pope GS, Sauer MJ, Darbre PD, Oestrogenic activity of parabens in MCF7 human breast cancer cells, J. Steroid Biochem. Mol. Biol 80 (2002) 49–60, 10.1016/S0960-0760(01)00174-1. [PubMed: 11867263]

[14]. Darbre PD, Byford JR, Shaw LE, Horton RA, Pope GS, Sauer MJ, Oestrogenic activity of isobutylparaben in vitro and in vivo, J. Appl. Toxicol 22 (2002) 219–226, 10.1002/jat.860. [PubMed: 12210538]

[15]. Darbre PD, Byford JR, Shaw LE, Hall S, Coldham NG, Pope GS, Sauer MJ, Eostrogenic activity of benzylparaben, J. Appl. Toxicol 23 (2003) 43–51, 10.1002/jat.886. [PubMed: 12518336]

[16]. Okubo T, Yokoyama Y, Kano K, Kano I, ER-dependent estrogenic activity of parabens assessed by proliferation of human breast cancer MCF-7 cells and expression of ERa and PR, Food Chem. Toxicol 39 (2001) 1225–1232, 10.1016/S0278-6915(01)00073-4. [PubMed: 11696396]

[17]. Darbre PD, Harvey PW, Paraben esters: review of recent studies of endocrine toxicity, absorption, esterase and human exposure, and discussion of potential human health risks Paraben esters: review of recent á3endocrine toxicity, J. Appl. Toxicol. J. Appl. Toxicol 28 (2008) 561–578, 10.1002/jat.1358. [PubMed: 18484575]

[18]. Miller D, Wheals BB, Beresford N, Sumpter JP, Estrogenic activity of phenolic additives determined by an in vitro yeast bioassay, Environ. Health Perspect. 109 (2001) 133–138 . doi:sc271_5_1835 [pii]. [PubMed: 11266322]

[19]. Oishi S, Effects of butyl paraben on the male reproductive system in mice, Arch. Toxicol. 76 (2002) 423–429, 10.1007/s00204-002-0360-8. [PubMed: 12111007]

[20]. Oishi S, Lack of spermatotoxic effects of methyl and ethyl esters of p-hydroxybenzoic acid in rats, Food Chem. Toxicol. 42 (2004) 1845–1849, 10.1016/j.fct.2004.06.015. [PubMed: 15350682]

[21]. Oishi S, Effects of butylparaben on the male reproductive system in rats, Toxicol. Ind. Health 17 (2001) 31–39, 10.1191/0748233701th093oa. [PubMed: 12004923]

[22]. Kang K-S, Che J-H, Ryu D-Y, Kim T-W, Li G-X, Lee Y-S, Decreased sperm number and motile activity on the F1 offspring maternally exposed to butyl p-Hydroxybenzoic acid (Butyl paraben), J. Vet. Med. Sci. 64 (2002) 227–235 (Accessed December 12, 2017), https://www.jstage.jst.go.jp/article/jvms/64/3/64_3_227/_pdf/-char/en. [PubMed: 11999442]

[23]. Nishihama Y, Toshima H, Yoshinaga J, Mizumoto Y, Yoneyama M, Nakajima D, Shiraishi H, Tokuoka S, Paraben exposure and semen quality of Japanese male partners of sub fertile couples, Environ. Health Prev. Med 22 (2017), 10.1186/s12199-017-0618-7.

[24]. Meeker JD, Yang T, Ye X, Calafat AM, Hauser R, Urinary concentrations of parabens and serum hormone levels, semen quality parameters, and sperm DNA damage, Environ. Health Perspect 119 (2011) 252–257, 10.1289/ehp.1002238. [PubMed: 20876036]

[25]. Morrissey RE, Lamb JC, Morris RW, Chapin RE, Gulati DK, Heindel JJ, Results and evaluations of 48 continuous breeding reproduction studies conducted in mice, Toxicol. Sci 13 (1989) 747–777, 10.1093/toxsci/13.4.747.

[26]. Chapin RE, Sloane RA, Reproductive assessment by continuous breeding: evolving study design and summaries of ninety studies, Env. Heal. Perspect 105 (1997) 199–395 (Accessed November 2, 2017), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1470239/pdf/envhper00326-0196.pdf.

[27]. NRC, Guide for the Care and Use of Laboratory Animals, Washington (DC) (2011), 10.17226/12910.

[28]. Cora MC, Kooistra L, Travlos G, Vaginal Cytology of the Laboratory Rat and Mouse:Review and Criteria for the Staging of the Estrous Cycle Using Stained Vaginal Smears, Toxicol. Pathol 43 (2015) 776–793, 10.1177/0192623315570339. [PubMed: 25739587]

[29]. Maronpot RR, Boorman GA, Interpretation of rodent hepatocellular proliferative alterations and hepatocellular tumors in chemical safety assessment, Toxicol. Pathol 10 (1982) 71–78, 10.1177/019262338201000210. [PubMed: 28094716]

[30]. Boorman GAJ, Montgomery CA Jr.,, Eustis SL, Wolfe MJ, McConnell EE, Hardisty, Quality assurance in pathology for rodent carcinogenicity studies, Handb. Carcinog. Test, Noyes Publications, Park Ridge, NJ, 1985.

[31]. Boorman GA, Haseman JK, Waters MD, Hardisty JF, Sills RC, Quality review procedures necessary for rodent pathology databases and toxicogenomic studies: the National Toxicology Program experience, Toxicol. Pathol 30 (2002) 88–92, 10.1080/01926230252824752. [PubMed: 11890481]

[32]. Datta S, Satten GA, Rank-sum tests for clustered data, J. Am. Stat. Assoc 100 (2005) 908–915, 10.1198/016214504000001583.

[33]. Jonckheere AR, A distribution-free k-sample test against ordered alternatives, Biometrika 41 (1954) 133–145, 10.2307/2333011.

[34]. Shirley E, A non-parametric equivalent of Williams' test for contrasting increasing dose levels of a treatment, Biometrics 33 (1977) 386–389, 10.2307/2529789. [PubMed: 884197]

[35]. Williams DA, A note on Shirley's nonparametric test for comparing several dose levels with a zero-dose control, Biometrics 42 (1986) 183–186, 10.2307/2531254. [PubMed: 3719054]

[36]. Dunn OJ, Multiple comparisons using rank sums, Technometrics 6 (1964) 241–252, 10.1080/00401706.1964.10490181.

[37]. Portier C, Bailer A, Testing for increased carcinogenicity using a survival-adjusted quantal response test, Fundam. Appl. Toxicol (1989) 731–737. [PubMed: 2744275]

[38]. Piegorsch W, Bailer A, Statistics for Environmental Biology and Toxicology, Chapman & Hall, New York, NY, USA, 1997.

[39]. Bailer AJ, Portier CJ, Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples, Biometrics 44 (2006) 417–431, 10.2307/2531856.

[40]. Bieler GS, Williams RL, Ratio estimates, the Delta method, and quantal response tests for increased carcinogenicity, Biometrics 49 (1993) 793, 10.2307/2532200. [PubMed: 8241374]

[41]. Nam J, A simple approximation for calculating sample sizes for detecting linear trend in proportions, Biometrics. 43 (1987) 701, 10.2307/2532006. [PubMed: 3663825]

[42]. Gart RE, Chu JJ, Tarone KC, Statistical issues in interpretation of chronic bioassay tests for carcinogenicity, J. Natl. Cancer Inst 62 (1979) 957–974. [PubMed: 285297]

[43]. Dixon J, W.J, Massey FJ, Introduction to Statistical Analysis, 2nd ed., McGraw-Hill Book Company, Inc., New York, 1957.

[44]. Roberts GK, Waidyanatha S, Kissling GE, Fletcher BL, Silinski MAR, Fennell TR, Cunny HC, Robinson VG, Blystone CR, Exposure to butyl paraben during gestation and lactation in Hsd:Sprague Dawley SD rats via dosed feed, Toxicol. Reports 3 (2016) 774–783, 10.1016/j.toxrep.2016.09.004.

[45]. Stoker TE, Parks LG, Gray LE, Cooper RL, Endocrine-disrupting chemicals: prepubertal exposures and effects on sexual maturation and thyroid function in the male rat. A focus on the EDSTAC recommendations, Crit. Rev. Toxicol 30 (2000) 197–252, 10.1080/10408440091159194. [PubMed: 10759431]

[46]. Lemini C, Jaimez R, Avila ME, Franco Y, Larrea F, Lemus AE, In vivo and in vitro estrogen bioactivities of alkyl parabens, Toxicol. Ind. Health 19 (2003) 69–79, 10.1191/0748233703th177oa. [PubMed: 15697177]

[47]. Shaw J, deCatanzaro D, Estrogenicity of parabens revisited: impact of parabens on early pregnancy and an uterotrophic assay in mice, Reprod. Toxicol 28 (2009) 26–31, 10.1016/j.reprotox.2009.03.003. [PubMed: 19490991]

[48]. Daston GP, Developmental toxicity evaluation of butylparaben in Sprague-Dawley rats, Birth Defects Res. Part B - Dev. Reprod. Toxicol 71 (2004) 296–302, 10.1002/bdrb.20016.

[49]. Taxvig C, Vinggaard AM, Hass U, Axelstad M, Boberg J, Hansen PR, Frederiksen H, Nellemann C, Do parabens have the ability to interfere with steroidogenesis? Toxicol. Sci 106 (2008) 206–213, 10.1093/toxsci/kfn148. [PubMed: 18648085]

[50]. Hoberman AM, Schreur DK, Leazer T, Daston GP, Carthew P, Re T, Loretz L, Mann P, Lack of effect of butylparaben and methylparaben on the reproductive system in male rats, Birth Defects Res. Part B - Dev. Reprod. Toxicol 83 (2008) 123–133, 10.1002/bdrb.20153.

[51]. Mathews JM, Brown SS, Patel PR, Black SR, Banks TT, Etheridge AS, Fennell TR, Snyder RW, Blystone CR, Waidyanatha S, Metabolism and disposition of [14C]n-butyl-p-hydroxybenzoate in male and female Harlan Sprague Dawley rats following oral administration and dermal application, Xenobiotica 43 (2013), 10.3109/00498254.2012.702935.

[52]. Holehan AM, Merry BJ, Lifetime breeding studies in fully fed and dietary restricted female CFY Sprague-Dawley rats. 1. Effect of age, housing conditions and diet on fecundity, Mech. Ageing Dev 33 (1985) 19–28, 10.1016/0047-6374(85)90106-X. [PubMed: 4079476]

[53]. Ashby J, 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 (2000) 35–47, 10.1002/(SICI)1099-1263(200001/02)20:1<35::AID-JAT633>3.0.CO;2-8. [PubMed: 10641015]

[54]. Kennedy GC, Mitra J, Body weight and food intake as initiating factors for puberty in the rat, J. Physiol 166 (1963) 408–418, 10.1113/jphysiol.1963.sp007112. [PubMed: 14031944]

[55]. Chapin RE, Gulati DK, Fail PA, Hope E, Russell SR, Heindel JJ, George JD, Grizzle TB, Teague JL, The effects of feed restriction on reproductive function in Swiss CD-I mice, Toxicol. Sci 20 (1993) 15–22, 10.1093/toxsci/20.1.15.

[56]. Biegel LB, Flaws JA, Hirshfield AN, O'Connor JC, Elliott GS, Ladies GS, Silbergeld EK, Van Pelt CS, Hurtt ME, Cook JC, Frame SR, 90-day feeding and one-generation reproduction study in Crl:CD BR rats with 17β-estradiol, Toxicol. Sci 44 (1998) 116–142, 10.1006/toxs.1998.2468. [PubMed: 9742652]

[57]. Smarr MM, Honda M, Kannan K, Chen Z, Kim S, Louis GMB, Male urinary biomarkers of antimicrobial exposure and bi-directional associations with semen quality parameters, Reprod. Toxicol 77 (2018) 103–108, 10.1016/j.reprotox.2018.02.008. [PubMed: 29474822]

[58]. Fisher JS, Turner KJ, Brown D, Sharpe RM, Effect of neonatal exposure to estrogenic compounds on development of the excurrent ducts of the rat testis through puberty to adulthood, Environ. Health Perspect 107 (1999) 397–405, 10.1289/ehp.99107397. [PubMed: 10210696]

[59]. Guerra MT, Sanabria M, Leite GAA, Borges CS, Cucielo MS, Anselmo-Franci JA, Foster WG, Kempinas WG, Maternal exposure to butyl paraben impairs testicular structure and sperm quality on male rats, Environ. Toxicol 32 (2017) 1273–1289, 10.1002/tox.22323. [PubMed: 27444704]

[60]. Mathews JM, Brown SS, Patel PR, Black SR, Banks TT, Etheridge AS, Fennell TR, Snyder RW, Blystone CR, Waidyanatha S, Metabolism and disposition of [14C]n-butyl-p-hydroxybenzoate in male and female Harlan Sprague Dawley rats following oral administration and dermal application, Xenobiotica 43 (2013) 169–181, 10.3109/00498254.2012.702935. [PubMed: 22830980]

[61]. Burkhardt WA, Guscetti F, Boretti FS, Todesco AI, Aldajarov N, Lutz TA, Reusch CE, Sieber-Ruckstuhl NS, Adrenocorticotropic hormone, but not trilostane, causes severe adrenal hemorrhage, vacuolization, and apoptosis in rats, Domest. Anim. Endocrinol 40 (2011) 155–164, https://doi.org/10.10l6/j.domaniend.2010.10.002. [PubMed: 21194873]

[62]. Melnick M, Parabens outlawed in childrens' products in Denmark, Time (2010).

[63]. SCCS (Scientific Committee on Consumer Safety), Opinion on parabens, SCCS/1348/10, (2011) https://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_041.pdf.

[64]. Zhang L, Dong L, Ding S, Qiao P, Wang C, Zhang M, Zhang L, Du Q, Li Y, Tang N, Chang B, Effects of n-butylparaben on steroidogenesis and spermatogenesis through changed E2 levels in male rat offspring, Environ. Toxicol. Pharmacol 37 (2014) 705–717, 10.1016/j.etap.2014.01.016. [PubMed: 24607685]

[65]. Cosmetic Ingredient Review (CIR), Amended Safety Assessment of Parabens As Used in Cosmetics, Washington D.C (2019) https://online.personalcarecouncil.org/ctfa-static/online/lists/cir-pdfs/FR746.pdf.

[66]. EFSA, Flavouring Group Evaluation 52 (FGE.52): consideration of hydroxy- and alkoxy-substituted benzyl derivatives evaluated by JECFA (57th meeting) structurally related to benzyl alcohols, benzaldehydes, a related acetal, benzoic acids, and related esters eval, EFSA J. 637 (2008) 1–69.
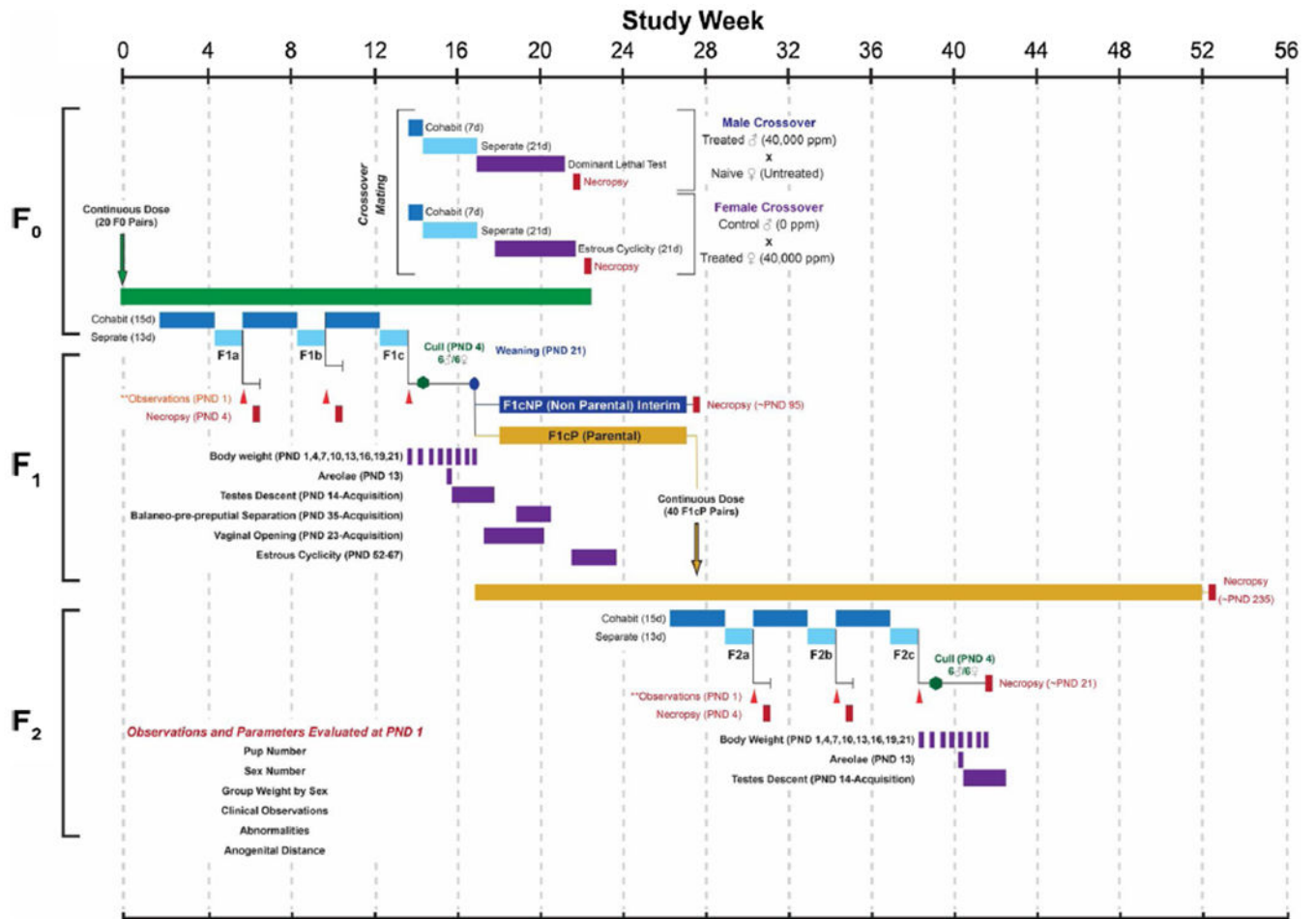
Fig. 1.
Representative diagram depicting overall RACB study design.
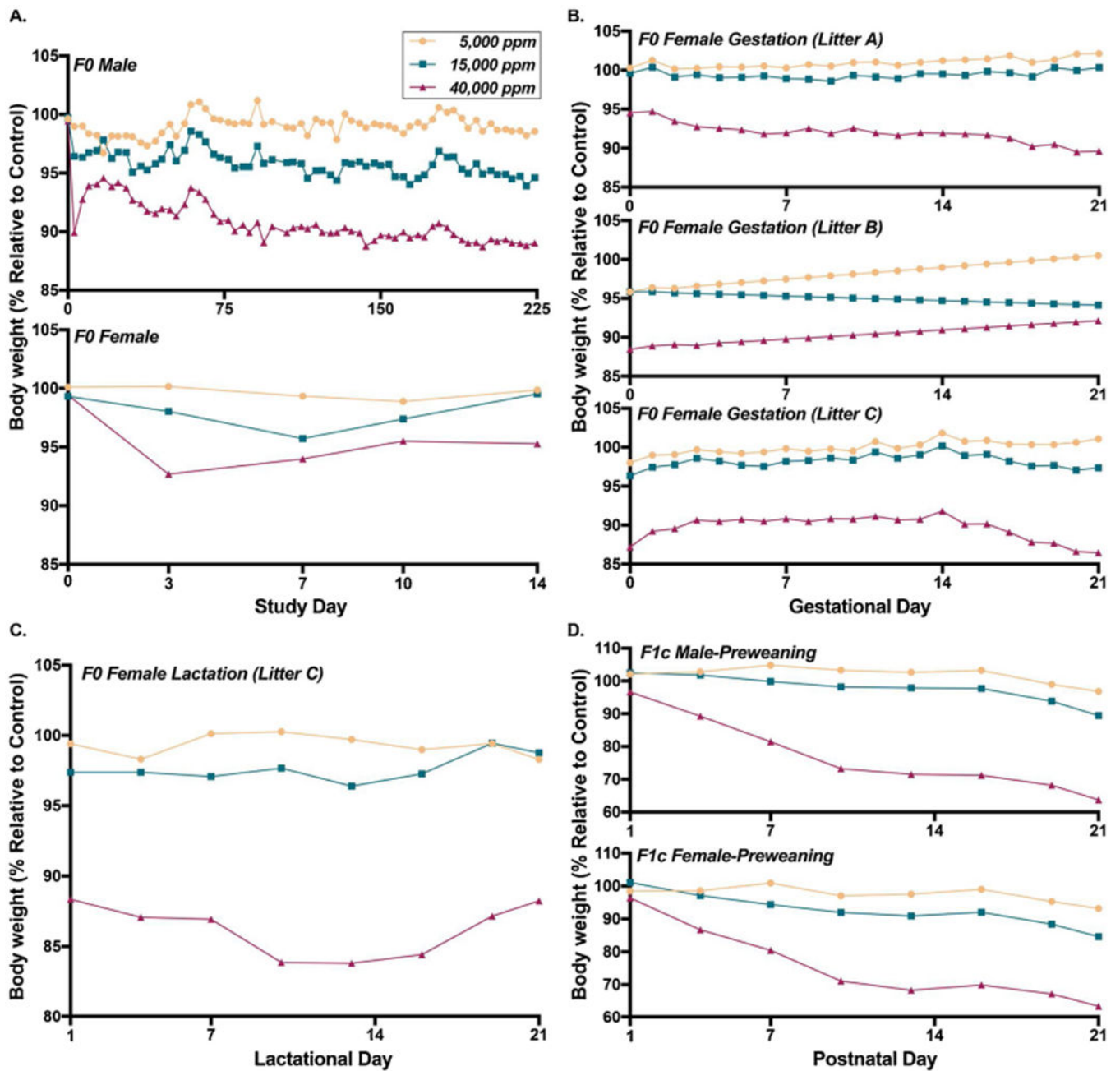
Fig. 2.
Summary of relative body weights (% control) of study animals measured across multiple generations/reproductive intervals, (A) F0 male/female, (B) F0 females during gestation intervals A, B, and C, (C) F0 females during lactation interval (litter C), and (D) F1c male/female pups prior to weaning.
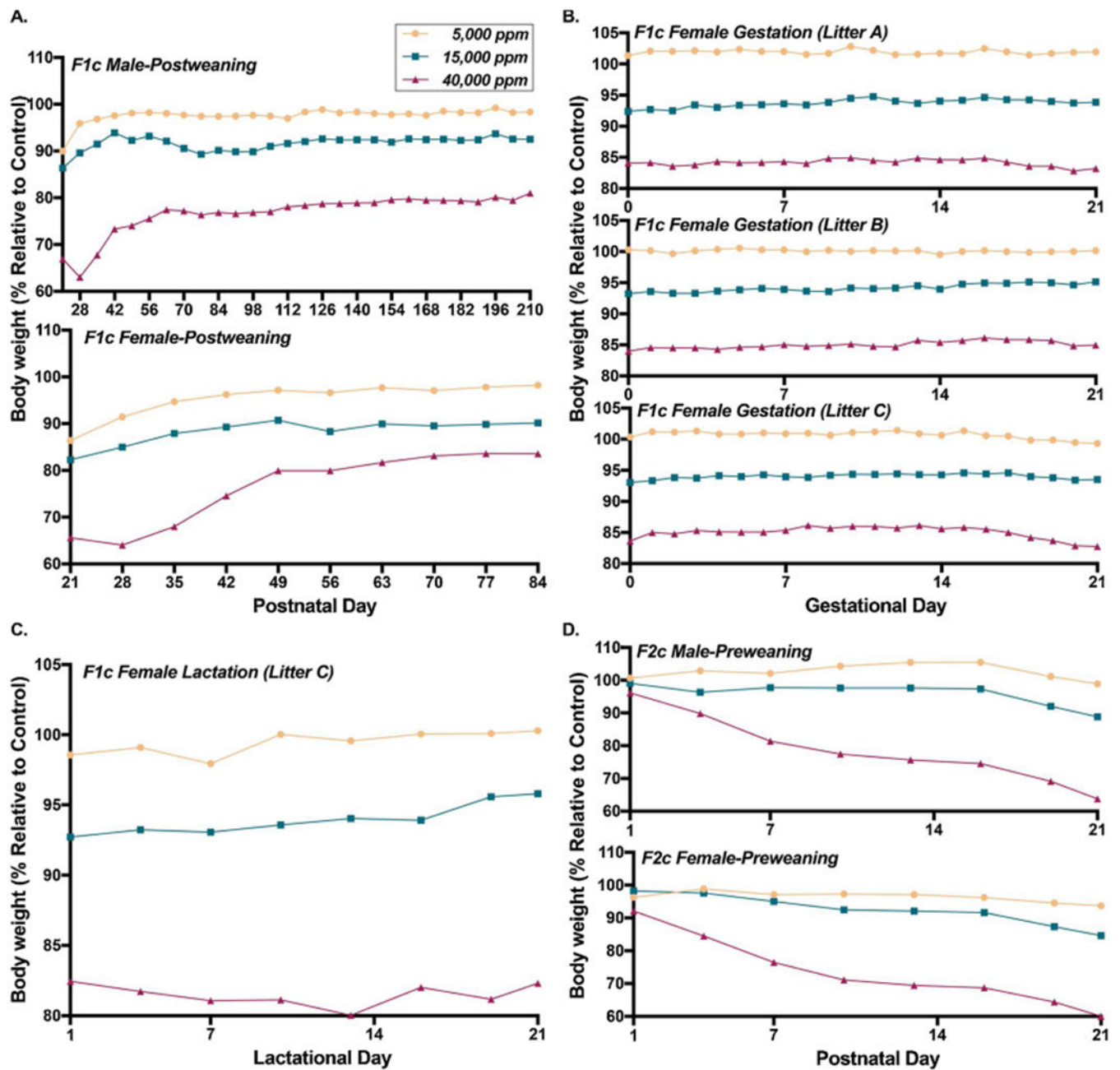
Fig. 3.
Summary of relative body weights (% control) of study animals measured across multiple generations/reproductive intervals, (A) F1c male/female (post-weaning), (B) F1c females during gestation intervals A, B, and C, (C) F1c females during lactation interval (litter C), and (D) F2c male/female pups prior to weaning.

Table 1

Butylparaben intake (mg/kg/day) based on feed consumption during the various intervals of the RACB study.

Generation/Study Interval	5000 ppm	15,000 ppm	40,000 ppm
F0			
Prior to Pairing (M)	324.8	1013.2	3095.4
Prior to Pairing (F)	338.7	1016.1	2599.5
Gestation[a]	335.6	990.8	3170.2
Lactation (PND 1–4)[a]	572.2	1763.8	5967.0
Lactation (PND 1–13)[b]	730.9	2062.1	6116.6
F1c			
Prior to Pairing (M)[c]	427.6	1380.5	4849.8
Prior to Pairing (F)[c]	467.6	1455.8	5214.6
F1cP			
Gestation[a]	343.2	1034.1	3025.2
Lactation (PND 1–4)[a]	546.3	1700.4	6325
Lactation (PND 1–13)[b]	686.8	2067.0	6709.4

[a]Average intake (mg/kg/day) of dams across the three breeding (A,B,C) periods.

[b]Intake of the dams (mg/kg/day) during lactation period of the C litter.

[c]Average intake (mg/kg/day) of all F1c males and females (parental and non-parental) post-weaning, prior to pairing or sacrifice (PND.35–84).

Table 2

Animal removal summary and rational during the BP RACB study.

Generation/Sex	Removal	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
F0					
Male	Study Start	22	22	22	22
	Moribund	2	1	1	0
	Scheduled Removal	20	21	21	22
Female	Study Start	22	22	22	22
	Moribund	1	4	1	0
	Scheduled Removal	21	18	21	22
F1c-Nonparental					
Male	Weaned	45	34	46	45
	Unscheduled Removal	0	0	0	25[a,b]
	Moribund	1	0	0	5
	Found Dead	1	0	0	9
	Scheduled Removal	43	34	46	6
Female	Weaned	40	35	41	34
	Unscheduled Removal	0	0	0	14[a,b]
	Moribund	1	0	0	6
	Found Dead	0	0	0	14
	Scheduled Removal	39	35	41	0
F1c-Parental					
Male	Weaned	40	40	40	26
	Moribund	1	3	1	0
	Scheduled Removal	39	37	39	26
Female	Weaned	40	40	40	26
	Moribund	3	2	4	0
	Found Dead	0	0	1	0
	Scheduled Removal	37	38	35	26
F2c					
Male	Scheduled Removal	121	133	130	97
Female	Scheduled Removal	117	132	114	79

[a] 40,000 ppm pups weaned on PND21 were unable to thrive. All removals occurred within 7 days of weaning. Postponed weaning of half of 40,000 ppm pups (PND 28) increased survival.

[b] Due excessive pup loss in 40,000 ppm offspring weaned on PND 21, surviving offspring, weaned on PND 28, were primarily assigned to the F1 parental cohort. A limited number of 40,000 ppm offspring were available for assignment to the nonparental cohort.

Table 3

Average reproductive performance across the three mating intervals (A, B, C) following BP exposure.

Generation	F0				F1c			
Endpoint	0 ppm	5000 ppm	15,000 ppm	40,000 ppm	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
Mating Pairs	22	19	22	22	38	35	37	26
Mated/Paired (%) Average[a]	97	96.8	97	98.5	80.7	83	87.7	93.6
Littered/Paired (%) Average[a]	89.4	81.7	90.9	83.3	73.6	75.5	82.6	92.3
Littered/Mated (%) Average[a]	92.2	84.6	93.8	84.7	90.7	90.3	93.8	98.7
Number of Litters/Pair Average[a,b]	2.6 ± 0.1	2.4 ± 0.2	2.7 ± 0.1	2.5 ± 0.1	2.2 ± 0.2	2.4 ± 0.1	2.4 ± 0.2	2.8 ± 0.1

[a] Data represent average across A, B, and C pairings.

[b] Data presented as average ± SEM. Comparisons are between generations for each dose group. No trend test was performed.

Table 4

Summary of littering parameters across the three mating intervals (A, B, C) following BP exposure.

Generation		F0				F1cP			
Endpoint	Litter	0 ppm	5000 ppm	15,000 ppm	40,000 ppm	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
Total Litter Size (PND 0)	A	13.3 ± 1.0**	14.1 ± 0.6	12.9 ± 0.7	11.7 ± 0.6**	13.8 ± 0.9*	14.3 ± 0.7	13.1 ± 0.7	11.9 ± 0.4*
	B	12.8 ± 1.0*	14.0 ± 0.9	14.0 ± 0.6	11.8 ± 0.4	13.4 ± 0.9	12.1 ± 1.0	12.5 ± 0.9	12.0 ± 0.6
	C	11.4 ± 0.8	13.6 ± 0.7	11.4 ± 0.9	9.4 ± 1.0	12.0 ± 1.0	11.4 ± 0.9	10.2 ± 1.0	8.6 ± 0.8
Live Litter Size (PND 0)	A	12.1 ± 1.0*	13.4 ± 0.6	12.3 ± 0.7	10.9 ± 0.7	11.9 ± 0.8	12.4 ± 0.8	12.2 ± 0.6	10.9 ± 0.6
	B	11.8 ± 1.0	12.8 ± 0.9	12.5 ± 0.6	11.1 ± 0.7	11.1 ± 0.9	10.7 ± 1.1	11.4 ± 1.0	11.1 ± 0.7
	C	10.4 ± 0.9	12.6 ± 0.6	11.0 ± 0.8	9.2 ± 1.0	10.4 ± 0.9	10.4 ± 0.9	9.2 ± 0.9	8.1 ± 0.8
Live Birth Ratio (PND 0)	A	0.85 ± 0.05	0.96 ± 0.02	0.95 ± 0.01	0.93 ± 0.02	0.88 ± 0.04	0.87 ± 0.04	0.93 ± 0.02	0.92 ± 0.04
	B	0.87 ± 0.05	0.91 ± 0.03	0.90 ± 0.02	0.92 ± 0.05	0.81 ± 0.04	0.85 ± 0.07	0.91 ± 0.03*	0.93 ± 0.02
	C	0.91 ± 0.04*	0.93 ± 0.02	0.97 ± 0.01	0.99 ± 0.01	0.86 ± 0.04	0.91 ± 0.04	0.91 ± 0.02	0.94 ± 0.02
Average Live Litter Size/Pair	Avg [a]	11.3 ± 0.7	12.9 ± 0.3	11.9 ± 0.4	10.6 ± 0.6	10.9 ± 0.6	10.7 ± 0.4††	10.8 ± 0.6	10.1 ± 0.4
Survival Ratio (PND 1–4)	A	0.93 ± 0.06	0.99 ± 0.01	0.98 ± 0.01	0.97 ± 0.01	0.94 ± 0.03	0.94 ± 0.03	0.93 ± 0.03	0.93 ± 0.01
	B	0.99 ± 0.01	0.96 ± 0.04	0.97 ± 0.02	0.96 ± 0.01	0.93 ± 0.01	0.96 ± 0.02	0.92 ± 0.04	0.97 ± 0.01
	C	0.99 ± 0.01	0.97 ± 0.01	0.99 ± 0.01	0.96 ± 0.03	0.98 ± 0.01	0.97 ± 0.01	0.92 ± 0.04	0.97 ± 0.01
Survival Ratio (PND 5–21)	C	0.98 ± 0.01	0.96 ± 0.02	0.97 ± 0.01	0.91 ± 0.05	0.97 ± 0.02	0.97 ± 0.01	0.95 ± 0.03	0.92 ± 0.03

Statistical significance for the control group (#) indicates a significant trend test. Statistical significance for a treatment group (*) indicates a significant pairwise test compared to the control group. F1 litter size and survival endpoints were analyzed using Jonckheere's test for trend and Shirley's or Dunn's methods for pairwise comparison of controls to dose groups. F2 litter size and survival endpoints were analyzed using the bootstrapped Jonckheere trend test. Pairwise comparisons were done using the Datta-Satten modified Wilcoxon tests with the Hommel adjustment for multiple comparisons.

Data presented as average ± SEM.

*
Statistically significant at p ≤ 0.05.

**
p ≤ 0.01.

[a]
Data represent average across A, B, and C pairings. Comparisons are between generations for each dose group. No trend test was performed.

††
p < 0.01.

Table 5

Summary of Developmental/Pubertal marker findings in study animals following BP exposure.

Generation	F1c				F2c			
Endpoint	0 ppm	5000 ppm	15,000 ppm	40,000 ppm	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
Males, Preweaning								
Examined, (No. litters)	98 (19)	76 (14)	102 (18)	72 (16)	121 (27)	133 (30)	130 (29)	98 (24)
Areolae/Nipples per Litter[a]	0.9 ± 0.06	0	0	0	0	0	0	0.4 ± 0.4
No. Pups with Areolae/Nipples, (%)[b]	2 (2.04)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (2.04)
No. Litters with Areolae/Nipples, (%)[b]	2 (10.53)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (4.17)
Day of Testes Descent (PND)[c]	15.3 ± 0.2	15.9 ± 0.4	15.4 ± 0.3	15.6 ± 0.3	15.5 ± 0.2**	15.5 ± 0.2	14.8 ± 0.1*	14.9 ± 0.2
Males, Postweaning								
Examined (No. litters)	86 (19)	74 (14)	93 (18)	32 (7)				
Age at BPS (PND)[c]	44.4 ± 0.4**	44.6 ± 0.6	46.0 ± 0.5	50.0 ± 0.8**				
Body Weight at BPS (g)[d]	198.8 ± 2.8**	194.6 ± 2.7	194.7 ± 2.9	175.7 ± 4.0**				
Body Weight at PND 21 (g)[d]	51.2 ± 1.3**	46.7 ± 1.3	44.0 ± 1.4**	36.0 ± 1.5**				
Adjusted Age at BPS[e,f]	45.8 ± 0.4*	45.0 ± 0.5	45.7 ± 0.5	47.7 ± 0.6				
Females, Postweaning								
Examined (No. litters)	89 (20)	78 (14)	87 (19)	26 (8)				
Age at VO (PND)[c]	34.5 ± 0.4**	35.0 ± 0.4	37.7 ± 0.4**	43.3 ± 1.1**				
Body Weight at VO (g)[d]	114.3 ± 1.7	111.1 ± 2.7	116.8 ± 2.9	117.8 ± 3.4				
Body Weight at PND 21 (g)[d]	50.1 ± 1.1**	43.7 ± 1.4**	41.3 ± 1.6**	35.5 ± 2.0**				
Adjusted Age at VO[e,f]	35.4 ± 0.4**	35.0 ± 0.4	37.4 ± 0.4**	42.1 ± 0.9**				

Statistical significance for the control group indicates a significant trend test. Statistical significance for a treatment group indicates a significant pairwise test compared to the control group.

*
Statistically significant at p   0.05.

**
p   0.01.

[a]
Statistical analysis for the F1 generation performed by Jonckheere (trend) and Shirley or Dunn (pairwise) tests. Statistical analysis for the F2 generation performed using a bootstrapped Jonckheere trend test; pairwise comparisons were done using the Datta-Satten modified Wilcoxon tests with Hommel adjustment for multiple comparisons.

[b] Statistical analysis for the F1 generation was performed using Cochran-Armitage (trend) and Fisher Exact (pairwise) tests. Statistical analysis for the F2 generation was performed using a Rao-Scott Cochran-Armitage test for both trend and pairwise tests.

[c] Means of litter means presented. Trend and pairwise tests were based on mixed models for day of attainment with dose as a covariate and a random effect for litter. The Dunnett-Hsu adjustment was used for multiple comparisons.

[d] Analysis of body weight at attainment and body weight at weaning were performed using mixed effects models with dose as covariate and a random effect for litter. The Dunnett-Hsu adjustment was used for multiple comparisons. Animals that attained during the observation period were used for analysis.

[e] Mean adjusted day of attainment was calculated from the mean of the litter means of the weaning weight-adjusted attainment days for individual pups. Trend and pairwise tests were based on mixed models for day of attainment with dose and weaning weight as covariates and a random effect for litter. The Dunnett-Hsu adjustment was used for multiple comparisons.

[f] Balaneo-pre-preputial Separation (BPS) and Vaginal Opening (VO) were adjusted using body weight at weaning as a covariate.

Table 6

Summary of sperm parameters of F0, F1cNP, and F1cP males following BP exposure.

Endpoint	Generation	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
No. Examined (Litters)	F0	20	21	21	22
	F1cNP	43 (16)	34 (12)	46 (18)	6 (5)[a]
	F1cP	39 (19)	37 (13)	39 (18)	26 (7)
Spermatid Head Count (x10^6)	F0	349.8 ± 9.5	327.4 ± 10.1	353.4 ± 15.8	340.7 ± 5.7
	F1cNP	251.2 ± 9.8	275.1 ± 10.2	267.8 ± 7.6	222.0 ± 11.1
	F1cP	372.2 ± 13.1	389.5 ± 6.2	386.5 ± 7.8	366.8 ± 9.2
Spermatid Head Count/gram testis (x10^6)	F0	172.8 ± 4.4	158.8 ± 5.5	167.4 ± 6.5	165.6 ± 3.4
	F1cNP	131.4 ± 4.1	140.3 ± 5.2	137.4 ± 4.1	131.5 ± 12.3
	F1cP	176.5 ± 4.1	179.1 ± 2.5	179.1 ± 2.9	181.5 ± 2.9
Motile (%)	F0	80.5 ± 2.0	80.4 ± 1.8	80.0 ± 2.2	78.2 ± 2.0
	F1cNP	69.3 ± 2.2	67.4 ± 2.6	69.4 ± 2.1	74.9 ± 3.0
	F1cP	70.2 ± 3.3	68.6 ± 3.0	72.2 ± 1.7	73.3 ± 2.6
Progressively Motile (%)	F0	65.1 ± 1.9	66.0 ± 1.9	65.7 ± 2.1	64.3 ± 2.0
	F1cNP	54.5 ± 2.1	53.6 ± 2.1	55.5 ± 1.7	61.1 ± 4.4
	F1cP	60.1 ± 3.1	60.2 ± 2.8	63.5 ± 1.5	64.9 ± 2.6
Sperm Count/Cauda (x10^6)	F0	222.9 ± 10.3	218.3 ± 9.8	224.9 ± 9.4	224.3 ± 8.2
	F1cNP	180.3 ± 9.5	177.5 ± 5.8	181.9 ± 6.1	150.8 ± 15.8
	F1cP	213.8 ± 8.7	222.5 ± 6.5	226.1 ± 6.8	234.2 ± 12.2
Sperm Count/gram Cauda (x10^6)	F0	817.4 ± 29.2	764.0 ± 30.8	790.3 ± 26.0	791.0 ± 23.9
	F1cNP	815.5 ± 23.5	808.6 ± 20.6	841.9 ± 20.6	818.8 ± 71.4
	F1cP	797.8 ± 19.6	795.8 ± 8.0	819.4 ± 22.1	861.6 ± 31.2

Statistical analysis for F0 data performed by Jonckheere (trend) and then a Shirley/Dunn pairwise test. Statistical analysis of F1 and/or F2 endpoints used a bootstrapped Jonckheere trend test and pairwise comparisons were done using the Datta-Satten modified Wilcoxon test with Hommel adjustment for multiple comparisons.

Data presented as average ± SEM.

[a] Only six animals were evaluated in this group due to pup loss related to failure to thrive postweaning.

Table 7

BP-associated effects on terminal body weights and liver weights in F0, F1cNP, and F1cP male and female rats.

Endpoint	Generation	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
Males					
No. Examined (Litters)	*F0*	20	21	21	22
	F1cNP	43 (16)	34 (12)	46 (18)	6 (5)[a]
	F1cP	39 (19)	37 (13)	39 (18)	26 (7)
Terminal Body Weight (g)	*F0*	503.6 ± 6.6**	497.6 ± 7.1	476.7 ± 5.0	448.8 ± 5.5
	F1cNP	386.0 ± 9.4**	388.1 ± 6.5	356.4 ± 6.1**	298.1 ± 16.0**
	F1cP	511.1 ± 9.5**	504.7 ± 5.3	474.1 ± 8.8**	409.8 ± 6.5**
Liver Absolute (g)	*F0*	18.84 ± 0.36*	19.27 ± 0.36	18.84 ± 0.52	17.85 ± 0.35
	F1cNP	16.58 ± 0.53**	17.23 ± 0.37	16.05 ± 0.34	14.52 ± 0.83
	F1cP	20.79 ± 0.56**	20.71 ± 0.54	19.21 ± 0.41*	17.77 ± 0.34**
Liver Relative (mg/g)[b]	*F0*	37.40 ± 0.44**	38.76 ± 0.60	39.48 ± 0.92*	39.75 ± 0.51*
	F1cNP	42.83 ± 0.49**	44.38 ± 0.70	45.03 ± 0.59*	48.67 ± 0.82**
	F1cP	40.63 ± 0.64*	41.02 ± 0.93	40.75 ± 0.54	43.36 ± 0.53*
Females					
No. Examined (Litters)	*F0*	21	18	21	22
	F1cNP	39 (17)	35 (14)	41 (15)	—[a]
	F1cP	27 (18)	30 (14)	30 (17)	24 (8)
Terminal Body Weight (g)	*F0*	297.1 ± 4.6**	297.8 ± 3.7	288.7 ± 2.8	286.5 ± 3.5*
	F1cNP	258.2 ± 3.7**	252.1 ± 5.4	230.3 ± 5.3**	—
	F1cP	356.5 ± 4.9**	356.0 ± 5.8	340.7 ± 6.5	293.9 ± 11.4**
Liver Absolute (g)	*F0*	10.17 ± 0.21	10.59 ± 0.21	10.23 ± 0.20	10.66 ± 0.27
	F1cNP	10.56 ± 0.27	10.81 ± 0.28	10.03 ± 0.39	—
	F1cP	17.31 ± 0.29	18.11 ± 0.48	18.02 ± 0.45	17.00 ± 0.99
Liver Relative (mg/g)[b]	*F0*	34.19 ± 0.34**	35.53 ± 0.40	35.44 ± 0.57	37.13 ± 0.63**
	F1cNP	40.79 ± 0.61*	42.83 ± 0.59	43.42 ± 0.98*	—

Endpoint	Generation	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
F1cP		48.59 ± 0.60 **	50.98 ± 1.32	52.93 ± 1.19 *	57.60 ± 1.53 **

Statistical analysis for F0 animals performed by Jonckheere (trend) and Williams or Dunnett (pairwise) tests. Statistical analysis for F1 animals was performed by linear mixed models, with dam ID as random effect, and using Dunnett-Hsu adjustment, for both trend and pairwise analyses. Statistical significance for the control group indicates a significant trend test. Statistical significance for a treatment group indicates a significant pairwise test compared to the control group.

Data presented as average ± SEM.

*
Statistically significant at p 0.05;

**
p 0.01.

*a*Reduced numbers of animals were evaluated in this group due to pup loss related to failure to thrive postweaning.

*b*Relative organ weights (mg) normalized to terminal body weight (g).

Table 8

Summary of incidences and severity of select non-neoplastic lesions in F0, F1cNP, and F1cP BP-exposed male and female rats.

Endpoint	Generation	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
Males					
No. Examined (Litters)	F0	22	22	22	22
	F1cNP	45 (16)	34 (12)	46 (18)	6 (5)^a
	F1cP	40 (19)	40 (14)	40 (18)	26 (7)
Liver					
Bile Duct; Hyperplasia	F0	0 (0 %)	1 (4.5 %) [1.0]^b	2 (9.1 %) [1.0]	0 (0 %)
	F1cNP	0 (0 %)**	0 (0 %)	9 (19.6 %) [1.0]	4 (66.7 %)* [1.0]
	F1cP	3 (7.5 %)** [1.0]	2 (5.0 %) [1.0]	12 (30.0 %)* [1.0]	19 (73.1 %)** [1.0]
Infiltration; Mononuclear Cell	F0	3 (13.6 %)** [1.0]	2 (9.1 %) [1.0]	19 (86.4 %)** [1.0]	22 (100.0 %)** [1.0]
	F1cNP	1 (2.2 %)** [1.0]	4 (11.8 %) [1.0]	29 (63.0 %)** [1.0]	0 (0 %)
	F1cP	0 (0 %)**	4 (10.0 %) [1.0]	32 (80.0 %)** [1.0]	16 (61.5 %)** [1.0]
Hepatocyte; Increased Mitosis	F0	0 (0 %)**	0 (0 %)	0 (0 %)	3 (13.6 %) [1.0]
	F1cNP	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
	F1cP	0 (0 %)**	0 (0 %)	0 (0 %)	5 (19.2 %)* [1.0]
Hepatocyte; Cytoplasmic Vacuolation	F0	0 (0 %)**	0 (0 %)	11 (50.0 %)** [1.0]	22 (100.0 %)** [1.3]
	F1cNP	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
	F1cP	0 (0 %)**	0 (0 %)	9 (22.5 %)* [1.0]	17 (65.4 %)** [1.2]
Females					
No. Examined (Litters)	F0	22	22	22	22
	F1cNP	40 (17)	35 (14)	41 (15)	—^a
	F1cP	40 (20)	40 (14)	40 (18)	26 (8)
Liver					
Bile Duct; Hyperplasia	F0	0 (0 %)*	0 (0 %)	0 (0 %)	2 (9.1 %) [1.0]
	F1cNP	0 (0 %)	1 (2.9 %) [1.0]	0 (0 %)	—
	F1cP	3 (7.5 %)** [1.0]	3 (7.5 %) [1.0]	5 (12.5 %) [1.0]	11 (42.3 %)** [1.0]

Endpoint	Generation	0 ppm	5000 ppm	15,000 ppm	40,000 ppm
Infiltration; Mononuclear Cell	F0	1 (4.5 %) ** [1.0]	1 (4.5 %) [1.0]	11 (50.0 %) ** [1.0]	15 (68.2 %) ** [1.1]
	F1cNP	0 (0 %) **	1 (2.9 %) [1.0]	27 (65.9 %) ** [1.0]	—
	F1cP	1 (2.5 %) * [1.0]	0 (0 %)	10 (25.0 %) * [1.1]	4 (15.4 %) [1.0]
Hepatocyte; Periportal Hypertrophy	F0	0 (0 %) **	2 (9.1 %) [1.0]	5 (22.7 %) * [1.0]	7 (31.8 %) ** [1.0]
	F1cNP	0 (0 %) *	1 (2.9 %) [1.0]	6 (14.6 %) [1.0]	—
	F1cP	0 (0 %) **	22 (55.0 %) ** [1.0]	28 (70.0 %) ** [1.1]	24 (92.3 %) ** [1.2]
Hepatocyte; Increased Mitosis	F0	0 (0 %) **	0 (0 %)	0 (0 %)	3 (13.6 %) [1.0]
	F1cNP	0 (0 %)	0 (0 %)	0 (0 %)	—
	F1cP	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
Hepatocyte; Cytoplasmic Inclusion	F0	0 (0 %)	0 (0 %)	0 (0 %)	1 (4.5 %) [1.0]
	F1cNP	0 (0 %)	0 (0 %)	0 (0 %)	—
	F1cP	0 (0 %) **	0 (0 %)	0 (0 %)	21 (80.8 %) ** [1.0]
Hepatocyte; Cytoplasmic Vacuolation	F0	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
	F1cNP	0 (0 %)	0 (0 %)	0 (0 %)	—
	F1cP	0 (0 %) *	0 (0 %)	0 (0 %)	6 (23.1 %) [1.0]
No. Examined (Litters)	F0	22	4 [c]	1 [c]	22
	F1cNP	40 (17)	35 (14)	41 (15)	–[a]
	F1cP	40 (20)	40 (14)	40 (18)	26 (8)
Adrenal Gland					
Adrenal Cortex – Zona Glomerulosa; Bilateral; Vacuolization	F0	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
	F1cNP	0 (0 %)	0 (0 %)	0 (0 %)	—
	F1cP	0 (0 %) **	2 (5.0 %) [1.0]	10 (25.0 %) ** [1.0]	23 (88.5 %) ** [1.0]

Statistical analysis for the F0 animals was performed using the Poly-3 trend and pairwise statistics.

Statistical analysis for the F1 Non-parental animals was performed by Cochran-Armitage test with a Rao-Scott modification for the random effect due to litter.

Statistical analysis for the F1 Parental animals was performed by Cochran-Armitage test with a poly-3 adjustment for age and a Rao-Scott modification for the random effect due to litter. Statistical significance for the control group indicates a significant trend test. Statistical significance for a treatment group indicates a significant pairwise test compared to the control group.

Data presented as Number of animals with lesion; (Percent of animals with lesion); [Mean severity].

*
Statistically significant at p   0.05;

**
p 0.01.

[a]Reduced numbers of animals were evaluated in this group due to pup loss related to failure to thrive postweaning.

[b]Average severity grade of lesions in affected animals: 1 = minimal, 2 = mild, 3 = moderate, 4 = marked.

[c]Microscopic evaluations of adrenal glands in the F0 low and mid exposure groups were limited to those with corresponding gross lesions observed.